



*applied sciences*

# Precision Dimensional Measurements

---

Edited by  
Kuang-Chao Fan and Liang-Chia Chen  
Printed Edition of the Special Issue Published in *Applied Sciences*

# Precision Dimensional Measurements



# Precision Dimensional Measurements

Special Issue Editors

**Kuang-Chao Fan**

**Liang-Chia Chen**

MDPI • Basel • Beijing • Wuhan • Barcelona • Belgrade



*Special Issue Editors*

Kuang-Chao Fan  
Dalian University of Technology  
China

Liang-Chia Chen  
National Taiwan University  
Taiwan

*Editorial Office*

MDPI  
St. Alban-Anlage 66  
4052 Basel, Switzerland

This is a reprint of articles from the Special Issue published online in the open access journal *Applied Sciences* (ISSN 2076-3417) from 2018 to 2019 (available at: [https://www.mdpi.com/journal/applsci/special\\_issues/Precision\\_Dimensional\\_Measurements](https://www.mdpi.com/journal/applsci/special_issues/Precision_Dimensional_Measurements))

For citation purposes, cite each article independently as indicated on the article page online and as indicated below:

LastName, A.A.; LastName, B.B.; LastName, C.C. Article Title. *Journal Name* **Year**, Article Number, Page Range.

**ISBN 978-3-03921-712-0 (Pbk)**

**ISBN 978-3-03921-713-7 (PDF)**

© 2019 by the authors. Articles in this book are Open Access and distributed under the Creative Commons Attribution (CC BY) license, which allows users to download, copy and build upon published articles, as long as the author and publisher are properly credited, which ensures maximum dissemination and a wider impact of our publications.

The book as a whole is distributed by MDPI under the terms and conditions of the Creative Commons license CC BY-NC-ND.

# Contents

<b>About the Special Issue Editors</b> . . . . .	<b>ix</b>
<b>Kuang-Chao Fan and Liang-Chia Chen</b> Special Issue on Precision Dimensional Measurements Reprinted from: <i>Appl. Sci.</i> <b>2019</b> , <i>9</i> , 3314, doi:10.3390/app9163314 . . . . .	<b>1</b>
<b>Meibao Wang, Xiaolin Zhang, Wenyan Tang and Jun Wang</b> A Structure for Accurately Determining the Mass and Center of Gravity of Rigid Bodies Reprinted from: <i>Appl. Sci.</i> <b>2019</b> , <i>9</i> , 2532, doi:10.3390/app9122532 . . . . .	<b>6</b>
<b>Tao Sun, Weiwei Zheng, Yingjie Yu, Ketao Yan, Anand Asundi and Sergiy Valukh</b> Algorithm for Surfaces Profiles and Thickness Variation Measurement of a Transparent Plate Using a Fizeau Interferometer with Wavelength Tuning Reprinted from: <i>Appl. Sci.</i> <b>2019</b> , <i>9</i> , 2349, doi:10.3390/app9112349 . . . . .	<b>19</b>
<b>Liang-Chia Chen, Duc-Hieu Duong and Chin-Sheng Chen</b> Optical 3-D Profilometry for Measuring Semiconductor Wafer Surfaces with Extremely Variant Reflectivities Reprinted from: <i>Appl. Sci.</i> <b>2019</b> , <i>9</i> , 2060, doi:10.3390/app9102060 . . . . .	<b>33</b>
<b>Shih-Ming Wang, Chun-Yi Lee, Hariyanto Gunawan and Chin-Cheng Yeh</b> An Accuracy-Efficiency-Power Consumption Hybrid Optimization Method for CNC Milling Process Reprinted from: <i>Appl. Sci.</i> <b>2019</b> , <i>9</i> , 1495, doi:10.3390/app9071495 . . . . .	<b>51</b>
<b>Shuyuan Gao, Rongyi Ji, Yao Li, Chun Liu, Junkai Shi, Yingling Pan and Weihu Zhou</b> Compensation of Frequency Drift in Frequency-Sweep Polarization-Modulation Ranging System Reprinted from: <i>Appl. Sci.</i> <b>2019</b> , <i>9</i> , 1243, doi:10.3390/app9061243 . . . . .	<b>71</b>
<b>Jie Li, Fengtao Yan, Lin Chen, Jie Yang and Yongjian Wan</b> Measurement Vertex Position of a Large Aspheric Mirror Reprinted from: <i>Appl. Sci.</i> <b>2019</b> , <i>9</i> , 1038, doi:10.3390/app9061038 . . . . .	<b>80</b>
<b>Hao Jiang, Zhao Ma, Honggang Gu, Xiuguo Chen and Shiyuan Liu</b> Characterization of Volume Gratings Based on Distributed Dielectric Constant Model Using Mueller Matrix Ellipsometry Reprinted from: <i>Appl. Sci.</i> <b>2019</b> , <i>9</i> , 698, doi:10.3390/app9040698 . . . . .	<b>86</b>
<b>Tung-Hsien Hsieh, Po-Yu Chen, Wen-Yuh Jywe, Guan-Wu Chen and Ming-Shi Wang</b> A Geometric Error Measurement System for Linear Guideway Assembly and Calibration Reprinted from: <i>Appl. Sci.</i> <b>2019</b> , <i>9</i> , 574, doi:10.3390/app9030574 . . . . .	<b>101</b>
<b>Chong Chen, Ryo Sato, Yuki Shimizu, Taku Nakamura, Hiraku Matsukuma and Wei Gao</b> A Method for Expansion of Z-Directional Measurement Range in a Mode-Locked Femtosecond Laser Chromatic Confocal Probe Reprinted from: <i>Appl. Sci.</i> <b>2019</b> , <i>9</i> , 454, doi:10.3390/app9030454 . . . . .	<b>120</b>
<b>Di Chang, Xu Xing, Pengcheng Hu, Jianing Wang and Jiubin Tan</b> Double-Diffracted Spatially Separated Heterodyne Grating Interferometer and Analysis on its Alignment Tolerance Reprinted from: <i>Appl. Sci.</i> <b>2019</b> , <i>9</i> , 263, doi:10.3390/app9020263 . . . . .	<b>134</b>

<b>Xingyuan Bian, Junning Cui, Yesheng Lu and Jiubin Tan</b> Ultraprecision Diameter Measurement of Small Holes with Large Depth-To-Diameter Ratios Based on Spherical Scattering Electrical-Field Probing Reprinted from: <i>Appl. Sci.</i> <b>2019</b> , <i>9</i> , 242, doi:10.3390/app9020242 . . . . .	145
<b>Shendong Shi, Linghui Yang, Jiarui Lin, Changyu Long, Rui Deng, Zhenyu Zhang and Jigui Zhu</b> Dynamic Measurement Error Modeling and Analysis in a Photoelectric Scanning Measurement Network Reprinted from: <i>Appl. Sci.</i> <b>2019</b> , <i>9</i> , 62, doi:0.3390/app9010062 . . . . .	159
<b>Yinbao Cheng, Zhongyu Wang, Xiaohuai Chen, Yaru Li, Hongyang Li, Hongli Li and Hanbin Wang</b> Evaluation and Optimization of Task-oriented Measurement Uncertainty for Coordinate Measuring Machines Based on Geometrical Product Specifications Reprinted from: <i>Appl. Sci.</i> <b>2019</b> , <i>9</i> , 6, doi:10.3390/app9010006 . . . . .	172
<b>Nianhan Wu, Wu Zhao, Xin Wang, Ye Tao and Zhengmeng Hou</b> A Novel Design of Through-Hole Depth On-Machine Optical Measuring Equipment for Automatic Drilling and Riveting Reprinted from: <i>Appl. Sci.</i> <b>2018</b> , <i>8</i> , 2671, doi:10.3390/app8122671 . . . . .	194
<b>Chao Chen, Xiuguo Chen, Yating Shi, Honggang Gu, Hao Jiang and Shiyuan Liu</b> Metrology of Nanostructures by Tomographic Mueller-Matrix Scatterometry Reprinted from: <i>Appl. Sci.</i> <b>2018</b> , <i>8</i> , 2583, doi:10.3390/app8122583 . . . . .	213
<b>Zurong Qiu, Haopeng Li, Wenchuan Hu, Chenglin Wang, Jiachen Liu and Qianhui Sun</b> Real-Time Tunnel Deformation Monitoring Technology Based on Laser and Machine Vision Reprinted from: <i>Appl. Sci.</i> <b>2018</b> , <i>8</i> , 2579, doi:10.3390/app8122579 . . . . .	225
<b>Yan Li, Yuyong Ma, Ye Tao and Zhengmeng Hou</b> Innovative Methodology of On-Line Point Cloud Data Compression for Free-Form Surface Scanning Measurement Reprinted from: <i>Appl. Sci.</i> <b>2018</b> , <i>8</i> , 2556, doi:10.3390/app8122556 . . . . .	248
<b>Liang-Chia Chen and Ching-Wen Liang</b> Novel Boundary Edge Detection for Accurate 3D Surface Profilometry Using Digital Image Correlation Reprinted from: <i>Appl. Sci.</i> <b>2018</b> , <i>8</i> , 2541, doi:10.3390/app8122541 . . . . .	266
<b>Xin Xiong, Yuki Shimizu, Xiuguo Chen, Hiraku Matsukuma and Wei Gao</b> Uncertainty Evaluation for Measurements of Pitch Deviation and Out-of-Flatness of Planar Scale Gratings by a Fizeau Interferometer in Littrow Configuration Reprinted from: <i>Appl. Sci.</i> <b>2018</b> , <i>8</i> , 2539, doi:10.3390/app8122539 . . . . .	281
<b>Xinghui Li, Yaping Shi, Xiang Xiao, Qian Zhou, Guan hao Wu, Haiou Lu and Kai Ni</b> Design and Testing of a Compact Optical Prism Module for Multi-Degree-of-Freedom Grating Interferometry Application Reprinted from: <i>Appl. Sci.</i> <b>2018</b> , <i>8</i> , 2495, doi:10.3390/app8122495 . . . . .	300
<b>Ye Wang, Hua Li, Bingjun Wan, Xiang Zhang and Gongbing Shan</b> Obtaining Vital Distances Using Wearable Inertial Measurement Unit for Real-Time, Biomechanical Feedback Training in Hammer-Throw Reprinted from: <i>Appl. Sci.</i> <b>2018</b> , <i>8</i> , 2470, doi:10.3390/app8122470 . . . . .	308

<b>Shilin Xiong, Yue Wang, Yawen Cai, Jiuli Liu, Jie Liu and Guanhao Wu</b> Calculating the Effective Center Wavelength for Heterodyne Interferometry of an Optical Frequency Comb Reprinted from: <i>Appl. Sci.</i> <b>2018</b> , <i>8</i> , 2465, doi:10.3390/app8122465 . . . . .	<b>317</b>
<b>Adam Gąska, Piotr Gąska, Maciej Gruza and Jerzy Śladek</b> Selection of Optimal Path Control Algorithms for Probe Heads Used on Five-Axis Measuring Systems Reprinted from: <i>Appl. Sci.</i> <b>2018</b> , <i>8</i> , 2455, doi:10.3390/app8122455 . . . . .	<b>324</b>
<b>Xiaolin Zhang, Hongjie Mao and Wenyan Tang</b> Demodulation Technique Based on Laser Interference for Weak Photo-Acoustic Signals on Water Surface Reprinted from: <i>Appl. Sci.</i> <b>2018</b> , <i>8</i> , 2423, doi:10.3390/app8122423 . . . . .	<b>336</b>
<b>Xiaokang Liu, Hui Zhang, Kai Peng, Qifu Tang and Ziran Chen</b> A High Precision Capacitive Linear Displacement Sensor with Time-Grating That Provides Absolute Positioning Capability Based on a Vernier-Type Structure Reprinted from: <i>Appl. Sci.</i> <b>2018</b> , <i>8</i> , 2419, doi:10.3390/app8122419 . . . . .	<b>348</b>
<b>Ingo Sieber, Allen Y. Yi and Ulrich Gengenbach</b> Metrology Data-Based Simulation of Freeform Optics Reprinted from: <i>Appl. Sci.</i> <b>2018</b> , <i>8</i> , 2338, doi:10.3390/app8122338 . . . . .	<b>359</b>
<b>Wei Qiu, Lu-Lu Ma, Hong-Tao Wang, Ren-Rong Liang, Yu-Cheng Zhao and Yun-Shen Zhou</b> Experimental Analyses on Multiscale Structural and Mechanical Properties of $\epsilon$ -Si/GeSi/C-Si Materials Reprinted from: <i>Appl. Sci.</i> <b>2018</b> , <i>8</i> , 2333, doi:10.3390/app8122333 . . . . .	<b>368</b>
<b>Guo-Ying Ren, Xing-Hua Qu and Shuang Ding</b> A Real-Time Measurement Method of Air Refractive Index Based on Special Material Etalon Reprinted from: <i>Appl. Sci.</i> <b>2018</b> , <i>8</i> , 2325, doi:10.3390/app8112325 . . . . .	<b>381</b>
<b>Xinghui Li, Haiou Lu, Qian Zhou, Guanhao Wu, Kai Ni and Xiaohao Wang</b> An Orthogonal Type Two-Axis Lloyd's Mirror for Holographic Fabrication of Two-Dimensional Planar Scale Gratings with Large Area Reprinted from: <i>Appl. Sci.</i> <b>2018</b> , <i>8</i> , 2283, doi:10.3390/app8112283 . . . . .	<b>388</b>
<b>Chuanchen Bao, Qibo Feng and Jiakun Li</b> Simultaneous Measurement Method and Error Analysis of the Six Degrees-of-Freedom Motion Errors of a Rotary Axis Reprinted from: <i>Appl. Sci.</i> <b>2018</b> , <i>8</i> , 2232, doi:10.3390/app8112232 . . . . .	<b>399</b>
<b>Yindi Cai, Zhifeng Lou, Siying Ling, Bo-syun Liao and Kuang-chao Fan</b> Development of a Compact Three-Degree-of-Freedom Laser Measurement System with Self-Wavelength Correction for Displacement Feedback of a Nanopositioning Stage Reprinted from: <i>Appl. Sci.</i> <b>2018</b> , <i>8</i> , 2209, doi:10.3390/app8112209 . . . . .	<b>410</b>
<b>Tianlong Yang, Qiancheng Zhao, Xian Wang and Quan Zhou</b> Sub-Pixel Chessboard Corner Localization for Camera Calibration and Pose Estimation Reprinted from: <i>Appl. Sci.</i> <b>2018</b> , <i>8</i> , 2118, doi:10.3390/app8112118 . . . . .	<b>424</b>



<b>Dingfu Chen, Junzheng Peng, Sergiy Valyukh, Anand Asundi and Yingjie Yu</b> Measurement of High Numerical Aperture Cylindrical Surface with Iterative Stitching Algorithm Reprinted from: <i>Appl. Sci.</i> <b>2018</b> , <i>8</i> , 2092, doi:10.3390/app8112092 . . . . .	<b>443</b>
<b>Xing-Ting Xiong, Xing-Hua Qu and Fu-Min Zhang</b> Error Correction for FSI-Based System without Cooperative Target Using an Adaptive Filtering Method and a Phase-Matching Mosaic Algorithm Reprinted from: <i>Appl. Sci.</i> <b>2018</b> , <i>8</i> , 1954, doi:10.3390/app8101954 . . . . .	<b>455</b>
<b>Yufu Qu, Renju Peng, Jialin Hao, Hui Pan, Jiebin Niu and Jie Jiang</b> Influence of Illumination Polarization and Target Structure on Measurement Sensitivity of Through-Focus Scanning Optical Microscopy Reprinted from: <i>Appl. Sci.</i> <b>2018</b> , <i>8</i> , 1819, doi:10.3390/app8101819 . . . . .	<b>468</b>
<b>Yanlin He, Lianqing Zhu, Guangkai Sun, Mingxin Yu and Mingli Dong</b> Design, Measurement and Shape Reconstruction of Soft Surgical Actuator Based on Fiber Bragg Gratings Reprinted from: <i>Appl. Sci.</i> <b>2018</b> , <i>8</i> , 1773, doi:10.3390/app8101773 . . . . .	<b>480</b>
<b>Rui-Jun Li, Peng-Yu Wang, Dan-Dong Li, Kuang-Chao Fan, Fang-Fang Liu, Li-Juan Chen and Qiang-Xian Huang</b> Precision Manufacturing of Patterned Beryllium Bronze Leaf Springs via Chemical Etching Reprinted from: <i>Appl. Sci.</i> <b>2018</b> , <i>8</i> , 1476, doi:10.3390/app8091476 . . . . .	<b>495</b>
<b>Na Zhao, Changku Sun and Peng Wang</b> Calibration Method of Orthogonally Splitting Imaging Pose Sensor Based on General Imaging Model Reprinted from: <i>Appl. Sci.</i> <b>2018</b> , <i>8</i> , 1399, doi:10.3390/app8081399 . . . . .	<b>510</b>
<b>Shanwei Yang, Zhan Gao, Haihui Ruan, Chenjia Gao, Xu Wang, Xiang Sun and Xin Wen</b> Non-Contact and Real-Time Measurement of Kolsky Bar with Temporal Speckle Interferometry Reprinted from: <i>Appl. Sci.</i> <b>2018</b> , <i>8</i> , 808, doi:10.3390/app8050808 . . . . .	<b>523</b>
<b>Kaiyi Zhu, Hongfang Chen, Shulian Zhang, Zhaoyao Shi, Yun Wang and Yidong Tan</b> Frequency-Shifted Optical Feedback Measurement Technologies Using a Solid-State Microchip Laser Reprinted from: <i>Appl. Sci.</i> <b>2019</b> , <i>9</i> , 109, doi:10.3390/app9010109 . . . . .	<b>538</b>
<b>Wenwen Liu, Penghao Hu and Kuangchao Fan</b> Comparison of Current Five-Point Cylindricity Error Separation Techniques Reprinted from: <i>Appl. Sci.</i> <b>2018</b> , <i>8</i> , 1946, doi:10.3390/app8101946 . . . . .	<b>565</b>

## About the Special Issue Editors

**Kuang-Chao Fan**, Professor, received his Ph.D. degree in mechanical engineering from the University of Manchester Institute of Science and Technology (UMIST) in 1984 and the DEng of University of Manchester in 2014. Originally serving as a professor of mechanical engineering at National Taiwan University from 1989 to 2015, he became a Cheng Kong Scholar at Hefei University of Technology in 2001. He is a Fellow of SME, ISNM, CSME, and CIAE. His current research interests include manufacturing metrology, machine tool metrology, robot calibration, and micro/nano measurement. He has published more than 170 journal papers and 250 conference papers, and he holds more than 40 national and international patents. He has also received many academic honors and awards. Currently, he is working with Dalian University of Technology and serves as the editor of four SCI journals.

**Liang-Chia Chen**, Professor, is currently working as a professor at the Department of Mechanical Engineering of National Taiwan University (NTU), Taiwan. His major research fields are in precision metrology and manufacturing, automated optical inspection (AOI), optomechanics instrumentation, and 3D machine vision and algorithms for industrial automation. He has published one textbook, four book chapters, more than 120 referred journal papers, and 65 invention patents internationally. In 2015–2017, he served as the president of the international committee of measurement and instrumentation (ICMI), and he currently serves as a committee board member in ICMI and IMEKO. He is also participating as a member in SPIE, SME, the Institution of Engineers of Australia (IEA), and Chinese Institute of Engineers (CIE) and is a board member of the Taiwan Society of Precision Engineering (TSPE), Taiwan Society of Automation technology, Automated Optical Inspection Equipment Association (AOIEA), and Chinese Metrology Society (CMS).



Editorial

# Special Issue on Precision Dimensional Measurements

Kuang-Chao Fan <sup>1,2,\*</sup>  and Liang-Chia Chen <sup>2,\*</sup> 

<sup>1</sup> School of Mechanical Engineering, Dalian University of Technology, 2 Linggong Rd., Dalian 116023, China

<sup>2</sup> Department of Mechanical Engineering, National Taiwan University, 1, Sec. 4, Roosevelt Rd., Taipei 10617, Taiwan

\* Correspondence: fan@ntu.edu.tw (K.-C.F.); lchen@ntu.edu.tw (L.-C.C.)

Received: 2 August 2019; Accepted: 6 August 2019; Published: 12 August 2019



## 1. Introduction

Precision dimensional measurements always play the critical role in workshop quality control. Although there are many instruments on the market suitable for a variety of measuring demands, such as lengths, angles, regular forms, free forms, and special geometries, not all products can be easily measured by current instruments and technologies. Some cutting-edge manufacturing processes have produced many difficult to measure parts, such as complex geometry, 3D forms, mini-to-micro-sized parts, ultrahigh precision, high aspect ratio, large-scaled, soft surface, variable reflectance surface, etc. Some new measurement sciences and technologies are urgently needed to cope with these demands.

Ever since this Special Issue was announced in May 2018, it has received a great attention and many positive responses from worldwide experts in measurement sciences and instrumentation, especially China. There were 87 papers submitted to this special issue, and 41 papers were accepted. Such a high acceptance rate reveals the good quality and interesting subjects contributed from authors of renowned universities and institutes. These published papers can be categorized into four main groups, including surface profile and form measurements, sensors and probes development, laser-based interferometer systems, and other measurement systems. Details are introduced in the following.

## 2. Surface Profile and Form Error Measurements

The first paper on 3D surface profile measurement is introduced by Qu and his colleagues at Beihang University on the measurement sensitivity of through-focus scanning optical microscopy (TSOM), which involves scanning a target through the focus and capturing of a series of images [1]. It was found that for enhanced sensitivity, illumination polarization should be perpendicular to the target texture. Professor Gao's group at Tohoku University presents a different method to increase Z-measurement range of 3D profile using mode-locked femtosecond laser chromatic confocal probe [2]. They also provide an uncertainty evaluation method for measurements of pitch deviation and out-of-flatness of planar scale gratings by a Fizeau Interferometer in Littrow configuration [3]. A collaborative work between Karlsruhe Institute of Technology of Germany and Ohio State University of USA presents a generalized approach using measurement data to enhance the simulation model for designing freeform optics [4]. The boundary of a surface is difficult to measure by normal microscopy. Professor Chen of National Taiwan University employed the digital image correlation (DIC) technique to assist accurate edge detection of 3D surface profilometry [5]. He also proposes a new moiré projection triangulation method using a dual optical sensing configuration by engaging two optical sensors at two different viewing angles for achieving simultaneous full-field surface profilometry for measuring semiconductor wafer surfaces with extremely variant reflectivities [6]. An international cooperative work by Shanghai University (Sun et al.), Nanyang Technological University (Asundi) and Linköping University (Valukh) presents the algorithm for surfaces profiles and thickness variation

measurement of a transparent plate using a Fizeau interferometer with wavelength tuning [7]. Another useful algorithm of iterative stitching was also developed by this team on high numerical aperture cylindrical surface [8]. Li et al. proposes an innovative methodology of cloud data compression for free-form surface scanning measurement. It can save substantial data memory while remaining in good quality [9]. We are pleased to have the contribution by Professor Liu at HUST on the Mueller matrix technique for nanostructure scatterometry [10] and ellipsometry [11]. Li et al. deals with the method of measuring a large aspherical mirror and identifying its vertex position [12]. Liu et al. summarizes and compares the current five-point cylindrical error separation method with proximity sensors [13]. This provides a guideline for accurate evaluation of cylindrical errors.

### **3. Sensors and Probes**

There are some papers focusing on the fabrication of new sensors and probes. Li's group at Hefei University of Technology successfully developed a chemical etching method to fabricate different shapes of leaf springs [14]. This is a key component of any touch-trigger probe on a coordinate measuring machine (CMM). Professor Tan's group at HIT (Harbin Institute of Technology) introduces a novel spherical scattering electrical-field contact probe [15]. This makes the measurement of ultraprecision diameter of small holes with large depth-to-diameter ratio possible. Image processing technique is often used in visual measurement. Camera calibration is a primary task to assure measurement accuracy. Zhao [16] and Yang [17] address respective methods to calibrate image pose sensors to subpixel level. An interesting time-grating capacitive displacement sensor is introduced by Professor Liu of Chongqing University of Technology [18]. It has the capability for absolute positioning based on a vernier-type structure. He et al. propose a real-time 3D shape reconstruction method of soft surgical actuator which has an embedded optical fiber with two Fiber Bragg Grating (FBG) sensors [19].

### **4. Laser Interferometer Length Measurements**

The distance measurement based on various laser interferometer systems is a highlight of the Special Issue. This part includes incremental distance interferometer, absolute distance interferometer and grating diffraction interferometer. For the miniaturization of laser interferometer, a compact laser diode-based interferometer was developed by Cai et al. of Dalian University of Technology [20]. The wavelength cannot be determined by the Edlén Equation, which is mainly used for He-Ne lasers. Therefore, a grating diffraction method is employed to real-time detect the wavelength. Ren et al. from Tianjin University developed a different method to measure air refractive index based on the special material Etalon [21]. Professor Zhang's group of Tsinghua University describes the unique frequency-shifted optical feedback measurement technologies using a solid-state microchip laser [22]. In the area of absolute distance measurement, the frequency scanning (FS) laser is the main laser source. Xiong et al. calculated the effective center wavelength for heterodyne interferometry of an optical frequency comb [23]. Gao et al. propose a technique to compensate the FS interferometer [24] and Xiong et al. propose an adaptive filtering method to correct errors due to interfering signals of FS interferometer [25]. In the area of grating diffraction-based interferometer, a double diffraction heterodyne interferometer was developed by Chang [26]. Li et al. of Tsinghua University designed two innovate key components for planar grating interferometers, namely the Lloy's mirror [27] and the prism module [28].

### **5. CMM and Machine Tool Error Measurements**

Coordinate measuring machines (CMMs) provide 3D measurement capability with probes. Cheng et al. evaluated and optimized task-oriented measurement uncertainty for CMMs based on geometrical product specifications [29]. Gaska et al. searched the optimal path control algorithm for probe heads used on five-axis CMMs [30]. Wu et al. developed on-machine optical measuring equipment for drilled holes [31].

There are some reports on the geometrical error measurement of machine tools. Professor Jywe's group of National Formosa University developed a geometric error measurement system to assist the linear guideway assembly process [32]. Professor Feng's group of Beijing Jiaotong University propose a simultaneous six-degree-of-freedom error measurement system for rotary stages [33]. Wang considered a hybrid optimization method to optimize the power consumption during CNC (computer numerically controlled) milling process [34].

## 6. Other Measurement Systems

In addition to dimensional measurements, there are some other purposes of measurement. Because they are interesting, we also consider the following for inclusion in this Special Issue. Wang studied the biomechanical feedback system in training of hammer-throw [35]. A wearable inertia measurement unit has been developed that could help obtain the vital distance. An efficient real-time tunnel deformation monitoring technology based on laser and machine vision was developed by Qiu et al. [36]. In the large-scale metrology, Shi et al. use photoelectric scanning measurement network to model dynamic errors [37]. Wang designed a structure for accurately determining the mass and center of gravity of rigid bodies [38]. Yang et al. developed non-contact and real-time measurement of high strain rate Kolsky bar with temporal speckle interferometry [39]. Zhang et al. present a demodulation technique for weak photo-acoustic signals of laser interferometer on water surface [40]. The last paper was written by Qiu et al. on the experimental analyses on multiscale structural and mechanical properties of  $\epsilon$ -Si/GeSi/C-Si materials [41]. Strained silicon ( $\epsilon$ -Si) is a promising material that could extend Moore's law by enhancing electron mobility. The layer components and thicknesses were measured using energy-dispersive spectroscopy and scanning-electron microscopy. Crystal and lattice characters were observed by using high-resolution transmission-electron microscopy and micro-Raman spectroscopy.

**Acknowledgments:** This issue would not be successful without the contributions of earnest authors, professional reviewers, and dedicated editorial team of *Applied Sciences*. We are grateful to all authors, who spent time to prepare the manuscript and respond comments from the reviewers, no matter what the final decisions of the submitted manuscripts were. Special thanks also to all reviewers, who always provided fast and responsible services to the requests. Finally, we take this opportunity to express our gratitude to the editorial team of *Applied Sciences*, and special thanks to Daria Shi, and Xiaoyan Chen both from MDPI Branch Office, Beijing.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Qu, Y.; Peng, R.; Hao, J.; Pan, H.; Niu, J.; Jiang, J. Influence of Illumination Polarization and Target Structure on Measurement Sensitivity of Through-Focus Scanning Optical Microscopy. *Appl. Sci.* **2018**, *8*, 1819. [[CrossRef](#)]
2. Chen, C.; Sato, R.; Shimizu, Y.; Nakamura, T.; Matsukuma, H.; Gao, W. A Method for Expansion of Z-Directional Measurement Range in a Mode-Locked Femtosecond Laser Chromatic Confocal Probe. *Appl. Sci.* **2019**, *9*, 454. [[CrossRef](#)]
3. Xiong, X.; Shimizu, Y.; Chen, X.; Matsukuma, H.; Gao, W. Uncertainty Evaluation for Measurements of Pitch Deviation and Out-of-Flatness of Planar Scale Gratings by a Fizeau Interferometer in Littrow Configuration. *Appl. Sci.* **2018**, *8*, 2539. [[CrossRef](#)]
4. Sieber, I.; Yi, A.; Gengenbach, U. Metrology Data-Based Simulation of Freeform Optics. *Appl. Sci.* **2018**, *8*, 2338. [[CrossRef](#)]
5. Chen, L.; Liang, C. Novel Boundary Edge Detection for Accurate 3D Surface Profilometry Using Digital Image Correlation. *Appl. Sci.* **2018**, *8*, 2541. [[CrossRef](#)]
6. Chen, L.; Duong, D.; Chen, C. Optical 3-D Profilometry for Measuring Semiconductor Wafer Surfaces with Extremely Variant Reflectivities. *Appl. Sci.* **2019**, *9*, 2060. [[CrossRef](#)]
7. Sun, T.; Zheng, W.; Yu, Y.; Yan, K.; Asundi, A.; Valukh, S. Algorithm for Surfaces Profiles and Thickness Variation Measurement of a Transparent Plate Using a Fizeau Interferometer with Wavelength Tuning. *Appl. Sci.* **2019**, *9*, 2349. [[CrossRef](#)]

8. Chen, D.; Peng, J.; Valyukh, S.; Asundi, A.; Yu, Y. Measurement of High Numerical Aperture Cylindrical Surface with Iterative Stitching Algorithm. *Appl. Sci.* **2018**, *8*, 2092. [[CrossRef](#)]
9. Li, Y.; Ma, Y.; Tao, Y.; Hou, Z. Innovative Methodology of On-Line Point Cloud Data Compression for Free-Form Surface Scanning Measurement. *Appl. Sci.* **2018**, *8*, 2556. [[CrossRef](#)]
10. Chen, C.; Chen, X.; Shi, Y.; Gu, H.; Jiang, H.; Liu, S. Metrology of Nanostructures by Tomographic Mueller-Matrix Scatterometry. *Appl. Sci.* **2018**, *8*, 2583. [[CrossRef](#)]
11. Jiang, H.; Ma, Z.; Gu, H.; Chen, X.; Liu, S. Characterization of Volume Gratings Based on Distributed Dielectric Constant Model Using Mueller Matrix Ellipsometry. *Appl. Sci.* **2019**, *9*, 698. [[CrossRef](#)]
12. Li, J.; Yan, F.; Chen, L.; Yang, J.; Wan, Y. Measurement Vertex Position of a Large Aspheric Mirror. *Appl. Sci.* **2019**, *9*, 1038. [[CrossRef](#)]
13. Liu, W.; Hu, P.; Fan, K. Comparison of Current Five-Point Cylindricity Error Separation Techniques. *Appl. Sci.* **2018**, *8*, 1946. [[CrossRef](#)]
14. Li, R.; Wang, P.; Li, D.; Fan, K.; Liu, F.; Chen, L.; Huang, Q. Precision Manufacturing of Patterned Beryllium Bronze Leaf Springs via Chemical Etching. *Appl. Sci.* **2018**, *8*, 1476. [[CrossRef](#)]
15. Bian, X.; Cui, J.; Lu, Y.; Tan, J. Ultraprecision Diameter Measurement of Small Holes with Large Depth-To-Diameter Ratios Based on Spherical Scattering Electrical-Field Probing. *Appl. Sci.* **2019**, *9*, 242. [[CrossRef](#)]
16. Zhao, N.; Sun, C.; Wang, P. Calibration Method of Orthogonally Splitting Imaging Pose Sensor Based on General Imaging Model. *Appl. Sci.* **2018**, *8*, 1399. [[CrossRef](#)]
17. Yang, T.; Zhao, Q.; Wang, X.; Zhou, Q. Sub-Pixel Chessboard Corner Localization for Camera Calibration and Pose Estimation. *Appl. Sci.* **2018**, *8*, 2118. [[CrossRef](#)]
18. Liu, X.; Zhang, H.; Peng, K.; Tang, Q.; Chen, Z. A High Precision Capacitive Linear Displacement Sensor with Time-Grating that Provides Absolute Positioning Capability Based on a Vernier-Type Structure. *Appl. Sci.* **2018**, *8*, 2419. [[CrossRef](#)]
19. He, Y.; Zhu, L.; Sun, G.; Yu, M.; Dong, M. Design, Measurement and Shape Reconstruction of Soft Surgical Actuator Based on Fiber Bragg Gratings. *Appl. Sci.* **2018**, *8*, 1773. [[CrossRef](#)]
20. Cai, Y.; Lou, Z.; Ling, S.; Liao, B.; Fan, K. Development of a Compact Three-Degree-of-Freedom Laser Measurement System with Self-Wavelength Correction for Displacement Feedback of a Nanopositioning Stage. *Appl. Sci.* **2018**, *8*, 2209. [[CrossRef](#)]
21. Ren, G.; Qu, X.; Ding, S. A Real-Time Measurement Method of Air Refractive Index Based on Special Material Etalon. *Appl. Sci.* **2018**, *8*, 2325. [[CrossRef](#)]
22. Zhu, K.; Chen, H.; Zhang, S.; Shi, Z.; Wang, Y.; Tan, Y. Frequency-Shifted Optical Feedback Measurement Technologies Using a Solid-State Microchip Laser. *Appl. Sci.* **2019**, *9*, 109. [[CrossRef](#)]
23. Xiong, S.; Yang, Z.; Leo, L.; Wu, G. Calculating the Effective Center Wavelength for Heterodyne Interferometry of an Optical Frequency Comb. *Appl. Sci.* **2018**, *8*, 2465. [[CrossRef](#)]
24. Gao, S.; Ji, R.; Li, Y.; Liu, C.; Shi, J.; Pan, Y.; Zhou, W. Compensation of Frequency Drift in Frequency-Sweep Polarization-Modulation Ranging System. *Appl. Sci.* **2019**, *9*, 1243. [[CrossRef](#)]
25. Xiong, X.; Qu, X.; Zhang, F. Error Correction for FSI-Based System without Cooperative Target Using an Adaptive Filtering Method and a Phase-Matching Mosaic Algorithm. *Appl. Sci.* **2018**, *8*, 1954. [[CrossRef](#)]
26. Chang, D.; Xing, X.; Hu, P.; Wang, J.; Tan, J. Double-Diffracted Spatially Separated Heterodyne Grating Interferometer and Analysis on its Alignment Tolerance. *Appl. Sci.* **2019**, *9*, 263. [[CrossRef](#)]
27. Li, X.; Lu, H.; Zhou, Q.; Wu, G.; Ni, K.; Wang, X. An Orthogonal Type Two-Axis Lloyd's Mirror for Holographic Fabrication of Two-Dimensional Planar Scale Gratings with Large Area. *Appl. Sci.* **2018**, *8*, 2283. [[CrossRef](#)]
28. Li, X.; Shi, Y.; Xiao, X.; Zhou, Q.; Wu, G.; Lu, H.; Ni, K. Design and Testing of a Compact Optical Prism Module for Multi-Degree-of-Freedom Grating Interferometry Application. *Appl. Sci.* **2018**, *8*, 2495. [[CrossRef](#)]
29. Cheng, Y.; Wang, Z.; Chen, X.; Li, Y.; Li, H.; Li, H.Y.; Wang, H. Evaluation and Optimization of Task-oriented Measurement Uncertainty for Coordinate Measuring Machines Based on Geometrical Product Specifications. *Appl. Sci.* **2019**, *9*, 6. [[CrossRef](#)]
30. Gařka, A.; Gařka, P.; Gruza, M.; Sladek, J. Selection of Optimal Path Control Algorithms for Probe Heads Used on Five-Axis Measuring Systems. *Appl. Sci.* **2018**, *8*, 2455. [[CrossRef](#)]
31. Wu, N.; Zhao, W.; Wang, X.; Tao, Y.; Hou, Z. A Novel Design of Through-Hole Depth On-Machine Optical Measuring Equipment for Automatic Drilling and Riveting. *Appl. Sci.* **2018**, *8*, 2671. [[CrossRef](#)]

32. Hsieh, T.; Chen, P.; Jywe, W.; Chen, G.; Wang, M. A Geometric Error Measurement System for Linear Guideway Assembly and Calibration. *Appl. Sci.* **2019**, *9*, 574. [[CrossRef](#)]
33. Bao, C.; Feng, Q.; Li, J. Simultaneous Measurement Method and Error Analysis of the Six Degrees-of-Freedom Motion Errors of a Rotary Axis. *Appl. Sci.* **2018**, *8*, 2232. [[CrossRef](#)]
34. Wang, S.; Lee, C.; Gunawan, H.; Yeh, C. An Accuracy-Efficiency-Power Consumption Hybrid Optimization Method for CNC Milling Process. *Appl. Sci.* **2019**, *9*, 1495. [[CrossRef](#)]
35. Wang, Y.; Li, H.; Wan, B.; Zhang, X.; Shan, G. Obtaining Vital Distances Using Wearable Inertial Measurement Unit for Real-Time, Biomechanical Feedback Training in Hammer-Throw. *Appl. Sci.* **2018**, *8*, 2470. [[CrossRef](#)]
36. Qiu, Z.; Li, H.; Hu, W.; Wang, C.; Liu, J.; Sun, Q. Real-Time Tunnel Deformation Monitoring Technology Based on Laser and Machine Vision. *Appl. Sci.* **2018**, *8*, 2579. [[CrossRef](#)]
37. Shi, S.; Yang, L.; Lin, J.; Long, C.; Deng, R.; Zhang, Z.; Zhu, J. Dynamic Measurement Error Modeling and Analysis in a Photoelectric Scanning Measurement Network. *Appl. Sci.* **2019**, *9*, 62. [[CrossRef](#)]
38. Wang, M.; Zhang, X.; Tang, W.; Wang, J. A Structure for Accurately Determining the Mass and Center of Gravity of Rigid Bodies. *Appl. Sci.* **2019**, *9*, 2532. [[CrossRef](#)]
39. Yang, S.; Gao, Z.; Ruan, H.; Gao, C.; Wang, X.; Sun, X.; Wen, X. Non-Contact and Real-Time Measurement of Kolsky Bar with Temporal Speckle Interferometry. *Appl. Sci.* **2018**, *8*, 808. [[CrossRef](#)]
40. Zhang, X.; Mao, H.; Tang, W. Demodulation Technique Based on Laser Interference for Weak Photo-Acoustic Signals on Water Surface. *Appl. Sci.* **2018**, *8*, 2423. [[CrossRef](#)]
41. Qiu, W.; Ma, L.; Wang, H.; Liang, R.; Zhao, Y.; Zhou, Y. Experimental Analyses on Multiscale Structural and Mechanical Properties of  $\epsilon$ -Si/GeSi/C-Si Materials. *Appl. Sci.* **2018**, *8*, 2333. [[CrossRef](#)]




© 2019 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).



Article

# A Structure for Accurately Determining the Mass and Center of Gravity of Rigid Bodies

Meibao Wang , Xiaolin Zhang \*, Wenyan Tang and Jun Wang

School of Instrumentation Science and Engineering, Harbin Institute of Technology, Harbin 150001, China; meibao\_wang@163.com (M.W.); tangwy@hit.edu.cn (W.T.); wang\_jun@hit.edu.cn (J.W.)

\* Correspondence: zhangxiaolin@hit.edu.cn; Tel.: +86-139-3626-1257

Received: 24 May 2019; Accepted: 19 June 2019; Published: 21 June 2019



**Abstract:** Measuring the mass and Center of Gravity (CG) of rigid bodies with a multi-point weighing method is widely used nowadays. Traditional methods usually include two parts with a certain location, i.e., a fixed platform and a mobile platform. In this paper, a novel structure is proposed to adjust the mobile platform for eliminating side forces which may load on the load cells. In addition, closed-form equations are formulated to evaluate the performance of the structure, and transformation matrices are used to estimate the characteristics of the structure. Simulation results demonstrate that repeatability of the proposed structure is higher than the traditional one and there are no side forces. Moreover, the measurement results show that the relative error of mass was within 0.05%, and the error of CG was within  $\pm 0.3$  mm. The structure presented in this paper provides a foundation for practical applications.

**Keywords:** mass; center of gravity; side force; load cell; mechanical structure; modeling

## 1. Introduction

Mass and Center of Gravity (CG) are prerequisite parameters when designing the dynamic performance of an aerospace vehicle, car, etc. [1,2]. When the shape of the vehicle is complicated, it is necessary to obtain mass and CG by experiments [3]. The mass can be achieved by weighing machines, for instance, load cells. There are several methods to measure the CG, which can be roughly classified into three categories as follows: (1) static method such as multi-point weighing method or unbalance moment method [1,4,5], (2) dynamic method such as spin balance method or inverted torsion pendulum method [6–9], and (3) other methods such as the photogrammetric technique for determination of the CG for large scale objects [10].

Each of these methods has its own strengths and weaknesses. For instance, the multi-point weighing method, which uses three or more load cells to support a test platform, has been widely used, and the reason is that this method can measure the mass and CG simultaneously. This method only depends on gravity force acting through the CG, and CG location is calculated by static equilibrium from the reaction forces on the load cells [1,4]. Due to these reasons, it is the most suitable method for some massive objects [4], and it is also the cheapest automatic system.

However, there are shortcomings in the multi-point weighing method, for example, a load cell is a spring and has a designed loading axis. In addition, the generated force against the load cell outside of this axis can result in errors and may shorten the operating life of the load cell [11–14]. Besides, the measurement accuracy of CG is related to the position of the loading force.

There are several types of load cells or bears to avoid incorrect loading, such as self-centering pendulum load cell, pendulum bearing, and pendulum supports. These load cells automatically guide the superstructure back to its original position when the load is introduced with lateral

displacement [15]. However, these methods are usually used for a single load cell. Towards the multi-point weighing method where three or more load cells need to be coupled, no related literatures report the measurement structure.

In order to achieve high accuracy when measuring mass and CG, optimizing the mechanical structure to make the force load on the designed axis is necessary. This paper focuses on designing a structure used in a multi-point weighing method to avoid lateral forces and get high repeatability. Moreover, mathematical modeling of the structure is further established to analyze the proposed method. Finally, simulation and experiment results are conducted to validate the effectiveness of the proposed method.

## 2. Background and Related Work

### 2.1. Method for Obtaining the Mass and CG of a Body

As shown in Figure 1, the multi-point weighing method is a static method, which includes three load cells and a support platform. These three load cells are placed on the same horizontal flat and angularly spaced by 120°. This proposed method depends only on gravity force acting through the CG [1,4].

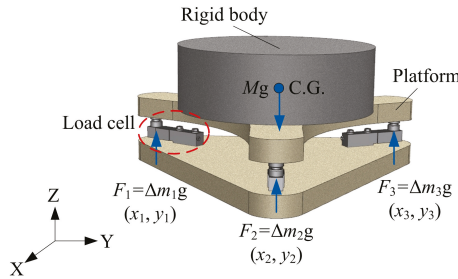


Figure 1. Diagram of the multi-point weighing method.

We record the readings of the three load cells, and the sum of the three load cell readings is the mass of the whole system. Similarly, we also record the mass of the system without the rigid body, and the mass of the rigid body is obtained from the difference of the readings shown in Equation (1).

$$M = \Delta m_1 + \Delta m_2 + \Delta m_3 \quad (1)$$

where:  $M$  is the mass of the rigid body, and  $\Delta m_1, \Delta m_2, \Delta m_3$  are the difference readings of the three load cells.

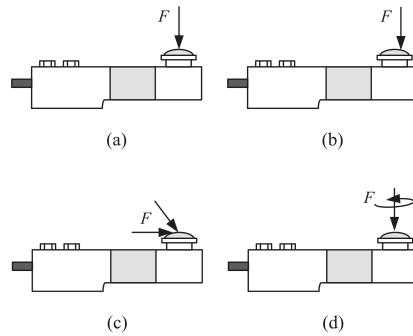
Based on the static equilibrium (first law), the CG location can be obtained by Equation (2).

$$\begin{cases} x = \frac{x_1 \cdot \Delta m_1 + x_2 \cdot \Delta m_2 + x_3 \cdot \Delta m_3}{M} \\ y = \frac{y_1 \cdot \Delta m_1 + y_2 \cdot \Delta m_2 + y_3 \cdot \Delta m_3}{M} \end{cases} \quad (2)$$

According to Equation (1) and Equation (2), there are two factors affecting the accuracy of the results.

(a) Accuracy of the load cell.

The load cell is sensitive to side forces, bending and torsional moments, which should be avoided in the measurement systems [11–16]. Figure 2 shows some examples of correct loading as well as incorrect loading methods on a load cell [13–15].



**Figure 2.** Correct loading on a load cell and some examples of incorrect loading. (a) Central load application; (b) Non-central load application; (c) Non-axial load application or side forces; (d) Moment-loading (torsion) load application.

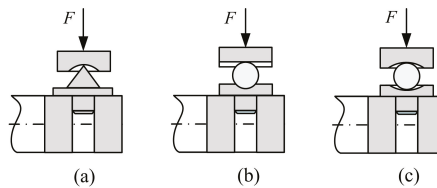
(b) Location of the reaction forces ( $F_1$ ,  $F_2$  and  $F_3$ ) of the load cells.

The load cell always has a designed axis of loading. If the location of the reaction force is out of this designed axis, it will cause errors of the load cells' readings, and errors of the CG results can be calculated according to Equation (2).

To avoid these issues, some efforts are done by designing the structure act on the load cell properly.

### 2.2. Traditional Structure and the Improved Structure

Figure 3 shows some types of structure. (a) Cone joint with conical pan, which is a traditional mounting and it is used for individual load cells, (b) and (c) are types of rocker bearings. The distortion of the load cell's reading is practically absent for rocker bearings, because in this case only slight rolling friction is present instead of a bending stress. However, the horizontal restraint of the rocker bearing is significantly less than fixed bearings. Rocker bearings are recommended if the position of the superstructure only changes horizontally [15].



**Figure 3.** The construction of the traditional structure. (a) Cone joint with conical pan; (b) Ball joint with groove; (c) Ball joint with fillet.

We have researched CG determination for many years, and one traditional method is placing three load cells on the same horizontal flat, as shown in Figure 4. In particular, each load cell has a ball groove and each groove holds a ball, making the center of the ball groove on the designed axis of the load cell. The upper platform has a similar construction. The ball can stay at the center of the ball groove because of the gravity.

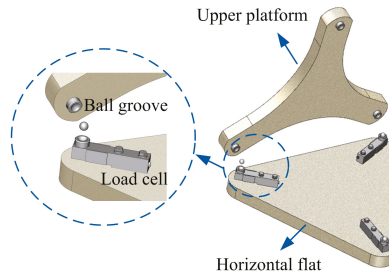


Figure 4. The construction of the traditional structure.

However, due to manufacturing tolerance and imperfect assembly, axis 2 is not the same as axis 1 shown in Figure 5, making it hard to guarantee  $F$  pass through each designed axis of the load cell. Moreover, reducing these errors will increase the costs significantly.

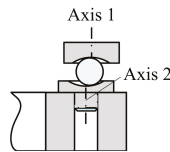


Figure 5. Axes of the ball grooves.

Regarding the above issue, an improved structure is proposed, as shown in Figure 6. Similar to Figure 4, three load cells rest on the base platform, each of them has a ball groove and each groove holds a ball, ensuring the center of the ball groove goes through the designed axis of the load cell. The base platform should be horizontal and this structure should minimize the friction.

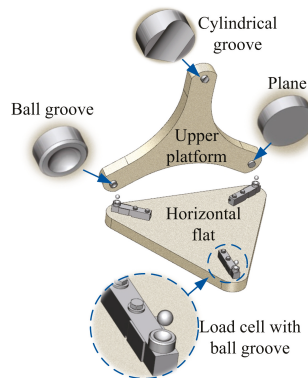


Figure 6. The construction of the new structure.

Specifically, the upper platform has one ball groove, one cylindrical groove and one planar surface. Because of the characteristic of the cylindrical groove and the plane, the upper platform can adjust its position and ensure  $F$  passes through the designed axis of the load cell without side force, even though these grooves may have position errors. A mathematical analysis of the structure will be performed in the next part.

### 3. Parametric Representation of the Structure

Here, for a convenient and simple explanation, the platform with the load cells is called the base-frame, and the upper platform is called the platform-frame [17]. The base-frame ( $O_B, X_B, Y_B, Z_B$ ) and the platform-frame ( $O_P, X_P, Y_P, Z_P$ ) are two right-handed orthonormal coordinate systems with  $(X_B, Y_B, Z_B)$  and  $(X_P, Y_P, Z_P)$  as their bases respectively, as shown in Figure 7. The three centers of ball grooves, which are held by load cells, always rest on the  $X_B Y_B$ -plane. The origin  $O_B$  of the base-platform is the intersection of the bisectors of the three angles of the base frame triangle. Coordinates of the centers of the ball grooves formulate the locations of the ball grooves.

The origin  $O_P$  of the platform-frame is the center of the ball groove shown in Figure 7, and the  $X_P$  axis is parallel with the axis of the cylindrical groove. The plane is parallel with the  $X_P Y_P$ -plane. In the neutral configuration, the platform-frame has a general position and rotation with respect to the base-frame.

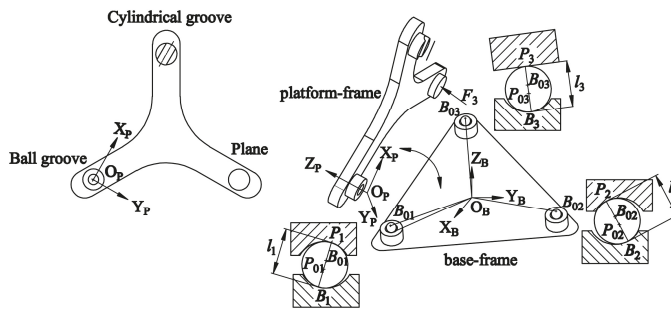


Figure 7. The schematic of the structure.

It is assumed that the external force acting on the platform-frame is known as well as the positions of the load cells, and it is assumed that  $B_i$  ( $i = 1, 2, 3$ ) and  $P_i$  ( $i = 1, 2, 3$ ) are unknowns. The significance of modeling is to analyze whether directions of  $\vec{B}_i P_i$  ( $i = 1, 2, 3$ ) are vertical to the  $X_B Y_B$ -plane, which means that there are no side forces.

To analyze the structure, the coordinates of  $P_i$  ( $i = 1, 2, 3$ ) or  $B_i$  ( $i = 1, 2, 3$ ) should represent the same coordinate system ( $O_B, X_B, Y_B, Z_B$ ), as shown in Figure 8, and  $P_i$  in base-frame ( $O_B, X_B, Y_B, Z_B$ ) is shown in Equation (3):

$$[P_i]_B = T_{BP} + R_{BP} \cdot P_i \quad (3)$$

where  $T_{BP} = [x, y, z]_B^T$  is the relative position vectors between  $O_P$ - $X_P Y_P Z_P$  and  $O_B$ - $X_B Y_B Z_B$ .  $R_{BP}$  is the matrix representation of the rotation between  $O_P$ - $X_P Y_P Z_P$  and  $O_B$ - $X_B Y_B Z_B$ .  $P_1$  is the coordinate of the contact point between the ball and the ball groove;  $P_2$  is the coordinate of the contact point between the ball and the cylindrical groove; and  $P_3$  is the coordinate of the contact point between the ball and the plane in the platform-frame.  $[P_i]_B$  is the coordinates of  $P_i$  ( $i = 1, 2, 3$ ) expressed in  $O_B$ - $X_B Y_B Z_B$ .

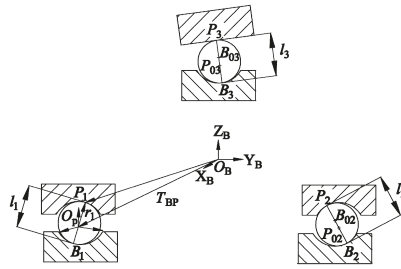


Figure 8. The vector loop between ball and ball groove.

The rotation matrix  $R_{BP}$  is expressed by:

$$R_{BP} = \begin{bmatrix} q_{11} & q_{12} & q_{13} \\ q_{21} & q_{22} & q_{23} \\ q_{31} & q_{32} & q_{33} \end{bmatrix} \tag{4}$$

where:

$$\begin{aligned} q_{11} &= \cos \gamma \cdot \cos \beta \\ q_{12} &= \cos \gamma \cdot \sin \beta \cdot \sin \alpha - \sin \gamma \cdot \cos \alpha \\ q_{13} &= \cos \gamma \cdot \sin \beta \cdot \cos \alpha + \sin \gamma \cdot \sin \alpha \\ q_{21} &= \sin \gamma \cdot \cos \beta \\ q_{22} &= \sin \gamma \cdot \sin \beta \cdot \sin \alpha + \cos \gamma \cdot \cos \alpha \\ q_{23} &= \sin \gamma \cdot \sin \beta \cdot \cos \alpha - \cos \gamma \cdot \sin \alpha \\ q_{31} &= -\sin \beta \\ q_{32} &= \cos \beta \cdot \sin \alpha \\ q_{33} &= \cos \beta \cdot \cos \alpha \end{aligned} \tag{5}$$

$\vec{l}_i$  is the direction of  $\vec{B_iP_i}$  ( $i = 1, 2, 3$ ), which can be described by Equation (6) [17]:

$$\vec{l}_i = [P_i]_B - B_i = (T_{BP} + R_{BP} \cdot P_i) - B_i \tag{6}$$

Ignore the effects of friction, the direction of  $\vec{B_iP_i}$  always passes through the centers of the balls. As a result, the magnitudes of  $\vec{l}_i$  are constant, which equal to the diameter of the balls, as shown in Equation (7):

$$\|\vec{l}_i\| = D \tag{7}$$

$F_i$  are the forces loading on the load cells, and the direction of  $F_i$  expressed in  $O_B-X_B Y_B Z_B$  can be expressed by the unit vector  $\vec{n}_{Bi}$ , as shown in Equation (8):

$$\vec{n}_{Bi} = \frac{\vec{l}_i}{\|\vec{l}_i\|} = \frac{1}{D} \cdot \vec{l}_i \tag{8}$$

While the direction of  $F_i$  expressed in  $O_P-X_P Y_P Z_P$  can be expressed by the unit vector  $\vec{n}_{Pi}$ , as shown in Equation (9):

$$\vec{n}_{Bi} = R_{BP} \cdot \vec{n}_{Pi} \tag{9}$$

The stability of the platform-frame must meet two conditions, namely, the force balance theorem and equilibrium of couples [18], which can be formulated in Equation (10):

$$\begin{bmatrix} \vec{n}_{p1} & \vec{n}_{p2} & \vec{n}_{p3} \\ \vec{r}_1 \times \vec{n}_{p1} & \vec{r}_2 \times \vec{n}_{p2} & \vec{r}_3 \times \vec{n}_{p3} \end{bmatrix} \begin{bmatrix} F_1 \\ F_2 \\ F_3 \end{bmatrix} = \begin{bmatrix} F \\ M \end{bmatrix} \quad (10)$$

where  $F$  and  $M$  are the external force and torque respectively, which act on the platform-frame as shown in Figure 9. Here, the  $\vec{r}_i$  ( $i = 1, 2, 3$ ) are formulated as Equation (11):

$$\vec{r}_i = O_p P_i = P_i(x_i, y_i, z_i) \quad (11)$$

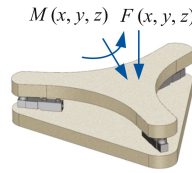


Figure 9. External force diagram of the measuring system.

Ignore the effects of friction; the  $F_i$  always passes through the centers of the ball grooves so that it rests on the base-frame, so do the  $\vec{n}_{Bi}$ . As a result, the coordinates  $B_{0i}$  of the centers of the three balls grooves are calculated by Equation (12):

$$B_{0i} = B_i + R_B \cdot \vec{n}_{Bi} \quad (12)$$

where  $R_B$  are the radiuses of ball grooves, and  $B_i$  are the contact points between the balls and the ball grooves.

Similarly, the orientation of  $F_1$  always passes through the center of the ball groove ( $P_{01x}, P_{01y}, P_{01z}$ ) which lies on the platform-frame, making the center of the ball groove equal to the origin (0, 0, 0) of the  $O_P-X_P Y_P Z_P$ , which can be formulated as Equation (13):

$$\begin{bmatrix} x_{p1} \\ y_{p1} \\ z_{p1} \end{bmatrix} - R_p \begin{bmatrix} n_{p1x} \\ n_{p1y} \\ n_{p1z} \end{bmatrix} = \begin{bmatrix} P_{01x} \\ P_{01y} \\ P_{01z} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \quad (13)$$

where  $(x_{p1}, y_{p1}, z_{p1})$  is the coordinate of  $P_1$ , which is the contact point between the ball and the ball groove;  $R_p$  is the radius of the ball groove;  $(n_{p1x}, n_{p1y}, n_{p1z})$  are the vectors of  $\vec{n}_{p1}$ .

The orientation of the  $\vec{n}_{p2}$  ( $n_{p2x}, n_{p2y}, n_{p2z}$ ) passes through the axis of the cylinder groove, because the axis of the platform ( $X_p$ ) is parallel with the axis of the cylinder groove, as a result, we can get Equation (14):

$$\begin{bmatrix} x_{p2} \\ y_{p2} \\ z_{p2} \end{bmatrix} - R_p \begin{bmatrix} n_{p2x} \\ n_{p2y} \\ n_{p2z} \end{bmatrix} = \begin{bmatrix} x_{p2} \\ 0 \\ 0 \end{bmatrix} \quad (14)$$

where  $(x_{p2}, y_{p2}, z_{p2})$  is the coordinate of  $P_2$ , which is the contact point between the ball and the cylindrical groove;  $R_p$  is the radius of the cylindrical groove.

Besides, the normal vector of the plane is the orientation of  $\vec{n}_{p3}$  ( $n_{p3x}, n_{p3y}, n_{p3z}$ ) and is given by Equation (15):

$$\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} n_{p3x} \\ n_{p3y} \\ n_{p3z} \end{bmatrix} \tag{15}$$

The solution described here uses the geometric model to calculate the relative position and orientation between the base-frame and the platform-frame. Analysis of the unknowns from Equation (3) to Equation (15) will help to analyze the characteristics of the improved structure.

#### 4. Simulation Results and Discussion

A simulation was carried out to demonstrate the feasibility of the proposed method. As illustrated in Figure 10, the solution employs a system of 27 equations. The corresponding unknown variables were solved iteratively by using a nonlinear numerical technique [19]. The inputs to the solver include  $B_{0i}, P_{01}, P_{02}, \vec{n}_{p3}, \vec{l}_i, F$ , and  $M$ . The outputs of the algorithm include  $B_i, P_i, F_i, T_{BP}$  and  $R_{BP}$ .

Known Variables	Number of Equations		Unknown Variables	Number of Equations
$B_{0i} (i=1, 2, 3)$	9	System of 27 Equations ⇒ Nonlinear Equation Solver	$B_i (x_{B_i}, y_{B_i}, z_{B_i}) (i=1, 2, 3)$	9
$P_{01} (0, 0, 0)$	3		$P_i (x_{P_i}, y_{P_i}, z_{P_i}) (i=1, 2, 3)$	9
$P_{02} (x_{P2}, 0, 0)$	3		$F_i (i=1, 2, 3)$	3
$\vec{n}_{p3} = (0, 0, 1)^T$	3		$T_{BP} = [x, y, z]_B^T$	3
$\ \vec{l}_i\  (i=1,2,3)$	3		$R_{BP} = [\alpha, \beta, \gamma]_B^T$	3
$F (x, y, z)$	3			
$M (x, y, z)$	3			

Figure 10. The solution for the proposed mathematical model.

With the proposed mathematical modeling, we analyze the traditional structure. Setting the offsets between axis 1 and axis 2, as shown in Figure 5, to 0 mm, 0.05 mm, 0.10 mm, . . . , 1.00 mm, the side forces loading on the three load cells were obtained and shown in Figure 11. The offsets of the loading axes between the designed axes are listed in Figure 12. The inputs are listed in Appendix B.

From Figures 11 and 12, we can see that the traditional structure provides high accuracy without lateral forces in the ideal case. However, when taking the manufacturing or assembling errors into account, i.e., making the axis of the ball grooves on the base-frame shift the axis of ball grooves on the platform-frame, the results show that the offsets of the loading axes between the designed axes increased, so do the side forces.

The improved structure was also analyzed. With the offset or not, the side forces loading on the three load cells were always less than 0.01 N, and the offsets of the loading axes between the designed axes were less than 0.001 mm, which means that the accuracy of the improved method is higher than the traditional one.



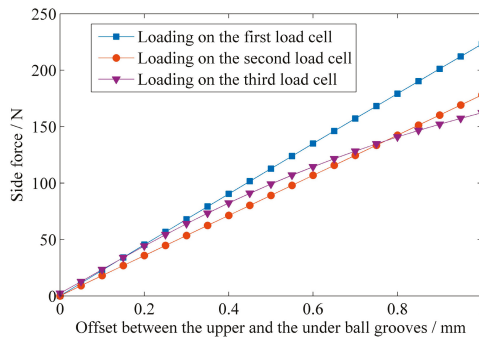


Figure 11. The simulation results of the side forces.

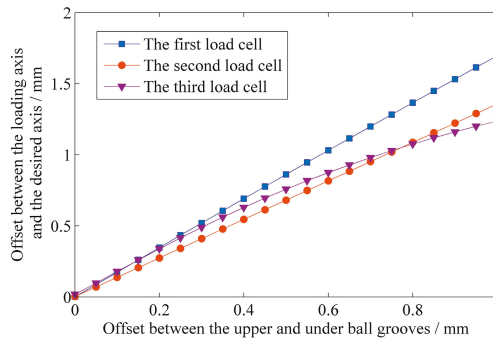


Figure 12. The simulation result of offsets between the loading axes and the designed axes.

### 5. Physical Experiments

Mass and CG of a sample were measured 20 times with the traditional structure and the improved one respectively, as shown in Figure 13, and the standard deviations of the data are listed in Table 1, which means that the new structure has higher repeatability. The accuracy of the load cell is 0.02%, and the range of the load cell is (0–220) kg.

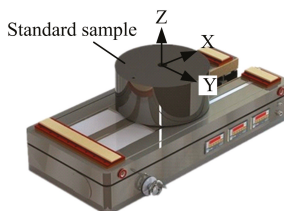


Figure 13. Experiment platform for measuring mass and CG.

**Table 1.** The standard deviations of the mass and the CG.

	Deviation of the Mass (kg)	Deviations of the CG (mm)	
		Along Axis of X	Along Axis of Y
The traditional structure	0.06	0.32	0.33
The improved structure	0.01	0.06	0.10

Mass and CG of a large-sized standard sample was measured by a system with the new structure as shown in Figure 14. Major dimensions of the system are shown in Table 2, and the specification and design requirements of the system are shown in Table 3.



**Figure 14.** Experimental platform for measuring the mass and CG.

**Table 2.** Major dimensions of the system.

Item	Dimensions
Height of the device	1680 mm
Length of the device	2850 mm
Width of the device	2140 mm
Accuracy of the load cells	0.03% (F.S.)

**Table 3.** Specification and design requirements of the system.

Item	Dimensions
Mass properties to be measured	Mass, CG along three orthogonal axes (X, Y, Z)
Range of CG measurement	(0–3000) mm
CG accuracy	±0.3 mm
Range of mass measurement	(500–1000) kg
Mass accuracy	±0.05%

The mass and CG of this sample have been calibrated by a metrological authority. According to the calibration result, the mass is 779.43 kg, and the offset of CG from the centroid of the sample is less than 0.05 mm. The standard sample was placed at 20 different positions of the measurement system. The mass errors measured by the system are shown in Figure 15, and the CG errors are shown in Figure 16.

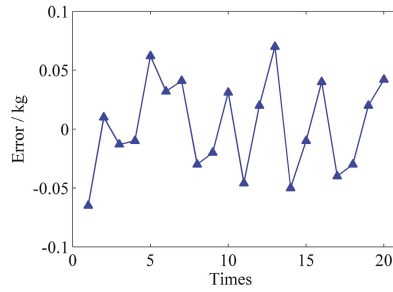


Figure 15. The measurement error of mass.

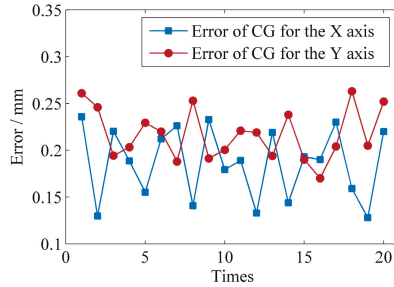


Figure 16. The measurement error of the CG.

According to the results of the mass measurement, the standard deviation of mass is 0.05 kg, and the maximum error is  $-0.07$  kg. Meanwhile, the standard deviation of the CG along X axis is 0.04 mm, and the standard deviation of the CG along Y axis is 0.05 mm, which means that the improved structure can be used to measure the mass and CG with high accuracy.

## 6. Conclusions and Outlook

An improved structure for measuring the mass and CG was proposed in this paper. In addition, the analysis of the structure was presented, which includes the position and orientation of the platform-frame. Simulation and experimental results illustrate that the repeatability of the proposed structure is higher than the traditional one. In future work, the deflection deformation of the balls and the friction between the balls and the grooves will be considered, and the optimal radiuses of the grooves and the balls will be considered too.

**Author Contributions:** Conceived the Method and Edited the Manuscript, X.Z.; Conceived the Method, Designed, Performed the Experiments and Wrote the Paper M.W. Conceived the Method and Designed the Experiments W.T. Performed the Experiments, J.W.

**Funding:** This study was funded by Shanghai Aerospace Science and Technology Innovation Foundation (SAST2015029).

**Acknowledgments:** The authors would like to thank the associate editor and the reviewers for their helpful comments, which improved the quality of this paper.

**Conflicts of Interest:** The authors declare no conflicts of interest.

Appendix A

Table A1. Variables and their description.

Variable ( $i = 1, 2, 3$ )	Description	Variable ( $i = 1, 2, 3$ )	Description
$B_{0i} (x_i, y_i, z_i)$	The coordinates of the ball grooves in $O_B-X_B Y_B Z_B$	$T_{BP} = [x, y, z]_B^T$	The relative position vectors between $O_P-X_P Y_P Z_P$ and $O_B-X_B Y_B Z_B$
$R_B$	The radiuses of the ball grooves on base-frame	$R_{BP} = [\alpha, \beta, \gamma]_B^T$	The rotation between $O_P-X_P Y_P Z_P$ and $O_B-X_B Y_B Z_B$
$O_B$	The origin of $O_B-X_B Y_B Z_B$	$\vec{r}_i$	The point of the force act upon platform-frame
$O_B-X_B Y_B Z_B$	The coordinate system of the base-frame	$F_i$	The forces between balls and grooves
$B_i$	The contact points between the balls and the ball grooves	$F(x, y, z)$	External force
$P_{01}$	The center of the ball groove on the platform-frame	$M(x, y, z)$	External torque
$R_P$	The radius of the ball groove on the platform-frame	$D$	The diameter of the balls
$O_P$	The origin of the platform-frame	$\vec{l}_i$	the direction of $F_i$
$O_P-X_P Y_P Z_P$	The coordinate system of the platform-frame	$\vec{n}_{Bi}$	The unit vector of $\vec{l}_i$
$P_i$	The coordinates of the contact points between the balls and the grooves on the platform-frame	$\vec{n}_{Pi}$	$\vec{n}_{Bi}$ expressed in $O_B-X_B Y_B Z_B$
$[P_i]_B$	The coordinates of $P_i$ expressed in $O_B-X_B Y_B Z_B$		

Appendix B

Table A2. Inputs and their values.

$B_{01} = (400, 0, 0)^T$	$P_{02} = (x_{p2}, 0, 0)^T$
$B_{02} = (-200, 346.41, 0)^T$	$\vec{n}_{p3} = (0, 0, 1)^T$
$B_{03} = (-200, -346.41, 0)^T$	$\ \vec{l}_i\  = 20$
$P_{01} = (0, 0, 0)^T$	$F = (0, 0, 600)^T$
$M = (0, 0, 0)^T$	$R_B = 15$
$R_P = 15$	

References

- Mondal, N.; Acharyya, S.; Saha, R.; Sanyal, D.; Majumdar, K. Optimum design of mounting components of a mass property measurement system. *Measurement* **2016**, *78*, 309–321. [CrossRef]
- Previati, G.; Gobbi, M.; Mastinu, G. Measurement of the mass properties of rigid bodies by means of multi-filar pendulums—Influence of test rig flexibility. *Mech. Syst. Signal Process.* **2019**, *121*, 31–43. [CrossRef]
- Tang, L.; Shangguan, W.B. An improved pendulum method for the determination of the center of gravity and inertia tensor for irregular-shaped bodies. *Measurement* **2011**, *44*, 1849–1859. [CrossRef]
- Gopinath, K.; Raghavendra, K.; Behera, M.K.; Rao, E.V.; Umakanth, M.; Gopinath, S. Product design aspects for design of accurate mass properties measurement system for aerospace vehicles. *Appl. Mech. Mater.* **2012**, *110–116*, 4712–4718. [CrossRef]
- Fabbri, A.; Molari, G. Static measurement of the centre of gravity height on narrow-track agricultural tractors. *Biosyst. Eng.* **2004**, *87*, 299–304. [CrossRef]
- Liu, Y.; Liang, B.; Xu, W.; Wang, X. A method for measuring the inertia properties of a rigid body using 3-URU parallel mechanism. *Mech. Syst. Signal Process.* **2019**, *123*, 174–191. [CrossRef]
- Brancati, R.; Russo, R.; Savino, S. Method and equipment for inertia parameter identification. *Mech. Syst. Signal Process.* **2010**, *24*, 29–40. [CrossRef]

8. Hou, Z.C.; Lu, Y.N.; Lao, Y.X.; Liu, D. A new trifilar pendulum approach to identify all inertia parameters of a rigid body or assembly. *Mech. Mach. Theory* **2009**, *44*, 1270–1280. [CrossRef]
9. Gobbi, M.; Mastinun, G.; Previati, G. A method for measuring the inertia properties of rigid bodies. *Mech. Syst. Signal Process.* **2011**, *25*, 305–318. [CrossRef]
10. NASA Technical Report Server. Photogrammetric Technique for Center of Gravity Determination. AIAA. Available online: <https://ntrs.nasa.gov/search.jsp?R=20120008800> (accessed on 23 April 2012).
11. NASA Technical Report Server. Training for a New Spacecraft Center of Gravity. Available online: <https://nasasearch.nasa.gov/search?query=center+of+garvity&affiliate=nasa&utf8=%E2%9C%93> (accessed on 1 May 2013).
12. Peterson, E. Control of Side-Loads and Errors When Weighing on Jacks. In Proceedings of the 59th Annual International Conference on Mass Properties Engineering, St. Louis, MO, USA, 5–7 June 2000.
13. OMEGA. Installing a Load Cell: Best Practices. Available online: <https://www.omega.com/technical-learning/load-cell-installation.html> (accessed on 1 September 2018).
14. OMEGA. Load Cell Installation Guide. Available online: <https://www.omega.co.uk/technical-learning/load-cell-installation.html> (accessed on 7 September 2016).
15. HBM. Load Application in Load Cells—Tips for Users. Available online: <https://www.hbm.com/en/3377/load-application-in-load-cells/> (accessed on 30 April 2018).
16. Wei, T.P.; Yang, X.X.; Yao, J.H.; Xu, H. The additional side force on the force transducer in the combinatorial load cell. *AMM* **2014**, *541–542*, 1327–1332. [CrossRef]
17. Afzali-Far, B.; Andersson, A.; Nilsson, K.; Lidström, P. Dynamic isotropy in 6-DOF kinematically constrained platforms by three elastic nodal joints. *Precis. Eng.* **2016**, *45*, 342–358. [CrossRef]
18. Schmiechen, P.; Slocum, A.H. Analysis of kinematic systems: A generalized approach. *Precis. Eng.* **1996**, *19*, 11–18. [CrossRef]
19. Barraja, M.; Vallance, R.R. Tolerancing kinematic couplings. *Precis. Eng.* **2005**, *29*, 101–112. [CrossRef]



© 2019 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).

Article

# Algorithm for Surfaces Profiles and Thickness Variation Measurement of a Transparent Plate Using a Fizeau Interferometer with Wavelength Tuning

Tao Sun <sup>1</sup>, Weiwei Zheng <sup>1</sup>, Yingjie Yu <sup>1,\*</sup>, Ketao Yan <sup>1</sup>, Anand Asundi <sup>2</sup> and Sergiy Valukh <sup>3</sup>

<sup>1</sup> Lab of Applied Optics and Metrology, Department of Precision Mechanical Engineering, Shanghai University, Shanghai 200072, China; sun\_tao07@shu.edu.cn (T.S.); wwdn951@shu.edu.cn (W.Z.); ketao\_yan@163.com (K.Y.)

<sup>2</sup> School of Mechanical and Aerospace Engineering, Nanyang Technological University, Singapore 639798, Singapore; anand.asundi@pmail.ntu.edu.sg

<sup>3</sup> Department of Physics, Chemistry and Biology, Linköping University, SE-581 83 Linköping, Sweden; serva@ifm.liu.se

\* Correspondence: yingjieyu@staff.shu.edu.cn; Tel.: +86-021-6613-0822

Received: 8 April 2019; Accepted: 4 June 2019; Published: 7 June 2019



**Featured Application:** The proposed method is used to separate the overlapped interferograms formed by parallel interfaces, specifically, surfaces of a transparent plate, and to obtain the front and rear surface profiles, as well as the thickness, simultaneously.

**Abstract:** An interferogram obtained from a transparent plate contains information on the profiles of both surfaces and on the thickness variation. The present work is devoted to the processing of interferograms of this type. The processing technique is based on a 36-step algorithm developed by the authors for characterization of transparent plates having approximately equal reflections from both sides. The algorithm utilizes weighted multi-step phase shifting that enables one not only separately to extract the front and rear surface profiles together with the thickness variation of the tested plate but also to suppress the coupling errors between the higher harmonics and phase-shift deviation. The proposed measuring method was studied on a wavelength tunable Fizeau interferometer. The tested sample had an optical thickness and surface profile deviations equal to 0.51  $\mu\text{m}$ , 1.38  $\mu\text{m}$  and 0.89  $\mu\text{m}$ , respectively. According to the results obtained using 10 repeated measurements, the root mean square (RMS) errors for determining both surface profiles did not exceed 1.5 nm. Experimental results show that the setup and presented 36-step algorithm are suitable for the measurement of a transparent plate of arbitrary thickness.

**Keywords:** Fizeau interferometry; wavelength tuning; separation of interferograms; characterization of a transparent plate; 36-step algorithm

---

## 1. Introduction

Being an optical element, transparent plates play an essential role in a large variety of optical components and applications [1,2]. The front and rear surface profiles together with the optical thickness are among the fundamental characteristics of a transparent plate. Precise control and monitoring of these characteristics by non-destructive and non-contact methods for samples with different sizes and thickness are often important from a technological point of view and are associated with the quality improvement of finished products. Several optical techniques have been developed for the determination of the thickness. Most of them are based on principles of ellipsometry and interferometry [3,4]. For example, these include the excess fraction method, digital phase-measuring

interferometry [5], Fourier transform profilometry [6], random tilt phase-shifting [7], and 3D profilometry based on modulation measurement [8]. White-light interferometry and confocal microscopy have also been used for optical thickness and profile measurement of transparent plates [9–12].

Depending on the spectrum of the light source used, the methods applied to measure thickness variation can be classified as monochromatic or spectral. The experimental setups of some of those methods include both a monochromatic laser and a spectral halogen lamp [13]. One of the disadvantages of using confocal microscopy is the limitation in the diameter of an observing aperture.

In addition, it is not suitable for measuring the distributions of the surface profiles and the thickness because this method is based on an assumption that a sample has a uniform thickness, it measures the sample on a point by point basis. White-light interferometry is time-consuming and involves a limited aperture. Moreover, the determination of large thicknesses requires a light source possessing a large coherence length.

Wavelength tuning interferometry (WTI) [14–25] is an effective method for application in surface metrology, which is primarily limited to front-surface reflected optical element measurements. WTI has also been exploited for separating overlapped interference signals formed by the reference surface of the interferometer used and both surfaces of the measured transparent plate in frequency domain. However, the WTI method [16] provides an air-gap length (the distance between the reference surface at Fizeau geometry to the front surface, Figure 1) which is less than the thickness of the sample. In order to facilitate its application in actual industrial measurements, the air-gap length must be large. In this case, it becomes difficult to measure a sample with a thickness less than 10 mm in an actual measurement process that uses Fizeau interferometer. Whereas a conventional phase-shifting interferometer [26] requires a stable phase detection method for only a single signal obtained from the surface to be tested. The challenge for a wavelength tuning interferometer is an extraction of information in several signal frequencies, when it is necessary to suppress the noise from the neighboring frequency signals.

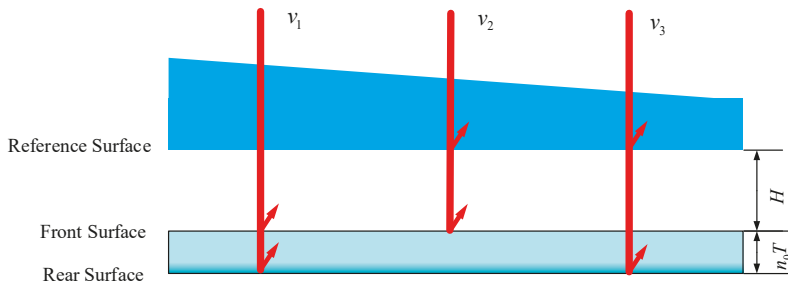


Figure 1. Diagram of the reflected rays forming the interferogram.

In this study, we report a developed interferometric phase-shifting method based on a 36-step algorithm to compensate for the phase-shift miscalibration and to suppress the signal contributions from other surfaces. Our algorithm for data processing realizes the separation of multi-surface superposition information. The method utilizes the spectral stretching of the Discrete Fourier Transform function and is meant to be used for measurements of transparent plates with arbitrary thicknesses. The experimental setup is based on the Fizeau interferometer with wavelength tuning.

## 2. Materials and Methods

When a beam of light is incident on a transparent plate, an interference pattern is obtained due to multiple reflections from the plate's surfaces. In order to measure the profile of the transparent plate, it

is necessary to extract from the registered interference pattern information caused by each surface [5]. The intensity distribution in the interferogram formed by transparent plate can be expressed as:

$$I(x, y, k) = A_0 + \sum_{n=1}^{\infty} A_n \cos(\varphi_n(x, y) + 2\pi v_n k), \tag{1}$$

where  $I(x, y, k)$  is the intensity of the  $k$ th record image in the coordinate  $(x, y)$ , where the coordinate plane  $x, y$  is parallel to the measured interfaces.  $A_0$  is the intensity of the background, and  $A_n$  is a contrast associated with intensity of the  $n$ th harmonic.  $\varphi_n$  is the phase of the  $n$ th harmonic components, which corresponds to the path difference between the sample surface and the reference surface of the interferometer. The accuracy (peak-to-valley value) is less than one twentieth of the wavelength,  $\lambda/20$ , generally. The modulation frequency is  $v_n$  and is proportional to the optical path difference (OPD) of each pair of interfering fringe.

For the signal generated by the interference due to the reflection from the reference and front-and-rear surfaces, the frequency  $v$  is given by:

$$v = \frac{2OPD}{\lambda_0^2} \cdot \frac{d\lambda}{dk}, \tag{2}$$

where  $\lambda_0$  is the central wavelength of the laser source, and  $\frac{d\lambda}{dk}$  is the wavelength-tuning rate per time unit.

In Figure 1, each frequency in the measurement can be described as follows

$$\begin{aligned} v_1 &= \frac{4n_0T}{\lambda_0^2} \cdot \frac{d\lambda}{dk} \\ v_2 &= \frac{4H}{\lambda_0^2} \cdot \frac{d\lambda}{dk} \\ v_3 &= \frac{4(n_0T+H)}{\lambda_0^2} \cdot \frac{d\lambda}{dk} \end{aligned} \tag{3}$$

where  $n_0$  is the refractive index of the measured components when  $\lambda = \lambda_0$  and  $T$  is the thickness of the transparent plate.  $H$  is length of the air-gap (the distance between the reference and the front surface).

Table 1 shows the relationship between the frequency information of the interference signals. We can detect the useful signals (front-and-rear surfaces and thickness information) from the superimposed interference patterns by setting the appropriate ratio  $M$  defined as  $M = H/(n_0T)$ . The value of  $M$  determines the application for the measurement of the transparent plate of arbitrary thickness. It is possible to develop a spectral stretching approach for obtaining the appropriate  $M$  value for a transparent plate of any thickness in the actual measurement process using the Fizeau interferometer with wavelength tuning.

**Table 1.** Frequency information of each measuring surface.

No.	$v_n/v_1$	Remark
$v_1$	1	thickness
$v_2$	$M$	Front interface
$v_3$	$M + 1$	Rear interface



The phase distribution can be calculated with the phase-shifting algorithm [27]. Consider a  $K$ -step phase-shifting algorithm, where the phase-steps are separated by  $K - 1$  equal intervals of  $\delta = 2\pi/N$  and  $N$  is an integer. Phase-detection formula is written as [27]:

$$\varphi(x, y) = \arctan \frac{\sum_{k=1}^K b_k I(x, y, k)}{\sum_{k=1}^K a_k I(x, y, k)}, \tag{4}$$

where  $a_k$  and  $b_k$  are the weighed values determined by the window function.  $\varphi(x, y)$  is the resulting phase, and  $I(x, y, k)$  given by Equation (1) is the intensity of the  $k$ th recorded image in the coordinate  $(x, y)$ .

According to the signal analysis theory [16], we developed a 36-step algorithm as an example and the phase-interval is  $\delta = \pi/4$ .

In measuring the wave-front of the front surface,

$$\begin{aligned} a_k &= \frac{1}{4}w(k) \bullet \cos\left[\frac{(k-22) \bullet M}{4}\right] \\ b_k &= \frac{1}{4}w(k) \bullet \sin\left[\frac{(k-22) \bullet M}{4}\right] \end{aligned} \tag{5}$$

in measuring the wave-front of the rear surface,

$$\begin{aligned} a_k &= \frac{1}{4}w(k) \bullet \cos\left[\frac{(k-22) \bullet (M+1)}{4}\right] \\ b_k &= \frac{1}{4}w(k) \bullet \sin\left[\frac{(k-22) \bullet (M+1)}{4}\right] \end{aligned} \tag{6}$$

in measuring the variation in the thickness,

$$\begin{aligned} a_k &= \frac{1}{4}w(k) \bullet \cos\left[\frac{(k-22)}{4}\right] \\ b_k &= \frac{1}{4}w(k) \bullet \sin\left[\frac{(k-22)}{4}\right] \end{aligned} \tag{7}$$

In the formulas above,  $w(k)$  gives the weighted values determined by a window function.

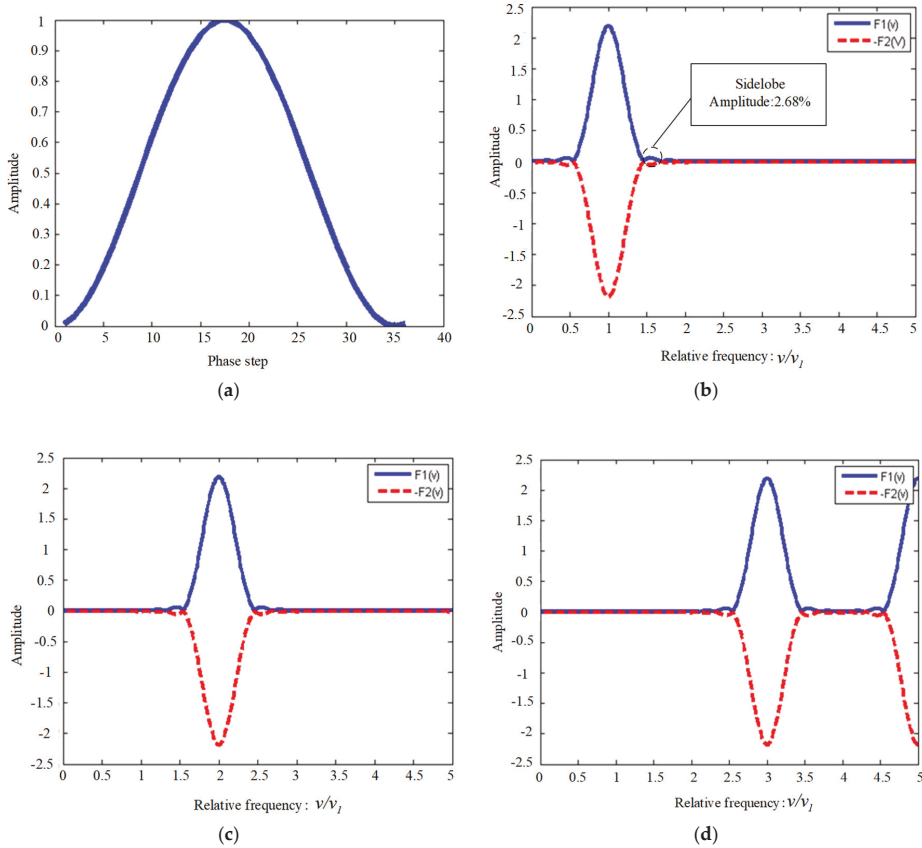
There are many kinds of window functions, such as Bartlett, Blackman, Hanning and Lifting Cosine [28–30]. In order to test the characteristics of the window functions, we use the following functions to estimate the restraint of the harmonic [16]:

$$\begin{aligned} F_1(v) &= \sum_{k=1}^{36} b_k \exp(-i\delta_k v/v_1) \\ F_2(v) &= \sum_{k=1}^{36} a_k \exp(-i\delta_k v/v_1) \end{aligned} \tag{8}$$

where  $i$  is the imaginary unit and  $v$  is the frequency variable. For the symmetrical property of the sampling amplitudes,  $F_1$  and  $F_2$  must be real functions.

Let us take the Hanning window as an example, its figure and evaluating function are shown in Figure 2. In Figure 2, it is assumed that  $M = 2$  and we do not consider the effect of the reflection coefficient. It can be seen that the tested phase can be separated from registered interference pattern in the frequency domain. From Figure 2b–d, it can also be seen that the sensitive of the algorithm for certain harmonious where the corresponding signals are non-zero. The amplitude of the sidelobe of the Hanning window function is approximately 2.68%, as shown in Figure 2b. This demonstrate that the Hanning window does not effectively suppress the side lobes, and information on their

frequencies will be introduced to affect the extraction accuracy of the information to be tested in multi-surface measurements.



**Figure 2.** The Hanning window function. (a) Shape of the Hanning window function; (b) evaluating function of the Hanning window in the thickness measurement; (c) evaluating function of the Hanning window in the front-surface measurement; (d) evaluating function of the Hanning window in rear-surface measurement.

In order to measure the highly reflective sample surface accurately, it is necessary to suppress the harmonic signals more effectively than the Hanning window function. The main factor for evaluating the validity of the window function is the amplitude of the sidelobe in the frequency domain. Therefore, the window function usually needs to be adjusted ulterior. In order to improve the restrain characteristics, the weighted values determined by the window function are subsequently processed. The processed results should satisfy the following equation [11]:

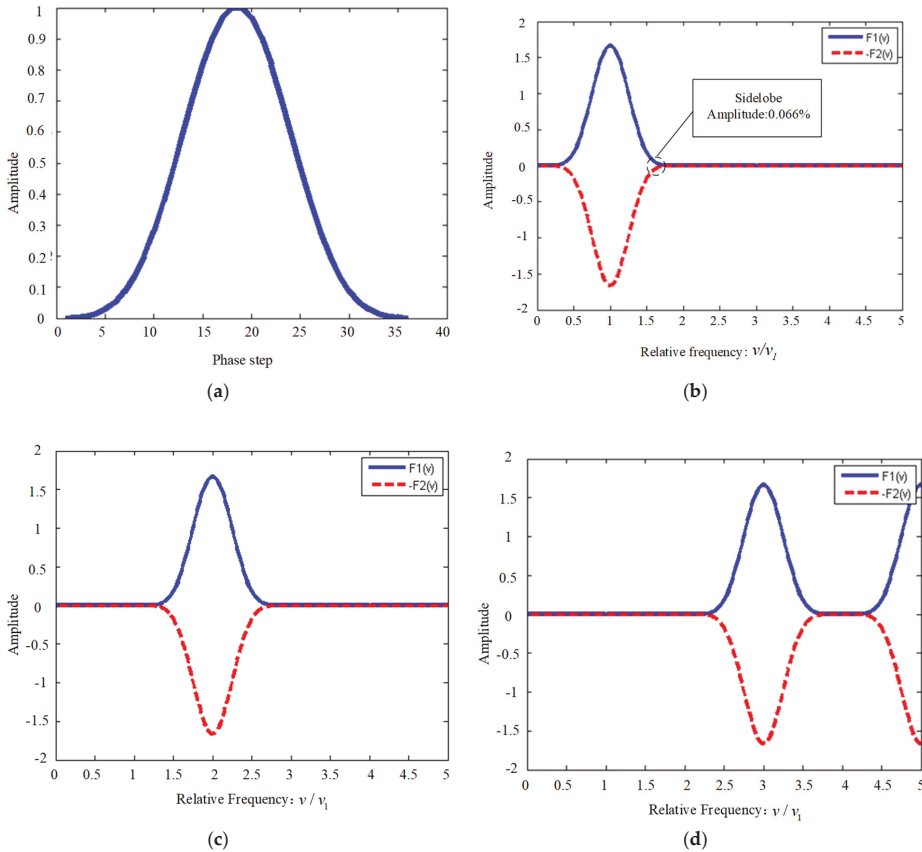
$$\begin{aligned}
 \sum_{k=1}^N a_k &= 0, \quad \sum_{k=1}^N b_k = 0 \\
 \sum_{k=1}^N b_k \sin(-\delta_k) &= \sum_{k=1}^N a_k \cos(\delta_k) \\
 \sum_{k=1}^N b_k \cos(\delta_k) &= \sum_{k=1}^N a_k \sin(-\delta_k)
 \end{aligned} \tag{9}$$

where  $\delta_k$  is the phase shifting value.

According to the phase extraction design algorithm-characteristic polynomial method [27] and Equations (5–7), it is necessary to develop a new window function to satisfy Equation (9). We define a 36-step window function based on the characteristic polynomial design criteria and the phase shift value  $\delta = 2\pi/N = \pi/4$  as follows:

$$w(36) = \frac{1}{2460} (1, 5, 15, 35, 70, 126, 210, 330, 490, 690, 926, 1190, 1470, 1750, 2010, 2260, 2380, 2460, 2460, 2380, 2260, 2010, 1750, 1470, 1190, 926, 690, 490, 330, 210, 126, 70, 35, 15, 5, 1) \quad (10)$$

Figure 3 shows the window function and its evaluating function at different frequencies, respectively. The amplitude of the sidelobe of new window function of the 36-step algorithm is suppressed by approximately 0.066%, as shown in Figure 3a, which is superior to those for the Hanning window (2.68%).



**Figure 3.** New window function of the 36-step algorithm. (a) Shape of the new window function of the 36-step algorithm defined by  $w(36)$ ; (b) new window function of the 36-step algorithm in the thickness measurement; (c) new window function of the 36-step algorithm in front-surface measurement; (d) new window function of the 36-step algorithm in the rear-surface measurement.

As mentioned above, in the actual measurement process it is difficult to measure a transparent plate with a thickness of less than 10 mm because of the limitation in the air-gap range for a Fizeau

type interferometer. In order to overcome this problem, we propose a novel method where a large value for  $M$  is used. This is achieved with the spectral stretching of the Discrete Fourier Transform function applied for the signal obtained for a transparent plate of an arbitrary thickness in the actual measurement process using the Fizeau interferometer with wavelength tuning. In this case

$$M = r \bullet N + 2, r = 0, 1, 2, 3 \dots, \quad (11)$$

where  $N$  is related to the phase-shifting value,  $\delta = 2\pi/N$ . For example, the phase-shift value is set as  $\delta = \pi/4$ , while  $N$  is 8. In this situation, one can measure the sample with  $M = 2, 10, 18, 26, 34 \dots$  as in Equation (11).  $H$  can be set as  $H = 3, 15, 27, 39, 51 \dots$  mm for the transparent plate with a thickness of 1 mm and a refractive index of 1.5.

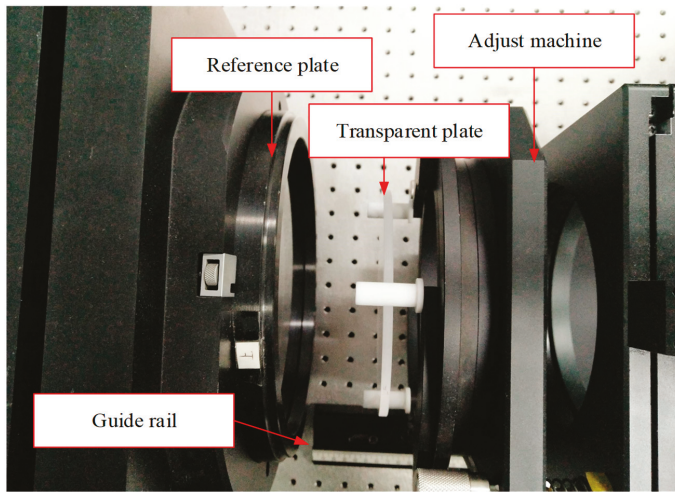
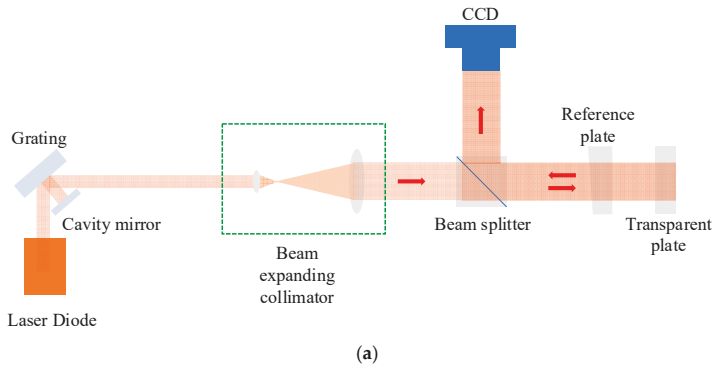
### 3. Results

To test the feasibility of the proposed method, an optical system based on the Fizeau interferometer was built. Figure 4b presents a photo of the measuring device. The sketch of the actual optical path system is shown in Figure 4a. A Littman external cavity (TLB-6804, Newport Corporation, Irvine, USA) consisting of a grating and a cavity mirror was used as the light source. The beam from the laser was directed at the reference surface and the transparent plate to be tested. The beams are reflected from the reference surface, front surface, and rear surface propagate along the same direction and form interference fringes on the CCD camera (The Imaging Source, Bremen, Germany) with the resolution of  $1024 \times 1024$  pixels. The sample is placed vertically on an adjusting machine, with an air-gap distance of  $H$ , which is approximately  $M$  times the length of the optical thickness of the transparent plate.

The central wavelength  $\lambda_0$  of the tunable laser is 632.8 nm. The required laser wavelength is selected by tuning the inclination of the diffraction grating. Such a wavelength control can be executed both manually and automatically by using a computer coupled with the corresponding electro-mechanical unit.

In our experiment, the wavelength was finely scanned from 632.67 nm to 632.97 nm and 36 interference images were recorded with an equal wavelength interval. The minimum fine-tuning resolution was 0.1 GHz, corresponding to  $\Delta\lambda \approx 0.5 \times 10^{-4}$  nm for  $\lambda_0 \approx 632.8$  nm. The stability of the light source wavelength was less than  $10^{-7}$  nm, which could reduce the error associated with the phase-shift during the recording time of 2 min.

In the experimental setup, the expanded collimated laser beam was then projected onto the front-and-rear surfaces of the tested sample through a beam splitter and the reference plate as illustrate in Figure 4a. The light reflected from the reference plate and the interfaces of the tested transparent plate in projected through the beam splitter on the CCD, where the interference field is registered. The example of the field distribution captured by the CCD is shown in Figure 5a.

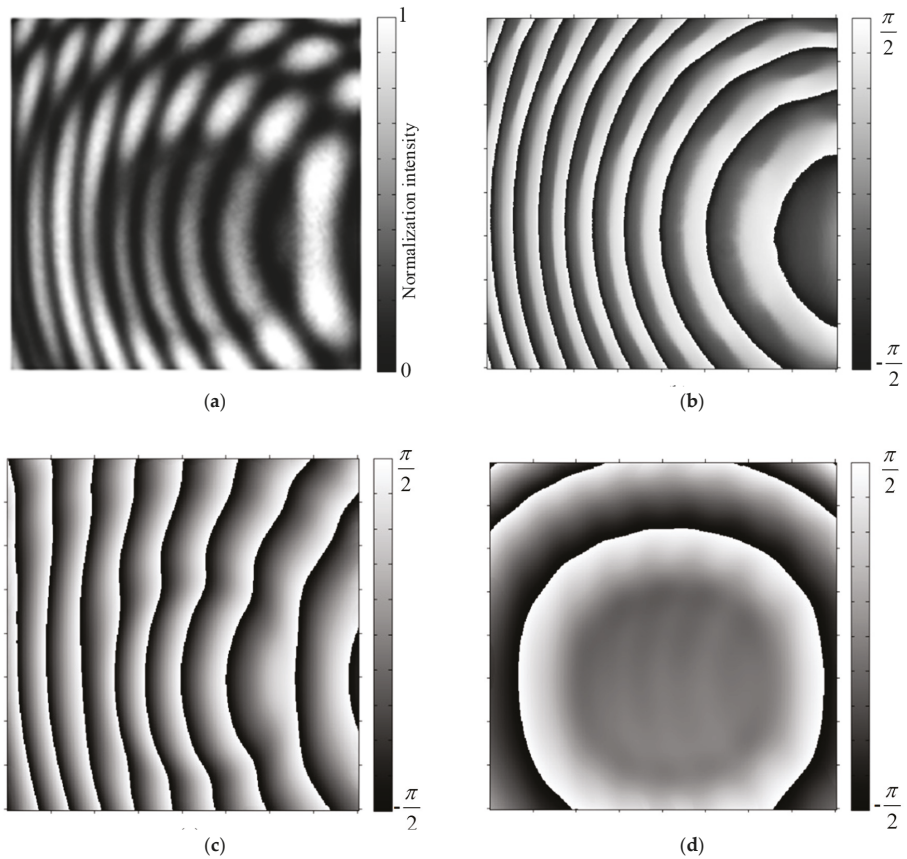


**Figure 4.** Experimental setup. (a) Schematic diagram of the interferometer used to measure the profiles of the front and rear surfaces; (b) laboratory photo of the measuring area of the transparent plate in a wavelength-tuning interferometer.

The tested sample in the experiment was a transparent plate with a thickness approximately 5 mm. The transparent plate was made of JGS1 material and its refractive index was approximately 1.45 at a wavelength of 633 nm. The distance between the reference plate and the front surface  $H$  was adjusted to 72.5 mm ( $M = 10$ ) using a guide rail and adjusting part.

In order to ensure the accuracy and stability of the experiment, the room temperature is stabilized to  $25 \pm 0.2$  °C. In this paper, we set the phase-shift as  $\delta_k = \frac{k \cdot \pi}{4}$  ( $k = 1, 2, 3 \dots$ ) in Equation (4). The diode laser coupled with the diffraction grating provided the interval of the wavelength change,  $\Delta\lambda = 3.5 \times 10^{-3}$  nm. The total wavelength-shift and phase-shift used in the experiment were 0.126 nm and  $9\pi$ , respectively.

By applying the 36-step algorithm described in Section 2, the information of the front and rear surfaces and thickness variation could be extracted, respectively. The corresponding shapes of the front and rear surfaces of the optical component were calculated after the phase unwrapping had occurred. To remove the high-frequency noise, the phase map was filtered using the windowed Fourier transform (WFT) method [31,32] as shown in Figure 5b–d.



**Figure 5.** Distributions of the registered and processed signals obtained in the experiment. (a) The fringe pattern from both surfaces of the transparent plate at the initial wavelength; (b) wrapped phase map of the front surface after the windowed Fourier transform (WFT) filter was applied; (c) wrapped phase map of the rear surface after the WFT filter was applied; (d) wrapped phase map of variation in the thickness after the WFT filter was applied.

After unwrapping and compensating for the tilt phase aberration of the phase map obtained from the both surfaces of the transparent plate, the profiles of both surfaces of the studied sample can be obtained according to the following relationship:

$$\delta = \frac{4\pi d(x, y)}{\lambda}, \tag{12}$$

where  $d$  describes the optical path difference (OPD). For the three phase maps generated from the reference surface and the front-rear surfaces of the tested transparent plate,  $\phi_f(x, y)$ ,  $\phi_r(x, y)$  and  $\phi_{f-r}(x, y)$  can be expressed as follows:

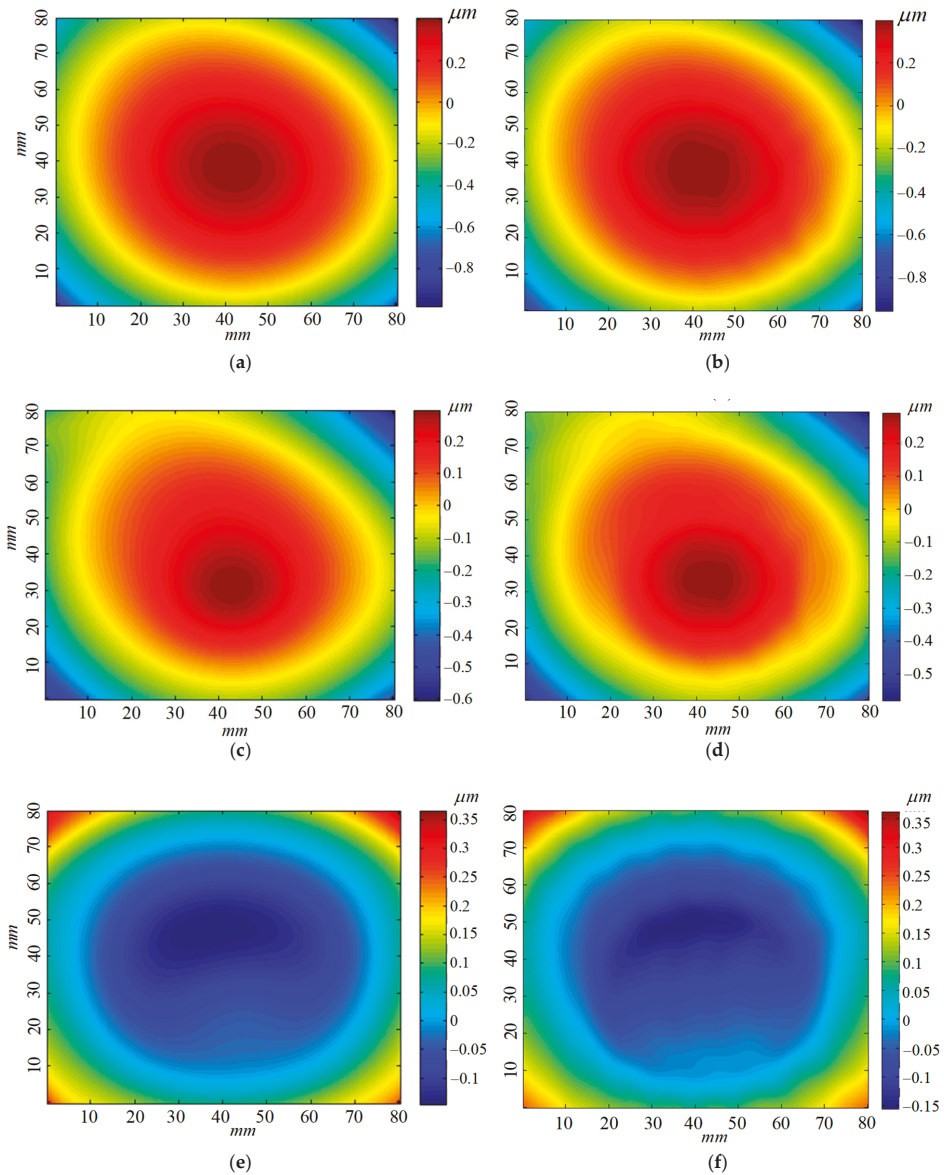
$$\begin{aligned} \phi_f(x, y) &= \frac{4\pi[H(x, y) + W_f(x, y)]}{\lambda} \\ \phi_r(x, y) &= \frac{4\pi[H(x, y) + W_f(x, y) + n_0(W_r(x, y) - W_f(x, y)) + n_0T(x, y)]}{\lambda} \\ \phi_{f-r}(x, y) &= \frac{4\pi[n_0T(x, y) + n_0(W_r(x, y) - W_f(x, y))]}{\lambda} \end{aligned} \tag{13}$$

where  $H$  represents the distance of the air-gap and  $W_f, W_r$  are front and rear surface profiles of the sample,  $n_0$  and  $T$  are the reflective index of the sample, which were assumed to have a uniform distribution, and the thickness of the sample, respectively.

The measured shape of the front surface, rear surface and variation in the thickness are shown in Figure 6a–c, respectively. The mean peak-to-valley (PV) values of the front-and-rear surfaces of the tested transparent plate and its thickness variation were found as 1.38  $\mu\text{m}$ , 0.89  $\mu\text{m}$  and 0.51  $\mu\text{m}$ , respectively, by carrying out 10 repeated measurements at the same position. The repeatability of the 36-step algorithm is confirmed by a pair of measurements taken successively within three days. According to the obtained results by 10 repeated measurements, the root mean square errors (RMSE) for determining both surfaces did not exceed 1.5 nm. The detailed data of the measurements are shown in Table 2. To test the reliability of the 36-step algorithm, the sample was measured by using a ZYGO interferometer (Zygo Corporate, Middlefield, USA), which is considered as a reliable testing instrument in industrial applications. However, it cannot measure the front surface and rear surface profiles simultaneously. Therefore, only one surface was measured, whereas the other surface was covered with Vaseline to suppress its reflection. Then, the other surface was measured while Vaseline covered the first surface. Figure 6b,d,f show the measured front-and-rear surface profiles and variation in the thickness respectively, using a ZYGO interferometer. The results indicate that the shapes measured with the proposed 36-step algorithm were consistent with the ZYGO interferometer. Quantificationally, the PV of front-and-rear surface profiles and variation in the thickness were 1.40  $\mu\text{m}$ , 0.88  $\mu\text{m}$  and 0.53  $\mu\text{m}$ , respectively.

**Table 2.** The result of the peak-to-valley (PV) value of three surfaces obtained by 10 repeated measurements using the 36-step algorithm ( $\mu\text{m}$ ).

	Measurement Number										Average Value	RMSE
	1	2	3	4	5	6	7	8	9	10		
front surface	1.3815	1.3809	1.3814	1.3824	1.3815	1.3803	1.3814	1.3838	1.3810	1.3802	1.3814	0.0010
rear surface	0.8932	0.8923	0.8943	0.8915	0.8924	0.8911	0.8932	0.8927	0.8900	0.8911	0.8922	0.0013
variation of thickness	0.5114	0.5126	0.5015	0.5127	0.5113	0.5112	0.5103	0.5139	0.5117	0.5108	0.5116	0.0011



**Figure 6.** The surface profiles measured with the 36-step algorithm and the ZYGO interferometer. (a) The profile of the front surface obtained with the 36-step algorithm; (b) the profile of the front surface obtained the ZYGO interferometer; (c) the profile of the rear surface obtained with the 36-step algorithm; (d) the profile of the rear surface obtained with the ZYGO interferometer; (e) the profile of the variation in the thickness obtained with the 36-step algorithm; (f) the profile of variation in the thickness obtained with the ZYGO interferometer.

#### 4. Discussion

The previous sections demonstrated the method for separating the superimposed fringe patterns generated as a result of reflection from two parallel surfaces of a tested optical component. The data



processing was based on the weighted Fourier transform technique, which is preferable to conventional phase-shifting methods. Despite the previously demonstrated advantages, there are some limitations to this approach. When the phase shift is nonlinear, the phase shift value  $\delta_m$  is a function of the phase-shift parameter. The phase shift value of the  $m$ th sample can be expressed as

$$\begin{aligned} \delta_m &= \delta_{0m}[1 + \varepsilon(\delta_{0m})] \\ &= \delta_{0m}[1 + \varepsilon_0 + \varepsilon_1 \frac{\delta_{0m}}{\pi} + \varepsilon_2 (\frac{\delta_{0m}}{\pi})^2 + \dots + \varepsilon_p (\frac{\delta_{0m}}{\pi})^p] \end{aligned} \tag{14}$$

where  $\varepsilon$  is the error coefficient of the phase-shift miscalibration and  $p$  is the maximum order of nonlinearity. Figure 7 shows the relationship between phase error coefficients of different order  $\varepsilon_0, \varepsilon_1, \varepsilon_2, \varepsilon_3$  and the phase shift error. According to this graph, the phase shift error increases with the increase of the coefficient order of the phase-shift error, where the first order linear coefficient  $\varepsilon_0$  has the greatest influence and the effects of the second and third nonlinear coefficient order decrease sequentially. In order to reduce the impact of phase-shift error on the accuracy of measurement, it is necessary to suppress the effects of low-order error coefficients, especially the first-order coefficient  $\varepsilon_0$ , second-order coefficient  $\varepsilon_1$  and third-order coefficient  $\varepsilon_2$ .

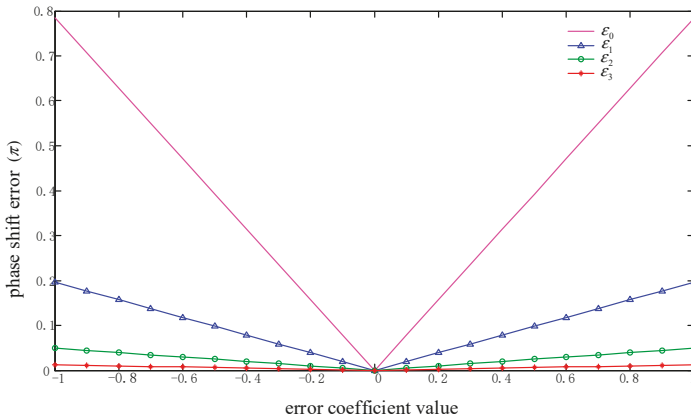


Figure 7. Relationship between phase error coefficients of different order and phase shift error.

Further work should focus on improving the phase-insensitive ability to meet the requirement for real-world applications. Adequate results can then be obtained using fewer interferograms.

### 5. Conclusions

This paper presented a new method for the simultaneous measurement of the front and rear surface profiles of a transparent plate. The simulation analysis and experimental results demonstrated the feasibility of the proposed method. A transparent plate with a thickness of approximately 5 mm was tested as the sample. According to the results, the RMSE using 10 repeat experiments in the determination of both surface profiles did not exceed 1.5 nm. A 36-step algorithm solves the problem regarding the need to restrict the thickness of the sample to a minimum value, for example, more than 5 mm, using the parameters detailed in the paper. This is a good result for a practical measurement application. In addition, a relatively simple experimental setup and reasonably high accuracy are among the advantages of the 36-step algorithm over other techniques [16,19]. Some issues, such as the calculation speed and the cost of a higher quality laser source, must be considered to improve the performance of the 36-step algorithm.

**Author Contributions:** Writing—Original manuscript Preparation and Software: T.S.; Formal Analysis, Investigation and Methodology: W.Z. and K.Y.; Project Administration, Y.Y.; Supervision: A.K.A.; Review and editing: Y.Y. and S.V.

**Funding:** The authors gratefully acknowledge the support of the National Science and Technology Major Project (project no. # 2016YFF0101905) and the National Natural Science Foundation of China (NSFC) (no. 51775326).

**Acknowledgments:** The authors would like to thank Paraskvi Tornari of Institute of Electronic Structure and Laser, Foundation for Research and Technology—Hellas and Wenjing Zhou from department of precision mechanical engineering of Shanghai University for help with this paper.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Schmitt, R.; Doerner, D. Measurement technology for the machine integrated determination of form deviations in optical surface. *CIPR Ann. Manuf. Technol.* **2006**, *55*, 559–562. [[CrossRef](#)]
2. Fang, F.Z.; Zhang, X.D.; Weckenman, A.; Zhang, G.X.; Evans, C. Manufacturing and measurements of freeform optics. *CIRP Ann. Manuf. Technol.* **2013**, *62*, 823–846. [[CrossRef](#)]
3. Valyukh, S.; Valyukh, I.; Skarp, K. Spectrophotometric determination of reflective liquid crystal cell parameters. *J. Appl. Phys.* **2006**, *99*, 8590. [[CrossRef](#)]
4. Kim, Y.; Hibino, K.; Sugita, N.; Mitsuishi, M. Absolute optical thickness measurement of transparent plate using excess fraction method and wavelength-tuning Fizeau interferometer. *Opt. Express* **2015**, *23*, 4065–4073. [[CrossRef](#)] [[PubMed](#)]
5. Ishii, M.Y.; Chen, J.; Tsujiuchi, J. Digital phase-measuring interferometry with a tunable laser diode. *Opt. Lett.* **1987**, *12*, 233–235. [[CrossRef](#)] [[PubMed](#)]
6. Takeda, M.; Mutoh, K. Fourier transform profilometry for the automatic measurement of 3-D object shapes. *Appl. Opt.* **1983**, *22*, 3977–3982. [[CrossRef](#)] [[PubMed](#)]
7. Liu, F.W.; Wu, Y.Q.; Wu, F. Phase shifting interferometry from two normalized interferograms with random tilt phase-shift. *Opt. Express* **2015**, *23*, 19932–19946. [[CrossRef](#)] [[PubMed](#)]
8. Su, X.Y.; Dou, Y.F.; Chen, Y.F.; Wang, Y. A flexible fast 3D profilometry based on modulation measurement. *Opt. Lasers Eng.* **2011**, *49*, 376–383. [[CrossRef](#)]
9. De Groot, P.; Deck, L. Surface profiling by analysis of white-light interferograms in the spatial frequency domain. *J. Mod. Opt.* **1995**, *42*, 389–401. [[CrossRef](#)]
10. Tsuruta, T.; Ichihara, Y. Accurate measurement of lens thickness by using white-light fringes. *J. Appl. Phys.* **1975**, *14*, 369–372. [[CrossRef](#)]
11. Kuo; Chuan, W.; Bou, Y.K.; Lai, C.M. Simultaneous measurement of refractive index and thickness of transparent material by dual-beam confocal microscopy. *Meas. Sci. Technol.* **2013**, *24*, 075003. [[CrossRef](#)]
12. Kim, S.; Na, J.; Kim, M.J.; Lee, B.H. Simultaneous measurement of refractive index and thickness by combining low-coherence interferometry and confocal optics. *Opt. Express* **2008**, *16*, 5516–5526. [[CrossRef](#)] [[PubMed](#)]
13. Liu, K.; Littman, M.G. Novel geometry for single-mode scanning of tunable lasers. *Opt. Lett.* **1981**, *6*, 117–118. [[CrossRef](#)] [[PubMed](#)]
14. Abou Zeid, Y.A.; Wiese, P. Interferometer with a wavelength-tuned diode laser for surface profilometry. *Meas. Sci.* **1998**, *9*, 1105–1110. [[CrossRef](#)]
15. Hibino, K.; Oreb, B.F.; Fairman, P.S. Improved algorithms for wavelength scanning interferometry: Application to the simultaneous measurement of surface topography and optical thickness variation in a transparent parallel plate. *Proc. SPIE* **2002**, 177–183. [[CrossRef](#)]
16. Okada, K.; Sakuta, H.; Ose, T. Separate measurements of surface shapes and refractive index inhomogeneity of an optical element using tunable-source phase shifting interferometry. *Appl. Opt.* **1990**, *23*, 3280–3285. [[CrossRef](#)] [[PubMed](#)]
17. Yu, Y.J.; Zhang, B.H.; Jiao, Y.F. Phase-shifting algorithm via wavelength tuning based on temporal Fourier transform. *Proc. SPIE* **2003**, 5188, 343. [[CrossRef](#)]
18. Hibino, K.; Mitsuishi, M.; Sugita, N.; Hanayama, R.; Kim, Y. Multiple-surface interferometry of highly reflective wafer by wavelength tuning. *Opt. Express* **2014**, *22*, 21145–21156. [[CrossRef](#)]
19. Deck, L.L. Multiple surface phase shifting interferometry. *Proc. SPIE* **2001**, 424–431. [[CrossRef](#)]

20. Kim, Y.; Hibino, K.; Sugita, N.; Mitsuishi, M. Interferometric profile measurement of optical-thickness by wavelength tuning with suppression of spatially uniform error. *Opt. Express* **2018**, *26*, 10870–10878. [[CrossRef](#)]
21. Hibino, K.; Oreb, B.F.; Fairman, P.S.; Burke, J. Simultaneous measurement of surface shape and variation in optical thickness of a transparent parallel plate in wavelength scanning Fizeau interferometer. *Appl. Opt.* **2004**, *43*, 1241–1249. [[CrossRef](#)] [[PubMed](#)]
22. De Groot, P. Method and System for Profiling Objects Having Multiple Reflective Surfaces using Wavelength Tuning Phase-Shifting Interferometry. U.S. Patent 6 359 692, 19 January 2002.
23. De Groot, P. Measurement of transparent plates with wavelength-tuned phase-shifting interferometry. *Appl. Opt.* **2000**, *39*, 2658–2663. [[CrossRef](#)] [[PubMed](#)]
24. Hibino, K.; Oreb, B.F.; Fairman, P.S. Wavelength-scanning interferometry of a transparent parallel plate with refractive-index dispersion. *Appl. Opt.* **2003**, *42*, 3888–3895. [[CrossRef](#)] [[PubMed](#)]
25. Sun, T.; Zheng, W.W.; Yu, Y.J.; Asundi, A.K.; Valyukh, S. Determination of surface profiles of transparent plates by means of laser interferometry with wavelength tuning. *Opt. Lasers Eng.* **2019**, *115*, 59–66. [[CrossRef](#)]
26. Kim, Y.; Hibino, K.; Sugita, N.; Mitsuishi, M. Optical thickness measurement of mask blank glass plate by the excess fraction method using wavelength-tuning interferometer. *Opt. Lasers Eng.* **2013**, *51*, 1173–1178. [[CrossRef](#)]
27. Surrel, Y. Design of algorithms for measurements by the use of phase stepping. *Appl. Opt.* **1996**, *35*, 51–60. [[CrossRef](#)] [[PubMed](#)]
28. Harris, F.J. On the use of windows for harmonic analysis with the discrete Fourier transform. *Proc. IEEE* **1978**, *66*, 51–83. [[CrossRef](#)]
29. Servin, M.; Estrada, J.C.; Quiroga, J.A. Spectral analysis of phase shifting algorithms. *Opt. Express* **2009**, *17*, 16423–16428. [[CrossRef](#)]
30. Servin, M.; Estrada, J.C.; Quiroga, J.A. The general theory of phase shifting algorithms. *Opt. Express* **2009**, *17*, 21867–21881. [[CrossRef](#)]
31. Qian, K.M. Windowed Fourier transform for fringe pattern analysis. *Appl. Opt.* **2004**, *43*, 52695–52702. [[CrossRef](#)]
32. Qian, K.M. Two-dimensional windowed Fourier transform for fringe pattern analysis: Principles, applications and implementations. *Opt. Lasers Eng.* **2007**, *45*, 304–317. [[CrossRef](#)]



© 2019 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).

Article

# Optical 3-D Profilometry for Measuring Semiconductor Wafer Surfaces with Extremely Variant Reflectivities

Liang-Chia Chen <sup>1,2,\*</sup> , Duc-Hieu Duong <sup>2</sup> and Chin-Sheng Chen <sup>2</sup> 

<sup>1</sup> Department of Mechanical Engineering, National Taiwan University, No. 1, Section 4, Roosevelt Rd, Da'an District, Taipei City 10617, Taiwan

<sup>2</sup> Graduate Institute of Automation Technology, College of Mechanical and Electrical Engineering, National Taipei University of Technology, No. 1, Section 3, Zhong-Xiao E. Rd, Da'an District, Taipei City 10608, Taiwan; duongduchieu85@gmail.com (D.-H.D.); saint@mail.ntut.edu.tw (C.-S.C.)

\* Correspondence: lchen@ntu.edu.tw; Tel.: +886-2-23620032

Received: 2 April 2019; Accepted: 14 May 2019; Published: 19 May 2019



**Abstract:** A new surface profilometry technique is proposed for profiling a wafer surface with both diffuse and specular reflective properties. Most moiré projection scanning techniques using triangulation principle work effectively on diffuse reflective surfaces, on which the reflected light beams are assumed to be well captured by optical sensors. In reality, this assumption is no longer valid when measuring a semiconductor wafer surface having both diffuse and specular reflectivities. To resolve the above problem, the proposed technique uses a dual optical sensing configuration by engaging two optical sensors at two different viewing angles, with one acquiring diffuse reflective light and the other detecting at the same time specular surface light for achieving simultaneous full-field surface profilometry. The deformed fringes measured by both sensors could be further transformed into a 3-D profile and merged seamlessly for full-field surface reconstruction. Several calibration targets and industrial parts were measured to evaluate the feasibility and accuracy of the developed technique. Experimental results showed that the technique can effectively detect diffuse and specular light with repeatability of one standard deviation below 0.3  $\mu\text{m}$  on a specular surface and 2.0  $\mu\text{m}$  on a diffuse wafer surface when the vertical measuring range reaches 1.0 mm. The present findings indicate that the proposed technique is effective for 3-D microscale surface profilometry in in-situ semiconductor automated optical inspection (AOI).

**Keywords:** semiconductor; surface profilometry; moiré projection; 3-D measurement; automated optical inspection (AOI)

## 1. Introduction

Measuring the 3-D profile of an object surface has become increasingly important in industrial automation, particularly in in-situ product inspection. 3-D surface measurement could be generally classified into two major categories: tactile and non-contact measuring conditions. Coordinate Measuring Machine (CMM) is a general system using both contact and non-contact techniques to measure 3-D profiles. Tactile sensing cannot accurately measure soft object or thin-film surfaces as the probe may scratch or deform the measured surface. In contrast, non-contact techniques can effectively measure surfaces with high spatial resolution and are useful for measuring soft or flexible objects. For non-contact measurement, optical triangulation techniques, such as stereo vision, laser scanning, and structured light techniques, are widely used in industries. However, optical techniques often require uniform surface light reflection to ensure reliable sensing. The absence of such condition would pose difficulty for using optical triangulation techniques to measure a surface with both diffuse and

specular reflectivities. One of the significant challenges in optical surface measurement lies in dealing with reflectivity variations of the object surfaces to be tested, such as a semiconductor wafer surface being engraved or marked by laser etching. Measuring objects with high reflectivity variations, such as highly shiny or scattering surfaces, is a critical issue to ensuring measurement accuracy and product quality assurance under automated optical inspection (AOI). Poor 3-D profiling results are common when encountering high reflectance variances on the wafer surface to be tested or reconstructed for 3-D printing or other purposes.

One of the popular 3-D optical surface measurement techniques is structured light projection, which is widely employed in industrial testing applications. The techniques basically work according to the triangulation measurement principle with the assumption of uniform light reflectivity on the tested object for ensuring capture of the deformed fringe reflecting from the tested surface. However, in reality, such assumption is not applicable in face of high reflectivity variations on the surface. For example, most industrial IC chips are fabricated using a silicon wafer substrate with an extremely shiny surface and are also embedded with metal bumps having a scattering surface reflectance. Under such condition, the reflected lights from the tested surface are rather complex and emitting in different directions. Consequently, the detecting sensor cannot receive accurate phase information for reconstructing surface morphology.

Structured fringe projection is a full-field optical technique for measuring 3-D object surfaces. Owing to its advantages of fast speed, high accuracy, being non-destructive, and full-field testing, fringe projection technique (FPP) has become one of the most promising 3-D surface profiling techniques in 3-D AOI. For obtaining the object shape, the common 3-D measurement techniques include mainly phase shifting profilometry (PSP) [1] and Fourier transform profilometry (FTP) [2]. In general, these techniques project structured light patterns, also called fringes, onto the object surface. Owing to optical path difference (OPD) between inspecting beams, the fringe is deformed according to the 3-D shape of the object surface. To reconstruct the 3-D shape of the object surface, one phase map is computed using a wrapping algorithm, and object height can be extracted using the phase unwrapping algorithm [3] and phase-to-height transformation. PSP is widely employed in industrial inspection because of its promising accuracy and high spatial resolution [4]. It uses a time-resolved wrapping strategy to extract phase information for reconstruction of object profiles. Different from PSP, FTP records distorted fringe images and analyzes object height information using phase modulation and demodulation to acquire a phase map from distorted fringes. The unique advantage of FTP is its one-shot capability of 3-D surface imaging. However, for exacting 3-D profile reconstruction, one of the key issues in FTP is to extract accurately the first-order spectrum from the spectral domain. The extraction would require an adequate band-pass filter to obtain accurate spectral information from the first-order spectrum and separate it from the background spectrum (zero-order spectrum). Conditions for separating these spectrum regions have been investigated by considering the minimum distance between spectrum regions [5–7]. The single elliptic band-pass filter [8] and double orthogonal elliptic band-pass filter [9,10] have been developed to extract accurately the first-order spectrum from the spectral domain. Nevertheless, FTP still lacks a robust algorithm for detecting the first-order spectrum to reconstruct accurately 3-D profiles. Apart from these methods, scanning white light interferometry (SWLI) is powerful for its phase unambiguity and nano-scale depth resolution [11]. It has been widely applied for precise sub-micro-scale and nano-scale surface profilometry, especially in semiconductor microstructure metrology. The method is mainly used for review or off-line metrology since its scanning efficiency is limited by vertical scanning required.

A major challenge for 3-D scanning techniques using FPP is the presence of specular reflectance on the tested object surface. Specular highlights occur on object surfaces when the specular element of reflection is dominant. Specular highlight can easily saturate the detected image and seriously deteriorate image quality. In the literature, several techniques have been proposed for resolving the problem when measuring specular targets. All these techniques can effectively reduce specular effects to different extents but cannot resolve the problem completely. One of the proposed techniques was

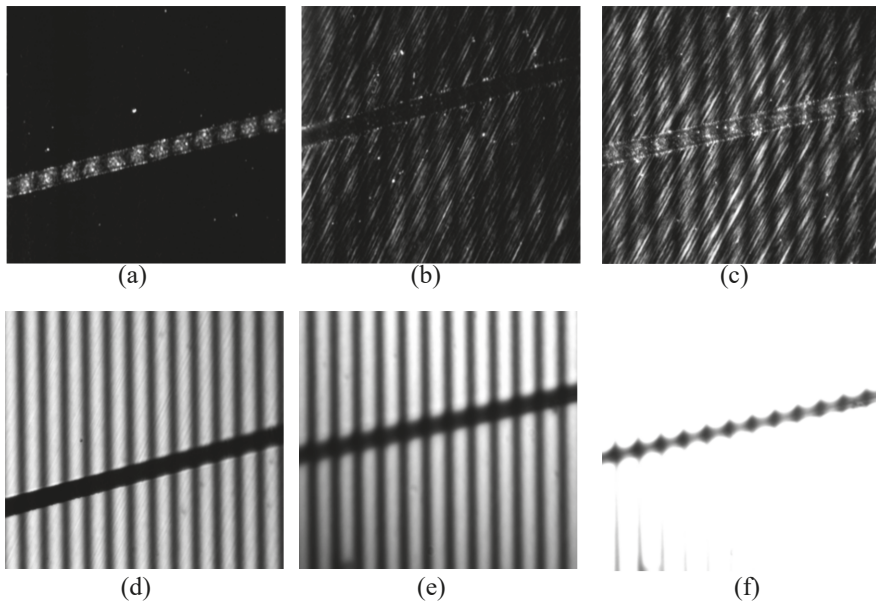
to separate specular components from diffuse components and remove the specular components using polarizers [12–15]. Tan and Ikeuchi relied on image color properties when transferring image color to the hue domain [16]. This technique effectively eliminated image-saturated regions mainly because the saturation value of every pixel was made constant for achieving maximum chromaticity while keeping their hue. Other techniques, such as using multi-exposure time [17] or projecting special structured light with strip edges [18] have been experimented for handling non-uniform light reflection. Phase measuring deflectometry (PMD) is a specular surface measuring technique using a Thin-Film-Transistor Liquid-Crystal Display (TFT LCD) screen as the pattern-projecting light source [19,20]. The maximum detectable surface curvature is currently limited by the LCD light dispersion angle. More recently, single-shot deflectometry using a single composite pattern was developed for one-shot measurement of three-dimensional imaging using Fourier transform (FT) and spatial carrier-frequency phase-shifting (SCPS) techniques [21]. However, generally for PMD, due to demanding computation in slope integration, the absolute shape measurement may not be accurately achieved. Thus, apart from visible moiré fringe projection, Wiedenmann et al. used high-power infrared laser pattern to project on the measured object surface, in which the projected laser power was partly transferred into dispersive heat [22]. Since heat dispersion can be generated from a highly specular surface, extremely dark or transparent materials can be observed clearly by a proper heat sensor [23]. Phase shifting technique can be employed to measure 3-D profile according to distortion of the heat pattern on the object surface. The current limitation of the method is its low measuring accuracy, which still requires further improvement. Up to date, some light separation strategies using light property analysis were developed, including separation specular reflection using color analysis [3], examining chromaticity and intensity value of specular and diffuse pixels [16], employing polarization to separate reflection components [24] or using environmental structured illumination [25] techniques. Although these methods have been effective to some extents in separating diffuse from specular light, they basically assume either uniform surface reflectance or specific surface light reflectance conditions, which may not be applicable for many in-situ testing applications.

Many semiconductor parts with variant reflectivities need to be inspected in in-situ manufacturing processes for strict dimensional compliance. These parts often comprise a body (such as epoxy and silicon), metal bumps, and pads (such as lead and solder), each of which may have complicated geometric shapes with different orientations with respect to the detecting optical sensor [26]. Moreover, the most challenging task is to overcome extremely variant wafer surface reflectivities, ranging widely from diffuse to specular conditions. In 3-D inspection, the 3-D surface profilometric technique of the tested objects often projects a structured light pattern with a desirable assumption that a non-absorbent and perfect diffusely scattered model of the surface under test (SUT) is observed. However, in reality, this assumption may not be true for many actual cases. In optical triangulation detection, viewing the measured surface from a sensing angle can pose a huge measuring uncertainty in ensuring detection of reflecting light due to variant surface reflectance properties. The extent of reflecting light detected by imaging sensors can significantly affect the success of surface measurement. Most existing techniques discussed above can resolve the light specular problem only for a surface with uniform reflectivity. For objects with non-uniform surface reflectivity, the captured image contrast becomes poor or even non-detectable.

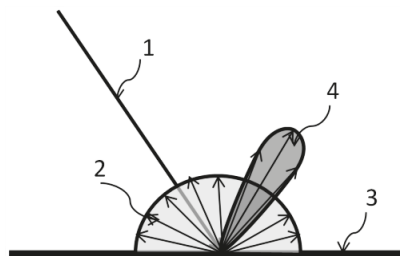
Therefore, it is of need to develop an effective technique that can overcome the complex reflectivity problem encountered in optical surface profilometry, especially for in-situ semiconductor wafer surface inspection. The article proposes a dual sensing strategy to overcome the 3-D surface profilometric problem encountered in semiconductor wafer surface. The rest of this paper is organized as follows. Section 2 presents the proposed method for 3-D profilometry using dual sensing. The experimental design and measured results are shown and analyzed in Section 3. Finally, the conclusions are summarized in Section 4.

## 2. Developed 3-D Profilometry Using Dual Sensing

In structured fringe projection, the angle between the projection and the sensing detection plays an important role in determining the reflecting direction. Figure 1a shows the fringe image captured from a semiconductor wafer surface detected using a laser-etched microscale groove with the viewing angle set at zero. As can be seen, the groove area is rather clear but the remaining wafer specular area is almost non-detectable. Therefore, the groove area can be seen with an adequate exposure time because diffuse light is generally weak. With larger viewing angle, the wafer surface becomes clearer with a better image contrast (see Figure 1b,c). Illustration shown from Figure 1b–d represents a transition in imaging with increase in viewing angle. When the reflectance relation follows the light reflectance rule, the fringe contrast on the wafer surface reaches its maximum as illustrated in Figure 1e. Figure 1f represents the case when the sensor becomes over saturated owing to excessive exposure time. Figure 1 illustrates the critical reflectance condition of the tested surface under inspection. Figure 2 shows the general reflection model of diffuse and specular components when the light is incident from an angle.



**Figure 1.** Images of wafer surface etched with a laser-etched groove and captured at different viewing angles: (a) viewing angle  $\pm 1^\circ$ , exposure time 420 ms; (b) viewing angle  $12 \pm 1^\circ$ , exposure time 200 ms; (c) viewing angle  $25 \pm 1^\circ$ , exposure time 80 ms; (d) viewing angle  $45 \pm 1^\circ$ , exposure time 20 ms; and (e) viewing angle  $45 \pm 1^\circ$ , exposure time 50 ms.



**Figure 2.** Reflection model of diffuse and specular components.

To resolve the above issue, this article presents a new optical measuring configuration to capture both of specular and diffuse light reflecting from wafer surfaces. A dual sensing strategy is proposed to detect the light reflection from wafer surfaces when structured fringe projection is used to measure and reconstruct 3-D surfaces. The proposed system design and measuring principle are detailed in the following sessions.

### 2.1. Optical System Setup

Figure 3 illustrates the optical system configuration of the developed 3-D profilometry. As can be seen, it includes a low-coherent light source (1), a diffuser (2), which is collimated by a lens module (3), a sinusoidal grating (4), and a linear translator (5) for generating an accurate shifted phase. To achieve full-field surface profilometry using PSP, the light source is a low-coherent light source such as LED (light-emitting diode), instead of a coherent one like lasers. In the system, the structured light generated by the incident optical module passes through a telecentric lens (6) and is projected onto the measured surface (8) defined by the reference plane (7). The reflected light passes through two objectives (12 and 13) and is captured by two individual optical sensing cameras (10 and 11). The first objective (12) is arranged in vertical direction to capture the diffuse light while the second objective is set to a reflection angle with respect to the incident projection light angle, so that specular light can be well captured. With such an optical configuration, both diffuse and specular light beams reflected can be well detected by their corresponding objectives. When the projected sinusoidal fringe (9) hits the measured wafer surface, it becomes a deformed fringe according to the phase difference generated by the optical path difference (OPD) between the object profile and the reference plane. The deformed fringe image is transferred to a computer (14) and the phase map is then computed with multi-phase shifting and the wrapping principle using a set of deformed fringe images. Subsequently, phase unwrapping and phase-to-height transformation are performed to extract the profile of the measured surface.

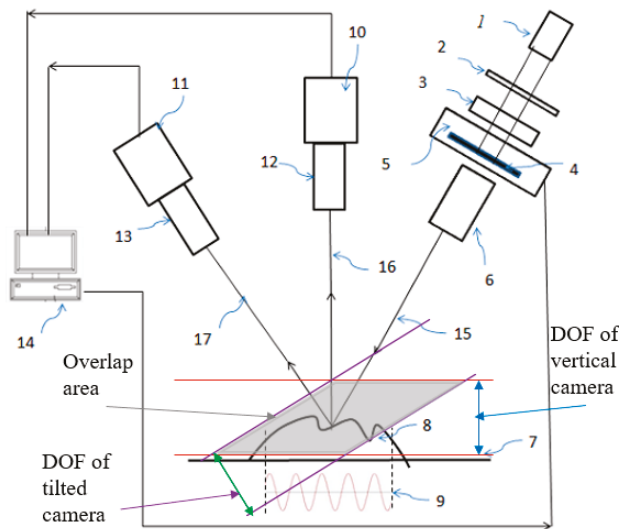


Figure 3. Configuration of the developed dual-sensing measurement system.

All the optical lenses employed in the developed system are telecentric to ensure the formation of a constant image size along the depth of field (DOF) of the measurement. When the measured surface located inside the overlapped with the DOF of the two cameras (indicated as the gray zone in Figure 3), the captured image size can be kept constant, so the FOV is kept as a rectangular or square shape according to the image sensor’s design. Meanwhile, to keep the focus plane of the imaging



sensor to locate at the same focus plane with the projecting light side, the image sensor can tilted with the principle optical axis by an adequate angle to compensate the tilting effect. Thus, both of the focus planes (light projection and imaging) are kept at the same focus plane and the image shape are rectangular or square in the same FOV. Meanwhile, an adequate camera calibration can be performed to identify all the intrinsic and extrinsic parameters of the coordinate transform, so 3D point cloud can be transformed with image coordinates accurately.

## 2.2. Measuring Principle

### 2.2.1. Fundamentals in PSP

Using PSP technique for 3-D shape measurement involves several procedures, including projecting sinusoidal fringe patterns onto the object surface, taking the images of the reflected distorted pattern on the object surface, and then calculating the phase map from these captured images. Height information is extracted from the phase map by phase unwrapping and phase-to-height transformation. The calibration coefficients are then employed to transform the 3-D coordinates in image coordinates into world coordinates  $(x, y, z)$  of each measured point. The phase  $\phi(x, y)$  can be computed using light intensities of captured images set by phase-shifted fringe projection. The five-step phase-shifting technique is commonly used because of its accuracy and efficiency. The intensities of the five channels in each pattern can be modeled as follows:

$$I_i(x, y) = I_b(x, y) + I_m(x, y) \cos[\phi(x, y) + \delta_i]_{(i=1\sim5)} \quad (1)$$

where  $\phi(x, y)$  is the phase value used by the phase-shifting algorithm in  $(x, y)$  location,  $\delta_i$  is the phase-shift,  $I_i(x, y)$  denotes the intensities of the five captured images for each pixel position,  $I_m(x, y)$  and  $I_b(x, y)$  are the intensity modulation and the average intensity for each pixel, respectively.

The phase value of each pixel  $\phi(x, y)$  can be resolved using Equation (1) as:

$$\phi(x, y) = -\tan^{-1} \left[ \frac{\sum_{i=1}^N I_i(x, y) \sin(\delta_i)}{\sum_{i=1}^N I_i(x, y) \cos(\delta_i)} \right] \quad (2)$$

where  $\delta_i, i = 0, 1, 2, \dots, N$ , are the phase-shift values of each channel.

For the five-step phase-shifting technique ( $N = 5$ ) and  $\delta_1 = 0^\circ, \delta_2 = 90^\circ, \delta_3 = 180^\circ, \delta_4 = 270^\circ, \delta_5 = 360^\circ$ , the phase value  $\phi(x, y)$  can be computed using the intensity of each channel and the corresponding phase-shift  $\delta_i$  from the five-step phase-shifting technique shown in Equation (2).

With the phase value  $\phi(x, y)$  being a function of arctangent, the phase value can be obtained in a range from  $-\pi$  to  $\pi$  for phase wrapping. Consequently, if the pixel-based phase difference is larger than  $2\pi$ , the wrapping phase becomes discontinuous and ambiguous. To obtain a continuous phase map, a phase unwrapping process for subtracting or adding the multiple values of  $2\pi$  is needed to reconstruct the non-ambiguous phase map. The Gold-Stein phase unwrapping algorithm was employed to extract the unwrapped phase map [3]. Then, the unwrapped phase difference can be transferred to a height map using the following phase to height transformation.

The phase-to-height transformation is performed using the triangulation technique. The height information of object surface  $h(x, y)$  of each pixel is carried by its phase difference value  $\Delta\phi(x, y)$ , which is measured by the phase difference between the measured phase on the object surface and the reference plane. Since the measured height is generally much smaller than the projection height of the optical system and the projecting objective is telecentric, the object height has a linear relation with the unwrapped phase difference as shown in Equation (3). Thus, the phase value can be transformed by a linear transformation coefficient  $K$ , and the height information of the object is then given as:

$$h(x, y) = K\Delta\phi(x, y) \tag{3}$$

### 2.2.2. Dual-Sensing Technique

As mentioned above, when the projected light beam hits the tested wafer surface, the reflected light beams may contain both diffuse and specular components (see Figure 2). The diffuse components are mainly formed by reflected light beams from micro structures on wafer surface with high surface roughness. Each reflected beam may interfere with other beams. However, there is no light interference due to low light coherence. In the optical design, the diffuse light intensity is reflected at a wide reflecting angle and can be detected by a vertical sensing camera. On the other hand, specular light is generated by the reflection of incident light from an object with a shiny reflective surface on a flat wafer surface having very low surface roughness. According to the law of reflection, specular light can be detected by an optical objective located at a reflectance angle equal to that of the incident light.

To capture both diffuse and specular components simultaneously, a dual-sensing optical system is developed. In the configuration, a camera (10) is dedicated to capturing diffuse components with another camera (11) designated for capturing specular components. To reconstruct the 3-D shape of the measured surface, the reflected fringe images representing both diffuse and specular beams are simultaneously captured using the proposed system, illustrated in Figure 3. By controlling the adequate duration of light exposure required by the individual cameras, two deformed fringe images with high contrast can be detected. The entire 3-D shape of the measured target surface can be effectively synthesized from the two individual images detected and reconstructed by accurate pixel spatial mapping between two sensors. Details of the proposed measuring system are illustrated by the flowchart shown in Figure 4.

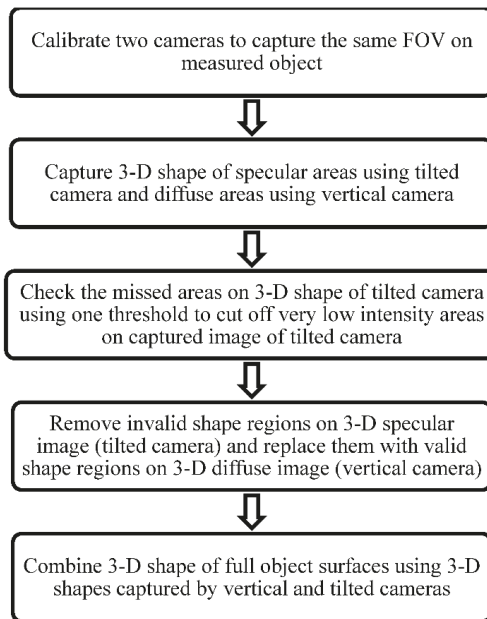


Figure 4. Flowchart of the proposed measuring technique.

In the proposed technique, the 3-D shape region detected and established from diffuse light can be well synthesized with the 3-D shape region detected and established from specular light. The two surface regions detected can be further merged seamlessly using a camera calibration procedure.

A special artifact target is designed and employed as a special mapping target. Two cameras can be calibrated and merged into the same measuring field of view (FOV) on the tested surface. Details of the proposed calibration procedure are illustrated by the flowchart shown in Figure 5.

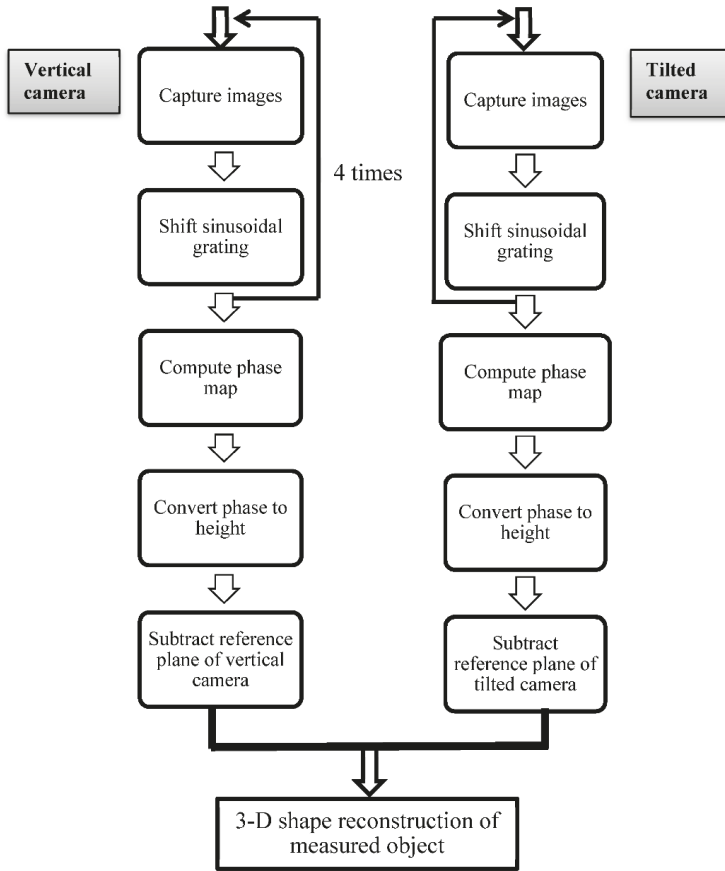


Figure 5. Flowchart of the proposed calibration procedure.

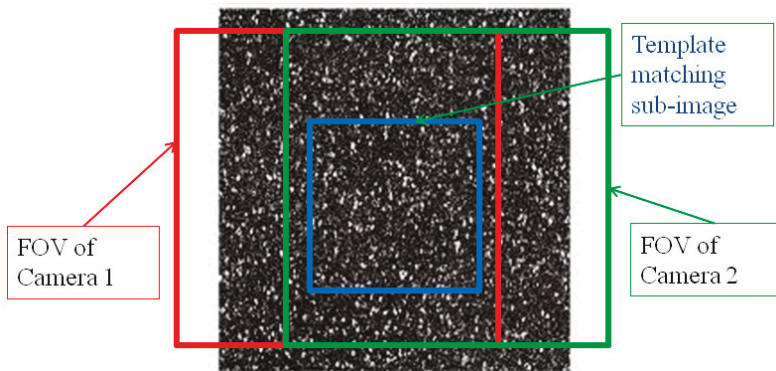
### 2.2.3. Dual Camera Mapping

To achieve precise mapping and calibration, an automatic and accurate mapping algorithm was developed to align the two cameras into the same FOV on the measured object. Figure 6 illustrates the proposed idea of dual camera mapping using random speckles. When mapping images captured by the two cameras, a unique speckle pattern is used and a cross-correlation algorithm is employed to compute the corresponding coefficient. In this study, the following normalized cross-correlation (NCC) algorithm is employed to compute the correlation between images captured by the two cameras.

$$R(x, y) = \frac{\sum_{x',y'} (T(x', y') \cdot I'(x + x', y + y'))}{\sqrt{\sum_{x',y'} T(x', y')^2 \cdot \sum_{x',y'} I'(x + x', y + y')^2}} \quad (4)$$

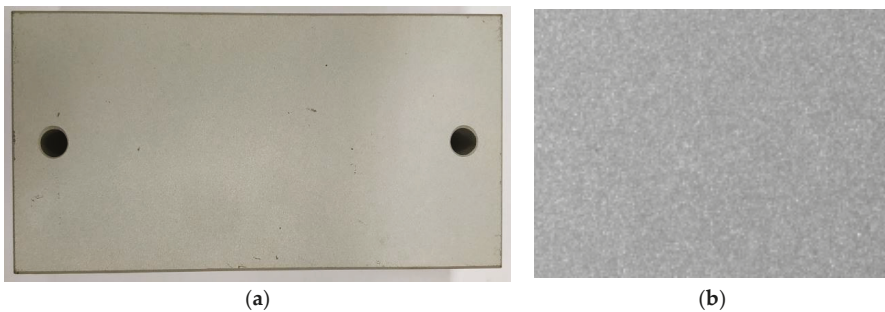
where:

$$I'(x + x', y + y') = I(x + x', y + y') - \frac{1}{w \cdot h} \cdot \sum_{x'',y''} I'(x + x'', y + y'')$$



**Figure 6.** Proposed idea of dual camera mapping using random speckles.

The correlation value ranges from 0.0 to 1.0. The matching score of 1.0 means 100% matching between sub-image and template image. To ensure that the dual camera mapping technique worked well, a correlation threshold is needed to estimate the matching score. It is also important to note that the sub-image size is set to be large enough to retain the accuracy of the matching process. When the sub-image size is too small, the robustness of the correlation technique cannot be assured. In practice, the sub-image size is chosen as a quarter of the captured FOV to achieve good matching and reasonable computation efficiency. To achieve precise image mapping between the two images, a random speckle pattern can be prepared and employed by projecting uniform blue light on an electrostatic-coating aluminum surface with micro random patterns on it. The unique and random speckle pattern can provide accurate image registration for image mapping. Figure 7 shows the aluminum surface target designed and one of the speckle images prepared for image registration.



**Figure 7.** Calibration target design for dual camera image mapping: (a) white painted aluminum target; (b) captured speckle image using the vertical camera.

Figure 8 shows the flowchart of the proposed dual camera mapping technique. One sub-image (a quarter of captured image) of the vertical camera is cropped in the image center and selected as a template image. The matching score between the tilted image and template image (vertical camera) is determined using the NCC algorithm. The shifted value between the template image and that captured by the tilted camera can be calculated along the X and Y directions of image coordinates. This parameter can accurately represent the spatial location difference between the FOV of the tilted and the vertical cameras. This process is repeated until the parameter converged to a preset threshold for completion.

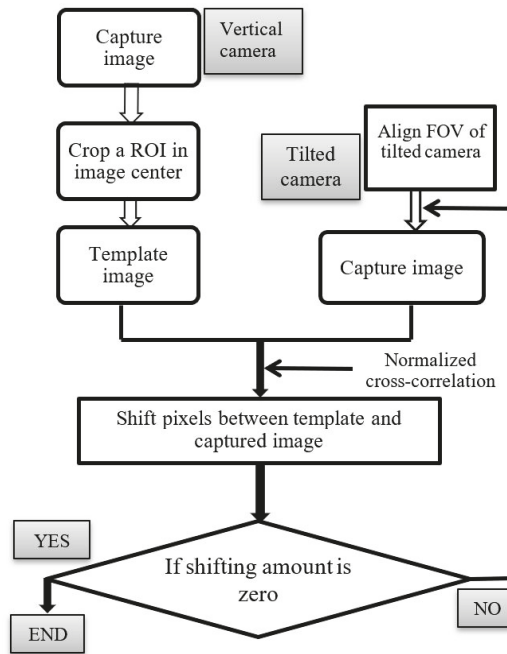
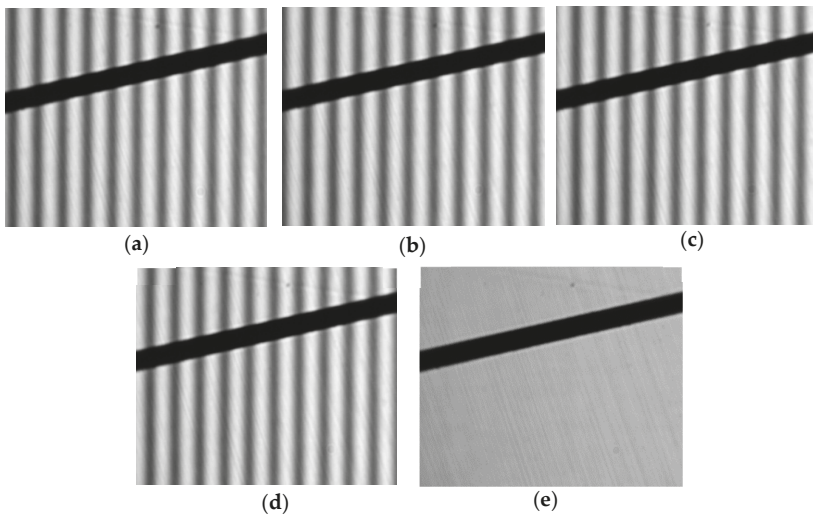


Figure 8. Flowchart of the proposed dual camera mapping procedure.

#### 2.2.4. Segmentation of Specular and Diffuse Images

An image averaging algorithm is developed to extract specular and diffuse measuring regions with high S/N ratio from the acquired deformed fringe images. Removal of the projected fringe from the acquired fringe is essential for segmentation because these deformed fringes contain light intensity modulation, thus making segmentation difficult. In phase shifting, several phase-shifted sinusoidal patterns are projected onto the tested object and their corresponding deformed fringe images are acquired sequentially. A background image is defined here as an illuminated object image without structured fringes. To obtain the background image of these deformed fringe images, the image averaging algorithm can be performed by simply summing all the phase-shifted images to remove the fringe modulation and generate the background image.

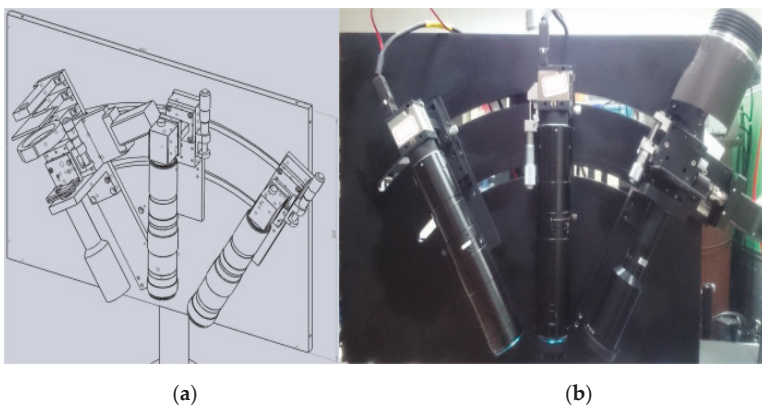
It is important to note that the above algorithm can reliably work even when the reflectivity of the tested surface is extremely variant. The algorithm operates on every individual pixel for the summing; hence, the result is generally insensitive to the pixel intensity. In an experimental test on the algorithm, the sinusoidal projection grating was made by a precise lithography process and a semiconductor’s wafer surface having laser-etched groves was tested. A precise linear stage with positioning repeatability less than 1  $\mu\text{m}$  was adopted to perform four-step phase shifting. Figure 9a–d illustrates the captured images from the four-step fringe projection with phase-shift 0°, 90°, 180°, and 270°, respectively. As can be seen, all the projected fringes are effectively removed with the image averaging algorithm, and the illustrated object image (also called the background image) is extracted. The experiment result shown in Figure 9 was acquired by the tilted camera for specular light detection on a smooth wafer surface. An excellent image contrast between specular and diffuse measured regions can be achieved in the background image. The same can be performed on the diffuse images to achieve equal effectiveness.



**Figure 9.** Phase-shifted images acquired from four-step phase shifting operation and image summing algorithm: (a) Phase-shift 0°; (b) Phase-shift 90°; (c) Phase-shift 180°; (d) Phase-shift 270°; and (e) Image result of average 4 channels.

### 3. Experimental Design and Result Analyses

To demonstrate the feasibility and measurement performance of the proposed technique, a 3-D measurement system consisting of two sensing cameras, a single-frequency transmission grating with a period of 24 line pairs per millimeter, and a step motor with positioning accuracy less than 1  $\mu\text{m}$  was developed. Two telecentric lenses with the same optical specification are integrated in front of the two cameras as the optical objective to collimate the reflected diffuse and specular light from the measured object surface. The light module was powered by a 30-Watt white light LED, in which the LED is mounted on a specially designed mechanical radiator for good heat dissipation. A SolidWorks design and hardware of the housing and optical system were implemented and shown in Figure 10a,b, respectively. The measuring FOV can reach  $1.0 \times 1.0 \text{ mm}$  for every measurement. A maximum measuring depth of the system can reach 1.0 mm which is sufficient for most microstructure profile inspection.



**Figure 10.** Developed system: (a) SolidWorks design; (b) system hardware.

Figure 11a,b shows the fringe images of the tested wafer surface captured using the tilted and vertical cameras, respectively. Figure 11c,e illustrates the phase map and one height cross-section of the specular wafer surface measured using the tilted camera. Similarly, the phase map and one height cross-section for the laser-etched groove surface regions are measured using the vertical camera simultaneously and shown in Figure 11d,f. The 3-D profile of groove surface region can be well measured and reconstructed (see Figure 11h). It is important to note that the 3-D measured shape area of the groove region has random noises due to the non-detectable diffuse light from the laser groove surface region. Nevertheless, this region can be totally recovered and reconstructed by another vertical camera which can capture the diffuse light from the groove region. Two detected 3-D profiles, shown in Figure 11g,h, are extracted from their respective original images by the image summing algorithm and further mapped by the camera calibration technique precisely to reconstruct a full-field 3-D profile of the measured wafer, shown in Figure 11i. These results indicate that the full-field 3-D wafer surface can be measured and reconstructed accurately using the developed technique.

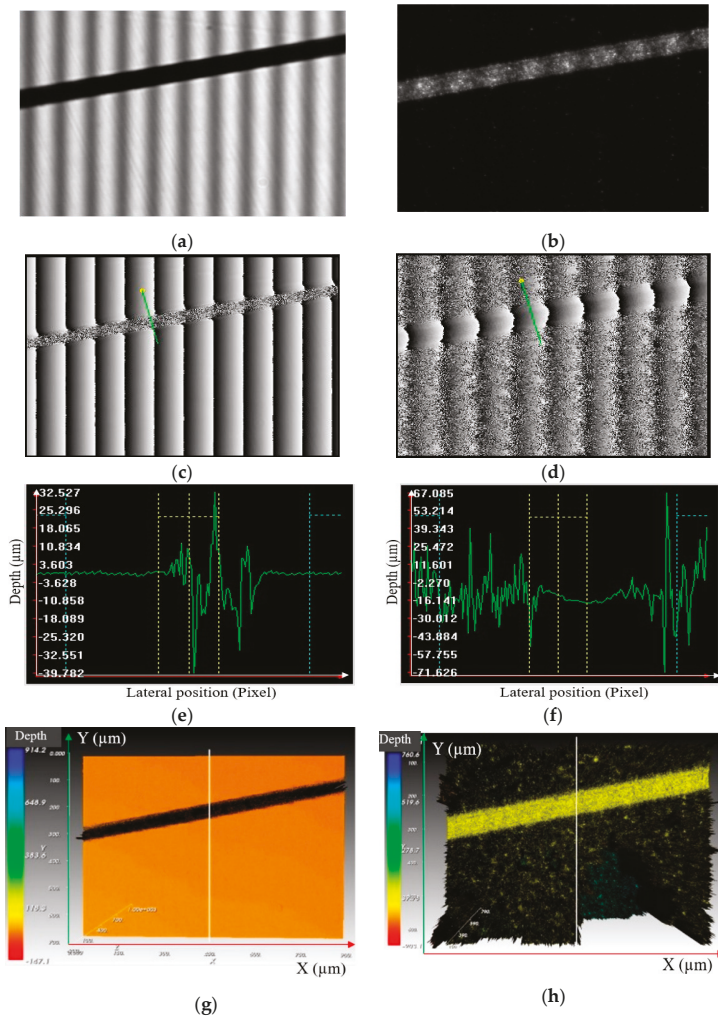
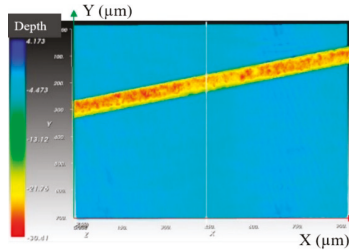


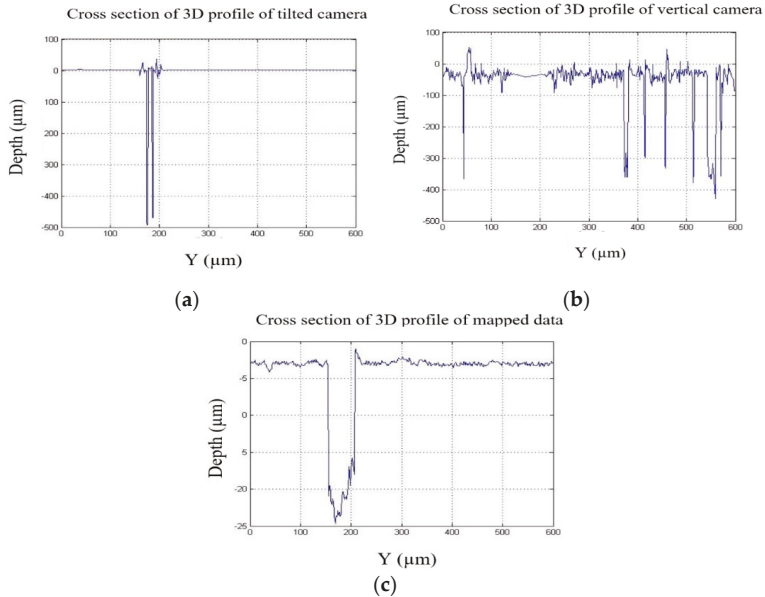
Figure 11. Cont.



(i)

**Figure 11.** Measurement results of the tested wafer sample shown in Figure 1: (a) captured deformed fringe image of specular light, (b) captured deformed fringe image of diffuse light, (c) wrapped phase map of (b,d) wrapped phase map of (c,e) reconstructed cross-section of (c,f) reconstructed cross-section of (d,g) reconstructed 3-D map of the smooth wafer surface, (h) reconstructed 3-D map of the laser-etched groove surface, and (i) reconstructed 3-D shape of the whole tested wafer surface.

Figure 12 exhibits the measured height cross-section along one lateral section on the tested wafer being captured by the tilted and vertical cameras. The standard deviation of the wafer surface region and groove area are evaluated for measurement precision. In a 30-time repeatability test, the standard deviation in the measured data region, shown in Figure 12a,b, are 0.265  $\mu\text{m}$  and 1.937  $\mu\text{m}$ , respectively.



(a)

(b)

(c)

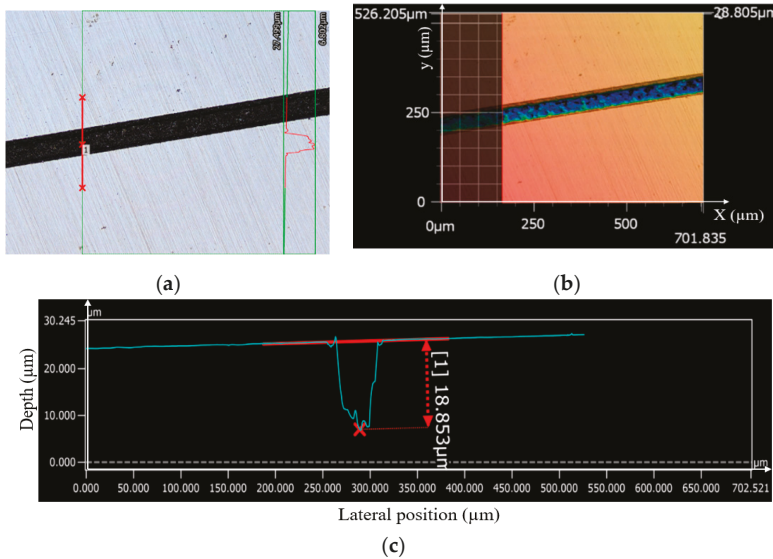
**Figure 12.** Height cross-section of a surface cross-section on the wafer in Figure 1: (a) reconstructed cross-section profile by capturing specular light. (b) reconstructed cross-section profile by capturing diffuse light, and (c) mapped cross-section profile.

The results obtained show that the measurement repeatability can be achieved for less than 2.0  $\mu\text{m}$  when measuring a rough laser-machined wafer surface. With the proposed technique, the measured average step height of the cutting groove on the wafer is 19.01  $\mu\text{m}$ .

To check the measurement accuracy, the measured result was compared with a pre-calibrated laser confocal microscope. Figure 13 illustrates the measurement result obtained using a laser confocal

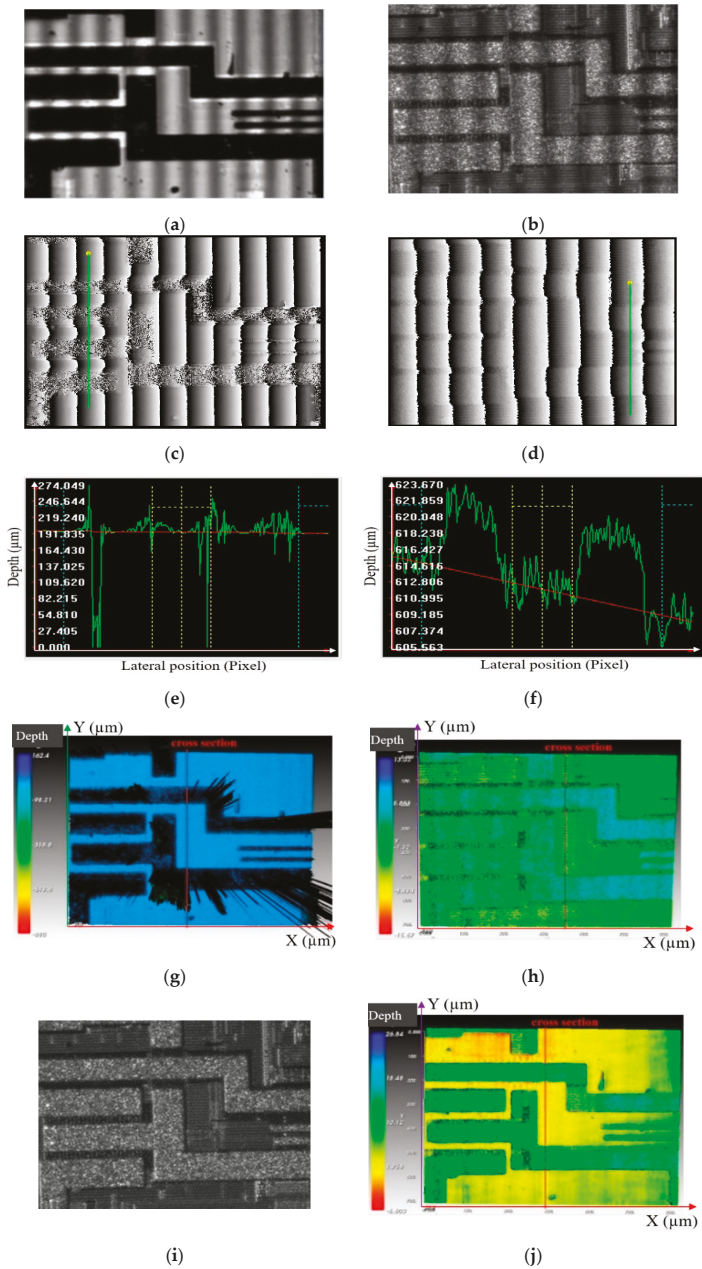


microscope (KEYENCE VK-X1000) on the same measured wafer surface region. The measured average step height of the laser-etched groove in a 30-time measurement was 18.853  $\mu\text{m}$ . With the developed method, the measured average step height on the same wafer surface was 20.21  $\mu\text{m}$ , in which the measurement deviation in this case was 1.357  $\mu\text{m}$ .



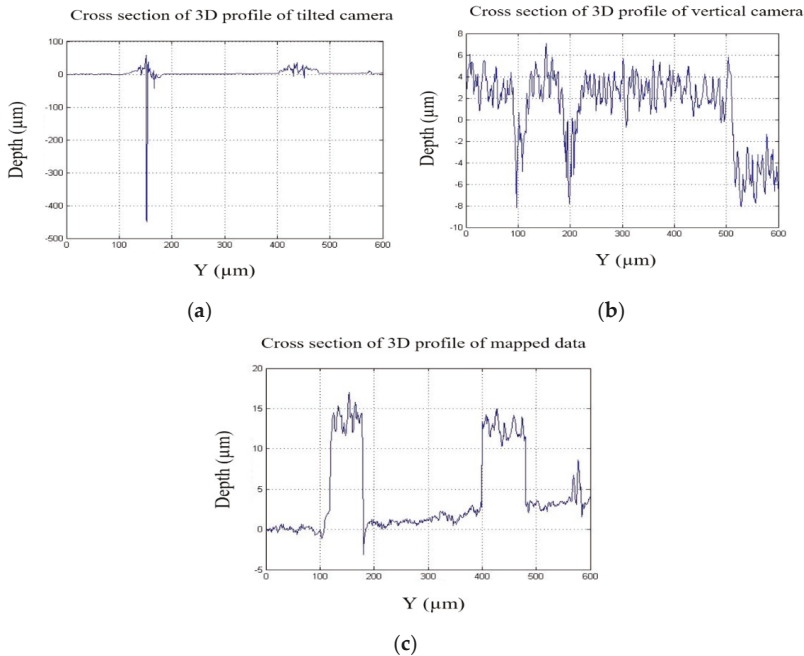
**Figure 13.** Measurement results of the wafer obtained using a laser confocal microscope (KEYENCE VK-X1000): (a) 2D captured image of the wafer, (b) 3-D measured profile of wafer, and (c) one cross-section of the 3-D profile.

To further check the feasibility of the developed method on an industrial semiconductor parts with variant surface reflectivity, an IC chip was selected for measuring its surface profile. Figure 14 shows the measured results on an industrial IC surface with extreme reflectivity variance between the substrate and the metal pads. The substrate image detected by specular light is shown in Figure 14a while the diffuse image detected by diffuse light is illustrated in Figure 14b. Neither of these two images can cover the entire surface detection. With the developed method, both specular and diffuse images can be detected simultaneously and reconstructed accurately. The phase map and one surface cross-section profile can be reconstructed for smooth substrate surface regions, as seen in Figure 14c,e, respectively. Similarly, Figure 14d,f represent respectively the phase map and one surface cross-section profile of the metal-pad regions. Compared with the developed method, the traditional fringe projection method will find excessive noises measured around the IC chip's pads which are mainly reflected diffuse light areas, as seen in Figure 14g. By using the developed technique, the surface measurement of the IC chip can be successfully obtained, in which both the pad and substrate areas can be reconstructed simultaneously and accurately, as shown in Figure 14j. In other words, a significant improvement in surface profilometry of objects with substantial variance of surface reflectance can be achieved with the developed approach.



**Figure 14.** Measurement results of industrial IC chip: (a) image captured by camera receiving specular light; (b) image captured by vertical camera; (c) wrapped phase map of tilted camera; (d) wrapped phase map of vertical camera; (e) cross-section on wrapped phase map of tilted camera; (f) cross-section on wrapped phase map of vertical camera; (g) 3-D shape of IC chip measured by tilted camera; (h) 3-D shape of IC chip measured by vertical camera; (i) measured ROI of industrial IC chip; and, (j) 3-D shape of measured ROI in (i).

Figure 15a,b illustrate the cross-section of height map at Column 400 on the IC chip imaged by the two cameras in charge of capturing specular and diffuse light, respectively. With the mapping algorithm, the merged cross-section profile can be obtained, as seen in Figure 15c. The measurement standard deviations of the specular and diffuse surfaces on the IC chip are 0.602 and 1.285  $\mu\text{m}$ , respectively. With the proposed technique, the measured average step height of the metal pads on the IC chip is 13.2  $\mu\text{m}$ .



**Figure 15.** Cross-section profile measured on IC chip shown in Figure 14: (a) cross-section profile reconstructed from specular light components; (b) cross-section profile reconstructed from diffuse light components; (c) 3-D cross-section surface mapped by proposed mapping algorithm.

#### 4. Conclusions

In this study, a new dual-sensing measuring technique for 3-D surface profilometry is developed and proved effective for measuring objects with extreme surface reflectance variance. The key element of the developed technique is its new optical configuration for simultaneous detection of both diffuse and specular light reflected from various surface regions on a tested object. Moreover, an accurate image segmentation and mapping algorithm is developed to extract effectively detected image regions and merge them seamlessly into an accurate 3-D surface map. Experimental results indicate that the proposed technique can successfully achieve simultaneous measurement with repeatability of one standard deviation below 0.3  $\mu\text{m}$  on a specular surface and 2.0  $\mu\text{m}$  in a total vertical measurable range of 1.0 mm on a diffuse surface. Specular light detection is more accurate than diffuse light detection in 3-D surface profilometry, which is mainly influenced by the SNR (signal-to-noise ratio) of the deformed fringe image. The developed technique can provide an effective wafer surface measuring technique for in-situ AOI, especially in 3-D full-field surface profilometry.

**Author Contributions:** Conceptualization, L.-C.C.; Methodology, L.-C.C. and D.-H.D.; Software, L.-C.C. and D.-H.D.; Validation, L.-C.C. and D.-H.D.; Formal Analysis, L.-C.C. and D.-H.D.; Investigation, L.-C.C., D.-H.D. and C.-S.C.; Resources, L.-C.C.; Data Curation, L.-C.C. and D.-H.D.; Writing-Original Draft Preparation, L.-C.C.

and D.-H.D.; Writing-Review & Editing, L.-C.C.; Visualization, L.-C.C.; Supervision, L.-C.C. and C.-S.C.; Project Administration, L.-C.C. and C.-S.C.; Funding Acquisition, L.-C.C.

**Funding:** This research was funded by Ministry of Science and Technology, Taiwan, under grant no. MOST 107-2218-E-002-060 and MOST 107-2218-E-002-002.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Gåsvik, K.J. *Optical Metrology*, 3rd ed.; John Wiley & Sons Ltd., Ed.; Wiley Online Library: Chichester, UK, 2003; pp. 173–190.
2. Takeda, M.; Mutoh, K. Fourier transform profilometry for the automatic measurement of 3-D object shapes. *Appl. Opt.* **1983**, *22*, 3977–3982. [[CrossRef](#)] [[PubMed](#)]
3. Ghiglia, D.C.; Pritt, M.D. *Two-Dimensional Phase Unwrapping: Theory, Algorithms, and Software*; John Wiley & Sons Ltd., Ed.; Wiley Online Library: New York, NY, USA, 1998; ISBN 978-0-471-24935-1.
4. Chen, L.-C.; Nguyen, X.-L.; Ho, H.-W. High-speed 3-D machine vision employing Fourier Transform Profilometry with digital tilting-fringe projection. In Proceedings of the 2008 IEEE Workshop on Advanced robotics and Its Social Impacts, Taipei, Taiwan, 23–25 August 2008; pp. 1–6.
5. Gdeisat, M.A.; Abid, A.; Burton, D.R.; Lalor, M.J.; Lilley, F.; Moore, C.; Qudeisat, M. Spatial and temporal carrier fringe pattern demodulation using the one-dimensional continuous wavelet transform: Recent progress, challenges, and suggested developments. *Opt. Lasers Eng.* **2009**, *47*, 1348–1361. [[CrossRef](#)]
6. Gorecki, C. Interferogram analysis using a Fourier transform method for automatic 3D surface measurement. *Pure Appl. Opt. J. Eur. Opt. Soc. Part A* **1992**, *1*, 103. [[CrossRef](#)] [[PubMed](#)]
7. Lin, J.-F.; Su, X. Two-dimensional Fourier transform profilometry for the automatic measurement of three-dimensional object shapes. *Opt. Eng.* **1995**, *34*, 34–36.
8. Chen, L.-C.; Ho, H.-W.; Nguyen, X.-L. Fourier transform profilometry (FTP) using an innovative band-pass filter for accurate 3-D surface reconstruction. *Opt. Lasers Eng.* **2010**, *48*, 182–190. [[CrossRef](#)]
9. Chen, L.-C.; Hai, H.H. Fourier transform profilometry employing novel orthogonal elliptic band-pass filtering for accurate 3-D surface reconstruction. *Precis. Eng.* **2014**, *38*, 512–524. [[CrossRef](#)]
10. Chen, L.-C.; Ho, H.-W. Method for Acquiring Phase Information and System for Measuring Three Dimensional Surface Profiles. Patent US8385629 B2, 26 February 2013.
11. Wyant, J.C. White light interferometry. *Proc. SPIE* **2002**, *4737*, 98–107.
12. Artusi, A.; Banterle, F.; Chetverikov, D. A survey of specular removal methods. In *Computer Graphics Forum*; Wiley Online Library: New York, NY, USA, 2011; Volume 30, pp. 2208–2230.
13. Kim, D.W.; Lin, S.; Hong, K.-S.; Shum, H.Y. Variational specular separation using color and polarization. In Proceedings of the IAPR Workshop on Machine Vision Applications, Nara, Japan, 11–13 December 2002; pp. 176–179.
14. Umeyama, S.; Godin, G. Separation of diffuse and specular components of surface reflection by use of polarization and statistical analysis of images. *IEEE Trans. Pattern Anal. Mach. Intell.* **2004**, *26*, 639–647. [[CrossRef](#)] [[PubMed](#)]
15. Ma, W.-C.; Hawkins, T.; Peers, P.; Chabert, C.-F.; Weiss, M.; Debevec, P. Rapid Acquisition of Specular and Diffuse Normal Maps from Polarized Spherical Gradient Illumination. In Proceedings of the 18th Eurographics Conference on Rendering Techniques, Grenoble, France, 25–27 June 2007; Eurographics Association: Aire-la-Ville, Switzerland, 2007; pp. 183–194.
16. Tan, R.T.; Ikeuchi, K. Separating reflection components of textured surfaces using a single image. *IEEE Trans. Pattern Anal. Mach. Intell.* **2005**, *27*, 178–193. [[CrossRef](#)] [[PubMed](#)]
17. Zhang, S.; Yau, S.-T. High dynamic range scanning technique. In Proceedings of the Mechanical Engineering Conference Presentations, San Diego, CA, USA, 27–31 July 2008.
18. Song, Z.; Chung, R.; Zhang, X. An Accurate and Robust Strip-Edge-Based Structured Light Means for Shiny Surface Micromasurement in 3-D. *IEEE Trans. Ind. Electron.* **2013**, *60*, 1023–1032. [[CrossRef](#)]
19. Knauer, M.C.; Kaminski, J.; Hausler, G. Phase measuring deflectometry: A new approach to measure specular free-form surfaces. *Proc. SPIE* **2004**, *5457*, 5411–5457.
20. Bothe, T.; Li, W.; von Kopylow, C.; Juptner, W.P.O. High-resolution 3D shape measurement on specular surfaces by fringe reflection. *Proc. SPIE* **2004**, *5457*, 5412–5457.

21. Nguyen, M.T.; Ghim, Y.-S.; Rhee, H.-G. Single-shot deflectometry for dynamic 3D surface profile measurement by modified spatial-carrier frequency phase-shifting method. *Sci. Rep.* **2019**, *9*, 3157. [[CrossRef](#)] [[PubMed](#)]
22. Wiedenmann, E.; Scholz, T.; Schott, R.; Tusch, J.; Wolf, A. First Utilization of Energy Transfer in Structured Light Projection—Infrared 3D Scanner. *Key Eng. Mater.* **2014**, *613*, 141–150. [[CrossRef](#)]
23. Ihrke, I.; Kutulakos, K.; Lensch, H.P.A.; Magnor, W.H. State of the Art in Transparent and Specular Object Reconstruction. In Proceedings of the STAR Proc. of Eurographics, Crete, Greece, 14–18 April 2008.
24. Wolff, L.B. Using polarization to separate reflection components. In Proceedings of the Proceedings CVPR '89: IEEE Computer Society Conference on Computer Vision and Pattern Recognition, San Diego, CA, USA, 4–8 June 1989; pp. 363–369.
25. Lamond, B.; Peers, P.; Ghosh, A.; Debevec, P.D. Image-based Separation of Diffuse and Specular Reflections using Environmental Structured Illumination. In Proceedings of the 2009 IEEE International Conference on Computational Photography (ICCP), San Francisco, CA, USA, 16–17 April 2009; IEEE: San Francisco, CA, USA, 2009.
26. SCILLC. *Soldering and Mounting Techniques. Reference manual*; Rev. 11, SOLDERRM/D; On semiconductor Ltd.: Phoenix, AZ, USA, March 2016; pp. 1–82.



© 2019 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).

Article

# An Accuracy-Efficiency-Power Consumption Hybrid Optimization Method for CNC Milling Process

Shih-Ming Wang <sup>1,2,\*</sup>, Chun-Yi Lee <sup>1</sup>, Hariyanto Gunawan <sup>1</sup> and Chin-Cheng Yeh <sup>1</sup>

<sup>1</sup> Department of Mechanical Engineering, Chung Yuan Christian University, Taoyuan 32023, Taiwan; sasuke0673@gmail.com (C.-Y.L.); harrywey@cycu.edu.tw (H.G.); jano.yeh@shl-group.com (C.-C.Y.)

<sup>2</sup> Technology Center, Walsin Lihwa Corporation, Taipei 110, Taiwan

\* Correspondence: shihming@cycu.edu.tw

Received: 3 March 2019; Accepted: 8 April 2019; Published: 10 April 2019



**Abstract:** This study proposes a hybrid optimization method which can help users to find optimal cutting parameters which will provide better efficiency and lower power consumption for a milling process. Empirical models including performance-power consumption characteristic curves of servo motors were built, and an optimization algorithm adopting the empirical models with procedure guiding function was developed. The empirical models were built based on the measurements from planned machining experiments with different combination of machining parameters including spindle speed, feedrate, and chip load, etc. After integrating the models and algorithm, an optimization system with human machine interface, which has procedure guiding function, was developed. The system can recommend optimal machining parameters for a milling process for shorter machining time and lower electricity costs based on the original machining parameters. Finally, cutting experiments were conducted to verify the proposed system, and the results showed that the proposed method can effectively enhance efficiency by 42.06% and save 34.74% in machining costs through reducing machining time and electrical power consumption.

**Keywords:** optimization; power consumption; machining efficiency; machining cost

## 1. Introduction

Increasing power efficiency in production is a prime target for many companies, for a number of different reasons. Power consumption is an important factor in operating cost. An efficiency power consumption system with high-performance output and higher productivity requiring less power is expected. In machine tool production, Computer Numerical Control (CNC) machine is usually used to produce the precision product. When the machining capacity is higher and machining time is longer, significantly higher power consumption will occur. In the past, enterprises use the shortest machining time to increase machine utilization and productivity. However, the shortest machining time does not always mean the lowest cost, because more power consumption is needed to finish the machining in a shorter time. As known, the energy crisis is accelerating, and energy costs are rapidly increasing so it will be beneficial if machining time was shorter and simultaneously, the power consumption lower. Moreover, most of the factories and power companies have signed an electric power contract, so if electricity consumption exceeds the contract, the cost of electricity will be greatly increased. Therefore, if the machine power consumption characteristics can be known, it can help the planning of production scheduling and cost.

In recent years, research on power consumption has grown rapidly. Gontarz et al. [1] introduced a new physical modeling method to predict the effects of power consumption and thermal effect for different modes of the turning machine tool and its components. A modeling framework was established and used to simulate power consumption. Abdoli et al. [2] explored the relationship

between cutting parameters and machine power consumption. The power consumption model was established based on this relationship, and the model can be used for production line planning. Diaz et al. [3] estimated the cutting power consumption and spindle motor power consumption through an empirical formula. Mori et al. [4] suggested reducing the power consumption by reducing the machining time or synchronizing the spindle speed acceleration/deceleration and the feedrate system. Draganescu et al. [5] proposed a model to calculate power consumption, machine efficiency and material removal rate for metal cutting. To establish the relationship between the above parameters, the experimental data and response surface methodology (RSM) were used. Kara et al. [6] proposed an empirical model to describe the relationship between power consumption of the material removal process and process variables. The power consumption model for calculating the material removal rate and the amount of cutting was developed.

To carry out the above development, extensive experimental research and data analysis of machine power consumption characteristics are necessary. Many companies will collect various data, such as production data, quality data, production environment data, etc., to analyze production management and develop decision based policy. However, the enterprises rarely collect production parameters and power consumption data for analysis. Hence, when an enterprise wants to control power consumption, they lack the necessary data support. Although some enterprises are concerned about power consumption, it is manually collected and calculated. In addition, the power consumption data acquisition and analysis method is lacking. Some researches related to process monitoring have developed, for example, Dimla et al. [7] used various sensor signals to develop cutting monitoring of metal cutting. Usui et al. [8] proposed an analytical method based on orthogonal cutting data for machining and tool wear characteristics, to predict various tool shape and cutting conditions. Prickett et al. [9] used a machine controller to monitor the process and detect tool breakage signals. Cai et al. [10] proposed a feedback method to predict tool wear, milling error and online compensation. Li et al. [11] used sensor technology to detect tool wear and faults, and predict the impact of tool wear to surface quality. Choi et al. [12] used an intelligent online system to monitor tool wear.

Although some researches on power consumption have been performed, they usually only focus on power consumption or processing independently, and still lack a method which can synchronously optimize the machining accuracy, production efficiency, and power saving. The objective of this research is to establish the hybrid optimization of machining accuracy, production efficiency, and power consumption. The characteristics of motor power consumption during machining were investigated. Furthermore, the relationship between machining parameters and motor power consumption was created. Moreover, the optimization algorithm of production efficiency and power saving was built. Finally, the hybrid optimization system was created by using Visual C# language. With this system, the user can import the Numerical Control (NC) program to optimize the parameters. Then the system will analyze and calculate the optimal machining time, power consumption, and recommend the machining parameters for better efficiency, lower power consumption, with the same product quality.

## **2. Algorithm and Method**

To determine the optimal machining parameters for better machining efficiency and lower power consumption without losing machining accuracy, it is necessary to understand the relationship between machine power consumption characteristics, machining parameters, and machining quality. Empirical models were first built based on the data (power consumption vs. motor speed or power consumption vs. machining feedrate) collected from the designed experiments. To separate the influence of cutting in motor power consumption, the data collection experiments were divided into two categories: (1) measurements without cutting load and (2) measurements with cutting load. Based on the collected data, a motor power consumption characteristics model with respect to machining parameters was built. Based on the model, the algorithm to search the machining parameter for shorter machining time and lower power consumption was developed. When a milling process and expected improvement of machining efficiency are given, the associated machining feedrate and spindle speed will be calculated

based on the theoretical cutting equation, and used as initial values to search based on the motor power consumption characteristics model for the optimal parameters which have a lower motor power consumption.

To analyze the relationship between electrical power consumption and machining conditions, the investigation and analysis of spindle motor and servo motor were important. The analysis process, first, used various machining parameters to observe the motor power consumption to understand the motor characteristics and power consumption, then established the relationship model between machining parameters and power consumption, the synchronous optimization algorithm of production efficiency and power consumption was established, and a human machine interface (HMI) was designed, to become the optimization manufacturing assisted system. The main function is the optimization of machining parameters to obtain power and production efficiency without decreasing the machining accuracy. Figure 1 shows the flowchart of the research.

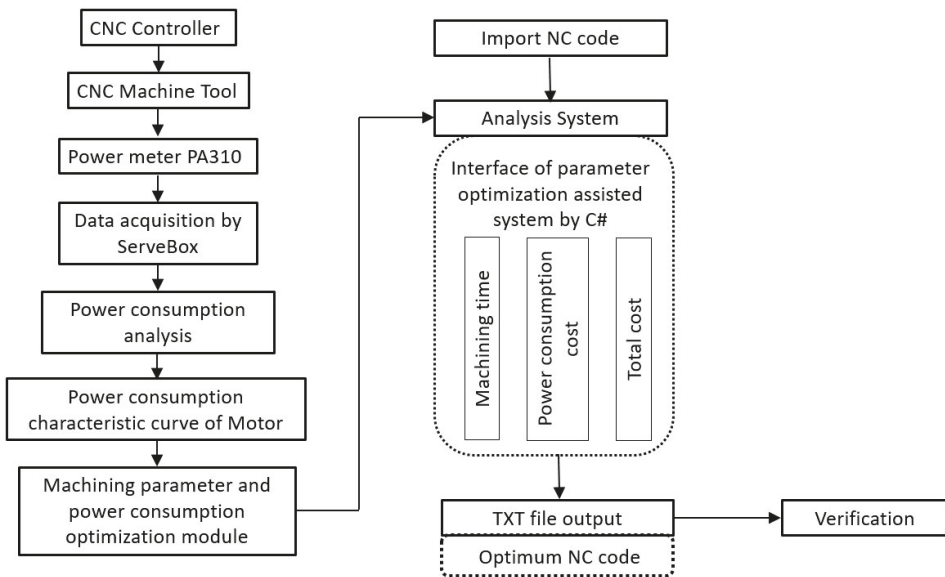


Figure 1. Flowchart of research. CNC: Computer Numerical Control; NC: Numerical Control.

Different materials with different machining conditions will cause different electrical power consumption, and the reasonable chip load of the cutter for different materials will be different too. Although the method and algorithm proposed in this study can be applied to different materials, the performance–power consumption characteristic curves of servo motors (such as power consumption vs. motor speed, or power consumption vs. machining feedrate) need to be investigated and built for different materials as core important data/information for implementation.

### 2.1. Optimization of Power Consumption

The cutting speed in the cutting process is proportional to the spindle motor speed, and the cutting feedrate is proportional to the servo motor speed. To shorten the machining time for improving machining efficiency, the feedrate of the servo motor needs to be increased. However, increasing the feedrate of the servo motor will increase the chip load and cutting resistance, hence, it will affect the tool life and machining accuracy. To avoid this effect, the spindle speed needs to be increased at the same time, so the chip load is kept maintained. However, the increasing of spindle speed and feedrate cause an increase in power consumption. In this study, power consumption is measured for different



spindle speeds, then the influence of machining parameters on machine power consumption is created for the establishment of an optimized design algorithm.

According to the spindle speed–motor power consumption characteristic curve and feedrate–motor power consumption characteristic curve, the model of the relationship between machining parameters and power consumption can be established by the polynomial curve fitting method, then analyzed to determine the influence of each parameter on machine power consumption and the influence of the combination of parameters on machine power consumption. To reduce machining time by increasing the servo feedrate (i.e., increasing the servo motor speed) without affecting the tool life and machining accuracy, it is necessary to maintain the cutting load per tooth (chip load), so the spindle speed needs to be increased.

In the optimization procedure, first of all, the user needs to roughly define the expected increase of feedrate for better machining efficiency, and the system will then search the optimal feedrate close to the expected feedrate, which has lower electrical power consumption for the related servo motors, based on the pre-built empirical model. After the optimal feedrate is selected, the spindle speed is calculated to remain the same chip load for the cutter to avoid deteriorating the cutter life. Since the optimal feedrate was selected based on the goal of efficiency improvement, the optimal combination of parameters can provide better efficiency, better electrical power consumption, and almost the same tool life. The relationship of chip load, spindle speed, and feedrate is expressed in Equation (1).

$$C_p = F / (N \times z) \tag{1}$$

where  $C_p$  is the chip load (mm/tooth),  $F$  is the feedrate (mm/min),  $N$  is the spindle speed (rpm), and  $z$  is the number of teeth.

When optimizing machining, the required spindle speed is entered to the machining parameters–power consumption module to calculate the power consumption change. To make the power saving and machining cost of production easy to understand for industry, the machine power consumption is converted into electricity cost. The machining efficiency can be determined as machining cost saving, which can be obtained by multiply the machining time saving and machining cost as shown in Equation (2).

$$M_{cr} = M_t \times \text{hr} \tag{2}$$

where  $M_{cr}$  is machining cost saving,  $M_t$  is machining time saving (hr), and hr is machining cost (NTD/hr).

Figure 2 shows the optimization algorithm of machining and power consumption. If the electricity cost of the shortened machining time exceeds the basic electricity cost, the system re-finds the machining parameters to obtain optimal machining efficiency and power consumption based on Equation (1) and the spindle speed–motor power consumption module and the feedrate–motor power consumption module. The final machining time must be less than the initial machining time to achieve lower cost.

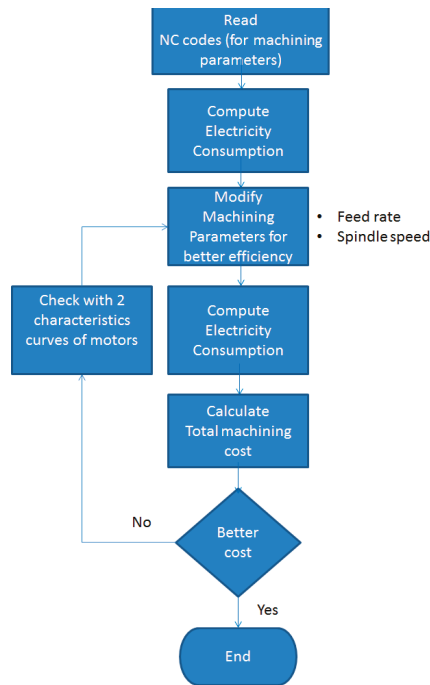


Figure 2. Optimization algorithm of machining and power consumption.

After establishing the empirical models for the relationship between machining parameters and power consumption, an optimization system was developed in visual C# language. The user can enter the machining information (such as original NC program, machining precision requirement, tool diameter, etc.) and the expected improvement in machining efficiency from the human machine interface (HMI). After that, the system will automatically calculate the recommended optimal machining parameters. In the procedure, the system will guide the user to enter the expected feedrate for optimization, and then the system will re-calculate and find the optimal machining parameters. Finally, the optimal parameters will be saved as a CSV file or TXT file, and meanwhile, the parameters can also be uploaded to the cloud database with a JSON data format. Through the Ethernet and Fanuc Open CNC API Specification (FOCAS), two-way communication between the CNC machine controller and hybrid optimization module was established for data acquisition and transferring command to the CNC controller of a machine tool.

## 2.2. Hybrid Optimization System

The HMI is designed and becomes the optimization manufacturing assisted system. The system can calculate the machining time, machining cost, and total cost as shown in Figure 3. The user needs to input regular electricity fees, base machining costs, and NC program, then the system will calculate the amount of three axes movements, and according to feedrate–power consumption model, analyze the machining time, machining cost, and total cost. The machining cost is calculated from power consumption (kW) multiplied machining time (hr). The user can enter the required machining parameters according to their needs, then the system will calculate the machining time as shown in Figure 3. The system can recommend optimal parameters, and the user can select this recommended parameter and directly modify and export the optimized parameters as the NC program.

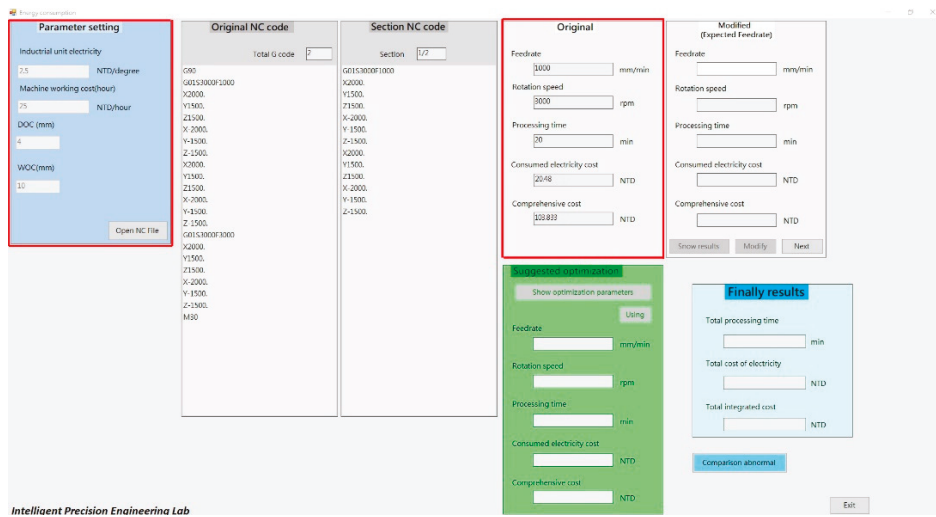


Figure 3. Machining efficiency and power consumption hybrid optimization human machine interface (HMI).

### 3. Experiments

#### 3.1. Experiment Design

Basically, higher speed will produce higher power consumption, so an experiment with various speeds of spindle motor and X-, Y-, Z-axis servo motor were carried out to study the characteristic of the motor power consumption. The experiment was divided into two parts: without cutting load and with cutting load conditions. Assuming that the spindle motor speed sets 500 until 3000 RPM with the increasing step of 500, a power consumption measurement is performed for the spindle motor under these spindle speeds, to obtain the relationship between spindle speed and power consumption. Similarly, assuming that the servo feedrate sets 200, 500, 1000, 1500, 2000, 2500, 3000, 3500, 4000, 4500, 5000, 5500, 6000 mm/min, a power consumption measurement is performed for the servo motor under these servo feedrates, to obtain the relationship between servo feedrate and power consumption.

For investigation of power consumption with cutting load, because of different cutting conditions, the cutting load is also different, hence, the motor will have different power consumptions. Therefore, different cutting parameters were taken to investigate the relationship between cutting loading and power consumption. The cutting conditions are shown in Table 1. The machining parameters are according to the tool material, hardness of workpiece material, milling speed, and material removal rate (MRR). By using the equation  $V = \pi \times D \times N / 1000$ , we can obtain the required spindle speed, and using equation  $F = f \times z \times N$ , we can obtain the feedrate (mm/min), where V is the milling speed (m/min); D is the diameter of tool (mm); N is the spindle speed (rpm); F is the feedrate (mm/min); f is the material removal rate (mm/t); z is the number of teeth (t).

Table 1. Cutting condition.

Workpiece Material	Cast Iron (HB244)
Workpiece dimensions	550 × 100 × 100 mm
Cutting tool	Carbide end mills
No. of teeth	4 teeth
Tool diameter	25 mm

The machining parameters are shown in Table 2. Eight kinds of spindle speed, three kinds of feedrate, three kinds of feed per tool were taken, so a total of 24 combinations of parameters. Each parameter was tested five times.

**Table 2.** Machining parameters.

Subject	Experiment 1	Experiment 2
Milling speed (m/min)	30~60	30~60
Chip load (mm/t)	0.12, 0.15, 0.18	0.12, 0.15, 0.18
Spindle speed (rpm)	450~750	1000~2750
Feed rate (mm/min)	192~540	480~1980
Width of Cut (mm)	20	20
Depth of Cut (mm)	4	4

### 3.2. Instrument and Data Acquisition

A CNC horizontal milling machine Dah Lih MCH-500 with X, Y, Z axis stroke  $750 \times 680 \times 600$  mm, spindle maximum speed 10,000 rpm, equipped with Fanuc controller is used in this experiment. The spindle and three servo motor power consumption are captured using Archmeter PA310 power meter. The power meters PA310 (hereinafter referred to as the power meter) were mounted on the spindle and three servo motor of R, S, T wire. RS232 cable is used to connect the motors and computer, then the power consumption data are captured via Servebox and stored in the computer.

## 4. Results and Discussions

From the experiment data, the analysis was conducted to investigate the relationship between cutting parameters and power consumption, and, thus, establish the regression equation using the polynomial regression method to achieve the purpose of optimizing machining efficiency and low power consumption.

### 4.1. Power Consumption Analysis without Cutting Load

According to the experiment scheme in the previous section, the experimental data can be divided into spindle and three servo motors. In the spindle experiments, six machining parameters were taken from 500 to 3000 rpm, with an increasing step every 500 rpm. The power consumption for each spindle speed parameter is shown in Figure 4. The experiment 1 to 3 and 4 to 6 were measured on different dates. It showed each power consumption value is very close to each other, around 0.00006 to 0.00009 kW. The power consumption does not significantly increase when the spindle speed increases. The reason is perhaps that the spindle is only in the idle state rather than the cutting state, so there is no cutting resistance, hence, increasing the spindle speed will not significantly affect the spindle power consumption.

For X-axis servo motor, 13 combinations of feedrate were carried out from 200 mm/min until 6000 mm/min. The ommencinh feedrate was 200 mm/min and after 500 mm/min, the feedrate was increased every 500 mm/min up to 6000 mm/min. Figure 5 shows the resulting power consumption for different feedrates. The experiment was repeated three times, and the results were reproducible. However, only a few of data showed a difference of about 0.005 kW, even at the feedrate of 4500 mm/min, the maximum difference was 0.01 kW. As seen in Figure 5, the power consumption increase rate is about 0.00125 kW per 500mm/min when the feedrate is 200 to1500 mm/min, but at feedrate of 2000 mm/min it reduces. When the feedrate is in the range of 2000 to 4000 mm/min, the power consumption trend is higher with an increase rate of about 0.0025 kW per 500 mm/min. However, at the feedrate of 4000 mm/min the power consumption is significantly reduced, but in the feedrate range of 4500 to 6000 mm/min the power consumption rises quickly with increasing

value of 0.005 kW per 500 mm/min. The increasing trend is higher than the first two feedrate sections. It can be concluded that increasing feedrate will increase power consumption, but when the feedrate is 2000 mm/min and 4500 mm/min the power consumption is lower. Moreover, it recommends to use these two feedrates to obtain lower power consumption.

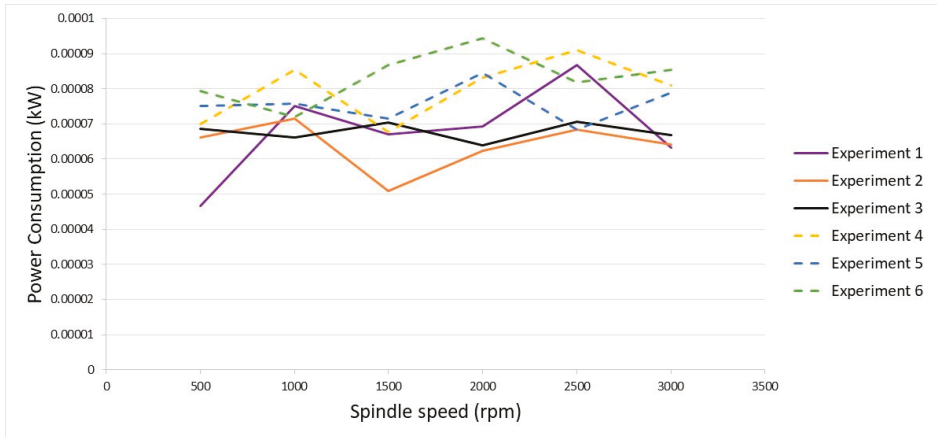


Figure 4. Relationship between spindle speed vs. power consumption w/o cutting load.

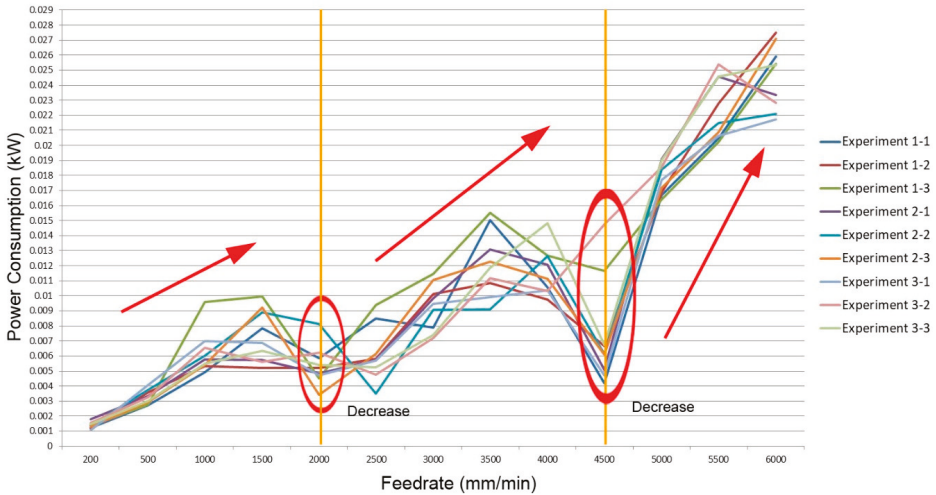


Figure 5. Relationship between X-axis motor feedrate vs. power consumption w/o cutting load.

Since the power consumption at the feedrate of 4500 mm/min is large, this data is excluded in the average calculation. Figure 6 shows the average calculation result, and each section regression equation can be obtained as follow:

$$y = -0.0014x^2 + 0.0078x - 0.0029 \tag{3}$$

$$y = -0.0003x^3 + 0.0021x^2 - 0.0031x + 0.0063 \tag{4}$$

$$y = -0.0016x^2 + 0.0128x - 0.0014 \tag{5}$$

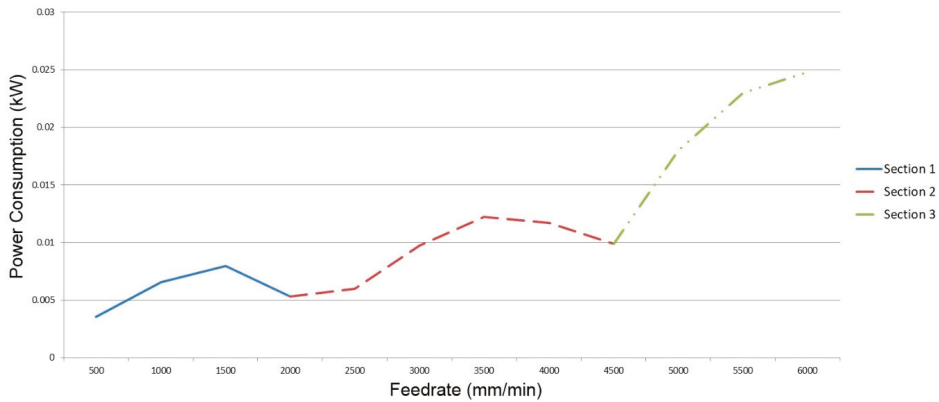


Figure 6. Curve of average X-axis motor feedrate vs. power consumption w/o cutting load.

For Y-axis servo motor, the parameter is similar to X-axis servo motor experiment, commencing at 200 mm/min and rising to 6000 mm/min with an increasing step of 500 mm/min. Figure 7 shows the relationship Y-axis motor upward direction with power consumption for different feedrates. It is seen the power consumption rate for feedrate range 200 to 1000 mm/min is about 0.0025 kW per 500 mm/min but smaller at a feedrate of 1500 mm/min. When the feedrate range is between 1500 and 2000 mm/min the power consumption increase rate is 0.005 kW per 500 mm/min, but a feedrate of 2500 to 3000 mm/min exhibits a decreasing trend. When the feedrate is between 3000 and 5500 mm/min the power consumption gradually increases from 0.008 kW up to 0.015 kW. From the results, the power consumption tends to increase with feedrate increase, but a feedrate of 1500 mm/min, 3000 mm/min, and 6000 mm/min showed lower power consumption. Therefore, these three feedrates can be considered to obtain lower power consumption without decrease machining efficiency.

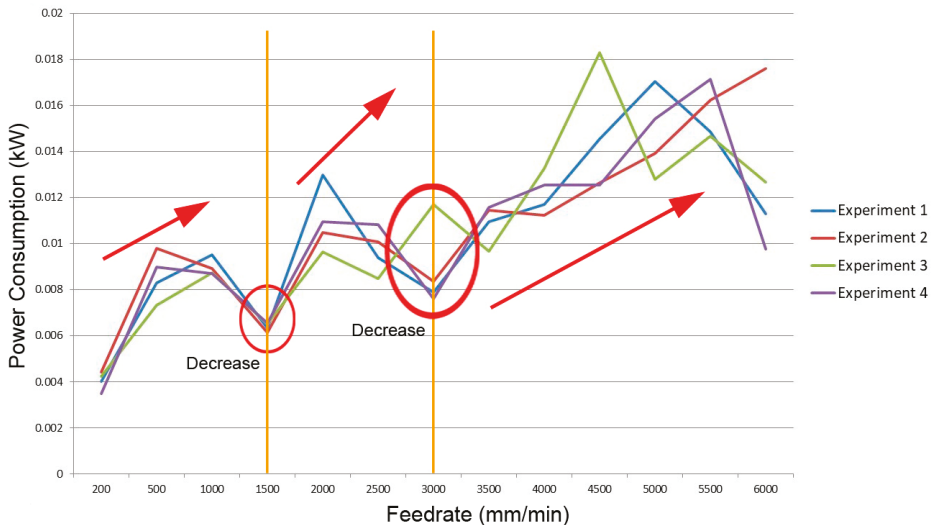


Figure 7. Relationship between Y-axis motor (up direction) feedrate vs. power consumption w/o cutting load.

Figure 8 shows the power consumption of Y-axis motor downward direction. It is seen the power consumption at a feedrate of 200 mm/min and 500 mm/min is greater than the feedrate of 1000

3500 mm/min. This perhaps is due to the influence of spindle weight. To maintain the low feedrate, the current must be increased to overcome the impact of spindle weight. When the feedrate is over 3500 mm/min, the power consumption increases with the increasing rate of 0.002 kW per 500 mm/min. As seen in Figure 8, the feedrate range of 1000 to 3500 mm/min provided lower power consumption without loss of machining efficiency, so this feedrate range can be considered for machining. Compare to up direction power consumption, the down direction power consumption is lower. This is because more power is needed to overcome spindle weight during toward up.

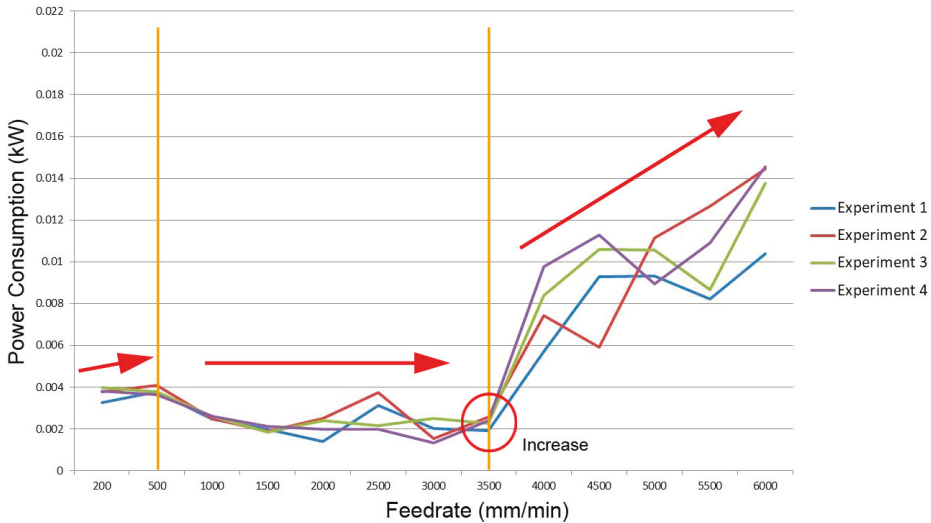


Figure 8. Relationship between Y-axis motor (down direction) feedrate vs. power consumption w/o cutting load.

Figure 9 shows the average calculation of Y-axis up direction power consumption, then the regression equation can be obtained as follows: Equation (6) for a feedrate of 500 to 1500 mm/min, Equation (7) for a feedrate of 1500 to 2000 mm/min, and Equation (8) for a feedrate of 2000 to 6000 mm/min.

$$y = -0.0015x^2 + 0.0048x + 0.0053 \tag{6}$$

$$y = 0.0046x + 0.0017 \tag{7}$$

$$y = -0.0001x^3 + 0.0016x^2 - 0.0058x + 0.0155 \tag{8}$$

Figure 10 shows the average calculation of Y-axis down direction power consumption, and the regression equation can be obtained as follows: Equation (9) for a feedrate of 500 to 3500 mm/min, and Equation (10) for a feedrate of 3500 to 6000 mm/min.

$$y = 0.0001x^2 - 0.0009 + 0.0044 \tag{9}$$

$$y = 0.0003x^3 - 0.0038x^2 + 0.0145 - 0.0087 \tag{10}$$

Regarding Z-axis servo motor, the experiment parameters are similar to X- and Y-axis, 13 parameter combination feedrates from 200 mm/min up to 6000 mm/min with increasing step 500 mm/min. Figure 11 shows the experiment results for different feedrates. As seen, the results are mostly reproducible. Only a few of the data have a 0.005 kW difference, such as at the feedrate of 5500 the difference is 0.012 kW. As seen in Figure 11, the power consumption rate at a feedrate of 200 to 1500 mm/min is 0.001 kW per 500 mm/min, but at a feedrate of 2000 mm/min the power consumption

is decreased. When the feedrate range is 2500 to 3500 mm/min, the power consumption rate increases about 0.002 kW per 500 mm/min. When the feedrate is 4000 mm/min and 4500 mm/min, the power consumption is reduced, but at a feedrate range of 4500 to 6000 mm/min the power consumption rate increase is about 0.003 kW per 500 mm/min. It can be seen that the overall power consumption tends to increase with the increasing feedrate, but at a feedrate of 2000 mm/min, 2500 mm/min, 4000 mm/min, and 4500 mm/min the power consumption is lower. These feedrates can be considered as the machining parameter to obtain lower power consumption without production efficiency loss.

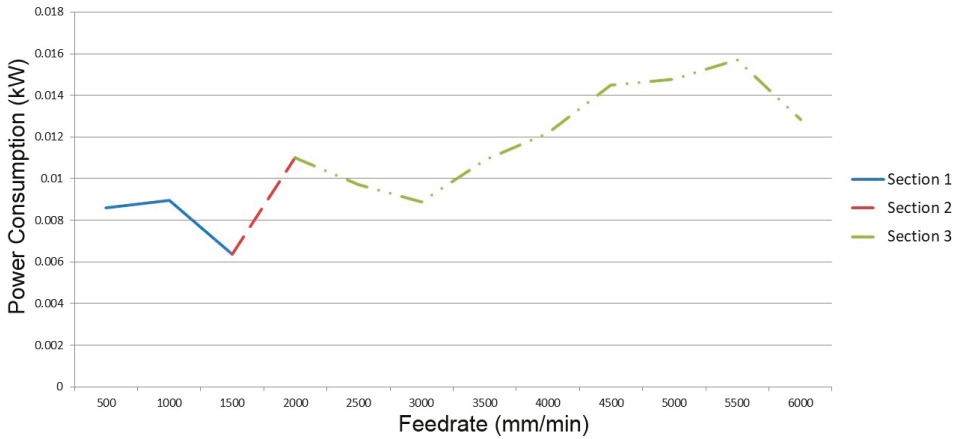


Figure 9. Curve of average Y-axis motor (up direction) feedrate vs. power consumption w/o cutting load.

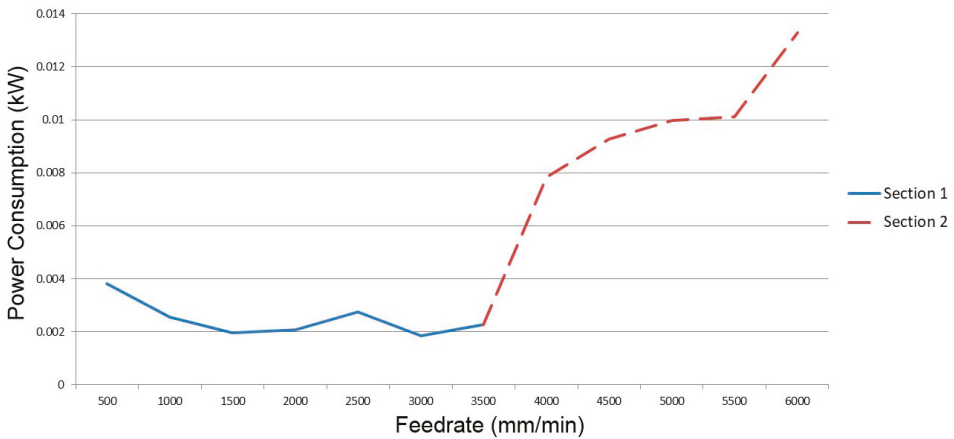


Figure 10. Curve of average Y-axis motor (down direction) feedrate vs. power consumption w/o cutting load.



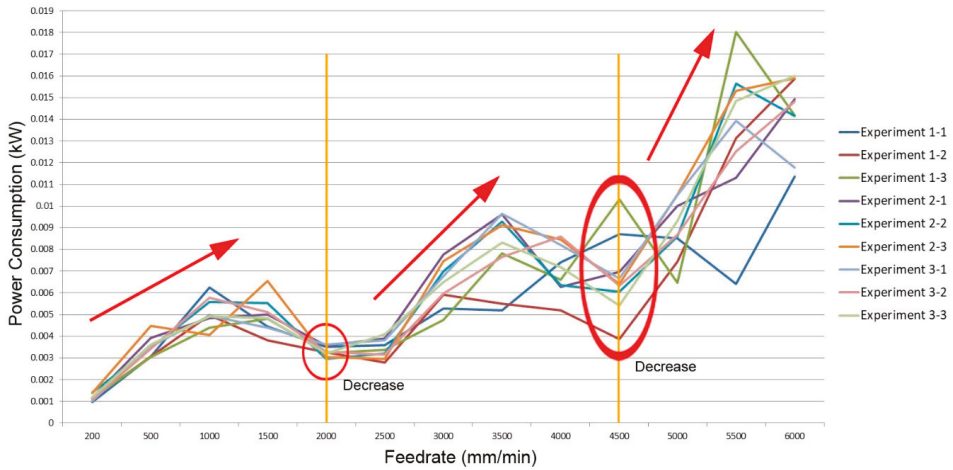


Figure 11. Relationship between Z-axis motor feedrate vs. power consumption w/o cutting load.

Since the power consumption at a feedrate of 5500 mm/min has a larger difference, it will be excluded from the average calculation analysis. Figure 12 shows the average calculation result, and the regression equation of each section can be obtained as follow: Equation (11) for a feedrate of 500 mm/min to 2000mm/min, Equation (12) for a feedrate of 2000 to 4500 mm/min, Equation (13) for a feedrate of 4500 to 6000 mm/min.

$$y = -0.0008x^2 + 0.0039x + 0.0004 \tag{11}$$

$$y = -0.0001x^3 + 0.0014x^2 - 0.0014x + 0.0034 \tag{12}$$

$$y = -0.0003x^2 + 0.0043x + 0.0024 \tag{13}$$

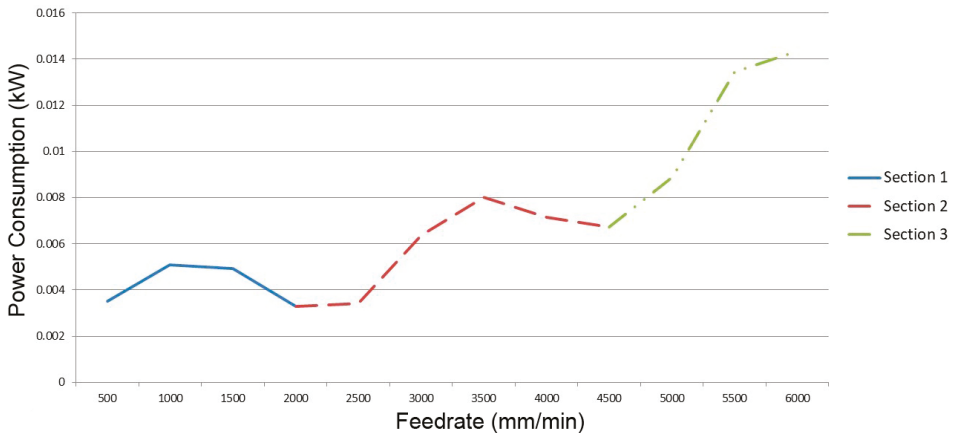


Figure 12. Curve of average Z-axis motor feedrate vs. power consumption w/o cutting load.

As seen in Figure 12, when the feedrate is 500 mm/min, the power consumption is 0.004 kW. If the feedrate is increased to 1500 mm/min, the power consumption is 0.005 kW. When the feedrate increases to 3000 mm/min, the power consumption is 0.006 kW. If in the beginning, the machining used a feedrate of 500 mm/min, the power consumption is 0.004 kW per second, so for an hour

the power consumption  $0.004 \text{ kW/sec} \times 3,600 \text{ sec/hr} = 14.4 \text{ kWh}$ . When the machining feedrate is  $1500 \text{ mm/min}$  for which the machining efficiency increases three times, the power consumption is  $0.005 \text{ kW/sec} \times 3,600 \text{ sec/hr} = 18 \text{ kWh}$ . When the machining feedrate is  $3000 \text{ mm/min}$  for which the machining efficiency increase six times, the power consumption is  $0.006 \text{ kW/sec} \times 3,600 \text{ sec/hr} = 21.6 \text{ kWh}$ . If we assume for a feedrate of  $500 \text{ mm/min}$ , the machining time needed is half an hour, the power consumption is  $0.5 \text{ hr} \times 14.4 \text{ kWh} = 7.2 \text{ kWh}$ . If we use the feedrate of  $1500 \text{ mm/min}$ , the machining time needed is 10 min, then the power consumption is  $3 \text{ kWh}$ . Hence, compared to a feedrate of  $500 \text{ mm/min}$ , the power saving is  $4.2 \text{ kWh}$  and also the machining time is three times faster.

#### 4.2. Power Consumption Analysis with Cutting Load

Since the experiment under cutting condition was repeated five times for each parameter, the average result will be used in the analysis. The result of experiment 1 is shown in Figure 13. The power consumption trend of X-axis motor for different chip loads is seen. Under the same spindle speed, a higher chip load ( $0.18 \text{ mm/t}$ ) did not consume higher power than the other two smaller chip loads. For eight different spindle speeds, there are four power consumptions of the  $0.18 \text{ mm/t}$  chip load smaller than the power consumption of the  $0.15 \text{ mm/t}$  chip load. Under the fixed spindle speed, the feedrate determines the chip load. Similarly, to maintain the same chip load, the feedrate must be proportional to the spindle speed. The chip load not only determines the spindle cutting load but also influences power consumption. The results do not exhibit significant affection for spindle power consumption due to the small difference of chip load.

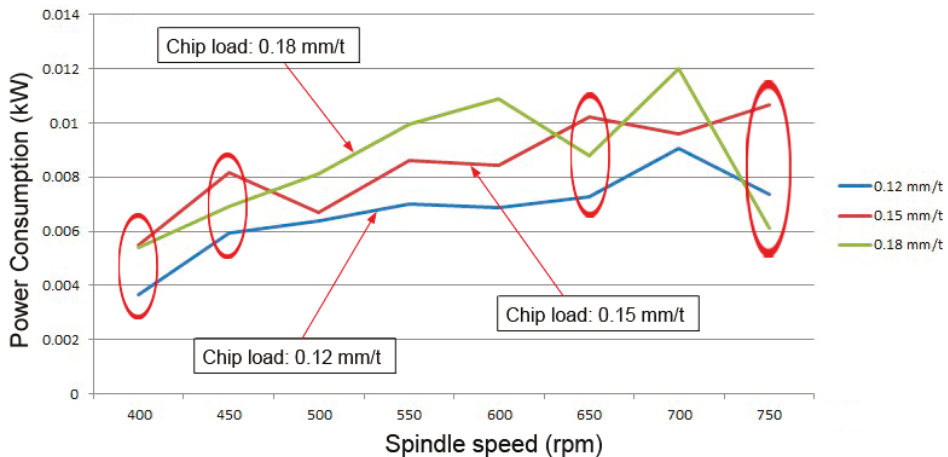


Figure 13. X-axis motor power consumption vs. Spindle speed for experiment 1.

In experiment 2, a larger interval of spindle speed was used to obtain a larger feedrate difference and observe the power consumption characteristic of the feedrate servo motor. The power consumption results of experiment 2 are shown in Figure 14. It is seen, even when the spindle speed was increased to  $1000$  to  $2250 \text{ rpm}$  with the maximum feedrate of  $270 \text{ mm/min}$ , the power consumption of chip load of  $0.18 \text{ mm/t}$  is smaller than chip load of  $0.12 \text{ mm/t}$  and  $0.15 \text{ mm/t}$ . This is perhaps due to the width of cut  $20 \text{ mm}$  which has reached  $80\%$  of cutter diameter that is classified as heavy cutting, thus, generating large cutting resistance. Hence, the motor power consumption is increased. In addition, when the spindle speed was  $2250 \text{ rpm}$  and the feedrate  $1620 \text{ mm/min}$ , a spark was appearing during the cutting process, so it is not appropriate to increase the feedrate.

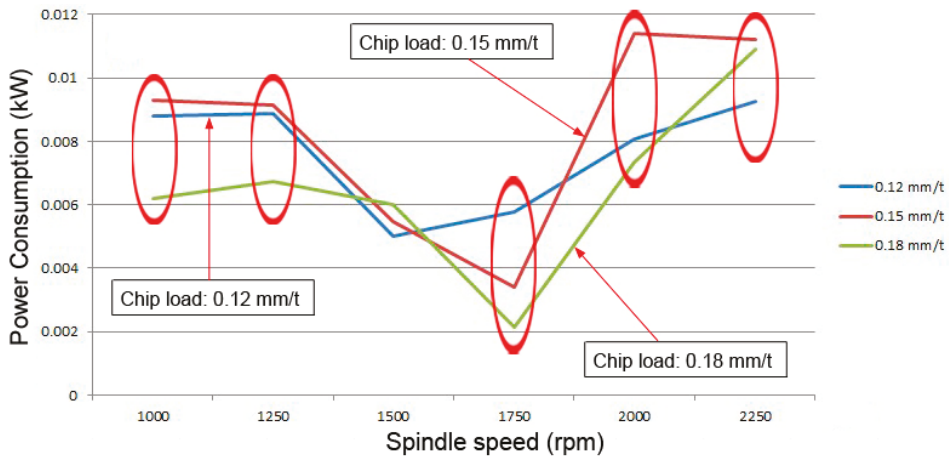


Figure 14. X-axis motor power consumption vs. Spindle speed range of 1000 to 2250 rpm.

### 4.3. Verification

#### 4.3.1. Power Consumption without Cutting Load

To verify the effectiveness of the proposed system, the electricity unit price and machining cost (hourly-rate) were set as 2.5 NTD/degree and 25 NTD/hour, and a 3-D straight line machining trajectory from (X2000, Y1500, Z1500) to (Z-1500, Y-1500, X-2000) was planned. The cutting path with the same feedrate was repeated twice. The original feedrate for machining was 1000 mm/min, and expected to improve for 3-times faster. Based on those parameters, the system calculated and recommend using a feedrate of 3000 mm/min as shown in Figure 15. Because the power consumption of three servo motors was quite small for a feedrate range of 1000 to 3000 mm/min (as shown in Figure 6, Figure 9, Figure 10, Figure 12), the electricity cost (about only 0.005 kW) is much less than the machining cost. Therefore, short machining time will be more dominant for the overall cost. Meanwhile, to remain the same chip load, the rotation speed of spindle was also increased from 300 rpm to 9000 rpm.

Figure 15. The proposed parameters information for w/o cutting load.

Power consumption amounts of the three axes servo motors with a feedrate of 1000 mm/min are shown in Figures 16–18. In the figures, those with very low power consumption (almost zero) represent that machine was moving at that time. The total machining time was about 20 min, the power consumption of X-, Y-, and Z-axis motor were 2.396, 3.619, and 2.178 kW, respectively. The total power consumption was 8.193 kW, and the power consumption cost was about 20.4825 NTD. The machining cost for 20-minutes is about 83.333 NTD. By summing up the power consumption cost and machining cost, the total cost for this machining process was about 103.833 NTD.

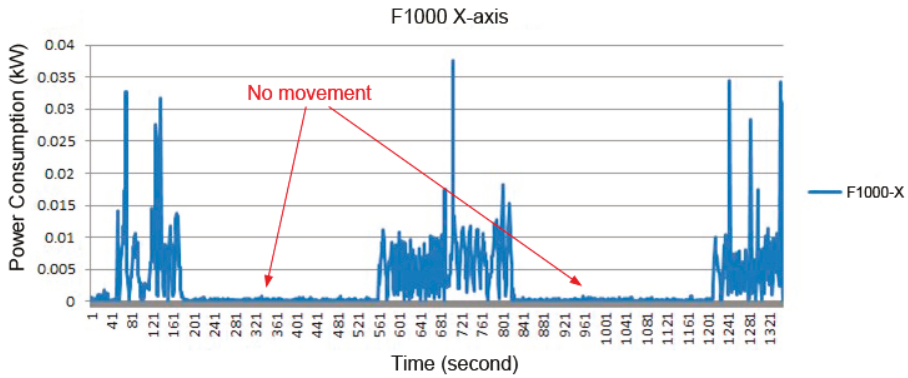


Figure 16. Power consumption data of X-axis with F1000.

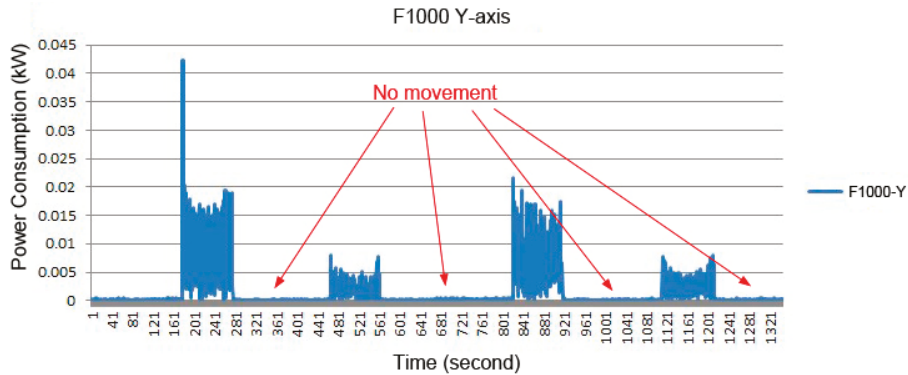


Figure 17. Power consumption data of Y-axis with F1000.

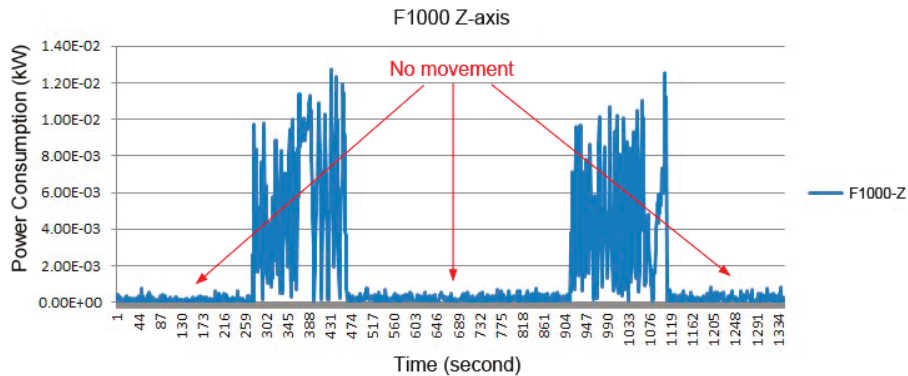


Figure 18. Power consumption data of Z-axis with F1000.

For a feedrate of 3000 mm/min, the power consumption amounts of the three axes servo motors are shown in Figures 19–21. The total machining time was about 6.667 min, and the power consumption of X-, Y-, and Z-axis were 0.7023, 1.5982, and 1.0192 kW, respectively. The total power consumption was about 3.32 kW, and the cost of the power consumption was 8.3 NTD. For 6.67 min, the machining cost was about 27.778 NTD. By summing up the power consumption cost and machining cost, the total cost for this machining process was about 36.078 NTD.

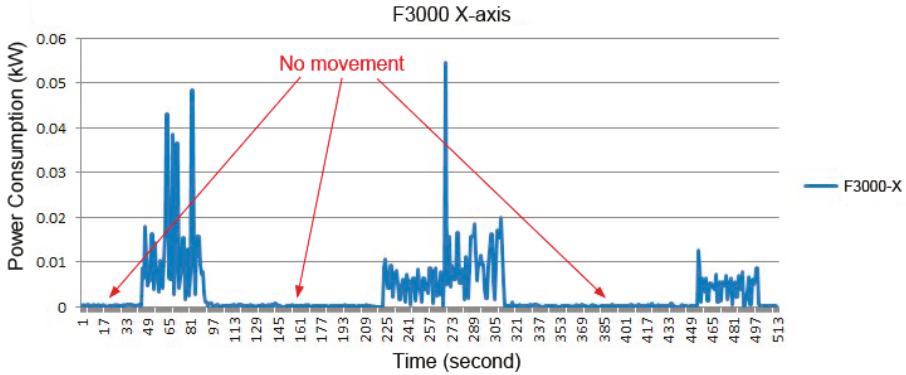


Figure 19. Power consumption data of X-axis with F3000.

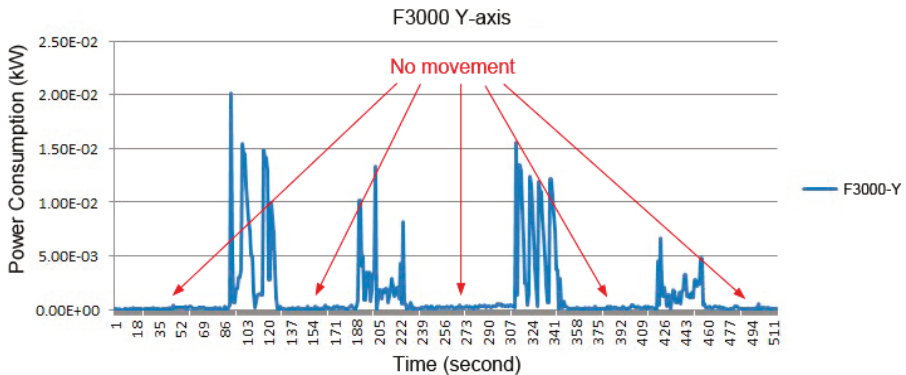


Figure 20. Power consumption data of Y-axis with F3000.

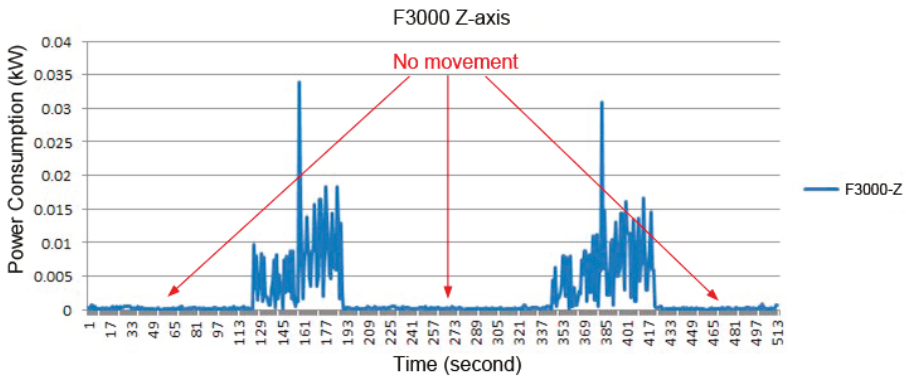


Figure 21. Power consumption data of Z-axis with F3000.

The comparison of power consumption between actual measured data and the estimated calculation is shown in Table 3. It can be seen that the actual power cost is 2 to 3 NTD (or equivalent to 1 to 2 kWh) higher than the estimated calculation. Due to the error between the generated characteristic curves and the actual values of the three axes servo motors, it is quite acceptable. It was also noted that the generated characteristic curves had good agreement with the actual measured values (difference between the actual values and the estimated values were only about 0.1 kWh).

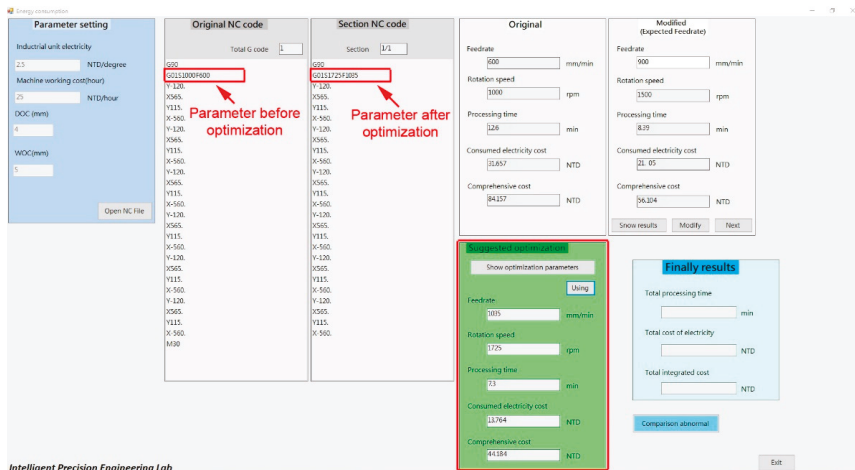
**Table 3.** Actual and estimated comparison for power consumption without cutting load.

Processing Source \ Information	Power Costs (NTD)	Machining Costs (NTD)	Total Costs (NTD)
Calculated by program (F1000)	17.85	83.333	101.183
Calculated with measured data (F1000)	20.482	83.333	103.833
Calculated by program (F3000)	8.3	27.778	36.078
Calculated with measured data (F3000)	8.3	27.778	36.078

As it was noted, when the feedrate was adjusted from 1000 mm/min to 3000 mm/min, the total power consumption cost was reduced from 21 NTD to 8.5 NTD, and the machining cost decreased from 83.3 NTD to 27.8 NTD; the total cost was lowered to 36.3 NTD from 104.4 NTD as well. It proved that when feedrate increased, the motor power consumption may also increase. However, if the machining time can be significantly shortened, then the total power consumption cost will still be significantly reduced.

### 4.3.2. Power Consumption with Cutting Load

The condition used for verification was assumed as electricity fee is 2.5 NTD/degree and the machine cost is 250 NTD/hour. For machining, a 4-mm depth of cut, and 5-mm width of cut, and 25-mm diameter end mill were used for machining iron. The original spindle speed and feedrate were 1000 rpm and 600 mm/min, respectively, the expected improvement of feedrate was at least 1.5 times faster. A straight line machining trajectory from (X565, Y-120.), (X-560, Y115.) was planned. The cutting path was repeated six times, and the total cutting length was 7560 mm. The system analyzed for the optimal feedrate according to motor power consumption characteristics curves. A feedrate of 1035 mm/min was finally chosen as the optimal value for the process (as shown in Figure 22).



**Figure 22.** The proposed parameters information for w/cutting load.

When a feedrate of 600 mm/min was used, the total machining time was about 12.6 min, and the total power consumption was about 12.663 kW. After converting, the power consumption cost was about 31.657 NTD, the machining cost was about 52.5 NTD, and total cost became 84.157 NTD. When the recommended feedrate of 1035 mm/min was used, the total machining time reduced to 7.3 min, the total power consumption was about 5.5056 kW. After converting, the power consumption cost was about 13.764 NTD, and the machining cost is 30.42 NTD. By summing up power consumption cost and machining cost, the total cost for this process was about 44.184 NTD. Figure 23 shows the power consumption data for using 600-mm/min feedrate. The X-axis actual measured power consumption was 8.485 kW (blue line), and the Y-axis actual measured power consumption was 4.178 kW (red line).

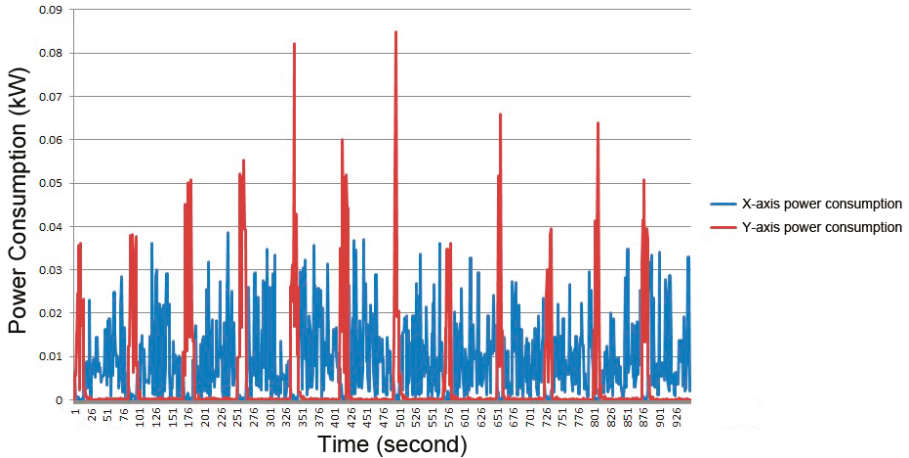


Figure 23. Actual power consumption with cutting load using 600-mm/min feedrate.

When a 1035-mm/min feedrate was used, the X-axis actual power consumption (measured) was 4.543 kW (blue line in Figure 24), and the Y-axis actual power consumption (measured) was 0.9626 kW (red line in Figure 24).

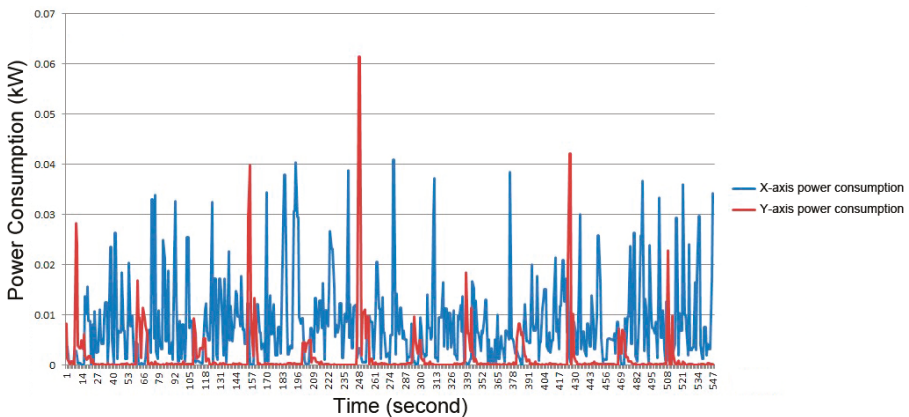


Figure 24. Actual Power consumption with cutting load using 1035-mm/min feedrate.

The comparison of actual measurement and estimated calculation of power consumption is shown in Table 4. It was noted that the estimated power consumption in X-axis and Y-axis were smaller than the actual measured data. For a feedrate of 600 mm/min, the actual total power consumption

for X-axis (3.325 kW) was higher than the estimate, also the Y-axis (1.226 kW) was higher than the estimated calculation value. For a 1035-mm/min feedrate, the actual total power consumption for X-axis (3.047 kW) was higher than the estimated value. In addition, the actual total power consumption of Y-axis (0.5809 kW) was higher than the estimated calculation value. The actual power consumption was higher than the estimation calculated by characteristics curves which may be because of the 24 starts and stops of the machine for each machining experiment due to repeating the machining process for six times. More energy was consumed for the machine starts and stops. In general, the recommended optimal parameters by the system could improve the machining efficiency and reduce the total power consumption cost. As it can be seen in Table 4, when the feedrate was adjusted from 600 mm/min to 1035 mm/min, the power consumption cost was reduced from 31.657 NTD to 13.764 NTD, and the total machining time also dropped from 12.6 min to 7.3 min. It proves that when the feedrate increase, even if the power consumption of each motor not always decreased, the save of machining time could reduce the total power consumption cost for synchronously improving the machining efficiency and power consumption.

**Table 4.** Actual and estimated comparison for power consumption with cutting load.

Information	X-axis Total Power Consumption (kW)	Y-axis Total Power Consumption (kW)	Power Costs (NTD)
Processing Source			
Calculated by program (F600)	5.16	2.952	20.28
Calculated with measured data (F600)	8.485	4.178	31.657
Calculated by program (F1035)	1.496	0.382	4.691
Calculated with measured data (F1035)	4.543	0.963	13.764

## 5. Conclusions

In this paper, an accuracy-efficiency-power consumption hybrid optimization method was established. The relationship between motor speed and power consumption with and without cutting load was created. The synchronous optimization algorithm was developed based on these relationship models. This study used motor-power consumption characteristic to develop optimization models and algorithm. According to the models and algorithm, an accuracy-efficiency-power consumption synchronous optimization system with a procedure-guided HMI was developed. Results of verification experiments showed that the system can help to select the optimal parameters for better machining efficiency and power consumption without losing machining accuracy.

**Author Contributions:** Conceptualization, S.-M.W.; methodology, S.-M.W. and C.-Y.L.; software, C.-Y.L.; validation, C.-Y.L. and C.-C.Y.; formal analysis, C.-Y.L. and H.G.; investigation, C.-Y.L. and S.-M.W.; resources, C.-Y.L. and C.-C.Y.; data curation, C.-Y.L.; writing, H.G. and S.-M.W.

**Acknowledgments:** This study was supported by Ministry of Science and Technology Taiwan under the grant number MOST 107-2221-E-033-066 and MOST 108-2634-F-033-001.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

- Gontarz, A.; Züst, S.; Weiss, L.; Wegener, K. *Energetic Machine Tool Modeling Approach for Energy Consumption Prediction*; Institute of Machine Tools and Manufacturing (IWF), Swiss Federal Institute of Technology: Zürich, Switzerland, 2012.
- Abdoli, S.; Semere, D.T. Investigation on Machine Tools Energy Consumptions. *Int. J. Mech. Aerosp. Ind. Mechatron. Eng.* **2014**, *8*, 1091–1098.
- Diaz, N.; Choi, S.; Helu, M.; Chen, Y.; Jayanathan, S.; Yasui, Y.; Kong, D.; Pavanaskar, S.; Dornfeld, D. Machine Tool Design and Operation Strategies for Green Manufacturing. In Proceedings of the 4th CIRP International Conference on High Performance Cutting, Gifu, Japan, 24–26 October 2010.



4. Mori, M.; Fujishima, M.; Inamasu, Y.; Oda, Y. A study on energy efficiency improvement for machine tools. *CIRP Ann. Manuf. Technol.* **2011**, *60*, 145–148. [[CrossRef](#)]
5. Dragannescu, F.; Gheorghe, M.; Doicin, C.V. Models of Machine Tool Efficiency and Specific Consumed Energy. *J. Mater. Process. Technol.* **2003**, *141*, 9–15. [[CrossRef](#)]
6. Kara, S.; Li, W. Unit process energy consumption models for material removal processes. *CIRP Ann. Manuf. Technol.* **2011**, *60*, 37–40. [[CrossRef](#)]
7. Dimla, S.D.E. Sensor signals for tool-wear monitoring in metal cutting operations—A review of methods. *Int. J. Mach. Tools Manuf.* **2000**, *40*, 1073–1098. [[CrossRef](#)]
8. Usui, E.; Shirakashi, T. Analytical prediction of cutting tool wear. *Wear* **1984**, *100*, 129–151. [[CrossRef](#)]
9. Prickett, P.W.; Johns, C. An overview of approaches to end milling tool monitoring. *Int. J. Mach. Tools Manuf.* **1999**, *39*, 105–122. [[CrossRef](#)]
10. Jun, C.H. The Study of Endmill Milling Error and Online Compensation. Master's Thesis, National Tsinghua University, Hsinchu, Taiwan, 1988.
11. Dan, L.; Mathew, J. Tool wear and failure monitoring techniques for turning—A review. *Int. J. Mach. Tools Manuf.* **1990**, *30*, 579–598. [[CrossRef](#)]
12. Choi, G.S.; Wang, Z.X.; Dornfield, D.A.; Tsujino, K. Development of an intelligent on-line tool wear monitoring system for turning operations. In Proceedings of the USA–Japan Symposium on Flexible Automation, A Pacific Rim Conference ISCIE, Kyoto, Japan, 9–13 July 1990.



© 2019 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).

Article

# Compensation of Frequency Drift in Frequency-Sweep Polarization-Modulation Ranging System

Shuyuan Gao <sup>1</sup>, Rongyi Ji <sup>2</sup>, Yao Li <sup>2</sup>, Chun Liu <sup>3</sup>, Junkai Shi <sup>2</sup>, Yingling Pan <sup>2</sup>  
and Weihu Zhou <sup>2,4,\*</sup>

- <sup>1</sup> School of Instrument Science and Optoelectronic Engineering, Hefei University of Technology, Hefei 230009, China; gaoshuyuan127@163.com
  - <sup>2</sup> Institute of Microelectronics of The Chinese Academy of Sciences, Beijing 100029, China; jirongyi@ime.ac.cn (R.J.); hustliyao@163.com (Y.L.); shijunkai@ime.ac.cn (J.S.); panyingling@ime.ac.cn (Y.P.)
  - <sup>3</sup> School of Optics and Photonics, Beijing University of Technology, Beijing, 100081, China; lc17273747@163.com
  - <sup>4</sup> University of Chinese Academy of Sciences, Beijing 100049, China
- \* Correspondence: zhouweihu@aoe.ac.cn

Received: 29 January 2019; Accepted: 20 March 2019; Published: 25 March 2019



**Featured Application:** Absolute distance measurement.

**Abstract:** In frequency-sweep polarization-modulation ranging, distance is determined by the frequency of modulated waves and the corresponding wavelength multiple when emitted and returned waves are in phase. However, measurement of the frequency and the wavelength multiple is affected by thermally induced phase delay of the polarized wave. In this article we systematically discuss the principle of the ranging method and analyze the influences of thermally induced phase delay. New approaches to measurement are proposed to eliminate the impact on frequency and the wavelength multiple. Theoretical analysis and experimental results proved the efficiency and applicability of the methods.

**Keywords:** absolute distance measurement; frequency-sweep polarization-modulation ranging; frequency drift; error compensation

## 1. Introduction

Absolute distance measurement (ADM) technologies are of significant application in scientific and engineering fields such as large equipment manufacturing, spacecraft fabrication, and general-purpose coordinate measurements [1–3]. Various methods of ADM have been developed over the past decades, such as phase shift-based ranging [4–6], frequency scanning interferometry-based ranging [7–10], multi-wavelength interference-based ranging [11–14], and femtosecond optical comb-based ranging [10,15–17]. In recent years, the latter three technologies have achieved great accuracy in laboratory, but the systems are complex and their performance needs to be enhanced in an industrial setting. The phase shift ranging method is a common technology in industry, in which the phase differences between continuous emitted and reference waves are compared. However, in phase shift ranging systems, the performance of phase discrimination is influenced by parasitic parameters in the circuit, and as a result ranging precision is limited.

The frequency-sweep polarization-modulation (FSPM) method is another phase ranging method in which modulation frequency is scanned and searched when emitted and returned waves are in phase, thus avoiding the environmental disturbance as well as shortage of phase discrimination [18–20].

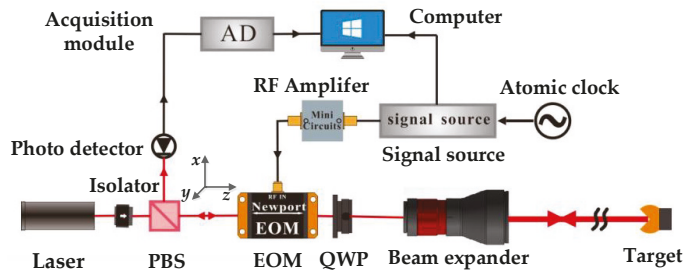
It was proposed by Kern & Co. Ltd. and first applied in the commercial distance meter ME5000 [21–23]. Due to its simple structure and high resolution, the FSPM ranging technique has the potential for small size and high performance. However, in the FSPM ranging system, the frequency is vulnerable to phase delay variation due to the sensitivity of crystals (e.g., LiNbO<sub>3</sub>) in the electro-optic modulator (EOM) to temperature disturbance. The heat dissipation of electronic components and laser irradiation will accelerate the temperature variation [24], thus leading to distance measurement errors.

In order to investigate this problem, a theoretical model was developed for the FSPM ranging systems. Based on the model, a correction approach for diminishing the thermally-induced errors in the system was proposed. A reciprocating sweeping method and multi-frequency interval method were respectively proposed to improve the measurement accuracy of frequency and the corresponding wavelength multiple. The validity of the method was proved by theoretical analysis and experiments. The article is organized as follows: In Section 2, the principle of the FSPM-based system is introduced, measurement errors due to phase delay are discussed, and the compensation methods are proposed. In Section 3, experiments are conducted to prove the effectiveness of the methods. Finally, a brief conclusion is given in Section 4.

## 2. Principle and Method

### 2.1. The Principle and Model

The frequency-sweep polarization-modulation ranging system consists of a continuous wave laser, an isolator, a polarizing beam splitter (PBS), an EOM, a quarter wave plate (QWP), a beam expander (EXP) and the target, as shown in Figure 1.



**Figure 1.** Schematic diagram of the polarization-modulation ranging system. EOM: electro-optic modulator; PBS: polarizing beam splitter; QWP: quarter wave plate.

To illustrate the working principle of the FSPM ranging method, a Cartesian coordinate system is established. The z-axis represents the direction of wave propagation, while the x-axis is the polarization direction of linearly polarized light generated by the laser. In the system, continuous polarized light emitted by the laser travels through an isolator and a PBS tipped 45 degrees from the x-axis, so equal amplitude of polarization components in the x-axis and y-axis is achieved when light travels into the EOM. The EOM generates phase retardation between the two components in a sinusoidal manner with linearly changed frequency. The QWP has a fast axis oriented at 45 degrees to the x-axis, so the light travels to the target and returns by the same path, to then be demodulated by the EOM. The phase difference between the polarization components when light transmits through the EOM successively can be respectively described as:

$$\begin{aligned} \varphi_1 &= k \sin(\omega t), \\ \varphi_2 &= k \sin(\omega t + \varphi) + \delta_T, \end{aligned} \tag{1}$$

where  $\omega$  is the modulation frequency,  $k$  represents the modulation depth,  $\delta_T$  represents the additional phase delay between the polarization components,  $\varphi$  is the phase shift of the modulation signal during the round-trip time, and  $\varphi = 4\pi f(t)nl/c$ , where  $f(t)$  is the modulation frequency,  $c$  represents the speed

of light in vacuum,  $n$  is the refractive index of the air, and  $l$  is the distance to be measured. Note that in the equation,  $0 < k \leq 1$  for linear modulation and  $\delta_T$  is small, typically less than  $\pi/2$ . With the amplitude of  $x$  and  $y$  components normalized, the intensity of light received by the detector can be written as:

$$I = 1 - \cos(\varphi_1 - \varphi_2) = 1 - \cos\left(2k \cos \frac{2\omega t + \varphi}{2} \sin \frac{\varphi}{2}\right) \cdot \cos \delta_T + \sin\left(2k \cos \frac{2\omega t + \varphi}{2} \sin \frac{\varphi}{2}\right) \cdot \cos \delta_T. \quad (2)$$

After simplification and filtering the high-frequency components, the equation can be rewritten as:

$$I = 1 - \left(1 - \frac{k^2}{2} + \frac{k^2}{2} \cos \varphi\right) \cdot \cos \delta_T. \quad (3)$$

Obviously, if  $\cos \delta_T$  is constant, there exists a cosine relationship between intensity and modulation frequency. Minimums of the equations By obtaining the corresponding frequencies  $f_x, f_{x+1}$  ( $k \in \mathbb{N}$ ) at adjacent minima, the distance can be determined by:

$$l = N \cdot \frac{c}{2nf_x}, \text{ in which } N = \left[ \frac{f_x}{f_{x+1} - f_x} \right], \quad (4)$$

where,  $N$  is the wavelength multiple,  $[ ]$  is the rounding operation,  $f_x / (f_{x+1} - f_x)$  is the initial value of the wavelength multiple, which comprises an integer and a fraction part, so the rounding operation is necessary. For convenience in the manuscript, The frequencies  $f_x$  and  $f_{x+1}$  will be defined as "in phase" frequency. Value of difference of adjacent in-phase frequencies i.e.,  $f_{x+1} - f_x$  will be defined as frequency interval  $f_b$ , which is ideally a constant when distance is invariant.

## 2.2. Influence of Thermally Induced Phase Delay and Traditional Correction Method

In a transverse modulator configuration, the effects of temperature variation on modulation characteristics are mainly as follows: first, due to the thermo-optic effect of the crystal, the refractive index of the crystal changes with temperature. More concretely, the change in the refractive index along the direction of light transmission causes the optical path variation of light, and the change in refractive index perpendicular to the direction of light transmission results in an additional phase delay of the polarization [25]. Second, due to the elasto-optical effect of the crystal, the stress caused by the thermal expansion of the crystal also changes the refractive index and the longitudinal length of the crystal. As a result, the thermally induced phase delay can be expressed as follows:

$$\delta_T = \frac{2\pi}{\lambda} \Delta n_t l + \frac{2\pi}{\lambda} n_t \Delta l, \quad (5)$$

where  $\lambda$  represents wavelength of light,  $n_t$  is the initial value of natural birefringence in the direction perpendicular to light transmission, and  $\Delta n_t$  is the thermally induced birefringence variation.  $l$  represents the initial value of the length of the crystal, and  $\Delta l$  is the thermally induced variation of the crystal length.

Several techniques have been applied to compensate for the thermally induced phase delay in the applications of EOM. An example of this is connecting a compensate crystal that has the same size and performance with the modulate crystal, with the optical axes of the crystals perpendicular to each other so that the phase delay in the two identical crystals are complementary [26]. Another way is to make the two crystals in series, with one half-wave plate placed in the middle of them to make the polarization plane of the polarized light rotate by  $\lambda/2$  [24]. In these methods, crystals with exactly the same parameters and their perfect assembly are the key to compensation, but that is difficult to achieve. In the FSPM ranging system, a quarter-wave plate is placed between the crystal and the target so that when light passes through the quarter-wave plate during emitting and returning, the plane of polarization can be rotated by  $\lambda/2$  to achieve the complementation [27,28]. In this structure,

the adjustment is simpler and the compensation is better. However, in the FSPM ranging system, a mismatch of polarizing devices and reflection of polarized light are inevitable and lead to the additional rotation of the polarization plane of light. Even a very small mismatch of polarization devices or any reflection in the system will result in the rotation of the polarization plane, and produce residual thermally induced phase delay.

In our experiment, the residual thermally induced phase delay was tested first. Since thermally induced phase delay varies slowly, it can be measured through a polarization analysis system (Thorlabs, PAX5710). At the same time, the temperature in the modulator chamber was tested with a high-precision temperature controller (Wavelength, Tc\_lab), with results shown in Figure 2. As is shown, with the increase of temperature, the phase delay slowly increased, accompanied by periodic oscillation, and when temperature growth slowed down, the oscillation speed also decreased.

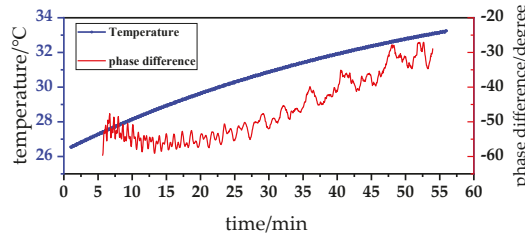


Figure 2. Thermally induced phase delay of round-trip polarization light in the system.

### 2.3. Reciprocating Sweeping-Based Measurement of “in Phase” Frequency

The change of additional phase delay accelerates or decelerates the polarization variation when the wave is modulated across the “in phase” frequency, thus destroying the symmetry of the intensity curve, leading to the shift of minima. The shift can be manifested in variation of the curve slope. Through the derivation of Equation (3), the slope of the intensity curve at the ideal minimum  $\varphi_0$  can be written as:

$$I'(\varphi_0) = \lim_{\Delta\varphi \rightarrow 0} \frac{I(\varphi_0 + \Delta\varphi) - I(\varphi_0 - \Delta\varphi)}{2\Delta\varphi} = \lim_{\Delta\varphi \rightarrow 0} \frac{-\left(1 - \frac{k^2}{2} + \frac{k^2}{2} \cos \Delta\varphi\right) (\sin \delta_0 \sin \Delta\delta)}{\Delta\varphi}, \quad (6)$$

where  $\delta_0$  is the thermally induced phase delay when  $\varphi = \varphi_0$ ,  $\Delta\varphi$  is the increment on point of  $\varphi_0$ ,  $\Delta\delta$  is the increment on point of  $\delta_0$ . According to the above equation,  $I'(\varphi_0)$  is a result of interaction between modulation depth, frequency scanning, and phase delay variation. Large modulation depths and frequency sweep speed are beneficial for eliminating the effect of thermally induced phase delay. However, modulation depth and sweep speed are often limited owing to equipment limitations, and a larger modulation depth requires a larger microwave voltage, thus aggravating the heating problem of the modulation circuit. Interestingly, we found that Equation (6) is an odd function about  $\varphi$ , so the orientation of the frequency shift can be controlled through frequency sweep directions. When modulation frequency is swept forward (i.e., from the lower frequency to the higher frequency),  $\Delta\varphi(t) > 0$ , otherwise,  $\Delta\varphi(t) < 0$ . Therefore, during the forward and backward sweep, the minimum drifts in opposite directions. Through averaging of such reciprocating sweeping, the drift errors can be diminished and frequency measurement stability can be improved. The method is further analyzed below.

Suppose the ideal frequency under test is  $f_0$ .  $f_1$  and  $f_2$  are frequencies very close to  $f_0$  and satisfy the equations  $f_1 = f_0 - \delta f$  and  $f_2 = f_0 + \delta f$ . If the system is not affected by thermally induced phase delay, obviously we get the equation  $I(f_1) = I(f_2)$ , then the exact value of  $f_0$  can be worked out through:  $f_0 = (f_1 + f_2)/2$ . However, due to thermally induced phase delay, during the forward sweeping,  $f_1$  and

$f_2$  respectively moves  $\Delta f_1^1$  and  $\Delta f_2^1$  towards lower frequency, then in the backward sweeping,  $f_1$  and  $f_2$  respectively moves  $\Delta f_1^2$  and  $\Delta f_2^2$  towards the opposite side, as shown in Figure 3.

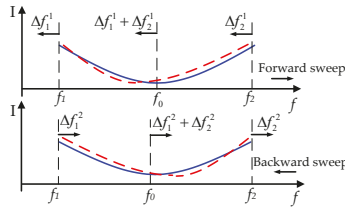


Figure 3. Influence of frequency drift during reciprocating sweeping.

Taking random measurement errors  $v$  into account, the extremum of the measurements are respectively expressed as:

$$\begin{aligned} f_0^1 &= \frac{(f_1 - \Delta f_1^1 \pm |v_1^1|) + (f_2 - \Delta f_2^1 \pm |v_2^1|)}{2}, \\ f_0^2 &= \frac{(f_1 + \Delta f_1^2 \pm |v_1^2|) + (f_2 + \Delta f_2^2 \pm |v_2^2|)}{2}. \end{aligned} \tag{7}$$

Through  $n$  times of such sweeps, the frequency can be obtained by:

$$f_0 = \frac{2n(f_1 + f_2) + \sum_{i=1}^n (\Delta f_1^{2n} - \Delta f_1^{2n-1}) + \sum_{i=1}^n (\Delta f_2^{2n} - \Delta f_2^{2n-1}) \pm \sum_{i=1}^{2n} |v_1^i| \pm \sum_{i=1}^{2n} |v_2^i|}{4n}. \tag{8}$$

During the time of reciprocating sweeping, variation of  $\delta$  is small so items of  $\sum_{i=1}^n (\Delta f_1^{2n} - \Delta f_1^{2n-1})$  and  $\sum_{i=1}^n (\Delta f_2^{2n} - \Delta f_2^{2n-1})$  can be diminished to some extent, and the random measurement error of frequency can be compensated by averaging.

#### 2.4. Multi-Frequency Interval-Based Measurement of Frequency Interval

Although frequency error can be partially eliminated through the reciprocating sweeping method, measurement of  $N$  requires a higher accuracy of frequency. The measurement error of  $N$  can be expressed as:

$$\Delta N = \left[ \frac{\Delta f_x}{f_x} - \frac{\Delta f_b \cdot f_x}{f_b^2} \right] = \left[ -\frac{\Delta f_b \cdot f_x}{f_b^2} \right]. \tag{9}$$

The above equation indicates that the absolute error of  $N$  is determined by the relative error of  $f_b$ . However, in the FSPM method, the impact of relative error of  $f_b$  is often  $2N$  times larger than the relative error of  $f_x$ , which means the accuracy of  $f_x$  may not meet the accuracy requirement of  $f_b$ . If  $f_x$  is around 2.5 GHz,  $l$  is about 100 m, a frequency measurement error of 2500 Hz only contributes a ranging error of 100  $\mu\text{m}$ , but the resulting relative error of  $f_b$  contributes a ranging error of 0.36 m. To ensure the accuracy of  $N$ ,  $\Delta f_b$  should be less than 90 Hz, which cannot be satisfied by the reciprocating sweeping method.

In order to reduce the measurement error of  $f_b$ , we used the characteristic of equal spacing of the “in phase” frequencies in our multi-frequency interval-based approach. For example,  $f_b$  can be expressed as follows:

$$f_b = \frac{(f_{x+u} - f_x)}{u}, \tag{10}$$

where  $f_x$  is an arbitrary “in phase” frequency in sweeping band,  $f_{x+u}$  is the  $u$ th “in phase” frequency on its upper side. Therefore,  $N$  can be rewritten as:

$$N = \left[ \frac{f_x}{f_b} \right] = \left[ \frac{u \cdot f_x}{f_{x+u} - f_x} \right]. \tag{11}$$

Hence, taking the frequency drift and random measurement error into account, the multi-frequency interval can be expressed as:

$$f_{x+u} - f_x = u \cdot f_b \pm \Delta f_u \pm (|v_{x+u}| + |v_x|), \tag{12}$$

where  $\Delta f_u$  represents the frequency drift. The measurement error of the wavelength multiple can be expressed as:

$$\Delta N = \left[ \frac{\Delta f_x}{f_b} - \frac{f_x \cdot \Delta(f_{x+u} - f_x)}{u \cdot f_b^2} \right] \approx \left[ -\frac{f_x \cdot \Delta(f_{x+u} - f_x)}{u \cdot f_b^2} \right] = \left[ -\frac{f_x \Delta(\pm \Delta f_u \pm (|v_{x+u}| + |v_x|))}{u \cdot f_b^2} \right]. \tag{13}$$

Therefore, errors caused by frequency drift and random measurement error are  $u$  times narrower than the general method.

In the above method,  $u$  can be determined by the following equation:

$$u \left[ \frac{f_{x+u} - f_x}{f_{x+1} - f_x} \right]. \tag{14}$$

According to the expression of  $\Delta u$ :

$$\Delta u = \left[ \frac{\Delta(f_{x+u} - f_x)}{f_b} - \frac{\Delta(f_{x+1} - f_x)}{u \cdot f_b} \right] = \left[ -\frac{\pm \Delta f_u \pm (|v_{x+u}| + |v_x|)}{f_b} - \frac{\pm \Delta f_1 \pm (|v_{x+1}| + |v_x|)}{u \cdot f_b} \right]. \tag{15}$$

The items are small quantities, that is,  $\pm \Delta f_u \pm (|v_{x+u}| + |v_x|) / f_b \ll 1$ ,  $\pm \Delta f_1 \pm (|v_{x+1}| + |v_x|) / (u f_b) \ll 1$ , so  $u$  can be accurately acquired.

### 3. Experimental Results and Discussion

In order to verify the effectiveness of above methods, the experimental system shown in Figure 1 was established and several experiments were conducted. In the system, a He-Ne laser (Newport R-32734) served as the light source, a high-frequency phase modulator (Newport, 4431,  $V_\pi$  is 40 V) served as the modulator, a signal generator (Keysight, N5171B) provided the modulation signal, the preset sweeping range was set from 2.500 to 2.520 GHz, with a sweep step size of 1 kHz. The residence time at each frequency was 20 ms. An avalanche detector (Thorlabs, APD410A) was used for the photoelectric conversion. A PXI oscilloscope (NI, 5171R) served as the acquisition and processing module. The target was approximately 80 m away from the crystal.

To evaluate the method of reciprocating sweep in compensation of thermally induced phase delay, an “in phase” frequency was obtained by different methods and results are shown in Figure 4. In the figure, black and red dots respectively denote the results obtained through forward and backward sweep, and blue dots denote the average of measurement results obtained by the reciprocating sweep. Results obtained through forward and backward sweeps drifted in opposite directions, and the forward sweeps had a larger statistical average than that of the backward sweeps. Through reciprocating sweeps, the results had a repeatability of 1727 Hz, which was much smaller than that of forward sweep (3130 Hz) and backward sweep (2503 Hz).

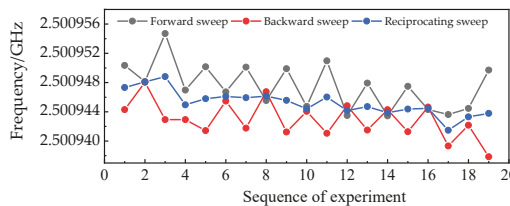


Figure 4. Experimental results of “in phase” frequency.

To validate our measurement method of  $N$ ,  $N$  was tested with multi-frequency intervals. The “in phase” frequency  $f_x$  near the lower edge of band was first obtained in the reciprocating sweep method, then, the  $u$ th “in phase” frequency on its upper side was obtained, and  $f_b$  and  $N$  were respectively determined according to Equations (10) and (11).  $u$  was set to change from 1 to 10 in seven groups of tests, each group of tests contained 10 independent measurements; the statistical results are shown in Figure 5. In the single-frequency interval method, the mean squares of  $f_b$  and  $N$  were 2384 Hz and 1.5, respectively. As  $u$  increased from 1 to 8, the mean square of  $f_b$  was reduced to 332 Hz, and the mean square of  $N$  was reduced to 0. Further increase of  $u$  would result in further reduction of the standard deviation of  $f_b$ , but in our experiment, the frequency interval number of 8 was sufficient for an accurate measurement of  $N$ .

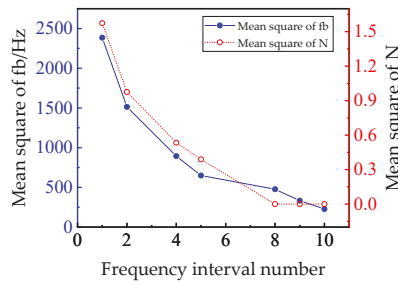


Figure 5. Mean square of  $f_b$  and  $N$  with different frequency interval numbers.

With the accurate measurement of  $N$ , distance could be determined through the “in phase” frequency and the corresponding  $N$ ; results are shown in Figure 6. The statistical average of the forward sweeps was 6  $\mu\text{m}$  larger than the backward sweeps. The statistical uncertainties of the forward sweeps and reciprocating sweeps were 87.1  $\mu\text{m}$  and 44.1  $\mu\text{m}$ , respectively.

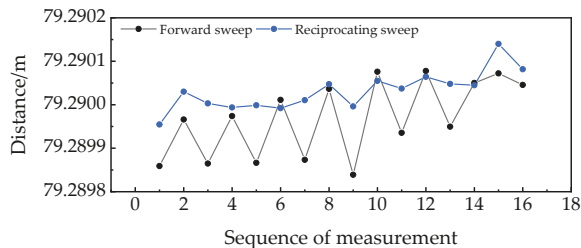
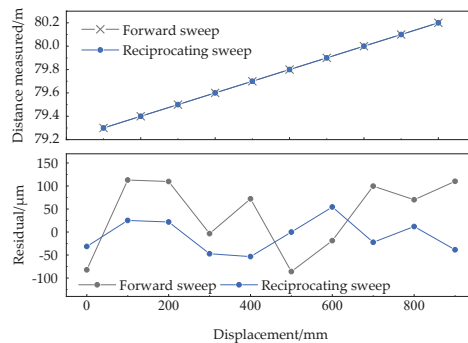


Figure 6. Measurement results under different sweep methods.

Finally, an uncertainty test was conducted using a dual-frequency laser interferometer. The target retroreflector was moved from 79.29 to 80.29 m by ten steps at an increment of 100 mm. A commercial heterodyne interferometer (5519, Keysight) was used to measure the displacement of the target. By the application of linear fitting, the residual error between the reference displacement and the measured distance could be obtained as shown in Figure 7. The residual error of the forward sweep was from  $-85.9$  to  $113 \mu\text{m}$ . With the reciprocating sweep method, the max residual error was  $54.4 \mu\text{m}$ . The relative error was attributed to the random measurement error of frequency, the temperature dependence of crystal properties, the air-path distance variation, and the instability of frequency reference. The random measurement error of frequency could be further reduced by reducing the sweep step size and increasing the signal-to-noise ratio of the system. The temperature dependence of crystal property included the refractive index and the length change on the optical path, and air-path distance variation was mainly caused by the variation of the air refractive index.





**Figure 7.** Residuals of the measured distance versus displacement from a commercial heterodyne interferometer.

#### 4. Conclusions

We systematically discussed the principle of the FSPM based ranging method and the impact of thermally induced phase delay on “in phase” frequency and wavelength multiple. To eliminate the frequency drift caused by the phase delay, a measurement method based on a reciprocating sweep was proposed. To obtain an accurate wavelength multiple, a multi-frequency interval approach was proposed according to the equal interval distribution of “in phase” frequencies. The results of our experiment proved the efficiency and applicability of the methods. The maximum deviation of the measured results between the FSPM-based system and the interferometer was 54.4  $\mu\text{m}$ .

**Author Contributions:** W.Z.; Conceptualization, S.G., Y.L. and R.J.; methodology, C.L.; validation, S.G., Y.L. and Y.P.; formal analysis, R.J.; resources, S.G.; writing—original draft preparation, Y.L., J.S., Y.P.; writing—review and editing, W.Z.; supervision.

**Funding:** This research was funded by Civil Aerospace Pre-research Project, grant number D020214.

**Conflicts of Interest:** The authors declare no conflict of interest.

#### References

- Gao, W.; Kim, S.W.; Bosse, H.; Haitjema, H.; Chen, Y.L.; Lu, X.D.; Knapp, W.; Weckenmann, A.; Estler, W.T.; Kunzmann, H. Measurement technologies for precision positioning. *Cirp Ann. Manuf. Technol.* **2015**, *64*, 773–796. [[CrossRef](#)]
- Parker, D.H. Nondestructive testing and monitoring of stiff large-scale structures by measuring 3D coordinates of cardinal points using electronic distance measurements in a trilateration architecture. In Proceedings of the Nondestructive Characterization and Monitoring of Advanced Materials, Aerospace, and Civil Infrastructure, Portland, OR, USA, 26–29 March 2017.
- Golygin, N.K.; Lysenko, V.G.; Khizhnyakov, V.A. Metrological support for opto-electronic coordinate measurements. *Meas. Tech.* **2017**, *59*, 1073–1077. [[CrossRef](#)]
- Servagent, N.; Bosch, T.; Lescure, M. Design of a phase-shifting optical feedback interferometer using an electrooptic modulator. *IEEE J. Sel. Top. Quantum Electron.* **2000**, *6*, 798–802. [[CrossRef](#)]
- Fujima, I.; Iwasaki, S.; Seta, K. High-resolution distance meter using optical intensity modulation at 28 ghz. *Bull. NRLM* **1998**, *48*, 1049. [[CrossRef](#)]
- Castagnet, D.; Tap-Beteille, H.; Lescure, M. Avalanche-photodiode-based heterodyne optical head of a phase-shift laser range finder. *Opt. Eng.* **2006**, *45*, 828. [[CrossRef](#)]
- John, D.; Ben, H.; Andrew, J.L.; Andrew, J.L.; Armin, J.H.R.; Matthew, S.W. Multi-channel absolute distance measurement system with sub ppm-accuracy and 20 m range using frequency scanning interferometry and gas absorption cells. *Opt. Express* **2014**, *22*, 24869–24893.
- Xiong, X.-T.; Qu, X.-H.; Zhang, F.-M. Error correction for fsi-based system without cooperative target using an adaptive filtering method and a phase-matching mosaic algorithm. *Appl. Sci.* **2018**, *8*, 1954. [[CrossRef](#)]

9. Medhat, M.; Sobee, M.; Hussein, H.M.; Terra, O. Distance measurement using frequency scanning interferometry with mode-hoped laser. *Opt. Laser Technol.* **2016**, *80*, 209–213. [[CrossRef](#)]
10. Weimann, C.; Hoeller, F.; Schleitzer, Y.; Diez, C.A.; Spruck, B.; Freude, W.; Boeck, Y.; Koos, C. Measurement of length and position with frequency combs. *J. Phys. Conf. Ser.* **2015**, *605*, 012030. [[CrossRef](#)]
11. Kim, S.-W.; Kim, Y.-J.; Hyun, S.; Chun, B.J.; Jang, Y.-S. Recent advances in absolute distance measurements using femtosecond light pulses. In Proceedings of the Ninth International Symposium on Precision Engineering Measurements and Instrumentation, Changsha, China, 8–10 August 2014.
12. Tan, J.; Yang, H.; Hu, P.; Diao, X. Identification and elimination of half-synthetic wavelength error for multi-wavelength long absolute distance measurement. *Meas. Sci. Technol.* **2011**, *22*, 115301. [[CrossRef](#)]
13. Meiners-Hagen, K.; Schödel, R.; Pollinger, F.; Abou-Zeid, A. Multi-wavelength interferometry for length measurements using diode lasers. *Meas. Sci. Rev.* **2009**, *9*, 16–26. [[CrossRef](#)]
14. Sa, V.D.B.; Persijn, S.T.; Kok, G.J.; Zeitouny, M.G.; Bhattacharya, N. Many-wavelength interferometry with thousands of lasers for absolute distance measurement. *Phys. Rev. Lett.* **2012**, *108*, 183901.
15. Ji, R.; Hu, K.; Li, Y.; Gao, S.; Zhou, W. Study of femtosecond laser spectrally resolved interferometry distance measurement based on excess fraction method. In Proceedings of the Second International Conference on Photonics and Optical Engineering, Xi'an, China, 14–17 October 2017.
16. Li, Y.; Shi, J.; Wang, Y.; Ji, R.; Liu, D.; Zhou, W. Phase distortion correction in dual-comb ranging system. *Meas. Sci. Technol.* **2017**, *28*, 075201. [[CrossRef](#)]
17. Zhou, W.; Shi, J.; Ji, R.; Li, Y.; Liu, Y. High-precision distance measurement using femtosecond laser frequency comb. *Chin. J. Sci. Instrum.* **2017**.
18. Gao, S.; Ji, R.; Li, Y.; Cheng, Z.; Zhou, W. Experiments and error analysis of laser ranging based on frequency-sweep polarization modulation. *Proc. SPIE* **2016**, *23*, 1002322.
19. Hei, K.F.; Yu, J.L.; Wang, J.; Wang, W.R.; Jia, S.; Wu, Q.; Xue, J.Q. Variable frequency range finding technology based on double polarization modulation method and system implementation. *Acta Phys. Sin.* **2014**, *63*, 100602.
20. Xiao, Y.; Wang, J.; Wang, W.-R.; Wang, Z.-X.; Xie, T.-Y.; Yu, Y.; Xue, J.-Q. Relationship between modulation frequency and range accuracy in the double polarization modulation range finding system. *Acta Phys. Sin.* **2016**, *65*, 601.
21. Meier, D. Electro-Optical Measuring Device for Absolute Distances. U.S. Patent Application No. 5764360, 9 June 1998.
22. Meier, D. Method and Measuring Device for Measuring an Absolute Distance. U.S. Patent Application No. 7609387, 10 April 1990.
23. Meier, D. Electrooptical Distance Measuring Apparatus. U.S. Patent Application No. 4759623, 20 April 1988.
24. Sima, W.; Liu, T.; Yang, Q.; Han, R.; Sun, S. Temperature characteristics of pockels electro-optic voltage sensor with double crystal compensation. *AIP Adv.* **2016**, *6*, 055109. [[CrossRef](#)]
25. Wong, K.K. *Properties of Lithium Niobate*; IET: London, UK, 2002.
26. Zaky, S.G. Optical Modulator Having Compensation for Thermal and Space Charge Effects. U.S. Patent Application No. 3900247, 19 August 1975.
27. Sabatke, D.; Handorf, R.V.; Sullivan, J. Polarization and fold mirrors in application of the leica absolute distance meter. *Proc SPIE* **2009**, *7461*, 485–560.
28. Luethi, T.; Meier, D. Electro-Optical Distance-Measuring Unit. U.S. Patent Application No. 8289524, 6 November 2012.



© 2019 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).

Article

# Measurement Vertex Position of a Large Aspheric Mirror

Jie Li \*, Fengtao Yan, Lin Chen, Jie Yang and Yongjian Wan

Institute of Optics and Electronics, Chinese Academy of Sciences, Chengdu 610209, China; yanfengtao427@163.com (F.Y.); chenlinlg@163.com (L.C.); zkb@ioe.ac.cn (J.Y.); yjwan@ioe.ac.cn (Y.W.)

\* Correspondence: lijie@ioe.ac.cn

Received: 30 January 2019; Accepted: 4 March 2019; Published: 13 March 2019



**Abstract:** We present a novel method to accurately measure the vertex position of a large aspheric mirror with a phasing measuring interferometer, a laser tracker, and a micro alignment telescope. The method consists of rotating the mirror by 180° about its mechanical axis and measuring the change in displacement. Use of a micro alignment telescope eliminates the need to adjust the mirror during the measurement, eases the alignment of the testing system, and improves the vertex position measurement accuracy. Using this method, we measured the vertex position of an aspheric mirror 1 m in diameter and 2 m in radius of curvature. The vertex position measurement uncertainty is 88 μm.

**Keywords:** optical measurement; dimensional measurement; aspheric mirror; vertex position

## 1. Introduction

Accurate measurement of the vertex position of a large aspheric mirror is often critical to the performance of the optical system in which the mirror is used. Take the monolithic primary tertiary mirror as an example, the Large Synoptic Survey Telescope (LSST) uses a unique opto-mechanical design that places the primary and tertiary mirrors on a single glass substrate, the alignment of primary and tertiary mirrors is controlled at the time of manufacture; the tolerance in lateral displacement of the two vertices is 1 mm [1–3]. Another example is the space telescope, which has very limited alignment space, and the vertex eccentricity of the primary mirror is often required to within less than a millimeter from nominal.

The displacement of the vertex causes shear of the spherical aberration terms, resulting in coma in surface form [4]. However, it is difficult to determine vertex position from measurement of surface form because of low sensitivity. Take a parabolic mirror with 1.3 m in diameter and 3.5 m in radius of curvature (RoC) as an example, an error of 1 mm in the measurement of the vertex position only induces 2 μm (Peak-to-Valley, PV) surface form error.

A variety of methods exists to measure the vertex position of an aspheric mirror. The mechanical method uses a coordinate measuring system, such as a coordinate measuring machine (CMM) or laser tracker to measure the surface coordinates and calculate the best-fit vertex position; the mechanical method is less efficient and often not applicable for large polished mirrors. The second method uses a phase measuring interferometer (PMI) and a null corrector. When the test optics are aligned to the mirror surface, the null corrector accurately defines the optical axis. The position and orientation of the null corrector and the position of the mirror's mechanical axis can be measured by a portable coordinate measuring system (CMS, such as a laser tracker), and the vertex position can be subsequently calculated. A shortcoming of this method is that a small error in the measurement of the null corrector's orientation results in a large error in the calculation of the vertex position due to Abbe error [5]. The third method consists of rotating the mirror by 180 degrees about its mechanical axis and measuring the change in coma. Two drawbacks associated with this method are as follows. First, the method depends on a high

accuracy rotating table and precise alignment of the test piece so that the PMI can measure the mirror surface both at 0 degrees and 180 degrees without adjustment. Second, the method relies on stable positioning of the test optics, which is difficult in large-scale dimensional measurement.

The need to carefully match the vertex position of a large aspheric mirror presents an interesting challenge for the optics shop. Accurately measuring the vertex position using conventional methods is extremely difficult. In the following sections, we present a technique to use a PMI, a laser tracker and a micro alignment telescope (MAT) to perform vertex position measurement on an aspheric mirror. The method eliminates the aforementioned drawbacks, and accurate measurement of the vertex position is achieved.

## 2. Methods

The method is actually an improvement of the rotating method. Instead of measuring the change in coma with the PMI, the method only uses the PMI for alignment and directly measures the displacement with the MAT.

In this method, a PMI and a null corrector which generates an aspheric wavefront are used to build a virtual axis. The virtual axis will always pass through the mirror’s vertex when the test optics are well aligned to the mirror surface at each rotating angle.

A laser tracker is used to align a visible target for MAT to the mirror’s mechanical center (often defined by the outer diameter). A laser tracker is a portable CMS that utilizes a distance measuring interferometer (DMI) and two rotary encoders to track and measure the location of a sphere-mounted retro-reflector (SMR).

The MAT allows the user to set reference lines of sight for applications such as the alignment and measurement of bearings and bores. By use of two built in optical micrometers, two dimensional displacements can be measured with high accuracy from a reference line of sight. The MAT is moved together with the test optics so that it can maintain relative position and orientation with the virtual axis. If the vertex is not at mirror’s mechanical center, the MAT could observe target displacement after rotating the mirror and re-aligned to the test optics. Thus, the vertex position can be determined with two dimensional displacements from the MAT.

The measurement sequence is as follows (see Figure 1):

1. Use a laser tracker to measure the mechanical center of the mirror and align a target (for MAT) to mirror’s mechanical center, record the eccentricity  $x_{target}, y_{target}$ .
2. Deploy the PMI, null corrector such that the optical axis of test optics coincides as closely as possible with mirror’s optical axis, read the coma terms  $c_{x_0}, c_{y_0}$ , and calculate alignment deviation  $x_{coma_0}, y_{coma_0}$  with mirror’s parameter.
3. Deploy the MAT, place it along with PMI and test optics (on the same 5-axis table), and aim at the target center, record the reading of two micrometers:  $x_{mat_0}, y_{mat_0}$ .
4. Rotate the mirror by 180 degrees, adjust 5-axis table to re-align optical path and calculate alignment deviation  $x_{coma_{180}}, y_{coma_{180}}$  with coma terms.
5. Adjust the MAT’s micrometers, aim at the target center, and record the reading of micrometers  $x_{mat_{180}}, y_{mat_{180}}$ .
6. The displacements between the vertex position and mirror’s mechanical center can be calculated as:

$$\begin{aligned} x_{vertex} &= (x_{mat_{180}} - x_{coma_{180}} + x_{coma_0} - x_{mat_0})/2 + x_{target} \\ y_{vertex} &= (y_{mat_{180}} - y_{coma_{180}} + y_{coma_0} - y_{mat_0})/2 + y_{target} \end{aligned} \tag{1}$$

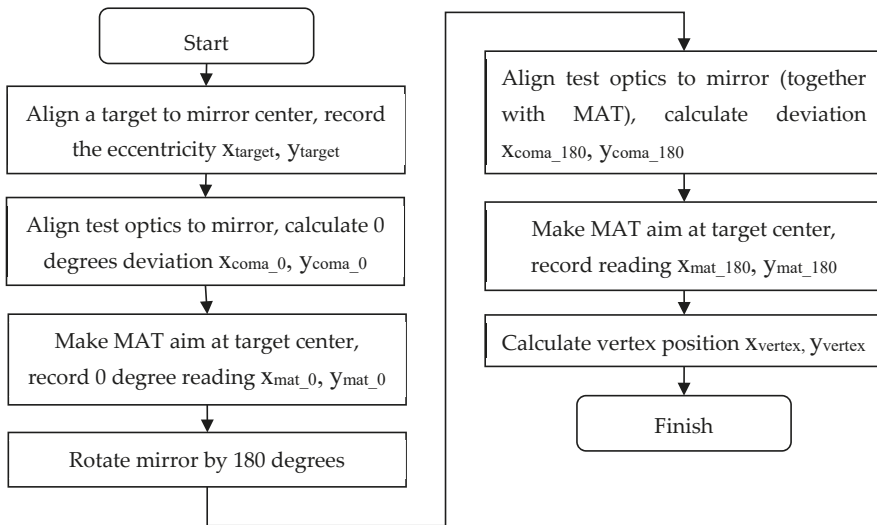


Figure 1. Vertex position measurement process.

### 3. Experiment and Results

The preliminary experiment is carried out on a concave aspheric mirror 1 m in diameter and 2.1 m in RoC. The mirror is supported by a whiffletree structure on a steel plate and can be rotated manually. The form error of the mirror is  $0.1\lambda$  (rms, with piston, tilt, power and coma terms removed,  $\lambda = 632.8$  nm).

The optical target for MAT is mounted on a hollow reference cylinder (see Figure 2). An OGP SmartScope ZIP 250 vision measuring system (maximum permissible error (MPE):  $1.8 + 6 L/1000 \mu\text{m}$ , where L is the measured length in mm) is used to measure the eccentricity between the target and reference cylinder. The vision measuring system measures the reference cylinder’s outer edge (roundness is  $3 \mu\text{m}$ ) and analyzes cylinder’s center position, then it aims at the central circular dot of the target with a crossline to determine the eccentricity. The eccentricity is measured and adjusted to  $3 \mu\text{m}$  in x direction and  $-2 \mu\text{m}$  in y direction. A rectangular light emitting diode (LED) is fixed in the reference cylinder and it provides uniform illumination light which is suitable for observation.

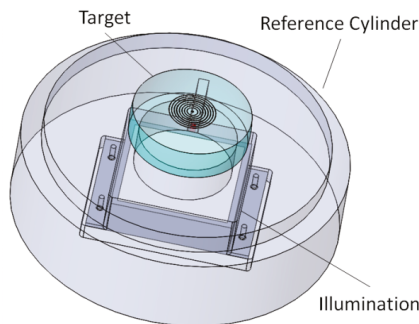


Figure 2. Target for MAT mounted on a hollow reference cylinder.

An API Radian laser tracker (range: 40 m) [6] is used to align the target to mirror’s mechanical center. The laser tracker is placed 2.5 m away from the mirror’s mechanical center. The laser tracker measures the outer edge of the mirror and the reference cylinder and analyzes center positions

respectively. The eccentricity of two center positions is repeatedly measured and adjusted. The MPE of laser tracker is  $10 + 5 L \mu\text{m}$  ( $L$  is measured length in m) and the eccentricity of two center positions is finally controlled within  $50 \mu\text{m}$ .

A Zygo GPI-XP/DPMI (test beam diameter: 4 inch) [7] is placed horizontally on the test tower with a  $45^\circ$  fold mirror which makes easy to align the optical path of heavy PMI system (see Figure 3). To align the optical path, a large 5-axis table which enables precision translation in the X, Y and Z axes, as well as tip-tilt adjustment for the PMI measurement is used. The refractive null corrector is also placed on the large 5-axis table and it is aligned to the PMI with a small 5-axis table by using interference fringe in the first place, then the large 5-axis table is precisely adjusted to align the optical path until the residual coma and defocus terms are within  $0.2\lambda$ .

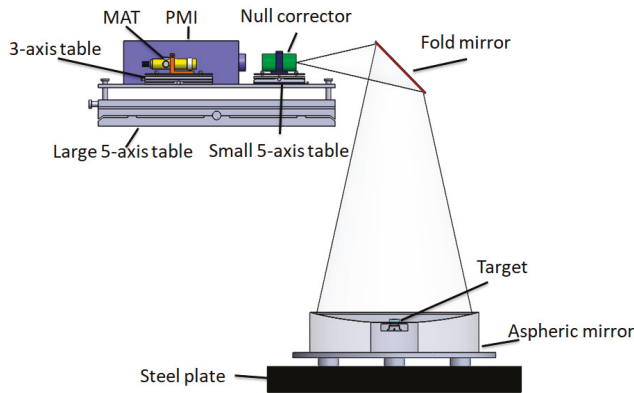


Figure 3. Setup to measure vertex position of aspheric mirror.

A MAT with built in optical micrometers (graduations:  $0.02 \text{ mm}$ , error:  $\pm 0.002 \text{ mm}$ ) from Taylor Hobson [8] is used to observe the lateral displacements of the target, as shown in Figure 3, the MAT is placed beside the PMI on a small 3-axis table, 3 m away from the mirror. The angle between the PMI axis and MAT axis is approximately 7 degrees in x direction, due to ‘Cosine error’, the actual displacement in x direction was 1.01 times the measured displacement. The measurement results are listed in Table 1.

Table 1. Results of the vertex position measurement.  $x_{\text{tar-ref}}$ ,  $y_{\text{tar-ref}}$  are displacements between the target center and reference cylinder center in x, y directions;  $x_{\text{ref-mir}}$ ,  $y_{\text{ref-mir}}$  are displacements between the reference cylinder center and mirror center in x, y directions.

Item	0 Degrees ( $\mu\text{m}$ )	180 Degrees ( $\mu\text{m}$ )
$x_{\text{tar-ref}}$	3	–
$y_{\text{tar-ref}}$	–2	–
$x_{\text{ref-mir}}$	21	–
$y_{\text{ref-mir}}$	42	–
$x_{\text{coma}}$	13	–7
$y_{\text{coma}}$	–6	14
$x_{\text{mat}}$	40	120
$y_{\text{mat}}$	140	320
$X_{\text{vertex}}$		74
$Y_{\text{vertex}}$		120

The measurement uncertainty analysis is as follows:

The measurement model for the vertex position is shown in Equation (1), while in the experiment,

$$\begin{aligned} x_{\text{target}} &= x_{\text{tar-ref}} + x_{\text{ref-mir}} \\ y_{\text{target}} &= y_{\text{tar-ref}} + y_{\text{ref-mir}} \end{aligned} \quad (2)$$

The uncertainty model of the vertex position measurement is shown in Equation (3) [9]:

$$u_{\text{vertex}}^2 = u_{\text{mat}}^2/2 + u_{\text{coma}}^2/2 + u_{\text{tar-ref}}^2 + u_{\text{ref-mir}}^2 \quad (3)$$

According to the MPE value of the SmartScope ZIP 250 (assigning rectangular distribution), the displacement between the target center and reference cylinder center measurement uncertainty is (Type B evaluation of uncertainty):

$$u_{\text{tar-ref}} = 2 \mu\text{m} \quad (4)$$

According to the MPE value of the API Radian laser tracker (assigning rectangular distribution), the displacement between the reference cylinder center and mirror center measurement uncertainty induced by the laser tracker is (Type B evaluation of uncertainty):

$$u_{\text{ref-mir\_lt}} = 14 \mu\text{m} \quad (5)$$

The displacement between the reference cylinder center and the mirror center measurement uncertainty evaluated by statistical analysis of 5 measurements is (Type A evaluation of uncertainty):

$$u_{\text{ref-mir\_repeatability}} = 24 \mu\text{m} \quad (6)$$

According to the measurement experience, the surface coma measurement uncertainty is 60 nm, to this mirror, the correlation coefficient of measured coma and lateral displacement is 227. The measurement uncertainty induced by coma measurement is (Type B evaluation of uncertainty):

$$u_{\text{coma}} = 14 \mu\text{m} \quad (7)$$

The uncertainty arises from the MAT reading is evaluated by measuring a target which mounted on a linear slider at a 3 m range. The statistical analysis of the repeatability and ranging test gives (Type A evaluation of uncertainty):

$$u_{\text{mat\_reading}} = 10 \mu\text{m} \quad (8)$$

The drift of relative orientation between the MAT axis and test optics axis is evaluated by the repeatability test. The statistical analysis of 6 times reading from the PMI and MAT (in an hour) gives (Type A evaluation of uncertainty):

$$u_{\text{mat\_drift}} = 45 \mu\text{m} \quad (9)$$

The combined uncertainty of the vertex position measurement is as follows:

$$u_{\text{vertex}} = ((u_{\text{mat\_reading}}^2 + u_{\text{mat\_drift}}^2)/2 + u_{\text{coma}}^2/2 + u_{\text{tar-ref}}^2 + u_{\text{ref-mir\_lt}}^2 + u_{\text{ref-mir\_repeatability}}^2)^{0.5} = 44 \mu\text{m} \quad (10)$$

The expanded uncertainty ( $k = 2$ ) is as follows:

$$U_{\text{vertex}} = k \times u_{\text{vertex}} = 88 \mu\text{m} \quad (11)$$

The analysis indicates the vertex position expanded measurement uncertainty is 88  $\mu\text{m}$ . For comparison, the mirror is also measured with a CMM (MPE:  $2.2 + L/350 \mu\text{m}$ , where L is measured length in mm). The vertex position x displacement is  $-90 \mu\text{m}$  and y displacement is  $170 \mu\text{m}$ . The vertex position measurement uncertainty of the CMM method is evaluated using statistical way (rotate the mirror 3 times by 90 degrees and use a CMM to measure the vertex position at 4 directions) and

vertex position measurement uncertainty is 0.29 mm in x direction and 0.12 mm in y direction. Thus, the result of the proposed method is within the uncertainty of the CMM measurement.

#### 4. Conclusions

We presented a method to use a PMI, a laser tracker, and a MAT to accurately measure the vertex position of an aspheric mirror. The MAT is adjusted with the PMI to maintain relative position and orientation, and it measures the displacements of the target placed on the mechanical center of the mirror which makes vertex position measurement more direct. This is a convenient way to measure the vertex position of an aspheric mirror, especially suitable for controlling the mirror's vertex position during the polishing process. A validation experiment is performed on a concave aspheric mirror 1 m in diameter and 2.1 m in RoC and the experimental results are compared with that of CMM, the results have confirmed the feasibility of the proposed method.

**Author Contributions:** Methodology, L.J.; validation, L.J., C.L., and Y.J.; writing—original draft preparation, L.J.; investigation, Y.T.; resources, W.J.

**Funding:** This research was funded by Sichuan Province Science and Technology Support Program, grant number “2017GZ0072”.

**Conflicts of Interest:** The authors declare no conflict of interest.

#### References

1. Sweeney, D.W.; Stepp, L.M. Overview of the Large Synoptic Survey Telescope project. In Proceedings of the SPIE Astronomical Telescopes + Instrumentation, Orlando, FL, USA, 24 May 2006; Volume 6267, p. 626706.
2. Tuell, M.T.; Martin, H.M.; Burge, J.H.; Gressler, W.J.; Zhao, C. Optical testing of the LSST combined primary/tertiary mirror. In Proceedings of the SPIE Astronomical Telescopes + Instrumentation, San Diego, CA, USA, 27 June 2010; Volume 7739, p. 77392V.
3. Martin, H.M.; Angel, J.R.P.; Angeli, G.Z.; Burge, J.H.; Gressler, W.; Kim, D.W.; Kingsley, J.S.; Law, K.; Liang, M.; Neill, D.; et al. Manufacture and final tests of the LSST monolithic primary/tertiary mirror. In Proceedings of the SPIE Astronomical Telescopes + Instrumentation, Edinburgh, UK, 26 June 2016; Volume 9912, p. 99120X.
4. Martin, H.M.; Burge, J.H.; Cuerden, B.; Davison, W.B.; Kingsley, J.S.; Lutz, R.D.; Miller, S.M.; Tuell, M. Manufacture of a combined primary and tertiary mirror for the Large Synoptic Survey Telescope. In Proceedings of the SPIE Advanced Optical and Mechanical Technologies in Telescopes and Instrumentation, Marseille, France, 23 June 2008; Volume 7018, p. 70180G.
5. Leach, R. Abbe Error/Offset. In *CIRP Encyclopedia of Production Engineering*; Springer: Berlin, Germany, 2014; pp. 1–4. [[CrossRef](#)]
6. API Radian Laser Tracker. Available online: <https://www.apisensor.com/products/3d-laser-tracker-systems/radian> (accessed on 26 January 2019).
7. Zygo Phase Measurement Interferometers. Available online: <https://www.zygo.com/?/met/interferometers> (accessed on 26 January 2019).
8. Taylor Hobson Micro Alignment Telescope. Available online: <https://www.taylor-hobson.com/products/alignment-level/micro-alignment-telescopes/micro-alignment-telescopes-with-built-in-optical-micrometers> (accessed on 26 January 2019).
9. JCGM100: 2008. Evaluation of Measurement Data. Available online: <http://www.iso.org/sites/JCGM/GUM/JCGM100/C045315e-html/C045315e.html?csnumber=50461> (accessed on 26 January 2019).



© 2019 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).



Article

# Characterization of Volume Gratings Based on Distributed Dielectric Constant Model Using Mueller Matrix Ellipsometry

Hao Jiang \*, Zhao Ma, Honggang Gu, Xiuguo Chen and Shiyuan Liu \*

State Key Laboratory for Digital Manufacturing Equipment and Technology,  
Huazhong University of Science and Technology, Wuhan 430074, China; zhaoma@hust.edu.cn (Z.M.);  
hongganggu@hust.edu.cn (H.G.); xiuguochen@hust.edu.cn (X.C.)

\* Correspondence: hjiang@hust.edu.cn (H.J.); shyliu@hust.edu.cn (S.L.)

Received: 17 January 2019; Accepted: 14 February 2019; Published: 18 February 2019



**Abstract:** Volume grating is a key optical component due to its comprehensive applications. Other than the common grating structures, volume grating is essentially a predesigned refractive index distribution recorded in materials, which raises the challenges of metrology. Although we have demonstrated the potential application of ellipsometry for volume grating characterization, it has been limited due to the absence of general forward model reflecting the refractive index distribution. Herein, we introduced a distributed dielectric constant based rigorous coupled-wave analysis (RCWA) model to interpret the interaction between the incident light and volume grating, with which the Mueller matrix can be calculated. Combining with a regression analysis with the objective to match the measured Mueller matrices with minimum mean square error (MSE), the parameters of the dielectric constant distribution function can be determined. The proposed method has been demonstrated using a series of simulations of measuring the volume gratings with different dielectric constant distribution functions. Further demonstration has been carried out by experimental measurements on volume holographic gratings recorded in the composite of polymer and zinc sulfide (ZnS) nanoparticles. By directly fitting the spatiotemporal concentration of the nanoparticles, the diffusion coefficient has been further evaluated, which is consistent to the result reported in our previous investigations.

**Keywords:** ellipsometry; volume grating; nanostructure metrology; distributed dielectric constant model; holography

---

## 1. Introduction

Volume gratings, which are formed by introducing a periodic refractive index modulation within the volume of a bulk material, are of great importance and popularly used in optical physics for data storage [1], optical elements [2], as well as optical communications [3]. Many kinds of materials, such as crystals [4], fused silica [5,6], photo-thermo-refractive (PTR) glass [7–9], and polymers [10,11], have been employed as the recording media. In order to achieve larger refractive index modulation as well as higher data storage density in flexible devices with lightweight [12–14], polymer nanocomposites, with doped nanoparticles [15–24], has attracted special attentions.

Other than common grating structures that can be easily observed using techniques such as microscopy, volume grating does not exhibit structural characteristics such as grating groove and ridge physically. Therefore, metrology of a volume grating is challenging. Sabel and Zschocher avoided the limit of resolution and achieved the image of the volume phase gratings recorded in polymer using an optical microscopy [25]. Braun and co-workers had tried to monitor the eventual status of nanoparticle distribution in the holographic gratings using transmission electron microscopy

(TEM) [26]. Although insightful information has been achieved, it is highly desirable to observe the volume gratings non-destructively. In practice, diffraction efficiency has been used comprehensively by many researchers to evaluate volume gratings [27,28]. However, since the spatial distribution is not the only factor affecting the diffraction efficiency [29], it was not always a good predictor on nanoparticle distribution. Chen et al. proposed a method based on angular selectivity curves to measure the refractive index modulation (RIM, i.e., the refractive index difference between bright and dark regions) of a volume grating recorded in photo-thermo-refractive glass and achieved high precision [9]. Butcher et al. measured the thickness, RIM and duty ratio of a volume diffraction grating using a commercial Fourier-transform spectrometer (FTS) combined with multi-incident angle measurement [30].

Recently, considering the advantages such as nondestructive to the samples, high sensitivity on anisotropy, as well as the capability of dealing with depolarization, we introduced Mueller matrix ellipsometry (MME) to characterize the volume holographic gratings recorded in a composite of poly(acrylate-co-acrylamide) and 5-nm zinc sulfide (ZnS) nanoparticles, and further studied the process of nanoparticle diffusion upon holography [31]. The time-dependent parameters had been achieved, such as the bright and dark region width, refractive index and nanoparticle volume fractions, which pave the way for the quantitative study of the nanoparticle diffusion process. However, due to the absence of general refractive index distribution model, an assumption of rectangular cross-section of the grating has been made, which may degrade the fidelity of the metrology, especially in the cases when the distribution of refractive index is not sinusoidal [32,33] or significant absorption exists [34].

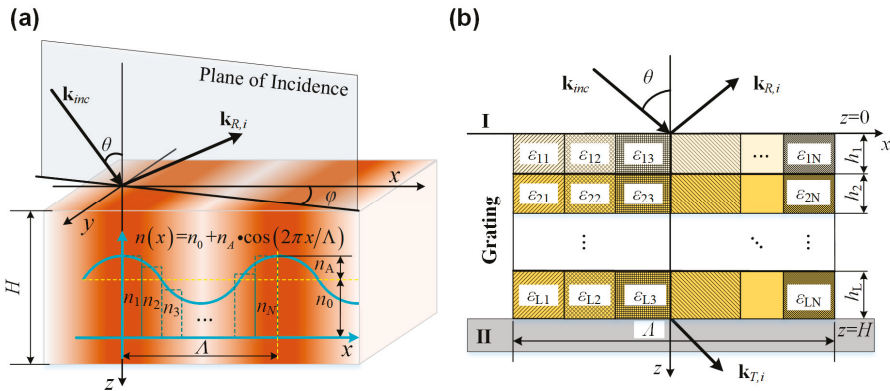
In this work, based on the rigorous coupled-wave analysis (RCWA) theory [35], we proposed a distributed dielectric constant model to interpret the interaction between the incident light and volume grating and calculate the Mueller matrices. In such a model, the dielectric constant of the volume grating is described by a general periodic spatial function, and two-dimensional discretization in one period have been carried out. By a regression analysis with the objective to best match the measured Mueller matrices, the refractive index distribution function can be reconstructed, as well as the thickness of the grating. The proposed method has been demonstrated using a series of simulations of measuring the volume gratings with different dielectric constant distribution functions. Measurement experiments have been carried out on the volume holographic gratings recorded in a composite of polymer and 5-nm zinc sulfide (ZnS) particles with different recording time of 5 s, 10 s, 15 s, 20 s, 25 s, 30 s, 35 s and 40 s. The rational results demonstrate the validity of the proposed method. By directly fitting the spatiotemporal concentration of ZnS nanoparticles, an apparent diffusion coefficient of  $2.18 \times 10^{-15} \text{ m}^2 \text{ s}^{-1}$  has been achieved, which agree with our previously reported results of  $2.0 \times 10^{-15} \text{ m}^2 \text{ s}^{-1}$  [31].

## 2. Methods

Ellipsometry is a typical model-based technology, which usually involves a multiparameter forward model and the solution of an inverse problem. The former describes the interaction between the probing light and the sample, while the objective of the latter is fitting the measured data with the theoretical outputs of the forward model. Determining an appropriate model to accurately calculate the polarization state change of the incident beam induced by the light-nanostructure interaction is the premise of a successful metrology.

Different to common grating structures consists of ridge and groove that are typically made of two different materials (such as air and grating material), the volume grating is essentially a specific distribution of refractive indices in the material, which can be depicted as Figure 1a. The fundamental principle of such distribution formation is based on the photoinduced property changes like photopolymerization induced diffusion, or ultraviolet radiation and thermal development, giving rise to expected refractive index distribution in the recording material, such as polymer-nanoparticle composites as well as PTR glasses, etc. As in a polymer-nanoparticle system, the photo-active monomer consumption occurs in bright regions during holography because of photopolymerization, which results in the movement of monomers towards the bright regions.

Consequently, the nanoparticles will be squeezed into the dark regions due to the chemical potential effects. As a result, a sinusoidal refractive index distribution as shown in Figure 1a is expected in the formed polymer nanocomposite. In order to accurately describe such a system, a distributed dielectric constant-based RCWA is proposed as the forward model for volume grating characterization.



**Figure 1.** Scheme of the (a) spatial refractive index distribution in a volume grating; and (b) the corresponding distributed dielectric constant model used for rigorous coupled-wave analysis.

2.1. Distributed Dielectric Constant-Based RCWA

For a volume grating whose refractive index distributed in the volume grating can be defined by an arbitrary periodic spatial function  $n(x)$  as shown Figure 1a, a RCWA model based on a distributed refractive index model can be discretized as Figure 1b. The geometric domain is divided into three regions, i.e. region I, II and the grating region, and the coordinate system is defined as shown in Figure 1. Angle of incident  $\theta$  and angle of azimuth  $\varphi$  are defined as shown in Figure 1a as well. If a specific function of refractive index distribution can be predetermined, such as the sinusoidal distribution shown in Figure 1a, the material optical properties can be quantified by some parameters, such as the pitch  $\Lambda$ , the height of grating region  $H$ , the initial refractive index  $n_0$ , and the amplitude of the maximum refractive index variation  $n_A$ . Regardless of magnetic material, the relationship between refractive index  $n$  and relative dielectric constant  $\epsilon$  (i.e. permittivity) is that  $\epsilon = n^2$ .

At first, the dielectric material in a period can be uniformly discretized into  $N$  units along  $x$  axis. Without losing generality, supposing the distribution can be described by a sinusoidal function, the refractive indices at an arbitrary position  $x$  can be achieved using interpolation. At the same time, if the optical property varies along  $z$  direction, a commonly used layer-by-layer discretization along  $z$  axis can be carried out as well. As shown in Figure 1b, the grating region is sliced into  $L$  layers in  $z$  direction and  $N$  units in  $x$  direction. Then, the  $x$ -dependent relative dielectric constant at  $l$ -th layer can be expanded as Fourier series

$$\epsilon_l(x) = \sum_g \epsilon_{l,g} \exp\left(j \frac{2\pi g}{\Lambda} x\right), \tag{1}$$

where  $\epsilon_{l,g}$  is the  $g$ -th component of the Fourier expansion, which can be obtained as

$$\epsilon_{l,g} = \begin{cases} \frac{1}{N} \sum_{n=1}^N \epsilon_{ln} & g = 0 \\ -\frac{1}{j2\pi g} \left[ \sum_{n=1}^{N-1} (\epsilon_{ln} - \epsilon_{l(n+1)}) \exp\left(-j \frac{2\pi n g}{N}\right) + \epsilon_{lN} - \epsilon_{l1} \right] & g \neq 0 \end{cases}, \tag{2}$$

where  $N$  is the number of units in one period. In the same way, the reciprocal of dielectric constant can be expanded as

$$\frac{1}{\epsilon_l(x)} = \sum_g \left(\frac{1}{\epsilon}\right)_{l,g} \exp\left(j\frac{2\pi g}{\Lambda}x\right), \tag{3}$$

where

$$\left(\frac{1}{\epsilon}\right)_{l,g} = \begin{cases} \frac{1}{N} \sum_{n=1}^N \left(\frac{1}{\epsilon}\right)_{ln} & g = 0 \\ \frac{1}{-j2\pi g} \left[ \sum_{n=1}^{N-1} \left(\left(\frac{1}{\epsilon}\right)_{ln} - \left(\frac{1}{\epsilon}\right)_{l(n+1)}\right) \exp\left(-j\frac{2\pi n g}{N}\right) + \left(\frac{1}{\epsilon}\right)_{lN} - \left(\frac{1}{\epsilon}\right)_{l1} \right] & g \neq 0 \end{cases} \tag{4}$$

The electrical field expressions of region I ( $z < 0$ ) and II ( $z > H$ ) can be achieved as Equations (5) and (6) by Rayleigh expansion

$$\mathbf{E}_I = \mathbf{E}_{inc} + \sum_i [R_i \exp(-jk_{R,i} \bullet \mathbf{r})], \tag{5}$$

$$\mathbf{E}_{II} = \sum_i \{T_i \exp[-jk_{T,i} \bullet (\mathbf{r} - H\hat{z})]\}, \tag{6}$$

where  $\mathbf{E}_{inc}$  is normalized incident electrical field;  $H$  is total height of the volume grating;  $\mathbf{R}_i$  is the amplitude of  $i$ -th diffractive wave;  $\mathbf{T}_i$  is the amplitude of  $i$ -th transmit wave;  $\mathbf{r}$  is the position vector of an arbitrary point on the plane wave front; and  $\mathbf{k}_{R,i}$  is the wave vector of  $i$ -th transmit wave.

According to Floquet theorem, the following conditions should be satisfied,

$$\mathbf{k}_{\rho,i} = k_{xi}\hat{x} + k_{yi}\hat{y} + k_{z,i}\hat{z}; \quad \rho = R, T, \tag{7}$$

$$k_{xi} = k_{inc,x} - i\frac{2\pi}{\Lambda}, \tag{8}$$

$$k_{yi} = k_{inc,y}, \tag{9}$$

where  $i$  is the diffraction order,  $\mathbf{k}_{\rho,z,i}$  is  $z$  component of the wave vector.

In the grating region ( $0 < z < H$ ), the electric field  $\mathbf{E}_l$  and magnetic field  $\mathbf{H}_l$  can be written as Floquet-Fourier series as

$$\begin{aligned} \mathbf{E}_l &= \sum_i S_{l,i}(z) \exp[-j(k_{xi}x + k_{yi}y)] \\ &= \sum_i \left\{ S_{l,xi}(z)\hat{x} + S_{l,yi}(z)\hat{y} + S_{l,zi}(z)\hat{z} \right\} \exp[-j(k_{xi}x + k_{yi}y)] \end{aligned} \tag{10}$$

$$\begin{aligned} \mathbf{H}_l &= \left(\frac{\epsilon_0}{\mu_0}\right)^{\frac{1}{2}} \sum_i U_{l,i}(z) \exp[-j(k_{xi}x + k_{yi}y)] \\ &= \left(\frac{\epsilon_0}{\mu_0}\right)^{\frac{1}{2}} \sum_i \left\{ U_{l,xi}(z)\hat{x} + U_{l,yi}(z)\hat{y} + U_{l,zi}(z)\hat{z} \right\} \exp[-j(k_{xi}x + k_{yi}y)] \end{aligned} \tag{11}$$

where  $S_{l,i}(z)$  and  $U_{l,i}(z)$  are the amplitudes of the electric and magnetic fields of the  $l$ -th layer in the grating region. In the grating region, the electric and magnetic fields satisfy the Maxwell equations

$$\nabla \times \mathbf{E}_l = -jk_0\mu\mathbf{H}'_l, \tag{12}$$

$$\nabla \times \mathbf{H}'_l = jk_0\epsilon(x)\mathbf{E}_l, \tag{13}$$

where  $k_0$  is the wave number in the free space, given by  $k_0 = 2\pi/\lambda_0$  and  $\lambda_0$  is the wavelength of the incident wave in free space;  $\epsilon_0$  and  $\mu_0$  are the electric permittivity and magnetic permeability in the free space and

$$\mathbf{H}'_l = \sqrt{\mu_0/\epsilon_0}\mathbf{H}_l. \tag{14}$$

Then submit Equations (10) and (11) to Maxwell equations Equations (12) and (13), the coupled wave equation in matrix form can be achieved as

$$\frac{\partial}{\partial z} \begin{bmatrix} \mathbf{S}_{l,x} \\ \mathbf{S}_{l,y} \\ \mathbf{U}_{l,x} \\ \mathbf{U}_{l,y} \end{bmatrix} = jk_0 \begin{bmatrix} 0 & 0 & \mathbf{K}_x \mathbf{E}_{l-tp}^{-1} \mathbf{K}_y & \mathbf{I} - \mathbf{K}_y \mathbf{E}_{l-tp}^{-1} \mathbf{K}_y \\ 0 & 0 & \mathbf{K}_x \mathbf{E}_{l-tp}^{-1} \mathbf{K}_x - \mathbf{I} & -\mathbf{K}_x \mathbf{E}_{l-tp}^{-1} \mathbf{K}_y \\ \mathbf{K}_x \mathbf{K}_y & \mathbf{F}_l^{-1} - \mathbf{K}_y^2 & 0 & 0 \\ \mathbf{K}_x^2 - \mathbf{E}_{l-tp} & -\mathbf{K}_y \mathbf{K}_x & 0 & 0 \end{bmatrix} \begin{bmatrix} \mathbf{S}_{l,x} \\ \mathbf{S}_{l,y} \\ \mathbf{U}_{l,x} \\ \mathbf{U}_{l,y} \end{bmatrix}, \quad (15)$$

where  $\mathbf{K}_x$  and  $\mathbf{K}_y$  are diagonal matrices, whose diagonal elements are  $k_{xi}/k_0$  and  $k_{yi}/k_0$ ;  $\mathbf{I}$  is identity matrix with the same dimension of  $\mathbf{K}$ .  $\mathbf{E}_{l-tp}$  is the Toeplitz matrix consists of the Fourier coefficients of the relative dielectric constant at the  $l$ -th layer.  $\mathbf{F}_l$  is the Toeplitz matrix corresponding to the reciprocal of the relative dielectric constant.

Combined with the boundary conditions and enhanced transmit matrix, the distribution of the electric field  $\mathbf{R}$  and magnetic field  $\mathbf{T}$  can be achieved by solving Equation (15). Further, the reflected and transmit electric fields of s- and p- polarized light can be achieved. If the transmission is considered, the amplitudes of the electric fields and the components of Jones matrix can be expressed as

$$t_{pp} = \frac{E_{tp}}{E_{ip}}, \quad (16)$$

$$t_{sp} = \frac{E_{ts}}{E_{ip}}, \quad (17)$$

$$t_{ss} = \frac{E_{ts}}{E_{is}}, \quad (18)$$

$$t_{ps} = \frac{E_{tp}}{E_{is}}, \quad (19)$$

And then, Mueller matrix can be calculated as

$$\mathbf{M} = \mathbf{A} (\mathbf{J} \otimes \mathbf{J}^*) \mathbf{A}^{-1}. \quad (20)$$

where

$$\mathbf{A} = \begin{bmatrix} 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & -1 \\ 0 & 1 & 1 & 0 \\ 0 & j & -j & 0 \end{bmatrix}, \quad (21)$$

$$\mathbf{J} = \begin{bmatrix} t_{pp} & t_{ps} \\ t_{sp} & t_{ss} \end{bmatrix}. \quad (22)$$

## 2.2. Inverse Problem Solving

It is insufficient to successfully obtain the measurands only with the Mueller matrices measured by MME and the Mueller matrices calculated using the forward model. An inverse problem solving process needs to be applied to find out the appropriated values for the measurands, which are able to fit the measured Mueller matrices with minimum MSE. A weighted least-squares regression analysis method (Levenberg–Marquardt algorithm) [36] is performed and the weighted mean square error function is defined as

$$\chi_r^2 = \frac{1}{15Q - P} \sum_{q=1}^Q \sum_{u,v=1}^4 \left[ \frac{m_{uv,k}^{meas} - m_{uv,q}^{calc}}{\sigma(m_{uv,q})} \right]^2, \quad (23)$$

where  $q$  stands for the  $q$ -th spectral point in  $Q$  spectral points in total, subscript indices  $u$  and  $v$  represent all the Mueller matrix elements normalized to  $m_{11}$  except  $m_{11}$ ,  $P$  is the measurands number,  $m_{uv,q}$  with

superscript *meas* and *calc* denote the measured and calculated Mueller matrix elements, respectively, and  $\sigma(m_{uv,q})$  is the estimated standard deviation associated with  $m_{uv,q}$ .

In this work, the measurands could be the parameters need to be determined in refractive index distribution function, the geometric parameters of the volume grating samples, the measurement configuration conditions such as  $\theta$  and  $\varphi$ , and the thickness  $H$ .

### 3. Simulation

In order to demonstrate the generality and validity of the proposed method, a series of volume grating measurements are simulated, with different dielectric constant distribution functions. Without losing generality, three typical distribution functions such as periodic binary distribution, periodic linear distribution, as well as continuous sinusoidal distribution are selected. These functions are selected because most of the complex functions can be easily assembled by the linear combinations of these simple functions. For the convenience of comparing the effects of function type on the Mueller matrix spectra, the nominal parameters of sample setting and measurement configuration are shared. The angle of incidence  $\theta$  is fixed as  $25^\circ$ , and the angle of azimuth  $\varphi$  is set as  $20^\circ$ . The substrates are glasses. The pitch of the dielectric distribution  $\Lambda$  is fix as 800 nm. For each test case of different function type, two thickness settings  $H$  are examined—3  $\mu\text{m}$  and 5  $\mu\text{m}$ , respectively. Considering the pitch  $\Lambda$  can be usually well controlled by the process of grating fabrication, we fixed it as a constant. Since the thickness  $H$ , angle of incidence  $\theta$ , and the angle of azimuth  $\varphi$  are easily varied parameters, we need reconstruct these parameters in the fitting process.

The general procedure of the simulation can be divided into three steps. At first, according to the trues values of parameters, the theoretical Mueller matrices will be calculated. Then, random errors with signal-noise-ratio  $SNR = 10000$  will be injected into the theoretical Mueller matrices to generate a set of measured Mueller matrices. At last, by a regression started with the initial values of parameters will be carried with the objective to best fit the measured Mueller matrices. When the optimal solution is achieved, we achieved the measured values of these parameters as well as corresponding fitted Mueller matrices.

To avoid the unphysical refractive index spectra, we quoted the correlations between the refractive index and nanoparticles weight fractions  $\omega$  reported in Ref. [31], so that the appropriate refractive index curve can be selected according to the value of nanoparticles weight fraction. Then, the dielectric constant distribution function can be represented by the distribution function of the nanoparticle concentration.

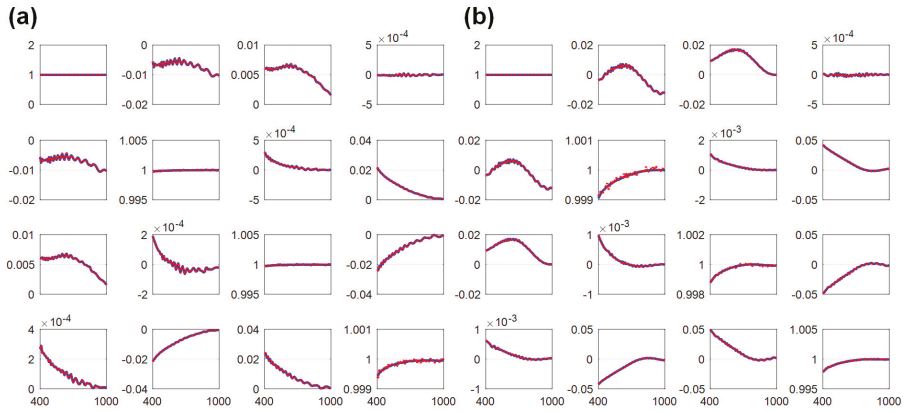
#### 3.1. Periodic Binary Distribution

In the first case, binary distribution defined by Equation (24) is used.

$$\omega = \begin{cases} \omega_a & (0 < x < f\Lambda) \\ \omega_b & (f\Lambda < x < \Lambda) \end{cases} \quad (24)$$

where  $\omega_a$  and  $\omega_b$  represent different nanoparticle concentration at different position in one pitch.  $f$  is the duty cycle of such a binary grating, which is set as 0.5 in our simulation.

In this case, the parameters  $\omega_a$  and  $\omega_b$  are measurands we need to determine. As mentioned in the previous paragraph,  $H$ ,  $\theta$ , and  $\varphi$  are the parameters needs to be verified. The measured and fitted Mueller matrix spectra are shown in Figure 2, and the corresponding parameters are listed in Table 1.



**Figure 2.** Measured Mueller matrix spectra (blue) and fitted Mueller matrix spectra (red) for the volume gratings whose dielectric constant is a periodic binary distribution, with thicknesses of (a) 3 μm and (b) 5 μm, respectively.

**Table 1.** Simulation results of case 1.

<i>H</i> (μm)		$\omega_a$	$\omega_b$	<i>H</i>	$\theta$	$\varphi$
3	Initial value	0.16	0.38	3.1	23.5	18
	Measured value	0.28	0.44	2.99	24.79	19.14
	True values	0.26	0.42	3	25	20
5	Initial value	0.16	0.38	5.1	23.5	18
	Measured value	0.29	0.45	4.98	24.92	19.98
	True values	0.26	0.42	5	25	20

As shown in Figure 2, although random noise has been injected into the “measured” Mueller matrices, the fitting is good. In Figure 2a,b, nonzero Mueller matrix elements in the off-diagonal elements are observed, clearly showing the anisotropy of the samples. As shown in Table 1, although the initial values of measurands  $\omega_a$  and  $\omega_b$  are given with large relative deviations to their true value, the measured values converge to their true values, which demonstrate the accuracy and robustness of the proposed method. Although the setting parameters *H*,  $\theta$ , and  $\varphi$  are fitted with the same algorithm, their measured values vary in a quite small range and converge to their true values as well.

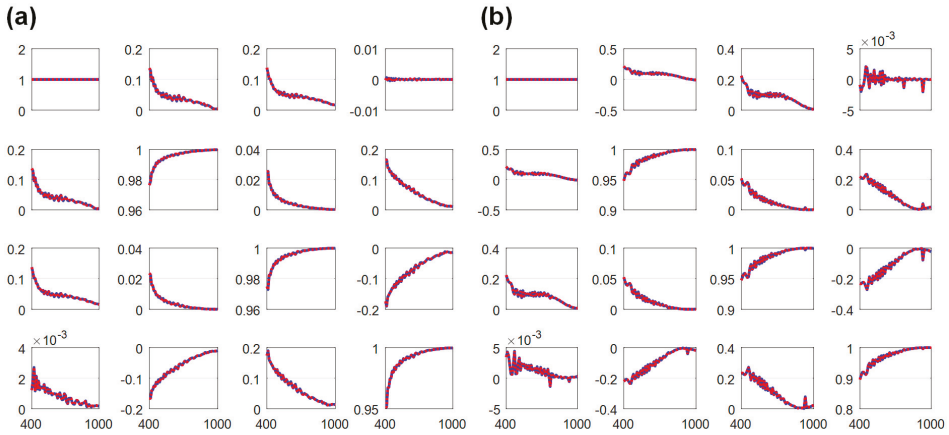
3.2. Periodic Linear Distribution

In the second test case, the fidelity of the proposed method for a periodic linear dielectric constant distribution will be examined. The distribution function is defined as

$$\omega = \omega_b + \omega_k x \quad (0 < x < \Lambda), \tag{25}$$

where  $\omega_b$  and  $\omega_k$  represent the intercept and the slope of the linear relationship in one pitch, respectively.

In this case, the parameters  $\omega_b$  and  $\omega_k$  are measurands we need to determine. Due to the same reason, *H*,  $\theta$ , and  $\varphi$  are the parameters needs to be verified. The measured and fitted Mueller matrix spectra are shown in Figure 3, and the corresponding parameters are listed in Table 2.



**Figure 3.** Measured Mueller matrix spectra (blue) and fitted Mueller matrix spectra (red) for the volume gratings whose dielectric constant is a periodic linear distribution, with thicknesses of (a) 3 μm and (b) 5 μm, respectively.

**Table 2.** Simulation results of case 2.

<i>H</i> (μm)		$\omega_b$	$\omega_k$	<i>H</i>	$\theta$	$\varphi$
3	Initial value	0.18	1.1	2.8	23.5	18
	Measured value	0.20	1.02	2.85	23.56	18.97
	True values	0.2	1	3	25	20
5	Initial value	0.18	1.1	4.8	23.5	18
	Measured value	0.21	1.00	4.98	23.75	20.01
	True values	0.2	1	5	25	20

Good-fitting results have been achieved as well, as shown in Figure 3. It is worth to note that the anisotropy encoded in the Mueller matrix spectra shown in Figure 3 are much more significant than the observations from Figure 2. This is because in the second case, the dielectric constant distributed in one pitch is not symmetric. The results list in Table 2 show the accuracy of the proposed method when the local optical constants are linear to the position.

### 3.3. Continuous Sinusoidal Distribution

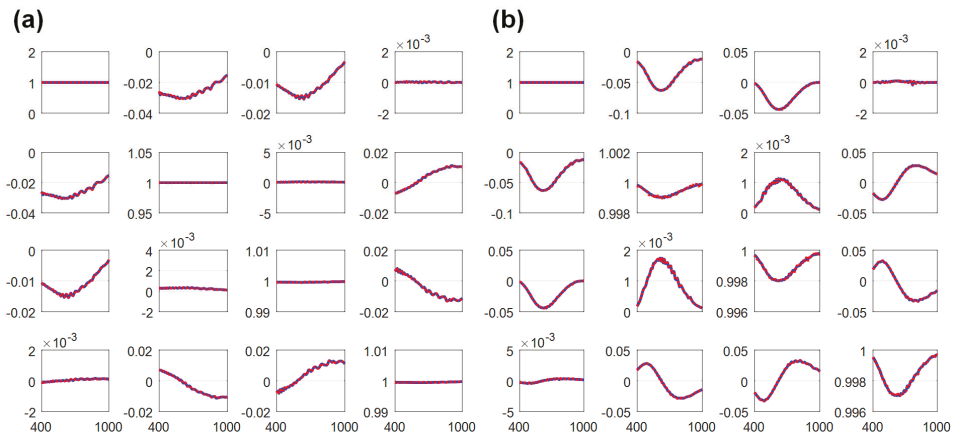
In the last case study, we studied the possibility of determining a continuous dielectric constant distribution using the proposed method, because in most of the existing volume holographic gratings, the dielectric constants are considered sinusoidally distributed. The distribution function is defined as

$$\omega = \omega_0 + \omega_A \sin(x) \quad (0 < x < \Lambda), \tag{26}$$

where  $\omega_0$  and  $\omega_A$  represent the average value and variation amplitude of nanoparticle concentration in one pitch, respectively.

In this case, the parameters  $\omega_0$  and  $\omega_A$  are measurands we need to determine. *H*,  $\theta$ , and  $\varphi$  are reconstructed as well. The measured and fitted Mueller matrix spectra are shown in Figure 4, and the corresponding parameters are listed in Table 3.





**Figure 4.** Measured Mueller matrix spectra (blue) and fitted Mueller matrix spectra (red) for the volume gratings whose dielectric constant is a continuous sinusoidal distribution, with thicknesses of (a) 3  $\mu\text{m}$  and (b) 5  $\mu\text{m}$ , respectively.

**Table 3.** Simulation results of case 3.

$H$ ( $\mu\text{m}$ )		$\omega_0$	$\omega_A$	$H$	$\theta$	$\varphi$
3	Initial value	0.3	0.1	3.1	23.5	18
	Measured value	0.36	0.15	3.00	25.00	20.00
	True values	0.36	0.15	3	25	20
5	Initial value	0.34	0.13	5.1	23.5	18
	Measured value	0.39	0.15	5.01	24.99	20.00
	True values	0.36	0.15	5	25	20

Since the nanoparticle distribution in one pitch is symmetric again, the achieved anisotropy in the spectrum is not as significant as in the second case, where the values of the off-diagonal elements shown in Figure 4 are at the similar level as in Figure 2. If further attention is paid to the results list in Table 3, a significant improvement on the accuracy can be observed comparing to the precious two cases. It may be attributed to the less singular points for the case of a continuous distribution, which decreases the numerical errors.

The above simulations demonstrate the feasibility, effectiveness, robustness of the proposed method. It is worth to point out, more complex distribution function with more parameters need to be determined will not noticeably degrade the performance of the proposed method, since the number of wavelength points in the spectrum usually is enough for obtaining the solution. If we compare the Mueller matrix spectra reported in Figures 3–5, significant differences can be distinguished, although the important parameters such as thickness, pitch, angle of incidence and azimuth, even similar interval of the refractive index changes are shared. Such high sensitivity can be used to verify the distribution function type selection, i.e., an inappropriate selection of function type is highly possible to result in the failure of regression.

## 4. Experiment

### 4.1. Volume Gratings Preparation

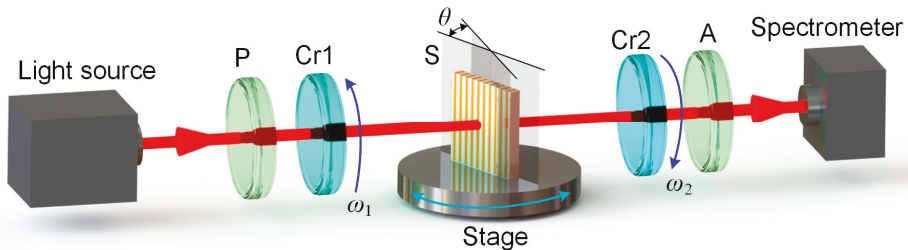
The volume gratings were prepared following the procedures introduces in our previous work [31]. The ZnS nanoparticles are about 5 nm in diameter, and were synthesized using pot reaction in oil bath followed by a purifying process [23,31]. Then the nanoparticles were dried in vacuum at room

temperature for 2 hours. Homogeneous holographic mixtures were prepared by ultrasonication at 30 °C for 50 min with the concentration of ZnS nanoparticles as 22.6 vol% using the recipe given in [31].

To form the holographic gratings, the mixtures were inject into the cavity formed by two parallel glasses, and the thickness of the sample is controlled by the silica spacers with a diameter of ~ 8  $\mu\text{m}$ . Volume gratings were recorded in the cell by two separate 442 nm He-Cd laser beams with equal intensity of 5  $\text{mW}/\text{cm}^2$  irradiating on the cells to form a sinusoidal hologram with period of 800 nm [23,31]. Different samples were prepared by varying the duration of the irradiation of 5 s, 10 s, 15 s, 20 s, 25 s, 30 s, 35 s, and 40 s respectively. Postcure in UV with intensity of 20  $\text{mW}/\text{cm}^2$  for 10 minutes was implemented at last to fix the grating structures.

#### 4.2. Experimental Setup

A Mueller matrix ellipsometer (ME-L ellipsometer, Wuhan Eoptics Technology Co., China) [37] will be used to measure the volume gratings. The MME has a dual rotating compensator configuration, whose layout in order of light propagation is shown in Figure 5, where P and A are the polarizer and analyzer, respectively; Cr1 and Cr2 refer to the first and second rotating compensators who are rotating with a fixed ratio of angular velocity; and S represents the holographic grating. With this instrument, the 16 Mueller matrix elements can be obtained in a single measurement. The spectral range of the instrument covers from 200 to 1000 nm. Since the light reflections from the glass surface and the sample are difficult to be differentiated by the detector in reflection mode, the transmission mode, i.e. straight through mode, was selected in our experiments. The samples are placed on the stage as shown in Figure 5. When the sample is rotated around the direction of light propagation, different azimuthal angle can be achieved. If the stage is rotated, the incident angle  $\theta$  can be arbitrary selected.

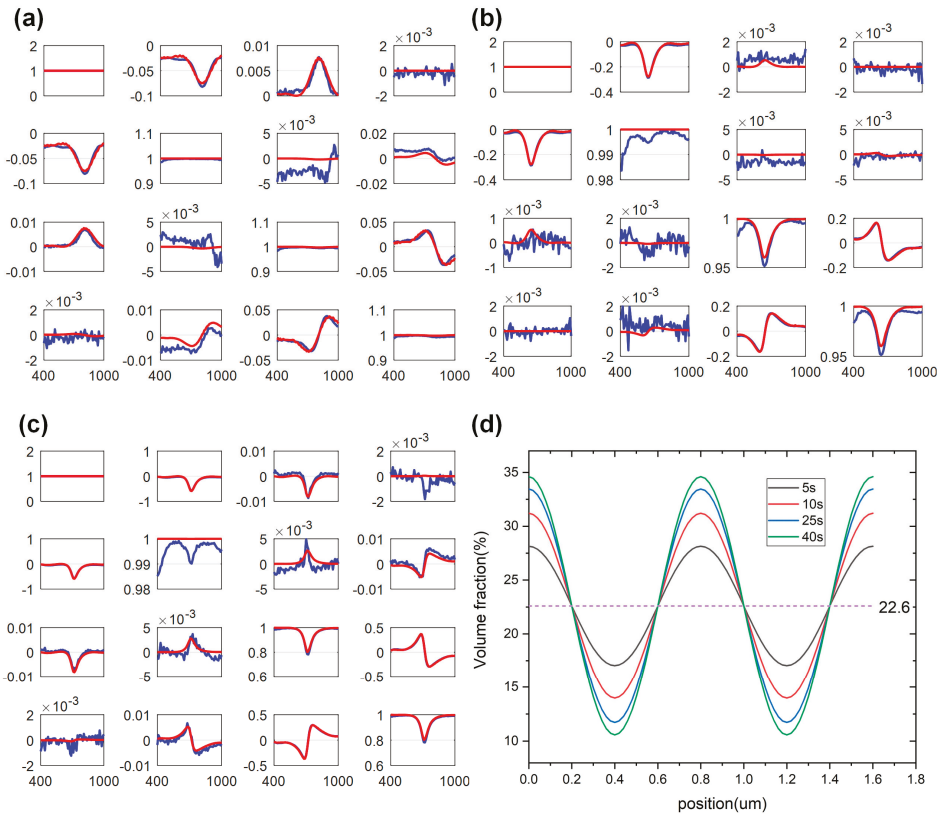


**Figure 5.** Scheme of the experimental setup based on a dual rotating-compensator Mueller matrix ellipsometer (MME) in the transmission mode.

#### 4.3. Measurement Results and Discussions

Measurement experiments are carried out on the volume holographic gratings prepared in Section 4.1. The angle of incidence  $\theta$  for the probing light was fixed at approximate 25°. Since the manual manipulation cannot guarantee the exact value of angle of incidence, we set it as a measurand with an initial value of 25°. The azimuthal angle  $\varphi$  was set as 0°, which indicates the incident plane is perpendicular to the gratings. In this case, the off-diagonal elements in the measured Mueller matrices should be zero, which can be applied to check the azimuthal angle settings in measurements. Same as angle of incidence, the azimuthal angle was set as a measurand with initial value of 0° to avoid the error introduced by manual manipulation. In analysis, the spectral range was from 400 to 1000 nm with an increment of 5 nm, and the distributed dielectric constant model based RCWA proposed in Section 2.1 was used to calculate the Mueller matrices. The distributed refractive index is assumed sinusoidal, which is defined as  $n(x)$  shown in Figure 1a. The number of retained orders in the truncated Fourier series was 12. Since the preliminary study [31] on the sample revealed that no significant refractive index variation along  $z$  direction had been observed because both the nanoparticles and the polymers are not absorbing materials, we set the layer number  $L$  to be 1 to improve the calculation

efficiency and the measurement accuracy. In this case, the amplitude of refractive index curve  $n_A$ , and the thickness of samples were measured simultaneously. Since the linear correlations between the refractive index and ZnS nanoparticles volume fraction have been achieved in our previous work [31], the distributed refractive indices can be directly converted into the volume fractions distributions. Figure 6 shows the comparison of calculated and measured Mueller matrix spectra when recording time  $t = 5$  s, 10 s, and 40 s, as well as the achieved spatial volume fraction distributions corresponding to  $t = 5$  s, 10 s, 25 s, and 40 s in the nanocomposite. More detailed measurement results including amplitude of concentration variation  $C_A$ , thickness of samples  $H$ , the angle of incidence  $\theta$ , the angle of azimuth  $\varphi$ , as well as the  $MSE$  of fitting are listed in Table 4. It is worth to point out that the pitch of the volume holographic grating is assumed 800 nm and the local nanoparticle concentration is ideally sinusoidally distributed.



**Figure 6.** Results of calculated (red) and measured (blue) Mueller matrix spectra in the transmission mode when recording time (a)  $t = 5$ s; (b)  $t = 10$ s; (c)  $t = 40$ s; and (d) measured spatial refractive index distribution curve of a volume grating at different exposure time.

**Table 4.** Measured results of volume gratings with different exposure time.

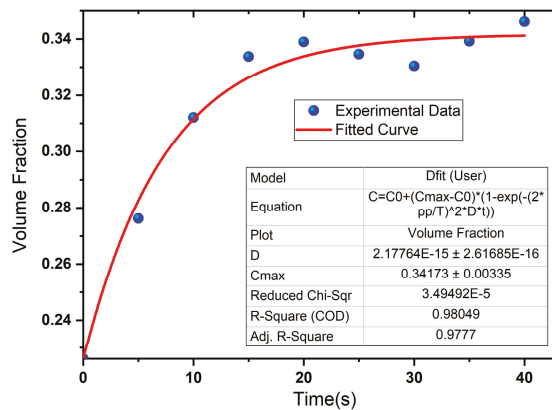
time(s)	$C_0$	$C_A$	$H$ (μm)	$\theta$ (°)	$\varphi$ (°)	MSE
5	0.226	0.055	8.2052	30.5508	−3.6431	4.8585
10		0.086	9.2286	23.8851	−0.0558	4.6035
15		0.108	8.4533	26.7676	−1.22687	6.1378
20		0.113	9.375	28.3695	−0.7261	14.0584
25		0.109	9.8429	25.6022	1.7953	21.8013
30		0.104	8.2554	26.5644	−2.9308	2.9309
35		0.113	6.5781	25.2278	−0.9119	6.004
40		0.120	8.763	26.7677	0.3869	6.7657

As shown in Figure 6a–c, the measured Mueller matrix fits well with the calculated matrix. All the Mueller matrices have similar characteristics, which indicates that the formed volume gratings are consistent. If some of the specific Mueller matrix elements such as  $m_{12}$  are selected for a further analysis, the depth of the dip become larger, which indicates that the characteristics of Bragg grating become more and more obvious. This is rational because with the increase of the recording time, the sinusoidal distribution of the refractive index is more obvious. Such a phenomenon is more intuitively shown in Figure 6d. It can be clearly observed in Figure 6d that the volume fractions in the dark regions are increasing, while the volume fractions in the bright regions are decreasing, with the process of holographic recording. Since the nominal azimuthal angle is selected as  $0^\circ$ , the Mueller matrix elements observed in the off-diagonal blocks are close to 0. If we correlate the results shown in Figure 6a–c with  $\varphi$  reported in Table 4, the relative lager oscillations observed in Figure 6a can be attributed to the azimuthal angle setting error. Since the RIM is much larger when recording time  $t$  is 40s, even though the reconstructed azimuthal angle is as small as  $0.3869^\circ$ , relative more significant anisotropy, i.e. nonzero off-diagonal elements, can be observed.

In order to further demonstrate the fidelity of the proposed method, we further investigated the time-dependent volume fraction changes during holography. With the benefit of the distribution function of the refractive index, the spatiotemporal concentration function of ZnS nanoparticles expressed in terms of position  $x$  and time  $t$ ,  $C(x, t)$  defined as Equation (24) [23] can be fitted directly.

$$C(x, t) = C_0 + (C_{max} - C_0) \sin\left(\frac{2\pi}{\lambda}x\right) \left(1 - e^{-(2\pi/\lambda)^2Dt}\right), \tag{27}$$

where  $C_{max}$  and  $C_0$  are the maximum and average nanoparticles concentration, respectively. The fitted curve can be achieved as shown in Figure 7.



**Figure 7.** Fitting result of spatiotemporal concentration of ZnS nanoparticles.

The apparent diffusion coefficient,  $D_a$ , achieved from the curve fitting is  $2.18 \times 10^{-15} \text{ m}^2 \text{ s}^{-1}$ , which agree to the previous results  $2.0 \times 10^{-15} \text{ m}^2 \text{ s}^{-1}$  we reported. Such an apparent diffusion coefficient achieved using a different model exhibits again that the apparent diffusion coefficient is 3 orders lower than the initial diffusion coefficient ( $3.4 \times 10^{-12} \text{ m}^2 \text{ s}^{-1}$ ) predicted by the Stokes–Einstein diffusion equation, which has been interpreted using the rapid increase of the mixture viscosity during polymerization [31].

## 5. Conclusions

In order to appropriately reflect the distribution of refractive indices in the volume grating so that the grating can be accurately characterized, a distributed dielectric constant-based RCWA is proposed as a forward model that can be used for ellipsometry. A set of measurement experiments is carried out on the volume gratings recorded in the composite of polymer and 5-nm ZnS nanoparticles with a different holographic recording time for demonstration. With the proposed model, parameters of the spatial refractive index distribution curve of a volume grating can be quantified, which also enables the quantitative determination of the spatiotemporal concentration function. Good agreement of the experimental results to the values we previously reported demonstrates the fidelity of the proposed method.

**Author Contributions:** Conceptualization, H.J.; methodology, H.J.; software, Z.M.; validation, H.J., and Z.M.; formal analysis, H.J., and Z.M.; writing—original draft preparation, H.J.; writing—review and editing, H.G., X.C., and S.L.; supervision, H.J. and S.L.; project administration, H.J. and S.L.; funding acquisition, H.J. and S.L.

**Funding:** This research was funded by the National Natural Science Foundation of China (Grant Nos. 51575214, 51525502, 51727809, and 51805193), the National Key Research and Development Plan (Grant No. 2017YFF0204705); the National Science and Technology Major Project of China (Grant No. 2017ZX02101006-004), and the Natural Science Foundation of Hubei Province of China (Grant Nos. 2018CFA057 and 2018CFB559).

**Conflicts of Interest:** The authors declare no conflicts of interest.

## References

1. Heifetz, A.; Shen, J.T.; Lee, J.K.; Tripathi, R.; Shahriar, M.S. Translation-invariant object recognition system using an optical correlator and a super-parallel holographic RAM. *Opt. Eng.* **2006**, *45*, 025201. [[CrossRef](#)]
2. Fernandez, R.; Gallego, S.; Marquez, A.; Frances, J.; Navarro, V.; Pascual, I. Diffractive lenses recorded in absorbent photopolymers. *Opt. Express* **2016**, *24*, 1559–1572. [[CrossRef](#)] [[PubMed](#)]
3. Iazikov, D.; Greiner, C.M.; Mossberg, T.W. Integrated holographic filters for flat passband optical multiplexers. *Opt. Express* **2006**, *14*, 3497–3502. [[CrossRef](#)] [[PubMed](#)]
4. Aksenova, K.A.; Angervaks, A.E.; Shcheulin, A.S.; Ryskin, A.I. Holographic characteristics of calcium fluoride crystals in the IR region. *J. Opt. Technol.* **2015**, *82*, 760. [[CrossRef](#)]
5. Mikutis, M.; Kudrius, T.; Šlekys, G.; Paipulas, D.; Juodkazis, S. High 90% efficiency Bragg gratings formed in fused silica by femtosecond Gauss-Bessel laser beams. *Opt. Mater. Express* **2013**, *3*, 1862. [[CrossRef](#)]
6. Zhang, Y.J.; Zhang, G.D.; Chen, C.L.; Stoian, R.; Cheng, G.H. Transmission volume phase holographic gratings in photo-thermo-refractive glass written with femtosecond laser Bessel beams. *Opt. Mater. Express* **2016**, *6*, 3491. [[CrossRef](#)]
7. Lumeau, J.; Glebov, L. Effect of the refractive index change kinetics of photosensitive materials on the diffraction efficiency of reflecting Bragg gratings. *Appl. Opt.* **2013**, *52*, 3993–3997. [[CrossRef](#)]
8. Dmitry, K.; Sergei, I.; Victoria, K.; Martti, S.; Yuri, S.; Nikolay, N. Thermal stability of volume bragg gratings in chloride photo-thermo-refractive glass after femtosecond laser bleaching. *Opt. Lett.* **2018**, *43*, 1083.
9. Chen, P.; He, D.; Jin, Y.; Chen, J.; Zhao, J.; Xu, J. Method for precise evaluation of refractive index modulation amplitude inside the volume bragg grating recorded in photo-thermo-refractive glass. *Opt. Express* **2018**, *26*, 157. [[CrossRef](#)]
10. Guo, J.; Gleeson, M.R.; Sheridan, J.T. A review of the optimisation of photopolymer materials for holographic data storage. *Phys. Res. Int.* **2012**, *2012*, 803439. [[CrossRef](#)]
11. Sabel, T.; Zschocher, M. Transition of refractive index contrast in course of grating growth. *Sci. Rep.* **2013**, *3*, 2552. [[CrossRef](#)]

12. Ji, R.; Fu, S.; Zhang, X.; Han, X.; Liu, S.; Wang, X.; Liu, Y. Fluorescent holographic fringes with a surface relief structure based on merocyanine aggregation driven by blue-violet laser. *Sci. Rep.* **2018**, *8*, 3818. [[CrossRef](#)] [[PubMed](#)]
13. De Sio, L.; Lloyd, P.F.; Tabiryan, N.V.; Bunning, T.J. Hidden gratings in holographic liquid crystal polymer-dispersed liquid crystal films. *ACS Appl. Mater. Interfaces* **2018**, *10*, 13107–13112. [[CrossRef](#)] [[PubMed](#)]
14. Ni, M.; Peng, H.; Xie, X. Structure regulation and performance of holographic polymer dispersed liquid crystals. *Acta Polym. Sin.* **2017**, *48*, 1557–1573.
15. Vaia, R.A.; Dennis, C.L.; Natarajan, L.V.; Tondiglia, V.P.; Tomlin, D.W.; Bunning, T.J. One-step, micrometer-scale organization of nano- and mesoparticles using holographic photopolymerization: A generic technique. *Adv. Mater.* **2001**, *13*, 1570–1574. [[CrossRef](#)]
16. Garnweitner, G.; Goldenberg, L.M.; Sakhno, O.V.; Antonietti, M.; Niederberger, M.; Stumpe, J. Large-scale synthesis of organophilic zirconia nanoparticles and their application in organic-inorganic nanocomposites for efficient volume holography. *Small* **2007**, *3*, 1626–1632. [[CrossRef](#)]
17. Sanchez, C.; Escuti, M.J.; van Heesch, C.C.; Bastiaansen, W.M.; Broer, D.J.; Loos, J.; Nussbaumer, R. TiO<sub>2</sub> nanoparticle-photopolymer holographic recording. *Adv. Funct. Mater.* **2005**, *15*, 1623–1629. [[CrossRef](#)]
18. Suzuki, N.; Tomita, Y.; Kojima, T. Holographic recording in TiO<sub>2</sub> nanoparticle-dispersed methacrylate photopolymer films. *Appl. Phys. Lett.* **2002**, *81*, 4121–4123. [[CrossRef](#)]
19. Smirnova, T.N.; Sakhno, O.V.; Yezhov, P.V.; Kokhtych, L.M.; Goldenberg, L.M.; Stumpe, J. Amplified spontaneous emission in polymer-CdSe/ZnS-nanocrystal DFB structures produced by the holographic method. *Nanotechnology* **2009**, *20*, 245707. [[CrossRef](#)] [[PubMed](#)]
20. Ostrowski, A.M.; Naydenova, L.; Toal, V. Light-induced redistribution of Si-MFI zeolite nanoparticles in acrylamide-based photopolymer holographic gratings. *J. Opt. A Pure Appl. Opt.* **2009**, *11*, 034004. [[CrossRef](#)]
21. Berberova, N.; Daskalova, D.; Strijkova, V.; Kostadinova, D.; Nazarova, D.; Nedelchev, L.; Stoykova, E.; Marinova, V.; Chi, C.; Lin, S. Polarization holographic recording in thin films of pure azopolymer and azopolymer based hybrid materials. *Opt. Mater.* **2017**, *64*, 212–216. [[CrossRef](#)]
22. Lü, C.; Cheng, Y.; Liu, Y.; Liu, F.; Yang, B. A facile route to ZnS–polymer nanocomposite optical materials with high nanophase content via  $\gamma$ -ray irradiation initiated bulk polymerization. *Adv. Mater.* **2006**, *18*, 1188–1192. [[CrossRef](#)]
23. Ni, M.; Peng, H.; Liao, Y.; Yang, Z.; Xue, Z.; Xie, X. 3D image storage in photopolymer/ZnS nanocomposites tailored by “photoinhibitor”. *Macromolecules* **2015**, *48*, 2958–2966. [[CrossRef](#)]
24. Peng, H.; Bi, S.; Ni, M.; Xie, X.; Liao, Y.; Zhou, X.; Xue, Z.; Zhu, J.; Wei, Y.; Bowman, C.N.; et al. Monochromatic visible light “photoinhibitor”: Janus-faced initiation and inhibition for storage of colored 3D images. *J. Am. Chem. Soc.* **2014**, *136*, 8855–8858. [[CrossRef](#)] [[PubMed](#)]
25. Sabel, T.; Zschocher, M. Imaging of Volume Phase Gratings in a Photosensitive Polymer, Recorded in Transmission and Reflection Geometry. *Appl. Sci.* **2014**, *4*, 19–27. [[CrossRef](#)]
26. Juhl, A.T.; Busbee, J.D.; Koval, J.J.; Natarajan, L.V.; Tondiglia, V.P.; Vaia, R.A.; Bunning, T.J.; Braun, P.V. Holographically directed assembly of polymer nanocomposites. *ACS Nano* **2010**, *4*, 5953–5961. [[CrossRef](#)]
27. Óscar, M.; Calvo, M.L.; Rodrigo, J.A.; Cheben, P.; Monte, F.D. Diffusion study in tailored gratings recorded in photopolymer glass with high refractive index species. *Appl. Phys. Lett.* **2007**, *91*, 141115.
28. Klepp, J.; Tomita, Y.; Pruner, C.; Kohlbrecher, J.; Fally, M. Three-port beam splitter for cold neutrons using holographic nanoparticle-polymer composite diffraction gratings. *Appl. Phys. Lett.* **2012**, *101*, 154104. [[CrossRef](#)]
29. Goldenberg, L.M.; Sakhno, O.V.; Smirnova, T.N.; Helliwell, P.; Chechik, V.; Stumpe, J. Holographic composites with gold nanoparticles: Nanoparticles promote polymer segregation. *Chem. Mater.* **2008**, *20*, 4619–4627. [[CrossRef](#)]
30. Butcher, H.L.; MacLachlan, D.G.; Lee, D.; Brownsword, R.A.; Thomson, R.R.; Weidmann, D. Mid-infrared volume diffraction gratings in IG2 chalcogenide glass: Fabrication, characterization, and theoretical verification. *Proc. SPIE* **2018**, *10528*, 105280R.
31. Jiang, H.; Peng, H.; Chen, G.; Gu, H.; Chen, X.; Liao, Y.; Liu, S.; Xie, X. Nondestructive investigation on the nanocomposite ordering upon holography using Mueller matrix ellipsometry. *Eur. Polym. J.* **2019**, *110*, 123–129. [[CrossRef](#)]

32. Mokhov, S.; Ott, D.; Smirnov, V.; Divliansky, I.; Zeldovich, B.; Glebov, L. Moiré apodized reflective volume Bragg grating. *Opt. Eng.* **2018**, *57*, 037106. [[CrossRef](#)]
33. Li, H.; Qi, Y.; Guo, C.; Malallah, R.; Sheridan, J.T. Holographic characterization of diffraction grating modulation in photopolymers. *Appl. Opt.* **2018**, *57*, E107–E117. [[CrossRef](#)] [[PubMed](#)]
34. Li, C.; Cao, L.; He, Q.; Jin, G. Holographic kinetics for mixed volume gratings in gold nanoparticles doped photopolymer. *Opt. Express* **2018**, *22*, 5017–5028. [[CrossRef](#)] [[PubMed](#)]
35. Moharam, M.G.; Gaylord, T.K. Rigorous coupled-wave analysis of planar-grating diffraction. *J. Opt. Soc. Am.* **1981**, *71*, 811–818. [[CrossRef](#)]
36. Press, W.H.; Teukolsky, S.A.; Vetterling, W.T.; Flannery, B.P. *Numerical Recipes: The Art of Scientific Computing*; Cambridge University Press: Cambridge, UK, 2007.
37. Liu, S.; Chen, X.; Zhang, C. Development of a broadband Mueller matrix ellipsometer as a powerful tool for nanostructure metrology. *Thin Solid Films* **2015**, *584*, 176–185. [[CrossRef](#)]



© 2019 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).

Article

# A Geometric Error Measurement System for Linear Guideway Assembly and Calibration

Tung-Hsien Hsieh <sup>1</sup>, Po-Yu Chen <sup>2,\*</sup>, Wen-Yuh Jywe <sup>1</sup>, Guan-Wu Chen <sup>1</sup> and Ming-Shi Wang <sup>2</sup>

<sup>1</sup> Smart Machine and Intelligent Manufacturing Research Center, National Formosa University, Yunlin 632, Taiwan; p98951078@gmail.com (T.-H.H.); jywe@nfu.edu.tw (W.-Y.J.); elroy\_tw@hotmail.com (G.-W.C.)

<sup>2</sup> Department of Engineering Science, National Cheng Kung University, Tainan 701, Taiwan; mswang@mail.ncku.edu.tw

\* Correspondence: jamie.pychen@gmail.com; Tel.: +886-5-631-3464; Fax: +886-5-631-5401

Received: 13 December 2018; Accepted: 1 February 2019; Published: 10 February 2019



**Abstract:** Geometric errors, such as straightness, perpendicularity, and parallelism errors are determinant factors of both the accuracy and service life of a linear guideway. In this study, a multipurpose geometric error measurement system was mainly composed of a laser source and an in-lab-developed optical module is proposed. Two adjustment methods were used for the in-lab-developed optical module to calibrate the altitude angle of the pentaprism: The first one is designed for ease of operation based on Michelson principle using a laser interferometer as the light receiver, and the second is aimed at high calibration repeatability based on the autocollimator principle using the quadrant detector (QD) to replace the light receiver. The result shows that the residual errors of the horizontal straightness and the vertical straightness are within  $\pm 1.3 \mu\text{m}$  and  $\pm 5.3 \mu\text{m}$ , respectively, when referred to as the commercial laser interferometer. Additionally, the residual errors of perpendicularity and parallelism are within  $\pm 1.2 \mu\text{m}$  and  $\pm 0.1 \mu\text{m}$ , respectively, when referred to as the granite reference blocks

**Keywords:** linear guideway; geometric errors; pentaprism; machine tool

## 1. Introduction

The development and manufacture of high-accuracy, large-scale, or long-travelling platforms have gradually become the mainstream trend of the machine tool industry. For the case of machine tools with long-travelling linear guideways, the geometric errors originated from assembly processes can obviously result in damaged transmission components and worn casting parts, and consequently, reduce the accuracy and service life.

The traditional measurement tools used for geometric error inspection of a linear guideway are granite reference blocks (square, tri-square, straight edge, parallels, etc.) plus a dial indicator, which are heavy, easily scratched, single-purpose, and low-accuracy compared to laser instruments, especially for the case of long-travelling measurements. A laser interferometer is the most reliable instrument for error measurement of displacement, straightness, yaw, and pitch; however, it can only measure one error at one time with a time-consuming setup process.

In recent years, numerous geometric error measurement systems based on multi-beam interference with flat mirrors, or multi-beam laser with position sensing detector (PSD) and/or quadrant detector (QD) have been developed to simultaneously obtain multiple errors during one test for both off-machine and on-machine measurements.

The typical solution based on multi-beam interference adopts a laser measuring system with a dual flat-mirror. Sommargren [1] proposed a dual measurement interferometer that is able to concurrently measure both linear and angular displacements for wafer stage metrology. Nakamura et al. [2] utilized



four interferometers and one corner cube for measuring three dimensional (3-D) coordinates of a microscopic scanning stage. Lee et al. [3] proposed a method with two laser interferometers to measure wafer planar positioning errors and a laser interferometer to measure angular errors.

Menq et al. [4] designed an interferometer system for the measurement of  $x$ - $y$  motion errors of wafer stages. In this system, incident beams are tilted with certain angle using a wedge prism, and an additional corner cube and a plane mirror are placed off the stage to help return the measuring beam. Zhang et al. [5] developed a laser interferometric system for real-time measurements of six-axis motions of a magnetic levitation stage by utilizing two corner cubes, six plane mirrors, and six polarization beam splitters (PBS). Jywe et al. [6] utilized two laser interferometers and two plane mirrors to measure three degrees of freedom (3-DOF) of the dual-axis nano-positioning stage, and successfully obtained the nano-scale positioning accuracy of linear and rotation motions.

For the geometric error measurement systems based on multi-beam with PSD and/or QD, Ni et al. [7,8] proposed a laser optical multi-degree-of-freedom measurement (MDFM) system, which utilized two sets of PSDs to calibrate five degrees of freedom (5-DOF) of coordinate measuring machines. Shimizu et al. [9] proposed an optical measurement system using one laser interferometer and three QDs for the measurement of 6-DOF linear motion errors of a machine tool table. Chou et al. [10] also developed an MDFM system based on two dimensional (2-D) charge-coupled device (CCD) cameras instead of QD in order to correct the geometric errors of coordinate measuring machines.

For the research on multi-beam measurement devices integrated with PSD or QD, Fan et al. [11] proposed a multi-function error measuring system using three sets of quadrant detectors for inspecting the 5-DOF of CNC machine tools. Fan et al. [12] developed a six degrees of freedom (6-DOF) measuring system by using four laser Doppler scales, two L-shaped plane mirrors, one long right-angle mirror, and two QDs for the measurement of six motion errors of a wafer stage. Jywe et al. [13] proposed a simple and low-cost technique using a laser diode, a one dimensional (1-D) grating, two PSDs, and a reference rotary table with good repeatability to measure the four degrees of freedom (4-DOF) errors of a rotary table for a 360° full circle.

You et al. [14] proposed a straightness error measurement device based on common-path compensation for the elimination of laser beam drift by utilizing an optical module consisting of a PSD and a QD. Jywe et al. [15] developed a novel optical calibration system using two sets of QDs and one ball lens module to obtain the total error associated with simultaneous multi-axis movements of CNC machine tools. Huang et al. [16] utilized a method for measuring 5-DOF errors of a moving stage with a monolithic prism and three PSDs. Yan et al. [17] proposed a laser straightness interferometer system that is able to compensate the rotational error and simultaneously detect the 6-DOF error of a linear stage by using one PSD, two QDs, one Wollaston prism, one corner cube, and other optical components.

In view of the literature, the repeatability of geometric error measurements using a pentaprism could be further improved by introducing a calibration mechanism to detect the altitude angle between the incident beam and the pentaprism. Therefore, two methods are proposed in this study for calibrating the above altitude angle: The first method is designed for ease of use based on Michelson principle using a laser interferometer as the light receiver, and the second one is aimed for high calibration repeatability based on autocollimator principle using QD to replace the light receiver. The literature mentioned above (e.g., Ni et al. [7], Fan et al. [11], Huang et al. [16], and Yan et al. [17]) proposed an MDFM system for single axis motion, respectively. However, these systems were composed with large numbers of optical lenses or optical components, which were difficult to setup and use in fabrication site of the factory. Compared with the reduced optical components compact module designed for the proposed system, which is capable of applying to the large linear guideway assembly process of various double-column, three-axis, and five-axis machine tools, the measurement range was up to 6000 mm and the horizontal and vertical straightness errors can be obtained in one measurement. Additionally, there is no need to reset the system if the laser source is sheltered during the measurement process.

## 2. Measurement Principles

In this study, a commercial laser interferometer (5529A, Agilent Tech., Santa Clara, CA, USA), which is commonly utilized as a measuring instrument by machine tool industry, is used as a dual-frequency laser source. The PSDs used in this study are dual lateral detectors (DLS-10, OSI Optoelectronics, Hawthorne, CA, USA), which are utilized for detecting the 2-D position deviation of the incident beam spots on the surface of each PSD. The calibration experiment result shows that the repeatability of the dual lateral detector is up to 1.0  $\mu\text{m}$  with a distance of one meter.

### 2.1. Pentaprism Module

In addition to dual laser interferometer and PSD, a pentaprism module with a lifting mechanism was also developed for the geometric error measurement in this study. The pentaprism (BPP-12.7, Newport, Irvine, CA, USA), a five-sided reflecting prism, was utilized to deflect the incident laser beam by exactly  $90^\circ$ , regardless of the incident angle. For keeping the maximum intensity of the reflected beam, the light receiver of the laser interferometer and a plane mirror were utilized to ensure the altitude angle between the pentaprism module and the laser source.

The setup process of the pentaprism module is described as follows: The plane mirror placed above the pentaprism is first lifted up to receive the laser beam by adjusting the lifting mechanism of the pentaprism module for refracting the incident laser beam back to the optical field receiver of the laser source. The altitude angle between the laser source and the pentaprism module is ensured by keeping the maximum optical intensity of the backed beam. After ensuring the altitude angle with the maximum reflected beam intensity, the pentaprism is then lifted up to receive the laser beam by again adjusting the lifting platform of the pentaprism module.

The optical arrangement of the beam module and the pentaprism module is shown in Figure 1.

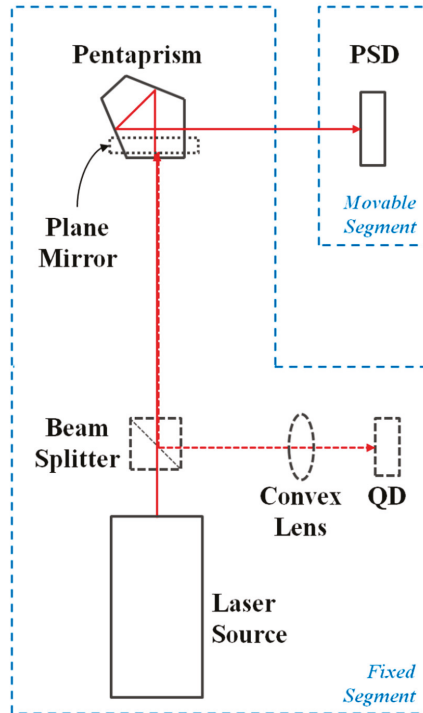
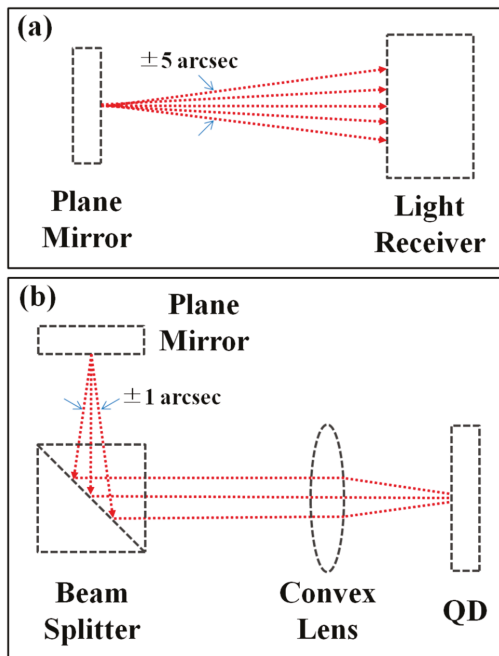


Figure 1. System optical arrangement.

For the commercial laser interferometer, the coincidence between the reflected and the incident beams could be achieved by perpendicularly adjusting the altitude angle of the incident beam to the plane mirror by its light receiver according to the Michelson principle. However, with the feedback of calibration results, it was confirmed that the tolerance error of the “100% coincidence” measured by the commercial laser interferometer was about  $\pm 5$  arcsec, which could result in poor measuring repeatability, especially for the case of long-range measurements, as shown in Figure 2a. Therefore, an altitude angle adjustment module composed of a beam splitter, a convex lens, and a QD was proposed based on the autocollimator principle [18] for obtaining higher calibration repeatability, as shown in Figures 2b and 3. With the proposed module, the reflected beam was deflected by the beam splitter and passed through a convex lens onto QD, and the tolerance error of the “100% coincidence” could be reduced to  $\pm 1$  arcsec.

Testing results showed that the perpendicularity repeatability ( $3\sigma$ ) with a measuring distance of 2 m was greatly reduced from 5.7 arcsec ( $27.63 \mu\text{m}/\text{m}$ ) to 1.6 arcsec ( $7.56 \mu\text{m}/\text{m}$ ), proving the feasibility of the proposed altitude angle calibration module. Practically, the altitude angle between the laser source and the pentaprism could be calibrated using the first method based on the Michelson principle mentioned above for laboratories and factories owing to the Agilent 5529A interferometer, and the second method, based on autocollimator principle mentioned above, could be used for high calibration repeatability applications.



**Figure 2.** The optical path of altitude angle calibration. (a) The tolerance error of the commercial laser interferometer and (b) the proposed module based on the autocollimator principle could reduce the tolerance error of the commercial laser interferometer.

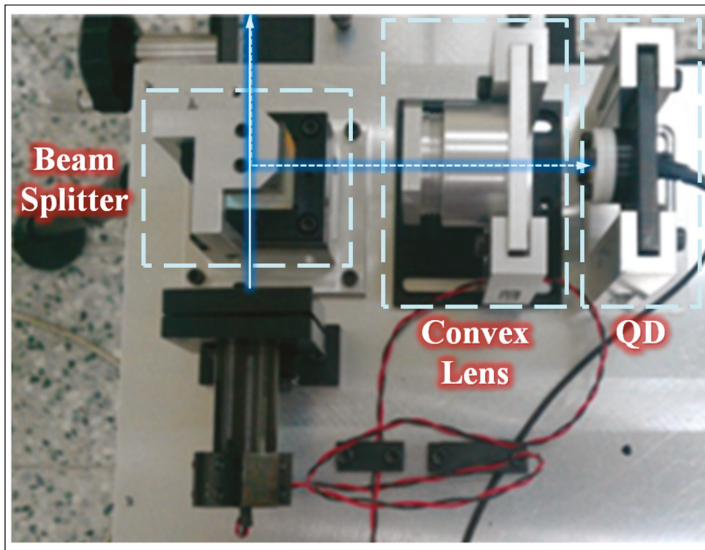


Figure 3. The in-lab-developed module for altitude angle calibration.

### 2.2. Straightness Measurement Module

The proposed straightness measurement module is composed of a PSD (Dual lateral), a signal processor, a data acquisition card (USB-6210, National Instruments Co., Taipei, Taiwan), and a laser source, as shown in Figure 4. The straightness error is the extent of the actual path deviating from a straight line, which can be divided into horizontal and vertical straightness errors [19].

The linear least-squares method is used for analyzing the geometric errors in this study, which is to find the equation of the straight line that minimizes the sum of the squares of deviations calculated from PSD-measured points:

$$y = ax + b \quad (1)$$

The straightness error  $\varepsilon_l$  can be expressed as:

$$\varepsilon_l = (E_i)_{max} - (E_i)_{min} \quad (2)$$

### 2.3. Perpendicularity Measurement Module

Figure 5 shows the proposed perpendicularity measurement system, which was mainly composed of the straightness module and the pentaprism module, as mentioned in previous sections. The perpendicularity measurement process is described as follows:

First, the deviation angle of the incident beam  $\theta_1$  (as shown in Figure 6) was calculated via the least-squares analysis of the measured points detected by the PSD set upon the linear guideway A.

Secondly, the PSD was moved and remounted on the linear guideway B; thereupon, the pentaprism module was set at the cross-point of linear guideway A and the extended line of linear guideway B, allowing the incident beam be refracted by  $90^\circ$  and projected on the re-set PSD.

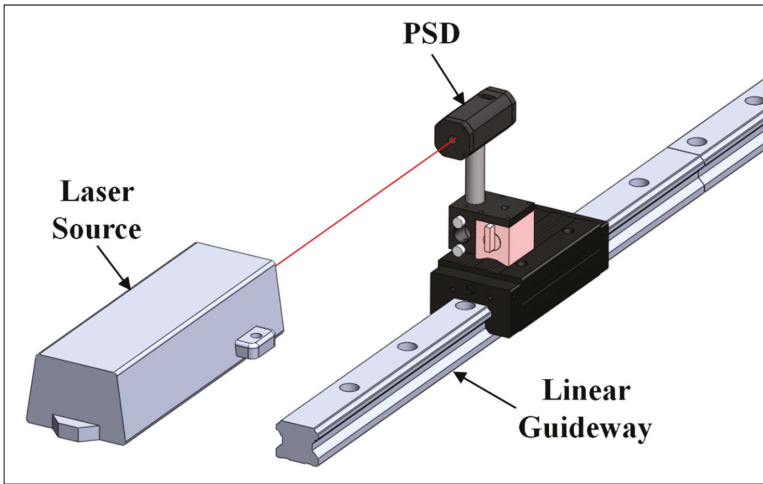


Figure 4. Schematic diagram of the straightness module.

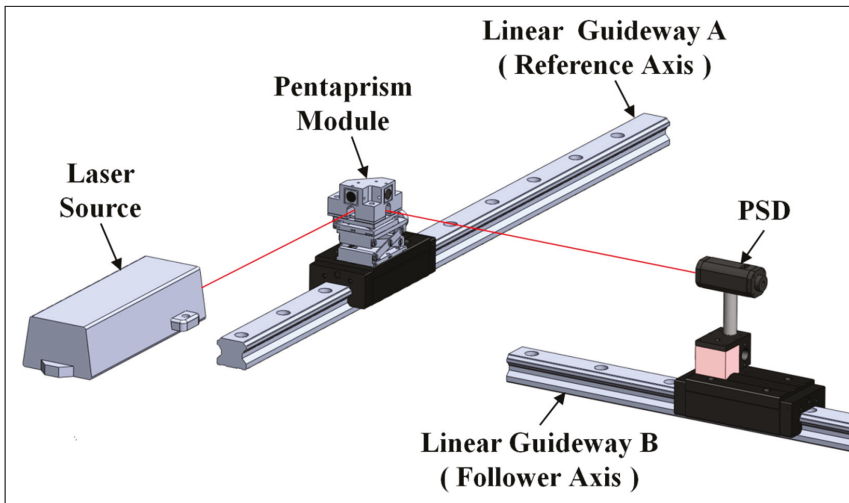


Figure 5. Schematic diagram of the perpendicularity module.

Finally, the deviation angle of the refracted beam  $\theta_2$  was also calculated via the least-squares analysis of the PSD-measured points.

The perpendicularity error  $\varepsilon_{per}$  can be obtained by the arctangent relation of  $\theta_1$ ,  $\theta_2$ , measured distance  $D$ , and the pentaprism error  $\alpha$  ( $\pm 30$  arcsec), as shown in Figure 6:

$$\varepsilon_{per} = D \tan^{-1}[(\theta_2 - \theta_1) - \alpha] \quad (3)$$

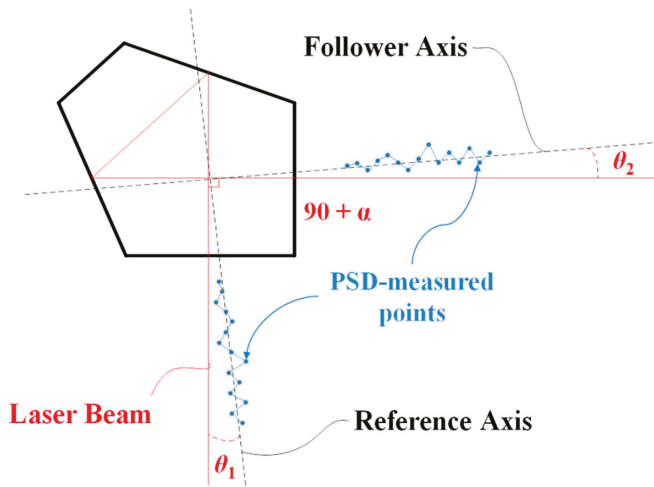


Figure 6. Beam path of perpendicularity measurement.

2.4. Parallelism Measurement Module

The measurement and setup principle of the parallelism is almost the same as that of perpendicularity described in the previous section; the only difference is the beam path utilized in the measurement, as shown in Figure 7.

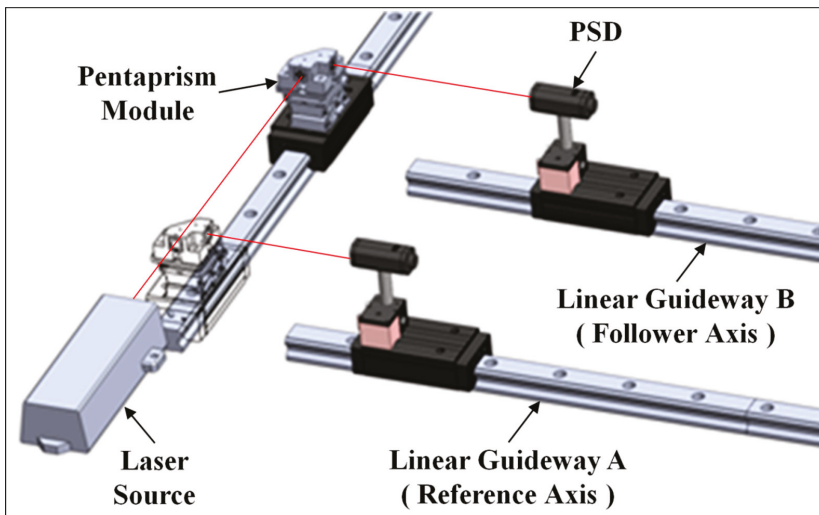


Figure 7. Schematic diagram of the parallelism module.

The deviation angles of the refracted beam  $\theta_{21}$  and  $\theta_{22}$  were also calculated via the least-squares analysis of the measured points detected by the PSD set upon the linear guideway, respectively. The perpendicularity error  $\epsilon_{para}$  can be obtained by the arctangent relation of  $\theta_{21}$ ,  $\theta_{22}$ , and the measured distance  $D$ , as shown in Figure 8:

$$\epsilon_{para} = D \tan^{-1}[(\theta_{22} - \theta_{21})] \tag{4}$$

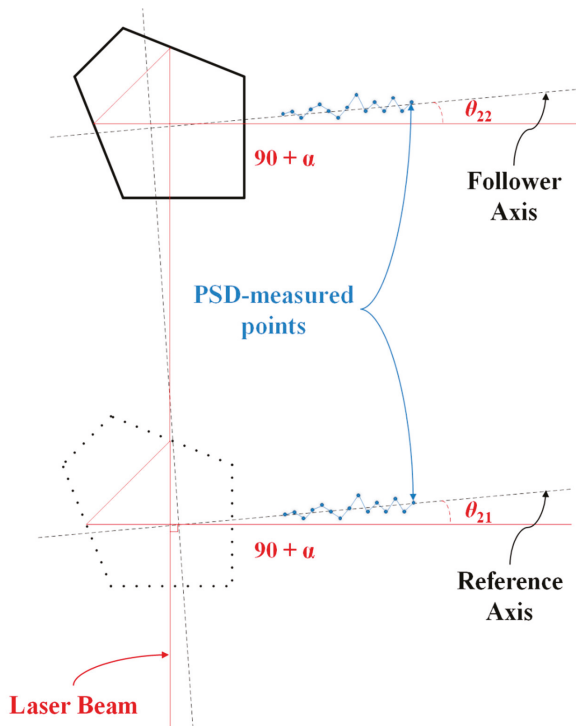


Figure 8. Beam path of parallelism measurement.

### 3. Uncertainty Analysis

The proposed system comprises one laser source, one pentaprism module, and one PSD module; therefore, any setup errors among any of these components absolutely influence the system uncertainty. The details are explained below.

#### 3.1. Laser Source Setup

During the measuring process, the laser source setup could be affected by the unstable foundation, machine operation, crane movement, staff walking, etc. The longer the distance is, the worse the measurement results; therefore, there is a need of an error analysis toward the laser source.

As depicted in Figure 9, suppose the distance between the laser source and the fixed end is  $D_{lp}$ , the error angle of laser source at vertical direction ( $\theta_{lv}$ ) will generate an error  $\delta_{lv}$ :

$$\delta_l = D_{lp} \tan \theta_l \tag{5}$$

Table 1 lists the setup errors in the vertical and horizontal directions of the laser source due to surface ground vibration with various measurement distances according to Equation (5).

Table 1. Error analysis of the laser source at vertical or horizontal directions.

$D_{lp}$ (mm)	$\theta_{lv}$ or $\theta_{lh}$ (arcsec)	$\delta_{lv}$ or $\delta_{lh}$ ( $\mu\text{m}$ )
1000	0.1	0.5
1000	0.5	2.4
1000	1	4.8

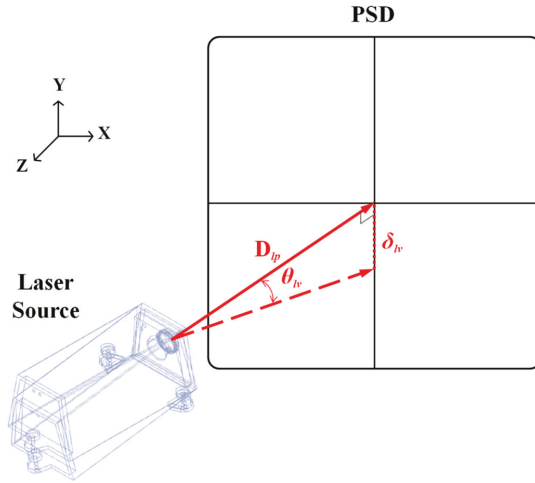


Figure 9. Vertical setup error of the laser source.

3.2. PSD Setup

There is usually an angle error due to the setup of PSD, and it will result in a cosine error in the vertical or horizontal direction, as shown in Figure 10. Assume  $F_v$  is the error measured in the vertical direction from the center of the PSD detection area, and  $F'_v$  is the error originated from PSD setup when an angle deviation  $\theta_{pv}$  exists, we get:

$$\delta_l = F'_v - F_v = F(\sec \theta_{pv} - 1) \tag{6}$$

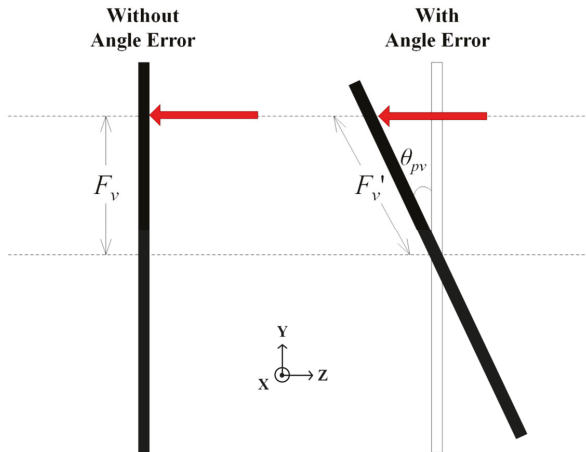


Figure 10. Position sensing detector (PSD) setup error.

Table 2 lists the setup errors in the vertical and horizontal directions of the laser PSD due to surface ground vibration with various measurement distances according to Equation (6).



Table 2. Error analysis of the PSD setup.

$F_v$ or $F_h$ ( $\mu\text{m}$ )	$\theta_{pv}$ or $\theta_{ph}$ (degree)	$\delta_{pv}$ or $\delta_{ph}$ ( $\mu\text{m}$ )
100	0.5	0.004
100	1	0.015
100	5	0.382

### 3.3. Pentaprism Setup

The setup error originated from the pentaprism module used in the perpendicularity and parallelism measurement was analyzed, as depicted in Figure 11. Suppose the distance between the laser source and the fixed end is  $D_{pt}$ , the error angle of the laser source in the vertical direction ( $\theta_{tv}$ ) will generate an error  $\delta_{tv}$ :

$$\delta_{tv} = D_{pt} \tan \theta_{tv} \tag{7}$$

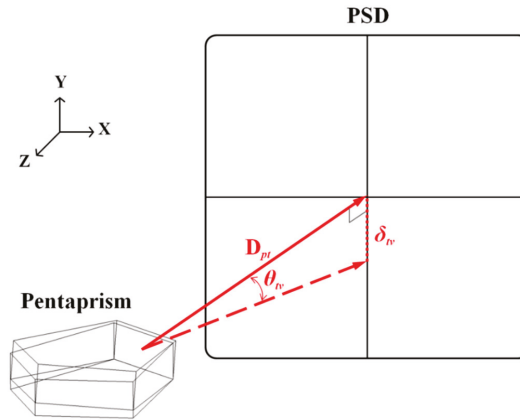


Figure 11. Vertical setup error of the pentaprism.

The error analysis table of pentaprism in the vertical or horizontal direction is the same as Table 1.

The measurement uncertainty of the proposed system can be obtained via partial differentiation of the linear displacements,  $\delta_h$  or  $\delta_v$ , along the horizontal or vertical direction ( $x$ -axis or  $y$ -axis), which were, respectively, measured by the position deviation of the light spot along the  $x$ -axis or  $y$ -axis of PSD. The linear displacements ( $\delta_h$  and  $\delta_v$ ) can be expressed as:

$$\delta_h = f_h(D_{lp}, \theta_{ph}, D_{pt}) \tag{8}$$

$$\delta_v = f_v(D_{lp}, \theta_{pv}, D_{pt}) \tag{9}$$

Take the partial derivatives of  $\delta$  in Equations (8) and (9) with respect to the corresponding uncertain factors  $D_l$ ,  $\theta_p$ , and  $D_t$ , we get:

$$\begin{aligned} d\delta_h &= \frac{\partial f_h}{\partial D_{lp}} dD_{lp} + \frac{\partial f_h}{\partial \theta_{ph}} d\theta_{ph} + \frac{\partial f_h}{\partial D_{pt}} dD_{pt} \\ &= \tan \theta_{lh} dD_{lp} + F \tan \theta_{ph} \sec \theta_{ph} d\theta_{ph} + \tan \theta_{th} dD_{pt} \end{aligned} \tag{10}$$

$$\begin{aligned} d\delta_v &= \frac{\partial f_v}{\partial D_{lp}} dD_{lp} + \frac{\partial f_v}{\partial \theta_{pv}} d\theta_{pv} + \frac{\partial f_v}{\partial D_{pt}} dD_{pt} \\ &= \tan \theta_{lv} dD_{lp} + F \tan \theta_{pv} \sec \theta_{pv} d\theta_{pv} + \tan \theta_{tv} dD_{pt} \end{aligned} \tag{11}$$

The horizontal and vertical straightness measurement uncertainties ( $\sigma_{\delta_h}$  and  $\sigma_{\delta_v}$ ) of compound errors can be expressed as:

$$\sigma_{\delta_h} = \pm \sqrt{\left(\frac{\partial f_h}{\partial D_{lp}} \sigma D_{lp}\right)^2 + \left(\frac{\partial f_h}{\partial \theta_{ph}} \sigma \theta_{ph}\right)^2 + \left(\frac{\partial f_h}{\partial D_{pt}} \sigma D_{pt}\right)^2} \quad (12)$$

$$\sigma_{\delta_v} = \pm \sqrt{\left(\frac{\partial f_v}{\partial D_{lp}} \sigma D_{lp}\right)^2 + \left(\frac{\partial f_v}{\partial \theta_{pv}} \sigma \theta_{pv}\right)^2 + \left(\frac{\partial f_v}{\partial D_{pt}} \sigma D_{pt}\right)^2} \quad (13)$$

In the present study, for vertical vibration in the environment is always higher than horizontal vibration; therefore, the error angle of laser source at vertical direction can be assumed to three times of the horizontal direction as: The horizontal direction  $\theta_{lh}$ : 20 arcsec and the vertical direction  $\theta_{lv}$ : 60 arcsec; otherwise the travel distance of the laser beam passes through the pentaprism were much longer, therefore, the error angle of pentaprism can be assumed to be twice that of the laser source as: The horizontal direction  $\theta_{lh}$ : 40 arcsec, and the vertical direction  $\theta_{lv}$ : 120 arcsec. Based on the calibration experiment, the angle setup error of PSD at the horizontal and vertical direction,  $\theta_{ph}$  and  $\theta_{pv}$ , can be assumed to be less than  $\pm 2.5^\circ$ , the PSD measurement error  $F$  can be assumed to be less than  $\pm 100 \mu\text{m}$ , and the setup errors of the measuring distance,  $D_{lp}$  and  $D_{lt}$ , can be both assumed to be 1 m. Thus, the measurement uncertainties,  $\sigma_{\delta_h}$  and  $\sigma_{\delta_v}$ , were estimated to be  $\pm 1.02$  and  $\pm 3.06 \mu\text{m}$  from Equations (12) and (13), respectively.

#### 4. System Verification

##### 4.1. Altitude Angle between Laser Source and Pentaprism

The measurement error originated from the pentaprism module setup in the perpendicularity and parallelism measurement was analyzed by an autocollimator and a mirror both set behind the pentaprism, as shown in Figure 12. The uncertainty analysis of the pentaprism setup in the different angle lists in Table 3.

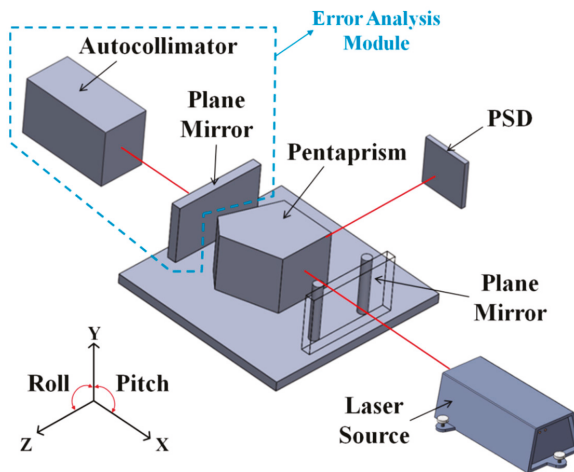


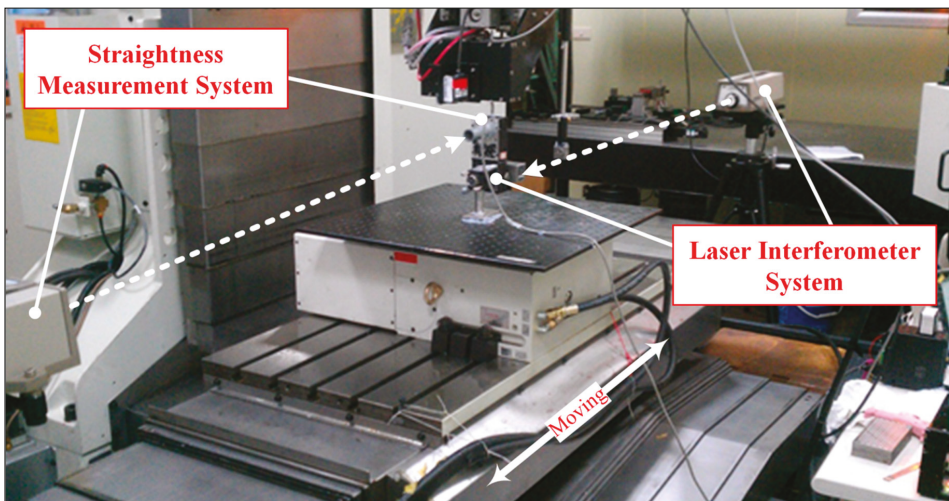
Figure 12. Schematic diagram of the pentaprism setup error detection.

**Table 3.** Uncertainty analysis of the pentaprism setup.

Errors on PSD	Pitch (arcsec)					Standard Deviation (arcsec)
	Altitude Angle Error	0	10	20	50	
Pitch (arcsec)	0	0.76	0.70	4.19	6.15	±2.3
Yaw (arcsec)		1.67	2.43	1.95	2.81	±0.4

4.2. Straightness Measurement

The verification of the proposed straightness system was conducted by using a dual laser interferometer (5529A, Agilent Tech., Santa Clara, CA, USA) as a reference standard to simultaneously measured straightness errors of a three-axis machine tool (KSC-611, KENT, New Taipei City, Taiwan) with the measurement distance of 1000 mm in a well air-conditioned laboratory, as shown in Figure 13.



**Figure 13.** Verification setup of straightness measurement.

The results of horizontal straightness measurements are shown in Figure 14. The average horizontal straightness errors measured by 5529A and the proposed system were found to be  $7.2 \pm 0.2 \mu\text{m}$  and  $8.2 \pm 0.5 \mu\text{m}$ , respectively. The results of vertical straightness measurements are shown in Figure 15. It was observed that the average vertical straightness errors measured by 5529A and the proposed system were  $5.3 \pm 0.3 \mu\text{m}$  and  $9.2 \pm 1.7 \mu\text{m}$ , respectively. The residual error of the horizontal straightness and vertical straightness are found to be  $\pm 1.3 \mu\text{m}$  and  $\pm 5.3 \mu\text{m}$ , respectively.

4.3. Perpendicularity Measurement

The verification of proposed perpendicularity system was conducted by using one square block (CVA239, Keysight Tech., Taipei, Taiwan), one dial indicator (GT1453sp, Girod-Tast Instrument, 2738 Court, Switzerland), and two linear guideways (both 400 mm) in a well air-conditioned laboratory. The dial indicator was utilized to examine the measurement results of the square block.

The linear guideway was first aligned parallel to each other using the square block and then measured by the proposed system, as shown in Figures 16 and 17. The average perpendicularity error measured by the square block and the proposed system were  $4.0 \pm 0.8 \mu\text{m}$  and  $2.7 \pm 0.9 \mu\text{m}$ , respectively.

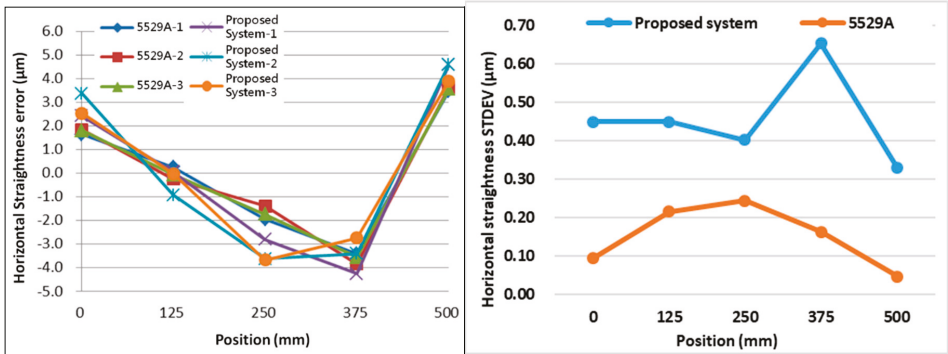


Figure 14. Results of horizontal straightness measurement.

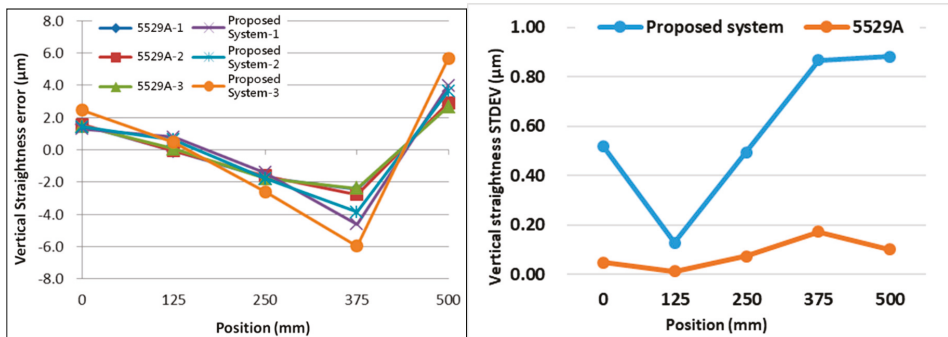


Figure 15. Results of vertical straightness measurement.

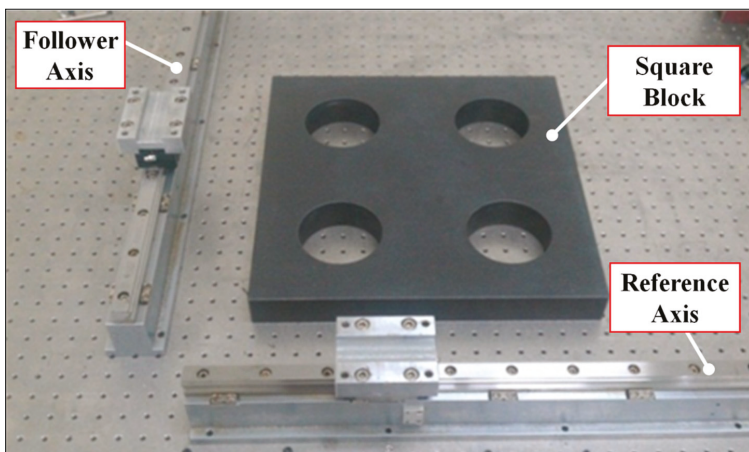


Figure 16. Verification setup of perpendicularity measurement using a square block.

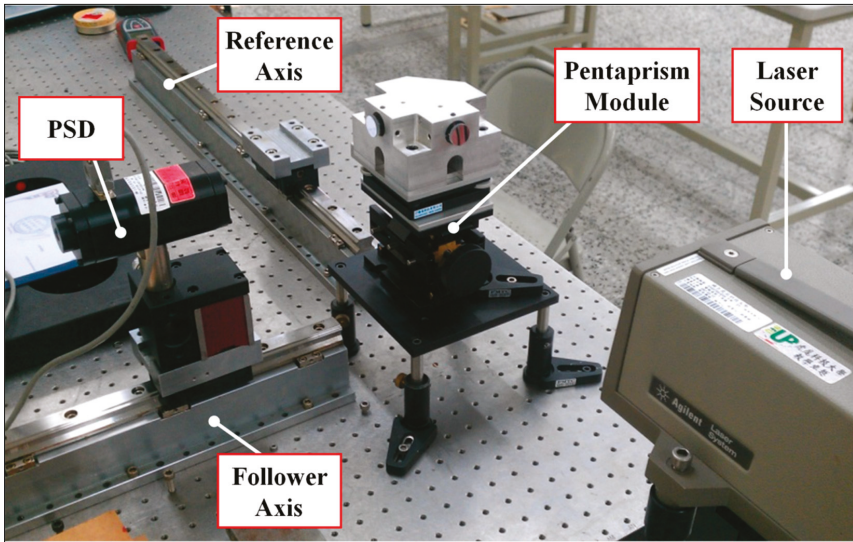


Figure 17. Verification setup of perpendicularity measurement using the proposed system.

#### 4.4. Parallelism Measurement

The verification of the proposed parallelism system was also conducted by using one square block, one dial indicator, and two linear guideways (both 400 mm) in a well air-conditioned laboratory. Figures 18 and 19 show the verification setup of the square block and the proposed system, respectively.

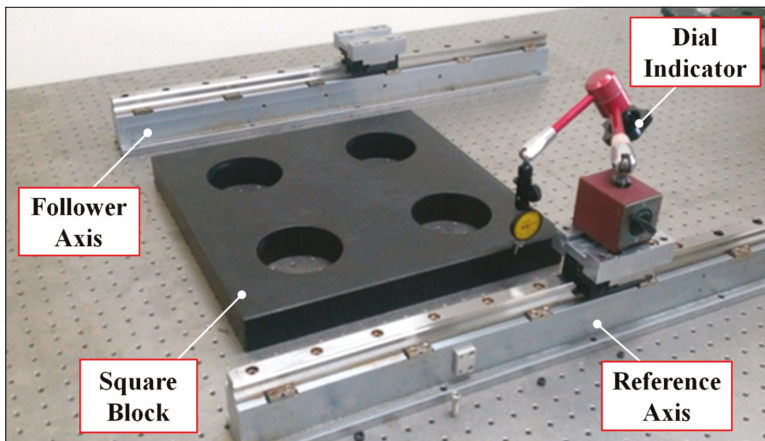
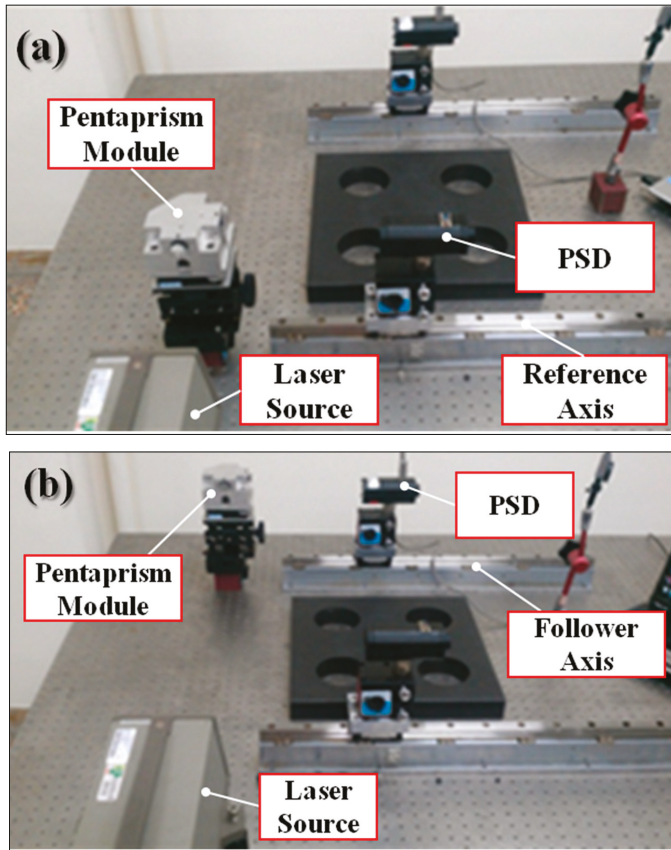


Figure 18. Verification setup of parallelism measurement using a square block.



**Figure 19.** Verification setup of parallelism measurement using the proposed system: (a) Reference axis and (b) follower axis.

The linear guideway was first aligned parallel to each other using a square block and then measured by the proposed system, as shown in Figures 17 and 18. The average parallelism error measured by the square block and the proposed system were  $11.3 \pm 0.5 \mu\text{m}$  and  $11.2 \pm 0.5 \mu\text{m}$ , respectively.

Table 4 shows the verification results of the long-travelling geometric error measurement system; obviously, the measurement difference  $\varepsilon_{Dif}$  between the commercial instrument/tool and the proposed system was very small, except for the vertical straightness error ( $>5 \mu\text{m}$ ). This phenomenon may originate from the vertical ground vibration of the laser source, which mainly depends on the measurement environment, as discussed in the above uncertainty analysis.

**Table 4.** Verification results of the proposed system.

	Measurement Distance (mm)	Commercial Instrument/Tool		Proposed System		Difference ( $\varepsilon_{Dif}$ *)
		Error ( $\varepsilon_I, \mu\text{m}$ )	Repeatability ( $\varepsilon_{IR}, \mu\text{m}$ )	Error ( $\varepsilon_s, \mu\text{m}$ )	Repeatability ( $\varepsilon_{SR}, \mu\text{m}$ )	
Straightness (Horizontal)		7.2	$\pm 0.2$	8.2	$\pm 0.5$	1.3
Straightness (Vertical)	1000	5.3	$\pm 0.3$	9.2	$\pm 1.7$	5.3
Perpendicularity	400	4.0	$\pm 0.8$	2.7	$\pm 0.9$	1.2
Parallelism	400	11.3	$\pm 0.5$	11.2	$\pm 0.5$	0.1

$$* \varepsilon_{Dif} = |(\varepsilon_I + |\varepsilon_{IR}|) - (\varepsilon_s + |\varepsilon_{SR}|)|.$$

4.5. Application of Proposed System

The proposed system was used in several Taiwanese domestic machine tool and precision stage manufacturer. The application included straightness, perpendicularity, and parallelism measurement.

The first case was for the straightness measurement of machine tool in fabrication site. We compared the proposed system with commercial straightness measurement equipment (ProLine, Status Pro Maschinenmesstechnik GmbH, Bochum, Germany). Total measurement range was 6000 mm, the application setup is shown in Figure 20, the proposed system installed in parallel with the ProLine on the saddle of the machine tool. The measurement results are shown in Table 5.

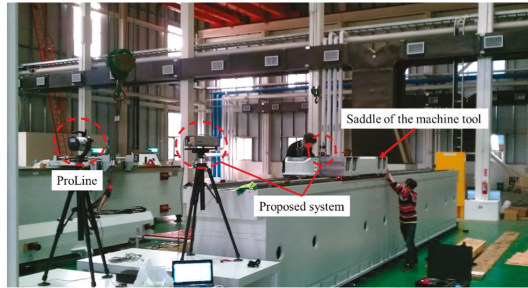


Figure 20. Case 1 straightness error measurement setup.

Table 5. Case 1 straightness error measurement results.

	Straightness Error (Horizontal)	Straightness Error (Vertical)
ProLine	16 $\mu\text{m}$	16 $\mu\text{m}$
Proposed system	13.97 $\mu\text{m}$	11.21 $\mu\text{m}$

The second case was for the straightness measurement of double-column machine tool. We compared the proposed system with commercial laser interferometer (5529A, Agilent Tech., Santa Clara, CA, USA). The total measurement range was 2000 mm; the application setup is shown in Figure 21. The measurement results are shown in Table 6.

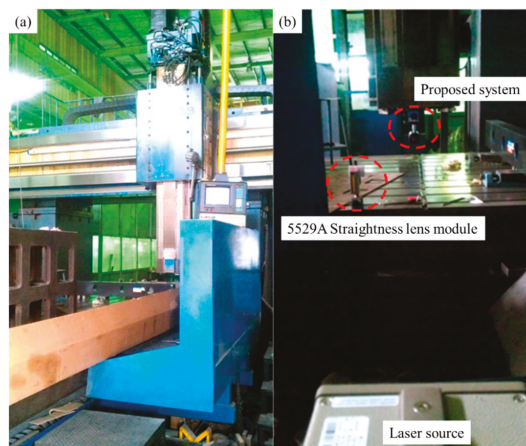
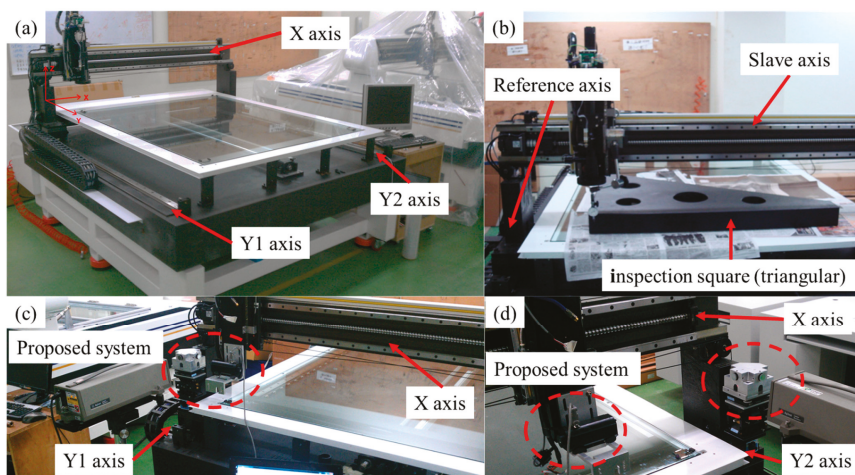


Figure 21. Case 2 straightness error measurement setup on double column machine tool, (a) the double column machine tool and (b) the setup of 5529A laser interferometer straightness module and proposed system.

**Table 6.** Case 2 straightness error measurement results.

	Straightness Error (Horizontal)	Straightness Error (Vertical)
5529A	34 $\mu\text{m}$	24 $\mu\text{m}$
Proposed system	36 $\mu\text{m}$	28 $\mu\text{m}$

The third case was for the perpendicularity measurement of AOI equipment. In this case, the reference measurement value was provided by the third company with the inspection square (triangular) (1200 mm grade 0, OPUS, Taoyuan, Taiwan). The total measurement range was 1100 mm; the application setup is shown in Figure 22. The measurement results are shown in Table 7.



**Figure 22.** Case 3 perpendicular error measurement setup. (a) The Automated optical inspection (AOI) equipment to be measured, (b) the inspection square (triangular) measurement setup, (c) the  $Y1 \perp X$  measurement setup of proposed system, and (d) the  $Y2 \perp X$  measurement setup of proposed system.

**Table 7.** Case 3 perpendicularity error measurement results.

	Two-Axis Perpendicularity Displacement Error	Two-Axis Perpendicularity Angle Error
Inspection square (triangular)	5 $\mu\text{m}/1.1 \text{ m}$	N/A
Proposed system $Y1 \perp X$	14.3 $\mu\text{m}/1.1 \text{ m}$	2.68 arcsec
Proposed system $Y2 \perp X$	14.1 $\mu\text{m}/1.1 \text{ m}$	-2.66 arcsec

The fourth case was for the parallel error measurement of the machine tool. In this case, the fourth company provided the reference measurement value. The total measurement range was 1600 mm; the application setup is shown in Figure 23. The measurement results are shown in Table 8.

**Table 8.** The parallelism error measurement results.

	Parallelism Angle Error	Parallelism Displacement Error
Company	3.3 arcsec	16 $\mu\text{m}/1.6 \text{ m}$
Proposed system	1.09 arcsec	5.29 $\mu\text{m}/1.6 \text{ m}$



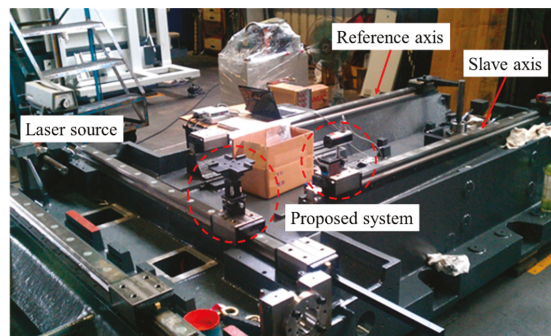


Figure 23. Case 4 parallelism error measurement setup.

## 5. Conclusions

In this study, a multipurpose geometric error measurement system mainly composed of a laser source and a self-developed optical module has been developed for the measurement of straightness, perpendicularity, and parallelism errors for a linear guideway assembly process.

This includes two adjustment methods using an in-lab-developed optical module capable of calibrating the altitude angle of the pentaprism: The first method is designed for ease of operations based on Michelson principle and uses a laser interferometer as the light receiver, and the second one is aimed for high calibration-repeatability based on autocollimator principle using QD to replace the light receiver. The measured data are analyzed by the least-squares method in order to obtain corresponding geometric errors; meanwhile, the system uncertainty analysis is also conducted for the laser source, PSD, and pentaprism module for evaluating the system performance.

The features of the multipurpose geometric error measurement system presented in this paper are described as follows:

- (a) There is no need to reset the system if the laser source is sheltered during the measurement process;
- (b) The horizontal and vertical straightness errors can be obtained in one measurement;
- (c) The system is capable of applying to the linear guideway assembly process of various double-column, three-axis, and five-axis machine tools;
- (d) The residual error of the horizontal straightness, vertical straightness, perpendicularity, and parallelism are found to be  $\pm 1.3 \mu\text{m}$ ,  $\pm 5.3 \mu\text{m}$ ,  $\pm 1.2 \mu\text{m}$ , and  $\pm 0.1 \mu\text{m}$ , respectively, which are better compared to those detected by the commercial laser interferometer and granite reference blocks.

**Author Contributions:** The authors contributed equally to the manuscript.

**Acknowledgments:** The present work was funded by the Ministry of Economic Affairs (MoEA) and the Ministry of Science and Technology (MoST) of Taiwan (Republic of China), the authors would like gratefully express their sincere acknowledgment to MoEA and MoST.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Sommargren, G.E. Linear/Angular Displacement Interferometer for Wafer Stage Metrology. *Proc. SPIE* **1989**, *1088*, 268–273.
2. Nakamura, O.; Goto, M. Four-beam laser interferometry for three-dimensional microscopic coordinate measurement. *Appl. Opt.* **1994**, *33*, 31–36. [[CrossRef](#)] [[PubMed](#)]
3. Lee, C.-W.; Kim, S.-W. An ultraprecision stage for alignment of wafers in advanced microlithography. *Precis. Eng.* **1997**, *21*, 113–122. [[CrossRef](#)]

4. Menq, C.H.; Zhang, J.H.; Shi, J. Design and development of an interferometer with improved angular tolerance and its application to x-y theta measurement. *Rev. Sci. Instrum.* **2000**, *71*, 4633–4638. [[CrossRef](#)]
5. Zhang, Z.; Menq, C.H. Laser interferometric system for six-axis motion measurement. *Rev. Sci. Instrum.* **2007**, *78*, 1–8. [[CrossRef](#)] [[PubMed](#)]
6. Liu, C.-H.; Jywe, W.-Y.; Jeng, Y.-R.; Hsu, T.-H.; Li, Y.-t. Design and control of a long-traveling nano-positioning stage. *Precis. Eng.* **2010**, *34*, 497–506. [[CrossRef](#)]
7. Ni, J.; Huang, P.S.; Wu, S.M. A Multi-Degree-of-Freedom Measuring System for CMM Geometric Errors. *J. Eng. Ind.* **1992**, *114*, 362–369.
8. Huang, P.S.; Ni, J. On-line error compensation of coordinate measuring machines. *Int. J. Mach. Tools Manuf.* **1995**, *35*, 725–738. [[CrossRef](#)]
9. Shimizu, S.; Lee, H.S.; Imai, N. Simultaneous Measuring Method of Table Motion Error in 6 Degrees of Freedom. *Int. J. Jpn. Soc. Precis. Eng.* **1994**, *28*, 273–274.
10. Chou, C.; Chou, L.-Y.; Peng, C.-K.; Huang, Y.-C.; Fan, K.-C. CCD-based CMM Geometrical error measurement using fourier phase shift algorithm. *Int. J. Mach. Tools Manuf.* **1997**, *37*, 579–590. [[CrossRef](#)]
11. Fan, K.C.; Chen, M.J.; Huang, W.M. A six-degree-of-freedom measurement system for the motion accuracy of linear stages. *Int. J. Mach. Tools Manuf.* **1998**, *38*, 155–164. [[CrossRef](#)]
12. Fan, K.-C.; Chen, M.-J. A 6-degree-of-freedom measurement system for the accuracy of X-Y stages. *Precis. Eng.* **2000**, *24*, 15–23. [[CrossRef](#)]
13. Jywe, W.; Chen, C.J.; Hsieh, W.H.; Lin, P.D.; Jwo, H.H.; Yang, T.Y. A novel simple and low cost 4 degree of freedom angular indexing calibrating technique for a precision rotary table. *Int. J. Mach. Tools Manuf.* **2007**, *47*, 1978–1987. [[CrossRef](#)]
14. You, F.-L.; Feng, Q.-B.; Zhang, B. Straightness error measurement based on common-path compensation for laser beam drift. *Opt. Precis. Eng.* **2011**, *3*, 004.
15. Jywe, W.; Hsu, T.-H.; Liu, C.H. Non-bar, an optical calibration system for five-axis CNC machine tools. *Int. J. Mach. Tools Manuf.* **2012**, *59*, 16–23. [[CrossRef](#)]
16. Huang, P.; Li, Y.; Wei, H.; Ren, L.; Zhao, S. Five-degrees-of-freedom measurement system based on a monolithic prism and phase-sensitive detection technique. *Appl. Opt.* **2013**, *52*, 6607–6615. [[CrossRef](#)] [[PubMed](#)]
17. Chen, B.Y.; Xu, B.; Yan, L.P.; Zhang, E.Z.; Liu, Y.N. Laser straightness interferometer system with rotational error compensation and simultaneous measurement of six degrees of freedom error parameters. *Opt. Express* **2015**, *23*, 9052–9073. [[CrossRef](#)] [[PubMed](#)]
18. Hsieh, T.H.; Jywe, W.Y.; Chen, S.L.; Liu, C.H.; Huang, H.L. Note: Development of a high resolution six-degrees-of-freedom optical vibrometer for precision stage. *Rev. Sci. Instrum.* **2011**, *82*, 056101. [[CrossRef](#)] [[PubMed](#)]
19. Gao, W.; Arai, Y.; Shibuya, A.; Kiyono, S.; Park, C.H. Measurement of multi-degree-of-freedom error motions of a precision linear air-bearing stage. *Precis. Eng.* **2006**, *30*, 96–103. [[CrossRef](#)]



© 2019 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).

Article

# A Method for Expansion of Z-Directional Measurement Range in a Mode-Locked Femtosecond Laser Chromatic Confocal Probe

Chong Chen, Ryo Sato, Yuki Shimizu \*, Taku Nakamura, Hiraku Matsukuma and Wei Gao

Precision Nanometrology Laboratory, Department of Finemechanics, Tohoku University, Sendai 980-8579, Japan; chongchen@nano.mech.tohoku.ac.jp (C.C.); satoryo@nano.mech.tohoku.ac.jp (R.S.); nakamura@nano.mech.tohoku.ac.jp (T.N.); hiraku.matsukuma@nano.mech.tohoku.ac.jp (H.M.); gaowei@cc.mech.tohoku.ac.jp (W.G.)

\* Correspondence: yuki.shimizu@nano.mech.tohoku.ac.jp; Tel.: +81-22-795-6950

Received: 13 December 2018; Accepted: 22 January 2019; Published: 29 January 2019



**Abstract:** A method is proposed to expand the Z-directional measurement range of a fiber-based dual-detector chromatic confocal probe with a mode-locked femtosecond laser source. In the dual-detector chromatic confocal probe, the Z-directional displacement of a measurement target is derived from the peak wavelength in the normalized intensity ratio from the two light intensities obtained by the two identical fiber detectors. In this paper, a new method utilizing the main-lobe and side-lobes of axial responses acquired from both the normalized intensity ratio  $I_a$  and the invert normalized intensity ratio  $I_n$ , which is the inverse of  $I_a$ , is proposed to obtain the seamless relationship between the peak wavelength and the Z-directional displacement of a measurement target. Theoretical calculations and experimental investigation are carried out to demonstrate the feasibility of the proposed measurement range expansion method.

**Keywords:** chromatic confocal probe; femtosecond laser; measurement range expansion; side-lobe

## 1. Introduction

Confocal microscopy [1,2] with a confocal probe has outstanding image formation properties in depth sectioning imaging and optical tomographic imaging [3–5], which makes it a potentially powerful tool for image formation in industrial and medical engineering fields, such as three-dimensional (3D) fine structure image formation [4,5] and living body observation of cells [6,7], compared with traditional microscopy. It is also expected to measure large-area 3D microstructured surfaces [8,9]. Meanwhile, the relatively small measurement range of a confocal probe with a monochromatic laser source has prevented the further application of confocal microscopy in these areas. Therefore, the extension of the measurement range of a confocal probe is critical and crucial for expansion of the application range of the confocal microscopy. For the extension of the measurement range of a confocal probe, several efforts have already been made so far [10–20]. Among them, enhancement of the chromatic dispersion has been proven as a way to realize the extension of the measurement range without any axial scanning of a target object or a measuring head of a confocal probe [18–20]. The employment of a diffractive optical element (DOE) is a candidate method to enhance the chromatic aberration, and five-fold extension of measurement range has been achieved [10]. Other DOEs for enlargement of the chromatic dispersion, such as a micro-lens array [13] and a Fresnel lens [14], have achieved enlarged measurement ranges of 210  $\mu\text{m}$  and 200  $\mu\text{m}$ , respectively. With a diffractive lens and objective lenses [11], an expansion of a measurement range has also been achieved [11,16]. Furthermore, by utilizing a pupil filter technique, such as super-resolution pupil filter [15] or annular pupil filter [17], for the re-shaping of a measuring beam, measurement ranges have also been enlarged to 40  $\mu\text{m}$  and 14  $\mu\text{m}$ , respectively. However,

the above methods need to employ additional optical components in the optical setup, which could affect the performances of original optical confocal setups. Meanwhile, employing a broadband laser source is another approach to expand the measurement range of a chromatic confocal probe [18,19], and several methods have been proposed so far. With the employment of a supercontinuum laser source, over 16-fold expanded measurement range has been acquired [19]. However, non-smoothness of the laser spectrum, which means non-uniformity of the spectral intensity distribution, has restricted the full exploration of the whole spectrum of the supercontinuum laser source for a wider axial measurement range in a chromatic confocal probe. In responding to the background described above, a fiber-based dual-detector chromatic confocal probe with a mode-locked femtosecond laser source has been developed [3]. The proposed method has demonstrated that the influence of non-smoothness of the mode-locked femtosecond laser source can be compensated while expanding the measurement range over the whole spectrum of the laser by employing an axial response, defined as the intensity ratio of the two different confocal signals.

In this paper, as the second step of research, a new signal processing method for the axial response curve from the fiber-based dual-detector chromatic confocal probe with a mode-locked femtosecond laser source is proposed. The method is designed to further expand the measurement range without modifying the optical setup of the dual-detector chromatic confocal probe. Due to the diffraction effect of a point detector in confocal microscopy [20], there are always a main-lobe and side-lobes in the axial response of a confocal probe, in which main-lobe stands for the highest power intensity distribution curve of the axial response, while side-lobes stand for local maxima power intensity distribution curves of the axial response. Since the presence of the side-lobes not only reduces the light intensity of the main-lobe but also degrades the imaging quality and depth resolution of confocal microscopy, a variety of methods have been proposed so far to reduce [21–29], or even remove [30,31], the side-lobes in the axial response. On the contrary, in this paper, these annoying side-lobes are utilized to expand the measurement range of the dual-detector chromatic confocal probe with a mode-locked femtosecond laser source. Theoretical analysis and computer simulation, as well as some experiments, are carried out to demonstrate the feasibility of the proposed method for the extension of the measurement range.

## 2. Principle of the Extension of the Measurement Range

### 2.1. Imaging Principle

Figure 1 shows a schematic of the optical configuration for the fiber-based dual-detector chromatic confocal probe with a mode-locked femtosecond laser source [3]. A mode-locked femtosecond laser source is employed, while single-mode step index fibers are used as the signal-propagating media. The  $k$ th optical mode in the mode-locked femtosecond laser has a specific frequency  $\nu_k$ , which corresponds to the deterministic working wavelength  $\lambda_k$ , which can be expressed as follows by using the carrier envelope offset frequency  $\nu_{ceo}$  and the pulse repetition rate  $\nu_{rep}$  [32–34]:

$$\nu_k = \nu_{ceo} + k \cdot \nu_{rep} \tag{1}$$

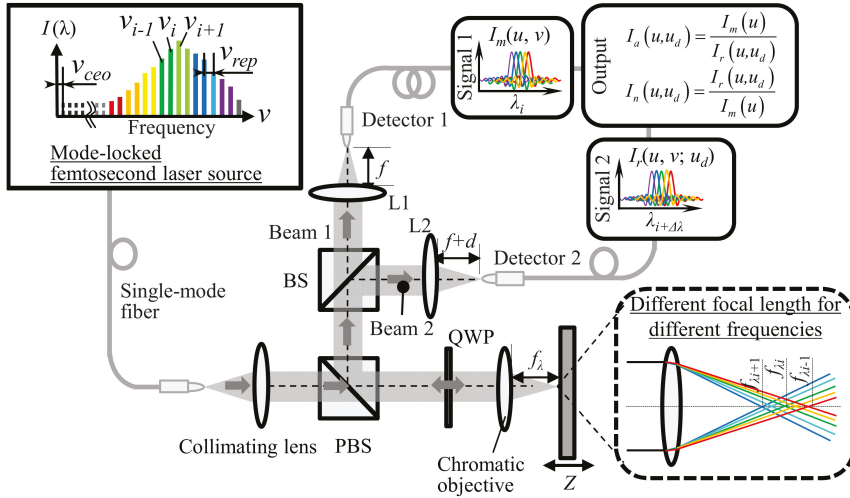
$$\lambda_k = \frac{c}{\nu_k} \tag{2}$$

where  $c$  is the speed of light in vacuum. For the image formation of confocal microscopy, the point spread function (PSF), and the pupil function of imaging lens are critical and are used in the equation of the final axial response of the confocal microscopy. The equation of the PSF, pupil function, and the equation of the effective PSF are given as follows [3,35–37]:

$$h_i(u, v) = \int_0^1 P_i(u, \rho) J_0(\rho v) \rho d\rho, i = 1, 2, \tag{3}$$

$$P_i = \begin{cases} 1, & \text{in pupil} \\ 0, & \text{else} \end{cases}, \quad i = 1, 2 \quad (4)$$

$$h_{i\_eff}(u, v) = h_i(u, v) \otimes e_1(u, v), \quad i = 1, 2 \quad (5)$$



**Figure 1.** Schematic diagram of the fiber-based dual-detector chromatic confocal configuration with a mode-locked femtosecond laser source.

The complex amplitude  $U_f$  on the confocal plane is therefore given as follows:

$$U_f(u, v) = h_{1\_eff}(u, v)h_{2\_eff}(u, v) \otimes t(u, v) \quad (6)$$

where  $h_i(u, v)$  and  $P_i(u, \rho)$  are the PSF and the pupil function of the two imaging lenses, respectively,  $\rho$  being the normalized radius of the chromatic objective lens,  $\otimes$  standing for the convolution operation,  $J_0(\cdot)$  being the Bessel function of the first kind of zero order,  $h_{i\_eff}$  standing for the effective PSF of the two imaging lenses,  $e_1$  being the fundamental transverse orthonormal modal field. It should be noted that  $u$  and  $v$  are the optical coordinates of the confocal setup associated with the real coordinates  $z$  and  $r$ , respectively, and can be described by the following equations [3,36,37]:

$$v = \frac{2\pi}{\lambda} r \sin\beta \quad (7)$$

$$u = \frac{2\pi}{\lambda} z \sin^2\beta \quad (8)$$

where  $\lambda$  and  $\sin\beta$  indicate the light wavelength and the numerical aperture of the chromatic objective lens, respectively. Since the whole confocal configuration is coherent, the light intensity  $I$  obtained by the detector placed on the focal plane of the objective lens is given by the following equation:

$$I = |h_{1\_eff}(u, v)h_{2\_eff}(u, v) \otimes t(u, v)|^2 \quad (9)$$

where  $t(u, v)$  is the amplitude reflection or transmittance of the object. Based on the above analyses, the light intensity distribution  $I_m(u)$  on the focal plane of the objective lens, where the fiber detector is placed, is given as follows:

$$I_m(u) = \left| \int_0^1 \exp(ju\rho^2) P_{1\_eff} P_{2\_eff} \rho d\rho \right|^2 \tag{10}$$

In order to overcome the non-smoothness of the laser spectrum in the mode-locked femtosecond laser source, a dual-detector differential confocal configuration shown in Figure 1 has been proposed [3]. The light intensity distribution  $I_r(u, u_d)$  on the plane with a distance  $d$  from the focal plane of the objective lens is given as follows:

$$I_r(u, u_d) = \left| \int_0^1 \exp\left[j\left(u + \frac{u_d}{2}\right)\rho^2\right] P_{1\_eff} P_{2\_eff} \rho d\rho \right|^2 \tag{11}$$

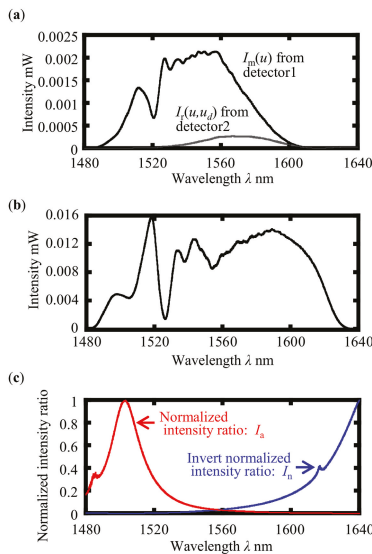
where  $u_d$  is the optical distance associated with the defocus  $d$  that can be described as follows:

$$u_d = \frac{2\pi}{\lambda} d \sin^2 \beta \tag{12}$$

Figure 2a shows typical spectra of the light intensities  $I_m(u)$  and  $I_r(u, u_d)$  obtained by the two identical fiber detectors. As can be seen in the figure, it is difficult to distinguish a peak in the spectra, since a mode-locked femtosecond laser source has a non-smooth spectrum, as shown in Figure 2b. The spectrum shown in Figure 2b is from a specific femtosecond laser used in the following experiments and its non-smooth spectrum is not the same as the other generic laser sources. To address the aforementioned problem, the normalized intensity ratio  $I_a(u, u_d)$  and the invert normalized intensity ratio  $I_n(u, u_d)$  of the obtained signal intensities, determined by the following equations, are employed in the proposed method:

$$I_a(u, u_d) = \frac{I_m(u)}{I_r(u, u_d)} \tag{13}$$

$$I_n(u, u_d) = \frac{I_r(u, u_d)}{I_m(u)} \tag{14}$$



**Figure 2.** Typical normalized intensity ratio obtained by the dual-detector chromatic confocal configuration with a mode-locked femtosecond laser source. (a) Light intensities obtained by the two identical fiber detectors with a target position of 40 μm and a defocus  $d$  of 150 μm; (b) Working spectrum of the mode-locked femtosecond laser source; (c) Normalized intensity ratio and invert normalized intensity ratio obtained from the two light intensities.

Figure 2c shows the normalized intensity ratio and the invert normalized intensity ratio obtained from the spectra of the obtained light intensities shown in Figure 2a, with a target position of 40 μm and a defocus  $d$  of 150 μm, in which main-lobes can clearly be distinguished in both the normalized intensity ratio  $I_a$  and the invert normalized intensity ratio  $I_n$ , but side-lobes are inconspicuous. Therefore, in subsequent simulations and experiments, to make good use of the side-lobes to extend the measurement range, we set a defocus  $d$  to be 110 μm so that the side-lobes can be visible. As can be seen in the figure, main peaks can clearly be distinguished in both the axial response and the reversed axial response; this enables the proposed fiber-based dual-detector chromatic confocal probe to have a high resolution of 30 nm and good linearity for displacement measurement along the depth direction.

### 2.2. Expansion of the Measurement Range

The measurement range of the proposed fiber-based dual-detector chromatic confocal probe is determined by the spectral bandwidth of the mode-locked laser source and the chromatic objective employed in the optical setup. A thin plano-convex lens made of N-SF11 is employed in the developed optical setup as the chromatic objective. According to the thin-lens equation [3,35], the relationship between a focal length of the chromatic objective  $f_\lambda$  and a refractive index  $N_\lambda$  is given as follows:

$$f_\lambda = \frac{1}{(N_\lambda - 1)(1/r_1 - 1/r_2)} \tag{15}$$

where  $r_1$  and  $r_2$  are curvature radii of the chromatic objective lens. The refractive index  $N_\lambda$  for the light wavelength  $\lambda$  can be obtained from the following Sellmeier equation [3]:

$$N_\lambda = \sqrt{1 + \frac{B_1\lambda^2}{\lambda^2 - C_1} + \frac{B_2\lambda^2}{\lambda^2 - C_2} + \frac{B_3\lambda^2}{\lambda^2 - C_3}} \tag{16}$$

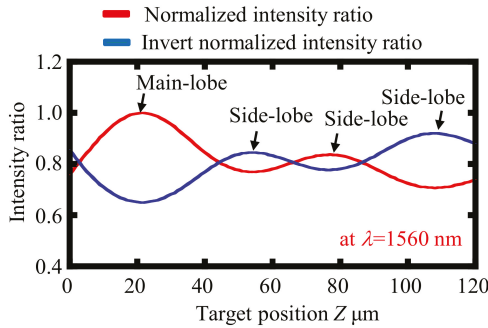
where  $B_i$  and  $C_i$  are the parameters associated with N-SF11. The measurement range  $\Delta Z$  of the chromatic confocal probe is therefore given as follows:

$$\Delta Z = |f_{\lambda_1} - f_{\lambda_2}| \tag{17}$$

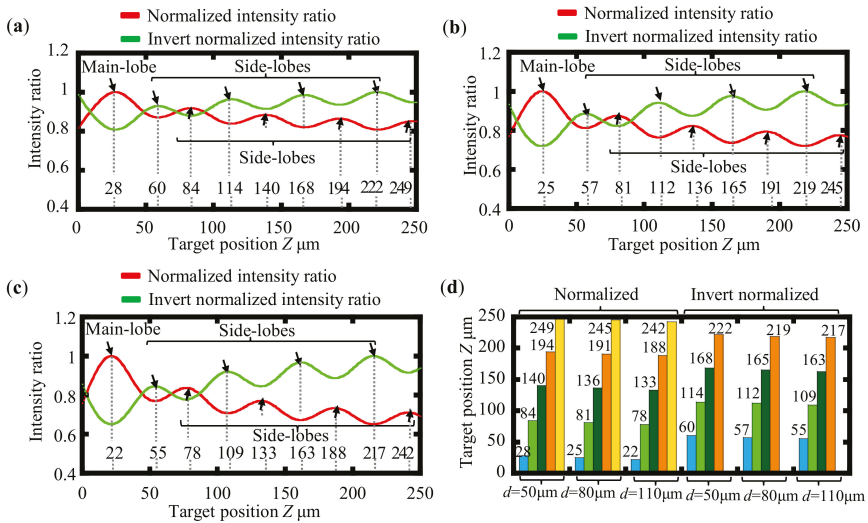
where  $f_{\lambda_1}$  and  $f_{\lambda_2}$  are the longest and shortest working wavelengths of the mode-locked femtosecond laser source, respectively. In the previous work by the authors [3], the measurement range of the developed chromatic confocal probe is derived from the linear relationship between the target position and the peak wavelength detected in the normalized intensity ratio.

To further expand the measurement range of the fiber-based dual-detector chromatic confocal probe without any modifications on the optical setup, in this paper, a new method employing both the main-lobes and the side-lobes of the axial responses obtained from the normalized intensity ratio  $I_a(u, u_d)$  and the invert normalized intensity ratio  $I_n(u, u_d)$  is proposed. Figure 3 shows a variation of the normalized intensity  $I_a$  as a function of the Z-directional position of a measurement target at a specific wavelength  $\lambda$ , which is referred to as the axial response [3]. Figure 3 also shows the axial response obtained from the invert normalized intensity ratio at the same wavelength  $\lambda$ . As can be seen in the figures, not only the main-lobes but also the side-lobes of the axial response proposed in our previous work [3], which cannot be observed in the single axial responses of  $I_m$  and  $I_r$  without the normalization process, can clearly be distinguished in the axial responses. This is another benefit of our newly proposed axial response; the side-lobes are clearer compared with the single axial response expected from a Lorentzian shape, in which the side-lobe is fuzzy and cannot be identified easily. Figure 4a–c show the normalized intensity ratio  $I_a(u, u_d)$  and the invert normalized intensity ratio  $I_n(u, u_d)$  with different defocus  $d$ , and Figure 4d shows the side-lobe spacing and the main-lobe to side-lobe spacing with respect to different defocus  $d$ . In order to make good use of the side-lobes to extend the measurement range, we set defocus  $d$  to be 110 μm, in which the side-lobes are more

visible and identifiable than the cases where  $d$  is set to be within a range from 50  $\mu\text{m}$  to 80  $\mu\text{m}$ . Consequently, the employment of the side-lobes has a possibility of expanding the measurement range of the fiber-based dual-detector chromatic confocal probe, and theoretical calculations are therefore carried out in the following.



**Figure 3.** A typical axial response obtained by the dual-detector chromatic confocal probe with a mode-locked femtosecond laser source with axial response of the normalized intensity ratio and axial response of the invert normalized intensity ratio, with target positions ranging from 0 to 120  $\mu\text{m}$  and a defocus  $d$  of 110  $\mu\text{m}$ .



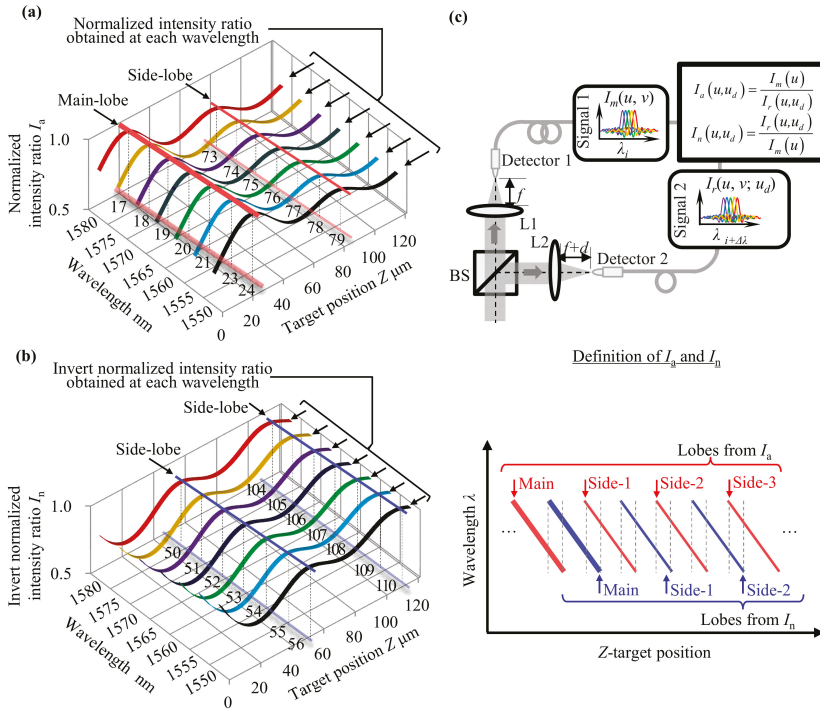
**Figure 4.** The main-lobes and the side-lobes extracted from the normalized intensity ratio and the invert normalized intensity ratio obtained at central wavelength  $\lambda = 1560 \text{ nm}$ , together with side-lobe spacing and main-lobe to side-lobe spacing with target positions ranging from 0 to 250  $\mu\text{m}$  and defocus  $d$  ranging 50  $\mu\text{m}$  to 110  $\mu\text{m}$ . (a) Main-lobe and side-lobes obtained from normalized and invert normalized intensity ratio at central wavelength with  $d = 50 \mu\text{m}$ . (b) Main-lobe and side-lobes obtained from normalized and invert normalized intensity ratio at central wavelength with  $d = 80 \mu\text{m}$ . (c) Main-lobe and side-lobes obtained from normalized and invert normalized intensity ratio at central wavelength with  $d = 110 \mu\text{m}$ . (d) The side-lobe spacing and main-lobe to side-lobe spacing at central wavelength with  $d = 50 \mu\text{m}$ , 80  $\mu\text{m}$ , and 110  $\mu\text{m}$ .

Regarding the non-smoothness of the laser spectrum shown in Figure 2b, axial responses of the fiber-based dual-detector chromatic confocal probe are estimated based on Equations (13) and (14), in



which the proposed axial response is defined as the intensity ratio of the signal 1 and signal 2, not as the single axial response anymore. The distance between the main-lobe and the side-lobe is therefore determined by the relative magnitude of the ratios of signal 1 and signal 2 at the same target position. Furthermore, the relative height of the main-lobe and the side-lobes is also controlled by the relative magnitude of the ratios of signal 1 and signal 2 at the same target position, which is further controlled by the defocus  $d$  that is shown in Equation (11). As illustrated in Equation (11), when we set a defocus  $d$  to be zero, theoretically the values of Equation (13) and Equation (14) always become 1.

For the calculations, the parameters summarized in Table 1 are employed to express the effect of the chromatic objective. Figure 5a,b show the axial responses from the normalized intensity ratio and the invert normalized intensity ratio, respectively, calculated at the wavelengths from 1550 nm to 1580 nm. As can be seen in the figures, the main-lobe and the side-lobes of the axial response from the invert normalized intensity ratio are found to be shifted along the axial ( $Z$ -) direction with respect to those from the normalized intensity ratio. In the previous work by the authors [3], only the information from the main-lobe in the normalized intensity ratio has been employed to detect a displacement in the axial direction. As a result, the measurement range has been limited to be approximately 40  $\mu\text{m}$ . Meanwhile, by using the information from not only the main-lobe but also the side-lobes in the axial responses from both the intensity ratios, the measurement range can be expanded; as can be seen in Figure 5c, the gap of the neighboring lobes (including the main-lobe and side-lobes) in the axial response of  $I_a(u, u_d)$  can be covered by the corresponding lobes of the axial response of  $I_n(u, u_d)$ .

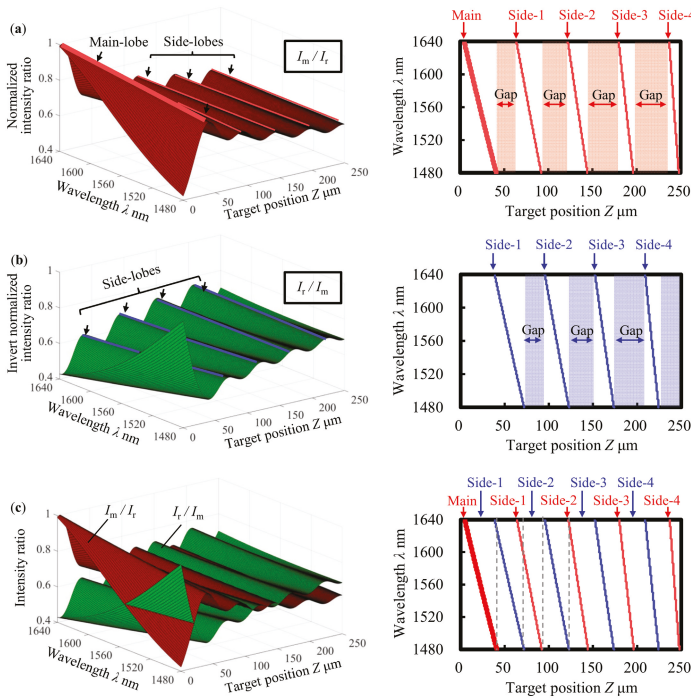


**Figure 5.** The main-lobe and the side-lobes in the axial response obtained by the dual-detector chromatic confocal probe with a mode-locked femtosecond laser source with target positions ranging from 0 to 120  $\mu\text{m}$  and a defocus  $d$  of 110  $\mu\text{m}$ . (a) Axial response of the normalized intensity ratio with shadow of the main-lobe and side-lobe and the corresponding target position values. (b) Axial response of the invert normalized intensity ratio with shadow of the side-lobes and the corresponding target position values. (c) The main-lobes and the side-lobes in the axial responses.

**Table 1.** Parameters employed for the theoretical calculation.

Items	Symbol	Value
Lens material	-	N-SF11
Surface radii of imaging lens	$r_1$	7.85 mm
	$r_2$	$\infty$
<u>Sellmeier coefficients</u>		
$B_1$	1.73759695	$C_1$ 0.013188707
$B_2$	0.313747346	$C_2$ 0.062306814
$B_3$	1.89878101	$C_3$ 155.23629

Theoretical calculation is expanded to a working range of the mode-locked laser of 1.48  $\mu\text{m}$ –1.64  $\mu\text{m}$  and a target position over 250  $\mu\text{m}$ . Figure 6a,b show the axial responses of the normalized intensity ratio  $I_a$  and the invert normalized intensity ratio  $I_n$ . As can be seen in the figure, side-lobes are clearly observed over a target position range of 250  $\mu\text{m}$  in both of the axial responses. Meanwhile, the existence of a gap between each neighboring lobes is also confirmed in both of the axial responses. In addition, the theoretical calculation results show that the gap becomes wider as the increase of the target position. Figure 6c shows the main-lobes and the side-lobes of the normalized intensity ratio  $I_a$  plotted together with those of the invert normalized intensity ratio  $I_n$ . As can be seen in the figure, the two axial responses are almost 180° out of phase, and the gaps between each neighboring lobes in the axial response of the normalized intensity ratio can be covered by the lobes in the axial response of the invert normalized intensity ratio over a target position range of 150  $\mu\text{m}$ .



**Figure 6.** The main-lobe and the side-lobes in the axial responses obtained by the dual-detector chromatic confocal probe with a mode-locked femtosecond laser source. (a) Axial response of the normalized intensity ratio with target positions ranging from 0 to 250  $\mu\text{m}$  and a defocus  $d$  of 110  $\mu\text{m}$ . (b) Axial response of the invert normalized intensity ratio. (c) The main-lobes and the side-lobes in the axial responses.

It should be noted that in order to utilize the side-lobes of the axial responses as illustrated in Figure 6c for the expansion of the axial measurement range, calculation of partial derivatives of Equation (13) is necessary:

$$\frac{\partial I_a}{\partial z} = \frac{\partial I_a}{\partial u} \frac{\partial u}{\partial z} = \frac{4\pi \sin^2 \beta \{ I_r \sqrt{I_m} \exp(ju\rho^2) - I_m \sqrt{I_r} \exp[j(u + \frac{u_d}{2})\rho^2] \}}{I_r^2 \lambda} \quad (18)$$

In order to obtain the relative maximum intensity values for wavelength-to-displacement encoding of the side-lobes, let Equation (13) be equal to zero and get the relationship between the target position  $z$  and wavelength  $\lambda$ :

$$\frac{\partial I_a}{\partial z} = \frac{\partial I_a}{\partial u} \frac{\partial u}{\partial z} = 0 \quad (19)$$

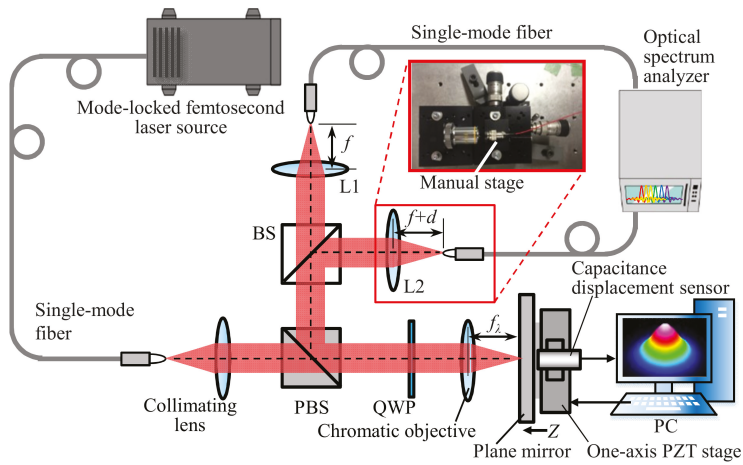
As illustrated in Figure 6c, the main-lobe and side-lobes are employed in the wavelength-to-displacement encoding, but gaps between main-lobe and side-lobes make the final measurement range discontinuous. In order to eliminate the gaps, side-lobes of the axial response of the invert normalized intensity ratio are employed. Therefore, the following relationship should also be satisfied:

$$\frac{\partial I_n}{\partial z} = \frac{\partial I_n}{\partial u} \frac{\partial u}{\partial z} = 0 \quad (20)$$

### 3. Experiments and Discussions

#### 3.1. Experimental Setup

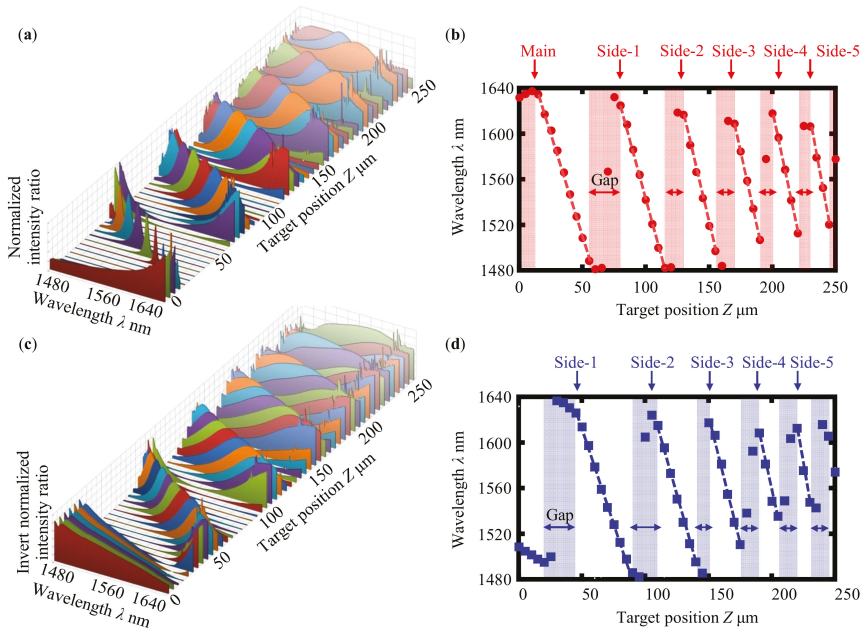
Figure 7 shows a schematic of the fiber-based dual-detector chromatic confocal probe developed in the previous study by the authors [3]. A mode-locked femtosecond laser (C-Fiber, MenloSystems) with a repetition rate  $\nu_{rep}$  of 100 MHz and a spectral range of 1480–1640 nm was employed as the light source. Single-mode step index fibers were employed for the light propagation and for the detectors together with an optical spectrum analyzer (AQ6370D, Yokogawa Electric Corp.). The laser beam from the laser source was transmitted through the single-mode step index fiber, and the output laser beam from the fiber was then collimated by a collimating lens to generate a collimated beam with a beam diameter of approximately 3.6 mm. The collimated laser beam was then made to pass through a polarized beam splitter (PBS) and a quarter wave plate (QWP). By utilizing a chromatic objective (Edmund Optics, #67-484), the collimated laser beam was focused on a plane mirror mounted on a one-axis piezoelectric (PZT) positioning stage. The combination of the PBS and the QWP made sure that the light source was effectively isolated from the laser beam reflected from the plane mirror. A capacitance displacement sensor, whose measurement axis was aligned to coincide with the laser axis, was located behind the plane mirror to measure the target position of the mirror along the Z-direction. The propagating direction of the reflected laser beam was then bent 90 degrees by the PBS and was divided into two sub-beams by a non-polarizing beam splitter (BS). By utilizing an achromatic lens L1 with a focal length  $f$  of 16.6 mm, one of the sub-beams (sub-beam1) was coupled into one end face of a single-mode step index fiber, having a core diameter of 8.2  $\mu\text{m}$  and a numerical aperture (NA) of 0.14. The other end face of the fiber was connected to the optical spectrum analyzer. Meanwhile, another sub-beam (sub-beam2) was coupled into one end face of another single-mode step index fiber, which was identical to that for the sub-beam1, by using another achromatic lens L2 identical to L1. It should be noted that the end face of the fiber for sub-beam2 was mounted on a precision manual stage system to give a defocus  $d$  to the end face of the fiber, with respect to the back focal plane of L2.



**Figure 7.** A schematic of the setup for the dual-detector chromatic confocal probe with a mode-locked femtosecond laser source.

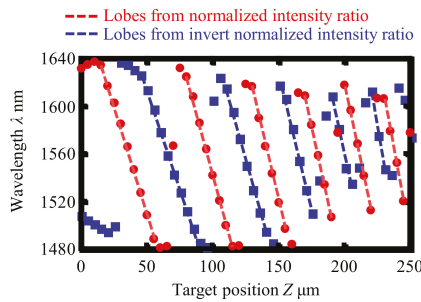
### 3.2. Experiments

Experiments were carried out to investigate the feasibility of the proposed measurement range expansion method. The plane mirror was made to move along the Z-direction in a step of 2  $\mu\text{m}$  to obtain the axial responses. It should be noted that the displacement of the plane mirror was measured by a capacitive displacement sensor simultaneously for comparison. Peak wavelengths were calculated from the centroid positions of the corresponding normalized intensity ratio, as well as the invert normalized intensity ratio at each Z-position of the mirror. Figure 8a shows the normalized intensity ratio obtained at each Z-position over a Z-directional scanning range of 250  $\mu\text{m}$ . As can be seen in Figure 8a, not only the main-lobe but also the side-lobes were clearly observed. Although the peak widths of the sub-lobes were found to become wider as the increase of the Z-displacement of the plane mirror, the sharpness of the sub-lobes was enough to be used in the wavelength-to-displacement encoding. Figure 8b shows the peak wavelength obtained from the normalized intensity ratio in Figure 8a at each Z-directional position of the plane mirror. As can be seen in the figure, a linear relationship between the peak wavelength and the axial Z-directional displacement of the plane mirror was found for the main-lobe and each of the side-lobes. Meanwhile, as predicted in the results of the above theoretical calculations, a gap between each of the neighboring lobes was found. Figure 8c shows the invert normalized intensity ratio obtained at each Z-position, which was calculated from the same intensities employed to calculate the corresponding normalized intensity ratio shown in Figure 8a. As can be seen in Figure 8c, the main-lobe and the side-lobes were clearly observed in the obtained invert normalized intensity ratio. In the same manner, as the case of normalized intensity ratio  $I_a$ , the relationship between the peak wavelength and the axial Z-directional displacement of the flat mirror was investigated by extracting the peak wavelengths in the invert normalized intensity ratio  $I_n$  shown in Figure 8c. The result is summarized in Figure 8d. As predicted in the results of the above theoretical calculations, the peak frequency in the invert normalized intensity ratio at a specific Z-directional position of the plane mirror was found to shift along the axial direction with respect to that of the normalized intensity ratio shown in Figure 8b.



**Figure 8.** The main-lobes and the side-lobes extracted from the normalized intensity ratio and the invert normalized intensity ratio obtained at each Z-directional position of the plane mirror, with target positions ranging from 0 to 250  $\mu\text{m}$  and a defocus  $d$  of 110  $\mu\text{m}$ . (a) Normalized intensity ratio obtained at each Z-position over scanning range of 250  $\mu\text{m}$ . (b) The lobes extracted from the normalized intensity ratio. (c) Invert normalized intensity ratio obtained at each Z-position over scanning range of 250  $\mu\text{m}$ . (d) The lobes extracted from the invert normalized intensity ratio.

Figure 9 summarizes the wavelength-to-displacement encoding curves obtained from the axial responses of the normalized intensity ratio  $I_a$  and that of the invert normalized intensity ratio  $I_n$ . As can be seen in the figure, a gap between each neighboring lobe in the axial response of the normalized intensity ratio was successfully covered by the axial response from the invert normalized intensity ratio. These experimental results demonstrated that measurement range of the fiber-based dual-detector chromatic confocal probe was successfully expanded from 40  $\mu\text{m}$  to 250  $\mu\text{m}$  by the proposed method without any modification on the optical setup of the developed confocal probe.



**Figure 9.** Wavelength-to-displacement encoding curves obtained from the axial responses of the normalized intensity ratio and that of the invert normalized intensity ratio, with target positions ranging from 0 to 250  $\mu\text{m}$  and a defocus  $d$  of 110  $\mu\text{m}$ .

#### 4. Conclusions

A new method for expanding the Z-directional measurement range of a fiber-based dual-detector chromatic confocal probe with a mode-locked femtosecond laser source has been proposed. The dual-detector chromatic confocal probe has been designed to detect the axial position of a measurement target by using the normalized intensity ratio calculated from light intensities obtained by the two identical fiber detectors. In the proposed method, for the expansion of the measurement range of the dual-detector chromatic confocal probe, invert normalized intensity ratio, which is the inverse of the normalized intensity ratio, has also been employed to realize seamless measurement with the main-lobes and the side-lobes of the axial responses obtained from both the normalized intensity ratio and the invert normalized intensity ratio. Theoretical calculations and experiments have been carried out to verify the feasibility of the proposed method, and the results have demonstrated that the measurement range can be expanded to 250  $\mu\text{m}$ , which is four times wider than that of the conventional dual-detector chromatic confocal probe developed in the previous work by the authors.

It should be noted that this paper is mainly focused on the proposal of the method for expansion of the Z-directional measurement, and subsequent experimental work will be carried out in 3D microstructure imaging in our future research. It should also be pointed out that in order to make good use of the side-lobes to extend the Z-directional measurement range, we set a defocus  $d$  to be 110  $\mu\text{m}$ , which achieves a bigger full width at half maximum (FWHM) value compared with the FWHM value when the defocus  $d$  is set to be 150  $\mu\text{m}$  in our previous work [3]. That means there will be a reduction in axial resolution. However, when the defocus  $d$  is fixed, the axial resolution will not be affected during the measurement range expansion processing because our expansion method is performed without modifying the optical setup of the dual-detector chromatic confocal probe, and therefore will not change the point spread function of our confocal imaging system theoretically. Furthermore, future research will also focus on the depth measurement provided by the optical phase and optical complex amplitude of the dual-detector chromatic confocal probe, compared with the depth measurement given by the optical intensity in this paper. Measurement uncertainty analysis together with the identification of the absolute optical wavelengths based on the optical frequency comb will also be carried out in the next stage of research.

**Author Contributions:** Conceptualization, W.G.; methodology, C.C. and T.N.; software, C.C. and T.N.; validation, C.C., H.M., and R.S.; formal analysis, C.C., T.N., and R.S.; investigation, C.C., T.N., and R.S.; resources, C.C. and R.S.; data curation, C.C. and R.S.; writing—original draft preparation, C.C.; writing—review and editing, Y.S., H.M., and C.C.; visualization, W.G. and Y.S.; supervision, W.G.; project administration, W.G.; funding acquisition, W.G. and Y.S.

**Funding:** This research was funded by Japan Society for the Promotion of Science (JSPS).

**Conflicts of Interest:** The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, and in the decision to publish the results.

#### References

1. Gao, W.; Kim, S.W.; Bosse, H.; Haitjema, H.; Chen, Y.L.; Lu, X.D.; Knapp, W.; Weckenmann, A.; Estler, W.T.; Kunzmann, H. Measurement technologies for precision positioning. *CIRP Ann. Manuf. Technol.* **2015**, *64*, 773–796. [[CrossRef](#)]
2. Butler, D.; Horsfall, A.; Hrynevych, M.; Kearney, P.D.; Nugent, K.A. Confocal profilometer with nanometric vertical resolution. *Opt. Commun.* **1993**, *100*, 87–92. [[CrossRef](#)]
3. Chen, X.G.; Nakamura, T.; Shimizu, Y.; Chen, C.; Chen, Y.L.; Matsukuma, H.; Gao, W. A chromatic confocal probe with a mode-locked femtosecond laser source. *Opt. Laser Technol.* **2018**, *103*, 359–366. [[CrossRef](#)]
4. Chen, L.C.; Chang, Y.W. Innovative simultaneous confocal full-field 3D surface profilometry for in situ automatic optical inspection (AOI). *Meas. Sci. Technol.* **2013**, *21*, 1–12.

5. Bohn, S.; Sperlich, K.; Allgeier, S.; Bartschat, A.; Prakasam, R.; Reichert, K.M.; Stolz, H.; Guthoff, R.; Mikut, R.; Kohler, B.; et al. Cellular in vivo 3D imaging of the cornea by confocal laser scanning microscopy. *Biol. Opt. Express* **2018**, *9*, 2511–2525. [[CrossRef](#)] [[PubMed](#)]
6. Stephens, D.J.; Allan, V.J. Light Microscopy Techniques for Live Cell Imaging. *Science* **2003**, *300*, 82–86. [[CrossRef](#)] [[PubMed](#)]
7. Hickey, P.C.; Swift, S.R.; Roca, M.G.; Read, N.D. Live-cell imaging of filamentous fungi using vital fluorescent dyes and confocal microscopy. *Methods Microbiol.* **2004**, *34*, 63–87.
8. Gao, W.; Araki, T.; Kiyono, S.; Okazaki, Y.; Yamanaka, M. Precision nano-fabrication and evaluation of a large area sinusoidal grid surface for a surface encoder. *Precis. Eng.* **2003**, *27*, 289–298. [[CrossRef](#)]
9. Shimizu, Y.; Aihara, R.; Mano, K.; Chen, C.; Chen, Y.L.; Chen, X.; Gao, W. Design and testing of a compact non-orthogonal two-axis Lloyd’s mirror interferometer for fabrication of large-area two-dimensional scale grating. *Precis. Eng.* **2018**, *52*, 138–151. [[CrossRef](#)]
10. Berkovic, G.; Shafira, E.; Golubb, M.A.; Brill, M.; Shurmanb, V. Multi-wavelength fiber-optic confocal position sensor with diffractive optics for enhanced measurement range. *Proc. SPIE* **2007**, *6619*, 66190U.
11. Dobson, S.L.; Sun, P.C.; Fainman, Y. Diffractive lenses for chromatic confocal imaging. *Appl. Opt.* **1997**, *36*, 4744–4748. [[CrossRef](#)] [[PubMed](#)]
12. Shafir, E.; Berkovic, G. Expanding the realm of fiber optic confocal sensing for probing position, displacement, and velocity. *Appl. Opt.* **2006**, *45*, 7772–7777. [[CrossRef](#)] [[PubMed](#)]
13. Tiziani, H.J.; Achi, R.; Kramer, R.N. Chromatic confocal microscopy with microlenses. *J. Mod. Opt.* **1995**, *43*, 155–163. [[CrossRef](#)]
14. Garzon, R.J.; Meneses, J.; Tribillon, G.; Gharbi, T.; Plata, A. Chromatic confocal microscopy by means of continuum light generated through a standard single-mode fibre. *J. Opt. A Pure Appl. Opt.* **2004**, *6*, 544–548.
15. Zhao, W.Q.; Tan, J.B.; Qiu, L.R. Bipolar absolute differential confocal approach to higher spatial resolution. *Opt. Exp.* **2004**, *12*, 5013–5021. [[CrossRef](#)]
16. Lin, D.J.; Liu, Z.Y.; Zhang, R.; Yan, J.Q.; Yin, C.Y.; Xu, Y. Step-height measurement by means of a dual-frequency interferometric confocal microscope. *Appl. Opt.* **2004**, *43*, 1472–1479. [[CrossRef](#)] [[PubMed](#)]
17. Zhao, W.Q.; Tan, J.B.; Qiu, L.R. Improvement of confocal microscope performance by shaped annular beam and heterodyne confocal techniques. *Optik* **2005**, *116*, 111–117. [[CrossRef](#)]
18. Tan, J.B.; Liu, J.; Wang, Y.H. Differential confocal microscopy with a wide measuring range based on polychromatic illumination. *Meas. Sci. Technol.* **2010**, *21*, 1–6. [[CrossRef](#)]
19. Shi, K.B.; Nam, S.H.; Li, P.; Yin, S.Z.; Liu, Z.W. Wavelength division multiplexed confocal microscopy using supercontinuum. *Opt. Commun.* **2006**, *263*, 156–162. [[CrossRef](#)]
20. Zhao, W.Q.; Jiang, Q.; Qiu, L.R.; Liu, D.L. Dual-axes differential confocal microscopy with high axial resolution and long working distance. *Opt. Commun.* **2011**, *284*, 15–19. [[CrossRef](#)]
21. Sheppard, C.J.R.; Gu, M. Aberration compensation in confocal microscopy. *Appl. Opt.* **1991**, *30*, 3563–3568. [[CrossRef](#)] [[PubMed](#)]
22. Martinez-Corral, M.; Caballero, M.T.; Pons, A.; Andres, P. Sidelobe decline in single-photon 4Pi microscopy by Toraldo rings. *Micron* **2003**, *34*, 319–325. [[CrossRef](#)]
23. Suhre, D.R.; Gupta, N. Acousto-optic tunable filter sidelobe analysis and reduction with telecentric confocal optics. *Appl. Opt.* **2005**, *47*, 5797–5801. [[CrossRef](#)]
24. Segawa, S.; Kozawa, Y.; Sato, S. Resolution enhancement of confocal microscopy by subtraction method with vector beams. *Opt. Lett.* **2014**, *39*, 3118–3121. [[CrossRef](#)]
25. Nagorni, M.; Hell, S.W. 4Pi-Confocal microscopy provides three-dimensional images of the microtubule network with 100- to 150-nm resolution. *J. Struct. Biol.* **1998**, *123*, 236–247. [[CrossRef](#)] [[PubMed](#)]
26. Cox, I.J.; Sheppard, C.J.R.; Wilson, T. Improvement in resolution by nearly confocal microscopy. *Appl. Opt.* **1982**, *21*, 778–781. [[CrossRef](#)] [[PubMed](#)]
27. Brakenho, G.J.; Blom, P.; Barend, P. Confocal scanning light microscopy with high aperture immersion lenses. *J. Microsc.* **1979**, *117*, 219–232. [[CrossRef](#)]
28. Neil, M.A.A.; Juskaitis, R.; Wilson, T.; Laczik, Z.J. Optimized pupil-plane filters for confocal microscope point-spread function engineering. *Opt. Lett.* **2000**, *25*, 245–247. [[CrossRef](#)]
29. Wang, T.D.; Contag, H.C.; Mandella, M.J.; Chan, N.Y.; Kino, G.S. Dual-axes confocal microscopy with post-objective scanning and low-coherence heterodyne detection. *Opt. Lett.* **2003**, *28*, 1915–1917. [[CrossRef](#)]

30. Gweon, D.G.; Daejeon.; Kang, D.K. Confocal Self-Interference Microscopy from Which Side-Lobe Has Been Removed. U.S. Patent No. 7324.273 B2, 29 January 2008.
31. Stelzer, E.H.K.; Lindek, S. Fundamental reduction of the observation volume in far-field light microscopy by detection orthogonal to the illumination axis: Confocal theta microscopy. *Opt. Commun.* **1994**, *111*, 536–547. [[CrossRef](#)]
32. Jones, D.; Diddams, S.A.; Ranka, J.K.; Stentz, A.; Windeler, R.S.; Hall, J.L.; Cundiff, S.T. Carrier-envelope phase control of femtosecond mode-locked lasers and direct optical frequency synthesis. *Science* **2000**, *288*, 635–639. [[CrossRef](#)] [[PubMed](#)]
33. Chen, Y.L.; Shimizu, Y.; Kudo, Y.; Ito, S.; Gao, W. Mode-locked laser autocollimator with an expanded measurement range. *Opt. Exp.* **2016**, *24*, 15554–15569. [[CrossRef](#)] [[PubMed](#)]
34. Shimizu, Y.; Kudo, Y.; Chen, Y.L.; Ito, S.; Gao, W. An optical lever by using a mode-locked laser for angle measurement. *Precis. Eng.* **2017**, *47*, 72–80. [[CrossRef](#)]
35. Wilson, T.; Sheppard, C. *Theory and Practice of Scanning Optical Microscopy*; Academic Press: London, UK, 1984; pp. 17–20.
36. Wilson, T.; Carlini, A.R. Size of the detector in confocal imaging systems. *Opt. Lett.* **1987**, *12*, 227–229. [[CrossRef](#)] [[PubMed](#)]
37. Kimura, S.; Wilson, T. Confocal scanning optical microscope using single-mode fiber for signal detection. *Appl. Opt.* **1991**, *30*, 2143–2149. [[CrossRef](#)] [[PubMed](#)]




© 2019 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).



Article

# Double-Diffracted Spatially Separated Heterodyne Grating Interferometer and Analysis on its Alignment Tolerance

Di Chang <sup>1,2,†</sup>, Xu Xing <sup>1,2,†</sup>, Pengcheng Hu <sup>1,2,\*</sup> , Jianing Wang <sup>1,2</sup> and Jiubin Tan <sup>1,2</sup>

<sup>1</sup> School of Instrumentation Science and Engineering, Harbin Institute of Technology, Harbin 150080, China; 16B901026@stu.hit.edu.cn (D.C.); da-lucifer@outlook.com (X.X.); 18S001057@stu.hit.edu.cn (J.W.); jbtan@hit.edu.cn (J.T.)

<sup>2</sup> Key Lab of Ultraprecision Intelligent Instrumentation, Ministry of Industry and Information Technology, Harbin 150080, China

\* Correspondence: hupc@hit.edu.cn; Tel.: +86-451-86412041-803

† These authors contribute equally to this work.

Received: 5 December 2018; Accepted: 8 January 2019; Published: 13 January 2019



**Abstract:** An optical configuration of double-diffracted spatially separated heterodyne grating interferometer with a mechanical fixture was designed. To further investigate its features and provide robust measurements, the alignment tolerance in double-diffracted spatially separated heterodyne grating interferometer was qualitatively and quantitatively analyzed. Except for the offset error causing no influence on the interfering signal, the effect of the other four errors, roll, yaw, pitch angles, and stand-off error were geometrically analyzed and mathematically modeled. The simulation result quantified the position mismatches of output beams in a double-diffracted configuration and found the crucial structural parameters related to the intensity of interfering signals. Experiments based on the grating interferometer with a mechanical fixture and the same optical configuration built by independent optical components were implemented, whose results agreed with the simulation. Besides, the results showed that the proposed grating interferometer structure could tolerate the  $\pm 1100$  arcsec roll movement,  $\pm 440$  arcsec yaw movement,  $\pm 280$  arcsec pitch movement, and  $\pm 0.6$  mm stand-off error when -10 dB intensity loss is afforded.

**Keywords:** grating interferometer; laser encoder; spatially separated heterodyne interferometry; alignment tolerance

## 1. Introduction

Nowadays, development in modern technology and industry requires more precise displacement measurement and nanopositioning sensors. Taking semiconductor manufacturing as an example, because of the decreasing node size, measuring the spatial position and attitude of a wafer stage in lithography will be a challenge for high accuracy and stability [1].

The trend towards a stable sub-nanometer accuracy measurement propels grating interferometer, a substitution for a laser interferometer, to be the solution of displacement measurement in a lithography machine. Advantages of the grating interferometer mainly have two aspects.

One is the resistance to ambient disturbance, especially to the fluctuation of the air refractive index, which makes it convenient to achieve accuracy in the sub-nanometer scale. Referring to the work of Gao et al., the measuring uncertainty of a laser interferometer is mainly caused by air refractive index error and wavelength error, while the influences of a grating interferometer are manufacturing errors and thermal expansion of the scale [2]. Compared with those influencing factors in a laser interferometer, the manufacturing errors and thermal expansion of the grating are easier to control by

the use of zero thermal expansion materials and the development of a fabrication technique of gratings. Recently, Deng et al. achieved linear displacement measurement with 68.6 pm resolution by using eightfold optical construction equipping a 1780 lines/mm high-density grating [3].

Another advantage is that the utilization of optical components such as large planar gratings makes grating interferometer easy to implement an additional measuring dimension. Considering a measured rigid body, its spatial degree of freedoms (DOF) include three translational DOFs and three rotational DOFs. The grating interferometer has been widely used in multi-DOF metrology systems. For instance, Huang et al. proposed a 3-DOF homodyne laser encoder using grating interferometry and quadrant detectors (QD) to measure movement along the x-direction and rotation around the x-axis and y-axis; the angular measuring ranges were  $\pm 100$  arcsec [4]. Lee et al. designed a similar 6-DOF optical encoder with  $\pm 200$  arcsec angular measuring ranges [5]. Li et al. reported researches on 6-DOF homodyne surface encoders in a multi-probe all-interferometric configuration and a grating-QD combination, the angular measuring ranges of which were over 60 arcsec [6,7]. The limitation of these ranges mentioned above could be attributed to angular mismatches—the beam misalignment caused by the rotations of the grating.

As a common method to attain a high fold factor, the double-diffraction structure also has the feature of effectively enhancing the alignment tolerance (also written head-to-scale tolerance in some references) [8–10]. Feng et al. proposed a 2-DOF double-diffracted heterodyne grating interferometer in common optical path configuration. However, the specific alignment tolerance was not mentioned [9]. Because the output beams are parallel with the incident beams, the double-diffraction structure is capable of turning the angular mismatches into the position mismatches—the deviations of output beams. However, it is worth emphasizing that the position mismatches will still cause intensity loss of interfering beams so that further analyses still need to be investigated.

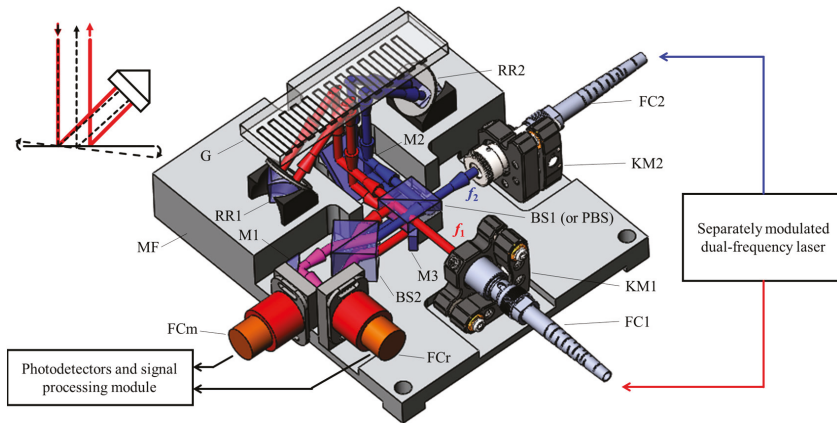
Currently, as a newly proposed concept, spatially separated heterodyne grating interferometer, which focuses on eliminating the periodic nonlinear errors from several to tens of nanometers, has become an effective approach to sub-nanometer scale measurement accuracy [11–13]. Compared with a typical common optical path grating interferometer, being spatially separated means it has two parallel modulated laser beams with frequency differences. To further improve its performance and lay a foundation for future multi-DOF measurements, the alignment tolerance of spatially separated heterodyne grating interferometer was analyzed on the structure with a mechanical fixture of our design. Qualitative analysis, quantitative simulations and experiments are implemented and described in the following sections.

## 2. Optical Configuration and Mechanical Fixture of the Proposed Grating Interferometer

Based on the concept of spatially separated heterodyne grating interferometer reported in our previous study [12], the optical configuration of the proposed grating interferometer is shown in Figure 1. Considering the real sizes of all components, several changes and adjustments were made to form a compact layout.

From a typical separately modulated heterodyne laser source, beams at different frequencies were transferred into the optical reading head via polarization maintaining optical fibers. Instead of a parallel distribution, two fiber collimators (FC1 and FC2) in vertical distribution guaranteed enough space for mounting and adjusting. The beams were divided into two parts, the reference part and measurement part, in the beam splitter (BS1). Then, the reference beam was combined in the upper part of the beam splitter (BS2), reflected at the surface of the mirror (M1), entered into the fiber coupler (FCr), and transferred to the corresponding photodetector by a multi-mode optical fiber. Another part of the beams reflected at the mirror (M2) to change directions for a compact layout and diffracted by the linear grating (G) at a normal incident angle. In the diffraction, called the first diffraction here, each of the beams generated a pair of  $\pm 1^{\text{st}}$  order diffracted beams following their diffraction angles. Choosing a  $+1^{\text{st}}$  order and a  $-1^{\text{st}}$  order diffracted beam with different modulated frequencies  $f_1$  and  $f_2$  propagating symmetrically, two corner-cube retro-reflectors (RR1 and RR2) were replaced to cause

a U-turn of the chosen beams. Then, following the grating equation, these beams entered the grating at oblique incident angles and were diffracted again. To be unambiguous, this diffraction was called the second diffraction. Obviously, only two of all the diffracted beams spread away from the grating vertically, with fourfold Doppler frequency shifts. It is worth emphasizing that the retro-reflectors guarantee that the selected diffracted beams of the second diffraction are parallel with the incident beams of the first diffraction when unexpected tip and tilt of grating occur. After being reflected in proper order by mirrors M2 and M3, the measurement beam was generated in the lower part of BS2. Similar to the reference beam, it entered the fiber coupler (FCm) and transferred to its photodetector by another multi-mode optical fiber. Displacement information, contained in Doppler frequency shifts, was detected by photodetectors and extracted from electrical signals by the processing board.



**Figure 1.** Optical configuration and mechanical fixture of the proposed double-diffracted spatially separated heterodyne grating interferometer. FC1 and FC2, optical fiber collimators; BS1 and BS2, non-polarized beam splitters; PBS, polarized beam splitter; M1, M2, and M3, mirrors; RR1 and RR2, retroreflectors; G, linear grating; FCr and FCm, optical fiber couplers; KM1 and KM2, kinematic mounts; MF, mechanical fixture. The colored markers  $f_1$  and  $f_2$  represent different modulated frequencies.

Compared to our previous study [13], the biggest improvement in optical configuration was substituting a polarized beam splitter (PBS) and half wave plate by the beam splitter BS1 and the mirror M3. The advantages were that not only that less beam splitter was used, but also that it was a polarized-insensitive optical configuration. In particular, it could be configured into a polarized or non-polarized interferometer by using a PBS or a BS at the place of BS1, decreasing the requirement of the laser source.

The mechanical fixture is also displayed in Figure 1. As shown, a pedestal with holes for mounting is designed for supporting all optical and mechanical components. Two optical fiber collimators were fixed on matched kinematic mounts fastening onto the pedestal. The parallelism of beams was adjusted to several arc-seconds by kinematic mounts so that the assembling error could be omitted in simulations and experiments below. Optical components, including beam splitters, mirrors and corner-cube retro-reflectors, were glued on the pedestal by ultraviolet curable adhesive. Moreover, two optical fiber couplers mounted on custom mechanical components (red and orange columns) were also glued onto the pedestal. The grating interferometer assembly was convenient for mounting on moving stages or linear guide rails.

### 3. Analysis of Alignment Tolerance on the Double-Diffracted Grating Interferometer

When a relative movement occurred between the fixed grating and moving interferometer, rotation angles of the interferometer would not be constant, but would fluctuate in a small range.

The angles, distinguished by roll, yaw and pitch from the axes they rotate about, are illustrated in Figure 2. Besides, the slight translation in y and z directions, are defined as offset and stand-off errors respectively [14].

Also indicated by Figure 2, it is easy to visualize that the offset error is uninfluential as long as the spots remain in the ruling area. Actually, in a linear or planar grating interferometer, the translational movements along x and y directions do not cause errors. However, errors in the other four DOFs will worsen the interference signal as their values increase.

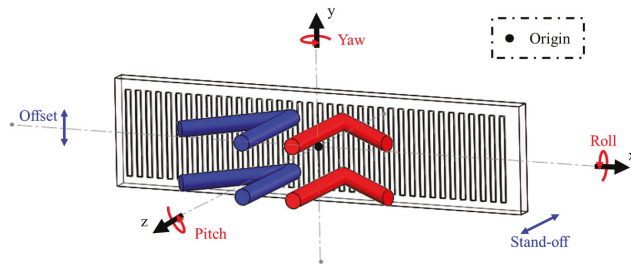


Figure 2. Distinguishing the misalignment angles (roll, yaw, and pitch), stand-off and offset.

### 3.1. Analysis of the Stand-off Error

As Figure 3 shows, the stand-off error changes the distances between the grating and the reading-head while keeping the incident beams vertical. The diffracted beams of the first diffraction influenced by stand-off (shown in red or blue) and the original position of the diffracted beams (shown in transparent white) are parallel but with a deviation. The diffracted beams entered the retro-reflectors respectively, creating spot offsets in the surfaces of the RRs. Considering the center symmetry feature of corner-cube retro-reflectors [15], the output beams, as the oblique incident beams of the second diffraction, follow the offset and propagate in the opposite directions. Finally, the second diffraction beams vertically transmit away from the grating with symmetrical deviation. Qualitatively, the bigger stand-off error is, the larger the deviation that occurs. After the second diffraction, in a beam-combined prism implemented by a NPBS in the 45° position, the beams with symmetrical deviation can still entirely overlap (but with a deviation). Only when the actual interfering zone exceeds the boundary of the PD, can the stand-off error have no effect on the interference signal.

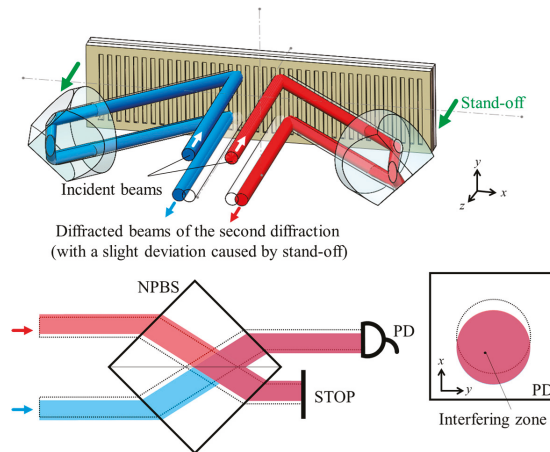
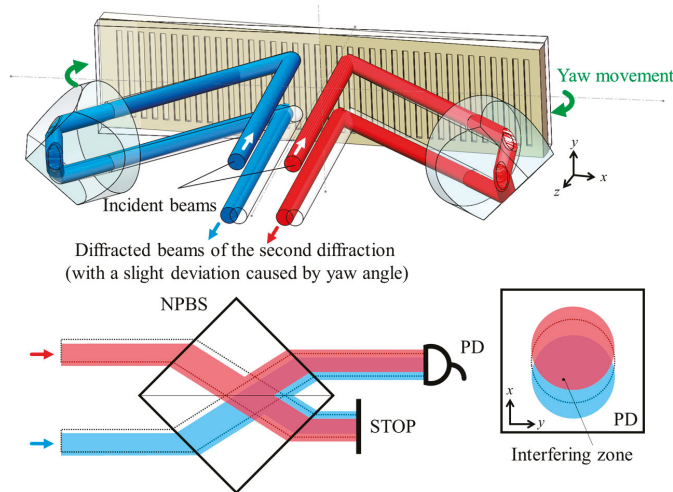


Figure 3. Geometrical analysis of the stand-off error and its influence on the measurement beam caused by spots deviation.

### 3.2. Analysis of the Roll, Yaw, and Pitch Angles

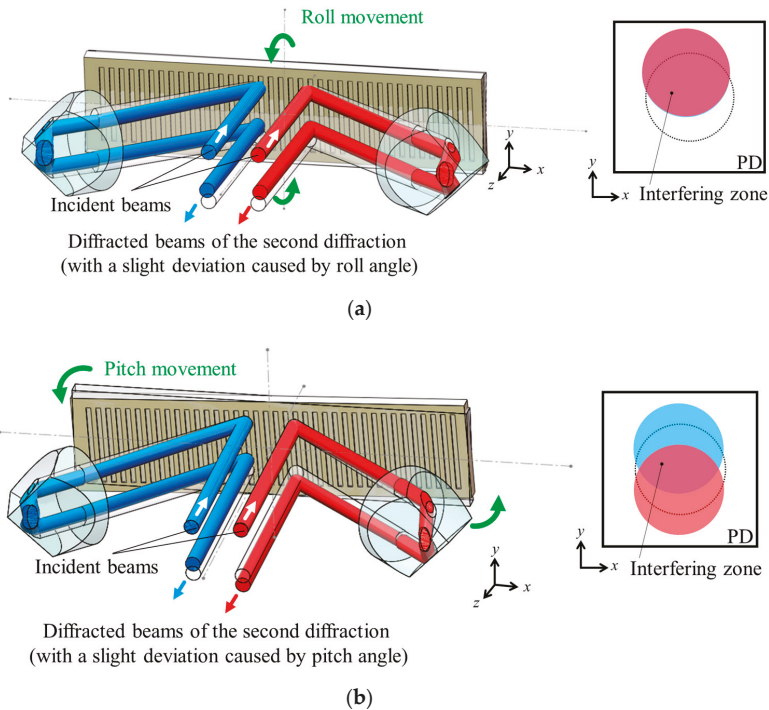
Compared to the translational errors discussed above, these three rotational errors are more complex with greater impact. Different from translational movement, rotation involves a rotating center. However, in the configuration of a spatially separated grating interferometer, more than one beam exists. Thus, a rotation center at the origin shown in Figure 2 is selected for the following analysis.

As Figure 4 shows, when a small yaw movement occurs, the incident angle will no longer be normal. Although the difference is slight, according to the grating equation, the absolute value of diffraction angles between +1<sup>st</sup> and -1<sup>st</sup> order diffracted beams are unequal. Similarly, the diffracted beams entered the retro-reflectors along their new angles respectively, creating offsets of spots on the surfaces of the RRs. After the second diffraction, deviations appear at the vertically transmitted beams. Through a beam-combined BS, the output measurement beams has a certain work-off. Although the deviations of the beams are in the same direction, the spots on the surface of PD move away from each other, resulting in a decreasing interference zone.



**Figure 4.** Geometrical analysis of the yaw movement and its influence on the measurement beam caused by spots deviation. (To be clear, the angles in the figure and their influence are exaggerated. Generally, these angles are no larger than several mrad or hundreds of arcseconds in actuality).

Similarly, the geometrical analyses and influences of roll and pitch angles are displayed in Figure 5. In these circumstances, the diffraction angle remains while the propagating directions are no longer parallel to the x-O-z plane of the coordinate. Caused by roll motion, the beams deviate in a similar direction in Figure 5a. However, the pitch motion leads to opposite deviations of the two beams in Figure 5b, resulting in a work-off of these two spots.



**Figure 5.** Geometrical analysis of the roll and pitch movements and their influences on the measurement beam caused by spots deviation. (a) Roll movement; (b) Pitch movement.

## 4. Simulations and Experiments

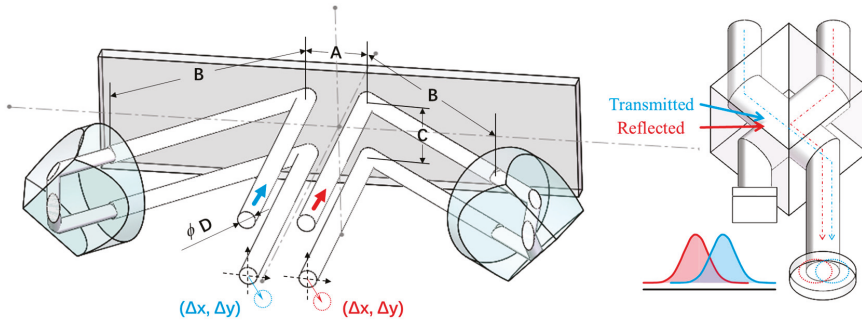
### 4.1. Simulations

To quantitatively describe the alignment tolerance, simulation codes for beam tracing were programmed with MATLAB. In the mathematical model, the O-x-y-z coordinate, optical interfaces, and vectors of incident beams were defined. The range of grating and apertures of the RRs are supposed to be infinite in the model, so that the influence and errors caused by out of clear apertures are not discussed.

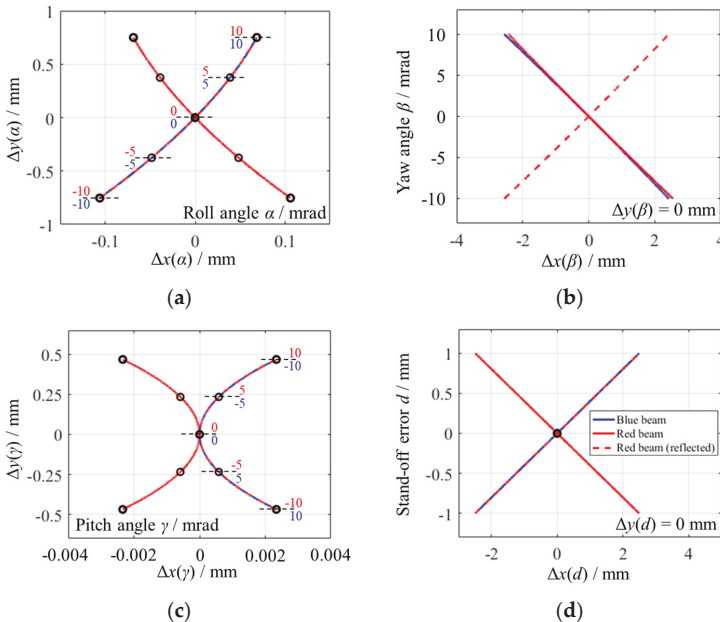
In Figure 6, four crucial structural parameters of the spatially separated heterodyne grating interferometer are marked. Specifically, A is the separating distance of the incident beams, B represents the lengths of diffracted beams from planar grating to the surfaces of RRs, C is the offset between the input and the returned output beams of RRs, and D is the beam diameter. Since the divergence of beams for interferometers are generally small (within a couple of mrad/hundreds of arcsec), only the Gaussian distribution of the intensity is considered while the divergence is ignored, which implies that the Gaussian beam seems to be a column in the simulation. The diameter D is assumed as the beam waist of the Gaussian beam. Therefore, the optical structure showing in Figure 1 could be digitized as  $A = 7$  mm,  $B = 30$  mm,  $C = 7$  mm, and  $D = 3$  mm. The influence of the stand-off error and three misalignment angles to such a structure are simulated; these results are demonstrated in Figure 7.

Figure 7 shows that the deviations between the actual and the ideal centers of the spots were continuously enlarged as the errors increased. The blue lines represent the positions of the blue spot in Figure 6. Similarly, the red solid lines depict the deviations of the red diffracted beam in Figure 6, and the red dashed lines describe the spot positions after reflection by the beam-combined BS. Because the roll and pitch angles created the planar movements of the beam centers, their curves in the x-O-y

plane with contour lines of the angles are portrayed in Figure 7a,c. Blue numbers and red numbers represent the angles of the blue lines and dashed red lines, respectively. However, only one-dimensional movements of the beams in the x-direction were caused by the yaw angle and the stand-off error; Figure 7b,d depicts the relationship in curves with an explanation  $\Delta y = 0$  mm. It is worth mentioning that the simulation results agree with the qualitative analysis in Chapter 3. Besides, the beam-combined BS placed at an angle of  $45^\circ$  could reduce the separation caused by the stand-off error and roll angle, but also deteriorate the influence of yaw and pitch angles. The work-off distances of the reflected red beams and the blue beams can be extracted from the distances at the same angles or errors between the corresponding lines.



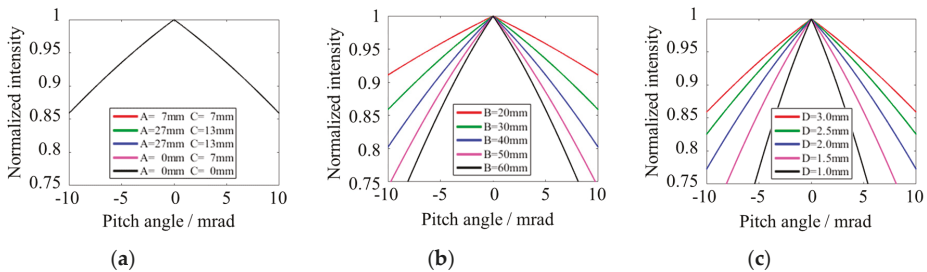
**Figure 6.** Mathematical model and the crucial structural parameters of the spatially separated heterodyne grating interferometer and the interfering intensity of the Gaussian beams.



**Figure 7.** The results of beam tracing with the variation of misalignment angles and the stand-off error shown by the curves of the beam centers. (a) Roll angle  $\alpha$ . (b) Yaw angle  $\beta$ . (c) Pitch angle  $\gamma$ . (d) Stand-off error  $d$ . These curves share the legend shown in figure (d).

In another simulation, the relationships of four specific structural parameters from *A* to *D* and their influence to interfering beams were investigated. Taking the pitch movement as an example, the simulation above has proved that the deviation of the two spots will decrease the overlapped region of the two Gaussian beams, and the larger the work-off distance *L* is, the lower the interfering intensity will be. Supposing that the Gaussian beams follow the same equation and only the parts inside the beam waists are calculated, the ratio of the overlapped region and the maximum volume (within the beam waist) is used to appraise the influence.

In Figure 8a, all these curves overlapped together, which means the rotational arm, equaling to  $L = (A^2 + C^2)^{1/2}/2$ , has no relation to the interfering beams. The common optical path heterodyne grating interferometer is reckoned as a special case of spatially separated configuration with the parameter *A* = 0. Thus, conclusions in this paper are also suitable for the common optical path configuration. According to Figure 8b, the parameter *B*, representing the length of diffracted beams, caused more changes of the interfering beams. It could also be attributed to a longer distance which generated a larger offset at a certain rotational angle. Figure 8c clearly illustrates that a bigger beam suffers larger spot deviations.



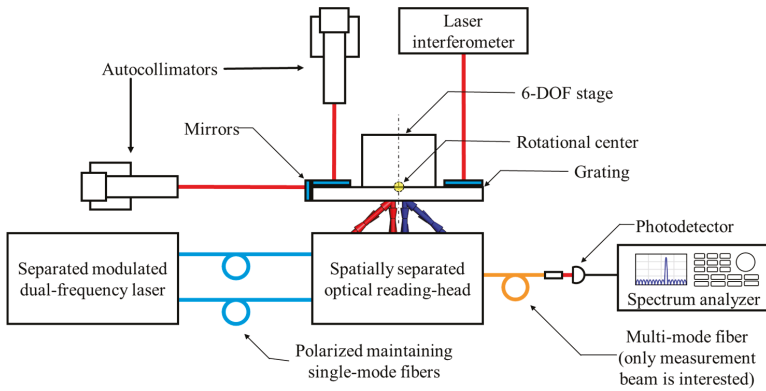
**Figure 8.** The relationship between the normalized intensity and pitch angle with different structure parameters. (a) Rotation arms are changed. (b) Lengths of the diffracted beams are changed. (c) Beam diameters are changed.

In summary, the simulation results demonstrate the influence of rotational angles and stand-off error on the deviation of measurement beams from the quantitative point of view. It also shows that the intensity of the interference signal is independent of the space between beams (parameters *A* and *C*) and is affected by the length of the diffracted beams and the spot diameter (parameters *B* and *D*).

4.2. Experiments

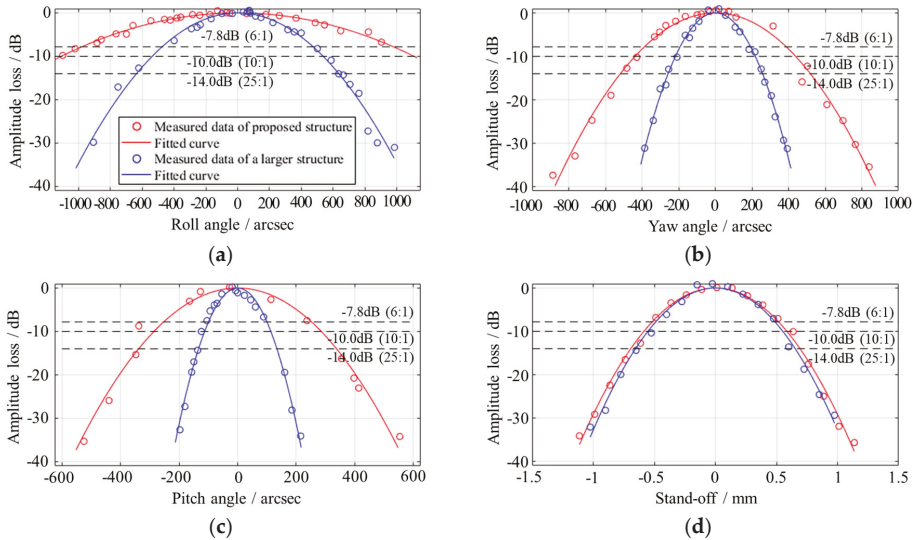
The experimental setup for testing the performance of alignment tolerance is shown in Figure 9. A fiber-delivered separated modulated laser was implemented by a single-frequency laser source with 780 nm wavelength (Sacher Lasertechnik Group TEC 500) and two acousto-optic modulators (AA Opto Electronic MT80). After transferring to the optical reading head, these beams diffracted at a custom-tailored 1 μm grating. The grating, assembling to a 6-DOF stage, could be adjusted to generate different types of errors. Autocollimators (Moeller-Wedel ELCOMAT 3000) were equipped to monitor the rotation angles, while the translational stand-off displacements were displayed by a homemade laser interferometer. Besides, it is worth emphasizing that the rotational center was located near the center of these beams. Finally, the measurement beam was delivered to the photodetector (FEMTO Messtechnik GmbH HCA-S-200M-SI) and spectrally analyzed with a spectrum analyzer (ADVANTEST R3131A). Since the experiment only focused on the measurement beam, the reference beam was not detected. The aperture of the photodiode was larger than the beam so that only the work-off of the beams were considered. As mentioned above, the parallelism of beams was adjusted by the kinematic mounts to exclude the assembling error.





**Figure 9.** Experimental setup for testing the performance of alignment tolerance.

A comparison of the proposed reading-head structure and a larger structure built with independent optical components were deployed in the experiments and are depicted in Figure 10. The plots from the former are colored red and those from the latter are blue. Since it is not clear enough to compare two sets of plots, curves were fitted to describe their trends. The type of the fitted curves was selected as the sum of two Gaussian functions, which also fits the curves in the simulation shown in Figure 8 well.



**Figure 10.** Experimental curves showing the relationships between the misalignment errors and amplitude loss of the interfering signal. (a) Roll angle. (b) Yaw angle. (c) Pitch angle. (d) Stand-off error. All four figures share the same legend as the first figure.

Considering the fluctuation of spectrum peaks and the amplitude differences of the assembled and independently built structures, Figure 10 selects the apex of the fitted curve as the origin and uses the amplitude loss as the standard for measuring the misalignment tolerance. Both in the simulations and experiments, the pitch angle leads to the most severe work-off distances and amplitude loss; the yaw angle causes a less severe deterioration, and the roll angle has the slightest effect, which proves that the work-off distance is one of the major factors of the amplitude loss. Compared with the

red curve and the blue one in Figure 10a–c, it is proved that the shorter the diffracted beams are, the stronger the tolerances to rotational angles are, which is also consistent with the simulation in Figure 8. It provides an available approach to improve the alignment tolerance, that is, reducing the size of the reading-head and the length of the diffracted beams. However, keeping large enough clear apertures while reducing the size and the beam lengths requires elaborate designs on the optical configurations. Besides, the stand-off error remains as Figure 10d depicts. As for the specific criteria of alignment tolerance, it is determined by the dynamic range of the photodetector, the ratio of the largest and smallest amplitudes can be acquired. Taking the Keysight receivers as instances, the dynamic ranges range from 6:1 to 25:1, depending on the AC/DC ratio [16]. Choosing a common ratio of 10:1 as a judgment, the designed structure stays accurate with the  $\pm 1100$  arcsec roll movement,  $\pm 440$  arcsec yaw movement,  $\pm 280$  arcsec pitch movement, and  $\pm 0.6$  mm stand-off error when  $-10$  dB intensity loss is afforded.

## 5. Conclusions

This paper analyzed the alignment tolerance of double-diffracted spatially separated heterodyne grating interferometer with a mechanical fixture of our design. The double-diffraction structure by corner-cube retro-reflectors doubles the optical fold factor and also converts the angular mismatch into a position mismatch, which improves the alignment tolerance. We analyzed the alignment tolerance from three aspects: quantitative research, qualitative simulation and experiments. The results depicted that the designed structure stays accurate with the  $\pm 1100$  arcsec roll movement,  $\pm 440$  arcsec yaw movement,  $\pm 280$  arcsec pitch movement and  $\pm 0.6$  mm stand-off error when  $-10$  dB intensity loss is afforded. Besides, the output results for different given crucial structural parameters shows that to achieve higher alignment tolerance, shorter diffracted beam length and wider laser beams are required. Since the separating distance does not affect the interfering intensity, the conclusions are also suitable for a common optical path configuration.

**Author Contributions:** D.C. established the simulation model, designed the experiments, and prepared the original draft. X.X. designed the structure and took part in the formal analysis. P.H. contributed to the conceptualization, project administration, and resources. J.W. took part in the experiments, data curation, and visualization. J.T. contributed to the supervision and funding acquisition.

**Funding:** The authors acknowledge gratefully the support of the National Natural Science Foundation of China (Project codes: 51605120, 51675138 and 61675058) and the National Science and Technology Major Project (Project code: 2017ZX02101006-005).

**Acknowledgments:** The authors cordially thank our co-workers Hongxing Yang, Haijin Fu, and Ruitao Yang for their helpful assistance and valuable suggestions.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Carballo, J.A.; Chan, W.T.J.; Gargini, P.A.; Kahng, A.B.; Nath, S. ITRS 2.0: Toward a re-framing of the semiconductor technology roadmap. In Proceedings of the 2014 IEEE 32nd International Conference on Computer Design, Seoul, South Korea, 19–22 October 2014; pp. 139–146.
2. Gao, W.; Kim, S.W.; Bosse, H.; Haitjema, H.; Chen, Y.L.; Lu, X.D.; Knapp, W.; Weckenmann, A.; Estler, W.T.; Kunzmann, H. Measurement technologies for precision positioning. *CIRP Ann.* **2015**, *64*, 773–796. [[CrossRef](#)]
3. Deng, J.; Yan, X.; Wei, C.; Lu, Y.; Li, M.; Xiang, X.; Jia, W.; Zhou, C. Eightfold optical encoder with high-density grating. *Appl. Opt.* **2018**, *57*, 2366–2375. [[CrossRef](#)] [[PubMed](#)]
4. Huang, H.-L.; Liu, C.-H.; Jywe, W.-Y.; Wang, M.-S.; Fang, T.-H. Development of a three-degree-of-freedom laser linear encoder for error measurement of a high precision stage. *Rev. Sci. Instrum.* **2007**, *78*, 066103. [[CrossRef](#)] [[PubMed](#)]
5. Lee, C.B.; Kim, G.H.; Lee, S.K. Design and construction of a single unit multi-function optical encoder for a six-degree-of-freedom motion error measurement in an ultraprecision linear stage. *Meas. Sci. Technol.* **2011**, *22*, 105901. [[CrossRef](#)]

6. Li, X.; Gao, W.; Muto, H.; Shimizu, Y.; Ito, S.; Dian, S. A six-degree-of-freedom surface encoder for precision positioning of a planar motion stage. *Precis. Eng.* **2013**, *37*, 771–781. [CrossRef]
7. Li, X.; Shimizu, Y.; Ito, T.; Cai, Y.; Ito, S.; Gao, W. Measurement of six-degree-of-freedom planar motions by using a multiprobe surface encoder. *Opt. Eng.* **2014**, *53*, 122405. [CrossRef]
8. Kao, C.-F.; Chang, C.C.; Lu, M.-H. Double-diffraction planar encoder by conjugate optics. *Opt. Eng.* **2005**, *44*, 023603.
9. Feng, C.; Zeng, L.; Wang, S. Heterodyne planar grating encoder with high alignment tolerance, especially insensitivity to grating tilts. In *Eighth International Symposium on Precision Engineering Measurement and Instrumentation, Proceedings of the SPIE, 8–11 August 2012, Chengdu, China*; SPIE Press: Bellingham, WA, USA, 2013; p. 87593L.
10. De Groot, P.J.; Badami, V.G.; Liesener, J. Concepts and geometries for the next generation of precision heterodyne optical encoders. In *Proceedings of the Annual Meeting of the American Society for Precision Engineering (ASPE 2016), Portland, OR, USA, 23–28 October 2016*; pp. 146–149.
11. Tan, J.; Hu, P.; Xing, X. Two-Degree-of-Freedom Anti-Optical Frequency Heterodyne Grating Interferometry and Measuring System. Chinese Patent CN 103,604,375A, 26 February 2014. (In Chinese)
12. Xing, X.; Chang, D.; Hu, P.; Tan, J. Spatial separated heterodyne grating interferometer for eliminating periodic nonlinear errors. *Opt. Express* **2017**, *25*, 31384. [CrossRef] [PubMed]
13. Guan, J.; Köchert, P.; Weichert, C.; Köning, R.; Siaudinyte, L.; Flügge, J. A differential interferometric heterodyne encoder with 30 picometer periodic nonlinearity and sub-nanometer stability. *Precis. Eng.* **2017**, *50*, 114–118. [CrossRef]
14. Cheng, F.; Fan, K. Linear diffraction grating interferometer with high alignment tolerance and high accuracy. *Appl. Opt.* **2011**, *50*, 4550. [CrossRef] [PubMed]
15. Liu, J.; Azzam, R.M.A. Polarization properties of corner-cube retroreflectors: Theory and experiment. *Appl. Opt.* **1997**, *36*, 1553. [CrossRef] [PubMed]
16. Keysight Technologies. High Performance Laser Interferometer Positioning for VMEbus Systems—Technical Overview. Available online: <https://literature.cdn.keysight.com/litweb/pdf/5965-1569E.pdf> (accessed on 22 October 2018).



© 2019 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).

Article

# Ultraprecision Diameter Measurement of Small Holes with Large Depth-To-Diameter Ratios Based on Spherical Scattering Electrical-Field Probing

Xingyuan Bian, Junning Cui \*, Yesheng Lu and Jiubin Tan

Center of Ultra-Precision Optoelectronic Instrument, Harbin Institute of Technology, Harbin 150080, China; bianxingyuan@hit.edu.cn (X.B.); yaleson\_l@126.com (Y.L.); jbtan@hit.edu.cn (J.T.)

\* Correspondence: cuijunning@126.com; Tel.: +86-451-8640-2258

Received: 6 December 2018; Accepted: 6 January 2019; Published: 10 January 2019



**Featured Application:** This measurement method has been successfully applied in the aviation, aerospace, and automotive industries, especially for the ultraprecision manufacturing of key parts and small structures with large aspect ratios.

**Abstract:** In order to solve the difficulty of precision measurement of small hole diameters with large depth-to-diameter ratios, a new measurement method based on spherical scattering electrical-field probing (SSEP) was developed. A spherical scattering electrical field with identical sensing characteristics in arbitrary spatial directions was formed to convert the micro gap between the probing-ball and the part being measured into an electrical signal. 3D non-contact probing, nanometer resolution, and approximate point probing—which are key properties for high measurement precision and large measurable depth-to-diameter ratios—were achieved. A specially designed hole diameter measuring machine (HDMM) was developed, and key techniques, including laser interferometry for macro displacement measurement of the probe, multi-degree-of-freedom adjustment of hole attitude, and measurement process planning, are described. Experiments were carried out using the HDMM and a probing sensor with a  $\phi 3$ -mm probing ball and a 150-mm-long stylus to verify the performance of the probing sensor and the measuring machine. The experimental results indicate that the resolution of the probing sensor was as small as 1 nm, and the expanded uncertainty of measurement result was  $0.2 \mu\text{m}$  ( $k = 2$ ) when a  $\phi 20$ -mm ring gauge standard was measured.

**Keywords:** small hole diameter; depth-to-diameter ratio; spherical scattering electrical-field probing; hole diameter measuring machine

## 1. Introduction

Deep and small holes with diameters in the range of 1–10 mm and large depth-to-diameter ratios are more and more widely used in the aviation, aerospace, and automotive industries [1]. The depth-to-diameter ratio of these precisely manufactured small holes often reaches a value up to several tens, even several hundred [2,3]. These deep and small holes present great challenges to the measurable depth-to-diameter ratio and the precision of existing measurement methods.

A contact probe can go deep into small holes with a long stylus or extension rod. However, stylus bending and contact deformation of the probing head as well as the part being measured significantly limit the measurable depth and precision. Some research has aimed to reduce the probing force to several mN [4,5], but high measurable depth and high precision cannot be achieved at the same time. Noncontact probing methods are promising solutions to this challenge [6–8]. Of all the noncontact probing methods, optical methods are more promising to obtain higher resolution and

precision [9–12]. Based on various “optical needle” principles, nanometer resolution can be easily achieved. However, applying optical probes in deep and small hole measurement is difficult because optical probes usually have only 1D nanometer resolution probing ability and are, thus, more suitable for 2D and  $2\frac{1}{2}$  D structure measurement, such as for the shape measurement of a shallow optical surface. A variety of micro and nano probes with nanometer resolution and high stylus aspect ratio have been developed [13–15]. However, these probes are relatively small and more suitable for the measurement of micro/nano structures. Additionally, their stylus length and measurable depth is limited to several millimeters.

Aside from the demand for new probing methods, other difficulties need to be solved to realize ultraprecision measurement of deep and small holes, such as multi-degree-of-freedom adjustment of hole attitude and macro displacement measurement of the probe.

In this paper, a patented ultraprecision measurement method [16] based on spherical scattering electrical-field probing (SSEP) is proposed, and a hole diameter measuring machine (HDMM) is specially designed and developed. Key techniques such as multi-degree-of-freedom adjustment of hole attitude, laser interferometry for macro displacement measurement of the probe, and hole diameter measurement process planning are described. Finally, experiments are carried out with the HDMM and a probe that is specifically developed for method verification.

## 2. Spherical Scattering Electrical-Field Probing Method

The model of small hole diameter measurement with a non-contact probe can be expressed as:

$$D = l + d + \delta_1 + \delta_2 \quad (1)$$

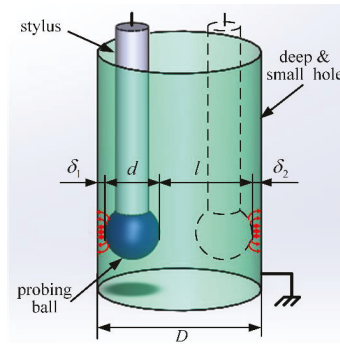
where  $D$  is the hole diameter to be measured;  $d$  is the probing ball diameter;  $l$  is the macro displacement of the probe when it is moved from the first probing point to the second along the measurement line of the diameter; and  $\delta_1$  and  $\delta_2$  are micro probing gaps between the probing ball and the hole sidewall at the two probing points.

In the proposed measurement method based on SSEP,  $d$  was calibrated beforehand,  $l$  was obtained using laser interferometry, and  $\delta_1$  and  $\delta_2$  were given by the SSEP probe. The quality of the diameter measurements depended on the ability to accurately determine  $\delta_1$  and  $\delta_2$ .

The principle of the SSEP method is shown in Figure 1. The probing ball and the small hole are electrically conductive. When the probing ball probes the internal sidewall of the hole, a dual conductor system is formed. By constructing certain boundary conditions—e.g., grounding the hole and keeping the potential of the probing ball constant—a spherical scattering electrical field is formed, and the electrical field between the probing ball and the hole sidewall, as well as the distribution of surface charge, is strongly affected by the micro probing gap. Based on this phenomenon, the micro gap between the probing ball and the hole sidewall is converted to a capacitive signal and detected with capacitive signal processing circuits [17].

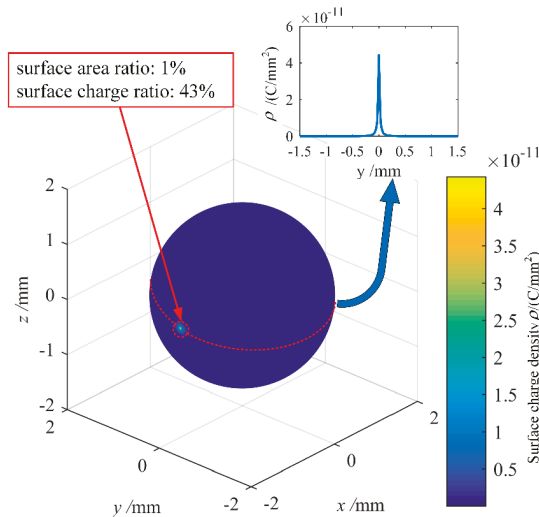
The sensing characteristics of the spherical scattering electrical field is identical in arbitrary spatial directions. Therefore, the measurement performance is not influenced by the relative attitude between the probe and the hole to be measured, which is often a key limitation in existing non-contact high-precision probing methods of small holes with large depth-to-diameter ratios.

Considering the actual demand in industry, only holes of conductor material are discussed—although the analysis shows that non-conductor material can also be probed using the SSEP method with much worse resolution.



**Figure 1.** Schematic diagram of the spherical scattering electrical-field probing (SSEP) method.

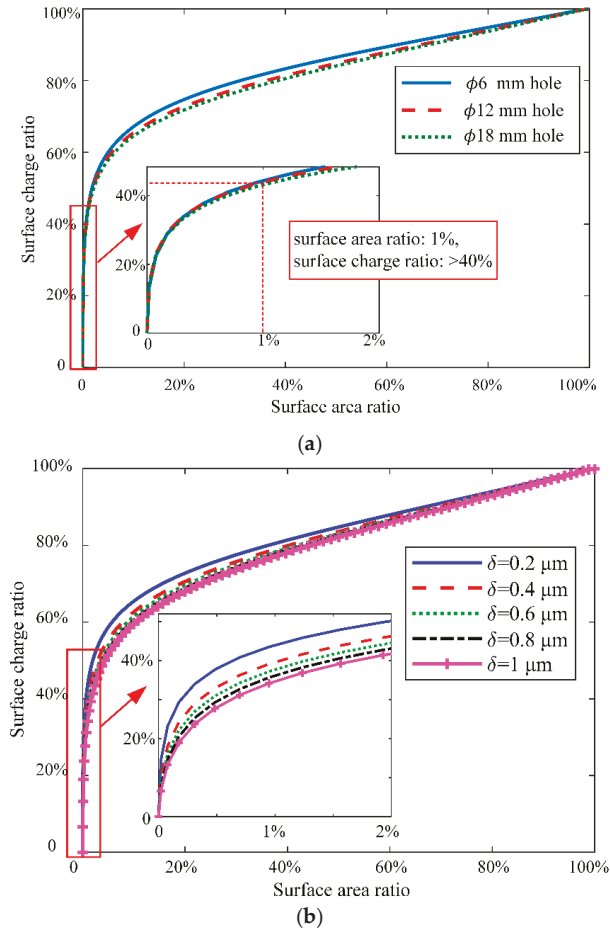
The bias electrical field was theoretically modeled using the seven-point finite difference method and non-uniform meshing in a spherical coordinate system [18]. The Laplace equation was solved for a model consisting of a  $\phi 3$ -mm conductor sphere as the probing ball and a conducting plane (yoz plane) as the surface to be measured. The potential of the sphere was 1V and the plane was grounded. The probing gap was set to 0.2  $\mu\text{m}$ . The simulation software was programmed by the authors with MATLAB (MathWorks, Natick, Massachusetts, US). The surface charge density characteristics of the probing ball are shown in Figure 2. It is shown that most surface charge on the probing ball was concentrated in a very small area closest to the part being measured, and 43% of the surface charge was concentrated in an area comprising 1% of the whole surface area of the probing ball. The charge density of the waist line of the probing ball is shown in the top right corner of Figure 2, and the curve is characterized by a single peak. This phenomenon indicates that the SSEP probe has a point probing capability. Therefore, the SSEP probe is capable of isotropic sensing and point-probing, both of which are ideal for the measurement of small holes with large depth-to-diameter ratios.



**Figure 2.** Surface charge density characteristics of a  $\phi 3$ -mm probing ball when a conductor plane (yoz plane) is probed.

Figure 3a shows the relationship between the surface area ratio, defined as the ratio of the charge concentrating area to the whole surface area of the probing ball, and the corresponding surface charge

ratio when a  $\phi 3$ -mm probing ball probes holes of different diameters with a  $0.2\text{-}\mu\text{m}$  probing gap. Figure 3b shows the relationship of the surface area ratio to surface charge ratio when a  $\phi 3$ -mm probing ball probes a  $\phi 12$ -mm hole with a probing gap from  $0.2\text{ }\mu\text{m}$  to  $1\text{ }\mu\text{m}$ . It can be observed that the surface charge concentration was similar when holes of different diameters were probed.



**Figure 3.** Relationship of the surface charge ratio to surface area ratio. (a) Simulation results with different holes (probing gap =  $0.2\text{ }\mu\text{m}$ ); (b) simulation results with different probing gaps (hole diameter =  $\phi 12$  mm).

Figure 4 shows the sensitivity characteristics of probing balls of different diameters when a plane was probed using capacitive signal processing techniques. The sensitivity was defined as the quotient of the change of capacitance and the corresponding change of the probing gap. The sensitivity increased with a decreasing probing gap ( $\delta$ ). The sensitivity increased more dramatically when  $\delta$  was below  $0.2\text{ }\mu\text{m}$ . Nanometer resolution could be achieved as a result of these characteristics.

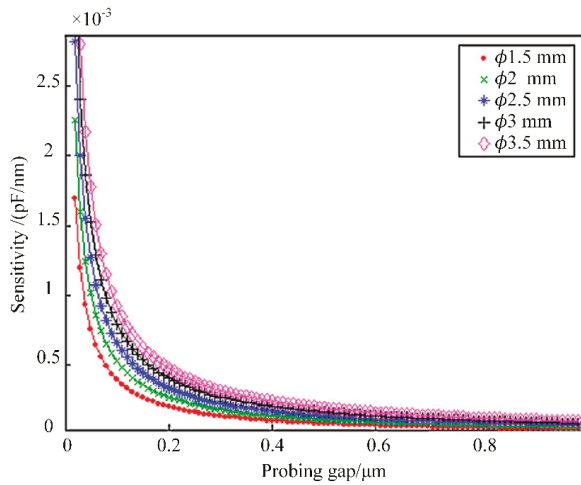


Figure 4. Sensitivity characteristics of probing balls of different diameters.

### 3. Development of the Hole Diameter Measuring Machine

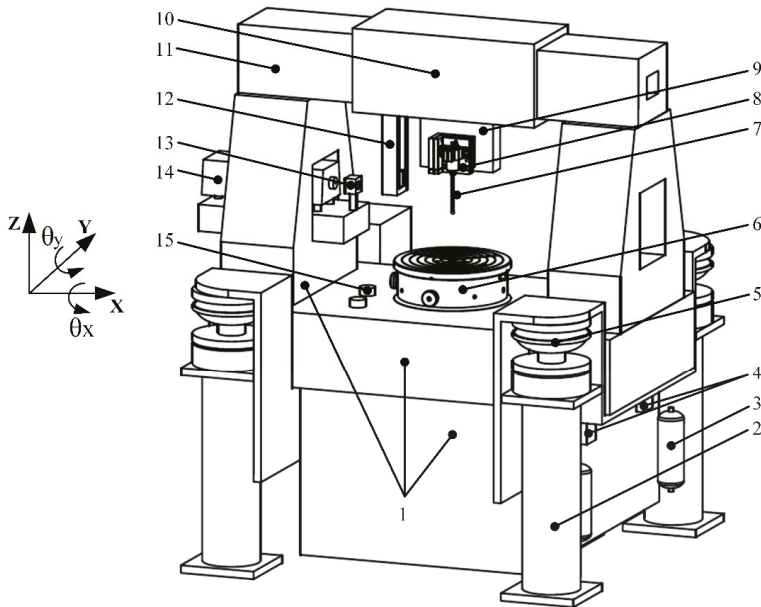
An HDMM was specially developed for ultraprecision diameter measurement of small holes with large depth-to-diameter ratios. Key issues, including machine structural design, the development of the probing sensor, the multi-degree-of-freedom attitude-adjusting worktable, and laser interferometry for macro probe displacement measurement, were subsequently described. The measurement process was optimally planned.

#### 3.1. Structural Design of the Measuring Machine

The structure of the HDMM is shown in Figure 5. A photo of the HDMM is shown in Figure 6. The machine had seven degrees of freedom of precision movement and positioning. The probe was mounted on a probe holder with two degrees of freedom (2D), which was mounted on an aerostatic linear slider. The 2D probe holder was used to precisely adjust the axis of the probe stylus to be parallel to the Z axis of the HDMM. The aerostatic linear slider could move horizontally in the X direction, which was also the direction of the measurement line of the hole diameter. The slider stroke in the X direction was 50 mm. A linear glass encoder with a resolution of 5 nm and an accuracy of  $\sim 1 \mu\text{m}$  after calibration was provided for the X slider. A laser interferometer was used to measure the macro displacement of the probe, i.e.,  $l$  in Equation (1). The part to be measured was mounted on a four-degrees-of-freedom (4D) worktable. The worktable could be controlled to move in the Y and Z directions, as well as to rotate in  $\theta_x$  (about the X axis) and  $\theta_y$  (about the Y axis) directions. The position and attitude of the hole to be measured could therefore be precisely adjusted with the worktable, which is further described in Section 3.3. The slider in the Z direction had a stroke of 150 mm. A linear glass encoder with a resolution of 5 nm and an accuracy of  $\sim 1 \mu\text{m}$  after calibration was also provided for the Z slider. The tilting angle of the worktable was adjusted with fine thread and servo motors. The tilting angle was evaluated according to pulse counts of the rotary encoders of the motors, and one count corresponded to  $\sim 0.005''$ . The maximum error of the tilting angle adjustment of the worktable was  $\sim 0.2''$ , well meeting the measurement requirement.

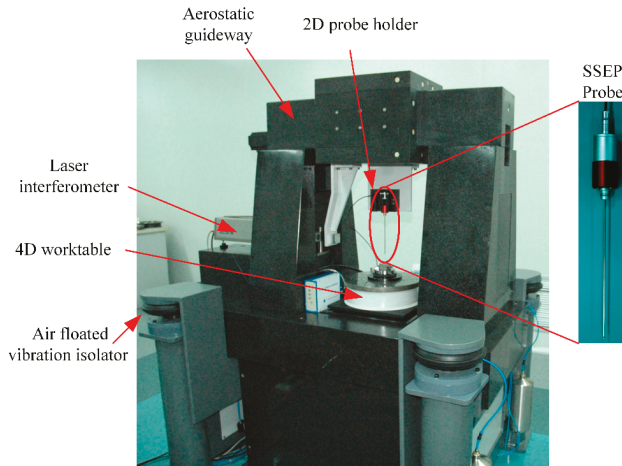
The measuring machine was supported by four air-floated vibration isolators. Environmental sensors were used to monitor the worktable temperature, air temperature, and air humidity to allow for compensation of the thermal expansion of the material and change of the air refractive index.





1. Machine frame; 2. Supporter; 3. Air tank for vibration isolator; 4. Pneumatic valve for isolator; 5. Air-floated vibration isolator; 6. Four-degree-of-freedom (4D) worktable; 7. SSEP Probe; 8. Two-degree-of-freedom (2D) probe holder; 9. Mounting rack; 10. Aerostatic linear slider; 11. Guideway; 12. Measurement mirror of interferometer and holder; 13. Interferometer; 14. Laser; 15. Environmental sensor.

**Figure 5.** Structural design of the hole diameter measuring machine (HDMM). SSEP: spherical scattering electrical-field probing.



**Figure 6.** Photo of the HDMM developed. SSEP: spherical scattering electrical-field probing.

### 3.2. Development of Probing Sensor

According to theoretical modeling results, when capacitive signal processing techniques are used after constructing certain boundary conditions for the spherical scattering electrical field, the capacitive

signal of the probe is on the order of a few pF [17]. As a result, weak capacitive signal processing circuits were designed and a signal modulation and demodulation method was used.

In order to avoid electromagnetic interference and ensure signal quality, a tri-coaxial active shielding structure was used in the SSEP sensor, which consisted of an SSEP probe and signal processing circuits, as shown in Figure 7. The inner shield, outer shield, and the signal conducting core wire of the stylus and probe cable comprised a tri-coaxial structure. The outer shield was grounded to shield off spatial electromagnetic interferences and eliminate the influence of parasitic capacitance. The inner shield was driven by a 1:1 amplifier and kept equipotential with the signal conducting wire to eliminate the influence of parasitic capacitance between them.

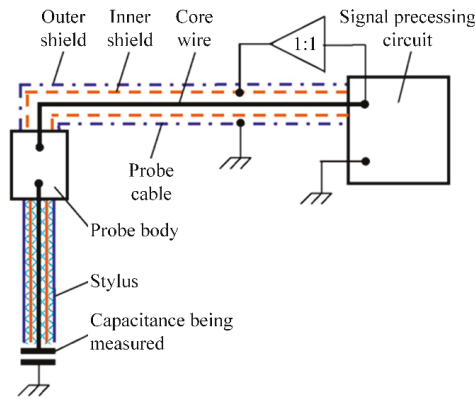


Figure 7. Tri-coaxial active shielding structure of the SSEP probing sensor.

A series of SSEP probes, with diameters ranging from  $\phi 1.5$  mm to  $\phi 3.5$  mm and stylus aspect ratios from 20 to 50, were developed. Figure 8 shows an SSEP probe prototype with a  $\phi 3$ -mm probing ball and a 150-mm-long stylus. Figure 9 shows the signal processing circuits. Nonlinearity correction, communication, and triggering control during probing was also conducted by the signal processing circuits.

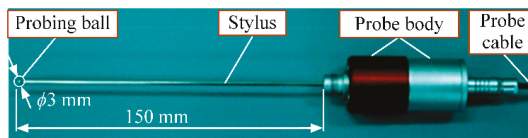


Figure 8. A SSEP probe prototype.

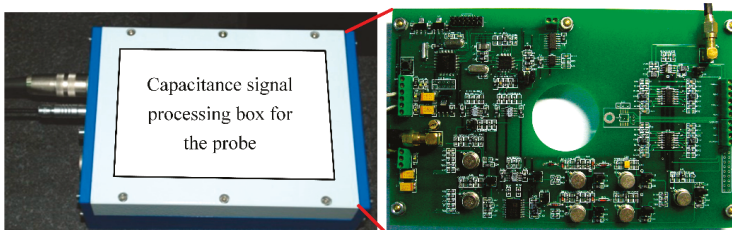


Figure 9. Signal processing circuits and box.

### 3.3. Four-Degree-of-Freedom Attitude-Adjusting Worktable

Precise and multi-degree-of-freedom adjustment of hole attitude also need to be thoroughly considered to realize ultraprecision measurement of deep and small hole diameters. The hole attitude must be precisely adjusted to make its axis parallel to the Z axis of the HDMM before diameter measurement. This was realized through a 4D attitude-adjusting worktable on which the part to be measured was mounted.

The inclination angles of the hole axis in the  $\theta_x$  and  $\theta_y$  directions with respect to the Z axis of the HDMM needed to be acquired to provide reference values for adjustment. The principle of determining an inclination angle from the 3D probing capability of the SSEP probe is shown in Figure 10. By successively probing the sidewalls of an upper section and a lower section,  $l_1$ ,  $l_2$ , and  $h$  could be obtained, and then the inclination angle  $\alpha$  could be calculated according to Equation (2).

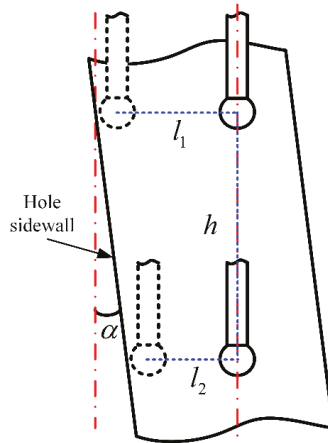


Figure 10. Principle of determining the inclination angle.

After acquiring inclination angles  $\alpha$  and  $\alpha'$  in two directions, the 4D worktable was rotated by  $-\alpha$  and  $-\alpha'$  in respective directions. The process of acquiring inclination angles and adjusting the worktable could be repeated until the inclination angles were smaller than a specified threshold. The experimental results show that two to three adjustments of the worktable are typically sufficient to reduce the inclination angle of the hole axis to a negligible value, thereby satisfying the measurement requirement.

It should be noted that the hole is assumed to be cylindrical here because the holes to be measured in the aviation, aerospace, and automotive industries are mostly cylindrical. Holes with a non-cylindrical shape are rarely seen. Also, because the precision of precisely manufactured holes in the aviation, aerospace, and automotive industries is usually in the micron order, misalignment resulting from differences at two different heights and the corresponding introduced measurement error is very small and can be ignored. Even for a truncated cone-shaped hole, the nominal value of its flank angle relative to its axis can be used to realize precision alignment.

$$\tan \alpha = \frac{|l_1 - l_2|}{h} \quad (2)$$

### 3.4. Macro Probe Displacement Measurement with Laser Interferometry

According to Equation (1), it is important to accurately obtain the macro displacement of the probe between the two probing points. A laser interferometer with a resolution of 1 nm was designed and integrated into the HDMM. According to calibration results, the relative length measurement error

of the interferometer was approximately  $\pm(1.5 \times 10^{-6})$ . As shown in Figures 5 and 6, the reflecting mirror was fixed on a holder mounted on the aerostatic guideway slider. To reduce the Abbe error, the laser interferometry was designed in such a way that the laser beam was incident on the center of the probing ball, and its direction was carefully adjusted to superpose the moving direction of the slider. In this way, the macro displacement of the probe could be accurately measured using laser interferometry.

### 3.5. Measurement Process Planning

The measurement process was planned to achieve fully automated hole attitude adjustment and diameter measurement of any cross-sections. The influence of personnel operations could be significantly reduced through automatic adjustment and measurement.

As shown in Figure 11, there were six major steps in the measuring process. First, the upper end face of the hole to be measured was determined with the SSEP probing sensor, and used as the depth reference of the cross-section hole diameter to be measured in the subsequent steps. Then, the hole attitude was finely adjusted using the principles described in Section 3.3. The part was moved vertically in the Z direction until the probing ball center was in the desired cross-section. The probe was then moved in the Y direction to find the inflection point, thus moving the probing ball center to the measurement line of the hole diameter to be measured.

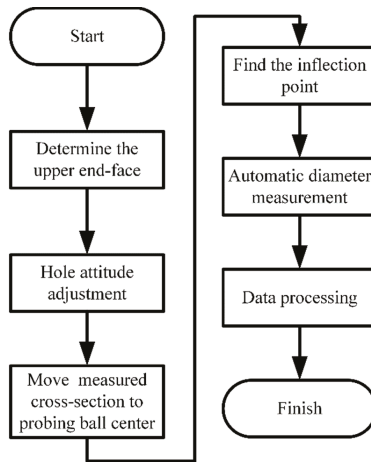


Figure 11. Diameter measurement process.

After the adjustment process above, automatic diameter measurement was carried out to obtain  $l$ ,  $\delta_1$ , and  $\delta_2$ . Measurements were repeated ten times, and ten measurement results were averaged.

## 4. Experiments and Measurement Uncertainty Analysis

### 4.1. Standards and Parts to Be Measured

To verify the performance of the SEPP method and the HDMM, a variety of standards and parts shown in Figure 12 were measured. The two deep and small hole standards were specially designed and used to verify the hole attitude adjustment method, planned measurement process, measurable depth, and depth-to-diameter ratio. The ring gauge standard was calibrated by NIM of China and Physikalisch Technische Bundesanstalt (PTB) of Germany. A tube part from a rocket engine was also measured. To this point, the smallest hole diameter measured with the HDMM was  $\phi 3$  mm, and the maximum hole depth measured was 150 mm.

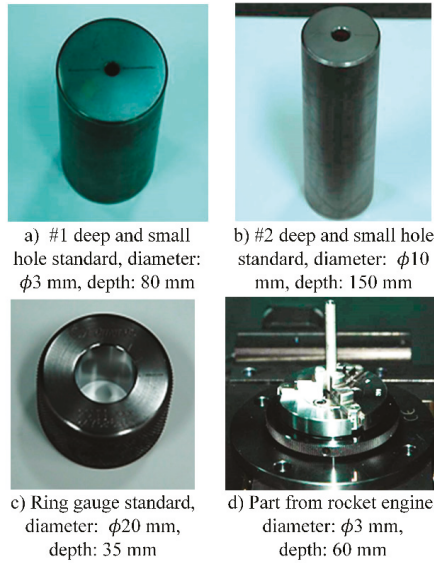


Figure 12. Standards and parts.

#### 4.2. Probe Resolution Experiment

Probe resolution experiments were carried out to verify the performance of an SSEP probing sensor. The experiments were conducted in a clean room with a temperature of  $20 \pm 0.5$  °C and humidity of  $45 \pm 5\%$  RH.

The probe with a  $\phi 3$ -mm probing ball shown in Figure 8 was used, and a metal plane part with surface roughness lower than 5 nm was measured. The probe was kept still, while the part was driven close to and away from the probing ball repeatedly with steps of 1 nm, 2 nm, and 3 nm by a piezoelectric stage (PI P-517.3CD). The smallest step that caused a perceptible change in the indication of the probe was defined as the resolution. The piezoelectric stage had a travel of 100  $\mu$ m, a closed-loop resolution of 1 nm, a repeatability of  $\pm 5$  nm, and a maximum nonlinearity of 7 nm in the full range. The experimental results are shown in Figure 13. The results indicate that the displacement resolution of the SSEP probe was as small as 1 nm.

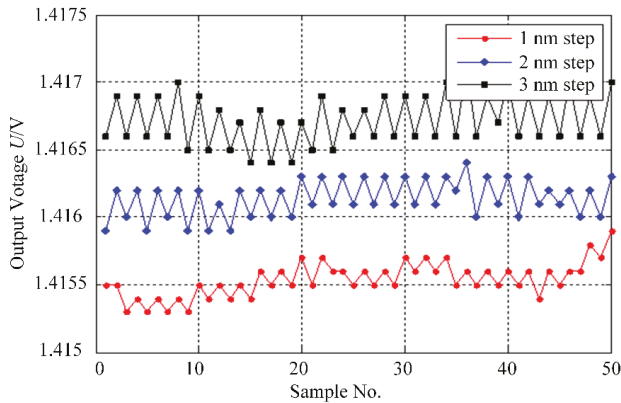


Figure 13. Resolution experiment results.

#### 4.3. Deep and Small Hole Diameter Measurement Experiment

The experiments were carried out using the specifically designed deep and small hole standards to verify the capability of measurable depth and depth-to-diameter ratio. Experimental results using the #2 standard and the  $\phi 3$ -mm probe in Figure 8 are shown in Figure 14. Diameters of six cross-sections at different depths were measured. The interval between adjacent cross-sections was 20 mm, and the maximum depth measured was 120 mm. Figure 14 indicates that the hole in #2 standard was generally conical with a slight trumpet shape at two ends.

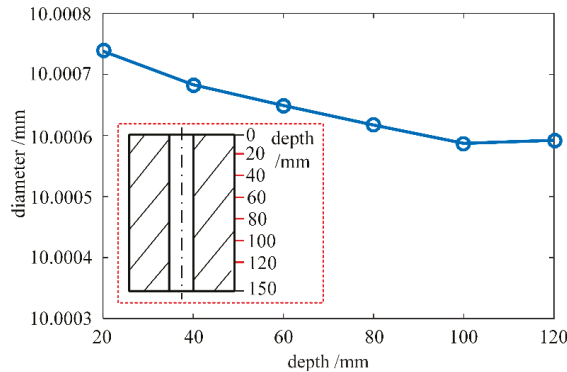


Figure 14. Diameter measurement results as a function of measured depth.

#### 4.4. Measurement Accuracy Experiment

Diameter measurement experiments were conducted in the clean room. The  $\phi 3$ -mm probe shown in Figure 8 was mounted on the HDMM. To validate the measuring method and the developed HDMM, a calibrated  $\phi 20$ -mm ring gauge standard with a depth of 35 mm from Edmunds Gages was used. The standard was mounted on the 4D worktable through a clamper, as shown in Figure 15. The measurement results are shown in Table 1. The average of ten diameter measurement results was 19.99998 mm. The diameter calibration results from NIM (National Institute of Metrology) of China, PTB of Germany, and the results obtained by the authors (HIT-UOI, Harbin Institute of Technology, center of Ultra-precision Optoelectronic Instrument) are compared in Table 2. The uncertainty intervals of all calibrations overlapped, indicating agreement.

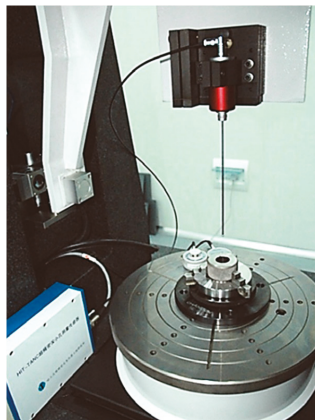


Figure 15. Measurement of  $\phi 20$ -mm ring gauge standard with the HDMM.

**Table 1.** Diameter measurement results of  $\phi 20$ -mm ring gauge standard.

Measurement Number	1	2	3	4	5
Measurement Result/mm	19.99993	20.00000	20.00001	20.00001	20.00001
Measurement Number	6	7	8	9	10
Measurement Result/mm	19.99998	19.99993	19.99995	19.99997	20.00001

**Table 2.** Comparison of diameter calibration results. PTB: Physikalisch Technische Bundesanstalt; NIM: National Institute of Metrology; HIT-UOI: Harbin Institute of Technology, center of Ultra-precision Optoelectronic Instrument.

Lab	Diameter Value D (mm)	Expand Uncertainty U ( $\mu\text{m}$ , $k = 2$ )
PTB	20.00000	0.1
NIM	20.00030	0.5
HIT-UOI	19.99998	0.2

#### 4.5. Measurement Uncertainty Analysis

The measurement model for small hole diameters is shown in Equation (1). For simplicity, the measurement model can be expressed as the sum of the macro displacement of the probe  $l$ , the probing ball diameter  $d$ , and two equal probing gaps  $\delta$  ( $\delta_1$  and  $\delta_2$  are assumed equal) as shown in Equation (3).

$$D = l + d + 2\delta \tag{3}$$

The uncertainty model of the diameter measurement is shown in Equation (4):

$$u_D^2 = u_l^2 + u_d^2 + 4u_\delta^2 \tag{4}$$

The sources and estimation of the standard uncertainty components are listed in Table 3. The uncertainty is evaluated according to the *Guide to the Estimation of Uncertainty in Calibration of Geometrical Measuring Equipment* (JJF 1130-2005). The expanded uncertainty of the diameter measurement result of the  $\phi 20$ -mm ring gauge standard is  $U = 0.2 \mu\text{m}$  ( $k = 2$ ).

**Table 3.** Sources and estimation of uncertainty components.

Symbol	Error Source and Type	Estimation/ $\mu\text{m}$	
$u_l$	uncertainty introduced by macro displacement measurement with interferometer	measurement repeatability (A <sup>1</sup> )	0.040
		interferometry measurement (B <sup>2</sup> )	0.002
		linearity of aerostatic guideway (B)	0.026
	$u_l = 0.048$		
$u_d$	uncertainty introduced by calibration of the probe ball diameter	standard (B)	0.031
		interferometry measurement (B)	0.040
		probing error (B)	0.036
	$u_d = 0.088$		
$u_\delta$	uncertainty introduced by probing	probing repeatability (A)	0.040
		probing resolution (B)	0.003
		nonlinearity of probing sensor (B)	0.006
		indication stability (B)	0.003
$u_\delta = 0.041$			
expanded uncertainty	$U = 0.2 \mu\text{m} (k = 2)$		

<sup>1</sup> Type A evaluation of uncertainty; <sup>2</sup> Type B evaluation of uncertainty.

## 5. Conclusions

An SSEP method was proposed for ultraprecision diameter measurement of small holes with large depth-to-diameter ratios and a specially designed measuring machine was developed. Experiments were carried out to verify the efficacy of the proposed method and the measurement performance of the machine. The resolution experiment results indicate that a resolution of 1 nm was achieved with a SSEP probe with a  $\phi 3$ -mm probing ball and a 150-mm stylus. The hole diameter measurement experiment results indicate that an expanded uncertainty of 0.2  $\mu\text{m}$  ( $k = 2$ ) was achieved when a standard with a  $\phi 20$ -mm nominal diameter was measured. It can be concluded that the proposed SSEP method is an effective solution for ultraprecision diameter measurement of small holes with large depth-to-diameter ratios. Future work will be focused on the application of the HDMM for micro holes with large aspect ratios and other inner structures.

**Author Contributions:** Conceptualization, J.C.; Methodology, J.C.; Software, Y.L.; Supervision, J.T.; Validation, X.B.; Writing—original draft, X.B.; Writing—review and editing, X.B.

**Funding:** This research was funded by [National Natural Science Foundation of China] grant number [51675139] and [Fundamental Research Funds for the Central Universities].

**Acknowledgments:** This research work is supported by the National Natural Science Foundation of China (Grant No. 51675139) and the Fundamental Research Funds for the Central Universities. The authors would like to thank the other members of the research team for their contributions to this study.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Bos, E.J.C. Aspects of tactile probing on the micro scale. *Precis. Eng.-J. Int. Soc. Precis. Eng. Nanotechnol.* **2011**, *35*, 228–240. [[CrossRef](#)]
2. Bauza, M.B.; Woody, S.C.; Woody, B.A.; Smith, S.T. Surface profilometry of high aspect ratio features. *Wear* **2011**, *271*, 519–522. [[CrossRef](#)]
3. Hansen, H.N.; Carneiro, K.; Haitjema, H.; De Chiffre, L. Dimensional micro and nano metrology. *CIRP Ann. Manuf. Technol.* **2006**, *55*, 721–743. [[CrossRef](#)]
4. Küng, A.; Meli, F.; Thalmann, R. Ultraprecision micro-cmm using a low force 3d touch probe. *Meas. Sci. Technol.* **2007**, *18*, 319. [[CrossRef](#)]




5. Takamasu, K.; Chih-Che, K.; Suzuki, A.; Hiraki, M.; Furutani, R.; Ozono, S. Development of pneumatic ball probe for measuring small hole. In Proceedings of the International Conference on Precision Engineering, Taipei, Taiwan, 20–22 November 1997.
6. Tian, G.Y.; Zhao, Z.X.; Baines, R.W.; Corcoran, P. A miniaturised sensor for deep hole diameter measurement. *Precis. Eng.* **1999**, *23*, 236–242. [[CrossRef](#)]
7. Sun, C.-K.; Wang, X.-B.; Liu, B.; Zheng, Y.-Z. Capacitance sensor measurement method for micro-aperture. *Nanotechnol. Precis. Eng.* **2006**, *2*, 103–107.
8. Grandy, D.; Koshy, P.; Klocke, F. Pneumatic non-contact roughness assessment of moving surfaces. *CIRP Ann.* **2009**, *58*, 515–518. [[CrossRef](#)]
9. Schmitt, R.; Pfeifer, T.; Depiereux, F.; König, N. Novel fiber-optical interferometer with miniaturized probe for in-hole measurements. *Optoelectron. Lett.* **2008**, *4*, 140–142. [[CrossRef](#)]
10. Zhao, W.; Tan, J.; Qiu, L.; Zou, L. A new laser heterodyne confocal probe for ultraprecision measurement of discontinuous contours. *Meas. Sci. Technol.* **2005**, *16*, 497. [[CrossRef](#)]
11. Heredia-Ortiz, M.; Patterson, E.A. On the industrial applications of moiré and fringe projection techniques. *Strai* **2003**, *39*, 95–100. [[CrossRef](#)]
12. Casavola, C.; Pappalettera, G.; Pappalettera, C. Design of a fiber optics fringe projector for 3d reconstruction of dental elements. In Proceedings of the 2014 Third Mediterranean Photonics Conference, Trani, Italy, 7–9 May 2014; pp. 1–3.
13. Muralikrishnan, B.; Stone, J.A.; Stoup, J.R. Fiber deflection probe for small hole metrology. *Precis. Eng.* **2006**, *30*, 154–164. [[CrossRef](#)]
14. Beomjoon, K.; Takahisa, M.; Tarik, B. The vibroscanning method for the measurement of micro-hole profiles. *Meas. Sci. Technol.* **1999**, *10*, 697.
15. Weckenmann, A.; Peggs, G.; Hoffmann, J. Probing systems for dimensional micro- and nano-metrology. *Meas. Sci. Technol.* **2006**, *17*, 504. [[CrossRef](#)]
16. Tan, J.; Cui, J.; Cui, J. Ultraprecision Non-Contact Three-Dimensional Probing System Based on Spherical Capacitive Plate. U.S. Patent 8,816,702, 26 August 2014.
17. Tan, J.; Cui, J. Ultraprecision 3d probing system based on spherical capacitive plate. *Sens. Actuators A Phys.* **2010**, *159*, 1–6. [[CrossRef](#)]
18. Bian, X.; Cui, J.; Lu, Y.; Tan, J. Bias electric field distribution analysis based on finite difference method with non-uniform grids for a non-contact tunneling current probe. In *10th International Symposium on Precision Engineering Measurement and Instrumentation*; SPIE: Kunming, China, 2018; in press.



© 2019 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).

Article

# Dynamic Measurement Error Modeling and Analysis in a Photoelectric Scanning Measurement Network

Shendong Shi <sup>1</sup>, Linghui Yang <sup>1,\*</sup>, Jiarui Lin <sup>1</sup>, Changyu Long <sup>2</sup>, Rui Deng <sup>1</sup>, Zhenyu Zhang <sup>1</sup> and Jigui Zhu <sup>1</sup>

<sup>1</sup> State Key Laboratory of Precision Measuring Technology and Instruments, Tianjin University, Tianjin 300072, China; ssd2168@tju.edu.cn (S.S.); linjr@tju.edu.cn (J.L.); denghsby@163.com (R.D.); zhang\_zhenyu@tju.edu.cn (Z.Z.); jiguzhu@tju.edu.cn (J.Z.)

<sup>2</sup> Beijing Institute of Spacecraft Environment Engineering, Beijing 100094, China; cylong\_cast@163.com

\* Correspondence: icelinker@tju.edu.cn; Tel.: +86-139-2029-7997

Received: 30 October 2018; Accepted: 19 December 2018; Published: 25 December 2018



**Featured Application:** A photoelectric scanning measurement network has the potential for three-dimensional (3-D) dynamic coordinate measurement in industrial manufacturing. It can be applied in aircraft assembly and leveling, digital shipbuilding, tooling monitoring, AGV navigation, and so on.

**Abstract:** A photoelectric scanning measurement network is a kind of distributed measurement system based on the principle of angle intersection, in which transmitters and photoelectric receivers are the main parts. The scanning lasers in transmitters emit signals and they are obtained by receivers at the measured points. Then the coordinate of the receiver can be calculated by the optimization algorithm. Its outstanding static measurement performance and network scalability capacity give it great potential in large-scale metrology. However, when it comes to moving targets, the angle intersection failure will produce a dynamic error, which limits its further application. Nowadays the research on error modeling and compensation is also insufficient though it has been the crucial concern. In this paper, we analyzed error causes and constructed a dynamic error model. Dynamic error characteristics and the law of propagation were discussed. The measurement uncertainty at different movement speeds was quantized through simulation experiments. To verify the error model, experiments were designed and the dynamic error was evaluated in practice. It matched well with simulations. The model was tested to be reasonable, and provided theoretical support for error compensation.

**Keywords:** Photoelectric scanning; angle intersection; dynamic error modeling; large-scale metrology

## 1. Introduction

The concept of “Industry 4.0” introduces the future trends in industrial development, and the growing interest in intelligent manufacturing has highlighted the need to research novel methods in large-scale metrology [1,2]. Integration of advanced measurement instruments with traditional technologies is recognized as an effective approach to achieve the prospect. Dimensional metrology in manufacturing, which is an important factor for quality insurance, has the characteristics of large measurement range, requirements for high precision and efficiency, and strong environmental interference. In recent years, the breakthroughs in optical technologies and low-cost, fast computers stimulated the wide use of laser tracker and photogrammetry systems. They both provide abundant solutions for three-dimension coordinate measurement [3–5].

A laser tracker is based on length measurement by interferometer and angle measurement by encoder. Leica AT series and Faro Vantage series are representatives in this field of research. With sophisticated environmental parameter compensation, they can gain an accuracy of  $15 \mu\text{m} + 6 \mu\text{m}/\text{m}$  and a measurement frequency of up to 1000 Hz. The measurement range is about 80 m. In most cases, it is seen as the benchmark in metrology and gives compensation to positioning error in other instruments [6,7]. However, the laser tracker has its limitations. Firstly, it only conducts single target measurement each time and the efficiency is relatively lower. For example, in large aircraft leveling, hundreds of points on the whole fuselage need to be measured, which may take one or more days. It is time-consuming and unacceptable. Secondly, in a large-scale space, a multi-station network is often necessary for full coverage of the measurement zone. The measurement error is prone to accumulate during coordinate system transformation and distance increase. The system performance will be affected to a great degree [8,9].

Close-range photogrammetry experienced a revolution at the beginning of the 2000s [10] and it has been widely used in applications ranging from automobile manufacturing to aerospace [11–15]. Monocular vision and binocular vision are common solutions for positioning and orientation in large-scale metrology and they enjoy an accuracy of submillimeter. Nevertheless, limited by the image processing and matching algorithm, photogrammetry methods have a relatively lower measurement frequency, which leads to unideal dynamic measurement capacity. Even worse, photogrammetry systems have strict requirements for the environment and measured objects. Light flection from the surface of objects, such as aircraft skin and body-in-white, may lead to low image quality and match failure [16].

In the past few years, we have also witnessed the rapid development of stationary distributed measurement network and it is now widely used in aircraft manufacturing, digital shipbuilding, and tooling monitoring. The workshop measurement positioning system (wMPS) is a representative composed of transmitters, photoelectric receivers, signal processors, and a terminal computer. Each transmitter contains a rotation head and a static base, as shown in Figure 1.

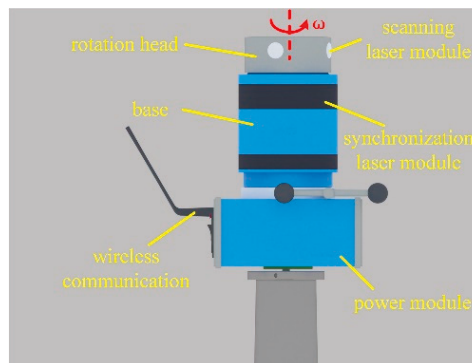


Figure 1. Transmitter configuration.

Two line-laser modules are embedded in the rotation head with an intersection angle of  $90^\circ$ . The head rotates about the z-axis of the local coordinate system at a speed of 1500–3000 rev/min. In the static base synchronization lasers are in circular distribution and they will shine when the rotation head goes across the pre-defined reference point in the encoder. The receiver at the measured point obtains the synchronization signal and record it as the start time. After that, laser module 1 and 2 will scan over the receiver in turn. The time difference between synchronization signal and scanning signals is captured through precise timing. It is in a linear relationship with two scanning angles. In this way, the spatial relationship between the transmitter and receiver can be determined solely. Combing signals from multiple transmitters, the coordinate of the receiver can be obtained by the

optimization algorithm [17]. At present, the research on static measurement performance of wMPS has been carried out and it gains an accuracy of 0.2 mm + 0.01 mm/m at a frequency of 20 Hz.

However, with the development of industrial automation, dynamic positioning has been increasingly required. In aircraft and ship assembly, large components are carried by industrial robots for docking. The real-time positions of components are necessary for feedback control. Besides, the relative position between different robots is also important to avoid the deformation caused by measurement asynchronization [18,19]. For indoor autonomous guided vehicle (AGV), accurate positioning in global positioning system (GPS) denied conditions is the foundation for navigation and control. On the one hand, wMPS has the potential for dynamic measurement, fast dynamic response, and good static stability. On the other hand, limited by the principle of angle intersection measurement, the error is inherent and non-negligible when the receiver is in motion. Dynamic error analysis and compensation has been the key to further improve the system performance and extend the application area.

In this article, we first introduce the dynamic error causes of wMPS. Then we construct the error model, discuss the error characteristics, and quantize the measurement uncertainty at different cases through simulation experiments. Finally, practical experiments are carried out to validate the proposed model.

## 2. Measurement Principle and Causes of Dynamic Error

### 2.1. Measurement Principle

Specifically, the measurement model of wMPS can be abstracted as two rotating laser planes. The normal vectors and z-axis intercepts of two laser planes at the reference point are described as:

$$\begin{bmatrix} a_{ij} & b_{ij} & c_{ij} & d_{ij} \end{bmatrix}^T \tag{1}$$

where  $i$  and  $j$  indicate the  $i$ 'th transmitter and  $j$ 'th laser module. They are constant after the transmitter is assembled.

When the two lasers scan across the receiver counterclockwise, the normal vectors can be expressed as:

$$\begin{bmatrix} a_{ij\theta} \\ b_{ij\theta} \\ c_{ij\theta} \\ d_{ij\theta} \end{bmatrix} = \begin{bmatrix} \cos(\omega_i t_{ij}) & -\sin(\omega_i t_{ij}) & 0 & 0 \\ \sin(\omega_i t_{ij}) & \cos(\omega_i t_{ij}) & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} a_{ij} \\ b_{ij} \\ c_{ij} \\ d_{ij} \end{bmatrix} \tag{2}$$

$\omega_i$  refers to the rotation speed of the  $i$ 'th transmitter, and  $t_{ij}$  refers to the time difference between the synchronization signal and  $j$ 'th scanning signal.

When a laser plane scans over a receiver, the receiver will locate on the plane and the constraint formula between the laser plane and receiver can be constructed:

$$\begin{bmatrix} a_{ij\theta} & b_{ij\theta} & c_{ij\theta} & d_{ij\theta} \end{bmatrix} \begin{bmatrix} R_{11} & R_{12} & R_{13} & T_x \\ R_{21} & R_{22} & R_{23} & T_y \\ R_{31} & R_{32} & R_{33} & T_z \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_g & y_g & z_g & 1 \end{bmatrix}^T = 0 \tag{3}$$

$R$  and  $T$  represent the rotation and translation components from the global coordinate system to the local coordinate system respectively. They are derived from measurement network calibration [17].  $\begin{bmatrix} x_g & y_g & z_g & 1 \end{bmatrix}$  represents the homogeneous coordinate of the receiver in the global coordinate system. There are three unknowns in the coordinate of each receiver and they will be calculable as long as more than three formulas are available. Consequently, signals from more than two transmitters

with different rotation speeds will be capable for coordinate calculation. The accuracy is positively related to the number of transmitters scanning across the receiver.

### 2.2. Causes of Dynamic Error

As stated in Section 2.1, laser signals from at least two transmitters with different rotation speeds are necessary for coordinate calculation. For each transmitter, a receiver will obtain a synchronization pulse and two scanning pulses in sequence as shown in Figure 2. In the measuring network, the synchronization pulses from different transmitters are irrelevant in the time domain.

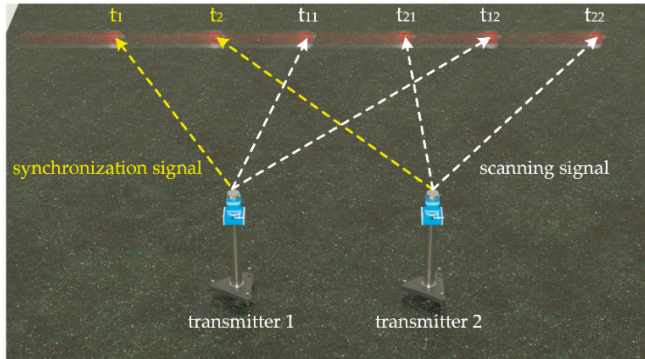


Figure 2. Time sequence of signals from two transmitters.

If the receiver is static, laser planes have a unique intersection point and the time difference between signals has no influence on measurement accuracy. However, for a moving receiver, in the period of obtaining signals from the two transmitters, i.e.,  $t_1$  to  $t_{22}$ , it has been at different positions and the trajectory during that is unknown. Calculating the coordinate of the receiver directly will lead to certain dynamic error.

We define single transmitter dynamic time error  $\delta t_{s_k}$  as the time difference between synchronization pulse and scanning pulse 1, pulse 2 ( $t_1$  and  $t_{11}$ ,  $t_1$  and  $t_{12}$ ), and the multi-transmitter dynamic error  $\delta t_{m_k}$  is defined as the time difference between synchronization pulses from different transmitters ( $t_1$  and  $t_2$ ). Note that multi-transmitter dynamic error  $\delta t_{m_k}$  is strictly related with rotation periods and it will accumulate over rotation cycles.

$$\delta t_{m_k} = \text{mod}(N(T_k - T_1), T_1) \tag{4}$$

$N$  represents the rotation cycles and  $T_k$  represents the rotation period of the transmitter  $k$ .  $T_1$  is the rotation period of the primary transmitter in the network.  $\delta t_{s_k}$  is determined by the distribution of two laser modules. They have an intersection angle of  $90^\circ$ .  $\delta t_{m_k}$  is determined by the rotation cycles of all transmitters in the network. When the measurement network is built, it remains unchanged.  $\delta t_{s_k}$  and  $\delta t_{m_k}$  are uncorrelated and lead to dynamic error  $\delta\theta_{s_k}$  and  $\delta\theta_{m_k}$  in scanning angle respectively.  $\delta\theta_{e_k}$  is the static angle error and it is related with the straightness error of the laser module and centering error of the receiver. It is uncorrelated with the two error resources above. Therefore, the total dynamic error of wMPS can be expressed in Equation (5):

$$\delta\theta_k = \delta\theta_{e_k} + \delta\theta_{s_k} + \delta\theta_{m_k} \tag{5}$$

The static measurement error of a single transmitter has been evaluated after it is assembled. It obeys the normal distribution and the standard deviation is about  $2''$ . Besides, the static error at each

measurement is independent and the covariance is zero. For the other two error resources, simulation experiments based on MATLAB are conducted for time error quantization.

Basing on the principle of intersection measurement, the coordinates of the receiver can be calculable if it obtains signals from two transmitters. However, there is no constraint redundancy and the accuracy and stability are relatively lower. With the increase of the transmitter amount, the accuracy will also be improved. In application, we generally construct a four-transmitter network to gain better performance. Taking it into account, a four transmitters network was built and they were in a “C” type layout, which had been tested to gain a higher accuracy than others [20]. The rotation speeds were 1600, 1800, 2000 and 2200 rev/min, respectively. A trajectory generator was applied to produce a spatial trajectory and the receiver moved along it at different speeds, i.e., 30 mm/s, 60 mm/s, and 120 mm/s. The dynamic time error  $\delta t_{s,k} + \delta t_{m,k}$ , i.e. the time difference between the first synchronization signal and the last scanning signal, was captured at each measurement. Experimental configuration and results are shown in Figures 3 and 4:

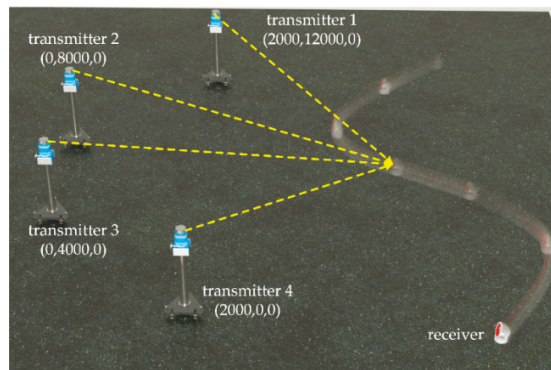


Figure 3. Configuration of simulation experiments.

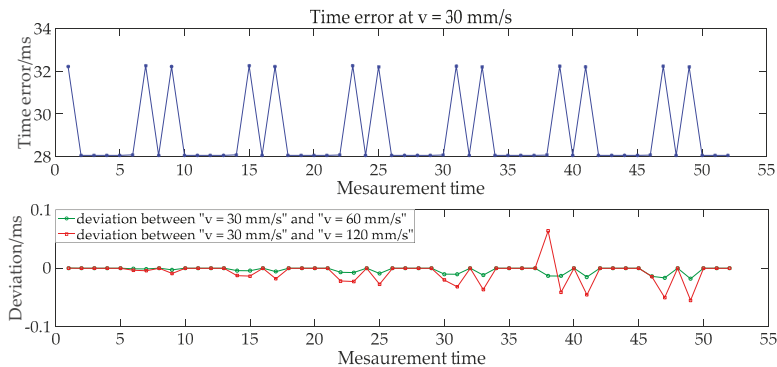


Figure 4. Dynamic time error in the four-transmitter network.

As displayed in Figure 4, the maximum and minimum error at  $v = 30$  mm/s are 32.24 ms and 28.06 ms. The time error corresponds to a periodic curve in the time-domain. The deviations at different speeds are lower than 0.05 ms and they are negligible in this research. Change the trajectory in subsequent simulations and we find that in the whole measurement zone the time error has great similarity with that of  $v = 30$  mm/s. It is dependent on the rotation speeds of transmitters in the network. With the increase of rotation speeds,  $\delta t_{s,k}$  and  $\delta t_{m,k}$  will decrease. However, the speed and position of the receiver have little influence on dynamic time error.

In wMPS we generally set a primary transmitter. The global coordinate system will be aligned with the coordinate system of the primary transmitter. The rotation and displacement relationship between the primary transmitter and others transmitters will be calibrated during the construction of the network. Each transmitter has a certain pose and position in the global coordinate system. On the basis of the conclusion above, we divide the movement speed of the receiver into three directions: radial speed  $v_r$ , tangential speed  $v_t$  and axial speed  $v_a$  in Figure 5. Here the radial, tangential, and axial directions mean those of each separate transmitter.

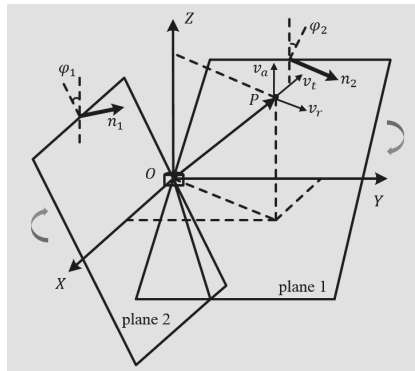


Figure 5. Transmitter model and speed decomposition.

According to the transmitter model, the radial speed has little influence on the dynamic error on account of the fast light speed. Therefore, the scanning angle error caused by time error can be expressed as:

$$\delta\theta_{s,k} + \delta\theta_{m,k} = \left( \frac{v_t}{\sqrt{x^2 + y^2}} + \frac{\tan \varphi \times v_a}{\sqrt{x^2 + y^2}} \right) (\delta t_{s,k} + \delta t_{m,k}) \tag{6}$$

$\begin{pmatrix} x & y & z \end{pmatrix}$  is the position of the receiver in the local coordinate system.  $\varphi$  refers to the intersection angle between the rotation axis and laser plane. In this way, the dynamic error resources in wMPS are presented.

### 3. Dynamic Error Modeling and Uncertainty Analysis

#### 3.1. Dynamic Error Modeling

In Section 2.2 we have analyzed the dynamic error mechanism. The measurement error in scanning angles causes the coordinate error of the moving receiver. The error propagation model is included in the following equations  $F(\mathbf{P}, \boldsymbol{\theta}) = 0$ :

$$\left\{ \begin{array}{l}
 F_1 = \begin{bmatrix} a_{11} & b_{11} & c_{11} & d_{11} \end{bmatrix} \begin{bmatrix} \cos \theta_{11} & -\sin \theta_{11} & 0 & 0 \\ \sin \theta_{11} & \cos \theta_{11} & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} R_1 & T_1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x_g & y_g & z_g & 1 \end{bmatrix}^T = 0 \\
 F_2 = \begin{bmatrix} a_{12} & b_{12} & c_{12} & d_{12} \end{bmatrix} \begin{bmatrix} \cos \theta_{12} & -\sin \theta_{12} & 0 & 0 \\ \sin \theta_{12} & \cos \theta_{12} & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} R_1 & T_1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x_g & y_g & z_g & 1 \end{bmatrix}^T = 0 \\
 \vdots \\
 F_{2n-1} = \begin{bmatrix} a_{n1} & b_{n1} & c_{n1} & d_{n1} \end{bmatrix} \begin{bmatrix} \cos \theta_{n1} & -\sin \theta_{n1} & 0 & 0 \\ \sin \theta_{n1} & \cos \theta_{n1} & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} R_n & T_n \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x_g & y_g & z_g & 1 \end{bmatrix}^T = 0 \\
 F_{2n-2} = \begin{bmatrix} a_{n2} & b_{n2} & c_{n2} & d_{n2} \end{bmatrix} \begin{bmatrix} \cos \theta_{n2} & -\sin \theta_{n2} & 0 & 0 \\ \sin \theta_{n2} & \cos \theta_{n2} & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} R_n & T_n \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x_g & y_g & z_g & 1 \end{bmatrix}^T = 0
 \end{array} \right. \tag{7}$$

The sources of coordinate uncertainty are scanning angles  $\theta_1, \theta_2, \dots, \theta_{2n-1}, \theta_{2n}$ . They are independent of each other and generally obey the normal distribution in Figure 6. The distribution center is determined by the dynamic error and the variance is determined by the static error.

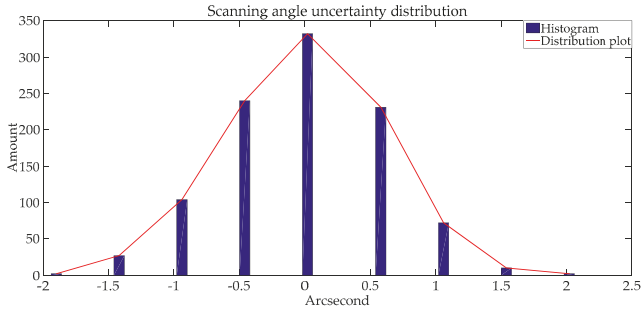


Figure 6. Scanning angle error distribution.

The covariance matrix  $\theta$ , also seen as the input, can be expressed as:

$$\theta = \begin{bmatrix} u^2(\theta_1) & & 0 \\ & \ddots & \\ 0 & & u^2(\theta_{2n}) \end{bmatrix} \tag{8}$$

According to the derivative rule for implicit functions [21],

$$\frac{dP}{d\theta} = -\left(\frac{dF}{dP}\right)^+ \frac{dF}{d\theta} \tag{9}$$

$P = \begin{bmatrix} x & y & z & 1 \end{bmatrix}$  and it is the homogeneous coordinate of the receiver in the global coordinate system.  $\frac{dF}{dP}$  and  $\frac{dF}{d\theta}$  are the Jacobian matrices of  $F$  relative to  $P$  and  $\theta$ .  $\left(\frac{dF}{dP}\right)^+$  is the generalized inverse matrix of  $\frac{dF}{dP}$ .



In simulations we found that the numerical stability of  $\theta$  is unsatisfactory and it is highly sensitive to round-off error in programming, so the other algorithm is tried which depends on the following steps [22]:

1. Divide  $\theta$  into a lower triangular matrix  $\lambda_C$  and its transposed matrix:  $\theta = \lambda_C^T \lambda_C$ .
2. Divide  $\frac{dF}{d\theta}$  into the product of an orthogonal matrix  $Q_C$  and an upper triangular matrix  $W_C$ :  $\frac{dF}{d\theta} = Q_C W_C$ .
3. Divide  $\frac{dF}{dP}$  into the product of a lower triangular matrix  $L_P$  and an upper triangular matrix  $W_P$ :  $\frac{dF}{dP} = L_P W_P$ .
4. Calculate matrix  $M_1$  through  $W_P^T M_1 = I$ ,  $M = \lambda_C W_C^T Q_C^T (L_P^T)^{-1} M_1$ .
5. Divide  $M$  into the product of an orthogonal matrix  $Q_M$  and an upper triangular matrix  $R_M$ :  $M = Q_M W_M$ .
6. Describe the uncertainty  $u_{xyz}$  as:  $u_{xyz} = u_P = W_M^T W_M$ .

### 3.2. Uncertainty Simulation and Analysis

Basing on the mathematical solution in Section 3.1, simulation experiments were carried out in MATLAB. A measurement network with four transmitters was built and the layout was the same as Section 2.2. A receiver moved at different speeds, 30, 60, and 120 mm/s respectively. We set 4800 points at three horizontal planes ( $z = 0$ ,  $z = 1000$ , and  $z = 2000$ ) uniformly as Figure 7. The coordinate uncertainty components of each point, i.e.,  $u_x$ ,  $u_y$  and  $u_z$ , were calculated. In Figures 8–10 we have displayed the compound uncertainty  $u$  which indicates the distance between the theoretical point and measurement result.

$$u = \sqrt{u_x^2 + u_y^2 + u_z^2} \tag{10}$$

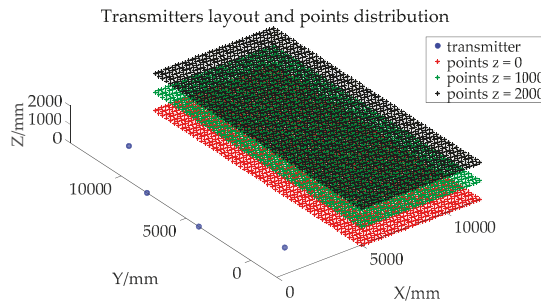


Figure 7. Points distribution in the measurement zone.

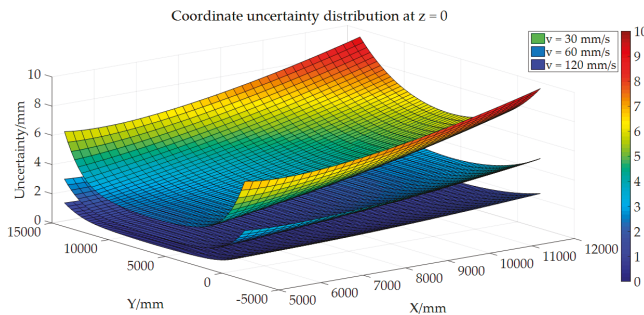


Figure 8. Coordinate uncertainty distribution of points at  $z = 0$ .

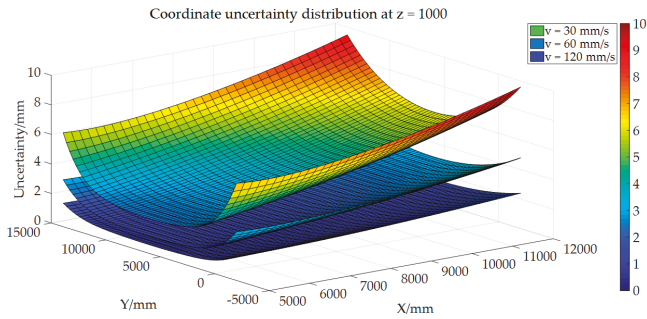


Figure 9. Coordinate uncertainty distribution of points at z = 1000.

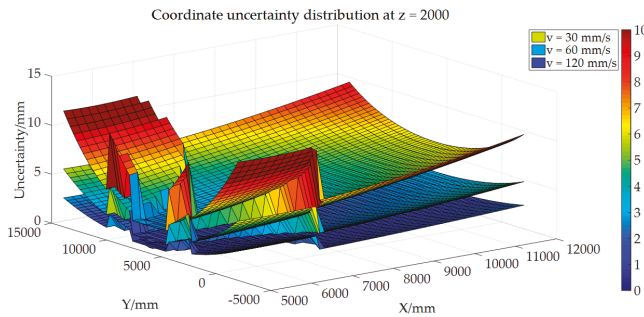


Figure 10. Coordinate uncertainty distribution of points at z = 2000.

We can draw a conclusion from Figures 8–10 that the dynamic error is in a linear relationship with the movement speed of the receiver. With the increase of receiver speed, the dynamic error will also get larger. Although in all cases it contains a constant static error, it occupies a small part in the total dynamic error. The single transmitter dynamic error and multi-transmitter dynamic error are dominating. Moreover, the layout of the network also influences the dynamic error. The optimal intersection angle for measurement is  $110^\circ$ . In the margin area, where the intersection angle is extremely small, the measurement error may be enlarged, which has the same characteristics as static conditions [20].

#### 4. Experimental Verification

To verify the mathematic model of wMPS dynamic error, practical experiments are designed and the configuration is displayed in Figure 11:

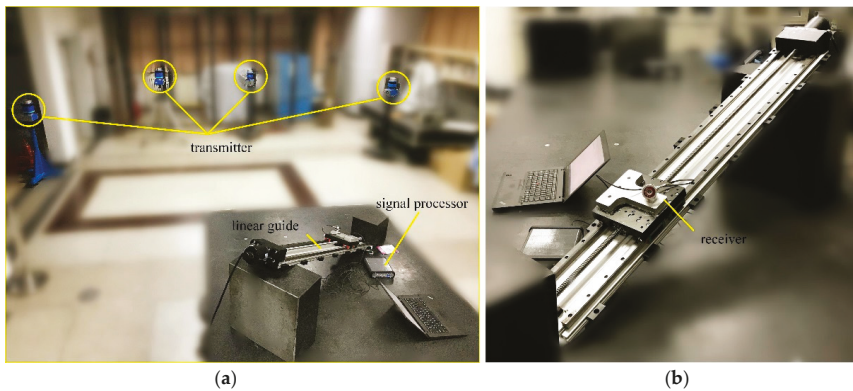


Figure 11. Experiment configuration: (a) Overview; (b) linear guide and receiver.

We first constructed a measurement network with four transmitters. The rotation speeds of transmitters were set as 1600, 1800, 2000, and 2200 rev/min respectively. They were consistent with those in simulations in Section 2.2. The orientation parameter of transmitters was calibrated according to a coordinate control network [17]. A linear guide with a maximum travel distance of 800 mm was applied to generate a standard linear trajectory and its straightness error was better than 0.07 mm after calibration with an interferometer. It is much smaller than dynamic error and the linear guide can be seen as a reference in experiments. A receiver was located on the platform and moves with it synchronously. Movement speeds were set as 30 mm/s, 60 mm/s, and 120 mm/s respectively. Experiments were carried out in two conditions successively: the linear guide was 5 m/7 m in front of the primary transmitter. To break the specificity in speed direction, in both conditions the linear guide had a horizontal angle of 45° and a vertical angle of 20° in the global coordinate system.

The real-time coordinates of the receiver during each motion were recorded. Then we fitted a spatial line. The dynamic error was evaluated by the straightness error and the root-mean-square errors at each case were also calculated in Figures 12 and 13.

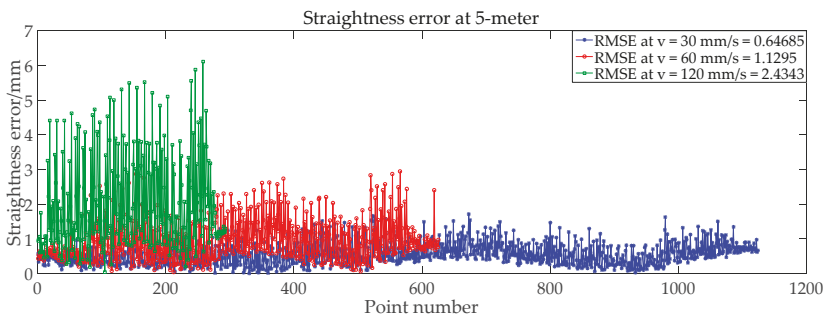


Figure 12. Straightness error of practical trajectory at 5-m at different speeds.

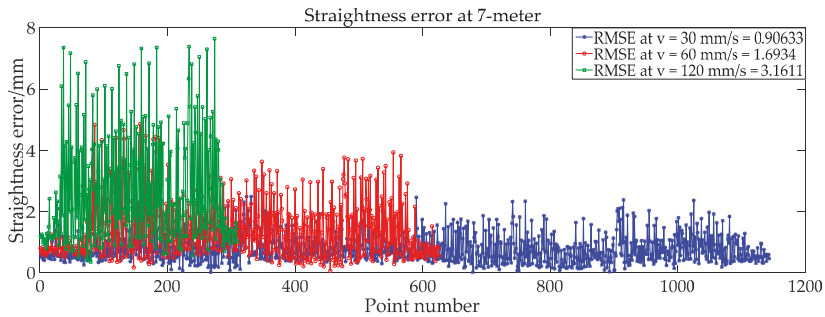


Figure 13. Straightness error of practical trajectory at 7-m at different speeds.

For further comparison, Table 1 lists the deviations between practical experiment results and corresponding simulation results. We can see that the experimental results match with the simulation results, and the maximum deviation between them is 0.091 mm. The dynamic error model of wMPS is validated to be reasonable.

Table 1. Deviation between simulation and practical experiments.

	Straightness Error/mm					
	Simulation Experiments		Practical Experiments		Deviation	
	5 m	7 m	5 m	7 m		
v = 30 mm/s	0.635	0.829	0.647	0.906	0.012	0.077
v = 60 mm/s	1.082	1.627	1.130	1.693	0.048	0.066
v = 120 mm/s	2.352	3.070	2.434	3.161	0.082	0.091

### 5. Conclusions

Dynamic coordinate measurement in industrial manufacturing is a popular issue which needs to be solved urgently. wMPS is now recognized as a promising solution because of its great potential. Nevertheless, the dynamic error caused by measurement principle gains little research and limits its further development. In this article, we first analyzed the mechanism of dynamic error. Then error modeling and simulations experiments about the uncertainty were carried out. The characteristics of the dynamic error were expounded. To validate the mathematical model, practical experiments were done and the results matched well with simulations. It laid the foundation for error compensation and performance improvement. It is noteworthy that we have made an assumption that the receiver moves at a constant speed. As stated in Section 1, wMPS has a measurement frequency of 20 Hz. It is often used in industrial manufacturing such as aircraft assembly and leveling, digital shipbuilding, and tooling monitoring. There is often no jump in acceleration for moving targets on such occasions. In a measurement cycle, i.e., 50 ms, the speed of the target can be seen as constant. The assumption is acceptable in practice.

In the future, we will focus our research on error compensation through hardware and software enhancement. The inertial measurement unit (IMU) can sense its own acceleration at a frequency of 400 Hz (such as STIM 300) and it will be embedded in the receiver. At each measurement cycle, the IMU will provide the real-time speed of the receiver. The model may be further optimized. Besides, the rotation speed of the transmitter has a great influence on the dynamic error. We will try more advanced shaft system and improve the rotation speed, which can also ensure a longer working life. As for software, a model-based compensation algorithm may be a good choice for dynamic error restriction. The development of machine learning provides novel ideas for our future work.

**Author Contributions:** Conceptualization, methodology, experiment design, and original draft, S.S. and L.Y.; experiment and analysis, R.D. and Z.Z.; software, J.L.; supervision and funding acquisition, J.Z. and C.L.; review and editing, S.S.

**Funding:** This research was funded by the National Key Research and Development Program of China, Grant No. 2017YFF0204802; National Natural Science Foundation of China, Grant Nos. 51721003, 51775380, 51705023 and Natural Science Foundation of Tianjin, Grant No. 16JCZDJC38100.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Zhou, K.L.; Liu, T.G.; Zhou, L.F. Industry 4.0: Towards future industrial opportunities and challenges. In Proceedings of the 12th International Conference on Fuzzy Systems and Knowledge Discovery (FSKD), Zhangjiajie, China, 15–17 August 2015; pp. 2147–2152.
2. Zhong, R.Y.; Xu, X.; Klotz, E.; Newman, S.T. Intelligent manufacturing in the context of industry 4.0: A review. *Engineering* **2017**, *3*, 616–630. [[CrossRef](#)]
3. Schmitt, R.H.; Peterek, M.; Morse, E.; Knapp, W.; Galetto, M.; Hartig, F.; Goch, G.; Hughes, B.; Forbes, A.; Estler, W.T. Advances in large-scale metrology—Review and future trends. *CIRP Ann.* **2016**, *65*, 643–665. [[CrossRef](#)]
4. Muralikrishnan, B.; Phillips, S.; Sawyer, D. Laser trackers for large-scale dimensional metrology: A review. *Precis. Eng. J. Int. Soc. Precis. Eng. Nanotechnol.* **2016**, *44*, 13–28. [[CrossRef](#)]
5. Franceschini, F.; Galetto, M.; Maisano, D.; Mastrogiacomo, L. Large-scale dimensional metrology (LSDM): From tapes and theodolites to multi-sensor systems. *Precis. Eng.* **2014**, *15*, 1739–1758. [[CrossRef](#)]
6. Aguado, S.; Santolaria, J.; Samper, D.; Velazquez, J.; Javierre, C.; Fernandez, A. Adequacy of technical and commercial alternatives applied to machine tool verification using laser tracker. *Appl. Sci.* **2016**, *6*, 16. [[CrossRef](#)]
7. Wang, Z.; Maropolous, P.G. Real-time error compensation of a three-axis machine tool using a laser tracker. *Int. J. Adv. Manuf. Technol.* **2013**, *69*, 919–933. [[CrossRef](#)]
8. Muralikrishnan, B.; Lee, V.; Blackburn, C.; Sawyer, D.; Phillips, S.; Ren, W.; Hughes, B. Assessing ranging errors as a function of azimuth in laser trackers and tracers. *Meas. Sci. Technol.* **2013**, *24*, 6. [[CrossRef](#)]
9. Hughes, B.; Forbes, A.; Lewis, A.; Sun, W.; Veal, D.; Nasr, K. Laser tracker error determination using a network measurement. *Meas. Sci. Technol.* **2011**, *22*, 12. [[CrossRef](#)]
10. Luhmann, T. Close range photogrammetry for industrial applications. *ISPRS J. Photogramm. Remote Sens.* **2010**, *65*, 558–569. [[CrossRef](#)]
11. Sun, B.; Zhu, J.G.; Yang, L.H.; Yang, S.R.; Guo, Y. Sensor for in-motion continuous 3d shape measurement based on dual line-scan cameras. *Sensors* **2016**, *16*, 15. [[CrossRef](#)] [[PubMed](#)]
12. Jung, K.; Kim, S.; Im, S.; Choi, T.; Chang, M. A photometric stereo using re-projected images for active stereo vision system. *Appl. Sci.* **2017**, *7*, 10. [[CrossRef](#)]
13. Montironi, M.A.; Castellini, P.; Stroppa, L.; Paone, N. Adaptive autonomous positioning of a robot vision system: Application to quality control on production lines. *Robot. Comput. Integr. Manuf.* **2014**, *30*, 489–498. [[CrossRef](#)]
14. Huang, S.R.; Shinya, K.; Bergstrom, N.; Yamakawa, Y.; Yamazaki, T.; Ishikawa, M. Dynamic compensation robot with a new high-speed vision system for flexible manufacturing. *Int. J. Adv. Manuf. Technol.* **2018**, *95*, 4523–4533. [[CrossRef](#)]
15. Di Leo, G.; Liguori, C.; Pietrosanto, A.; Sommella, P. A vision system for the online quality monitoring of industrial manufacturing. *Opt. Lasers Eng.* **2017**, *89*, 162–168. [[CrossRef](#)]
16. Liu, T.; Yin, S.B.; Guo, Y.; Zhu, J.G. Rapid global calibration technology for hybrid visual inspection system. *Sensors* **2017**, *17*, 16. [[CrossRef](#)] [[PubMed](#)]
17. Zhao, Z.Y.; Zhu, J.G.; Lin, J.R.; Yang, L.H.; Xue, B.; Xiong, Z. Transmitter parameter calibration of the workspace measurement and positioning system by using precise three-dimensional coordinate control network. *Opt. Eng.* **2014**, *53*, 8. [[CrossRef](#)]
18. Guo, F.Y.; Wang, Z.Q.; Kang, Y.G.; Li, X.N.; Chang, Z.P.; Wang, B.B. Positioning method and assembly precision for aircraft wing skin. *Proc. Inst. Mech. Eng. Part B J. Eng. Manuf.* **2018**, *232*, 317–327. [[CrossRef](#)]

19. Mei, Z.Y.; Maropoulos, P.G. Review of the application of flexible, measurement-assisted assembly technology in aircraft manufacturing. *Proc. Inst. Mech. Eng. Part B J. Eng. Manuf.* **2014**, *228*, 1185–1197. [[CrossRef](#)]
20. Guo, S.Y.; Lin, J.R.; Ren, Y.J.; Yang, L.H.; Zhu, J.G. Study of network topology effect on measurement accuracy for a distributed rotary-laser measurement system. *Opt. Eng.* **2017**, *56*, 8. [[CrossRef](#)]
21. Golub, G.H.; Van Loan, C.F. *Matrix Computations*, 4th ed.; The John Hopkins University Press: Baltimore, MD, USA, 2012; pp. 63–105.
22. Joint Committee for Guides in Metrology. *Evaluation of Measurement Data—Supplement 2 to the “Guide to the Expression of Uncertainty in Measurement”—Extension to Any Number of Output Quantities*; Joint Committee for Guides in Metrology: Paris, France, 2011.



© 2018 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).

Article

# Evaluation and Optimization of Task-oriented Measurement Uncertainty for Coordinate Measuring Machines Based on Geometrical Product Specifications

Yinbao Cheng <sup>1</sup>, Zhongyu Wang <sup>1,\*</sup>, Xiaohuai Chen <sup>2</sup>, Yaru Li <sup>1</sup>, Hongyang Li <sup>1</sup>, Hongli Li <sup>2</sup> and Hanbin Wang <sup>3</sup>

<sup>1</sup> School of Instrumentation and Optoelectronic Engineering, Beihang University, Beijing 100191, China; chengyinbao@buaa.edu.cn (Y.C.); Liyr022518@163.com (Y.L.); hongyangli@buaa.edu.cn (H.L.);

<sup>2</sup> School of Instrument Science & Opto-electronic Engineering, Hefei University of Technology, Hefei 230009, China; xhchenhf@163.com (X.C.); lihongli\_hgd@163.com (H.L.)

<sup>3</sup> Fujian Metrology Institute, Fuzhou 350003, China; wanghanbinbin@163.com

\* Correspondence: mewan@buaa.edu.cn; Tel.: +86-10-8233-8881

Received: 24 September 2018; Accepted: 17 December 2018; Published: 20 December 2018



**Abstract:** Measuring instruments are intended to be intelligent, precise, multi-functional and developing multidirectionally, scientific, and reasonable; the reliable evaluation of measurement uncertainty of precision instruments is also becoming more and more difficult, and the evaluation of the Coordinate Measuring Machines (CMM) measurement uncertainty is among the typical problems. Based on Geometric Product Specification (GPS), this paper has systematically studied the CMM uncertainty for evaluating the size and geometrical errors oriented toward measurement tasks, and thus has realized the rapid and reliable evaluation of the CMM uncertainty for task-oriented measurement. For overestimation of the CMM uncertainty for task-oriented measurements in the initial evaluation, a systematic optimization solution has been proposed. Finally, the feasibility and validity of the evaluation model and the optimization method have been verified by three different types of measurement examples of diameter, flatness and perpendicularity. It is typical and representative to systematically solve the problem of the CMM uncertainty for evaluating the measurement tasks targeted at dimensions and geometric errors, and the research contents can be effectively applied to solve the uncertainty evaluation problems of other precision instruments, which are of great practical significance not only for promoting the combination of modern uncertainty theory and practical applications but also for improving the application values of precision measurement instruments.

**Keywords:** measurement uncertainty; coordinate measuring machines; evaluation and optimization; geometrical product specifications

## 1. Introduction

Geometric measurement is the foundation of modern metrology, being the earliest and largest important branch in the field of measurement and also the foundation for the development of modern science and technology. At present, geometric measurement in various fields has developed different types of measurement techniques or instruments and has presented a trend of mutual integration; coordinate measurement technology is undoubtedly among the best [1,2]. Traditional geometric measurement instruments mainly use optical vernier technology to improve the measurement accuracy and resolution based on the measurement mode of geometric theory, so that the geometric quantity

cannot be expressed and transmitted digitally, which has been for a long time a difficult technical bottleneck overcome in the field of mechanical manufacturing. In essence, the presence of CMM has solved the problem that geometric measurement cannot be digitalized [3,4]. Coordinate measurement technology is the most common and fundamental measurement technology in the field of modern machinery manufacturing, especially in aerospace, automobile manufacturing, mold processing, and other industries. At present, the whole life cycle of the product development, design, processing, measurement, acceptance, use, maintenance, scrapping and so on must follow the GPS standard system in the field of manufacturing. This standard system covers many aspects such as product size, geometric form and surface appearance. The GPS can effectively eliminate some technical barriers in international trade, boasting an important role in promoting the integration of the global economy as an important technical base for modern manufacturing informatization and quality management. The new generation of the GPS system has overcome the problem of digital representation and transmission of geometric quantities. The traditional measurement mainly relies on optical vernier technology, in which the role of the computer is limited. The development of the CAD technology and the presence of CMM has essentially solved the problem that geometric quantities cannot be expressed and transmitted digitally, which makes it possible to initialize a new generation of the GPS system based on measuring mathematical theory. As a measuring instrument for geometric characteristics of products, CMM plays its role under the guidance of the GPS, boasting important measuring equipment essential to quality control in modern enterprises [5,6].

As an important parameter to characterize the quality of the measurement results, measurement uncertainty reflects the credibility of measurement results. To give scientific and proper evaluation of measurement uncertainty is an important factor to guarantee the development of modern measuring science [7,8]. CMM can complete the measurement of spatial geometric elements (including size, geometrical error parameters) more conveniently, featuring a large measurement range, high measurement efficiency and strong measurement versatility. However, CMM is an omnipotent geometric measurement instrument, so the diversity of its measurement strategies makes the process for evaluating the measurement uncertainty of different measurement tasks quite different from the results; simultaneously there are many error sources affecting CMM measurement uncertainty, and the transitive relationship is difficult to determine between such error sources and measurement results. Therefore, at present, CMM mostly cannot provide reports on the measurement uncertainty during measurement, but usually gives an estimated value of the measured quantity only [9].

Physikalisch-Technische Bundesanstalt (PTB), Germany, National Physical Laboratory (NPL), Britain, Istituto di Metrologia "G. Colonnetti" (IMGC), Italy, and other institutions have taken the CMM measurement uncertainty evaluation as an important research direction. PTB has proposed an expert system scheme for evaluating the uncertainty of a coordinate measuring machine and has researched the influence of the measurement strategy on the CMM uncertainty [10,11]. NPL has standardized the CMM measurement strategy so as to obtain accurate measurement results [12]. IMGC has carried out the research on evaluating the uncertainty of the task-oriented coordinate measuring machine by using computer simulation technology [13]. J. Sladek et al. have proposed an evaluation method of coordinate measurement uncertainty with the use of a virtual machine model based on the Monte Carlo method [14]. W. Jakubiec et al. have discussed the problem of evaluating the CMM measurement uncertainty by Type-B methods and carried out relevant research on CMM uncertainty evaluation based on the GPS specifications [15,16]. P.B. Dhanish et al. have studied the influence of coordinate point selection on the uncertainty of the CMM measurement "circle" [17].; J. Beaman and E. Morse have studied the uncertainty evaluation of the CMM-specific measurement tasks [18]. J. Feng et al. have studied the experimental problem of CMM uncertainty evaluation [19]. J.P. Kruth et al. have studied how to evaluate the uncertainty of the CMM shape measurement tasks based on the Monte Carlo method [20]. F. Aggogeri et al. have designed a simplified process for evaluating the CMM uncertainty by means of simulation experiments [21]. G.X. Zhang proposed a CMM error modeling method based on rigid body models and developed in-depth research on error identification and compensation [22].



R.G. Wilhelm et al. considered task-specific uncertainty as the measurement uncertainty associated with the measurement of a specific feature using a specific measurement plan [23]. H. Haitjema has discussed task-specific uncertainty estimation in dimensional metrology, and these principles can be used as well for e.g., roughness, roundness, cylindricity, flatness, and CMM measurements [24]. K. Takamasu et al. has formulated methods of estimating uncertainties using the coordinate measuring system after calibration [25].

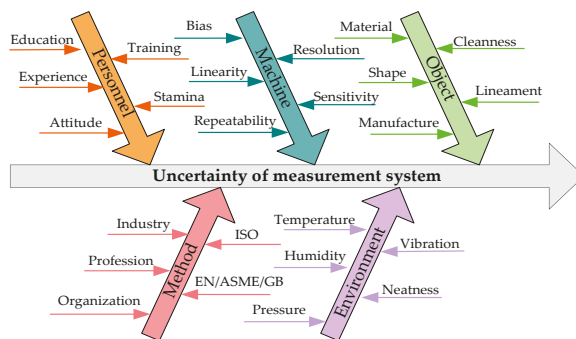
So far, the CMM task-oriented research has been carried out earlier, focusing on the uncertainty evaluation of specific measurement tasks in practical applications, however, there have been few research results on evaluating the uncertainty of the CMM system-wide measurement tasks targeting sizes and geometric errors, and most of such results have only involved individual aspects of the ISO15530 standards [26,27]. The intelligent evaluation of the CMM measurement uncertainty by virtue of computer simulation technology is an important trend of development in the future; it is especially worth discussing how to combine the virtual measurement technology with uncertainty evaluation, but the software for uncertainty evaluation has had its practical applications limited due to involvement in intellectual property rights. With the increasing improvement of the research on coordinate measurement technology, more theories and methods for CMM measurement uncertainty have been put forward continuously, the ISO15530 series of standards have been gradually formed for evaluating the CMM task-oriented measurement uncertainty and a variety of task-oriented uncertainty evaluation methods have been put forward, such as the strategy of applying repeated measurement, method of alternative measurement, computer simulation and expert analytical judgment. However, the GPS standard system still has a lot of standards to be added and still there is no content of the uncertainty evaluation model.

The simplicity, practicability and economic efficiency is very important in practical applications for evaluating the CMM measurement uncertainty. Within the framework of the GPS standard system [26–29], it is of great significance for improving the application value of the measuring instrument by studying the CMM uncertainty evaluation of task-oriented measurement for quickly and reliably evaluating the uncertainty of the measurement results.

## 2. Evaluation Method

### 2.1. Source Analysis of Uncertainty

During CMM measurement, all relevant factors may have an impact on the measurement results. The sources of uncertainty in the coordinate measuring system can be divided into five categories according to the analysis method of “personnel, machine, object, method and environment” commonly used in product quality management as shown in Figure 1: Uncertainty caused by CMM instrument’s own errors, measured workpiece, surveyors, measuring method, and external environment.



**Figure 1.** The uncertainty sources of the measurement system. EN: European Norm; ASME: American Society of Mechanical Engineers; GB: Chinese National Standards.

(1) Uncertainty caused by surveyors

In addition to professional qualities, the uncertainty caused by surveyors is mainly characterized by the differences of the measurement strategies caused by different understandings of the measurement specifications, such as clamping positioning, coordinate system establishment, sampling strategy, and probe configuration.

(2) Uncertainty caused by the instrumental errors

This refers to the uncertainty component produced by the deviation of the CMM's own metering characteristics from the ideal characteristics, including the uncertainty caused by design, standard quantity, detection system, dynamic characteristics, fitting and evaluation algorithm, and other factors, mainly including the impacts of 21 items of CMM mechanism errors and probe system errors, etc. on the measurement results [30]. Usually, CMM can correct 21 items of mechanism errors and calibrate the probe system. The CMM's own errors are reflected in the influence of residual system errors after correction and calibration.

To ensure that the CMM measurement accuracy meets the relevant requirements, it is usually necessary to carry out acceptance and re-inspection testing of the CMM measuring characteristics by means of the measurement calibration programs. In the GPS ISO10360 series standards [26,27], the CMM performance evaluation has been defined and relevant performance parameters and evaluation methods of the measuring instrument have been given, such as the indication, detection and scanning detection errors of the dimensional measurement as well as the dimensional and geometrical errors of the universal detection system. The technical parameters corresponding to the above-mentioned performance parameters are generally given before CMM is delivered, that is, the maximum permissible errors. In practical measurement, the Maximum Permissible Indication Error ( $E_{L, MPE}$ ) is mainly related to the errors of the distance and other dimensional elements, and the  $MPE_p$  indicates the error of the whole measurement system in a very small test space, generally affecting the form measurement.

(3) Uncertainty caused by measured workpiece

The geometric and physical characteristics of the workpiece itself will affect the measurement results, for example, the form, surface waviness and surface roughness of the measured elements have impacts on the measurement results together with the sampling strategy and probe configurations; the thermal expansion coefficient of the workpiece and its changes produce uncertainty in the temperature compensation; the force deformation is also affected by the workpiece positioning and clamping mode. Therefore, it is necessary to combine the specific measurement tasks to evaluate the uncertainty introduced by the measured workpiece.

(4) Uncertainty caused by measurement methods

In the actual measurement, measurement specifications only give general guidance and constraints for the measurement process, resulting in a certain degree of randomness in measurement methods. If currently there are no standards or specifications to clearly and meticulously stipulate the choice of the CMM sampling strategy and probe configurations, different sampling strategies and probe configurations will result in inconsistency of the measurement results.

(5) Uncertainty caused by environmental factors

The temperature, humidity, temperature spatial and temporal gradient, vibration, dust, and other environmental factors may have impacts on the measurement results. During CMM dimensional measurement, temperature compensation is usually needed. If the temperature changes, the CMM grating ruler and workpiece thermal expansion coefficient will all cause uncertainty through the process of temperature compensation.

The evaluation of measurement uncertainty should consider all factors that may affect the measurement results and focus on the specific measurement tasks. Based on the above analysis, it can be concluded that the sources of CMM task-oriented measurement uncertainty have the following characteristics: Wide sources and rich types, interconnection between sources of uncertainty, difficulty in quantifying sources of uncertainty and close correlation between sources of uncertainty

and measurement tasks. Because of the complexity and quantification difficulty of the sources of uncertainty, it is a key problem to establish an uncertainty evaluation model which can fully reflect the influence of the sources of uncertainty and their relationship on the measurement results.

## 2.2. Uncertainty Evaluation Model

The quantitative statistical analysis method can carry out direct statistics and analysis of the measurement results; the output and input quantities have identical units and the measurement results can be obtained without relying on any other quantity subject to a functional relationship with the quantity being measured; therefore, compared with the error traceability method, it features a simple transfer relation for uncertainty, it has convenient model use and it has richer adaptability to the requirements of simplicity, rapidness and practicability for task-oriented measurement uncertainty. However, disadvantageously, the quantitative statistical analysis method cannot clearly grasp the law of uncertainty transmission and may cause the consequence of "excessive estimation" of uncertainty components. Moreover, due to the complexity of error sources, it is difficult for general surveyors to ensure that no uncertainty source is repeated or missed in the analysis of error sources. The most obvious problem is that currently most of the relevant researches based on error traceability use indication errors to quantize individual points in the CMM space, and the indication errors essentially limit the measurement results, so the quantification of uncertainty in single point measurement is in itself characterized by the problem of "over estimation". The measurement system analysis method widely used in product quality management has summarized the characteristic indexes of the measurement of the whole measurement system as resolving power, bias, linearity, stability, repeatability, and reproducibility, which can comprehensively reflect the impact of uncertain system and random errors of the measurement system on the measurement results; the uncertainty modeling method based on statistics and analysis of the measurement characteristics has given the guiding ideology for evaluating the CMM task-oriented measurement uncertainty from the top design.

The measurement task of dimensional and geometrical errors mainly includes two essential problems, namely, how to extract the actual elements and how to evaluate the ideal elements. The former determines the detection method while the latter depends on the evaluation method. The new generation of uncertainty evaluation theory has summarized the uncertainty caused by the evaluation method into the category of "normative uncertainty". Therefore, the influence of the method for ideal element evaluation is not considered in the evaluation of "measurement uncertainty", which is particularly important for evaluating the measurement uncertainty of geometrical errors [31]. Thus, the sources of the CMM task-oriented uncertainty can be analyzed as follows:

### (1) Uncertainty components caused by bias and linearity

The influence of bias and linearity on the CMM measurement results is reflected in the uncertainty component  $u_E$  caused by the indication or detection errors of the measuring instrument. For the purpose of safety, the "overestimation" should be adopted, with the CMM  $E_{L, MPE}$  and  $MPE_P$  used to quantize  $u_E$ . In calibration, the CMM indication or detection errors have considered the influence of such factors as probe configurations of the measuring instrument, method for coordinate system establishment, measuring object clamping, space position and environment. Therefore, when  $E_{L, MPE}$  and  $MPE_P$  are used to quantize  $u_E$ , the influence of the above uncertainty sources is also included.

### (2) Uncertainty components caused by resolution and repeatability

There is a certain correlation between the resolution and repeatability of the instrument, so CMM only needs to consider the uncertainty component  $u_r$  caused by repeatability.

### (3) Uncertainty components caused by stability and reproducibility

The stability index usually has a significant influence on the electronic measuring instruments but less on the CMM geometric measurement; the stability is equivalent to the reproducibility caused by time variation, so this uncertainty component can be ignored.

The uncertainty component  $u_R$  caused by CMM task-oriented reproducibility indicates the consistency between the measurement results of the same measuring object when the measurement

conditions are changed, that is, different measurement conditions such as personnel changes and different measurement strategies have led to the variations of the measurement mean value for the same measurement task. In the actual uncertainty component quantization, the repeatability experiment is used to determine the influence of the component. For CMM, multiple groups of repeatability quantization experiments are completed by different surveyors according to the sampling strategy with the measuring points determining themselves based on the habits for measurement.

Therein, the uncertainty component caused by sampling strategy is affected by the number and distribution of the sampling points. Sampling point distribution determines the probability of extracting the extreme error point of the errors for measured elements when the number of sampling points is the same. It has become a consensus in CMM applications that minimal measurement uncertainty occurs when the sampling points are evenly distributed. The number of sampling points reflects the CMM ability to extract the form information of the measured elements. When there is a small number of sampling points, the measuring points will have a higher probability of excluding the extreme point of the form tolerance. From the point of view of information extraction only, the larger the number of sampling points, the better. However, the measurement time similar to contact triggered CMM will increase sharply as the number of sampling points increases, which has violated the CMM characteristics of measurement efficiency. At the same time, the excessive increase of the number of measuring points will multiply the impact of the CMM residual mechanism errors on the measurement results. Generally, for contact triggered CMM, the suitable number and preferable distribution of sampling points can be confirmed according to BS7172 [12]. As the evaluation software requires, the number of measurement points just needs to be greater than the minimum number of the points required by the mathematical requirements of the geometric elements to be measured.

Therefore, the primary model is as follows for evaluating the CMM task-oriented uncertainty based on the method for statistical analysis of measurement characteristics:

$$u_c = f(u_E, u_r, u_R) \tag{1}$$

Formula (1) shows three inputs, namely, the maximum permissible error  $\delta_E$ , measurement repeatability  $\delta_r$  and measurement reproducibility  $\delta_R$  of the instrument. The expected values are 0 for all inputs and all are the measurement characteristics of the measurement task output  $Y$ .  $y$  is the measurement estimated value of  $Y$ . Therefore, the analysis model is as follows for measurement uncertainty:

$$Y = y + \delta_E + \delta_r + \delta_R \tag{2}$$

Based on the Guide to the expression of uncertainty in measurement (GUM) [32], Formula (1) can be written as:

$$u_c \sqrt{u_E^2 + u_r^2 + u_R^2} \tag{3}$$

### 3. Modeling for Typical Task Uncertainty Evaluation

#### 3.1. Uncertainty Model for Dimensional Measurement Task

The CMM dimensional measurement tasks mainly include distance, diameter, radius and so on. In time of CMM acceptance and reexamination,  $E_{L,MPE}$  is used to express the CMM ability for dimensional measurement. Therefore, the uncertainty component  $u_E$  caused by the bias and linearity of the CMM dimensional measurement tasks is as follows:

$$u_E = \frac{E_{L,MPE}}{\sqrt{3}} \tag{4}$$

Under the same conditions, measure the target workpiece repeatedly and calculate the laboratory standard deviation of a single measurement via Bessel formula:

$$S = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (y_i - \bar{y})^2} \tag{5}$$

Where,  $n$  indicates the number of repeated measurements,  $y_i$  is the measured value of the  $i$ th measurement and  $\bar{y}$  is the average value of the repeated measurement column.

If  $N$  times' measurement mean is taken as the best estimation, the uncertainty component  $u_r$  caused by measurement repeatability is as follows:

$$u_r = \frac{S}{\sqrt{N}} = \sqrt{\frac{1}{N(n-1)} \sum_{i=1}^n (y_i - \bar{y})^2} \tag{6}$$

Several different surveyors can determine the measurement strategy and probe configurations according to their own habits for measurement, carry out  $m$  groups of independent measurements repeatedly of the dimensional parameters of the workpiece under test, and set the average value of Group  $j$  repeated measurement column as  $\bar{y}_j$ . Then, regard the mean value  $\bar{y}_j$  of the column consisting of  $m$  groups of measurements as a measurement column, work out the mean value  $\bar{\bar{y}}$  of the measurement column; the uncertainty component  $u_R$  caused by the reproducibility is as follows:

$$u_R = \sqrt{\frac{1}{(m-1)} \sum_{j=1}^m (\bar{y}_j - \bar{\bar{y}})^2} \tag{7}$$

Based on the above analysis, the universal model as follows can be achieved for evaluating the uncertainty of the CMM dimensional measurement tasks according to Formula (3):

$$u_{uc} \sqrt{\frac{E_L, MPE^2}{3} + \frac{1}{N(n-1)} \sum_{i=1}^n (y_i - \bar{y})^2 + \frac{1}{(m-1)} \sum_{j=1}^m (\bar{y}_j - \bar{\bar{y}})^2} \tag{8}$$

### 3.2. Uncertainty Model for Form Error Measurement Task

The form error is defined as the variation of the measured actual shape elements relative to the ideal shape elements [33]. Form measurement is not the specialty of the CMM measurement function unless there is a commitment in the measurement time. Compared with roundness meters, autocollimators, level meters and other special form tolerance instruments, CMM is slightly inferior in measurement accuracy, but its functional diversity can significantly improve the comprehensive efficiency of measurement. Therefore, CMM, when selected for form measurement, should be compatible with the competence of the measuring instrument, that is, the elements to be measured should be the task "measurable" by CMM when normal measurement conditions are satisfied, and the requirements for detection efficiency should be fully considered when sampling strategy is selected.

Different from dimensional measurement tasks, the form measurement belongs to miniature measurement and is comparatively sensitive to "overestimation" in time of uncertainty evaluation. Dimensions are characterized by the measurement of absolute values, but form error are considering relative changes in a very small test space. In the analysis of indication errors, linear impacts can be ignored, just focusing on the bias of the form measurement indications caused by residual system errors of the measuring instrument. The CMM ability for form detection is represented by  $MPE_P$  and is calibrated by the sphericity of the standard ball, which essentially reflects the comprehensive impact of the residual system errors on the form measurement results in different directions and at different

positions. It is more reliable for using  $MPE_P$  to evaluate the uncertainty component caused by the indication errors of the CMM form error measurement; the formula for qualification is as follows:

$$u_E = \frac{MPE_P}{\sqrt{3}} \tag{9}$$

Similarly, the uncertainty components caused by repeatability and reproducibility can be obtained according to Formula (6) and Formula (7). According to formula (3), the universal model as follows can be achieved for evaluating the uncertainty of the CMM form measurement tasks according to Formula (3):

$$uu_c \sqrt{\frac{MPE_P^2}{3} + \frac{1}{N(n-1)} \sum_{i=1}^n (y_i - \bar{y})^2 + \frac{1}{(m-1)} \sum_{j=1}^m (\bar{y}_j - \bar{\bar{y}})^2} \tag{10}$$

### 3.3. Uncertainty Model for Location and Orientation Errors Measurement Tasks

The location and orientation errors show the positional relationship between two or more geometric elements, i.e. the relationship between the measured elements and the reference elements. It is defined as the variation of the measured actual elements to the ideal elements with certain direction or location, able to be divided into orientation error and location error [33].

The form and position errors are both variations between the actual and ideal elements, but the methods are different for evaluation. Form measurement does not require reference elements. By sampling on the surface of the workpiece to be measured, enough sampling points of the measured elements can be obtained. Computer software can be used to fit lines, planes, circles or cylinders, and calculate the maximum variation from sampling points to fitting elements as the value of the current form error. However, the location and orientation errors are the positional relationship between two elements. It is necessary to measure the two elements, work out one fitting element as a benchmark and calculate the maximum variation from the other element to the benchmark as the measurement value of the current position error. In the principle of measurement, it is basically the same as the dimensional measurement, so the location and orientation error measurement can be concluded as miniature dimensional measurements. While evaluating the uncertainty of the CMM location and orientation tolerances measurement task,  $E_{L, MPE} = A + B \cdot L$  may be used to represent the influence of its indication error. However, the location and orientation errors are still greatly different from dimensional measurement due to the impacts of the instrument offset and linearity; different types of position errors are differently affected by the offset and linearity of the measuring instrument. Here follows the analysis and discussion of the uncertainty components for position errors caused by the indication errors of the measuring instrument.

CMM parallelism measurement task: First measure the reference plane, fit out its plane formula, then sample a characteristic point on the measured plane, calculate the longest distance  $l_{max}$  and the shortest distance  $l_{min}$  from the sampling point to the reference fitting plane, and the maximum distance difference is the parallelism tolerance  $t$  between the two planes:

$$t = l_{max} - l_{min} \tag{11}$$

Where, the influence of the indication errors is both  $MPE_E$  for the longest distance  $l_{max}$  and the shortest distance  $l_{min}$ , then

$$t = (l_{max} \pm E_{L, MPE}) - (l_{min} \pm E_{L, MPE}) \tag{12}$$

Fully consider the influence of  $E_{L, MPE}$  and the error range of the overestimated parallelism tolerance  $t$  to get:

$$t = (l_{max} - l_{min}) \pm 2E_{L, MPE} \tag{13}$$

If it is given that the composite distribution of two identical uniform distributions is a triangular distribution, the uncertainty component  $u_{EW_1}$  of the CMM parallelism measurement task caused by the indication error is as follows according to type B evaluation method [32]:

$$u_{EW_1} = \frac{2E_{L, MPE}}{\sqrt{6}} \tag{14}$$

The tolerance zone of angularity and perpendicularity is the range defined by two parallel planes whose spacing is equal to the tolerance value  $t$ . Both indicate the degree to which the measured element maintains a certain angle relative to the reference element. The difference is that the two parallel planes of angularity incline to the reference at a given theoretical angle (excluding  $0^\circ$ ,  $90^\circ$  and  $180^\circ$ ) and such two parallel planes are vertical (or parallel) to the reference. When measuring the angularity and perpendicularity, CMM will determine the fitting element (direction) according to the reference element and theoretical angle and calculate the difference of the limit distance  $t = L_1 - L_2$  between the measured elements and the fitting elements. Therein,  $L_1$  and  $L_2$  indicate the distances from the measured elements to the fitting elements.

Relative to the dimensional measurement,  $L_1$  and  $L_2$  both are miniature dimensions, not considering the influence of linearity but only the constant term of the measuring instrument deviation. The uncertainty  $u_{EW_2}$  of angularity and perpendicularity measurement tasks caused by the indication errors is as follows according to uncertainty type B evaluation method:

$$u_{EW_2} = \frac{2 \cdot A}{\sqrt{6}} \tag{15}$$

Where,  $A$  is the constant term of  $E_{L, MPE} = A + B \cdot L$ .

The position degrees can be divided into point, line and plane position degrees. The position degree is essentially two times the maximum distance between extracted elements (point, line and plane) and fitting elements (determined by reference elements and theoretical size), that is, the micro-level dimensional measurement, which can neglect the influence of dimension sizes and only consider the constant term of the deviation.

Similarly, coaxiality represents the maximum distance from the point of the measured element to the reference element (point or line) and is essentially a micro-level dimensional measurement. The indication error also only considers the constant term of the deviation.

The uncertainty component  $u_{EW_3}$  of the CMM position degree and coaxiality measurement tasks caused by the indication errors is as follows according to the uncertainty type B evaluation method:

$$u_{EW_3} = \frac{A}{\sqrt{3}} \tag{16}$$

Symmetry is essentially two times the maximum difference between the two extracted elements to be measured and the reference vertical distance, namely,  $t = 2(L_1 - L_2)$ . Essentially, it is the difference of the micro-level dimensions and the indication error only considers the constant term of the deviation. Then, the uncertainty component  $u_{EW_4}$  caused by the indication error is as follows for CMM symmetry measurement task:

$$u_{EW_4} = \frac{2 \cdot A}{\sqrt{6}} \tag{17}$$

Similarly, the uncertainty components caused by repeatability and reproducibility of position error measurements are Formula (6) and Formula (7). Substitute Formula (14)–Formula (17) into Formula (3) and a universal model can be obtained for evaluating the uncertainty of the CMM position measurement tasks.

#### 4. Method for Optimizing Measurement Uncertainty

##### 4.1. Secondary Optimal Evaluation of Uncertainty Components

The uncertainty component features "overestimation" in time of quantization, which is the possible upper bound of the uncertainty component under worse conditions to ensure the safety and reliability of the evaluation results. Referring to the new generation of GPS uncertainty evaluation theory, develop an uncertainty management program as shown in Figure 2 for a given task-oriented measurement process [34]. In this management program, the measurement tasks, conditions and methods, etc. as well as relevant matters related to the measurement process are given, so its core task is to optimize the estimation of the measurement uncertainty.

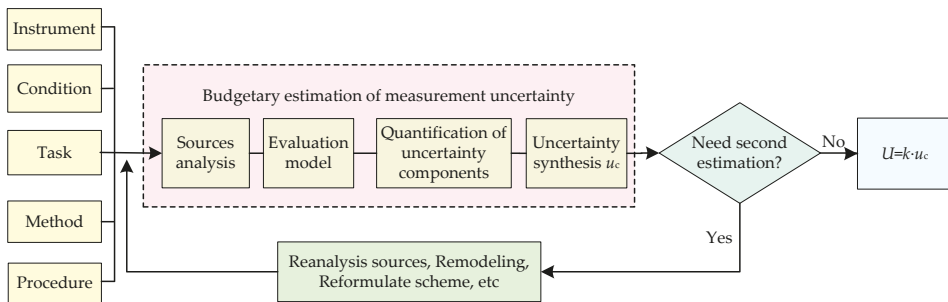


Figure 2. Management program of task-oriented uncertainty by definitized measurement process.

From the initial evaluation of the measurement uncertainty, we can find the dominant contribution factors to uncertainty, redesign the scheme for quantization of the dominant uncertainty component, make it more close to the actual situation of the uncertainty component and avoid the influence of excessive estimation. However, such an uncertainty component will somewhat increase the measurement time and economic costs during secondary optimal evaluation, and the surveyors should make corresponding improvements according to their own measurement needs. The optimal evaluation of the uncertainty components described in this section mainly cover the uncertainty components introduced by indication errors and reproducibility.

The uncertainty component caused by the indication errors as described in Section 3 is estimated according to the CMM maximum permissible error, i.e. the error limit. During secondary evaluation, physical standards similar to the object under test or calibrated workpiece can be adopted to calibrate the indication errors and thus to obtain a more accurate upper bound estimation of the indication error. See Table 1 for CMM  $E_{L=60, MPE}$  and  $MPE_p$  calibration results by using standard gauge blocks and master balls with calibration uncertainty able to be neglected. It is shown that the uncertainty component introduced by the indication errors after secondary evaluation has been significantly reduced compared with the initial evaluation.

Table 1. Calibration results of indication error and probing error of CMM.

Error	Error Limit	Uncertainty Component
$E_{L=60, MPE}$	3.24 $\mu\text{m}$ ( $L = 60 \text{ mm}$ )	1.87 $\mu\text{m}$
$E_{L=60}$	1.4 $\mu\text{m}$ ( $L = 60 \text{ mm}$ )	0.81 $\mu\text{m}$
$MPE_p$	3.5 $\mu\text{m}$	2.02 $\mu\text{m}$
$P$	1.2 $\mu\text{m}$	0.69 $\mu\text{m}$

The uncertainty component caused by reproducibility depends on the operation difference of the surveyor. The greater uncertainty component of reproducibility indicates that the operation



of the surveyor has a significant impact on the measurement results, and a more normalized and standardized measurement program file should be formulated to reasonably reduce the impacts on the evaluation results of reproducibility. For example, the secondary evaluation of the uncertainty component introduced by sampling strategy: When more prior information of the measurement task is available, reasonably and properly reduce the change range of sampling points; use spectral analysis to determine the best sampling strategy and estimate the changes of the measurement results from the difference between the actual measurement points and the optimal sampling points to evaluate the uncertainty caused by the sampling strategy. When the number of optimal sampling points is known, the influence of uncertainty introduced by this factor can be ignored.

#### 4.2. Real-Time Updating of Repeatability Uncertainty Component

The method of calculating uncertainty by means of statistical analysis based on experimental data is called the type A evaluation method [32]. The uncertainty components introduced by repeatability are usually obtained by experimental pre-evaluation and determined only by one experiment, so the information contained in the results is limited and poor in representativeness; moreover, the working states of instruments and workpiece may change with time during the experiment, so the uncertainty component determined by one evaluation experiment cannot reflect the latest information in the evaluation process.

If the uncertainty component caused by repeatability is not changed after evaluation, the evaluation results of the uncertainty will not sufficiently reflect the latest information in the process of measurement. For the measurement of the same batch of workpieces with the same machining accuracy, the cost is rather high for repetitive experiments of each workpiece respectively, so it is impossible to carry out a large number of uncertainty evaluation experiments at any time for batch products. If the uncertainty components caused by repeatability can be updated continuously in real time based on daily measurement data, then the current and historical information can be fully integrated into the uncertainty evaluation results and the latest status of the CMM same or similar measurement tasks can be reflected in real time.

Therefore, this paper has proposed using the Bayesian information fusion method to establish an information fusion model based on Bayesian formula and to achieve real-time and continuous updates of uncertainty components, so that the evaluation results of task-oriented CMM measurement uncertainty could reflect the latest information in the measurement system in real time, and improve the reliability of the uncertainty evaluation results.

The repeated measurements generally obey normal distribution and the prior distribution is identical to the posterior distribution in form during the process of information fusion. Therefore, the conjugate Bayesian method can be used to update the uncertainty components of repeatability continuously.

Set the measurement series of repetitive experiments as  $X = (x_1, x_2, x_3, \dots, x_n)$  and  $X \sim N(\theta, \sigma^2)$ . Then the Bessel formula can be used to work out the standard uncertainty component of repeatability in a single result:

$$u = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n(n-1)}} \tag{18}$$

If the number of measurements in the first repeatability experiment is  $n_0$  and the measurement result is  $X = (x_{01}, x_{02}, x_{03}, \dots, x_{0n})$ , according to the evaluation method of conjugate Bayesian uncertainty, the conjugate prior distribution of  $\sigma^2$  is as follows:

$$\pi(\sigma^2) = \frac{\sqrt{S_0^{n_0-1}}}{\sqrt{2^{n_0-1}} \Gamma(\frac{n_0-1}{2})} \left(\frac{1}{\sigma^2}\right)^{\frac{n_0+1}{2}} \exp\left(-\frac{S_0}{2\sigma^2}\right) \propto \left(\frac{1}{\sigma^2}\right)^{\frac{n_0+1}{2}} \exp\left(-\frac{S_0}{2\sigma^2}\right) \tag{19}$$

where,  $S_0 = \sum_{j=1}^n (x_{0j} - \bar{x}_0)^2$ .

Therefore, the uncertainty component of the first repeatability experiment is:

$$u_0 = \sqrt{\frac{S_0}{n_0 - 1}} \tag{20}$$

Assuming that the number of measurements in the second repeatability experiment is  $n_1$  and the measurement result is  $X = (x_{11}, x_{12}, x_{13}, \dots, x_{1n})$ , use the latest repeatability data to update  $\sigma^2$  and its likelihood function is as follows:

$$l(\sigma^2|x) \propto \left(\frac{1}{\sigma^2}\right)^{\frac{n_1}{2}} \exp\left(-\frac{S_1}{2\sigma^2}\right) \tag{21}$$

where,  $S_1 = \sum_{j=1}^n (x_{1j} - \bar{x}_1)^2$ .

Calculate the  $\sigma^2$  posteriori probability density function and its distribution according to the Bayes formula:

$$\pi(\sigma^2|x) \propto \pi(\sigma^2)l(\sigma^2|x) \propto (\sigma^2)^{-\frac{n_0+n_1-1}{2}-1} \exp\left(-\frac{S_0+S_1}{2\sigma^2}\right) \tag{22}$$

$$\pi(\sigma^2|x) \sim \Gamma^{-1}\left(\frac{n_0+n_1-1}{2}, \frac{S_0+S_1}{2\sigma^2}\right) \tag{23}$$

Therefore, the updated repeatability uncertainty component is as follows:

$$u_1 = \sqrt{\frac{S_0+S_1}{n_0+n_1-3}} \tag{24}$$

The general formula can be concluded as follows for updating the repeatability uncertainty component:

$$u_1 = \sqrt{\frac{(n_0-1)u_0^2 + \sum_{j=1}^{n_1} (x_{1j} - \bar{x}_1)^2}{n_0+n_1-3}} \tag{25}$$

### 5. Experimental Analysis

HEXAGON Micro-Hite 3D DCC CMM was used to measure the workpiece as shown in Figure 3. The measurement tasks include diameter, flatness and perpendicularity respectively. See Figure 4 for the experiment. The maximum permissible error of CMM is:  $E_{L, MPE} = (3 + L/250) \mu\text{m}$ ,  $MPE_P = 3.5 \mu\text{m}$ . The TesaStar-i touch trigger probe is applied to this experiment. The type of tip in routine repeated experiments is 3BY40, namely that the sphere diameter is 3 mm and the pole length is 40 mm. In routine reproducibility tests, the sampling point's numbers of plane and circle are nine, and the sampling points are evenly distributed.

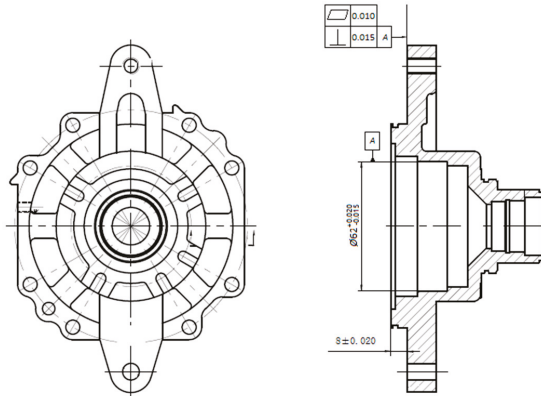


Figure 3. The drawing marking of the measured part.

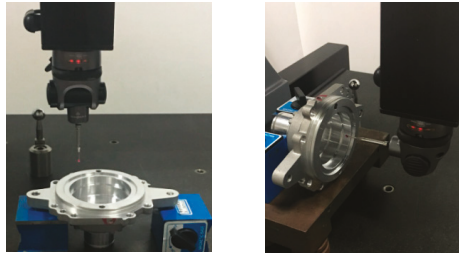


Figure 4. The experimental of measurement.

5.1. Example for Evaluating Uncertainty of Diameter Measurement

The height of the measured cylinder is 14 mm, the upper tolerance limit of diameter is +0.020 mm and the lower tolerance limit of the diameter is −0.015 mm. The measured cylinder is made by a milling-tool, its roughness is 0.8 μm, and the effect of surface roughness for diameter measurement is negligible.

According to Formula (4), the uncertainty component caused by indication errors can be calculated as follows:

$$u_E = \frac{E_{L, MPE}}{\sqrt{3}} = \frac{1}{\sqrt{3}} \times \left( 3 + \frac{62}{250} \right) = 1.875 \mu\text{m} \tag{26}$$

Under the condition of repeatability, complete 10 times' continuous and rapid measurement of the circle to be measured. See Table 2 for data on the repeatability experiment. Calculate the standard deviation of the repeatability experiment according to the experimental data in Table 2:

$$S_r = \sqrt{\frac{\sum_{i=1}^n (d_i - \bar{d})^2}{(n - 1)}} = 0.51 \mu\text{m} \tag{27}$$

Table 2. The experimental data of diameter measurement for repeatability detection.

No.	1	2	3	4	5	6	7	8	9	10
measured value $d_i$ (mm)	62.0010	62.0011	62.0010	61.9998	62.0011	62.0002	62.0008	61.9998	62.0007	62.0008

If the average value of three times' measurements is taken as the best estimate of the diameter measurement, the standard uncertainty caused by measurement repeatability is as follows:

$$u_r = \frac{S_r}{\sqrt{3}} = 0.294 \mu\text{m} \tag{28}$$

Carry out three groups of repeatability measurements independently by three surveyors with professional measurement knowledge and operation level according to their habits and in accordance with the measurement requirements. Try to have the sampling strategy, measurement starting point and other factors in the experiment kept as different as possible. The probe configurations that can be selected by the surveyors include 2BY20, 2BY40, 3BY20, 3BY40, 4BY20, and 4BY40. The number and distribution of sampling points conform to BS 7172 [12]. See Table 3 for the data in the reproducibility experiment:

**Table 3.** The experimental data of diameter measurement for reproducibility detection.

The three times measurement mean of surveyor A $\bar{d}_j$ (mm)	Group I 62.0012	Group II 61.9998	Group III 62.0013
The three times measurement mean of surveyor B $\bar{d}_j$ (mm)	Group IV 61.9998	Group V 62.0016	Group VI 62.0001
The three times measurement mean of surveyor C $\bar{d}_j$ (mm)	Group VII 62.0013	Group VIII 61.9996	Group IX 62.0016

Note:  $\bar{d}_j$  indicates the measurement mean of the reproducibility experiment in Group  $j$ .

Total mean of the measurement for calculating reproducibility:

$$\bar{d} = \frac{\sum_{j=1}^9 \bar{d}_j}{9} = 62.0007 \text{ mm} \tag{29}$$

According to Formula (7), the standard uncertainty caused by the reproducibility of the diameter measurement is as follows:

$$u_R = \sqrt{\frac{1}{(9-1)} \sum_{j=1}^9 (\bar{d}_j - \bar{d})^2} = 0.850 \mu\text{m} \tag{30}$$

Then see Table 4 for the uncertainty component of the diameter measurement:

**Table 4.** The uncertainty budget for the diameter measurement.

Standard Uncertainty	Source of Uncertainty	Evaluation Result
$u_E$	Indication error	1.875 $\mu\text{m}$
$u_r$	Repeatability	0.294 $\mu\text{m}$
$u_R$	Reproducibility	0.850 $\mu\text{m}$

According to Formula (3), the combined standard uncertainty of the diameter measurement task can be calculated as follows:

$$u_c = \sqrt{u_E^2 + u_r^2 + u_R^2} = 2.1 \mu\text{m} \tag{31}$$

According to GUM, take  $p = 95\%$   $k = 2$  and the expanded uncertainty is as follows:

$$U = k \times u_c = 4.2 \mu\text{m} \tag{32}$$

5.2. Example for Evaluating Uncertainty of Flatness Measurement

According to Formula (9), the uncertainty component caused by indication errors can be calculated as follows in flatness measurement:

$$u_E = \frac{MPE_P}{\sqrt{3}} = \frac{3.5}{\sqrt{3}} = 2.021 \mu\text{m} \tag{33}$$

Similarly, under the condition of repeatability, complete 10 times' continuous and rapid measurement of the plane to be measured, calculate the standard deviation of the repeatability experiment as  $S_r = 0.618 \mu\text{m}$ ; if the average value of three times' measurements is taken as the best estimate of the flatness measurement, then the standard uncertainty caused by measurement repeatability is as follows:

$$u_r = \frac{S_r}{\sqrt{3}} = 0.357 \mu\text{m} \tag{34}$$

Similarly, complete the repeatability measurements by three surveyors and try to have the sampling strategy and other factors kept as different as possible, thus obtaining the standard uncertainty caused by the reproducibility of flatness measurement as  $u_R = 0.915 \mu\text{m}$ .

Then see Table 5 for the uncertainty component for flatness measurement:

Table 5. The uncertainty budget for flatness measurement.

Standard Uncertainty	Source of Uncertainty	Evaluation Result
$u_E$	Indication error	2.021 $\mu\text{m}$
$u_r$	Repeatability	0.357 $\mu\text{m}$
$u_R$	Reproducibility	0.915 $\mu\text{m}$

The combined standard uncertainty is as follows for flatness measurement tasks:

$$u_c = \sqrt{u_E^2 + u_r^2 + u_R^2} = 2.3 \mu\text{m} \tag{35}$$

Similarly, the expanded uncertainty is as follows:

$$U = k \times u_c = 4.6 \mu\text{m} \tag{36}$$

5.3. Example for Evaluating Uncertainty of Perpendicularity Measurement

According to Formula (15), the uncertainty component caused by indication errors can be calculated as follows in perpendicularity measurement:

$$u_E = \frac{2 \cdot A}{\sqrt{6}} = \frac{2 \times 3}{\sqrt{6}} = 2.449 \mu\text{m} \tag{37}$$

Similarly, under the condition of repeatability, complete 10 times' continuous and rapid measurement of the measurement task, calculate the standard deviation of the repeatability experiment as  $S_r = 0.682 \mu\text{m}$ ; if the average value of three times' measurements is taken as the best estimate of the perpendicularity measurement, then the standard uncertainty caused by measurement repeatability is as follows:

$$u_r = \frac{S_r}{\sqrt{3}} = 0.394 \mu\text{m} \tag{38}$$

Similarly, complete the reproducibility measurements by three surveyors and try to have the sampling strategy and other factors kept as different as possible, thus obtaining the standard uncertainty caused by the reproducibility of perpendicularity measurement as  $u_R = 1.060 \mu\text{m}$ . See Table 6 for the uncertainty component for perpendicularity measurement:

**Table 6.** The uncertainty budget for perpendicularity measurement.

Standard Uncertainty	Source of Uncertainty	Evaluation Result
$u_E$	Indication error	2.449 $\mu\text{m}$
$u_r$	Repeatability	0.394 $\mu\text{m}$
$u_R$	Reproducibility	1.060 $\mu\text{m}$

The combined standard uncertainty is as follows for perpendicularity measurement tasks:

$$u_c = \sqrt{u_E^2 + u_r^2 + u_R^2} = 2.7 \mu\text{m} \tag{39}$$

Similarly, the expanded uncertainty is as follows:

$$U = k \cdot u_c = 5.4 \mu\text{m} \tag{40}$$

5.4. Result Analysis and Optimizing Uncertainty

After the measurement task is determined, it is necessary to select the measuring instruments and methods according to the accuracy requirements. During measurement, see Table 7 [35] for the requirements of the geometric parameters for measurement accuracy, which is particularly important for geometrical tolerances detection; the tolerance level is Level 5–6 for the measurement object as shown in Figure 3, so the measurement uncertainty should be less than 20% of the tolerance value as required and then the uncertainty should be less than 2.0  $\mu\text{m}$  and 3.0  $\mu\text{m}$  respectively in flatness and perpendicularity detections. Obviously, the measurement uncertainty given in Formula (31)–Formula (40) has the evaluation results not in conformity with the accuracy requirements of the measurement tasks. There are two major reasons for such inconformity as follows: Firstly, the production-oriented CMM is not the preferred measuring instrument for geometrical tolerances. Its ability for measuring geometrical tolerance is slightly inferior to that of the roundness meter, autocollimator and other special instruments; secondly, the CMM task-oriented uncertainty evaluation model described in Section 2.2 features “overestimation” to ensure the safety and reliability of the evaluation results. When CMM is selected as the measurement instrument to complete the measurement tasks, it is bound to optimize the conventional measurement scheme and its uncertainty evaluation model. Therefore, here follows the research on the method for optimizing and evaluating the uncertainty of CMM task-oriented measurement in this paper.

**Table 7.** Accuracy requirements of geometric sense measurement.

Tolerance Grade	0	1	2	3	4	5	6	7	8	9	10	11	12
<b>Proportion of Uncertainty to Tolerance</b>		33%		25%		20%		16%		12.5%		10%	

Take the flatness measurement in Section 3.4.2 as an example to illustrate how to optimize and evaluate the CMM task-oriented measurement uncertainty and to determine the target measurement uncertainty  $U_o = 2.0 \mu\text{m}$  of the measurement task according to the requirements in Table 7. Therefore, only if the uncertainty of flatness measurement  $U$  after optimal evaluation should be smaller than  $U_o$  can the accuracy requirements of the measurement be satisfied.

Use  $P = 1.2 \mu\text{m}$  calibrated in Table 1 to substitute as the upper limit of the indication error and then the uncertainty component caused by the indication error in flatness measurement is as follows:

$$u'_E = \frac{P}{\sqrt{3}} = \frac{1.2}{\sqrt{3}} = 0.69 \mu\text{m} \tag{41}$$

In the reproducibility experiment of the flatness measurement in Section 3.4.2, try to have the number and distribution of the sampling points on the plane to be measured and other factors kept somewhat different. See Table 8 for the data of reproducibility measurement.

$$u_R = \sqrt{\frac{1}{(9-1)} \sum_{j=1}^9 (\bar{t}_j - \bar{\bar{t}})^2} = 0.915 \mu\text{m} \tag{42}$$

**Table 8.** The first reproducibility measurement data of flatness.

The three times measurement mean of surveyor A $\bar{t}_j$ (mm)	Group I 0.0053	Group II 0.0062	Group III 0.0045
The three times measurement mean of surveyor B $\bar{t}_j$ (mm)	Group IV 0.0041	Group V 0.0058	Group VI 0.0065
The three times measurement mean of surveyor C $\bar{t}_j$ (mm)	Group VII 0.0047	Group VIII 0.0066	Group IX 0.0049

In time for secondary optimal evaluation, distribute the sampling points evenly, use 12 sampling points in all and adopt 4mm probes and 20 mm measuring rods uniformly; see Table 9 for the reproducibility experimental data after the experimental scheme has been optimized.

$$u'_R = \sqrt{\frac{1}{(9-1)} \sum_{j=1}^9 (\bar{t}_j - \bar{\bar{t}})^2} = 0.518 \mu\text{m} \tag{43}$$

**Table 9.** The optimized reproducibility measurement data of flatness.

The three times measurement mean of surveyor A $\bar{t}_j$ (mm)	Group I 0.0051	Group II 0.0058	Group III 0.0048
The three times measurement mean of surveyor B $\bar{t}_j$ (mm)	Group IV 0.0047	Group V 0.0058	Group VI 0.0055
The three times measurement mean of surveyor C $\bar{t}_j$ (mm)	Group VII 0.0049	Group VIII 0.0061	Group IX 0.0049

It can be seen that the standard uncertainty introduced by reproducibility has reduced from 0.915  $\mu\text{m}$  to  $u'_R = 0.518 \mu\text{m}$  by optimizing the experimental scheme.

Considering that the repeatability indexes have been updating by the latest measurement information continuously arising during the process of the measurement process, regard the repeatability uncertainty component  $u_r = 0.357 \mu\text{m}$  of the flatness measurement in Section 3.4.2 as prior information  $u_0$ , and the number of measurements contained in the prior information  $n_0 = 10$ .

Table 10 shows the daily measurement data of the flatness for three workpieces in the same batch. Take workpiece A as the sample information and the repeatability uncertainty component can be calculated as follows after information fusion:

$$u_{r1} = \sqrt{\frac{(n_0 - 1)u_{r0}^2 + \sum_{j=1}^{n_1} (t_{1j} - \bar{t}_1)^2}{n_0 + n_1 - 3}} = 0.339 \mu\text{m} \tag{44}$$

**Table 10.** The daily measurement data of flatness.

Measured Value	Workpiece A/ $t_{1j}$	Workpiece B/ $t_{2j}$	Workpiece C/ $t_{3j}$
$t_{i1}$	0.0057 mm	0.0051 mm	0.0060 mm
$t_{i2}$	0.0048 mm	0.0044 mm	0.0049 mm
$t_{i3}$	0.0051 mm	0.0054 mm	0.0053 mm
Mean value $\bar{t}_i$	0.0052 mm	0.0050 mm	0.0054 mm
$n_i$	3	3	3
Standard deviation	0.458 $\mu\text{m}$	0.513 $\mu\text{m}$	0.556 $\mu\text{m}$
$u_r$	0.265 $\mu\text{m}$	0.296 $\mu\text{m}$	0.321 $\mu\text{m}$

That is to say, after incorporating the daily measurement information, the repeatability index has been updated from 0.357  $\mu\text{m}$  to 0.339  $\mu\text{m}$ ; replace the prior information with the result of information fusion to prepare for the next update; then:

$$u_{r0}^* = u_{r1} = 0.339 \mu\text{m} \tag{45}$$

$$n_0^* = n_0 + n_1 - 2 = 11 \tag{46}$$

Based on the updated prior information, further integrate the measurement results of part B, then the latest production and measurement information can be continuously integrated to achieve real-time and continuous updates of the repeatability index. By fusing the daily measurement data of flatness, the uncertainty components introduced by the repeatability of the flatness measurement of workpieces B and part C can be as follows:

$$u_{r2} = \sqrt{\frac{(n_0^* - 1)(u_{r0}^*)^2 + \sum_{j=1}^{n_1} (t_{2j} - \bar{t}_2)^2}{n_0^* + n_2 - 3}} = 0.323 \mu\text{m} \tag{47}$$

$$u_{r3} = \sqrt{\frac{(n_0^{**} - 1)(u_{r0}^{**})^2 + \sum_{j=1}^{n_1} (t_{3j} - \bar{t}_3)^2}{n_0^{**} + n_3 - 3}} = 0.309 \mu\text{m} \tag{48}$$

If the repeatability index has not been updated, the uncertainty component caused by repeatability will always keep the evaluation results of the previous repeatability experiments. If only daily measurement data is adopted to evaluate the repeatability, then the sample data will be small in number and less representative.

See Table 11 for the results from comparing the repeatability uncertainty component estimated in Section 3.4.2, the repeatability uncertainty component of the workpiece daily measurements and the repeatability uncertainty component updated in real time using the Bayesian formula.

**Table 11.** The comparison of the repeatability evaluation results by different methods.

Repeatability	Workpiece A	Workpiece B	Workpiece C
Repeatability of prediction	0.357 $\mu\text{m}$	0.357 $\mu\text{m}$	0.357 $\mu\text{m}$
Repeatability of sample data	0.265 $\mu\text{m}$	0.296 $\mu\text{m}$	0.321 $\mu\text{m}$
Repeatability of real-time updates	0.339 $\mu\text{m}$	0.323 $\mu\text{m}$	0.309 $\mu\text{m}$

As can be seen from Table 11, the predicted repeatability based on 10 applied measurements, once determined, will remain static, so it can not reflect the latest information in daily measurements in real time. However, when the sample data of daily measurements is used to evaluate the uncertainty component caused by repeatability, the small sample size can only reflect limited information and is easily affected by accidental factors in the experiment. Using the Bayesian method to fully fuse the

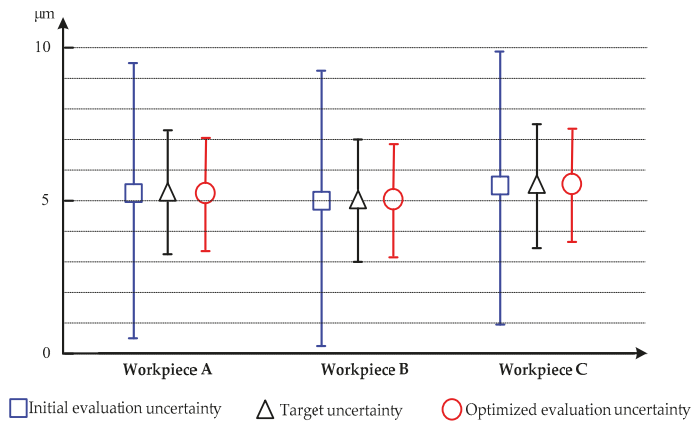


historical information and the current information for updating the uncertainty component introduced by repeatability in real time can integrate the latest measurement information into the evaluation results to reflect the latest trend of the random effects in the measurement system in time; moreover, the amount of information is comparatively great in the samples used for fully integrating the current and historical information, so the evaluation results will not easily be influenced by accidental factors and the repeatable uncertainty components tend to be stable and reliable.

See Table 12 and Figure 5 for estimations of the uncertainty optimal evaluation for the flatness measurement of the workpieces A, B and C in the same batch:

**Table 12.** The uncertainty budget of optimized evaluation for flatness measurement.

Uncertainty Components		Initial Evaluation Results	Optimized Evaluation Results		
Symbol	Sources		Workpiece A	Workpiece B	Workpiece C
$u_E$	Indication error	2.021 $\mu\text{m}$	0.69 $\mu\text{m}$	0.69 $\mu\text{m}$	0.69 $\mu\text{m}$
$u_r$	Repeatability	0.357 $\mu\text{m}$	0.339 $\mu\text{m}$	0.323 $\mu\text{m}$	0.309 $\mu\text{m}$
$u_R$	Reproducibility	0.915 $\mu\text{m}$	0.518 $\mu\text{m}$	0.518 $\mu\text{m}$	0.518 $\mu\text{m}$
Standard uncertainty $u_c$		2.3 $\mu\text{m}$	0.9 $\mu\text{m}$	0.9 $\mu\text{m}$	0.9 $\mu\text{m}$
Expanded uncertainty U (p = 95%)		4.6 $\mu\text{m}$	1.8 $\mu\text{m}$	1.8 $\mu\text{m}$	1.8 $\mu\text{m}$
Comparison with target uncertainty $U_o = 2.0 \mu\text{m}$		Excess	Less	Less	Less



**Figure 5.** Comparison of measurement uncertainty before and after optimization.

As can be seen from Figure 5, when the extended uncertainty of the first evaluation is regarded as the uncertainty evaluation result of the flatness measurement task, the requirements of the target uncertainty  $U_o$  cannot be satisfied; however, the extended uncertainty of the flatness measurement tasks for three workpieces to be measured after optimal evaluation can meet the accuracy requirements of the measurement tasks, but the sacrifice will be relatively high accordingly. Therefore, according to the basic principle of the task-oriented uncertainty optimal evaluation described in this section, the surveyors should make choices according to their own measurement conditions.

### 6. Conclusions

In the framework of the GPS standard system, the CMM task-oriented uncertainty evaluation has been studied. The main content is as follows:

(1) The difficulty of uncertainty evaluation for CMM measurement tasks oriented at dimensional and geometric errors has been solved systematically. Based on the systematic analysis of the CMM task-oriented uncertainty sources, a model for evaluating the CMM task-oriented measurement uncertainty based on the measurement system analysis has been proposed. Starting from the statistical

characteristic indexes of the measurement results, this model has somewhat comprehensively reflected the impact of the CMM measuring system uncertainty sources on measurement results.

(2) The quantization method and evaluation model for uncertainty components of such three different measurement tasks of the CMM as dimensions, geometrical errors have been studied. The uncertainty components caused by the indication errors of CMM different measurement tasks are quite different, and especially the form error and some position errors reflect the relative changes in the micro-sized space. The linear influence should be neglected when the uncertainty components are quantified. Moreover, the indication bias introduced by the residual system errors of the measuring instrument should also be quantified and characterized by reasonably selecting the error indicators according to the specific measurement tasks.

(3) The optimal evaluation of the CMM task-oriented uncertainty has been studied. The target uncertainty of geometric parameter detection must satisfy the requirements of the design tolerance. In order to ensure the safety and reliability of the evaluation results, there is an "excessive estimation" in the CMM task-oriented uncertainty evaluation, which has the evaluation results of uncertainty possibly not meet the requirements for measurement accuracy. At this time, the secondary optimal evaluation of the measurement uncertainty should be extremely important. The uncertainty source information of the measurement process should be further grasped by means of experiments and other prior information, as well as each uncertainty component should be re-quantified reasonably; thus the combined standard uncertainty would be reduced.

**Author Contributions:** Conceived the Method and Wrote the Paper, Y.C.; Data curation, H.L. and Y.L.; Performed some Confirmatory Experiments, Y.C., H.L. and H.W.; Resources, Z.W. and X.C.; Edited the Manuscript, Z.W.

**Funding:** This research was supported by National Key Research and Development Plan of China, Grant Number 2016YFF0203801; Fujian Province Public Scientific Research Institute Special Program of China, Grant Number 2018R1033-6; Natural Science Foundation of China, Grant Number 51575032.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Li, H.L.; Chen, X.H.; Cheng, Y.B.; Liu, H.D.; Wang, H.B.; Cheng, Z.Y.; Wang, H.T. Uncertainty modeling and evaluation of CMM task oriented measurement based on SVCMM. *Meas. Sci. Rev.* **2017**, *17*, 226–231. [[CrossRef](#)]
2. Cappetti, N.; Naddeo, A.; Vilecco, F. Fuzzy approach to measures correction on coordinate measuring machines: The case of hole-diameter verification. *Measurement* **2016**, *93*, 41–47. [[CrossRef](#)]
3. D'Amato, R.; Caja, J.; Maresca, P.; Gómez, E. Use of coordinate measuring machine to measure angles by geometric characterization of perpendicular planes. Estimating uncertainty. *Measurement* **2014**, *47*, 598–606. [[CrossRef](#)]
4. Vrba, I.; Palencar, R.; Hadzistevic, M.; Strbac, B.; Jokic, V.S.; Hodolic, J. Different approaches in uncertainty evaluation for measurement of complex surfaces using coordinate measuring machine. *Meas. Sci. Rev.* **2015**, *15*, 111–118. [[CrossRef](#)]
5. Li, R.J.; Fan, K.C.; Huang, Q.X.; Zhou, H.; Gong, E.M.; Xiang, M. A long-stroke 3D contact scanning probe for micro/nano coordinate measuring machine. *Precis. Eng.* **2016**, *43*, 220–229. [[CrossRef](#)]
6. Thalmann, R.; Meli, F.; Küng, A. State of the art of tactile micro coordinate metrology. *Appl. Sci.* **2016**, *6*, 150. [[CrossRef](#)]
7. Cheng, Y.B.; Chen, X.H.; Li, H.L.; Cheng, Z.Y.; Jiang, R.; Lü, J.; Fu, H.D. Analysis and comparison of bayesian methods for measurement uncertainty evaluation. *Math. Probl. Eng.* **2018**, *1*, 7509046. [[CrossRef](#)]
8. Bich, W. Revision of the 'Guide to the Expression of Uncertainty in Measurement'. Why and how. *Metrologia* **2014**, *51*, S155–S158. [[CrossRef](#)]
9. Gaska, P.; Gaska, A.; Gruga, M. Challenges for modeling of five-axis coordinate measuring systems. *Appl. Sci.* **2017**, *7*, 803. [[CrossRef](#)]
10. Schwenke, H.; Siebert, R.L.; Waldele, F.; Kunzmann, H. Assessment of Uncertainties in Dimensional Metrology by Monte Carlo Simulation: Proposal of a Modular and Visual Software. *CIRP Ann. Manuf. Technol.* **2000**, *49*, 395–398. [[CrossRef](#)]

11. Weckenmann, A.; Knauer, M.; Kunzmann, H. The Influence of Measurement Strategy on the Uncertainty of CMM-Measurements. *CIRP Ann. Manuf. Technol.* **1998**, *47*, 451–454. [[CrossRef](#)]
12. Advanced Manufacturing Technology Standards Policy Committee of Britain. *BS 7172: 1989, Guide to Assessment of Position, Size and Departure from Nominal Form of Geometric Features (Confirmed January 2010)*; AMT: London, UK, 2010.
13. Balsamo, A.; Di Ciommo, M.; Mugno, R.; Rebaglia, B.I.; Ricci, E.; Grellaz, R. Evaluation of CMM Uncertainty through Monte Carlo Simulations. *CIRP Ann. Manuf. Technol.* **1999**, *48*, 425–428. [[CrossRef](#)]
14. Sladek, J.; Gaška, A. Evaluation of coordinate measurement uncertainty with use of virtual machine model based on monte carlo method. *Measurement* **2012**, *45*, 1564–1575. [[CrossRef](#)]
15. Jakubiec, W. Estimation of uncertainty of coordinate measurements according to the Type B method. *Key Eng. Mater.* **2010**, *437*, 253–257. [[CrossRef](#)]
16. Jakubiec, W.; Plowucha, W. First Coordinate Measurements Uncertainty Evaluation Software Fully Consistent with the GPS Philosophy. *Procedia CIRP* **2013**, *10*, 317–322. [[CrossRef](#)]
17. Dhanish, P.B.; Mathew, J. Effect of CMM point coordinate uncertainty on uncertainties in determination of circular features. *Measurement* **2006**, *39*, 522–531. [[CrossRef](#)]
18. Jonathan, B.; Edward, M. Experimental evaluation of software estimates of task specific measurement uncertainty for CMMs. *Precis. Eng.* **2010**, *34*, 28–33. [[CrossRef](#)]
19. Feng, C.X.J.; Saal, A.L.; Salsbury, J.G.; Ness, A.R.; Lin, G.C.S. Design and analysis of experiments in CMM measurement uncertainty study. *Precis. Eng.* **2007**, *31*, 94–101. [[CrossRef](#)]
20. Kruth, J.P.; Gestel, N.V.; Bleys, P.; Welkenhuyzen, F. Uncertainty determination for CMMs by monte carlo simulation integrating feature form deviations. *CIRP Ann. Manuf. Technol.* **2009**, *58*, 463–466. [[CrossRef](#)]
21. Aggogeri, F.; Barbato, G.; Barini, E.M.; Genta, G.; Levi, R. Measurement uncertainty assessment of Coordinate Measuring Machines by simulation and planned experimentation. *CIRP J. Manuf. Sci. Technol.* **2011**, *4*, 51–56. [[CrossRef](#)]
22. Zhang, G.X. *Coordinate Measuring Machines*; Tianjin University Press: Tianjin, China, 1999.
23. Wilhelm, R.G.; Hocken, R.; Schwenke, H. Task specific uncertainty in coordinate measurement. *CIRP Ann. Manuf. Technol.* **2001**, *50*, 553–563. [[CrossRef](#)]
24. Haitjema, H. Task specific uncertainty estimation in dimensional metrology. *Int. J. Precis. Technol.* **2011**, *2*, 226–245. [[CrossRef](#)]
25. Takamasu, K.; Takahashi, S.; Abbe, M.; Furutani, R. Uncertainty estimation for coordinate metrology with effects of calibration and form deviation in strategy of measurement. *Meas. Sci. Technol.* **2008**, *19*, 084001. [[CrossRef](#)]
26. International Organization for Standardization. *ISO/TS 15530-1: 2013—Geometrical Product Specifications (GPS)—Coordinate Measuring Machines (CMM): Technique for Determining the Uncertainty of Measurement—Part 1: Overview and Metrological Characteristics*; ISO: Geneva, Switzerland, 2013.
27. International Organization for Standardization. *ISO 15530-3: 2011—Geometrical Product Specifications (GPS)—Coordinate Measuring Machines (CMM): Technique for Determining the Uncertainty of Measurement—Part 3: Use of Calibrated Workpieces or Standards*; ISO: Geneva, Switzerland, 2011.
28. International Organization for Standardization. *ISO 10360-1: 2000—Geometrical Product Specifications (GPS)—Acceptance and Reverification Tests for Coordinate Measuring Machines (CMM)—Part 1: Vocabulary*; ISO: Geneva, Switzerland, 2000.
29. International Organization for Standardization. *ISO 10360-2: 2009—Geometrical Product Specifications (GPS)—Acceptance and Reverification Tests for Coordinate Measuring Machines (CMM)—Part 2: CMMs Used for Measuring Linear Dimensions*; ISO: Geneva, Switzerland, 2009.
30. Li, R.J.; Xiang, M.; He, Y.X.; Fan, K.C.; Cheng, Z.Y.; Huang, Q.X.; Zhou, B. Development of a high-Precision touch-trigger probe using a single sensor. *Appl. Sci.* **2016**, *6*, 86. [[CrossRef](#)]
31. International Organization for Standardization. *ISO 17450-2: 2012—Geometrical Product Specifications (GPS)—General Concepts—Part 2: Basic Tenets, Specifications, Operators, Uncertainties and Ambiguities*; ISO: Geneva, Switzerland, 2012.
32. The Joint Committee for Guides in Metrology. *JCGM 100: 2008, Guide to the Expression of Uncertainty in Measurement, GUM 1995 with Minor Corrections*; BIPM: Paris, France, 2008.


33. International Organization for Standardization. *ISO 1101: 2017, Geometrical Product Specifications (GPS)—Geometrical Tolerancing—Tolerances of Form, Orientation, Location and Run-Out*; ISO: Geneva, Switzerland, 2017.
34. International Organization for Standardization. *ISO 14253-2: 2011—Geometrical Product Specifications (GPS)—Inspection by Measurement of Workpieces and Measuring Equipment—Part 2: Guidance for the Estimation of Uncertainty in GPS Measurement, in Calibration of Measuring Equipment and in Product Verification*; ISO: Geneva, Switzerland, 2011.
35. Standardization Administration of China. *GB/T 1958-2017, Geometrical Product Specifications (GPS)—Geometrical Tolerance—Verification*; SAC/TC: Beijing, China, 2017.



© 2018 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).

Article

# A Novel Design of Through-Hole Depth On-Machine Optical Measuring Equipment for Automatic Drilling and Riveting

Nianhan Wu <sup>1</sup>, Wu Zhao <sup>1</sup>, Xin Wang <sup>2</sup>, Ye Tao <sup>1,3,\*</sup> and Zhengmeng Hou <sup>3</sup> 

<sup>1</sup> School of Manufacturing Science and Engineering, Sichuan University, Chengdu 610065, China; nianhanwu@163.com (N.W.); zhaowu@scu.edu.cn (W.Z.)

<sup>2</sup> Aerospace Research Institute of Materials and Processing Technology, China Academy of Launch Vehicle Technology, Beijing 100076, China; shiningwoon@126.com

<sup>3</sup> Energy Research Center of Lower Saxony (EFZN), 38640 Goslar, Germany; hou@tu-clausthal.de

\* Correspondence: yetao@scu.edu.cn; Tel.: +49-0152-5155-4369

Received: 25 November 2018; Accepted: 13 December 2018; Published: 18 December 2018



**Featured Application:** On-machine optical measuring process of through-hole depth for automatic drilling and riveting system.

**Abstract:** In the aerospace manufacturing industry, it is impossible to achieve precise and efficient automatic drilling and riveting for largescale composite board parts. The bottleneck is that the depth detection of rivet holes still relies on manual operation, which seriously affects the assembly efficiency and stability of composite board parts. In order to realize accurate and efficient on-machine automatic measurement for through holes in the automatic drilling and riveting process of largescale composite board parts, this paper presents a novel hole depth measuring device. Its mechanical structure is developed based on our newly designed measurement scheme and optical path, the purpose of which is to convert the hole depth data into displacement data of the probe motion. Its electrical hardware consists of three units: a laser transceiver unit to pick up laser spots; a displacement measuring unit to capture the probe movement in real time; and a driving unit to achieve motion control of the probe. Finally, the experimental results indicated that the proposed method and device are capable of performing automatic measurements for through-hole depth. In addition, factors affecting the measuring accuracy and stability of the device are initially analyzed and discussed, which lay a foundation for subsequent research on error compensation and probe calibration.

**Keywords:** on-machine measurement; through-hole depth; image processing; automatic drilling and riveting; large-scale composite board; depth detection

## 1. Introduction

The aerospace manufacturing industry belongs to the field of high-end equipment manufacturing and requires a large number of high-tech technologies as technical support [1]. Among them, the precise and efficient assembly of largescale composite board parts is a representative technical difficulty in the aircraft assembly process [2]. Such parts typically exhibit complex curved surface structures, and their assembly requires extensive drilling and riveting operations [3]. It is worth noting that the depth of the rivet hole is different from the original design. This is because preparation processes such as bending, bonding, and solidification of the composite board parts cannot be controlled to perfection, resulting in mismatch in shape and thickness [4]. Therefore, it is especially necessary to adaptively adjust the length of the rivet, according to the actual size of the hole [5]. In other words, the automatic drilling

and riveting manipulator first needs to measure the current hole depth, and then select a rivet with appropriate length for riveting operations.

However, there is currently no equipment that can automatically measure the depth of through holes. At present, the riveting of large composite board parts still relies on manual operations, which include hole depth measurement, rivet selection, and riveting [6]. It is worth noting that largescale aerospace composite parts often have thousands of holes that need to be riveted [7]. Manual riveting consumes a lot of manpower, is inefficient, and is prone to human error [8]. Therefore, there is the urgent need for a method or apparatus that will enable on-machine automatic measurement of through-hole depth, offering vital technical support to automatic drilling and riveting of largescale composite board parts.

In-depth research on hole parameter measurement can be roughly divided into the following categories:

(i) Detection and recognition of hole position and hole shape. Shetty et al. [9] proposed a hole detection method based on machine vision to test the cooling holes of fan blades in aero engines. They developed a testing system based on machine vision and laser detection that guarantees the quality of porous components. The system recognizes the appearance of measured parts to detect pore size and distance between adjacent holes on parts. It can recognize the bottom of holes using a laser detection system. The purpose of the system is to detect the distance between holes and the diameter of holes, but not the depth of the holes. Baeg et al. [10] studied the identification and positioning of threaded holes. In order to realize automatic assembly in the field of automotive manufacturing, they designed a set of automatic identification and positioning system for thread holes. This method adopts machine vision, auxiliary light-emitting diode (LED), and auxiliary positioning laser generator to identify the threaded hole and get its three dimensions (3D) coordinates. The purpose is to identify the location of the hole without measuring its specific parameters. Usamentiaga et al. [11] specialized in the study of non-destructive testing of holes in honeycomb sandwich panels. The whole panel was thermally imaged and image analyzed in order to obtain the position and shape information of the hole on the honeycomb sandwich panel. The hole extracted in this study is mainly a position parameter, and the integrity of the hole, but does not involve the extraction of hole depth parameters.

(ii) Hole quality inspection. Gong et al. [12] focused on the field of deep hole measurement. Aimed at holes with a large depth-to-diameter ratio, they proposed an internal structure detection method based on micro vision. Charge-coupled device (CCD) sensors and endoscopes are used to capture images of the hole structure and components, extract the curve from the image, and then detect the quality of the hole structure. This research realizes quality inspection of the inner connector of the micro hole, and provides a method for its detection. However, its main purpose is quality inspection. It does not pay attention to the measurement of the dimension parameters of the hole itself, and its measurement of displacement or distance is inadequate. Bernard, Flaherty and O'Connor [13] conducted studies on the field of quality measurement of holes. The object of this study was a pipe with small boreholes. The boreholes are usually through-holes, but excessive drilling may occasionally lead to damage. Therefore, a detection method was designed, in which a luminescent fiber is inserted into the tube and CCD is used to collect light signals outside the tube, in order to detect the permeability of the hole. The research object is a hose part, which does not involve mechanical parts of other materials such as metal materials, and its characteristics acquired are hole permeability and the integrity of parts. Zavyalov [14] proposed a high curvature optical imaging lens to obtain 3D shapes of pores. The proposed method can generate images and detect the quality of holes quickly. However, limited by the measuring principle and lens curvature, this system makes measuring holes with small aperture and long depth difficult, and it also cannot offer specific hole depth data.

(iii) Measurement of hole depth parameters. In order to achieve laser drilling of bones, Quest, Gayer and Hering [15] proposed a method to measure the hole depth of bone drilling by laser, which is applied to the field of oral implants. With laser triangulation and tomography, the actual position of

laser ablation in bone can be calculated. This study provides an indirect method to measure the hole depth parameters of micro holes using optical systems. The measurement object is the skeleton, and the whole process needs to be combined with fault images for comprehensive analysis. Therefore, it cannot be directly applied to the rapid measurement of large mechanical parts, and cannot achieve through-hole depth measurement. Takeda et al. [16] developed a method that can obtain hole depth by using the fringe projection approach. To measure 3-D objects with large height discontinuities, they proposed a coaxial optical sensor system with a common image plane for pattern projection and observation. A grating pattern is constantly illuminated on the measuring object and the whole field of view is imaged back onto the CCD sensor through a specific optical path. The object then moves along the Co-image plane until it passes through. If the measured object has a hole, the fringe amplitude reaches its peak when the bottom and upper surfaces of the hole pass the Co-image. According to this feature, hole depth can be calculated by measuring the displacement of the object between the two peaks. The fringe projection approach can be used to measure not only hole depths, but also shapes of objects [17]. This study provides the possibility of simultaneously measuring of multiple hole depths. Even more, the fringe projection approach can be applied to complex measurement environments by using fiber optics [18]. However, since this method takes some time to perform a large number of calculations to obtain fringe amplitude peaks, it will have a slightly lower measurement efficiency, especially during online measurements. And the premise of this measurement method is that the measured hole needs to have a bottom surface. When it comes to a through hole, the fringe amplitude will only have one peak, which may make measuring through holes a little difficult. Lin, Powell and Jiang [19] carried out research in the field of real-time depth measurement of laser drilling. An optical system based on confocal principle was developed for laser drilling, which is used for real-time measurement of micro hole depth during laser drilling. The idea is to obtain the depth of laser drilling indirectly by obtaining the relevant parameters in the laser beam path. Similarly, Ho, Chiu and Chang [20] studied the relationship between the hole depth parameters of laser drilling and the intensity of plasma luminescence emitted during drilling. The above two methods focus on hole depth measurement during laser drilling, which are not suitable for depth measurement in other processes such as mechanical drilling.

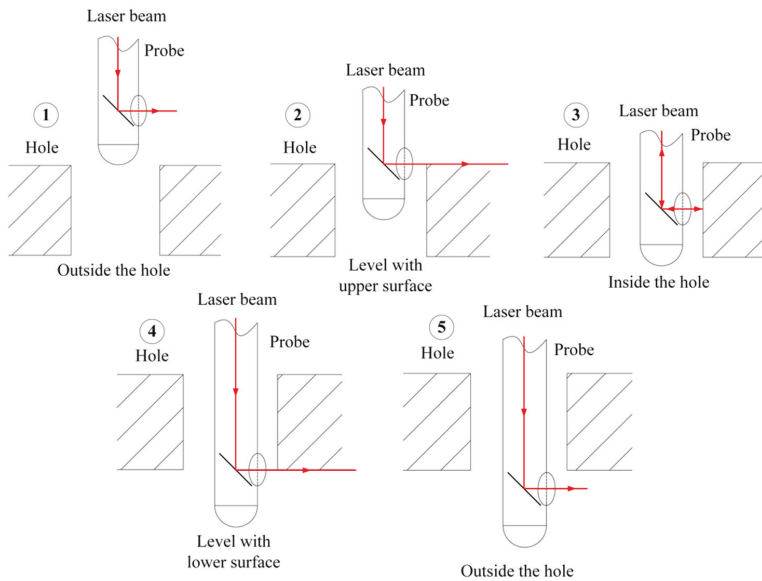
The above methods are mainly aimed at the spatial position identification of holes, quality inspection of hole shape, or hole depth measurement for laser drilling. However, existing research on the depth measurement of through holes is lacking. Existing methods can only measure the depth of blind holes using laser measurements or machine vision, and it is impossible to measure the depth of through holes, especially in the case where the depth/diameter ratio is large. At present, there are no reports on methods and instruments for on-machine measurement of through-hole depth in the field of automatic drilling and riveting. In view of the above problems, this paper carried out research on the on-machine measurement of through-hole depth. A new measurement method is proposed, a corresponding laser light path and novel measuring device with optical sensor and grid-capacitance sensor is designed, and its prototype is fabricated.

The remainder of this paper is organized as follows: In Section 2, the optical path design, device design, and the measuring method design of the hole depth on-machine optical measuring equipment are described in detail. Section 3 shows the experimental testing process and discusses data processing and analysis results. Finally, this paper draws conclusions in Section 4 and looks forward to follow-up research.

## **2. Approach Description**

### *2.1. Measurement Scheme*

In order to realize the measurement of through-hole depth, the measurement scheme is first designed and its overall workflow is introduced. As shown in Figure 1, the whole measurement process is divided into five steps:



**Figure 1.** Schematic diagram of the probe movement process.

The probe is in continuous movement during the measuring process. At the beginning of the measuring process, the device is first placed above the hole using mechanical equipment (e.g., mechanical arm, articulated arm-measuring machine); the axis of the device coincides with the axis of the hole. The direction of laser emission from the front of the probe is perpendicular to the axis of the probe. In Step 1, since the emitted laser beam is not projected onto the inner wall of the hole, the probe cannot sense the laser spot on the wall of the hole and continues to move downwards. In Step 2, the laser beam is even with the upper surface of the hole. As the probe continues to move downward, the laser beam is projected onto the wall of the hole and forms a laser spot. At this point, the sensor inside the probe detects the laser spot and immediately records the current position of the probe. In Step 3, the probe enters the hole, and continuously senses the laser spot signal and continues to feed downward [21]. In Step 4, the laser beam is even with the lower surface of the hole, and exits the hole. After that moment, the probe cannot sense the laser spot and immediately records its current position. The depth data of the through hole is obtained by the displacement difference of positions recorded in steps 2 and 4. In Step 5, the probe moves upward and returns to its initial position.

## 2.2. Optical Path Design

This section details how the laser signal of the device is projected from the laser transmitter onto the wall of the hole and ultimately reflected back to the CCD. Figure 2 shows the optical path designed for our measuring system. The laser emitted by the laser transmitter is at an angle of  $45^\circ$  to the plane of the beam splitter. Due to the working principle of the beam splitter, when the laser is directed at the beam splitter, half of the light is projected in the original direction, and the other half of the light is reflected [22]. Therefore, by using two parallel splitters, the laser beam emitted from the laser transmitter can be projected perpendicular to the wall of the measured hole. At this time, the laser beam reflected by the hole wall can also be projected onto the CCD through these two beam splitters. In order to converge the laser beam, which will gradually diverge in the projection process, a convex lens is placed at the laser exit end, as shown in Figure 2, and another one is placed in front of the CCD to focus the laser beam on it.



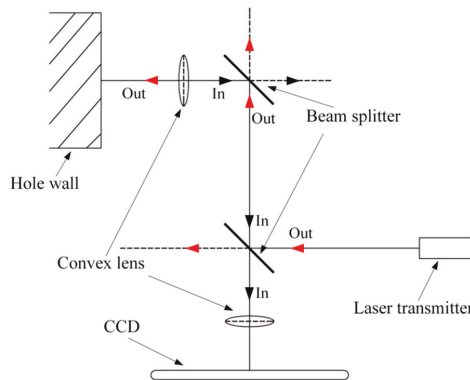


Figure 2. Optical path design of the measuring system.

A problem in the process is that the diameter of the probe is only 2 mm, which cannot accommodate a laser transmitter with diameter of 6 mm. Therefore, the ultimate goal of the optical path design is to project the laser beam emitted from the laser transmitter onto the hole wall. Beyond that, the light reflected back from the wall of the hole can be collected by the CCD through this optical path and processed in the next step.

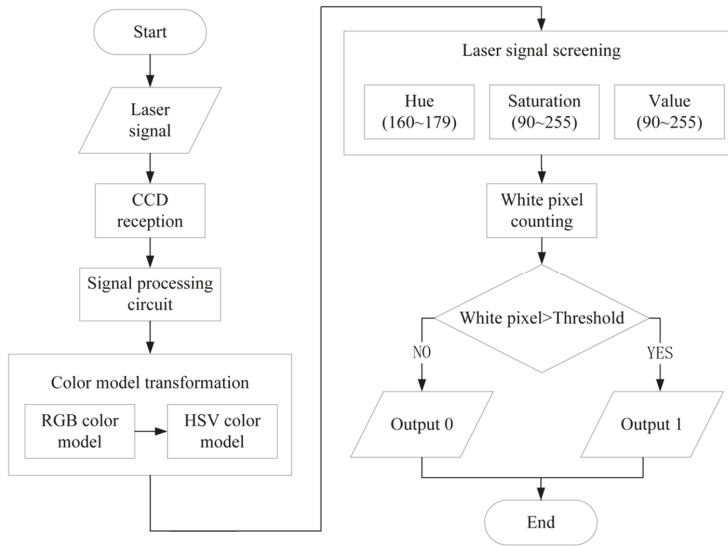
### 2.3. Image Processing

Based on the optical path designed above, this section will specifically describe how image processing is performed when the CCD receives the laser signal, and how to set the white pixel threshold for adjusting the actual measured value.

The device determines whether the probe enters the hole or protrudes out of the hole by acquiring the laser signal reflected back from the hole wall. After the laser signal is acquired by the CCD, it is imaged as a red spot by red-green-blue (RGB) colour model [23]. Since the CCD acquires optical signals for all external environments, there are laser signals and interference light signals from the external environment in the original image. Therefore, the ultimate goal of image processing is to identify the presence or absence of a red laser signal from the original image. Figure 3 shows the block diagram of imaging identification and screening process.

As shown in Figure 3, when the laser signal is projected onto the hole wall, the CCD receives the reflected laser signal and transmits the single frame image to the host computer for real-time processing [24]. Next, the image colour model needs to be transformed from RGB to hue-saturation-value (HSV). After that, we set the parameters of HSV model image to screen the laser wavelength produced by the laser transmitter. In order to facilitate calculation of the selected laser spot imaging, the image needs to be processed in two values. There are only white and black pixels in the binary image, and the white pixels represent the laser signal reflected by the hole wall. When white pixels are generated, this means that the laser signal is projected onto the wall of the hole to be measured. The host computer outputs a signal to the microcontroller unit (MCU) to show that the front end of the probe has entered the hole. In order to make the whole device have high fault tolerance and adaptability, it is necessary to set a threshold for the number of white pixels, which decides whether to send signals to MCU or not.

A digital image is composed of a number of pixels, and each pixel has its internal display principle, which includes parameters such as brightness and colour. Therefore, the colour composition and expression of the pixel form the colour model. The visible spectrum contains all the visible light subsets, each of which contains colour light for each wavelength of the visible spectral range to which this subset belongs. Different combinations of these visible light subsets form different image colours, and different colour models have different colour expressions.



**Figure 3.** The block diagram of imaging identification and screening process; CCD: Charge-coupled device.

The RGB colour model contains red, green, and blue colours, and the different ratios of the three colours blend other colours [25]. Since the colour expression method is the same as the colour formation principle of the actual display device, most of the display device and the imaging device adopt the RGB image colour model, and the image colour model directly extracted by the device designed in this paper is the RGB colour model. The HSV colour model is composed of Saturation, Hue, and Value. Compared with the RGB colour model, HSV colour model is more in line with the human eye’s recognition and computer processing of images. The HSV colour model explains the origin of colour and conforms to the physical principle [26]. Therefore, in image processing, the RGB colour model needs to first be transformed into an HSV colour model.

The RGB colour model is expressed using a three-dimensional rectangular coordinate system, and the three parameters of R, G and B correspond to X, Z and Y axes, respectively. The HSV colour model is expressed in a three-dimensional polar coordinate system. Hue, Saturation and Value correspond to Z axis, polar axis and rotation angle, respectively [27]. Therefore, the conversion between the two colour models is calculated as follows:

Hue calculation:

$$H = \begin{cases} \frac{G-B}{\max(R,G,B)-\min(R,G,B)} & \text{when } R = \max(R, G, B) \\ \frac{B-R}{\max(R,G,B)-\min(R,G,B)} & \text{when } G = \max(R, G, B) \\ \frac{R-G}{\max(R,G,B)-\min(R,G,B)} & \text{when } B = \max(R, G, B) \end{cases} . \quad (1)$$

Saturation calculation:

$$S = \begin{cases} 0 & \text{when } V = 0 \\ 1 - \frac{\min(R,G,B)}{\max(R,G,B)} & \text{when } V \neq 0 \end{cases} . \quad (2)$$

Value calculation:

$$V = \max(R, G, B). \quad (3)$$

Because the optical signal captured by CCD includes both laser signals and ambient light signals, it is necessary to screen the laser signal to distinguish it from the ambient light signal after the image is

transformed from the RGB color model to the HSV color model. The screening method adopted in this paper enables adjustment of the values of the three parameters of the HSV color model.

In the HSV colour model, the Hue adjustment range is 0 to 179, and as the value increases, the colour tends to turn red. The adjustment range of Saturation and Value are 0 to 255, and as the value of the parameter increases, Saturation decreases. The colour of the laser emitted by the laser transmitter in this device is bright red. Its Saturation is high because of its generation principle [28]. Therefore, the optimum range of the three parameters in image processing is Hue—160 to 179, Saturation—90 to 255, and Value—90 to 255.

After the image is screened, it is necessary to binarize the image to determine whether the laser signal is received or not. Binarization of the image means to turn a picture into a black and white picture [29]. In this image, the pixel has only two colours, black and white, and the value of the pixel is 0 or 255. When the value of the pixel is 0, it is a black pixel, and when the value of the pixel is 255, it is white. Thus, when no laser signal is reflected to the CCD, there will be no white pixel in the image. However, if a laser signal is received, a circular white spot appears in the image, which means that the laser signal is projected to the wall of the hole.

#### 2.4. Hardware Design

In this section, the hardware design of the proposed measuring device is described in detail, and the hardware interrelationship is explained. From a structural point of view, the device is divided into three parts: the outer shell, the inner shell, and the probe, as shown in Figure 4. The laser transmitter, the CCD sensor, the beam splitters, and the nut are in the inner shell. The MCU, the motor, the coupling, and the screw are in the outer shell. The fixed ruler and movable ruler of the capacitive grating displacement sensor are, respectively, installed in the inner shell and the outer shell. The laser transceiver unit comprises of the laser transmitter, the CCD and the beam splitter. The displacement measuring unit comprises of the capacitive grating displacement sensor, the MCU, the outer shell, and the inner shell. The driving unit comprises of an electric, a coupling, a screw, and a nut. The CCD is packaged with a convex lens and forms a miniature camera. The camera is fixed by the two sets of fixing frames above the beam splitter. The outer shell and the inner shell are nested structures, the motor is fixed to the end of the outer shell, and the output shaft of the motor is connected to the screw through the coupling. The nut is fixed on the inner casing and is threaded on the lead screw. The MCU for logic control is also disposed at the end of the outer shell.

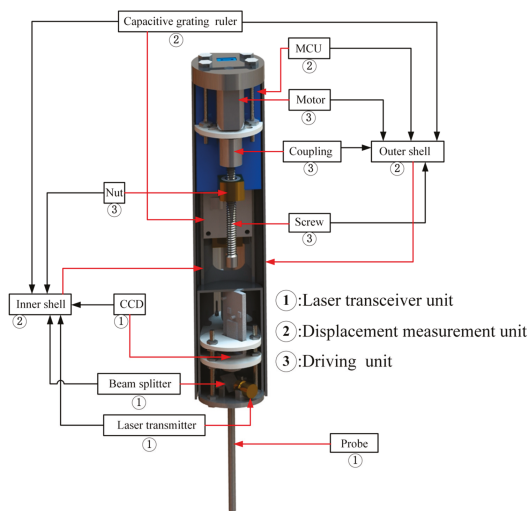


Figure 4. Internal structure of the measuring device; MCU: microcontroller unit.

From the perspective of electrical hardware, the measurement system is divided into three units: laser transceiver unit, displacement measuring unit and driving unit. In the laser transceiver unit, the laser is projected through the designed optical path to the hole wall and finally returned to the CCD. This unit is used for state determination during probe measurement. The drive unit is used to implement the telescopic movement of the probe, and during this process, the inner shell and the outer shell are relatively displaced. The function of the displacement measuring unit is to measure the relative displacement between the inner shell and the outer shell. The core component of this unit is the capacitive grating displacement sensor, the function of which is to measure the displacement of the probe. Finally, the hole depth parameter is obtained by intercepting the displacement recorded in the process of laser scanning the hole wall.

The structure design of the device is shown in Figure 5. The probe and the inner shell are hollow structures, and the diameter of the probe is only 3 mm, which enables the probe to penetrate into holes with smaller diameters. Even more, the probe can be further miniaturized by using smaller optical lenses. Thus, the device has the ability to measure the depth of holes with a large depth-to-diameter ratio. The designed optical system is placed in the inner shell and the probe. The laser transmitter cannot be placed in the small probe because of its large volume. Therefore, using the designed optical system, the laser emitted from the laser transmitter can be transferred from the inner shell of the device to the probe tip. The laser reflected from the hole wall is received by the CCD located in the inner shell. The outer shell and the inner shell of the device form a retractable structure through the screw and the nut. The inner shell moves forward or backward along the common axis of the outer shell and the inner shell is driven by the rotation of the motor. The capacitive grating scale (fixed ruler) and the capacitance grid ruler (movable ruler) are combined to become a capacitive grating displacement sensor [30], wherein, the movable ruler is arranged in the inner shell, and the fixed ruler is mounted on the outer shell. When the relative displacement between the inner shell and the outer shell occurs under the drive of the motor, the movable ruler slides on the fixed ruler and get the relative displacement data.

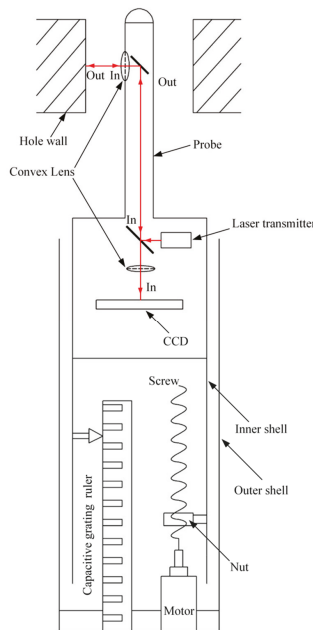


Figure 5. Schematic diagram of equipment structure design.

Figure 6 shows the hardware connection diagram of the measurement system. A separate direct current (DC) power supply is required for the laser transmitter and the motor. The MCU, the capacitive grating displacement sensor, and the CCD are directly connected to the host computer, and are electrically driven by the host computer. The laser emitted by the laser transmitter is reflected by the wall of the holes and is received by the CCD. After the optical signal received by the CCD is processed into image information, this image signal will be finally output to the host computer. Under the control of the MCU, the motor gradually pushes the probe into the hole to be tested. When the laser beam has completely scanned the entire hole, the motor drives the probe to retract and return to the original position. The capacitive grating displacement sensor transmits the displacement data to the host computer in real time. The port for clearing the displacement data of capacitive grating displacement sensor is connected with the MCU. When this port is at a low level, the displacement data transmitted by the capacitive grating displacement sensor is always zero; when the port is at a high level, the displacement data can be recorded normally.

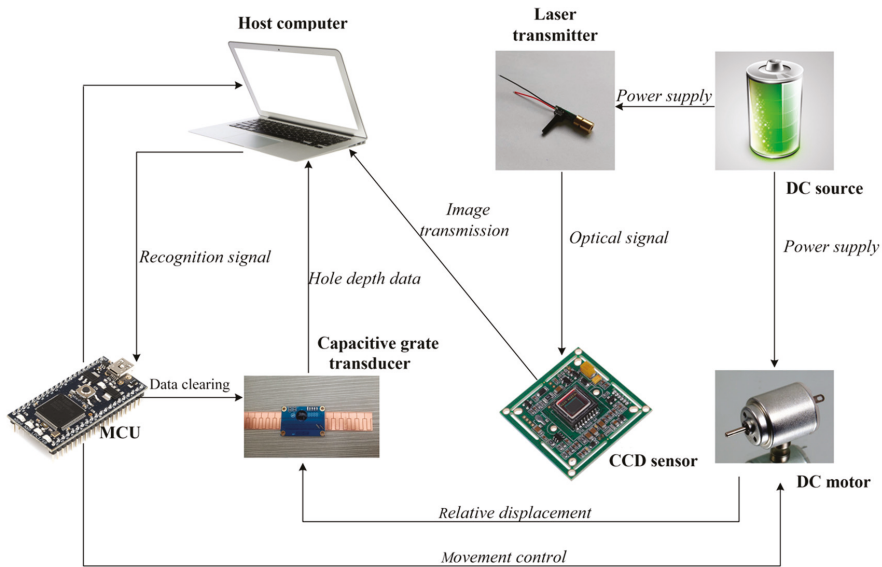


Figure 6. Hardware connection diagram of the measurement system; DC: direct current

### 2.5. Software Design

According to the measurement scheme designed in Section 2.1, the entire measurement process can be divided into five steps, as shown in Figure 1. In this section, the software design is implemented based on the proposed measurement scheme. Figure 7 shows a control flow chart of the measurement process. When the measurement starts, the motor rotates forward to drive the probe feeding forward. In this process, the MCU continues to judge the received signal. When the laser does not touch the hole wall, the signal received by the MCU is “0”, and the displacement data recorded by the capacitive grating displacement sensor is always zero, and the probe continues to feeding forward. When the laser touches the hole wall, the signal received by the MCU is “1”, and the capacitive grating displacement sensor will begin to record the displacement data of the probe. When the probe is about to extend the hole, the laser will break away from the hole wall again. In this process, the MCU will continue to judge whether the received signal is “0”. When the signal is not zero, this indicates that the probe has not yet extended the hole and the device is still in the measurement state. So the motor will continue rotating forward to drive the probe feeding forward. When the signal is “0”, it indicates that the probe has extended out of the hole. At this time, the capacitive grating displacement sensor will

stop recording displacement data. Then the motor will reverse to drive the probe back to its original position. The program will eventually output the displacement data recorded by the capacitive grating displacement sensor, which is the depth of the hole.

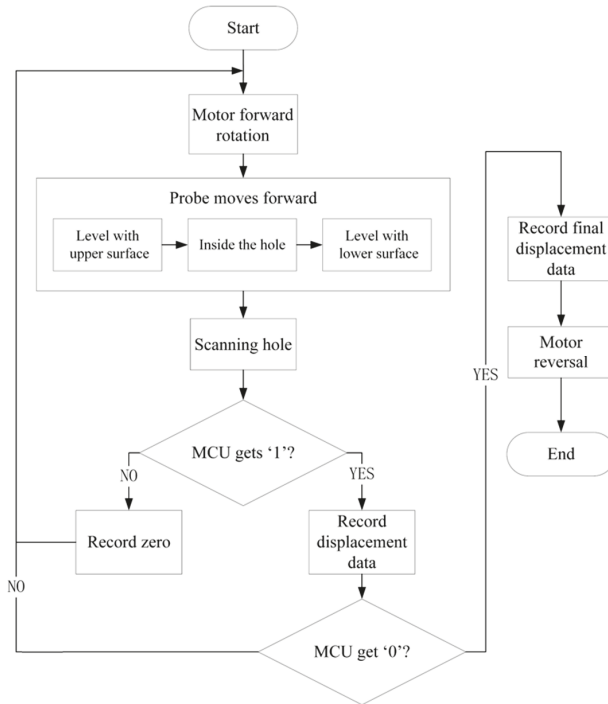


Figure 7. Control flow chart of the measurement process.

### 3. Experiment Results

In order to verify the feasibility of the developed methodology and equipment, some experiments were performed in this section. The three experiments that tested the accuracy and stability of the measuring device are: (i) measurement accuracy and repeatability verification; (ii) the effect of feed speed on the accuracy of measurement; and (iii) pixel threshold effect on measurement value.

As shown in Figure 8, the prototype of the measuring device used for the experimental test was fabricated and assembled. The experiments were carried out on a three-coordinate workbench, and the probe device was vertically fixed on the jig. The plate-like parts to be tested was fixed on the test bench below the probe, and the probe axis was parallel to the axis of the holes. The control unit used in this experiment was NXP’s LPC1768 microcontroller, which is based on ARM Cortex-M3 core. The wavelength of the laser diode used in the laser transmitter was 650 nm. The threshold of Value, Hue, and Saturation in the image screening process were both kept fixed. The resolution of the capacitive grating displacement sensor was 0.01 mm. The motor used here is a DC type with a rated voltage of 6V. The frame rate of image uptake is 30 frames per second. It is worth mentioning that in this study we only carry out preliminary experiments to get measurement accuracy and repeatability of the proposed device. The purpose is to verify the feasibility of the measurement method and the corresponding measuring instrument. The device was not compensated and optimized before measurement experiments. In a follow-up study, the accuracy compensation algorithm will be designed to further optimize the measurement accuracy.

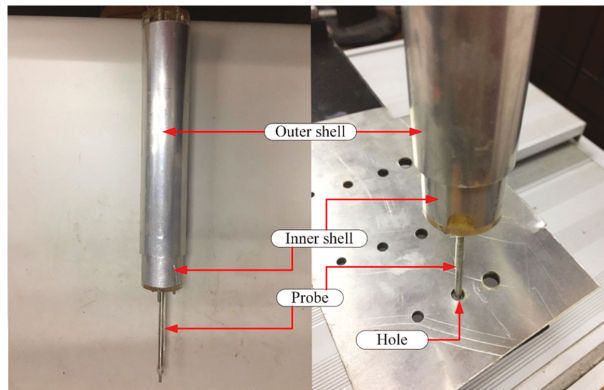


Figure 8. The entity of the device and the state of its measurement.

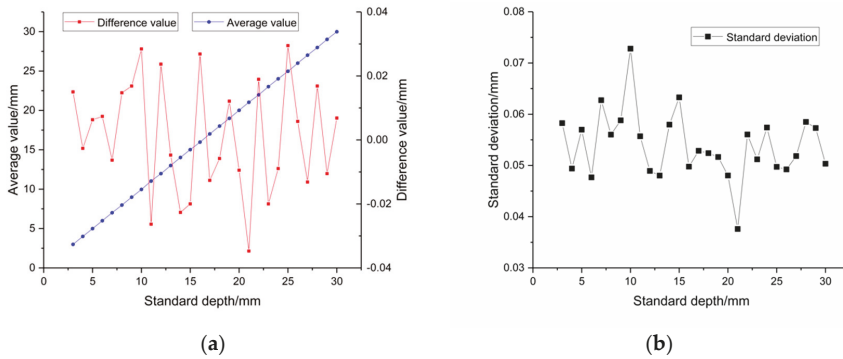
### 3.1. Measurement Accuracy and Repeatability Verification

The purpose of this experiment was to check the measurement accuracy and stability of the measuring device. The selected depths of holes are from 3 mm to 30 mm with an interval of 1 mm and the diameters of these holes are all 6 mm. We selected 28 holes with different depths as measured objects. The feed speed of the probe was fixed at 4 mm/s, and the white pixel threshold was set to 100. Each hole with a certain standard depth was repeatedly measured 50 times; thus, a total of 50 sets of measurement data were obtained. Table 1 intercepts part of the measurement data in this experiment; five groups of actual measured value were randomly selected for each hole with standard depth.

Table 1. Experimental data for measurement accuracy and repeatability verification.

Standard Depth (mm)	Five sets of Actual Measured Value (mm)				
	First	Second	Third	Fourth	Fifth
3	3.03	3.08	2.98	2.93	2.90
4	4.00	4.01	3.98	4.07	4.10
5	5.01	5.10	5.06	5.03	5.07
6	5.99	6.09	6.03	5.99	5.96
7	7.03	7.00	7.09	7.10	6.90
8	7.94	7.93	8.02	8.01	8.05
9	8.93	9.05	8.97	8.90	8.91
10	9.93	10.04	9.90	9.93	10.00
11	10.94	11.03	11.08	11.1	11.02
12	11.95	11.97	12.06	11.91	11.95
13	13.05	13.08	13.03	13.07	12.96
14	14.04	14.07	14.05	13.92	14.01
15	15.06	15.05	15.05	15.09	15.04
16	15.98	15.92	15.92	16.01	15.93
17	17.07	17.00	16.94	17.02	17.07
18	18.04	18.06	18.01	17.98	18.00
19	18.95	18.99	18.95	19.02	19.02
20	19.99	20.07	20.05	20.06	20.03
21	20.97	21.08	20.97	21.01	21.03
22	22.08	21.98	21.95	22.04	21.95
23	23.02	22.93	22.99	23.06	22.98
24	24.03	24.03	23.94	23.94	23.98
25	24.98	25.00	25.07	24.98	24.96
26	25.94	25.97	26.03	25.95	25.94
27	27.08	26.98	27.07	26.99	26.94
28	27.92	28.02	28.05	28.07	27.96
29	29.00	28.92	29.07	29.05	28.98
30	29.97	30.06	30.05	29.99	29.97

In order to analyse the measurement accuracy and stability of the device, Figure 9a gives the interpolated curves of average value and difference value under different standard depths, and Figure 9b shows the standard deviation under different standard depths in the repeatability experiment. As shown in Figure 9a, it can be seen from the curve average value that the measured results of this device can better fit the real value of hole depth, and its measuring structure is approximately a straight line. The average value of each node in the average value curve can be calculated by:



**Figure 9.** Measurement accuracy and repeatability analysis: (a) the average value and difference values under different standard depths; (b) the standard deviation curve of the repeatability experiment.

$$\mu = \frac{\sum_{n=1}^{50} S_n}{50} \tag{4}$$

The difference value curve in Figure 9a reveals the measurement error between the measured value and the true value under different standard depths. The horizontal axis is the standard depth of real hole, and the vertical axis is the error value. The error value fluctuates with the change of standard depth. The minimum error is 6 microns and the maximum error is 34 microns.

To verify the repeatability of the designed device, Figure 9b shows the standard deviation curve of the repeatability experiment. Its horizontal axis is the standard depth of the measured hole, and its vertical axis is the standard deviation of 50 sets of actual measured data. The standard deviation at each standard depth can be calculated by the following formula:

$$\delta = \sqrt{\frac{\sum_{n=1}^{50} (X_n - \mu)^2}{50}} \tag{5}$$

According to Figure 9a, standard deviation fluctuates continuously with the change of the standard hole depth. Its maximum value is 0.073 mm and its minimum value is 0.038 mm. In order to further verify the reliability of the device and determine whether the deviation is systematic, we can calculate the standard error of mean using the following formula:

$$SE = \frac{S}{\sqrt{N}} = \frac{0.073}{\sqrt{50}} = 0.010 \text{ mm} \tag{6}$$

where SE is the standard error of mean; S is the standard deviation; N is the number of measurements. The value of S = 0.073 mm is the maximum value of standard deviation, which is more able to explain the reliability of the device.



After the above calculation, the maximum value of standard error of mean was only 0.01 mm, judging by which the new approach and the device have high reliability. We will further explore ways to improve measurement repeatability in subsequent studies.

*3.2. The Effect of Feed Speed on Measurement Accuracy and Stability*

The function of laser signal reception of the device is realized by CCD and its signal processing circuit. The speed of image acquisition is limited by the performance of CCD. Moreover, due to the delay of the image processing program, some errors will appear in the process of measurement. When the laser is projected to the upper or lower surface of the hole (Step 2 and Step 4 as shown in Figure 1), displacement data should be recorded immediately. However, due to the response delay of the circuit and image processing algorithms, the starting point of the recorded data will be delayed, which will undoubtedly lead to measurement errors, and the magnitude of the error depends on the feed speed of probe. Therefore, from the above theoretical analysis, the faster the feed speed of the probe is, the greater the error will be.

**Table 2.** The effect of feed speed on the accuracy of measurement: actual measured value.

Feed Speed (mm/s)	Five Groups of Actual Measured Value (mm)				
	First	Second	Third	Fourth	Fifth
3	9.97	9.99	9.97	9.94	9.99
4	9.95	10.01	9.91	10.00	10.04
5	9.99	10.01	9.94	10.13	10.00
6	10.12	10.03	10.16	10.11	9.84
7	10.14	9.94	9.91	9.94	9.85
8	10.06	9.84	10.10	9.74	9.81
9	10.13	10.08	9.74	10.20	10.04
10	9.63	9.9	10.16	9.80	9.87
11	9.63	10.31	10.29	10.25	9.64
12	9.75	9.87	9.55	9.79	10.44

This experiment only uses the simplest way to find out the relationship between feed speed and measurement results, that is, to measure the same hole depth at different feed speeds for many points in time. This paper does not delve into the mechanism of the influence of feed speed on measurement stability and accuracy and how to compensate for it. These issues will be addressed in our subsequent studies. Table 2 lists five groups of actual measured values under a certain feed speed and gives their corresponding average value and standard deviation. In this experiment, the white pixel threshold is set to 100, and the hole with standard depth of 10 mm was measured 10 times under each feed speed (i.e., from 3 mm/s to 12 mm/s), and five groups of measurement data were randomly selected for each feed speed.

In order to observe the influence of the feed speed on the measurement accuracy, Figure 10 was drawn, according to the experimental data provided in Table 2. The horizontal axis is the feed speed of the probe, and the vertical axis is the actual measured hole depth. The standard hole depth in this experiment is 10 mm. The speed of the probe is adjusted by controlling the speed of the motor. The feed speed is from 3 mm/s to 12 mm/s. Five groups of measurement data are selected for each speed. Each group of measured data is represented by a curve with certain colour and node shape.

Figure 10 shows that the probe has higher stability when the feed speed is slower. As the feed speed increases, the measured data will gradually diverge, the fluctuation will become larger and the stability will be lower. Therefore, it can be concluded that the stability of the measuring device will decrease with the increase of feed speed of the probe. The specific relationship between the accuracy or stability of the probe and the feed speed will be analysed below.

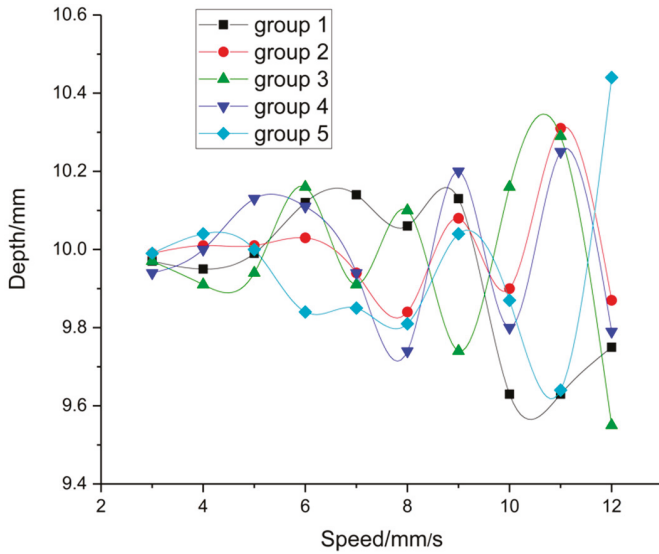


Figure 10. The depth data measured at different feed speeds.

Figure 11 shows the average value and standard deviation under different feed speeds. The average value of each node is the average of 5 groups of measured data. As can be seen from the figure, the measured value of the probe varies with the increase of feed speed. When the feed speed is less than 5 mm/s, the measured value is close to the actual value. When the feed speed continues to increase, the measured value will fluctuate and deviate from the true value.

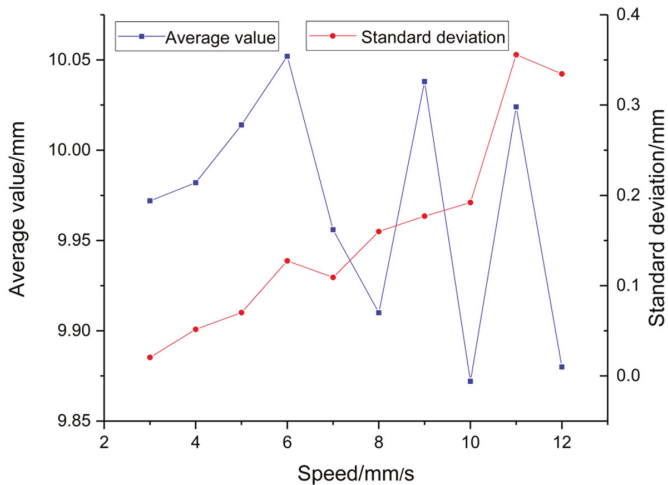


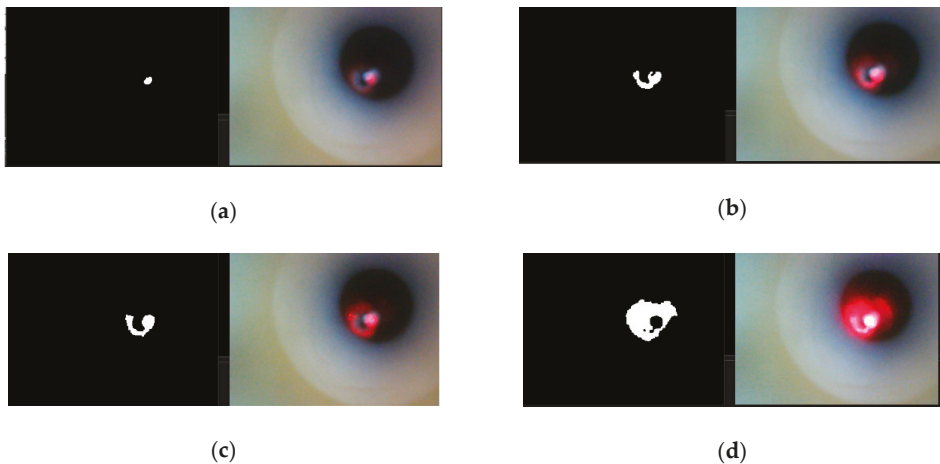
Figure 11. The average value and standard deviation under different feed speeds.

The standard deviation of the measuring device increases with the increase of feed speed, and the stability of the probe decreases accordingly. As can be seen in Figure 11, the standard deviation can be located within the range of 0.02 mm to 0.07 mm when the feed speed is less than 5 mm/s. The measurement error will continue to increase as the feed speed of the probe increases.

### 3.3. The Effect of Pixel Threshold on Measurement Accuracy and Stability

This experiment was conducted to test the relationship between the threshold value of the white pixels and the measured depth of holes. The size of the white pixel threshold will directly affect the measured results of the hole depth. The reason is that the laser is projected on the wall of the hole to form a laser spot, and the laser spot has a certain diameter. When the laser beam is just projected onto the edge of the hole wall, the laser spot is incomplete and it is a critical state. How the probe entering or leaving the hole is defined will directly affect the final measurement result.

The image of the laser spot signal received by the CCD sensor during the insertion of the probe into the hole is shown in Figure 12. At the very beginning, the laser beam just touches the upper surface of the hole and forms a very small laser spot [31]; later, about half of the laser beam is projected to the hole wall; after that, the laser spot continues to grow; and at last, all the laser beams are projected to the hole wall. The above processes are completed within a very short time. Therefore, knowing when to start triggering the capacitive grating displacement sensor to record the displacement data (i.e., how much the threshold of white pixels is set) will determine the final result of the measurement data. The relationship between the final measurement result and the white pixel threshold is explored below. The purpose of this experiment is only to find out the relationship between threshold setting and measurement results. There is no in-depth study on how threshold setting affects measurement stability and accuracy, and how to compensate for it. These issues will be addressed in follow-up studies.



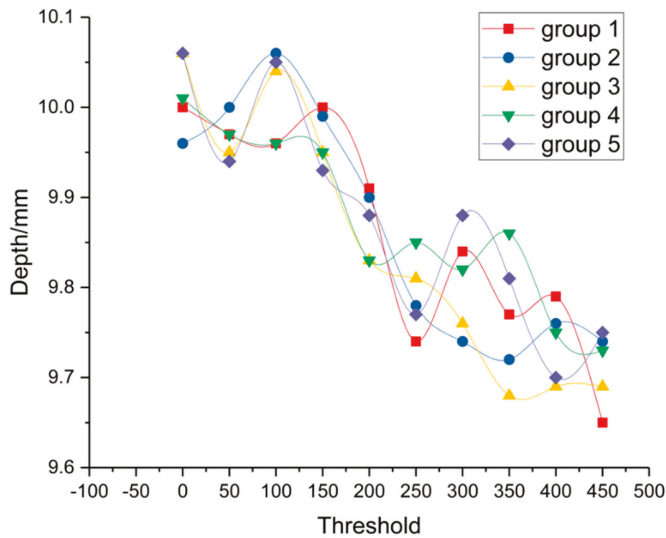
**Figure 12.** The image received by the CCD sensor of probe in the process of measurement: (a) Level with upper surface; (b) Half of the laser projection; (c) Most laser projection; (d) All laser projection.

Table 3 lists the raw data to reflect the relationship between the threshold value of the white pixels and the measured depth of holes. Ten groups of actual measured value were recorded under each selected pixel threshold in this experiment and five groups of actual measured value were randomly selected and listed in Table 3. The selected threshold of the pixel ranges from 0 to 450 with an interval of 50, the feed speed of the probe is fixed at 4 mm/s, and the standard depth of hole in this experiment is 10 mm.

**Table 3.** The effect of pixel threshold on measurement value: actual measured value.

Pixel Threshold	Five Groups of Actual Measured value (mm)				
	First	Second	Third	Fourth	Fifth
0	10	9.96	10.06	10.01	10.06
50	9.97	10	9.95	9.97	9.94
100	9.96	10.06	10.04	9.96	10.05
150	10	9.99	9.95	9.95	9.93
200	9.91	9.9	9.83	9.83	9.88
250	9.74	9.78	9.81	9.85	9.77
300	9.84	9.74	9.76	9.82	9.88
350	9.77	9.72	9.68	9.86	9.81
400	9.79	9.76	9.69	9.75	9.7
450	9.65	9.74	9.69	9.73	9.75

Figure 13 is drawn to observe the experimental results more visually. The horizontal axis is the threshold value of white pixels, and the vertical axis is the measured hole depth data. As can be seen from this figure, with the increase of the pixel threshold, the measured hole depth data gradually decreases. This trend is in line with expectations. The specific relationship between pixel threshold and measured results are further explained below.



**Figure 13.** The measurement data under different pixel thresholds.

As shown in Figure 14, the red curve is the average of 10 groups of actual measured data. It can be clearly seen that with the increase of pixel threshold, the final measurement value will gradually decrease. The blue curve in Figure 14 shows the standard deviation of the 10 groups of actual measured data. Judging from this graph, the standard deviation of the final measurement data is always fluctuating with the increase of pixel threshold, but all of them are below 0.08. The error fluctuation is within a normal range.

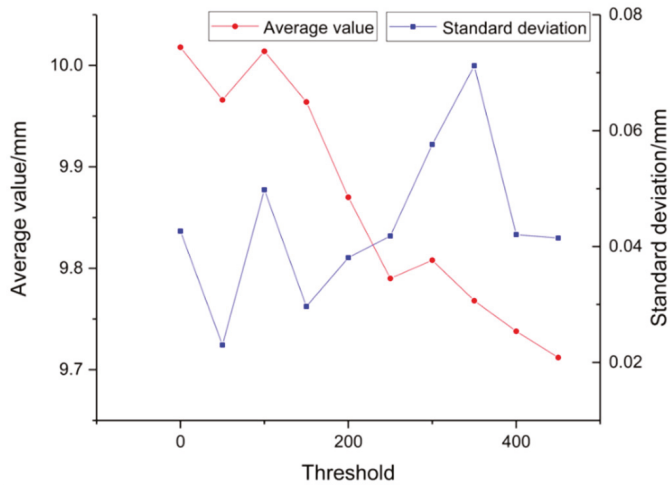


Figure 14. The average value and standard deviation in pixel threshold experiment.

#### 4. Analysis and Discussion

The above experimental results offers the following analyses:

1. According to the repetitive measurement experiment in Section 3.1, it can be seen that in the absence of error compensation algorithm and under measuring speed of 4 mm/s and white pixel threshold of 100, the measurement error of each time is about 0.05 mm and does not exceed 0.1 mm; the fluctuation range of standard deviation is between 0.038 mm and 0.073 mm. Above accuracy and stability of the designed measuring device are sufficient to meet the hole depth measuring requirements for automatic drilling and riveting of large-scale composite board parts.
2. Through the experiment on the effect of feed speed on measurement accuracy in Section 3.2, it can be seen that with the increase of probe feed speed, the measurement stability of the device decreases gradually. Considering the factors of measurement efficiency, the optimal probe feed speed is 5 mm/s, while ensuring its measurement stability.
3. Based on the effect of pixel threshold on measurement values in Section 3.3, it can be seen that with the increase of the pixel threshold, the measured hole depth gradually decreases, which is in line with expectations. Moreover, with the increase of pixel threshold, the stability of measurement fluctuates. Therefore, in order to bring the measuring result closer to the true value, the optimal pixel threshold in this experiment should be 100. In practical applications, the optimal pixel threshold needs to be adjusted according to the measurement object due to effects such as flash, burrs, iron filings, etc.
4. In follow-up studies, we can further design and optimize an error compensation algorithm and automatic calibration method, study the corresponding relationship between feed speed and measurement error and find an effective compensation method, study the mapping relationship between pixel threshold and measurement error and offer a compensation method, study the specific relationship between the diameter of the laser spot and the measurement accuracy, etc., so as to further improve the measurement accuracy and efficiency of the device.
5. This paper only focuses on the design of measuring methods and measuring devices, and preliminarily proves the feasibility of the proposed method. Follow-up research can integrate the prototype device into the automatic riveting manipulator arm and measure the actual engineering sample, and further explore the feasibility and stability of its engineering application.

## 5. Conclusions

In order to get rid of manual intervention in hole depth measurement and realize on-machine automatic measurement in the automatic drilling and riveting process of largescale composite board parts, a novel hole depth measuring device was designed and developed to achieve accurate and efficient automatic measurement of hole depth parameters. The mechanical structure of the measuring device was developed according to our newly designed measurement scheme and optical path. The structure consists of three parts, namely a probe, an inner casing and an outer casing, the purpose of which is to convert the hole depth data into displacement data of probe motion. The electrical hardware of the device consists of three units: a laser transceiver unit, a displacement measuring unit, and a driving unit. Among them, the laser transceiver unit involves optical path design and image processing; the displacement measuring unit involves the data communication between capacitive grating displacement sensor and the host computer; the driving unit relates to the motion control of the probe. Finally, the experimental results indicated that the proposed measurement method and device can accurately measure the hole depth parameters, and its accuracy and stability are sufficient to meet the hole depth measuring requirements for automatic drilling and riveting of largescale composite board parts. In particular, the influence of probe feed speed and white pixel threshold on the accuracy and stability of the measuring device was explored, which lays the foundation for subsequent research of an error compensation algorithm and a probe calibration method.

**Author Contributions:** All works with relation to this paper has been accomplished by all authors' efforts. Conceptualization, X.W. and Y.T.; methodology, Y.T.; software, N.W.; validation, Z.H.; formal analysis, N.W.; investigation, W.Z.; resources, X.W.; data curation, X.W.; writing—original draft preparation, N.W.; writing—review and editing, Y.T. and N.W.; visualization, Z.H.; supervision, W.Z.; project administration, W.Z.; funding acquisition, Y.T. and W.Z.

**Funding:** This research was funded by the National Natural Science Foundation of China (Grant Nos. 51505310, 51435011), the Key Research & Development Program of Sichuan Province of China (Grant No. 2018GZ0282) and the Science & Technology Ministry Innovation Method Program of China (Grant No. 2017IM040100).

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

- Dehoff, R.; Duty, C.; Peter, W. Case study: Additive manufacturing of aerospace brackets. *Adv. Mater. Process.* **2013**, *171*, 22–25. [[CrossRef](#)]
- Yu, L.; Zhang, Y.; Bi, Q.; Wang, Y. Research on surface normal measurement and adjustment in aircraft assembly. *Precis. Eng.* **2017**, *50*, 482–493. [[CrossRef](#)]
- Patterson, J.B.; Grenestedt, J.L. Manufacturing of a composite wing with internal structure in one cure cycle. *Compos. Struct.* **2018**, *206*, 601–609. [[CrossRef](#)]
- Ahmed, E.; Wan Badaruzzaman, W.H.; Wright, H.D. Two-way bending behavior of profiled steel sheet dry board composite panel system. *Thin-Walled Struct.* **2002**, *40*, 971–990. [[CrossRef](#)]
- Tian, W.; Zhou, Z.; Liao, W. Analysis and investigation of a rivet feeding tube in an aircraft automatic drilling and riveting system. *Int. J. Adv. Manuf. Technol.* **2016**, *82*, 973–983. [[CrossRef](#)]
- Bernard, T.; Bergmann, H.W.; Haberling, C.; Haldenwanger, H. Joining Technologies for Al-Foam—Al-Sheet Compound Structures. *Adv. Eng. Mater.* **2002**, *4*, 798–802. [[CrossRef](#)]
- Li, Y.; Li, N.; Gao, J. Tooling design and microwave curing technologies for the manufacturing of fiber-reinforced polymer composites in aerospace applications. *Int. J. Adv. Manuf. Technol.* **2014**, *70*, 591–606. [[CrossRef](#)]
- Hassan, A.; Maskin, M.; Tom, P.P.; Brayon, F.; Hlavac, P.; Mohamed, F. Operator response modeling and human error probability in TRIGA Mark II research reactor probabilistic safety assessment. *Ann. Nucl. Energy* **2017**, *102*, 179–189. [[CrossRef](#)]
- Shetty, D.; Eppes, T.; Campana, C.; Filburn, T.; Nazaryan, N. New Approach to the Inspection of Cooling Holes in Aero-Engines. *Opt. Lasers Eng.* **2009**, *47*, 686–694. [[CrossRef](#)]
- Baeg, M.H.; Baeg, S.H.; Moon, C.; Jeong, G.M.; Ahn, H.S.; Kim, D.H. A New Robotic 3D Inspection System of Automotive Screw Hole. *Int. J. Control Autom. Syst.* **2008**, *6*, 740–745. [[CrossRef](#)]

11. Usamentiaga, R.; Venegas, P.; Guerediaga, J.; Vega, L.; Lopez, I. Non-destructive inspection of drilled holes in reinforced honeycomb sandwich panels using active thermography. *Infrared Phys. Technol.* **2012**, *55*, 491–498. [[CrossRef](#)]
12. Gong, X.; Su, H.; Zhang, Z.T.; Xu, C. Vision-based quality inspection for components with small diameter and deep hole. In Proceedings of the IEEE International Conference on Mechatronics and Automation, Takamatsu, Japan, 6–9 August 2017; IEEE: New York, NY, USA, 2017; pp. 2016–2021.
13. Bernard, F.; Flaherty, T.; O'Connor, G.M. Novel fiber-based technique for inspection of holes in narrow-bore tubes. In Proceedings of the Optical Measurement Systems for Industrial Inspection VI, Munich, Germany, 15–18 June 2009; SPIE-The International Society for Optical Engineering: Bellingham, WA, USA, 2009; p. 7389.
14. Zavyalov, P. 3D Hole Inspection Using Lens with High Field Curvature. *Meas. Sci. Rev.* **2015**, *15*, 52–57. [[CrossRef](#)]
15. Quest, D.; Gayer, C.; Hering, P. Depth measurements of drilled holes in bone by laser triangulation for the field of oral implantology. *J. Appl. Phys.* **2012**, *111*, 013106. [[CrossRef](#)]
16. Takeda, M.; Aoki, T.; Miyamoto, Y.; Tanaka, H.; Gu, R.; Zhang, Z. Absolute three-dimensional shape measurements using coaxial and coimage plane optical systems and Fourier fringe analysis for focus detection. *Opt. Eng.* **2000**, *39*, 61–68. [[CrossRef](#)]
17. Heredia-Ortiz, M.; Patterson, E.A. On the industrial applications of moire and fringe projection techniques. *Strain* **2003**, *39*, 95–100. [[CrossRef](#)]
18. Casavola, C.; Pappalettera, G.; Pappalettere, C. Design of a fiber optics fringe projector for 3D reconstruction of dental elements. In Proceedings of the 3rd Mediterranean Photonics Conference, Trani, Italy, 7–9 May 2014. [[CrossRef](#)]
19. Lin, C.H.; Powell, R.A.; Jiang, L.; Xiao, H.; Chen, S.J.; Tsai, H.L. Real-time depth measurement for micro-holes drilled by lasers. *Meas. Sci. Technol.* **2010**, *21*, 025307. [[CrossRef](#)]
20. Ho, C.C.; Chiu, C.M.; Chang, Y.J.; Hsu, J.C.; Kuo, C.L. On-line depth measurement for laser-drilled holes based on the intensity of plasma emission. *Meas. Sci. Technol.* **2014**, *25*, 094007. [[CrossRef](#)]
21. Hung, Y.H.; Chien, H.L.; Lee, Y.C. Excimer Laser Three-Dimensional Micromachining Based on Image Projection and the Optical Diffraction Effect. *Appl. Sci.* **2018**, *8*, 1690. [[CrossRef](#)]
22. Agnesi, A.; Degiorgio, V. Beam splitter phase shifts: Wave optics approach. *Opt. Laser Technol.* **2017**, *95*, 72–73. [[CrossRef](#)]
23. Kumar, U.P.; Haifeng, W.; Mohan, N.K.; Kothiyal, M.P. White light interferometry for surface profiling with a colour CCD. *Opt. Lasers Eng.* **2012**, *50*, 1084–1088. [[CrossRef](#)]
24. Jahedsaravani, A.; Massinaei, M.; Marhaban, M.H. Development of a machine vision system for real-time monitoring and control of batch flotation process. *Int. J. Miner. Process.* **2017**, *167*, 16–26. [[CrossRef](#)]
25. Carney, M.N.; Johnston, W.M. A novel regression model from RGB image data to spectroradiometric correlates optimized for tooth colored shades. *J. Dent.* **2016**, *51*, 45–48. [[CrossRef](#)]
26. Chaves-Gonzalez, J.M.; Vega-Rodriguez, M.A.; Gomez-Pulido, J.A.; Sanchez-Perez, J.M. Detecting skin in face recognition systems: A colour spaces study. *Digit. Signal Proc.* **2010**, *20*, 806–823. [[CrossRef](#)]
27. Seo, D.K.; Kim, Y.H.; Eo, Y.D.; Park, W.Y. Learning-Based Colorization of Grayscale Aerial Images Using Random Forest Regression. *Appl. Sci.* **2018**, *8*, 1269. [[CrossRef](#)]
28. Pai, K.J.; Qin, L.D.; Lin, C.H.; Tang, S.Y. Start-Up Current Spike Mitigation of High-Power Laser Diode Driving Controller for Vehicle Headlamp Applications. *Appl. Sci.* **2018**, *8*, 1532. [[CrossRef](#)]
29. Chen, Q.; Sun, Q.S.; Heng, P.A.; Xia, D.S. A double-threshold image binarization method based on edge detector. *Pattern Recognit.* **2008**, *41*, 1254–1267. [[CrossRef](#)]
30. Nouira, H.; Vissiere, A.; Damak, M.; David, J.M. Investigation of the influence of the main error sources on the capacitive displacement measurements with cylindrical artefacts. *Precis. Eng.* **2013**, *37*, 721–737. [[CrossRef](#)]
31. Lee, D.; Cho, J.; Kim, C.H.; Lee, S.H. Application of laser spot cutting on spring contact probe for semiconductor package inspection. *Opt. Laser Technol.* **2017**, *97*, 90–96. [[CrossRef](#)]



Article

# Metrology of Nanostructures by Tomographic Mueller-Matrix Scatterometry

Chao Chen, Xiuguo Chen \*, Yating Shi, Honggang Gu, Hao Jiang  and Shiyuan Liu \*

State Key Laboratory for Digital Manufacturing Equipment and Technology, Huazhong University of Science and Technology, Wuhan 430074, China; chaochen@hust.edu.cn (C.C.); yatingshi@hust.edu.cn (Y.S.); hongganggu@hust.edu.cn (H.G.); hjiang@hust.edu.cn (H.J.)

\* Correspondence: xiuguochen@hust.edu.cn (X.C.); shyliu@hust.edu.cn (S.L.)

Received: 16 November 2018; Accepted: 10 December 2018; Published: 12 December 2018



**Abstract:** The development of necessary instrumentation and metrology at the nanoscale, especially fast, low-cost, and nondestructive metrology techniques, is of great significance for the realization of reliable and repeatable nanomanufacturing. In this work, we present the application of a homemade novel optical scatterometer called the tomographic Mueller-matrix scatterometer (TMS), for the measurement of photoresist gratings. The TMS adopts a dual rotating-compensator configuration and illuminates the nanostructure sequentially under test conditions by a plane wave, with varying illumination directions and records. For each illumination direction, the polarized scattered field along various directions of observation can be seen in the form of scattering Mueller matrices. That more scattering information is collected by TMS than conventional optical scatterometry ensures that it achieves better measurement sensitivity and accuracy. We also show the capability of TMS for determining both grating pitch and other structural parameters, which is incapable by current zeroth-order methods such as reflectometry- or ellipsometry-based scatterometry.

**Keywords:** ellipsometry; scatterometry; Mueller matrix; diffraction grating; inverse scattering; pitch measurement

---

## 1. Introduction

Nanomanufacturing involves manufacturing of products (including materials, structures, devices, and systems) with feature dimensions at the nanoscale [1]. It is an essential bridge between the newest nanoscience discoveries and real-world nanotechnology products. One critical challenge to the realization of nanomanufacturing is the development of the necessary metrology at the nanoscale [2]. Although scanning electron microscopy (SEM), atomic force microscopy (AFM), and transmission electron microscopy (TEM) have the ability to analyze extremely small targets, they are faced with high costs and throughput concerns, and are in general not well suited for integrated metrology applications.

In comparison with SEM, AFM, and TEM, optical scatterometry based on a reflectometer or an ellipsometer has become one of the most important techniques for measuring the critical dimension (CD) and overlay of nanostructures in semiconductor manufacturing, due to its inherent noncontact, nondestructive, time-effective, and relatively inexpensive merits [3–9]. It is different from conventional image-based metrology techniques, such as optical microscopy, in that the measurement in optical scatterometry is not a WYSIWYG (what you see is what you get) process and typically involves the solution of an inverse problem by fitting the measured data with a multiparameter model that describes the light-nanostructure interaction. Even so, it is not restricted by the well-known Abbe diffraction limit in image-based techniques and thus plays an important role in addressing devices with sub-wavelength feature sizes in the semiconductor industry.



With the ever-decreasing dimensions of advanced technology nodes (22 nm and beyond), there are also some challenges and limitations to optical scatterometry [10,11], such as the parameter correlation issue. In addition, optical scatterometry is mostly suitable for measuring repetitive dense structures, but infeasible for the measurement of isolated or generally non-periodic structures. To address the challenges or limitations in conventional optical scatterometry, several designs have been presented with the idea of trying to collect the scattering information about the nanostructure under test conditions as much as possible, such as with the goniometric optical scatter instrument [12–14], through-focus scanning optical microscopy [15], scatterfield microscopy [16], tomographic diffractive microscopy [17,18], and Fourier scatterometry [19,20]. Recently, we have also developed a novel instrument called the tomographic Mueller-matrix scatterometer (TMS) [21]. The TMS illuminates a sample sequentially by a plane wave with varying illumination directions (incidence angles  $0^\circ\sim 65.6^\circ$  and azimuthal angles  $0^\circ\sim 360^\circ$ ) and records. For each illumination direction, the polarized scattered field along various directions of observation (scattering angles  $0^\circ\sim 67^\circ$  and azimuthal angles  $0^\circ\sim 360^\circ$ ) can be seen in form of scattering Mueller matrices. The experiments performed on a Si grating had preliminarily demonstrated the potential of a TMS in nanostructure metrology [21].

Due to the strong correlation between pitch and other structural parameters, it is a common practice to predetermine grating pitch by another metrology tool such as AFM, or directly fix grating pitch to its nominal value in the solution of the inverse problem in optical scatterometry. In this work, we present the application of a TMS for the accurate reconstruction of lithographic patterns. We show the capability of a TMS for determining both pitch and other structural parameters, which is incapable by current zeroth-order methods such as reflectometry- or ellipsometry-based scatterometry.

The rest of this paper is organized as follows. Section 2 briefly introduces the principle of the TMS. Section 3 presents the associated methods for determining pitch and other structural parameters by TMS. Section 4 first describes the sample under measurement and then provides the details of measurement results by TMS. Finally, we draw some conclusions in Section 5.

## 2. Experimental Setup

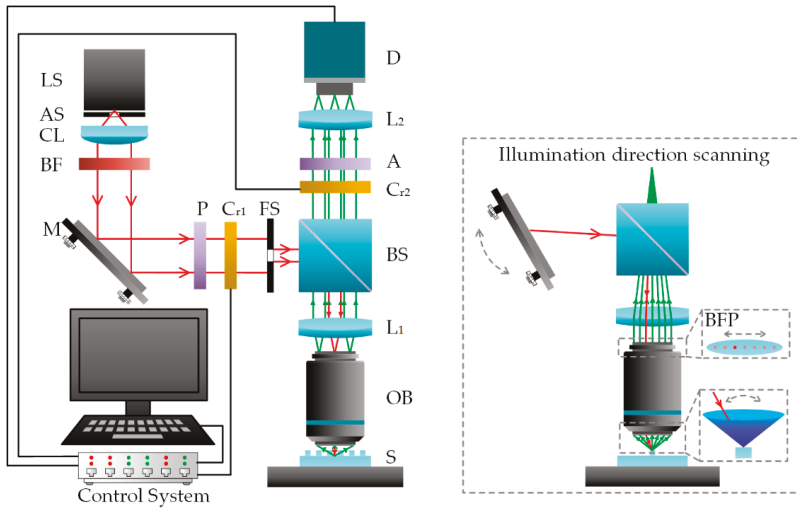
Figure 1 presents the scheme of the developed TMS, which is a combination of a Mueller matrix ellipsometer (MME) [22] and a reflection microscope. As shown in Figure 1, a broadband white light from a light source (LS; LDLS<sup>TM</sup>, EQ-99FC, Energetiq, Woburn, MA, USA) is collimated by a flat convex lens (CL) and then filtered to be a monochromatic light by a bandpass filter (BF; FL632.8-1, Thorlabs, Inc., Newton, NJ, USA). After reflected by a rotatable flat mirror (M), which functions as a scanner of illumination directions, the light beam passes through a polarization state generator (PSG) made up of a polarizer (P; PGT5012, Union Optic, Inc., Wuhan, China) and an optimally designed zero-order quartz crystal waveplate ( $C_{r1}$ ). Then, a non-polarizing beam splitter (BS) and an achromatic doublet ( $L_1$ ) guides light rays to be focused on the back focal plane (BFP) of a high numerical-aperture (NA) objective lens (OB; EC Epiplan-Apochromat  $50\times/0.95$  HD DIC, Zeiss, Inc., Oberkochen, Germany). The emerging light from the BFP of the OB is then parallelly incident upon the sample. Afterwards, the scattered light from the sample successively traverses the OB,  $L_1$ , BS, a polarization state analyzer that consists of the same polarization components as a PSG but in a reverse order, another achromatic doublet ( $L_2$ ), and is finally collected by a detector (D; PCO.edge 5.5, PCO, Inc., Kelheim, Germany).

As presented in the inset of Figure 1, the position of the focal point of the light beam on the BFP of the OB can be changed by rotating the flat mirror M, which further leads to the change of illumination direction on the sample. An epi-illumination setup is designed to collect the scattered-field distribution associated with each illumination direction by imaging the BFP of the OB. Thanks to the dual rotating-compensator configuration, a  $4\times 4$  Mueller matrix associated with each point on the BFP of the OB can be obtained. Since the 16 elements of a Mueller matrix contain all polarization information that one can extract from a linear polarization scattering process, the full polarization properties of the scattered field are thus achieved.

As for the Mueller matrix measurement, the Stokes vector  $S_{out}$  of the light beam detected by the detector can be expressed as the following Mueller matrix product with incident Stokes vector  $S_{in}$  [8,22]:

$$S_{out} = [M_A R(A_0)] [R(-C_{20}) M_{C2}(\delta_2) R(C_{20})] [M_t^{BS} M_b^{OB} M_S M_f^{OB} M_r^{BS}] \times [R(-C_{10}) M_{C1}(\delta_1) R(C_{10})] [R(-P_0) M_P] S_{in}, \quad (1)$$

where  $M_P$ ,  $M_{C1}(\delta_1)$ ,  $M_{C2}(\delta_2)$ , and  $M_A$  are the Mueller matrices of the polarizer P, the first compensator  $C_{r1}$ , the second compensator  $C_{r2}$ , and the analyzer A;  $M_r^{BS}$  and  $M_t^{BS}$  are reflection and transmission matrices of the BS, respectively.  $M_f^{OB}$  and  $M_b^{OB}$  represent Mueller matrices of the OB in the illumination and collection directions, respectively.  $M_S$  is the sample Mueller matrix;  $R(\cdot)$  is the Mueller rotation transformation matrix for rotation by an angle, which can be the initial transmission-axis orientations of the polarizer and analyzer,  $P_0$  and  $A_0$ , the initial fast-axis orientations of the 1st and 2nd rotating compensators,  $C_{10}$  and  $C_{20}$ .  $\delta_1$  and  $\delta_2$  are the phase retardances of 1st and 2nd compensators, respectively. Details about the calibration of  $P_0$ ,  $A_0$ ,  $C_{10}$ ,  $C_{20}$ ,  $\delta_1$  and  $\delta_2$ , as well as  $M_r^{BS}$ ,  $M_t^{BS}$ ,  $M_f^{OB}$  and  $M_b^{OB}$  can be found in Ref. [21] and are omitted here for the sake of brevity.



**Figure 1.** Scheme of the tomographic Mueller-matrix scatterometer (TMS). LS, light source; AS, aperture stop; CL, collimated lens; BF, bandpass filter; M, mirror; P and A, polarizer and analyzer;  $C_{r1}$  and  $C_{r2}$ , 1st and 2nd rotating compensators; FS, field stop; BS, Beam splitter;  $L_1$  and  $L_2$ , imaging lenses; OB, objective lens; S, sample; D, detector. The inset on the right side illustrates the principle of illumination direction scanning.

### 3. Methods

#### 3.1. Inverse Modeling

The inverse modeling includes the solution of a direct problem and an indirect problem. The direct problem involves the establishment of a model to simulate the theoretical Mueller matrices of a nanostructure. The indirect problem involves the reconstruction of the nanostructure profile with the objective of finding an optimal input to the above established model whose simulated Mueller matrices can best match the measured values.

As for the direction problem here, the rigorous coupled-wave analysis (RCWA) [23,24] is used to simulate theoretical Mueller matrices of a periodic nanostructure. Figure 2 shows the case of polarized light incidence for a one-dimensional grating structure. The grating is denoted by vector

$\mathbf{x} = [x_1, x_2, \dots, x_n]^T$  where  $x_1, x_2, \dots, x_n$  represent structural parameters such as top CD, height, sidewall angle, and so on. As shown in Figure 2b, the grating is divided by several different layers with boundaries joined together. Inverse modeling includes three steps: (1) derive electromagnetic expression from Maxwell’s equations; (2) expand the permittivity function and electromagnetic fields into Fourier series; (3) and match the tangential field components at boundaries between different layers, reducing the boundary-value problem to an algebraic problem. According to the reflection coefficients calculated by solving the algebraic problem, a  $2 \times 2$  Jones matrix  $\mathbf{J}$  associated with the diffracted light can be formulated by:

$$\begin{bmatrix} E_{rp} \\ E_{rs} \end{bmatrix} = \mathbf{J}(\mathbf{x}) \begin{bmatrix} E_{ip} \\ E_{is} \end{bmatrix} = \begin{bmatrix} r_{pp} & r_{ps} \\ r_{sp} & r_{ss} \end{bmatrix} \begin{bmatrix} E_{ip} \\ E_{is} \end{bmatrix}, \tag{2}$$

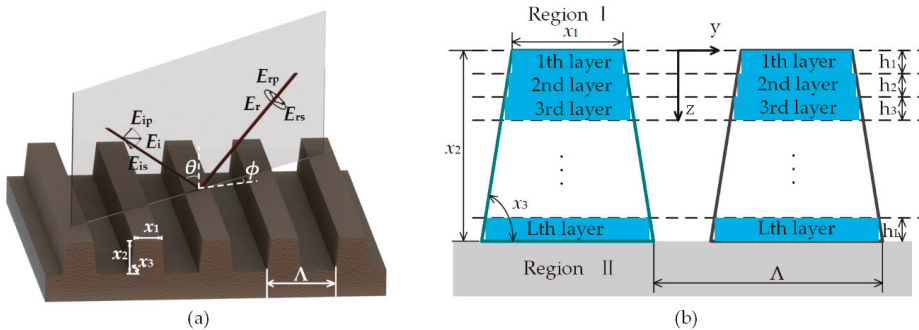
where  $E_{s,p}$  refers to the electric field component perpendicular and parallel to the plane of incidence, respectively. When depolarization could be ignored, the Mueller matrix of the sample can be obtained by:

$$\mathbf{M} = \mathbf{A}(\mathbf{J} \otimes \mathbf{J}^*)\mathbf{A}^{-1}, \tag{3}$$

where the symbol  $\otimes$  denotes the Kronecker product.  $\mathbf{J}^*$  is the complex conjugate of  $\mathbf{J}$ , and the matrix  $\mathbf{A}$  is given by:

$$\mathbf{A} = \begin{bmatrix} 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & -1 \\ 0 & 1 & 1 & 0 \\ 0 & i & -i & 0 \end{bmatrix}. \tag{4}$$

when depolarization could not be ignored, the sample Mueller matrix can be represented as the sum of multiple non-depolarizing Mueller matrices  $\mathbf{M} = \sum_i \mathbf{M}_{i,}$  and each non-depolarizing Mueller matrices can be calculated as mentioned above [25]. In practice, the Mueller matrix  $\mathbf{M}$  is usually normalized to the (1, 1)th element  $M_{11}$ , with the \*/normalized Mueller matrix elements being  $m_{ij} = M_{ij}/M_{11}$ .



**Figure 2.** (a) Representation of polarized light incidence for a one-dimensional grating; (b) layers division in inverse modeling using rigorous coupled-wave analysis.

As for the indirect problem, a weighted least-squares regression analysis (Levenberg-Marquardt algorithm [26]) is performed, during which the structural parameters under measurement are varied until the calculated and measured data match as close as possible. This is done by minimizing a weighted mean square function  $\chi_r^2$  defined by:

$$\chi_r^2 = \frac{1}{15N - P} \sum_{k=1}^N \sum_{j=1}^4 \left[ \frac{m_{ij,k}^{\text{exp}} - m_{ij,k}^{\text{calc}}(\mathbf{x})}{\sigma(m_{ij,k})} \right]^2, \tag{5}$$

where  $k$  indicates the  $k$ th data point, which could be a wavelength, incidence angle, or azimuthal angle point. From the total number  $N$ , indices  $i$  and  $j$  show all the Mueller matrix elements except  $m_{11}$  (normalized to  $m_{11}$ ).  $m_{ij,k}^{exp}$  denotes the Mueller matrix elements measured with the standard deviation of  $\sigma(m_{ij,k})$ , and  $m_{ij,k}^{calc}(\mathbf{x})$  denotes the calculated Mueller matrix elements associated with the vector  $\mathbf{x}$ . Since the Levenberg-Marquardt algorithm is apt to fall into local minima, proper starting values for the fitting parameters are required in the iteration, which are usually taken as their nominal values or as the SEM-measured values if available. In addition, the curve of  $\chi_r^2$  in a small parameter range centered on the extracted parameter values is usually plotted to check the uniqueness of the minimal point, to ensure that the found solution is the globally optimum solution.

3.2. Pitch Measurement

In scatterometric data analysis, the grating pitch is commonly fixed to extract other profile parameters. However, inaccurate pitch values would lead to a poor result. Moreover, it is difficult to set the pitch dimension  $\Lambda$  as a parameter, because it is coupled with other parameters which may make measurements imprecise. Next, we show how to measure grating pitch by TMS.

For grating periods  $\Lambda$  larger than  $\lambda/2$ , where  $\lambda$  is the wavelength of the incident light, the first-order diffracted light can be easily observed at an emergent angle  $\theta$ . Assuming  $\theta$  is the difference between the diffraction angles of the zeroth- and first-order diffracted light, according to the grating equation the grating pitch can be written as:

$$\Lambda = \frac{\lambda}{\sin \theta \cos \theta + (1 - \cos \theta) \sin \theta}. \tag{6}$$

It is easy to see the variable  $\theta$  become the only parameter to measure pitch. There are three ways to obtain the value of  $\theta$ . The first one is a so-called Littrow configuration [27]. The schematic drawing of the method is presented in Figure 3. The main point is to rotate the grating to ensure the incidence light coincides with the first-order diffracted light. Assumed rotating angle of the grating is  $\alpha$ , then it is simple to know  $\theta = 2\alpha$  and  $\theta = 90^\circ - \alpha$ . Equation (6) could be rewritten as:

$$\Lambda = \frac{\lambda}{2 \sin \alpha}. \tag{7}$$

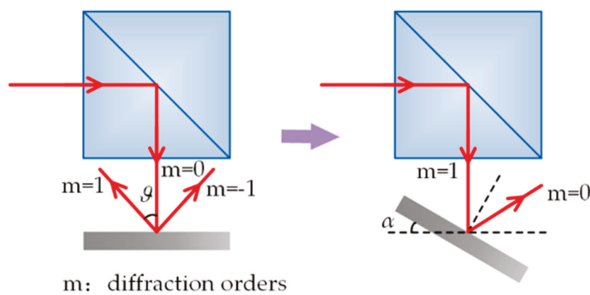


Figure 3. Littrow configuration for grating pitch measurement.

The second method is to use a grating whose pitch is known as a reference. Because a camera is used as a detector, the number of pixels can be acquired between the zeroth- and first-order diffracted light spot projected on the camera, set as  $d_1$  and  $d_2$  for reference and sample grating, respectively. Next, we can calculate the difference of diffraction angles  $\theta_1$  and angle of zeroth-order diffracted light  $\beta_0$ , which is equal to the incidence angle  $\theta$ , using the grating equation. When the sample grating

is illuminated by the same incidence light,  $\beta_0$  is same to that of the reference one. Afterwards, the difference of diffraction angles  $\theta_2$  for simple grating can be obtained through the geometric relation:

$$\frac{\tan \theta_2 - \tan \beta_0}{\tan \theta_1 - \tan \beta_0} = \frac{d_2}{d_1}. \tag{8}$$

In case of the influence of aberrations of the optical system, test targets such as concentric circle and crosshair grid targets are used to calibrate and compensate errors induced by the system. If there is no reference grating available, as a third method we could use two different lights whose wavelengths ( $\lambda_1$  and  $\lambda_2$ ) are known to illuminate sample grating at the same incidence angle. It is similar to the second method, thus there is no more detailed description. For this case, another geometric relation is:

$$\frac{\sin(\theta'_2 + \beta_0) + \sin \beta_0}{\sin(\theta'_1 + \beta_0) + \sin \beta_0} = \frac{\lambda_2}{\lambda_1}. \tag{9}$$

The intensity of diffracted light is recorded by the camera. However, for a real light beam, the diameter could not be extremely small, so on the camera there usually is a spot with several pixels. When we deal with the signal, the data from a pixel is chosen for analysis. For a practical optical system, there are inevitably aberrations and misplacements of optical components. Due to the asymmetrical intensity distribution, the pixel chosen would influence the result. Therefore, a pixel picked from a spot shouldn't be random. The approach is based on calculating the center of mass (COM),

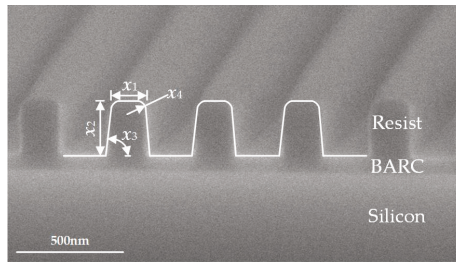
$$p = \frac{\sum_{i=1}^N \sum_{j=1}^M i I_{ij}}{\sum_{i=1}^N \sum_{j=1}^M I_{ij}}, \quad q = \frac{\sum_{i=1}^N \sum_{j=1}^M j I_{ij}}{\sum_{i=1}^N \sum_{j=1}^M I_{ij}}, \tag{10}$$

where  $I_{ij}$  is the intensity of pixel  $(i, j)$  in a  $N \times M$  image, and  $(p, q)$  is the COM of light spot.

## 4. Results and Discussion

### 4.1. Sample Description

The sample is a photoresist array on a bottom anti-reflective coating (BARC) layer deposited on a Si substrate. Figure 4 presents a cross-sectional SEM image of the photoresist grating structure. Nominal dimensions of the grating structure are as follows: grating pitch is 400 nm, top critical dimension  $x_1 = 200$  nm, grating height  $x_2 = 311$  nm, and sidewall angle  $x_3 = 90^\circ$ . The SEM image indicates that the top corner rounding could not be ignored, while the bottom corner rounding is unnoticeable. Therefore, the top corner rounding denoted by parameter  $x_4$  in Figure 4 is also taken into consideration. In data analysis, we fixed the grating pitch and let the parameters  $x_1 \sim x_4$  vary. As well, optical constants are necessary, and for the Si substrate, it was taken from literature [28]. The optical properties of the BARC layer were modeled using a Tauc-Lorentz model [29], the parameters of which were taken as  $A = 6.4378$  eV,  $C = 0.609$  eV,  $E_0 = 5.851$  eV,  $E_g = 4.914$  eV, and a Gaussian model [30], parameters of which were taken as  $A = 0.8567$  eV,  $\omega_0 = 0.3174$  eV,  $\sigma = 6.429$  eV, respectively. The optical properties of the resist layer were modeled using a Cody-Lorentz model [31], the parameters of which were taken as  $A = 2.261$  eV,  $\Gamma = 0.750$  eV,  $E_0 = 6.189$  eV,  $E_g = 4.867$  eV,  $E_t = 0.292$  eV,  $E_u = 2.000$  eV, and three Lorentz models [32], the parameters of which were taken as  $A = 0.7089$  eV,  $0.1628$  eV,  $0.1755$  eV,  $\Gamma = 0.3315$  eV,  $0.2055$  eV,  $0.3886$  eV,  $E_n = 5.414$  eV,  $4.404$  eV,  $4.786$  eV, respectively. Both the BARC and photoresist layers were predetermined by a commercial MME (ME-L, Eoptics Technology Co., Wuhan, China), which can provide full  $4 \times 4$  Mueller matrices in the spectral range of 200–1000 nm.



**Figure 4.** SEM micrograph and geometric model of the photoresist grating.

#### 4.2. Measurement of the Grating Pitch

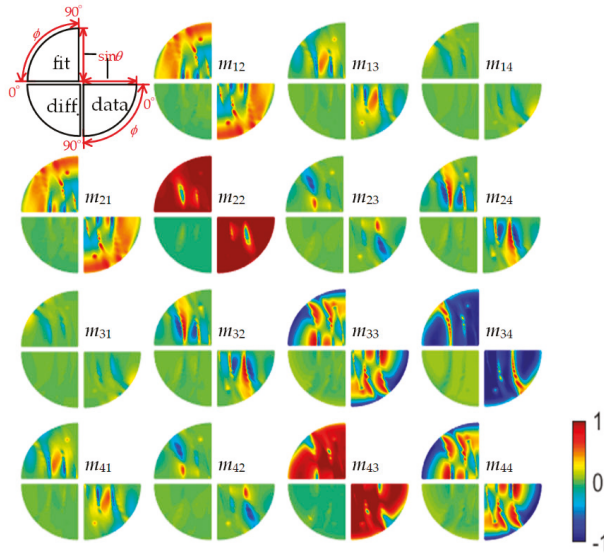
We used the Littrow configuration method to measure the grating pitch by TMS. We set the incidence angle at  $\theta = 0^\circ$ , thus the zeroth-order diffracted light is overlapped with the incidence light. Then, the sample stage was rotated so that the zeroth-order diffraction light deflected from its initial direction and the first-order diffracted light coincided with the incident light. The rotating angle of the stage was  $50.2^\circ$ , and the illumination wavelength used was 633 nm. Substituting the values into Equation (7), and the calculated pitch of the grating is about 412 nm, which is in good agreement with the SEM-measured pitch value of about 414 nm from Figure 4.

As described in Section 3.2, the Littrow configuration method is applicable when the grating pitch is greater than  $\lambda/2$ . In our case, for a wavelength  $\lambda = 633$  nm, the grating pitch cannot be less than 316.5 nm. To further extend the range of pitch measurement, a shorter wavelength is required. A bandpass filter with a 1 nm bandwidth (FL632.8-1, Thorlabs, Inc., Newton, NJ, USA) was employed in the measurement to ensure an accurate result, since the diffracted light associated with different wavelengths within the bandwidth will be undistinguished on the detector for a bandpass filter with a small bandwidth. Additionally, it should be noted that in the measurements of different rotating angles of the sample stage will lead to a minor change of spot size, which will further lead to a change in the number of grating periods covered by the illumination spot in the pitch measurement.

#### 4.3. Measurement of Grating Structural Parameters

To reduce errors caused by sample nonuniformity, the measurements of grating pitch and other structural parameters were carried out on the same grating region. Figure 5 presents the scattered field of the photoresist grating collected at the wavelength of 633 nm, the incidence angles varied from  $0^\circ$  to  $65^\circ$  at interval of  $1^\circ$ , azimuthal angles from  $0^\circ$  to  $90^\circ$  at interval of  $5^\circ$ . The reason why the azimuthal angle was limited to the range of  $0^\circ \sim 90^\circ$  rather than  $0^\circ \sim 360^\circ$  is that the Mueller matrices stay unchanged after a  $180^\circ$  rotation, due to rotation symmetry of the photoresist grating. The grating also has reflection symmetry relative to the plane that is perpendicular to the periodic direction. Since the photoresist grating sample is a periodic structure, the collected scattered field is actually a diffracted field. Due to the limitation of the NA of the objective lens in the TMS, for each illumination direction, only the zeroth-order diffracted light was collected. As shown in Figure 5, the panel associated with each Mueller matrix element consists of three sectorial parts, of which the lower right part corresponds to the TMS-measured data. The polar coordinate system of the sectorial part is set, as shown in the inset of Figure 5, with the radial coordinate being the sine of the incidence angle  $\sin\theta$  and the angular coordinate being the azimuthal angle  $\varphi$ . As can be observed, as the azimuthal angles change, the grating exhibits anisotropy with the  $2 \times 2$  off-diagonal Mueller matrix elements deviating from about 0 at  $\varphi = 0^\circ$ , and finally approaching to about 0 when  $\varphi = 90^\circ$ . We also calculated the depolarization index distribution associated with the TMS-measured data according to  $DI = \sqrt{[\text{Tr}(\mathbf{M}\mathbf{M}^T) - m_{11}^2]/3m_{11}^2}$  ( $0 \leq DI \leq 1$ ) [33], with  $DI = 1$  and  $DI = 0$  representing a totally non-depolarizing and a totally depolarizing Mueller matrix, respectively. The calculated

depolarization indices indicated that  $|DI - 1| < 0.028$ , so the depolarization effect could be ignored in the following data analysis.

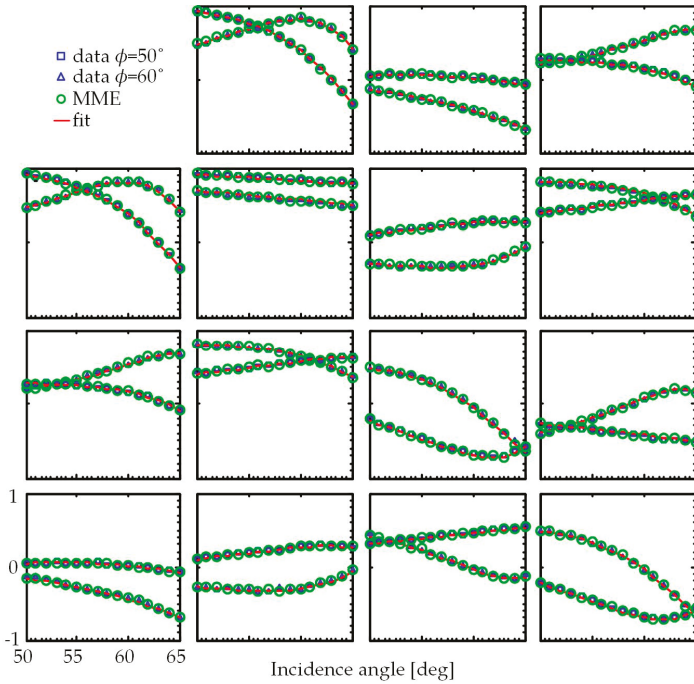


**Figure 5.** Mueller matrix representation of the zeroth-order diffracted field of the photoresist grating collected at the wavelength of 633 nm. The incidence angles varied from  $0^\circ$  to  $65^\circ$  at interval of  $1^\circ$ , and the azimuthal angles from  $0^\circ$  to  $90^\circ$  at interval of  $5^\circ$ . The panel associated with each Mueller matrix element (normalized to  $m_{11}$ ) consists of three sectorial parts. The lower right part corresponds to the TMS-measured data, the upper left part corresponds to the calculated best-fit data, and the lower left part presents the difference between the measured and calculated best-fit data. For each sectorial part, the polar coordinate system is set as shown in the inset with the radial coordinate being the sine of the incidence angle  $\sin\theta$  and the angular coordinate being the azimuthal angle  $\varphi$ .

It should be noted that the angle-resolved Mueller polarimetry [34] can also acquire the diffracted field of a grating sample similar to the TMS. However, it should be pointed out that the measurement modes of the angle-resolved Mueller polarimetry and TMS are inherently different. For the angle-resolved Mueller polarimetry, polarization properties are acquired simultaneously by illuminating a sample simultaneously at all possible incidence angles and imaging the BFP of a high NA objective lens. When the sample is a periodic dense structure, both the TMS and angle-resolved Mueller polarimetry are appropriate for characterizing the sample. However, for a case in which the pitch of a periodic dense structure is greater than  $\lambda/(2 \times NA)$ , or an isolated or non-periodic structure is detected, an overlapping between the polarized scattering information occurs at the BFP of the objective lens in the angle-resolved Mueller polarimetry. In comparison, the TMS overcomes these limits and can distinguish information appropriately.

As shown in Figure 5, the upper left one of the three sectorial parts of each panel presents the calculated best-fit data to the TMS-measured data. To make it clear, the difference between the TMS-measured data and the calculated best-fit data is presented in the lower left sectorial part of each panel. It was found that the difference of most points was less than 0.02. To make a comparison, we also provided the results measured by the commercial MME. As an example, Figure 6 presents the fitting result of the measured and calculated best-fit Mueller matrices of the photoresist grating at the wavelength of 633 nm, the incidence angles varied from  $50^\circ$  to  $65^\circ$  at an interval of  $1^\circ$ , and the azimuthal angles of  $50^\circ$  and  $60^\circ$ . Good agreement can be observed from Figure 6. As can also be observed, the TMS-measured results exhibit good agreement with those measured by MME. Some deviations

between the TMS- and MME-measured data may be induced by measurement errors. Still, the different sizes and locations of illumination spots between TMS and MME may induce an unequal number of grating periods and different areas detected by the systems, bringing about little disparity in the measured results.



**Figure 6.** Fitting result of the measured and calculated best-fit Mueller matrices of the photoresist grating at the wavelength of 633 nm. The incidence angles varied from 50° to 65° at interval of 1°, and the azimuthal angles of 50° and 60°. The red solid lines represent the calculated best-fit data, the blue rectangles and blue triangles represent the TMS-measured data at the azimuthal angles of 50° and 65°, respectively. The green circles represent the corresponding MME-measured data for a comparison.

Table 1 presents the comparison of fitting parameters of the photoresist grating extracted from TMS, MME, and SEM measurements. For TMS measurements, we present both the fitting parameter values before and after the pitch correction. For MME measurements, we only present the fitting parameter values after the pitch correction. As can be seen, after the pitch correction, the TMS measured parameters  $x_1$  and  $x_3$  match well with their nominal values and MME and SEM results. As for the parameter  $x_2$ , the TMS-measured result exhibits a discrepancy with its nominal value and SEM result, but shows good agreement with the MME measured result. The top corner rounding  $x_4$  also matches well with the MME result. This discrepancy is probably due to sample denaturation. We analyzed areas with structures and the assumption was supported by variations in the refractive index and extinction coefficient of the photoresist in comparison with our previous analysis [35]. A changed mixing layer existing between the resist and BARC layers could have introduced the measurement error. In addition, a possible modification of the structural cross-section due to the slicing process in SEM measurement (typically the focused ion beam), as well as the different testing areas of TMS and SEM on the sample, may be other reasons for the discrepancy. Even so, a noticeable improvement in accuracy of the extracted parameters can be recognized from the TMS-measured parameters before and



after the pitch correction. It is thereby necessary to acquire accurate pitch value before the extraction of other grating structural parameters, as it affects the results greatly.

**Table 1.** Comparison of the fitting parameters of the photoresist grating extracted from TMS, MME, and SEM measurements.

Parameter	Nominal Value	TMS		MME	SEM
		Before Correction	After Correction		
$x_1$ (nm)	200	206.1	204.2	204.7	203.4
$x_2$ (nm)	311	343.5	295.8	294.5	303.7
$x_3$ (deg)	90	88.9	89.2	89.7	89.5
$x_4$ (nm)	—	53.3	54.3	52.1	—

## 5. Conclusions

In this work, we have presented the measurement of a photoresist grating by a homemade TMS. We have shown the collection of the polarized scattered-field distributions of the photoresist grating by TMS at different illumination directions. To reconstruct the grating profile parameters, an inverse scattering problem solving method has been adopted to fit the measured data with a RCWA model. The reconstructed grating profile parameters have exhibited good agreement with those measured by a commercial MME and SEM. In addition, the comparison between the TMS-measured results with and without pitch correction has also demonstrated the capability of TMS for determining both the grating pitch and other structural parameters, which is impossible with current zeroth-order methods such as reflectometry- or ellipsometry-based scatterometry, due to the strong correlation between pitch and other structural parameters in the solution of the inverse scattering problem.

**Author Contributions:** Conceptualization, X.C.; methodology and formal analysis, C.C. and X.C.; writing—original draft preparation, C.C.; writing—review and editing, X.C., Y.S., H.G., H.J., and S.L.; supervision, X.C. and S.L.; project administration, X.C. and S.L.; funding acquisition, X.C. and S.L.

**Funding:** This research was funded by the National Natural Science Foundation of China (Grant Nos. 51727809, 51775217, 51525502, and 51475191), the National Science and Technology Major Project of China (Grant No. 2017ZX02101006-004), and the National Science Foundation of Hubei Province of China (Grant Nos. 2018CFB559 and 2018CFA057).

**Acknowledgments:** The authors thank Shanghai Micro Electronics Equipment Co., Ltd. (Shanghai, China) for preparing the grating sample.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

- Fang, F.Z.; Zhang, X.D.; Gao, W.; Guo, Y.B.; Byrne, G.; Hansen, H.N. Nanomanufacturing—Perspective and applications. *CIRP Ann. Manuf. Technol.* **2017**, *66*, 683–705. [[CrossRef](#)]
- Postek, M.T.; Lyons, K. Instrumentation, metrology, and standards: Key elements for the future of nanomanufacturing. *Proc. SPIE* **2007**, *6648*, 664802.
- Raymond, C.J. Scatterometry for semiconductor metrology. In *Handbook of Silicon Semiconductor Metrology*; Diebold, A.C., Ed.; CRC Press: Boca Raton, FL, USA, 2001; Chapter 18.
- Huang, H.T.; Kong, W.; Terry, F.L., Jr. Normal-incidence spectroscopic ellipsometry for critical dimension monitoring. *Appl. Phys. Lett.* **2001**, *78*, 3983–3985. [[CrossRef](#)]
- Wurm, M.; Endres, J.; Probst, J.; Schoengen, M.; Diener, A.; Bodermann, B. Metrology of nanoscale grating structures by UV scatterometry. *Opt. Express* **2017**, *25*, 2460–2468. [[CrossRef](#)] [[PubMed](#)]
- Novikova, T.; De Martino, A.; Bulkin, P.; Nguyen, Q.; Drévilion, B.; Popov, V.; Chumakov, A. Metrology of replicated diffractive optics with Mueller polarimetry in conical diffraction. *Opt. Express* **2007**, *15*, 2033–2046. [[CrossRef](#)] [[PubMed](#)]
- Kim, Y.N.; Paek, J.S.; Rabello, S.; Lee, S.; Hu, J.; Liu, Z.; Hao, Y.; McGahan, W. Device based in-chip critical dimension and overlay metrology. *Opt. Express* **2009**, *17*, 21336–21343. [[CrossRef](#)] [[PubMed](#)]

8. Liu, S.; Chen, X.; Zhang, C. Development of a broadband Mueller matrix ellipsometer as a powerful tool for nanostructure metrology. *Thin Solid Films* **2015**, *584*, 176–185. [[CrossRef](#)]
9. Diebold, A.C.; Antonelli, A.; Keller, N. Perspective: Perspective: Optical measurement of feature dimensions and shapes by scatterometry. *APL Mater.* **2018**, *6*, 058201. [[CrossRef](#)]
10. Boundary, B.; Solecky, E.; Vaid, A.; Bello, A.F.; Dai, X. Metrology capabilities and needs for 7 nm and 5 nm logic nodes. *Proc. SPIE* **2017**, *10145*, 101450G.
11. Orji, N.G.; Badaroglu, M.; Barnes, B.M.; Beitia, C.; Bunday, B.D.; Celano, U.; Kline, R.J.; Neisser, M.; Obeng, Y.; Vladar, A.E. Metrology for the next generation of semiconductor devices. *Nat. Electron.* **2018**, *1*, 532–547. [[CrossRef](#)]
12. Germer, T.A.; Asmail, C.C. Goniometric optical scatter instrument for out-of-plane ellipsometry measurements. *Rev. Sci. Instrum.* **1999**, *70*, 3688–3695. [[CrossRef](#)]
13. Liu, C.Y.; Liu, T.A.; Fu, W.E. Out-of-plane ellipsometry measurements of nanoparticles on surfaces for thin film coated wafer inspection. *Opt. Laser Technol.* **2010**, *42*, 902–910. [[CrossRef](#)]
14. Maria, J.; Aas, L.M.S.; Kildemo, M. In and out of incidence plane Mueller matrix scattering ellipsometry of rough mc-Si. *Thin Solid Films* **2014**, *571*, 399–404. [[CrossRef](#)]
15. Attota, R.; Germer, T.A.; Silver, R.M. Through-focus scanning-optical-microscope imaging method for nanoscale dimensional analysis. *Opt. Lett.* **2008**, *33*, 1990–1992. [[CrossRef](#)] [[PubMed](#)]
16. Silver, R.M.; Barnes, B.M.; Attota, R.; Jun, J.; Stocker, M.; Marx, E.; Patrick, H.J. Scatterfield microscopy for extending the limits of image-based optical metrology. *Appl. Opt.* **2007**, *46*, 4248–4257. [[CrossRef](#)] [[PubMed](#)]
17. Maire, G.; Drsek, F.; Girard, J.; Giovannini, H.; Talneau, A.; Konan, D.; Belkebir, K.; Chaumet, P.C.; Sentenac, A. Experimental demonstration of quantitative imaging beyond Abbe’s limit with optical diffraction tomography. *Phys. Rev. Lett.* **2009**, *102*, 213905. [[CrossRef](#)] [[PubMed](#)]
18. Zhang, T.; Ruan, Y.; Maire, G.; Sentenac, D.; Talneau, A.; Belkebir, K.; Chaumet, P.C.; Sentenac, A. Full-polarized tomographic diffraction microscopy achieves a resolution about one-fourth of the wavelength. *Phys. Rev. Lett.* **2013**, *111*, 243904. [[CrossRef](#)] [[PubMed](#)]
19. Boher, P.; Petit, J.; Leroux, T.; Foucher, J.; Desieres, Y.; Hazart, J.; Chaton, P. Optical Fourier transform scatterometry for LER and LWR metrology. *Proc. SPIE* **2005**, *5752*, 594526.
20. Petrik, P.; Kumar, N.; Fried, M.; Fodor, B.; Urbach, H.P. Fourier ellipsometry—An ellipsometric approach to Fourier scatterometry. *J. Eur. Opt. Soc. Rapid* **2015**, *10*, 15002. [[CrossRef](#)]
21. Tan, Y.; Chen, C.; Chen, X.; Du, W.; Gu, H.; Liu, S. Development of a tomographic Mueller-matrix scatterometer for nanostructure metrology. *Rev. Sci. Instrum.* **2018**, *89*, 073702. [[CrossRef](#)]
22. Li, J.; Ramanujam, B.; Collins, R.W. Dual rotating compensator ellipsometry: Theory and simulations. *Thin Solid Films* **2011**, *519*, 2725–2729. [[CrossRef](#)]
23. Moharam, M.G.; Grann, E.B.; Pommet, D.A.; Gaylord, T.K. Formulation for stable and efficient implementation of the rigorous coupled-wave analysis of binary gratings. *J. Opt. Soc. Am. A* **1995**, *12*, 1068–1076. [[CrossRef](#)]
24. Li, L. Use of Fourier series in the analysis of discontinuous periodic structures. *J. Opt. Soc. Am. A* **1996**, *13*, 1870–1876. [[CrossRef](#)]
25. Chen, X.; Zhang, C.; Liu, S. Depolarization effects from nanoimprinted grating structures as measured by Mueller matrix polarimetry. *Appl. Phys. Lett.* **2013**, *103*, 151605. [[CrossRef](#)]
26. Press, W.H.; Teukolsky, S.A.; Vetterling, W.T.; Flannery, B.P. *Numerical Recipes: The Art of Scientific Computing*, 3rd ed.; Cambridge University Press: Cambridge, UK, 2007; Chapter 15.
27. Chernoff, D.A.; Buhr, E.; Burkhead, D.L.; Diener, A. Picometer-scale accuracy in pitch metrology by optical diffraction and atomic force microscopy. *Proc. SPIE* **2008**, *6922*, 69223J.
28. Herzinger, C.M.; Johs, B.; McGahan, W.A.; Woollam, J.A.; Paulson, W. Ellipsometric determination of optical constants for silicon and thermally grown silicon dioxide via a multi-sample, multi-wavelength, multi-angle investigation. *J. Appl. Phys.* **1998**, *83*, 3323–3336. [[CrossRef](#)]
29. Jellison, G.E., Jr.; Modine, F.A. Parameterization of the optical functions of amorphous materials in the interband region. *Appl. Phys. Lett.* **1996**, *69*, 371–373. [[CrossRef](#)]
30. Meneses, D.D.S.; Malki, M.; Echegut, P. Structure and lattice dynamics of binary lead silicate glasses investigated by infrared spectroscopy. *J. Non-Cryst. Solids* **2006**, *352*, 769–776. [[CrossRef](#)]



31. Ferlauto, A.S.; Ferreira, G.M.; Pearce, J.M.; Wronski, C.R.; Collins, R.W.; Deng, X.; Ganguly, G. Analytical model for the optical functions of amorphous semiconductors from the near-infrared to ultraviolet: Applications in thin film photovoltaics. *J. Appl. Phys.* **2002**, *92*, 2424–2436. [[CrossRef](#)]
32. Synowicki, R.A. Spectroscopic ellipsometry characterization of indium tin oxide film microstructure and optical constants. *Thin Solid Films* **1998**, *313*, 394–397. [[CrossRef](#)]
33. Gil, J.J.; Bernabeu, E. Depolarization and polarization indices of an optical system. *Opt. Acta Int. J. Opt.* **1986**, *33*, 185–189. [[CrossRef](#)]
34. Ben Hatit, S.; Foldyna, M.; De Martino, A.; Drévilion, B. Angle-resolved Mueller polarimeter using a microscope objective. *Phys. Status Solidi A* **2008**, *205*, 743–747. [[CrossRef](#)]
35. Chen, X.; Du, W.; Yuan, K.; Chen, J.; Jiang, H.; Zhang, C.; Liu, S. Development of a spectroscopic Mueller matrix imaging ellipsometer for nanostructure metrology. *Rev. Sci. Instrum.* **2016**, *87*, 053707. [[CrossRef](#)] [[PubMed](#)]



© 2018 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).

Article

# Real-Time Tunnel Deformation Monitoring Technology Based on Laser and Machine Vision

Zurong Qiu , Haopeng Li , Wenchuan Hu, Chenglin Wang, Jiachen Liu and Qianhui Sun

State Key Laboratory of Precision Measuring Technology and Instruments, School of Precision Instruments and Opto-electronics Engineering, Tianjin University, Tianjin 300072, China; lhp19911119@126.com (H.L.); wenchuan\_tju@outlook.com (W.H.); me\_wangcl@tju.edu.cn (C.W.); liujiachen@zju.edu.cn (J.L.); sunqianhui@tju.edu.cn (Q.S.)

\* Correspondence: qiuzr@tju.edu.cn

Received: 23 October 2018; Accepted: 7 December 2018; Published: 11 December 2018



**Featured Application:** The proposed tunnel deformation monitoring system was designed for tunnel during construction in dusty and dark environment with low visibility, and it also can be applied for the tunnel that has already been built.

**Abstract:** Structural health monitoring is a topic of great concern in the world, and tunnel deformation monitoring is one of the important tasks. With the rapid developments in tunnel traffic infrastructure construction, engineers need a portable and real-time system to obtain the tunnel deformation during construction. This paper reports a novel method based on laser and machine vision to automatically measure tunnel deformation of multiple interest points in real time and effectively compensate for the environment vibration, and moreover it can overcome the influence of a dusty and dark tunnel environment in low visibility. An automatic and wireless real-time tunnel deformation monitoring system, which is based on laser and machine vision and can give early warnings for tunnel collapse accidents, is proposed. The proposed system uses a fixed laser beam as a monitoring reference. The image acquisition modules mounted on the measured points receive the laser spots and measure the tunnel accumulative deformation and instantaneous deformation velocity. Compensation methods are proposed to reduce measurement errors caused by laser beam feasibility, temperature, air refraction index, and wireless antenna attitude. The feasibility of the system is verified through tunnel tests. The accuracy of the detection system is better than 0.12 mm, the repeatability is less than 0.11 mm, and the minimum resolution is 10  $\mu\text{m}$ ; therefore, the proposed system is very suitable for real-time and automatic detection of tunnel deformation in low visibility during construction.

**Keywords:** structural health monitoring; real-time monitoring; tunnel deformation measurement; machine vision; laser beam; wireless; low visibility

---

## 1. Introduction

Structural health monitoring plays an indispensable role in diagnosing structure safety around the world [1], and tunnel deformation monitoring is one of the important tasks [2]. With the continuous expansion of the scale of tunnel traffic infrastructure construction, tunnel collapse accidents have occurred frequently in recent years [3]; this indicates that the traditional methods for monitoring tunnel deformation are unable to meet the new safety monitoring requirements for large tunnel traffic infrastructure [4], when real time becomes an important issue. In tunnel construction, the tunnel excavation surface is composed of tunnel face, side face, and roof face. Numerous workers are present on the excavation surface with dusty and dark environment in low visibility [5], and the tunnel

deformation is at a maximum; thus, monitoring the deformation of this fracture surface in real time is important. The current tunnel deformation monitoring technology, including non-machine vision measurement methods and machine vision measurement methods [6], however, is not effective enough. Most of the existing equipment is manually operated and susceptible to a dusty and dark environment in low visibility, and cannot meet real-time automatic monitoring and early warning requirements or provide timely warnings about dangerous deformation [7].

Non-machine vision measurement methods are widely used in tunnel deformation monitoring; however, they are susceptible to environmental and human factors, and take a long time. Tape extensometers and convergence gauge are unstable and operated manually; in addition, they are strongly influenced by environmental and human factors [8]. The leveling method mainly addresses the tunnel settlement value [9] and requires many monitoring points to be set; thus, the error accumulates with the increase in the tunnel length. The indoor global positioning system measurement method requires many reference signal transmitters; however, it is difficult to ensure the stability of the reference signal transmitters position. The traditional total station measurement method has a low degree of automation, exhibits strong subjectivity, and requires a fixed station; these features affect the tunnel construction period [10]. The new-generation total station measurement method allows stations to be freely monitored, compared with the traditional method [11,12], but it cannot be used to monitor all kinds of tunnels because of its high cost [13]. The fiber Bragg grating measurement method [14] can be used to monitor tunnel deformation constantly over a long period of time, but it is expensive, and the relevant equipment is easily damaged. Moreover, the equipment needs to be embedded into the structure or adhered onto the structure surface during monitoring, which results in a higher demand on the environment. A deformation monitoring system based on ultrasonic sensors [15] is automatic and in real time; however, it is only suitable for small- to middle-size tunnels, and it cannot be used to realize early warnings of dangerous deformation during the construction. Terrestrial laser scanning [16] achieves tunnel deformation monitoring; however, it takes a long time to obtain and process point cloud data, and cannot realize real-time monitoring.

A machine vision measurement method for tunnel deformation develops a lot in the last few years [17]; however, existing machine vision measurement methods are not suitable for real-time tunnel deformation monitoring during construction with dusty and dark environments. A three-dimensional reconstruction method based on machine vision [18] needs to take lots of photos and the data processing algorithms are complex; this takes a long time and cannot meet real-time automatic monitoring and early warning requirements. A digital close-range photogrammetric method [19] is susceptible to tunnel dust and light intensity, heavy dust and a dark environment during construction reduces photo quality and detection accuracy. A photogrammetric method for tunnel detection based on a CCD camera [20] has a higher requirement for brightness and lighting equipment; this requires external power supply; when using batteries, the equipment is heavy and not easy to install; when using an electric generator, extra great vibration is produced and this affects the result. The infrared photogrammetric method [21] reduces dust impact at the price of decreasing photo brightness; as a result, the detection accuracy is low and cannot be used for real-time monitoring. Photogrammetric techniques for monitoring tunnel deformation [22] are only suitable for tunnels without dust after excavation, and it is also susceptible to dusty and dark environments. A tunnel detection method combining photogrammetry and laser scanning [23] is susceptible to the set position of the equipment; when the equipment is set deviating from the tunnel axis, the detection error increases obviously [24]; this indicates that the method is not suitable for a tunnel during construction, since the set position has a conflict with tunnel construction.

This paper presents a real-time and automatic tunnel deformation monitoring system to realize real-time monitoring and early warning for dangerous tunnel deformation. Considering the dusty and dark tunneling environment, a fixed laser beam is used to transmit the monitoring reference from a stable point to a measured point through the dust and darkness in low visibility, and a shielded image acquisition module mounted on the measured points is used to acquire the laser spot image regardless

of the dust and darkness. However, because of the adverse tunnel environment, the measurement result is affected by laser beam feasibility, temperature, air refraction index, and wireless antenna attitude. These factors are carefully examined in this study, and compensation methods are proposed to reduce the measurement errors caused by them. The results show that the proposed system not only achieves good detection accuracy, repeatability and resolution, but also compensates for the shortcomings of the above methods, such as the low degree of automation, the lack of real-time measurement function and overcoming the influence of dusty and dark environments in low visibility.

## 2. Methodology

The tunnel collapse process is the deformation of the construction fracture surface in the tunnel, which takes a fixed point of the earth as a reference. At present, the new Austrian tunneling method [25], which includes designing an initial support and secondary lining to the construction fracture surface [26], is mostly adopted in tunnel construction. The secondary lining is constructed on the basis of the initial support when the deformation of the initial support becomes stable [27]. Therefore, the deformation of the secondary lining surface can be ignored; however, the deformation of the excavation surface and the edge part of the initial support is significant. The proposed system takes a fixed point of the earth as monitoring reference and measures the deformation of excavation surface. Given the good directivity and energy concentration of a laser, as well as the dusty and dark tunneling environment in low visibility, a laser beam is used to transmit the monitoring reference from the fixed point to the measured point through the dust and darkness, thereby forming an optical projection of the fixed point on the intersecting surface near the measured point, which is the optical equivalent reference, to monitor the tunnel deformation in the system. A sketch map of the system is shown in Figure 1.

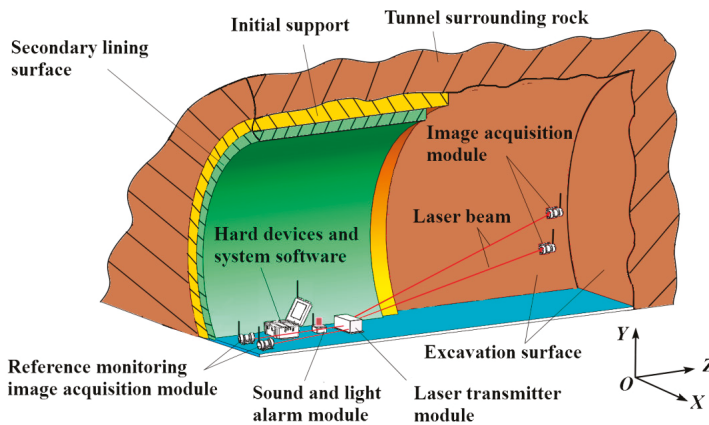


Figure 1. System principle.

A laser transmitter module, whose laser beam points to the intersecting surface near the measured points and forms an optical projection of the fixed point as a laser spot, is firmly fixed on the earth. To obtain the position relationship between the measured point and the fixed point, an image acquisition module mounted on the measured point and located perpendicularly to the laser beam is used to acquire the laser spot image of the laser beam. This module monitors the laser spot position in the image coordinate system with machine vision technology. The image acquisition mode is shown in Figure 2. The image sensor acquires the laser spot image through a laser receiving screen placed in front of the sensor.

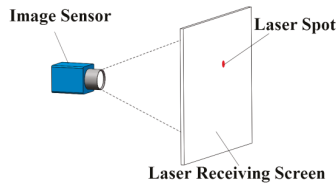


Figure 2. Image acquisition module.

The image acquisition module is fixed on the measured points. When the excavation surface sinks during the deformation, the image acquisition module will sink along with it, thereby changing the laser spot position in the image coordinate system. The change of the laser spot position in the image coordinate system reflects the deformation of the excavation surface, including accumulative deformation and instantaneous deformation velocity [28].

The mathematical model of this principle is shown in Figure 3. The geometric center of the laser is the coordinate origin, the tunnel driving direction is the  $z$ -axis, the horizontal direction is the  $x$ -axis, and the vertical axis is the  $y$ -axis. Therefore, a three-dimensional world coordinate system  $o$ - $xyz$  is established. The laser beam points to the laser receiving screen and forms a laser spot  $P$  on the screen.

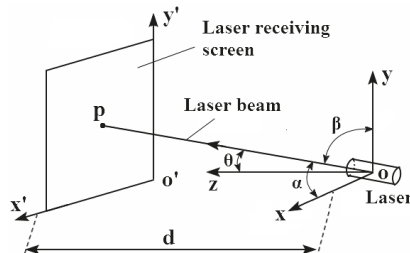


Figure 3. Mathematical model.

Considering the horizontal and vertical position information, the coordinates of the spot  $P$  in the world coordinate system  $o$ - $xyz$  are  $(x, y)$ . To examine the position information of laser spot  $P$  in the image coordinate system, an image coordinate system on the laser receiving screen is established, as shown in Figure 3. The  $o'x'$  axis is parallel to the  $ox$  axis, and the  $o'y'$  axis is parallel to the  $oy$  axis; the coordinates of laser spot  $P$  in the image coordinate system are  $(x', y')$ . The coordinate transformation of spot  $P$  in the two coordinate systems is shown in Equation (1). The constants  $c_0$  and  $c_1$  are coordinate conversion parameters, which are related to the relative position of the two coordinate systems:

$$\begin{cases} x' = x + c_0 \\ y' = y + c_1 \end{cases} \quad (1)$$

In the three-dimensional world coordinate system  $o$ - $xyz$ , the laser beam equation is shown as Equation (2), and the equation of the intersecting surface near the measured point is shown as Equation (3), where  $A, B, C$ , and  $d$  are constant parameters and are related to the laser position:

$$A \times x = B \times y = C \times z, \quad (2)$$

$$z = d. \quad (3)$$

According to Equations (1)–(3), the coordinates of the laser spot in the image coordinate system  $o'-x'y'$  can be obtained as shown in Equation (4):

$$\begin{cases} x' = \frac{C \times d}{A} + c_0 \\ y' = \frac{C \times d}{B} + c_1 \end{cases} \quad (4)$$

According to Equation (4), when the positions of laser and image acquisition module are invariant, that is, when the parameters  $A$ ,  $B$ ,  $C$ , and  $d$  are invariant, the coordinates of the laser spot in the image coordinate system  $\left(\frac{C \times d}{A} + c_0, \frac{C \times d}{B} + c_1\right)$  remain fixed. Therefore, the monitoring reference can be transmitted from the fixed point of the earth to the measured points. When the image acquisition module sinks during tunnel deformation, the image coordinate system moves relative to the world coordinate system  $o-xyz$ . Supposing that the vertical and horizontal deformation amounts of the image acquisition module are  $\Delta x$  and  $\Delta y$ , respectively, the laser spot position in the image coordinate system also shifts  $\Delta x$  and  $\Delta y$  in the vertical and horizontal directions, respectively. Therefore, when the image acquisition module shifts in the vertical and horizontal directions along with the measured point, the change of the laser spot position in the image coordinate system reflects the deformation.

In an actual situation, the laser beam is not absolutely perpendicular to the laser receiving screen. As a result, when the image acquisition module moves along the  $z$ -axis with the measured point, the laser spot will produce the displacement component in the vertical and horizontal directions. As shown in Figure 3,  $\alpha$  is the angle between the laser beam and the  $ox$  axis,  $\beta$  is the angle between the laser beam and the  $oy$  axis, and  $\theta$  is the angle between the laser beam and the  $oz$  axis. Assuming that the displacement change of image acquisition module along the  $oz$  axis is  $\Delta z$ , the errors of the horizontal and vertical deformation measurements are as shown in Equation (5):

$$\begin{cases} \Delta x = \Delta z \times \frac{\cos\alpha}{\cos\theta} \\ \Delta y = \Delta z \times \frac{\cos\beta}{\cos\theta} \end{cases} \quad (5)$$

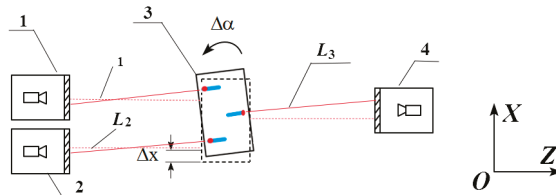
Assuming that the laser beam points obliquely to the receiving screen, the angles  $\alpha = 73.3^\circ$ ,  $\beta = 81.5^\circ$ , and  $\theta = 18.9^\circ$  (in fact, the tilt of the laser beam is not that particularly obvious). When the image acquisition module shifts 1 mm along the  $oz$  axis,  $\Delta x = 0.3037$  mm and  $\Delta y = 0.1562$  mm. Therefore, when the shifting displacement of the image acquisition module along the  $z$ -axis is not large (and in fact it is not large) [29], this error can be ignored.

The laser transmitter module is mounted on the fixed point of the earth distant from the construction site, and it should remain stable. However, in cases of emergency during construction, such as sudden vibrating, when the laser beam reference is shifted and rotated, the position of the laser spot in the image coordinate will change. Taking the  $x$ -axis coordinate as an example, if the displacements along the  $x$ -axis and the horizontal rotation angle of the laser beam reference are  $\Delta x$  and  $\Delta\alpha$ , respectively, the resulting laser spot position error  $\Delta x'$  is as shown in Equation (6):

$$\Delta x' = \Delta x + L \times \tan\Delta\alpha = \Delta x + L \times \Delta\alpha. \quad (6)$$

In order to reduce this error, a data compensation method is proposed here. The data compensation method uses two backward lasers in the laser transmitter module and two fixed distant reference monitoring image acquisition modules with higher stability to monitor the feasibility of the laser transmitter, as shown in Figure 4, where  $L_s$  represents the distance between the laser transmitter module and the image acquisition modules,  $\Delta\alpha$  represents the horizontal rotation angle of the laser transmitter module, and  $\Delta x$  represents the displacement along the  $x$ -axis of the laser transmitter module.





**Figure 4.** Sketch map. 1. reference monitoring image acquisition module A; 2. reference monitoring image acquisition module B; 3. laser transmitter module; 4. image acquisition module.

Given that  $L_s$  were dozens of meters, small displacement and rotation angle will have no effect on these  $L_s$ ; that is, the lengths of  $L_s$  were invariable before and after the position change of the laser transmitter module. With the two reference monitoring image acquisition modules, two laser spot position errors  $\Delta x'_1$  and  $\Delta x'_2$  can be obtained. The image acquisition module detecting displacement with error is  $\Delta x'_3$ . The equations of two variables  $\Delta x$  and  $\Delta \alpha$  were obtained, as shown in Equation (7):

$$\begin{cases} \Delta x'_1 = \Delta x + L_1 \times \Delta \alpha \\ \Delta x'_2 = \Delta x + L_2 \times \Delta \alpha \end{cases} \quad (7)$$

According to Equation (10),  $\Delta x$  and  $\Delta \alpha$  can be determined as shown in Equation (8):

$$\begin{cases} \Delta x = \Delta x'_1 - L_1 \times \frac{\Delta x'_1 - \Delta x'_2}{L_1 - L_2} \\ \Delta \alpha = \frac{\Delta x'_1 - \Delta x'_2}{L_1 - L_2} \end{cases} \quad (8)$$

Then, the detected displacement can be compensated in Equation (9):

$$\Delta x_3 = \Delta x'_3 - \left[ \left( \Delta x'_1 - L_1 \times \frac{\Delta x'_1 - \Delta x'_2}{L_1 - L_2} \right) + L_3 \times \frac{\Delta x'_1 - \Delta x'_2}{L_1 - L_2} \right] \quad (9)$$

Therefore, the detected displacement error caused by sudden vibration is compensated successfully.

### 3. System Composition

Figure 5 shows the main composition of the real-time tunnel deformation monitoring system, and Figure 6 shows the details. The working principle is briefly described as follows.

Install the system in the tunnel and turn it on. After calibration, adjust the direction of laser to the image acquisition module. The command to measure tunnel deformation is first generated at the industrial computer and sent to the wireless adaptor by a USB port. The communication between the adaptor and the image acquisition module is wireless. The wireless adaptor is shown in Figure 7. The image acquisition module receives the command and begins to capture laser spot images and sends them to the industrial computer through the corresponding wireless adaptor in real time. The software installed in the industrial computer receives the images and processes them, obtaining the tunnel accumulative deformation and instantaneous deformation velocity. If the deformation values exceed the safe threshold, another command is generated and sent to the sound and light alarm module and the command center through a wireless adaptor, giving an early warning to workers and managers.



Figure 5. Main composition of the real-time tunnel deformation monitoring system.

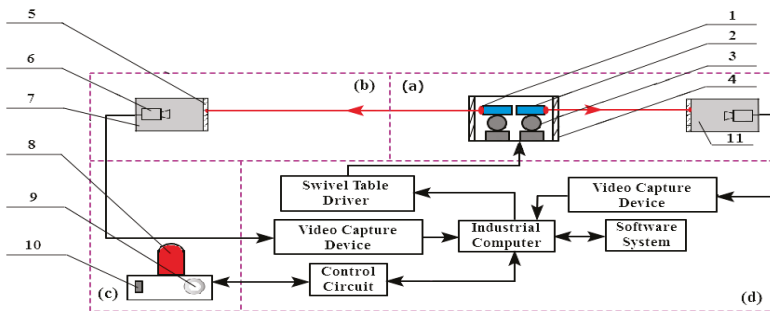


Figure 6. System composition: (a) laser transmitter module; (b) image acquisition module; (c) sound and light alarm module; (d) hard devices and system software. 1. forward laser; 2. backward laser; 3. electronically controlled two-dimensional swivel table; 4. organic glass window; 5. filter imaging screen; 6. high-definition industrial camera; 7. darkroom; 8. alarm light; 9. alarm horn; 10. power switch; 11. reference monitoring image acquisition module.



Figure 7. Wireless adaptor.

### 3.1. Laser Transmitter Module

All parts of this module, except the reference monitoring image acquisition module, are rigidly connected, thereby preventing the deformation of the module. The organic glass window and the dirt shroud form a sealed space, which can effectively prevent the interference of tunnel dust to the module and guarantee the permeability of the laser beam. The electronically controlled two-dimensional swivel table can adjust the horizontal and vertical angles of the laser beam, making the laser beam to point to the intersecting surface near the measured points. The reference monitoring image acquisition module is behind and distant from the other parts of the laser transmitter module, and it is mounted on a more stable fixed point away from the construction site, monitoring the feasibility of the forward laser beam.

### 3.2. Image Acquisition Module

A band-pass filter film [30] matching with the laser frequency is coated on the exterior surface of the filter imaging screen [31]; this allows the laser beam to go through the screen. A thin layer of diffuse reflection coating is present on the interior surface of the filter imaging screen [32]. On this layer, the laser beam forms a laser spot with Lambertian reflection. The HD industrial camera constantly takes photos of the laser spot, extracting the position information of the laser spot in the image coordinate system with machine vision. The band-pass filter film and darkroom [33] helps to reduce the stray light interference.

### 3.3. Sound and Light Alarm Module

When the tunnel accumulative deformation and instantaneous deformation velocity exceeds the set threshold, the module is turned on and provides an early warning to the command center through a sound and light alarm; this gives a timely and effective early warning against tunnel collapse accidents, thereby reducing casualties and property losses.

### 3.4. Hard Devices and System Software

The hard devices are used to realize the communication between each module and install the system software. The main functions of the software for real-time tunnel deformation monitoring system include measurement, control, communication, and data storage. Considering the multi-system operation, cross platform, and compatibility of the system software requirements, LabVIEW [34] 2010 by NI (Austin, TX, USA) is adopted as the software development platform. The communication between the laser transmitter module, image acquisition module, sound and light alarm module, and hard devices and system software is wireless. The flow chart of the system software is shown as Figure 8.

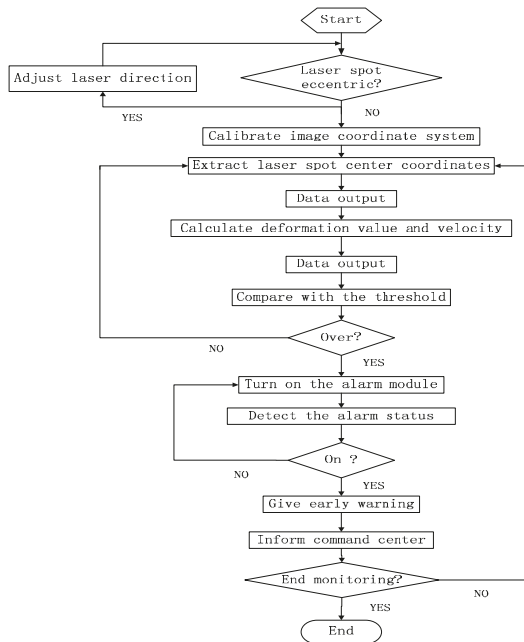
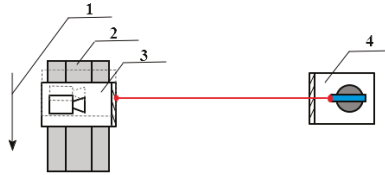


Figure 8. Flowchart of the system software.

## 4. Test Results and Analysis

### 4.1. Principle Verification Test

To verify the principle of the proposed tunnel deformation monitoring technology, a tunnel deformation process was designed and simulated as shown in Figure 9.

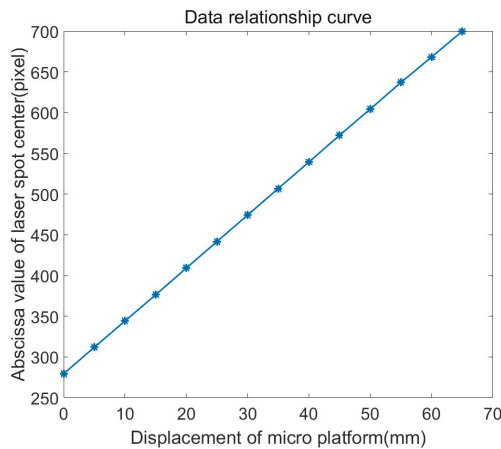


**Figure 9.** Sketch map of the principle verification experiment. 1. moving direction; 2. micro motion platform; 3. image acquisition module; 4. laser transmitter module.

The image acquisition module was placed on a high-precision micro motion platform, which is a one-dimensional motion platform with a 1  $\mu\text{m}$  accuracy to enable the image acquisition module to move with the moving platform and simulate the deformation process of the tunnel. The laser transmitter module was arranged on the other side, with the laser pointing to the filter imaging screen. If the change in the coordinates of the laser spot center [35] in the image coordinate system is proportional to the tunnel deformation, then the correctness of the tunnel deformation monitoring principle is proved.

In the experiment, the micro motion platform was moved at an interval of 5 mm, and the displacement of the micro motion platform and the corresponding abscissa values of the laser spot center in the image coordinate system in pixel units were recorded.

The data relation curve is shown in Figure 10. In this curve, the  $x$ -axis is the displacement of the micro motion platform, and the  $y$ -axis is the measured displacement, namely the abscissa value of the laser spot center.



**Figure 10.** Data relation curve.

According to the data relation curve, the displacement of the micro motion platform was linearly correlated to the measured displacement, and the linear correlation coefficient was close to 1, which demonstrates the correctness of the principle.

#### 4.2. Image Acquisition Module Calibration Experiment

The position information obtained by the image acquisition module is the plane two-dimensional data in the image coordinate system in pixels. Thus, the image coordinate system needs to be calibrated to restore the true deformation value of the measured point [36].

##### 4.2.1. Horizontal Calibration Experiment

Given the different positions of the measured points in the tunnel, the image acquisition module will often roll along the  $z$ -axis after it is installed. As a result, an angle will form between the  $x'$ -axis of the image coordinate system and the  $x$ -axis of the world coordinate system. Given this angle, the change in the position information of the laser spot center obtained by the image acquisition module cannot directly represent the settlement and the diameter convergence of the measured point. Therefore, the horizontality of the image coordinate system needs to be calibrated.

A simplified model for the relationship between the world coordinate system and the image coordinate system is shown in Figure 11 [37]. Here,  $o-xy$  represents the world coordinate system,  $o'-x'y'$  represents the image coordinate system, and  $\theta$  represents the rotation angle. The line  $l$  is a horizontal line, which is parallel to the  $ox$  axis and comes from a laser cast instrument. The laser cast instrument produced a horizontal laser line with a  $\pm 10''$  accuracy. Supposing that the slope of the line  $l$  in the image coordinate system is  $k$ , the relationship between  $\theta$  and  $k$  is as shown in Equation (10):

$$\theta = 180^\circ - \tan^{-1}k. \tag{10}$$

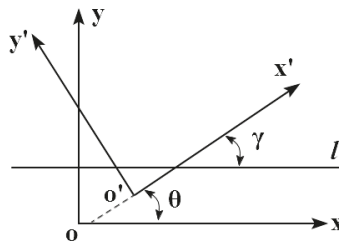


Figure 11. Simplified coordinate model.

After the rotation angle  $\theta$  was obtained, the horizontal calibration of the image coordinate system could be completed by rotating the image coordinate system by an angle  $\theta$  around point  $o'$ , as the rotation center in the reverse direction. The horizontal calibration process is shown in Figure 12.

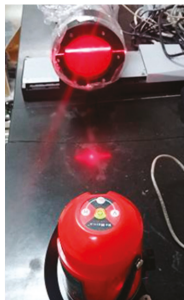


Figure 12. Horizontal calibration experiment.

Before the horizontal calibration, when the image acquisition module horizontally moved along the micro motion platform, as shown in Figure 9, both the  $x$ -axis and  $y$ -axis values of the laser spot

center in the image coordinate system changed with the movement. After the horizontal calibration was completed, only the  $x$ -axis value changed, thereby indicating that the horizontal calibration of the image coordinate system was successful. The micro motion platform moved at an interval of 5 mm. The experimental data curves are shown in Figure 13.

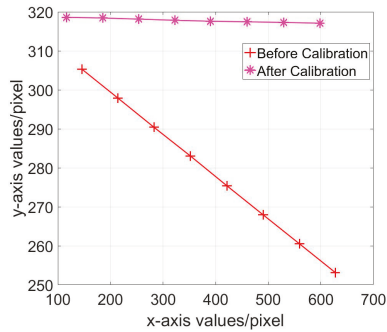


Figure 13. Coordinate change before and after horizontal calibration.

According to Figure 13, the displacement of the micro motion platform was from 0 mm to 70 mm, and the change in the  $y$ -axis value after calibration was as follows:

$$\Delta Y = 318.52 - 317.01 = 1.51 \text{ (pixel)}.$$

Therefore, within the permitted error, the  $y$ -axis values were basically unchanged, and the horizontal calibration of the image coordinate system was successful.

In the tunnel environment, a universal arm was used to hold the laser cast instrument during the horizontal calibration, as shown in Figures 14 and 15, and the calibration method was the same. Every time before using the system, the horizontal calibration was required to be done first.



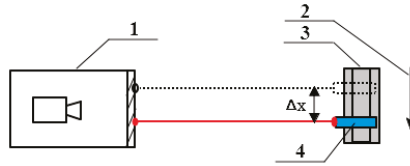
Figure 14. Universal arm.



Figure 15. Tunnel environment.

#### 4.2.2. Conversion Coefficient Calibration Experiment

The unit of spot center coordinates in the image coordinate system is pixel, which cannot represent the true displacement directly [38]. A sketch map of the conversion coefficient calibration experiment is shown in Figure 16. The laser was placed on the micro motion platform and pointed to the fixed image acquisition module, thereby allowing it to move with the micro motion platform.



**Figure 16.** Conversion coefficient calibration experiment. 1. image acquisition module; 2. movement direction; 3. micro motion platform; 4. laser.

To restore the true displacement of the micro motion platform, the horizontal coordinate conversion coefficient  $\alpha$  and the vertical coordinate conversion coefficient  $\beta$  needed to be determined [39]. The calculation formulas are shown in Equation (11):

$$\begin{cases} \Delta x = \alpha \times \Delta X \\ \Delta y = \beta \times \Delta Y \end{cases} \quad (11)$$

In Equation (11),  $(\Delta X, \Delta Y)$  represents the displacement of the laser spot in the image coordinate system in pixels, and  $(\Delta x, \Delta y)$  represents the real displacement in the world coordinate system. The  $x$ -axis conversion coefficient calibration experimental data are shown in Table 1.

**Table 1.**  $x$ -axis conversion coefficient calibration experimental data.

Displacement of Micro Motion Platform (mm)	Abscissa Value of Laser Spot Center (pixel)	$\Delta x$ (mm)	$\Delta X$ (pixel)
0	636.39	—	—
10	567.28	10	69.11
20	498.04	10	69.24
30	428.44	10	69.60
40	360.53	10	67.91
50	291.98	10	68.55
60	223.58	10	68.40
70	154.93	10	68.65

The  $x$ -axis conversion coefficient can be obtained with Equation (11), and the arithmetic mean value formula is shown as Equation (12):

$$\alpha = \frac{\sum_{i=1}^7 \frac{\Delta x_i}{\Delta X_i}}{7} = 0.1454. \quad (12)$$

The calibration process of the  $y$ -axis conversion coefficient  $\beta$  was the same as  $\alpha$ ; that is,  $\beta = 0.1454$ . The relationship between the positions of the filter imaging screen and the HD industrial camera is invariant. Once the conversion coefficients  $\alpha$  and  $\beta$  are determined, they are invariant and only need to be calibrated once.

#### 4.3. Data Processing Algorithm Comparison Test

Because of changes in the refractive index of air and other factors in a long distance of 30 m, the laser beam will slightly deflect during the propagation, and the position of the laser spot on the

filter imaging screen will vibrate. Given the isotropy of laser spot vibration, this experiment only addressed the  $x$ -axis coordinates of the spot, and the unit is pixels.

As shown in Figure 17, the laser spot position information of 1000 spot images was processed, wherein the abscissa axis was the acquisition time. As can be seen from the diagram, the vibration range of laser spot was big; therefore, when the laser propagation distance is long, the laser spot center data at a single time point could not be used to represent the monitoring reference. The position probability distribution of the 1000 spot images is shown in Figure 18. Although the laser spot was in vibration, the position probability distribution was approximately in accordance with Gaussian distribution; therefore, stable data was found to represent the monitoring reference by processing a set of data over a period of time.

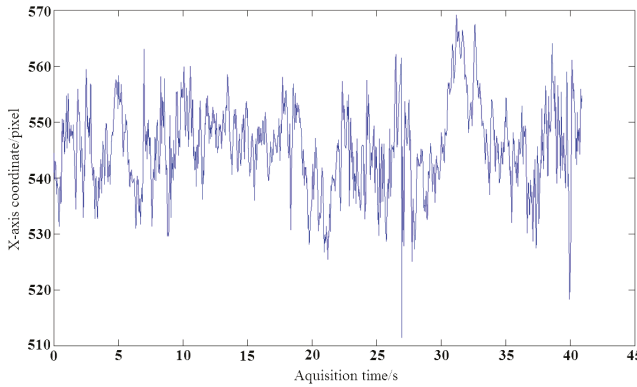


Figure 17. Laser spot position information of 1000 spot images.

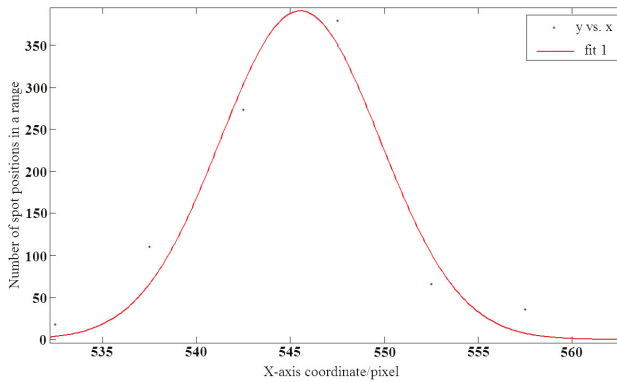


Figure 18. Position probability distribution.

Given the characteristics of the dataset, the following methods were used to process the data: (1) calculation of the extremum after Gauss fitting, (2) calculation of the expectation from original data, (3) calculation of the expectation after low-pass filtering, and (4) acquisition of the maximum probability data from data probability distribution map. The results are shown in Table 2. As can be seen, the fourth method has the highest accuracy; therefore, this method was used to process the laser spot position information.



Table 2. Results comparison.

Method	Accuracy (mm)
1. Calculation of the extremum after Gauss fitting	0.28
2. Calculation of the expectation from original data	0.42
3. Calculation of the expectation after low-pass filtering	0.21
4. Acquisition of the maximum probability data	0.12

4.4. Wireless Antenna Attitude Test

According to the principle of electromagnetic wave emission, the antenna has a certain angle when transmitting and receiving electromagnetic waves, as shown in Figure 19, where the signal is strongest in the main lobe range. When the antenna is arranged, if the antenna attitude can be adjusted to the right position, the intensity of the received signal will be enhanced.

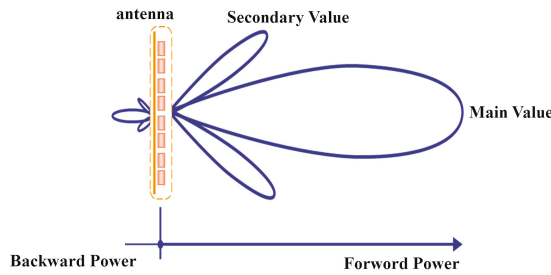


Figure 19. Schematic diagram of electromagnetic wave emission principle.

A spectrum analyzer was used as a wireless signal detection tool to analyze the size of the power received under different poses. Because the attitudes of the transmitting antenna and the receiving antenna were relative, the transmitting antenna was fixed horizontally during the experiment, and the attitude of the receiving antenna was changed.

As shown in Figure 20a, when the receiving antenna was parallel to the transmitting antenna, the maximum signal intensity was  $-23.63$  db. As shown in Figure 20b, when the receiving antenna was vertical to the transmitting antenna, the signal strength was minimum,  $-35.68$  db. Therefore, when using the instrument, the two antennas should be as parallel as possible to enhance the intensity of the received signal.

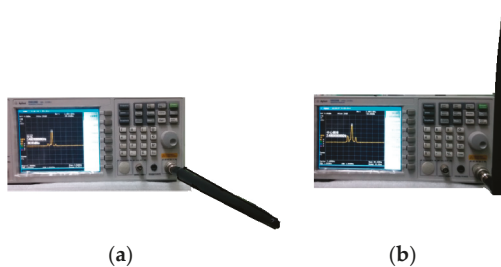


Figure 20. Receiving antenna power map. (a) Parallel attitude; (b) vertical attitude.

4.5. Temperature Change Test

The system needs to meet the requirement of real-time measurement without interruption for a long time; therefore, it is bound to be affected by the change of ambient temperature. However, the long distance laser reference propagation has a magnifying effect on the laser’s own angle change.

Since the maximum temperature difference in the tunnel during day and night is only 4 °C, the linear thermal deformation of the laser system can be ignored, but the angle change will have a large error. As shown in Figure 21, for a temperature difference 10 °C during 14 h, the x-axis coordinate difference of the spot position caused by thermal deformation can reach 38 pixels on a 50 m propagation distance. For the 4 °C temperature difference in the tunnel, the thermal deformation will cause less x-axis coordinate difference.

To reduce the effect of thermal deformation on the precision of the system, the mechanical structure was designed to be symmetrical to minimize the angle torsion caused by thermal deformation.

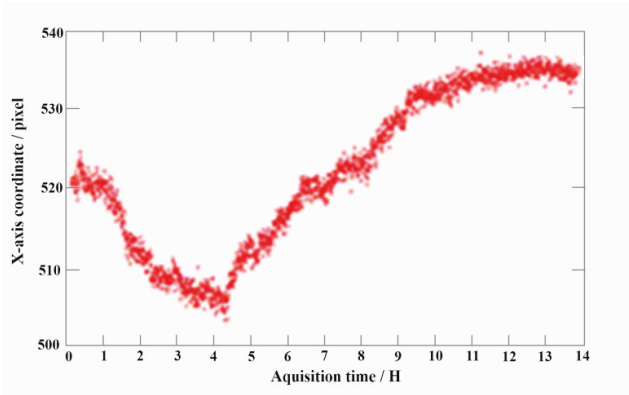


Figure 21. Result of data change caused by thermal deformation.

#### 4.6. System Accuracy Test Experiment

##### 4.6.1. Detection Accuracy Test Experiment

After the above tests were performed, the real-time tunnel deformation monitoring system was ready to operate. A high-precision micro motion platform was used to simulate the tunnel deformation process, and it moved at an interval of 5 mm. The obtained detection displacement is shown in Table 3.

Table 3. Detection accuracy test experimental data.

Position of Micro Motion Platform (mm)	Laser Spot Center before Calibration (pixel)	Laser Spot Center after Calibration (mm)	$x$ (mm)	$ \Delta x - 10 $ (mm)
(70, 0)	(598.27, 317.10)	(86.99, 46.11)	—	—
(60, 0)	(529.16, 317.29)	(76.94, 46.13)	10.05	0.05
(50, 0)	(459.92, 317.45)	(66.87, 46.15)	10.07	0.07
(40, 0)	(390.32, 317.60)	(56.75, 46.18)	10.12	0.12
(30, 0)	(322.41, 317.84)	(46.87, 46.21)	9.88	0.12
(20, 0)	(253.86, 318.14)	(36.91, 46.25)	9.96	0.04
(10, 0)	(185.46, 318.42)	(26.97, 46.29)	9.94	0.06
(0, 0)	(116.81, 318.61)	(16.88, 46.32)	10.09	0.09

According to Table 3, the monitoring system had a high detection accuracy better than 0.12 mm.

##### 4.6.2. Repeatability Accuracy Test Experiment

The micro motion platform performed a linear reciprocating motion at an interval of 10 mm. Experimental data are shown in Table 4.

Table 4. Repeatability accuracy test experimental data.

Position of Micro Motion Platform (mm)	System Detection Results		Position of Micro Motion Platform (mm)	System Detection Results	
	(x, y) (mm)	$  \Delta x -10 $ (mm)		(x, y) (mm)	$  \Delta x -10 $ (mm)
(70, 0)	(86.95, 46.09)	—	(60, 0)	(76.95, 46.13)	0.06
(60, 0)	(77.02, 46.12)	0.07	(70, 0)	(86.96, 46.11)	0.01
(50, 0)	(66.90, 46.15)	0.12	(60, 0)	(76.89, 46.12)	0.07
(40, 0)	(56.80, 46.19)	0.10	(50, 0)	(66.88, 46.16)	0.01
(30, 0)	(46.87, 46.21)	0.07	(40, 0)	(56.80, 46.18)	0.08
(20, 0)	(36.89, 46.23)	0.02	(30, 0)	(46.83, 46.22)	0.03
(10, 0)	(26.85, 46.27)	0.04	(20, 0)	(36.91, 46.25)	0.08
(0, 0)	(16.90, 46.32)	0.05	(10, 0)	(26.88, 46.28)	0.03
(10, 0)	(26.88, 46.28)	0.02	(0, 0)	(16.89, 46.33)	0.01
(20, 0)	(36.91, 46.25)	0.03	(10, 0)	(26.96, 46.29)	0.07
(30, 0)	(46.88, 46.22)	0.03	(20, 0)	(36.89, 46.26)	0.07
(40, 0)	(56.80, 46.20)	0.08	(30, 0)	(46.85, 46.23)	0.04
(50, 0)	(66.89, 46.14)	0.09	(40, 0)	(56.78, 46.18)	0.07

According to Table 4, the monitoring system had good stability and repeatability, and the repeatability accuracy was better than 0.11 mm.

#### 4.6.3. System Resolution

The micro motion platform moved slowly with the minimum displacement, that is, at an interval of 1  $\mu\text{m}$ . In the system accuracy test experiment, when the micro motion platform moved 10  $\mu\text{m}$ , the coordinates of the laser spot from the system software started to change, indicating that the system resolution was 10  $\mu\text{m}$ .

#### 4.7. Laser Transmitter Module Feasibility Test

In order to simulate the dusty and dark tunneling environment, this test was conducted on a foggy night as shown in Figure 22.

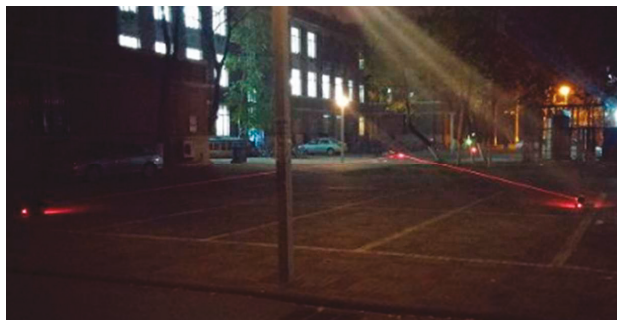
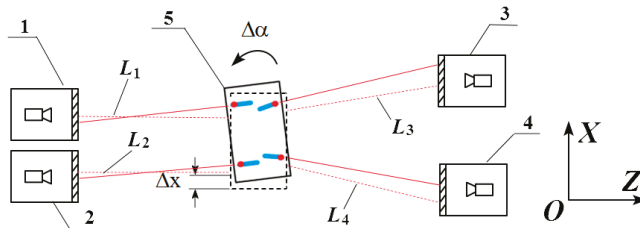


Figure 22. Test environment in foggy and dark night.

The test sketch map is shown in Figure 23.



**Figure 23.** Test sketch map. 1. reference monitoring image acquisition module A; 2. reference monitoring image acquisition module B; 3. image acquisition module I; 4. image acquisition module II. ( $L_1 = 40\text{ m}$ ,  $L_2 = 46\text{ m}$ ,  $L_3 = 35\text{ m}$ ,  $L_4 = 38\text{ m}$ ).

As shown in Figure 23, the laser transmitter module was located on a precision swivel table that was fixed on the micro motion platform. Two image acquisition modules and two reference monitoring image acquisition modules were placed on both sides of the laser transmitter module respectively, an  $L_s$  represents the distance between the laser transmitter module and the image acquisition modules,  $\Delta\alpha$  represents the horizontal rotation angle of the laser transmitter module, and  $\Delta x$  represents the displacement along the  $x$ -axis of the laser transmitter module.

In this test, the two reference monitoring modules and the two image acquisition modules were fixed, and the laser beam reference was shifted 12 mm and rotated 36°. According to Equation (9), the detected displacement can be compensated. The test data are shown in Table 5.

**Table 5.** Laser transmitter module feasibility test data (unit: mm).

Devices	Reference Monitoring Image Acquisition Modules		Image Acquisition Modules	
Number	1	2	3	4
Initial laser spot coordinates	(92.02, 48.13)	(69.27, 66.21)	(80.42, 60.30)	(73.89, 58.62)
Final laser spot coordinates	(97.18, 48.15)	(73.36, 66.18)	(62.22, 60.37)	(55.12, 58.70)
Horizontal displacement before compensation	5.16	4.09	18.20	18.71
Horizontal displacement after compensation	—	—	0.09	0.14

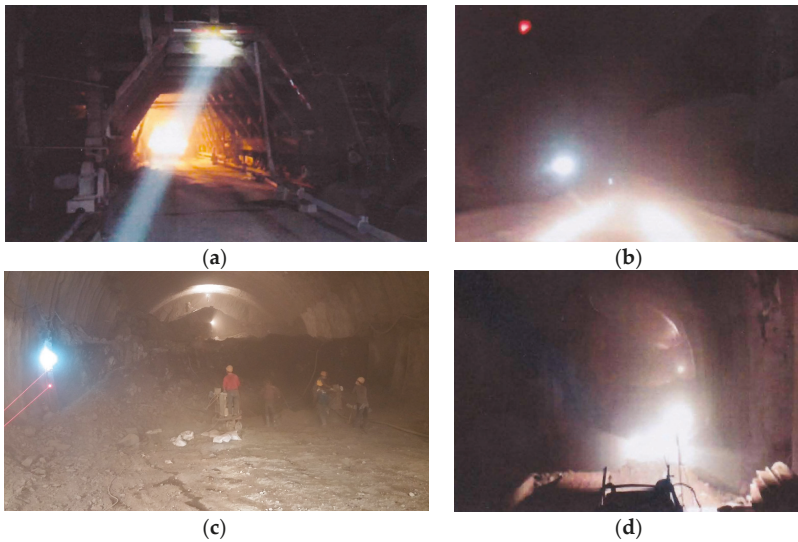
According to Table 5, this compensation method can effectively reduce the error caused by small sudden vibrations of the laser transmitter module.

#### 4.8. Field Test

To verify the stability and accuracy of the monitoring system, the system was taken to the Xingyuan Town Tunnel in Heilongjiang province for a field test.

As shown in Figure 24, the tunneling environment was adverse and filled with dust, darkness and sudden vibration during construction. Tunnel trolley, as shown in Figure 24a, ran between initial support and secondary lining; this may block the laser transmitting path, so that the laser transmitter module was fixed on the earth close to the side face; therefore, the proposed monitoring system had no conflict with construction as shown in Figure 24c.

As shown in Figure 24b,d, the dust is heavy, this may decrease the laser spot energy received by the image acquisition module, so that the laser transmitter module used 500 mw high power lasers with an optical lens to achieve better long-distance transmission effect. In addition, the distance from the laser transmitter module to the measured points was varying from 30 m to 80 m, and the laser spot energy decreased with increasing distance, so that a laser energy transmission test with difference receiving distances was conducted in the dusty tunnel to detect the image quality acquired by image acquisition module. The test data are shown as Table 6



**Figure 24.** Tunnel environment during construction. (a) Tunnel construction trolley; (b) dusty and dark tunneling environment; (c) deformation monitoring during construction; (d) tunnel face environment.

**Table 6.** Laser energy transmission test data.

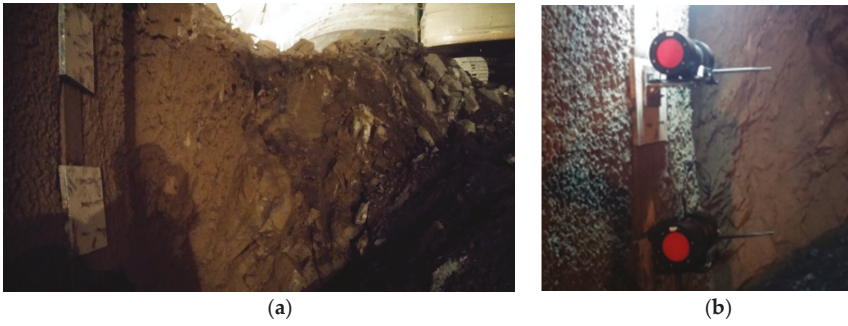
Receiving Distance (m)	Gray Value of Laser Spot	Receiving Distance (m)	Gray Value of Laser Spot	Receiving Distance (m)	Gray Value of Laser Spot
20	255	45	246	70	185
25	255	50	233	75	171
30	255	55	221	80	169
35	255	60	210	85	155
40	255	65	196	90	139

According to Table 6, laser spot energy decreased with increasing receiving distance, but the laser spot could still be obtained by image acquisition module, and the gray value of the laser spot met measurement requirements; this indicates that the proposed system is not susceptible to dusty and dark environment. The laser spot image obtained in 70 m receiving distance is shown in Figure 25a, and the high power laser with optical lens is shown in Figure 25b.



**Figure 25.** (a) Laser spot image in 70 m receiving distance; (b) high power laser with optical lens.

In order to verify the stability and accuracy of the monitoring system, a field test was conducted in the tunneling environment. The image acquisition module was fixed on the edge of initial support of the tunnel, as shown in Figure 26a.



**Figure 26.** (a) Measured points; (b) image acquisition module mounted on measured points.

As shown in Figure 27, a total station was used to synchronously measure the deformation of the same measured point in the tunnel every 1 h, while the system conducted real-time monitoring during construction as shown in Figure 28.



**Figure 27.** Total station measurement.



**Figure 28.** System measurement.

In the field test, a large number of data were obtained, and part of the system measurement results corresponding with the station measurement results are presented in Table 7.

Table 7. Partial field test data.

No.	Time	Deformation from Total Station (mm)	Deformation from Monitoring System (mm)
1	16:00	0	0
2	17:00	2.51	2.46
3	18:00	4.90	4.95
4	19:00	6.21	6.22
5	20:00	7.10	7.15
6	21:00	8.01	7.99

According to Table 6, the monitoring system and total station measurement results are close, which indicates that the monitoring system has good accuracy, stability, and repeatability in a field application.

## 5. Discussion

### 5.1. About Data Compensation Method for Environment Vibration

The data compensation method for environment vibration is effective for small and sudden vibration. When the vibration of the laser transmitter module is large and continues for a while, the spot position in the reference monitoring image acquisition modules will change and reflect the vibration; then, the system will stop the measurement for the measured points until the vibration of the laser transmitter module stops, filtering out the influence of the vibration on the accuracy of the system.

### 5.2. Drawbacks and Improvements of Simulation Environment

Because of the complex tunnel environment, the verification test environment needs to be discussed for such a complicated engineering mechanics problem.

- (1) For the principle verification test, image acquisition module calibration experiment and wireless antenna attitude test, a relative stable environment was needed; thus, these tests were conducted in the laboratory reasonably;
- (2) For the data processing test and temperature change test, a real tunnel environment was needed. However, because of the busy, dusty and vibrational construction sites, it was not only hard to study effects of single factor, but also inconvenient to conduct the tests in such a complex tunnel environment. Therefore, the tests were conducted in laboratory and studied without dust, and it would be better to simulate a dusty environment during the lab test;
- (3) For the system accuracy test, the deformation was simulated by a micro motion platform with a 1  $\mu\text{m}$  accuracy. The micro motion platform was placed on a table and moved horizontally. However, for tunnel settlement deformation, it settled vertically. Given the isotropy of the image sensor for vertical and horizontal direction, the test was reliable, and the field test verified this adequately. In addition, it would be better to use a vertical micro motion platform for the deformation simulation in a vertical direction;
- (4) For the laser transmitter feasibility test, this test was conducted in a foggy night; thus, the simulation environment was similar to a tunnel, a dusty and dark environment in low visibility.
- (5) For the field test, the environment is dusty and even wet. In order to prevent the interference of tunnel dust and water mist, dirt shrouds and waterproof devices were used in laser transmitter module, image acquisition module and some electric parts. To improve the permeability of the laser beam, an automatic device needs to be added to clear the mud consisting of dust and water mist on the laser receiving screen and laser transmitter module glass window.

### 5.3. About Diffraction Rings of Laser Spots

In order to improve a laser transmitting effect, a set of optical lens were used, and this produced additional diffraction rings, as shown in Figure 25a. Precisely because of the dusty environment, the diffraction effect decreased through the dust [40], and this had advantages in the image process.

### 5.4. About Application of the Proposed System

The proposed system was designed for tunnel deformation monitoring in construction, but it is also suitable for the tunnel that has already been built. However, because of the heavy traffic or trains' passage, an unavoidable noise is produced in operating tunnels, and the noise intensity is different and relates to traffic volume, vehicle speed and vehicle type.

In order to solve this problem in the application of operating tunnels, a noise prediction model based on traffic volume, vehicle speed and vehicle type needs to be studied [41]. With this prediction model, a data compensation method should be proposed to compensate for the measurement error caused by unavoidable noise.

## 6. Conclusions

This paper presents a novel method based on laser and machine vision to automatically measure tunnel deformation of multiple interest points in real time and effectively compensate for the environment vibration, and, moreover, it can overcome influence of dusty and dark environments in low visibility. An automatic and wireless real-time tunnel deformation monitoring system that is based on laser and machine vision, and can give early warnings for tunnel collapse accidents, is proposed. Compensation methods are proposed to reduce the measurement errors caused by laser beam feasibility, temperature, air refraction index, and wireless antenna attitude. Experimental results show that the proposed monitoring system has a high detection accuracy, which can compensate for the shortcomings of the traditional tunnel deformation monitoring method in highly automatic real-time monitoring. The proposed system can also avoid the conflict with tunnel construction, that is, it realizes real-time monitoring during construction. Moreover, the monitoring system has good stability and repeatability. The detection accuracy is better than 0.12 mm, the repeatability accuracy is less than 0.11 mm, and the minimum resolution is 10  $\mu\text{m}$ .

**Author Contributions:** Z.Q. conceived the method and revised the paper; H.L. designed the experiments and wrote the paper; W.H. and C.W. designed mechanical and electrical structure; W.H. and J.L. designed and conducted the field test; W.H., J.L., and Q.S. developed system software and improved the experimental results.

**Funding:** This research was financially supported by the National Natural Science Foundation of China (NSFC) (No: 51775378) and the National Key R&D Program of China (No.2017YFF0108102).

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Glišić, B.; Inaudi, D. Introduction to Structural Health Monitoring. In *Fibre Optic Methods for Structural Health Monitoring*; John Wiley & Sons, Ltd.: Hoboken, NJ, USA, 2007.
2. Brownjohn, J.M.W. Structural health monitoring of civil infrastructure. *Philos. Trans. Math. Phys. Eng. Sci.* **2007**, *365*, 589–622. [[CrossRef](#)] [[PubMed](#)]
3. Ding, L.Y.; Zhou, C.; Deng, Q.X.; Luo, H.B.; Ye, X.W.; Ni, Y.Q.; Guo, P. Real-time safety early warning system for cross passage construction in Yangtze Riverbed Metro Tunnel based on the internet of things. *Autom. Constr.* **2013**, *36*, 25–37. [[CrossRef](#)]
4. Jicheng, T. The present situation and developing trend of surveying instrument. *Mod. Surv. Mapp.* **2003**, *26*, 46–48. [[CrossRef](#)]
5. Kanaoka, C.; Furuuchi, M.; Inaba, J.; Myojo, T. Flow and dust concentration near working face of a tunnel under construction. *J. Aerosol Sci.* **2000**, *31*, 31–32. [[CrossRef](#)]



6. Yuan, W.; Xue, D. Review of Tunnel Lining Crack Detection Algorithms based on Machine Vision. *Chin. J. Sci. Instr.* **2017**, *38*, 3100–3111.
7. Han, D.; Xie, J. State-of-Arts of health monitoring techniques for long span bridges. *Bridge Constr.* **2002**, *6*, 69–73. [[CrossRef](#)]
8. Wang, J. *Tunnel Project Monitoring and Information Design Principles*; China Railway Publishing House: Beijing, China, 1990; pp. 19–40. ISBN 7-113-00917-4.
9. Zhang, S. *Algorithm Research in Image Processing and Recognition for Coded Level Ruler of Digital Levels*; Xi'an Technological College: Xi'an, China, 2004; Volume 24, pp. 353–355.
10. Zhang, P. *Study and Research of Digital Photogrammetry to Rock Mass Joint Distortion in the Underground Structures*; Wuhan University of Technology: Wuhan, China, 2007.
11. Yang, S.; Liu, W.; Wang, M. Study on the auto-total station system for monitoring analyzing and forecasting tunnel country rock deformation. *J. China Railw. Soc.* **2004**, *26*, 93–97. [[CrossRef](#)]
12. Yang, S.; Liu, W.; Shi, H.; Huang, F. A study on the theory and method of non-contact monitoring for tunnel rock deformation based on free stationing of a total station. *China Civ. Eng. J.* **2006**, *39*, 100–104. [[CrossRef](#)]
13. Chen, Y. *Study on Tunnel Deformation Monitoring Using Digital Image Processing Techniques*; Guangdong University of Technology: Guangzhou, China, 2014.
14. Zhao, X.; Qiu, H. Application of fiber bragg grating sensing technology to tunnel monitoring. *Chin. J. Rock Mech. Eng.* **2007**, *26*, 587–593. [[CrossRef](#)]
15. Xu, D.S.; Zhao, Y.M.; Liu, H.B.; Zhu, H.H. Deformation Monitoring of Metro Tunnel with a New Ultrasonic-Based System. *Sensors* **2017**, *17*, 1758. [[CrossRef](#)]
16. Yufeng, S.; Jun, Z.; Yingya, Z. Tunnel safety monitoring based on terrestrial laser scanning technology. *J. Southeast Univ.* **2013**, *43*, 246–249. [[CrossRef](#)]
17. Xu, Y.; Brownjohn, J.M.W. Review of machine-vision based methodologies for displacement measurement in civil structures. *J. Civ. Struct. Health Monit.* **2017**, *8*, 91–110. [[CrossRef](#)]
18. Wang, J.; Chen, D.; Zhao, T. Research of Three-Dimensional Reconstruction Based on Computer Vision. In Proceedings of the International Conference on Wireless Communications Networking & Mobile Computing, Leipzig, Germany, 21–25 September 2010.
19. Wang, L. *Tunnel Deformation Monitoring Based on Digital Close-Range Photogrammetry*; Chongqing Jiaotong University: Chongqing, China, 2012.
20. Huang, H.; Sun, L.; Xue, Y. Research progress of tunnel lining surface disease detection technology based on machine vision. *Mod. Tunn. Technol.* **2014**, *51*, 19–31. [[CrossRef](#)]
21. Sang, Z. *Research on Close-Range Photogrammetry Technology in Tunnel Deformation Monitoring*; Tongji University: Shanghai, China, 2008.
22. Scaioni, M.; Barazzetti, L.; Giussani, A.; Previtali, M.; Roncoroni, F.; Alba, M.I. Photogrammetric techniques for monitoring tunnel deformation. *Earth Sci. Inform.* **2014**, *7*, 83–95. [[CrossRef](#)]
23. Qi, C.; Kegang, H.; Shaohui, L.; Wu, T.; Guo, W.; Tang, Y. Development of a Modeling Method for Monitoring Tunnel Deformation Based on Active Panoramic Vision Technology. In Proceedings of the Sixth International Conference on Intelligent Systems Design & Engineering Applications, Guiyang, China, 18–19 August 2016.
24. Qi, C. *Research of the Modeling Method for Monitoring Tunnel Deformation Based on Active Stereo Omni-Directional Vision Sensor*; Zhejiang University of Technology: Zhejiang, China, 2017.
25. Wei, T.G. NATM Tunneling Construction Technology. *Value Eng.* **2013**, *32*, 120–121. [[CrossRef](#)]
26. Sui, X.Z.; Teng, W.Y. Analysis of the Interaction between Primary Support and the Second Lining of Large Span and Double Arch Tunnel. *Nat. Sci.* **2003**, *16*, 8–11. [[CrossRef](#)]
27. Sun, M.L.; Zhu, Y.Q. Research on construction time of secondary lining soft rock of large-deformation tunnel. *Chin. J. Rock Mech. Eng.* **2008**, *27*, 580–588. [[CrossRef](#)]
28. Chen, D. Research on Deformation Features and Stability Analysis of Tunnels in Carbonaceous Shale. *Nat. Sci.* **2011**, *30*, 58–64. [[CrossRef](#)]
29. Lai, J.; Fan, H.; Lai, H. In-situ monitoring and analysis of tunnel deformation law in weak loess. *Rock Soil Mech.* **2015**, *36*, 2003–2012. [[CrossRef](#)]
30. Zhao, H. *Evaporation and Research on Anti-Radiation of Bandpass Filter*; Harbin Institute of Technology: Harbin, China, 2014.
31. Bai, S.; Gu, P.; Liu, X.; Tang, J. Optical stability of thin film filters. *Acta Photonica Sin.* **2001**, *30*, 576–580.

32. Zhang, J.; Wang, S.; Wang, Z. Research on diffuse reflectivity of the reflective screen for a laser-guided weapon simulation system. *Electron. Opt. Control* **2005**, *12*, 27–30. [[CrossRef](#)]
33. Zhou, Z. *Automatic Measurement Technology for Shaft Parts Based on the Analysis of Sub-Pixel*; Jilin University: Jilin, China, 2016.
34. Kalkman, C.J. LabVIEW: A software system for data acquisition, data analysis, and instrument control. *J. Clin. Monit.* **1995**, *11*, 51–58. [[CrossRef](#)]
35. Wei, Z.; Yang, K. Laser spot analytic system based on CCD and LabView. *Laser J.* **2011**, *32*, 17–18. [[CrossRef](#)]
36. Singh, K.R.; Garg, A.; Upadhyay, A. A Novel Camera Calibration method for measurement of horizontal and vertical displacement of target object using LABVIEW. *Int. J. Adv. Eng. Res. Sci.* **2015**, *2*, 21–24.
37. Wu, B.F.; Chen, W.H.; Chang, C.W.; Liu, C.C.; Chen, C.J. Dynamic CCD camera calibration for traffic monitoring and vehicle applications. In Proceedings of the IEEE International Conference on Systems, Man and Cybernetics, Montreal, QC, Canada, 7–10 October 2007; pp. 1717–1722.
38. Hsu, C.C.; Lu, M.C.; Chin, K.W. Distance measurement based on pixel variation of CCD images. *ISA Trans.* **2009**, *48*, 389–395. [[CrossRef](#)] [[PubMed](#)]
39. Yi, Y.; Huang, J.; Long, M. Camera calibration method based on binocular vision research. *Int. J. Sci.* **2016**, *3*, 129–134.
40. Xu, Y.; Qiu, J.H. Focusing and Diffraction Effects of a Laser Beam through a Lens System. *Appl. Laser* **1991**, *11*, 51–54.
41. Qin, L.; Dong, L.-L.; Xu, W.-H.; Zhang, L.-D.; Leon, A.S. An intelligent luminance control method for tunnel lighting based on traffic volume. *Sustainability* **2017**, *9*, 2208. [[CrossRef](#)]



© 2018 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).

Article

# Innovative Methodology of On-Line Point Cloud Data Compression for Free-Form Surface Scanning Measurement

Yan Li <sup>1</sup>, Yuyong Ma <sup>2</sup>, Ye Tao <sup>1,3,\*</sup> and Zhengmeng Hou <sup>3</sup>

<sup>1</sup> School of Manufacturing Science and Engineering, Sichuan University, Chengdu 610065, China; liyan@scu.edu.cn

<sup>2</sup> Aerospace Research Institute of Materials and Processing Technology, China Academy of Launch Vehicle Technology, Beijing 100076, China; myy007391@casctx-1.net.cn

<sup>3</sup> Energy Research Center of Lower Saxony (EFZN), 38640 Goslar, Germany; hou@tu-clausthal.de

\* Correspondence: yetao@scu.edu.cn; Tel.: +49-0152-5155-4369

Received: 23 October 2018; Accepted: 3 December 2018; Published: 10 December 2018



**Featured Application:** On-line point cloud data compression process for 3D free-form surface contact or non-contact scanning measuring equipment.

**Abstract:** In order to obtain a highly accurate profile of a measured three-dimensional (3D) free-form surface, a scanning measuring device has to produce extremely dense point cloud data with a great sampling rate. Bottlenecks are created owing to inefficiencies in manipulating, storing and transferring these data, and parametric modelling from them is quite time-consuming work. In order to effectively compress the dense point cloud data obtained from a 3D free-form surface during the real-time scanning measuring process, this paper presents an innovative methodology of an on-line point cloud data compression algorithm for 3D free-form surface scanning measurement. It has the ability to identify and eliminate data redundancy caused by geometric feature similarity between adjacent scanning layers. At first, the new algorithm adopts the bi-Akima method to compress the initial point cloud data; next, the data redundancy existing in the compressed point cloud is further identified and eliminated; then, we can get the final compressed point cloud data. Finally, the experiment is conducted, and the results demonstrate that the proposed algorithm is capable of obtaining high-quality data compression results with higher data compression ratios than other existing on-line point cloud data compression/reduction methods.

**Keywords:** data compression; data reduction; free-form surface; point cloud; scanning measurement; redundancy identifying; redundancy eliminating; geometric feature similarity

## 1. Introduction

With the rapid development of modern industry, three-dimensional (3D) free-form surface parts are being utilized more and more widely. These involve, but are not limited to, aviation, aerospace, shipbuilding, automotive, biomedical and home appliance industries [1,2]. Recently, the automated 3D digitization of free-form surface objects has been widely applied in many areas, such as additive manufacturing (3D printing), rapid prototyping, reverse engineering, civil buildings, medical prosthetics and clinical diagnosis [3–13]. Scanning measurement is one of the key technologies for digitizing 3D physical models with free-form surfaces [14–17]. Unfortunately, in order to obtain a high-quality profile of a measured surface, scanning measuring devices have to produce massive amounts of point cloud data with great sampling rates, and not all these points are indispensable [18–20]. Bottlenecks arise from the inefficiencies of storing, manipulating and transferring them [21]. Furthermore, the parametric

modelling from these massive amount of point cloud data is a time-consuming task [22–24]. For this reason, compressing the measured point data while maintaining the required accuracy is a crucial task during the scanning measuring process [25]. Herein, the required accuracy is a threshold of distance, which is preset to a constant positive integer before the beginning of scanning measurement. The accuracy of a certain data compression algorithm is characterized by the distance from each sampled point in the initial dense point cloud data to the surface generated by the compressed point cloud. Describing a measured surface with the least point data while guaranteeing a certain data compression accuracy is always an expectation [26,27]. Therefore, a high-quality point cloud data compression algorithm for 3D free-form surface scanning measurement is being pursued constantly [28].

Experts and scholars around the world have been paying more and more attention to this issue, and a number of point cloud data compression/reduction algorithms for free-form/irregular surface scanning measurement have been developed. Lee et al. [29] proposed an algorithm for processing point cloud data obtained by laser scanning devices. This algorithm adopts a one-directional (1D) or bi-directional (2D) non-uniform grid to reduce the amount of point cloud data. Chen et al. [5] presented a data compression method based on a bi-directional point cloud slicing strategy for reverse engineering. This method can preserve local details (geometric features in both two parametric directions) when performing data compression. Ma and Cripps [30] proposed a new data compression algorithm for surface points to preserve the original surface points. The error metric is defined as the relative Hausdorff distance between two principal curvature vector sets for surface shape comparison. After comparison, the difference between the compressed data points and original data points can be obtained. Therefore, redundant points are repeatedly removed until the difference induced exceeds the specified tolerance. Smith, Petrova, and Schaefer [31] presented a progressive encoding and compression method for surfaces generated from point cloud data. At first, an octree is built whose nodes contain planes that are constructed as the least square fit of the data within that node. Then, this octree is pruned to remove redundant data while avoiding topological changes created by merging disjointed linear pieces. Morell et al. [32] presented a geometric 3D point cloud lossy compression system based on plane extraction, which represents the points of each scene plane as a Delaunay triangulation and a set of points/area information. This compression system can be customized to achieve different data compression or accuracy ratios. The above methods have focused on optimizing data compression quality based on building and processing polyhedral models or numerical iterative calculations. Nevertheless, they are all off-line data compression algorithms and can only compress the point cloud data of a whole measured surface after data acquisition. In other words, they cannot perform online data compression during real-time measurement. Data acquisition and data compression processes are completely separate. A large amount of redundant point cloud data occupies a great deal of storage space in scanning measuring devices. Moreover, the transmission and processing of point cloud data still takes up a significant amount of time and hardware resources.

This problem has attracted the attention of many scholars and engineers, and they have proposed quite a number of on-line point cloud data compression/reduction methods. Lu et al. [33] adopted the chordal method to compress point cloud data, and realized the on-line data compression of point cloud data during real-time scanning measurement for the first time. ElKott and Veldhuis [34] presented an automatic surface sampling approach based on scanning isoparametric lines. The sampling locations are confirmed by the deviations between the alternative geometry and sampled model, and the location of each sampling line is confirmed by the curvature of the sampled surface model. Wozniak, Balazinski, and Mayer [35] presented a point cloud data compression method based on fuzzy logic and the geometric solution of an arc at each measured point. This is an on-line data compression method and can be used in the surface scanning measuring process of coordinate measuring machines (CMMs). Jia et al. [36] proposed an on-line data compression method based on the equal-error chordal method and isochronous sampling. In order to solve the problem of massive data storage, dual-buffer and dual-thread dynamic storage is adopted. Tao et al. [37] found that the essence of all the above on-line point cloud data compression methods is the chordal method, which specifies that all discrete dense

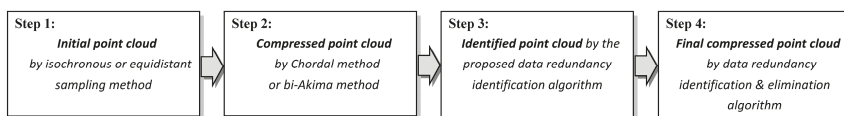
point sets are connected by straight segments. Therefore, the surface reconstructed by the compressed point cloud will be full of cusp points, and so we cannot obtain a smooth interpolated surface. In view of this limitation, they presented an on-line point cloud data extraction algorithm using bi-Akima spline interpolation.

Although the above methods implement on-line point cloud data compression, they can only eliminate data redundancy of the current scanning line. Nevertheless, most surface 3D scanning measuring devices adopt a layer-by-layer scanning path (e.g., contact scanning probes [38], laser triangle displacement sensors [39], linear structured light systems [40], industrial computed tomography (CT) systems [41], etc.), and adjacent scanning lines are extremely similar in shape. The geometric feature similarity between such scanning layers is bound to result in data redundancy, which makes it possible to further compress the point cloud data during the scanning measuring process. Therefore, this study focuses on identifying and eliminating this kind of data redundancy caused by geometric feature similarity between adjacent scanning layers. After that, the massive amount of point cloud data can be further compressed during the 3D free-form surface measuring process.

The contents of this paper consist of four sections. In Section 2, the innovative methodology of the on-line point cloud data compression algorithm for 3D free-form surface scanning measurement is described in detail. In Section 3, the proposed algorithm was tested in the real-time scanning measuring process and compared with existing methods. Finally, some conclusions are drawn from this paper in Section 4.

## 2. Innovative Methodology

As shown in Figure 1, the overall process of on-line point cloud data compression in this work consists of four steps. In Step 1, the initial point cloud flow is obtained by 3D scanning measuring devices using an isochronous [42] or equidistant sampling method and the layer-by-layer scanning path is adopted (Figure 2). In Step 2, the initial point cloud data flow is immediately compressed by the chordal method [36] or bi-Akima method [37], both of which compress the amount of point cloud data based on the data redundancy in the current single scanning layer. In Step 3, the data redundancy in the compressed point cloud which is obtained in the previous step is further identified. In Step 4, the identified redundant point data is eliminated, and then we can obtain the final compressed point cloud. At last, the final compressed data flow is transmitted to the storage space of the measurement system.



**Figure 1.** The overall process of on-line data compression for 3D free-form surface scanning measurement.

Herein, the real-time performance of the proposed data compression algorithm needs to be further analyzed and described. The path planning is performed before the start of the scanning measurement in Step 1. As shown in Figure 2, a layer-by-layer scanning path is adopted. The distance between the adjacent scanning layers is determined by the preset measuring accuracy. The measured surface is cut by the scanning layers to form a number of corresponding scanning lines. As shown in Figure 2, there are two planning modes for scanning directions: (i) the progressive scanning mode, and (ii) the S-type scanning mode. Regardless of the scanning mode, the measuring device in Step 1 will continuously transmit the initial point cloud data flow to the data compressor in Step 2. The compressor performs data compression immediately after receiving all initial point data of a single scanning layer, rather than waiting for the entire surface to be scanned before performing data compression. That is, each time the point cloud data in the current scanning layer is completely transmitted to the compressor, the subsequent data compression algorithm is executed immediately.

Therefore, the proposed data compression algorithm is essentially a quasi-real-time method, which we call an on-line data compression method.

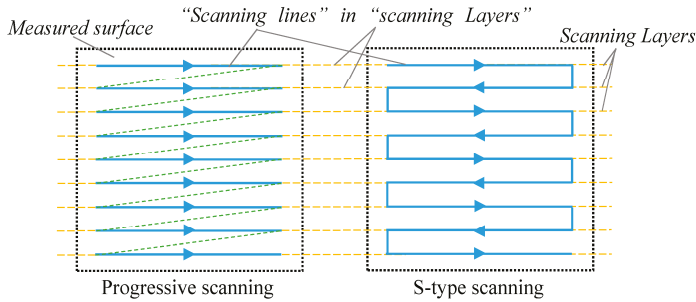


Figure 2. Layer-by-layer scanning path for 3D free-form surface scanning measurement.

The flow chart of this algorithm is illustrated in Figure 3, and its principle is described in detail as follows:

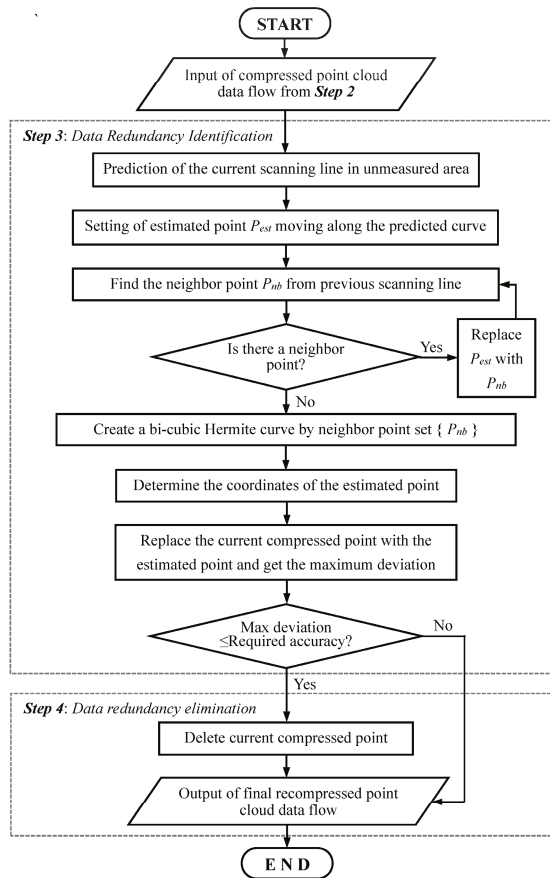


Figure 3. The flow chart of point cloud data redundancy identification and elimination algorithm.

### 2.1. Data Redundancy Identification

In order to identify redundant data points in the compressed point cloud data flow from Step 2, it is first necessary to predict the current scan line in the unmeasured area. Herein, the prediction is realized by Hermite extrapolation [43], and a predicted curve is created. The data redundancy identification algorithm is detailed as follows:

Figure 4 shows the schematic diagram of the data redundancy identification algorithm, in which line  $i$  is the current scanning line during the on-line measuring process, and  $P_{i,j}$  represents the  $j$ th point in scanning line  $i$ . If  $j \geq 2$ , a shape-preserving piecewise bicubic Hermite curve can be built to predict the shape and direction of the current scanning line; here, we name this the predicted curve, as shown in Figure 4. After that, suppose  $k$  is a positive integer and let  $1 \leq k < j$  and the coordinates of point  $P_{i,k}$  be  $(x_k, y_k, z_k)$ ; then, a series of specific Hermite interpolation polynomials can be determined by

$$\begin{cases} H_y(x) = y_k\alpha_k(x) + y_{k+1}\alpha_{k+1}(x) + y'_k\beta_k(x) + y'_{k+1}\beta_{k+1}(x) \\ H_z(x) = z_k\alpha_k(x) + z_{k+1}\alpha_{k+1}(x) + z'_k\beta_k(x) + z'_{k+1}\beta_{k+1}(x) \end{cases}, \quad (1)$$

where

$$\begin{cases} \alpha_k(x) = \left(1 + 2\frac{x-x_k}{x_{k+1}-x_k}\right)\left(\frac{x-x_{k+1}}{x_k-x_{k+1}}\right)^2 \\ \alpha_{k+1}(x) = \left(1 + 2\frac{x-x_{k+1}}{x_k-x_{k+1}}\right)\left(\frac{x-x_k}{x_{k+1}-x_k}\right)^2 \\ \beta_k(x) = (x-x_k)\left(\frac{x-x_{k+1}}{x_k-x_{k+1}}\right)^2 \\ \beta_{k+1}(x) = (x-x_{k+1})\left(\frac{x-x_k}{x_{k+1}-x_k}\right)^2 \end{cases}, \quad (2)$$

and the first derivatives of  $y'_k, y'_{k+1}, z'_k, z'_{k+1}$  can be estimated by the following formulas.

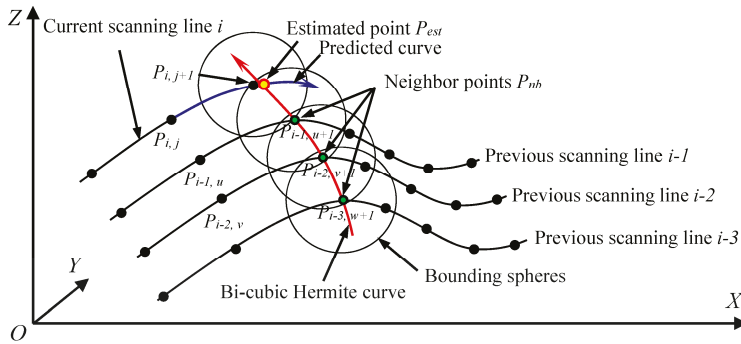


Figure 4. The schematic diagram of data redundancy identification.

When  $1 < k < j$ :

$$y'_k = f_y'(x_k) = \begin{cases} 0, \text{ if } \frac{y_{k+1}-y_k}{x_{k+1}-x_k} \cdot \frac{y_k-y_{k-1}}{x_k-x_{k-1}} < 0 \\ \frac{1}{2} \left( \frac{y_{k+1}-y_k}{x_{k+1}-x_k} + \frac{y_k-y_{k-1}}{x_k-x_{k-1}} \right), \text{ if } \frac{y_{k+1}-y_k}{x_{k+1}-x_k} \cdot \frac{y_k-y_{k-1}}{x_k-x_{k-1}} \geq 0 \end{cases}, \quad (3)$$

$$z'_k = f_z'(x_k) = \begin{cases} 0, \text{ if } \frac{z_{k+1}-z_k}{x_{k+1}-x_k} \cdot \frac{z_k-z_{k-1}}{x_k-x_{k-1}} < 0 \\ \frac{1}{2} \left( \frac{z_{k+1}-z_k}{x_{k+1}-x_k} + \frac{z_k-z_{k-1}}{x_k-x_{k-1}} \right), \text{ if } \frac{z_{k+1}-z_k}{x_{k+1}-x_k} \cdot \frac{z_k-z_{k-1}}{x_k-x_{k-1}} \geq 0 \end{cases}. \quad (4)$$

When  $k = 1$ :

$$y'_1 = \begin{cases} 0, \text{ if } d_y \cdot \frac{y_2-y_1}{x_2-x_1} < 0 \\ 3 \cdot \frac{y_2-y_1}{x_2-x_1}, \text{ if } |d_y| > 3 \left| \frac{y_2-y_1}{x_2-x_1} \right| \& \frac{(y_2-y_1)(y_3-y_2)}{(x_2-x_1)(x_3-x_2)} < 0 \end{cases}, \quad (5)$$

$$z'_1 = \begin{cases} 0, \text{ if } : d_z \cdot \frac{z_2-z_1}{x_2-x_1} < 0 \\ 3 \cdot \frac{z_2-z_1}{x_2-x_1}, \text{ if } : |d_z| > 3 \left| \frac{z_2-z_1}{x_2-x_1} \right| \& \frac{(z_2-z_1)(z_3-z_2)}{(x_2-x_1)(x_3-x_2)} < 0 \end{cases} \tag{6}$$

in which

$$\begin{cases} d_y = \frac{(x_3+x_2-2x_1)(y_2-y_1)}{(x_2-x_1)(x_3-x_1)} - \frac{(x_2-x_1)(y_3-y_2)}{(x_3-x_2)(x_3-x_1)} \\ d_z = \frac{(x_3+x_2-2x_1)(z_2-z_1)}{(x_2-x_1)(x_3-x_1)} - \frac{(x_2-x_1)(z_3-z_2)}{(x_3-x_2)(x_3-x_1)} \end{cases} \tag{7}$$

When  $k = j$ :

$$y'_j = \begin{cases} 0, \text{ if } : d_y \cdot \frac{y_j-y_{j-1}}{x_j-x_{j-1}} < 0 \\ 3 \cdot \frac{y_j-y_{j-1}}{x_j-x_{j-1}}, \text{ if } : |e_y| > 3 \left| \frac{y_j-y_{j-1}}{x_j-x_{j-1}} \right| \& \frac{(y_j-y_{j-1})(y_{j-1}-y_{j-2})}{(x_j-x_{j-1})(x_{j-1}-x_{j-2})} < 0 \end{cases} \tag{8}$$

$$z'_j = \begin{cases} 0, \text{ if } : d_y \cdot \frac{z_j-z_{j-1}}{x_j-x_{j-1}} < 0 \\ 3 \cdot \frac{z_j-z_{j-1}}{x_j-x_{j-1}}, \text{ if } : |e_z| > 3 \left| \frac{z_j-z_{j-1}}{x_j-x_{j-1}} \right| \& \frac{(z_j-z_{j-1})(z_{j-1}-z_{j-2})}{(x_j-x_{j-1})(x_{j-1}-x_{j-2})} < 0 \end{cases} \tag{9}$$

in which

$$\begin{cases} e_y = \frac{(2x_j-x_{j-1}-x_{j-2})(y_j-y_{j-1})}{(x_j-x_{j-1})(x_j-x_{j-2})} - \frac{(x_j-x_{j-1})(y_{j-1}-y_{j-2})}{(x_{j-1}-x_{j-2})(x_j-x_{j-2})} \\ e_z = \frac{(2x_j-x_{j-1}-x_{j-2})(z_j-z_{j-1})}{(x_j-x_{j-1})(x_j-x_{j-2})} - \frac{(x_j-x_{j-1})(z_{j-1}-z_{j-2})}{(x_{j-1}-x_{j-2})(x_j-x_{j-2})} \end{cases} \tag{10}$$

Herein, based on the compressed point cloud data flow from Step 2, the shape-preserving piecewise bicubic Hermite polynomials can be created according to the above algorithm. Then, Hermite extrapolation is performed to create a predicted curve, which is marked in blue as shown in Figure 4, and its analytical formula can be described as follows:

$$\begin{cases} H_y(x) = y_{j-1}\alpha_{j-1}(x) + y_j\alpha_j(x) + y'_{j-1}\beta_{j-1}(x) + y'_j\beta_j(x) \\ H_z(x) = z_{j-1}\alpha_{j-1}(x) + z_j\alpha_j(x) + z'_{j-1}\beta_{j-1}(x) + z'_j\beta_j(x) \end{cases} \tag{11}$$

After that, an estimated point  $P_{est}$  is created to move along the predicted curve with a stepping distance of  $\lambda$ .  $P_{i,j}$  is the starting point of  $P_{est}$ . Meanwhile, a bounding sphere is built with point  $P_{est}$  as the center. The radius of the sphere is

$$R_{sph} = \kappa h_{ls} \tag{12}$$

in which  $\kappa \in [1, 2]$  is the radius adjustment coefficient, and  $h_{ls}$  is the distance between two parallel scanning layers. As shown in Figure 4, the predicted curve with estimated point  $P_{est}$  are used to search for the neighbor point  $P_{nb}$  from the previous scanning line  $i-1$ . The necessary and sufficient conditions for point  $P_{nb}$  as the neighbor point of  $P_{est}$  are  $|\overline{P_{est}P_{nb}}| \leq R_{sph}$ , which means that  $P_{nb}$  is inside the bounding sphere with point  $P_{est}$  as its center. At the very beginning,  $P_{est}$  coincides with  $P_{i,j}$ . At this point, there are two possibilities: (i)  $P_{i-1,u}$  is inside the bounding sphere (i.e.,  $|\overline{P_{i-1,u}P_{i,j}}| \leq R_{sph}$ ), or (ii)  $P_{i-1,u}$  is outside the bounding sphere. In case (i),  $P_{i-1,u}$  is the first found neighbor point. As  $P_{est}$  moves along the scanning direction with a stepping distance of  $\lambda$ , if  $|\overline{P_{est}P_{i-1,u}}| < |\overline{P_{i,j}P_{i-1,u}}|$ , then  $P_{i-1,u}$  is the neighbor point of  $P_{est}$ ; otherwise, discard point  $P_{i-1,u}$ , as it is not the neighbor point of  $P_{est}$  but of  $P_{i,j}$ . In case (ii), there is no operation because no neighbor point has been found. After case (i) or case (ii) is completed, point  $P_{est}$  continues to move forward along the scanning direction until the neighbor point  $P_{nb}$  of  $P_{est}$  is found; if the neighbor point cannot be found, the search is stopped.

If the neighbor point  $P_{nb}$  is found in line  $i-1$  (e.g.,  $P_{i-1,u+1}$  in Figure 4), then a new bounding sphere is built with  $P_{i-1,u+1}$  as the center and the radius is  $R_{sph}$ . After that, we use this new bounding sphere to search for the neighbor point of  $P_{i-1,u+1}$  in line  $i-2$ ; and if the neighbor point cannot be found, we stop searching. Next, we take the new neighbor point in line  $i-2$  (e.g.,  $P_{i-2,v+1}$ ) as a new center to build a bounding sphere and repeat the above process until we find three neighbor points in different scanning lines (e.g.,  $P_{i-1,u+1}$ ,  $P_{i-2,v+1}$ ,  $P_{i-3,w+1}$  in Figure 4).



Based on the neighbor point set  $\{P_{i-1,u+1}, P_{i-2,v+1}, P_{i-3,w+1}\}$ , the coordinates of estimated point  $P_{est}$  can be fixed uniquely. As shown in Figure 4, a bi-cubic Hermite curve is built, and it can be expressed as

$$\begin{cases} H_x(y) = x_{i-2}\alpha_{i-2}(y) + x_{i-1}\alpha_{i-1}(y) + x'_{i-2}\beta_{i-2}(y) + x'_{i-1}\beta_{i-1}(y) \\ H_z(y) = z_{i-2}\alpha_{i-2}(y) + z_{i-1}\alpha_{i-1}(y) + z'_{i-2}\beta_{i-2}(y) + z'_{i-1}\beta_{i-1}(y) \end{cases} \quad (13)$$

in which  $y$  is an independent variable;  $\alpha_{i-1}(y), \alpha_{i-2}(y), \beta_{i-1}(y), \beta_{i-2}(y)$  are obtained by Equation (2);  $x'_{i-2}, z'_{i-2}$  are acquired by Equations (3) and (4); and  $x'_{i-1}, z'_{i-1}$  are obtained by Equations (8)–(10). Obviously, the bicubic Hermite curve must be in the curved surface with the equation

$$H_x(y) = x_{i-2}\alpha_{i-2}(y) + x_{i-1}\alpha_{i-1}(y) + x'_{i-2}\beta_{i-2}(y) + x'_{i-1}\beta_{i-1}(y), \quad (14)$$

and the predicted curve will pass through this curved surface. Therefore, estimated point  $P_{est}$  can be fixed at the intersection of the predicted curve and the curved surface which is described in Equation (14). That is, the coordinates of estimated point  $P_{est}(x_{est}, y_{est}, z_{est})$  can be determined by Equations (11) and (14).

### 2.2. Data Redundancy Elimination

After the coordinates of estimated point  $P_{est}$  are determined, we use  $P_{est}$  to replace  $P_{i,j+1}$  in scanning line  $i$ . Afterwards, the new point set that contains  $P_{est}$  is used for bi-Akima interpolation, and there is a deviation  $h_{i,k}$  between the interpolated curve and each initial sampled point  $Q_k$ , where  $i$  is the scanning line number and  $k$  is the serial number of initial point cloud in line  $i$ . As mentioned earlier, the initial point cloud is obtained by 3D scanning measuring devices using the isochronous or equidistant sampling method in Step 1 as shown in Figure 1. The deviation  $h_{i,k}$  can be obtained by

$$h_{i,k} = \text{MIN}_{x \in (X_j, X_{j+1})} (s) = \text{MIN}_{x \in (X_j, X_{j+1})} \left( \sqrt{(x - x_k)^2 + (y - y_k)^2 + (z - z_k)^2} \right), \quad (15)$$

where point  $Q_k(x_k, y_k, z_k)$  is an initial sampled point between  $P_{i,j}(X_j, Y_j, Z_j)$  and  $P_{i,j+1}(X_{j+1}, Y_{j+1}, Z_{j+1})$ , and  $P_{curv}(x, y, z)$  is the point in interpolated curve that makes the distance  $S$  shortest. Then, the max deviation  $d_{max}$  of the whole curve (i.e., from  $P_{i,1}$  to  $P_{est}$ ) can be calculated by the following formula:

$$d_{max} = \text{MAX}(h_{i,k}), \quad (16)$$

which is compared with the required accuracy  $\epsilon$ . If  $d_{max} > \epsilon$ , discard point  $P_{est}$ . If  $d_{max} < \epsilon$ , delete current compressed point  $P_{i,j+1}$  which is input from Step 2. Next, create an estimative flag  $F_{i,j+1} = 1$  to replace point  $P_{i,j+1}$ . This flag takes up only one bit of data storage space. After completing the above process, output the final compressed point cloud data flow, which contains the point coordinate and estimative flag information to the data storage devices. Afterwards, make  $j = j + 1$ , build a new shape-preserving piecewise bicubic Hermite curve to predict the shape and direction of the current scanning line, and create a new estimative point  $P_{est}$  to loop through the above data redundancy identification and elimination process until  $P_{i,j}$  is the end point of the current scanning line  $i$  or the data sampling is over. In addition, when  $P_{i,j}$  is the end point of line  $i$ , make  $i = i + 1$  and continue to loop the above algorithm until the measurement is completed.

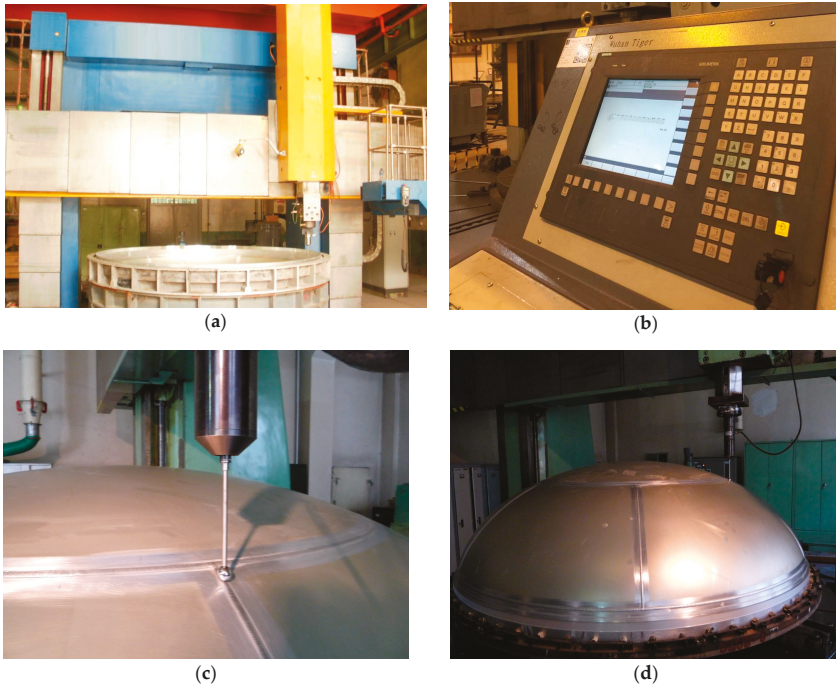
## 3. Experimental Results

In order to verify the feasibility of the proposed methodology, some experiments were performed in this section.

### 3.1. Test A

The on-line point cloud data compression algorithm was tested in the industrial real-time measuring process and compared with existing methods (chordal method and bi-Akima method).

The measuring system consists of a contact 3D scanning probe, a vertical lathe and a commercial computer numerical control (CNC) system of SINUMERIK 840D (Munich, Bayern, Germany) as shown in Figure 5. The proposed algorithm is integrated in the original equipment manufacturer (OEM) application that runs on the host computer of the CNC system. The product model of the contact 3D scanning probe is DIGIT-02 (Dalian, Liaoning Province, China). More detailed technical characteristics of the measuring instrument are shown in Table 1.



**Figure 5.** The measuring system and measured 3D free-form surface: (a) vertical lathes; (b) computer numerical control (CNC) system; (c) scanning probe; (d) half-ellipsoidal measured part.

**Table 1.** Detailed technical characteristics of the measuring system.

Technical Characteristics	Values
Scope of X axis	2400 mm
Positioning accuracy of X axis	0.019 mm/1000 mm
Repeatability of X axis	0.016 mm/1000 mm
Scope of Z axis	1200 mm
Positioning accuracy of Z axis	0.010 mm/1000 mm
Repeatability of Z axis	0.003 mm/1000 mm
Positioning accuracy of C axis	6.05"
Repeatability of C axis	2.22"
Measuring range of scanning probe	±1 mm
Accuracy of scanning probe	±8 μm
Repeatability of scanning probe	±4 μm
Stylus length of probe	100 mm/150 mm/200 mm
Contact force (with stylus of 200 mm)	1.6 N/mm
Weight of scanning probe	1.8 kg

The measured part is a half-ellipsoidal surface which is welded together by seven pieces of thin-walled aluminum alloy sheet, as shown in Figure 5d, with a semi-major axis of 1450 mm and

semi-minor axis of 950 mm. A rotational progressive scanning mode is adopted, and the layer spacing is 7 mm. Figure 6 shows the spatial distribution of the initial point cloud data. The isochronous sampling method is adopted and the number of initial sampling points is 272,638.

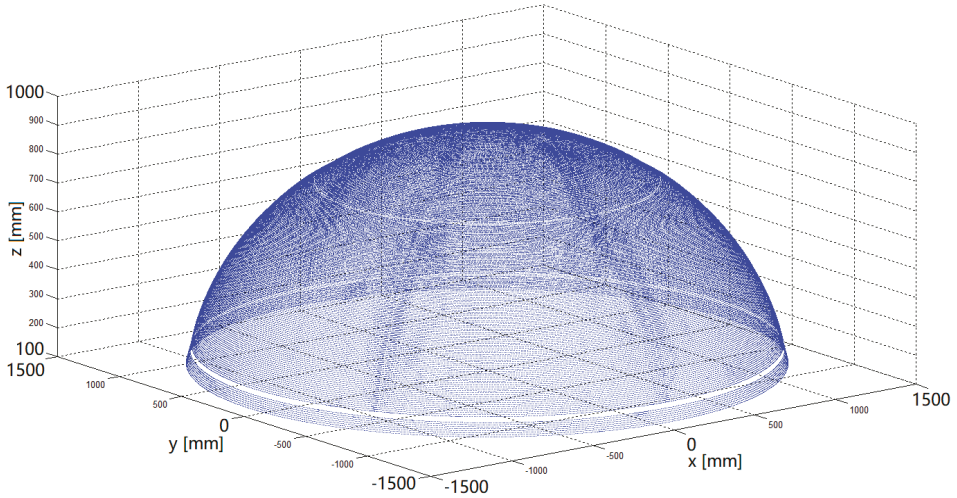


Figure 6. Spatial distribution of initial point cloud data.

Using the same initial point cloud data set as shown in Figure 6, the comparison of data compression performance is made between the proposed method, chordal method and bi-Akima method under different required accuracies (i.e., from 0.001 mm to 1 mm). Table 2 summarizes the results of the data compression performance including the number of points and data compression ratio, where the compression ratio is defined as the ratio between the uncompressed size and compressed size:

$$\text{Compression Ratio} = \frac{\text{Uncompressed Size}}{\text{Compressed Size}} = \frac{\text{Number of Initial Points}}{\text{Number of Compressed Points}} \quad (17)$$

Obviously, the proposed method has a higher data compression ratio than the chordal method and bi-Akima method, and the chordal method obtains the lowest data compression ratio under the same required accuracy. The number of data points obtained by the proposed method is about half of that obtained by the bi-Akima method under the same required accuracy.

Table 2. Compression performance under different required accuracies.

Required Accuracy (mm)	Number of Points			Compression Ratio		
	Chordal Method	Bi-Akima Method	Proposed Method	Chordal Method	Bi-Akima Method	Proposed Method
0.001	237,363	122,929	67,448	1.15	2.22	4.04
0.002	189,824	120,952	67,121	1.44	2.25	4.06
0.005	152,674	110,175	63,813	1.79	2.47	4.27
0.01	136,027	93,588	51,062	2.00	2.91	5.34
0.02	123,891	71,629	41,862	2.20	3.81	6.51
0.05	103,205	44,072	28,837	2.64	6.19	9.45
0.1	87,008	27,894	15,974	3.13	9.77	17.07
0.2	61,124	12,191	7102	4.46	22.36	38.39
0.5	28,473	5594	3140	9.58	48.74	86.83
1	9029	3969	2217	30.20	68.69	122.99

Figure 7 provides the comparison of the compression ratios between the three methods under the different required accuracies. With the decrease in accuracy requirements, the compression ratio increases for all methods; however, for all levels of required accuracy, our proposed compression method manifests a superior compression ratio than the other two methods. Obviously, the chordal method has the lowest data compression ratio. Therefore, we focus on comparing our proposed method with the bi-Akima method in the subsequent experiments.

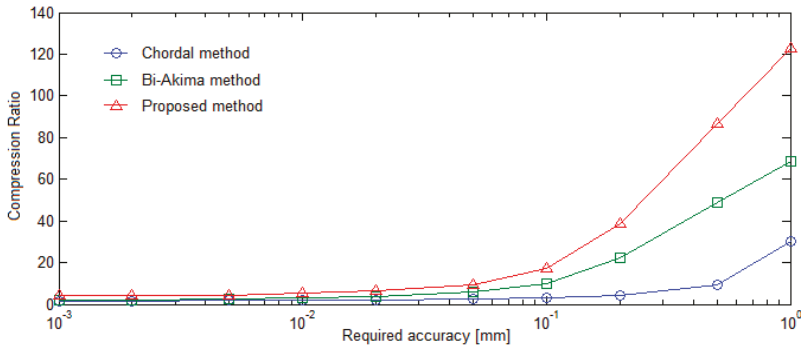
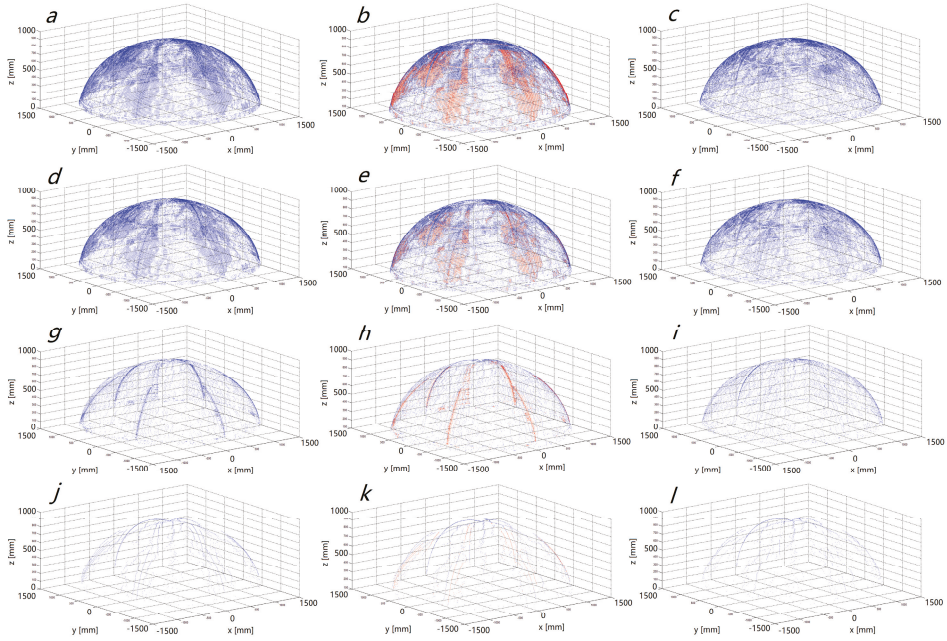


Figure 7. Data compression ratios under different required accuracies.

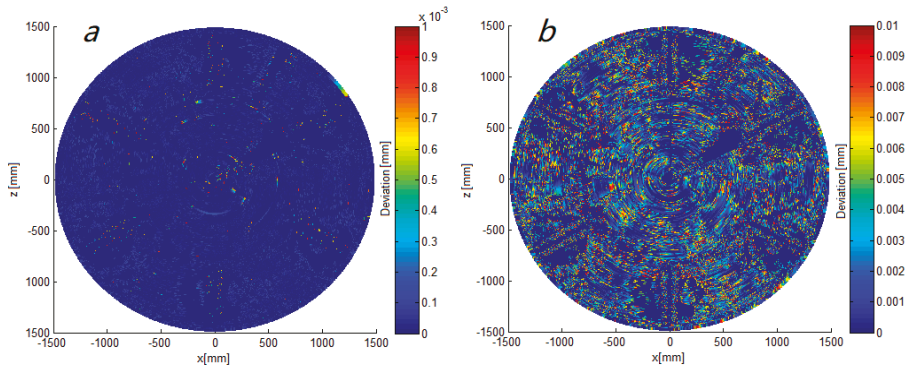
To make the comparison more vivid and intuitive, Figure 8 visually illustrates the difference between the proposed method and bi-Akima method by displaying spatial distributions of compressed point sets under different required accuracies. Subfigures a, d, g and j show the point cloud distribution compressed by bi-Akima method while subfigures b, e, h and k give the point cloud distribution after data redundancy identification by the proposed method, with the identified redundant points marked in red. In subfigures c, f, i and l, the identified redundant points are eliminated. These subfigures show the distributions of the final compressed point cloud data. By contrast, we can clearly observe the difference of point cloud density between these two methods under the same required accuracy. Take subfigures g–i, for example: when using the bi-Akima method, we can observe that there are many curves roughly along the welded region (Figure 8g), because the bi-Akima method can only deal with the point set in the current scanning line and the data redundancy outside the current scanning line cannot be eliminated. With the involvement of our proposed method, redundant data points are identified and marked in red (Figure 8h) and the data redundancy in the adjacent scanning layers is eliminated and the final compressed point cloud data is obtained (Figure 8i).

To verify the accuracy of the proposed algorithm, Figure 9 analyzes the spatial distribution of deviation between each initial sampled point and the interpolated surface obtained from the final compressed point cloud data under different required accuracies. As can be seen, all the deviations are within the allowable range of required accuracy. Our method can tightly control the deviation within the error tolerance range (i.e., the deviation between each initial sampled point and interpolation curve is less than or equal to the required accuracy). In addition, deviations are far lower than the required accuracy in most of the measured area. In Figure 9d, there is an interesting and noteworthy phenomenon: the upper right sector has a higher deviation. As mentioned earlier, the measured part is a large thin-walled surface which is welded together by seven pieces of aluminum alloy sheet (Figure 5d). The aluminum alloy sheet has a thickness of only 0.8 mm, but its size is very large (the semi-major axis of the ellipse is 1450 mm). The part has undergone great deformation after welding. There is a large and random deviation between each welded part and the original design size. According to past experience, the maximum deviation in a local section can even reach 3 mm. Consequently, we infer that the upper right sector has a higher deviation because of deformation in this area. In the case where the required accuracy is on the order of millimeters (e.g., required accuracy

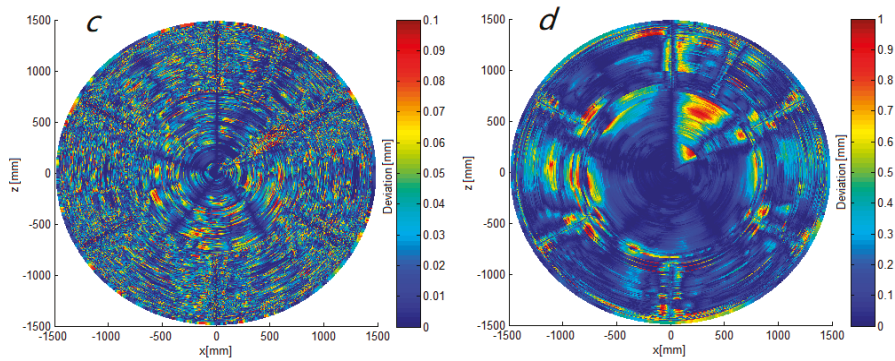
$\epsilon = 1$  mm in Figure 9d), the compressed point cloud data is very sparse. Therefore, this phenomenon is formed in a region where the point cloud density is low and the local deformation is large. However, in any case, the proposed method can tightly control the deviation within the preset range.



**Figure 8.** Spatial distributions of compressed point cloud data under different required accuracies  $\epsilon$ : (a) bi-Akima compression,  $\epsilon = 0.001$  mm; (b) redundancy identification,  $\epsilon = 0.001$  mm; (c) redundancy elimination,  $\epsilon = 0.001$  mm; (d) bi-Akima compression,  $\epsilon = 0.01$  mm; (e) redundancy identification,  $\epsilon = 0.01$  mm; (f) redundancy elimination,  $\epsilon = 0.01$  mm; (g) bi-Akima compression,  $\epsilon = 0.1$  mm; (h) redundancy identification,  $\epsilon = 0.1$  mm; (i) redundancy elimination,  $\epsilon = 0.1$  mm; (j) bi-Akima compression,  $\epsilon = 1$  mm; (k) redundancy identification,  $\epsilon = 1$  mm; (l) redundancy elimination,  $\epsilon = 1$  mm.



**Figure 9.** Cont.



**Figure 9.** Spatial distributions of deviation under different required accuracies  $\epsilon$ : (a)  $\epsilon = 0.001$  mm; (b)  $\epsilon = 0.01$  mm; (c)  $\epsilon = 0.1$  mm; (d)  $\epsilon = 1$  mm.

### 3.2. Test B

The overall structure of the model in Test A is relatively simple. In order to further verify the universality and adaptability of the proposed method, we chose a more complex surface model with a large number of details, edges and sharp features for experimentation. As shown in Figure 10, the tested model is a piece of jewelry, which is inlaid with 30 diamonds of different sizes.

Figure 11 shows the initial point cloud data acquisition result. The progressive scanning mode and equidistant sampling mode were adopted. Scanning lines are along the X-direction (horizontal direction). The distance between two adjacent scanning layers is 0.1 mm, and the distance between adjacent points is 0.05 mm in each scanning layer. The initial point number is 63,376.

The comparison is made between the proposed method and bi-Akima method under different required accuracies (i.e., from 0.001 mm to 1 mm). Table 3 gives the results of data compression performance, including the number of points and data compression ratio. Obviously, the proposed method has a higher data compression ratio than the bi-Akima method. The number of points obtained by the proposed method is about half of that obtained by bi-Akima method under the same required accuracy.

Figure 12 provides the comparison of the compression ratios between these two methods under different required accuracies. With the decrease in accuracy requirements, the compression ratio increases for all methods; however, for all levels of required accuracy, our proposed compression method manifests a superior compression ratio than the bi-Akima method.



**Figure 10.** The tested complex surface model: jewelry.

Figure 13 visually illustrates the difference between the proposed method and bi-Akima method by displaying spatial distributions of the compressed point sets under different required accuracies. Subfigures a, d, g and j show the point cloud distribution compressed by the bi-Akima method,

while subfigures b, e, h and k give the point cloud distribution after data redundancy identification by the proposed method, with the identified redundant points marked in red. In subfigures c, f, i and l, the identified redundant points are eliminated. These subfigures show the distributions of the final compressed point cloud data. By contrast, we can clearly observe the difference in point cloud density between these two methods under the same required accuracy. Take subfigures j, k and l, for example: when using the bi-Akima method, we can observe that there are many curves roughly along the vertical direction (Figure 13j). This is because the bi-Akima method can only deal with the point set in the current single scanning line, which is along the horizontal direction, and the data redundancy outside the current scanning line cannot be eliminated. With the involvement of our proposed method, redundant data points are identified and marked in red (Figure 13k), the data redundancy in adjacent scanning layers is eliminated and the final compressed point cloud data is obtained (Figure 13l).

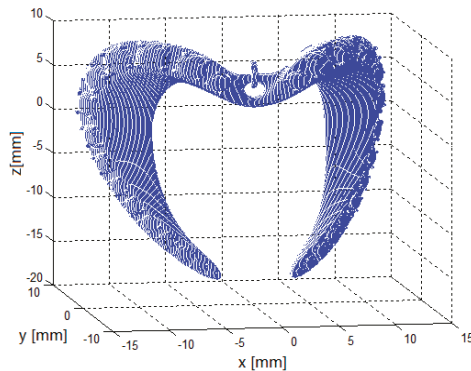


Figure 11. Spatial distribution of initial point cloud data.

Table 3. Compression performance under different required accuracies.

Required Accuracy (mm)	Number of Points		Compression Ratio	
	Bi-Akima Method	Proposed Method	Bi-Akima Method	Proposed Method
0.001	18,906	8516	3.35	7.44
0.002	16,857	7609	3.76	8.33
0.005	14,323	6563	4.42	9.66
0.01	12,432	5743	5.10	11.04
0.02	10,720	5007	5.91	12.66
0.05	8767	4232	7.23	14.98
0.1	7190	3535	8.81	17.93
0.2	5892	2974	10.76	21.31
0.5	4625	2412	13.70	26.28
1	4204	2213	15.08	28.64

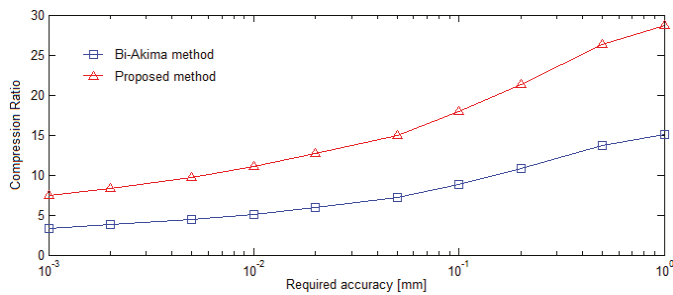
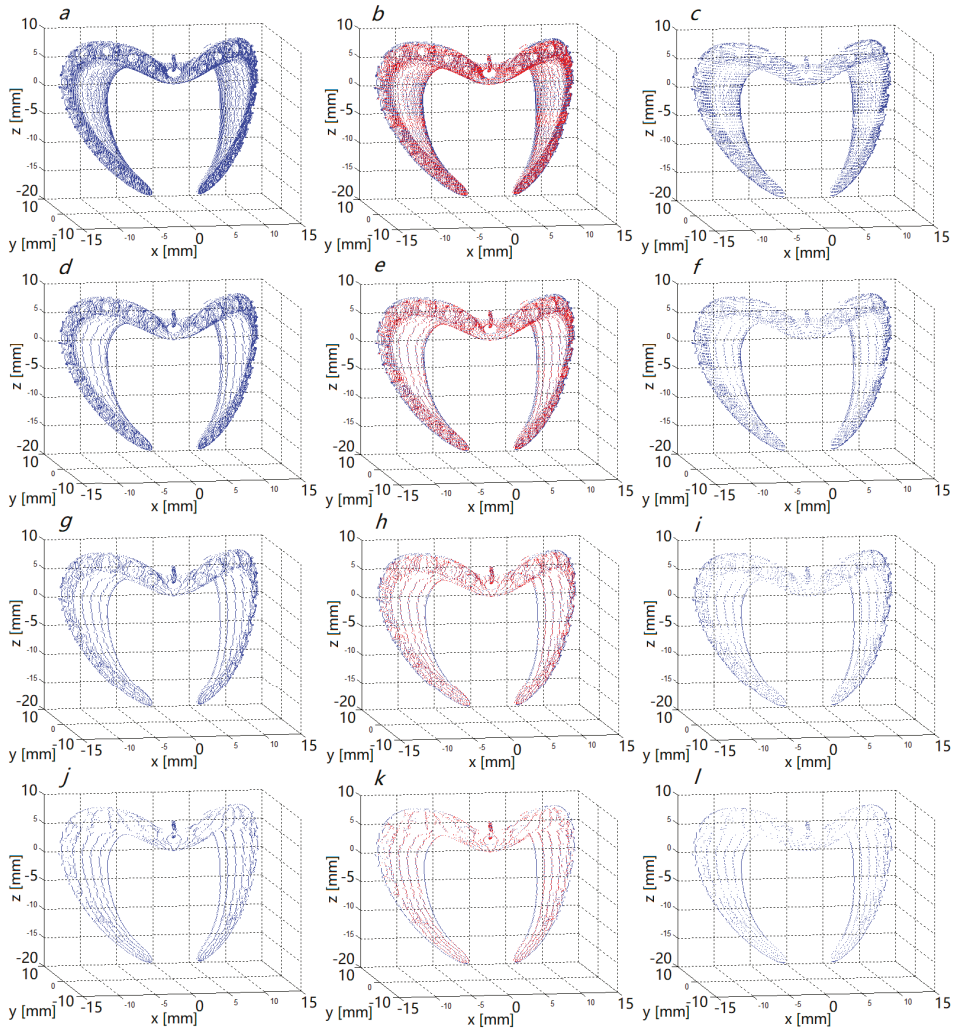


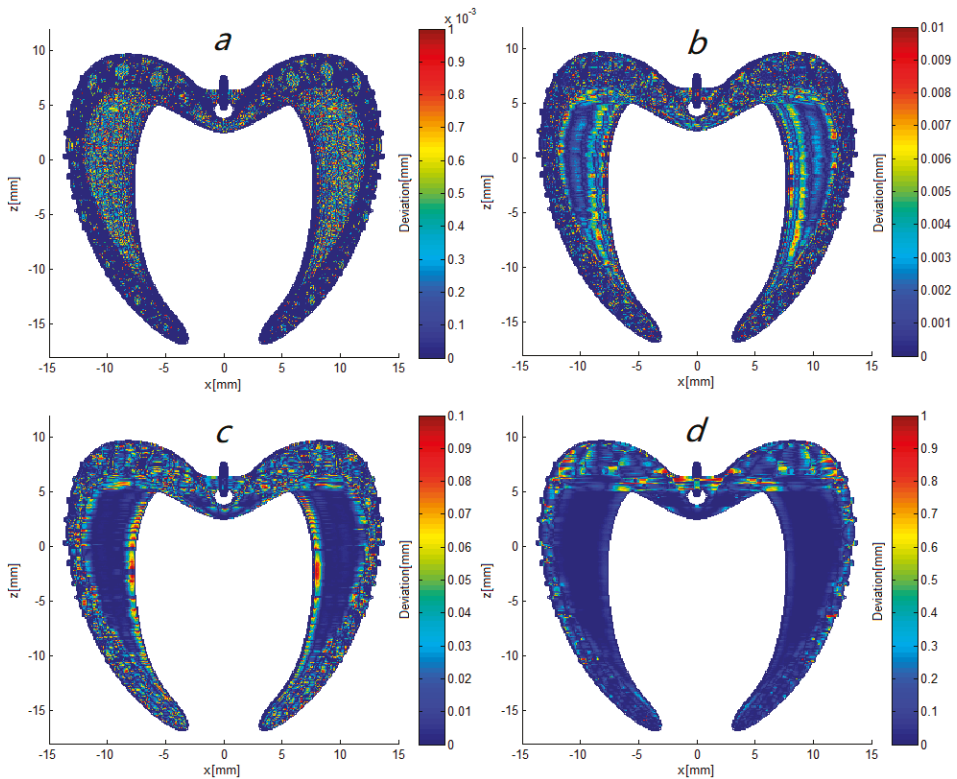
Figure 12. Data compression ratios under different required accuracies.



**Figure 13.** Spatial distributions of compressed point cloud data under different required accuracies  $\epsilon$ : (a) bi-Akima compression,  $\epsilon = 0.001$  mm; (b) redundancy identification,  $\epsilon = 0.001$  mm; (c) redundancy elimination,  $\epsilon = 0.001$  mm; (d) bi-Akima compression,  $\epsilon = 0.01$  mm; (e) redundancy identification,  $\epsilon = 0.01$  mm; (f) redundancy elimination,  $\epsilon = 0.01$  mm; (g) bi-Akima compression,  $\epsilon = 0.1$  mm; (h) redundancy identification,  $\epsilon = 0.1$  mm; (i) redundancy elimination,  $\epsilon = 0.1$  mm; (j) bi-Akima compression,  $\epsilon = 1$  mm; (k) redundancy identification,  $\epsilon = 1$  mm; (l) redundancy elimination,  $\epsilon = 1$  mm.

In order to verify the accuracy of the proposed algorithm, Figure 14 analyzes the spatial distribution of deviation between each initial sampled point and the interpolated surface obtained from the final compressed point cloud data under different required accuracies. As can be seen, all the deviations are within the allowable range of required accuracy, which proves that the proposed method can tightly control the deviation within the error tolerance range (i.e., the deviation between each initial sampled point and interpolation curve is less than or equal to the required accuracy). In addition, deviations are far less than the required accuracy in most of the measured area.





**Figure 14.** Spatial distributions of deviation under different required accuracies  $\epsilon$ : (a)  $\epsilon = 0.001$  mm; (b)  $\epsilon = 0.01$  mm; (c)  $\epsilon = 0.1$  mm; (d)  $\epsilon = 1$  mm.

#### 4. Discussion

The experimental results in Section 3 indicate that the proposed on-line point cloud data compression algorithm for free-form surface scanning measurement has the following features:

- It can further compress point cloud data and obtain a higher data compression ratio than the existing methods under the same required accuracy. Its compression performance is obviously superior to the bi-Akima and chordal methods;
- It is capable of tightly controlling the deviation within the error tolerance range, and deviations in most measured area are far less than the required accuracy;
- Test A preliminarily verifies the application feasibility of the proposed method in an industrial environment. Test B demonstrates that the method is equally effective for complex surfaces with a large number of details, edges and sharp features, and it has stable performance;
- The proposed method has the potential to be applied to industrial environments to replace traditional on-line point cloud data compression methods (bi-Akima and chordal methods). Its potential applications may be in the real-time measurement processes of scanning devices such as contact scanning probes, laser triangle displacement sensors, mobile laser scanners, linear structured light systems, industrial CT systems, etc. The application feasibility of this method needs to be further confirmed in subsequent case studies.

However, the proposed method is not perfect and still has the following limitations. In future work, the following aspects need to be further developed:

- This method can only handle 3D point cloud data streams and is not suitable for processing point cloud data containing additional high-dimensional information (e.g., 3D point cloud data with grayscale or color information). We will try to solve the above problem in our future research work;
- This method can only compress the point cloud data stream which is scanned layer by layer. If the 3D point cloud is randomly sampled and there are no regular scan lines (e.g., 3D measurement with speckle-structure light), our method cannot perform effective data compression. It is a huge challenge to solve the above problems.

## 5. Conclusions

In an attempt to effectively compress dense point cloud data obtained from a 3D free-form surface during the real-time scanning measuring process, this paper presents a novel on-line point cloud data compression algorithm which has the ability to identify and eliminate data redundancy caused by geometric feature similarity between adjacent scanning layers. At first, the new algorithm adopts the bi-Akima method to compress the initial point cloud data obtained by 3D scanning measuring devices. Next, the data redundancy in the compressed point cloud obtained in the previous stage is further identified and eliminated, and then we can obtain the final compressed point cloud data. Finally, the proposed on-line point cloud data compression algorithm was tested in the real-time scanning measuring process and compared with existing methods (the chordal method and bi-Akima method). The experimental results have preliminarily verified the application feasibility of our proposed method in industrial environment, and shown that it is capable of obtaining high-quality compressed point cloud data with a higher compression ratio than other existing methods. In particular, it can tightly control the deviation within the error tolerance range, which demonstrates the superior performance of the proposed algorithm. This algorithm could be used in the data acquisition process of 3D free-form surface scanning measurement to replace other existing on-line point cloud data compression/reduction methods.

**Author Contributions:** All work with relation to this paper has been accomplished by the efforts of all authors. Conceptualization, Y.L. and Y.T.; methodology, Z.H.; software, Y.T.; validation, Y.M. and Z.H.; formal analysis, Y.T.; investigation, Y.M.; resources, Y.M.; data curation, Y.M.; writing—original draft preparation, Y.T.; writing—review and editing, Y.T.; visualization, Z.H.; supervision, Y.L.; project administration, Y.L.; funding acquisition, Y.L.

**Funding:** This research was funded by the National Natural Science Foundation of China (Grant Nos. 51505310, 51435011), the Key Research and Development Program of Sichuan Province of China (Grant No. 2018GZ0282) and the Key Laboratory for Precision and Non-traditional Machining of Ministry of Education, Dalian University of Technology (Grant Nos. JMTZ201802, B201802).

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Galetto, M.; Vezzetti, E. Reverse engineering of free-form surfaces: A methodology for threshold definition in selective sampling. *J. Mach. Tools Manuf.* **2006**, *46*, 1079–1086. [[CrossRef](#)]
2. Han, Z.H.; Wang, Y.M.; Ma, X.H.; Liu, S.G.; Zhang, X.D.; Zhang, G.X. T-spline based unifying registration procedure for free-form surface workpieces in intelligent CMM. *Appl. Sci.* **2017**, *7*, 1092. [[CrossRef](#)]
3. Ngo, T.D.; Kashani, A.; Imbalzano, G.; Nguyen, K.T.Q.; Hui, D. Additive manufacturing (3D printing): A review of materials, methods, applications and challenges. *Compos. Pt. B Eng.* **2018**, *143*, 172–196. [[CrossRef](#)]
4. Liu, J.; Bai, D.; Chen, L. 3-D point cloud registration algorithm based on greedy projection triangulation. *Appl. Sci.* **2018**, *8*, 1776. [[CrossRef](#)]
5. Chen, L.; Jiang, Z.D.; Li, B.; Ding, J.J.; Zhang, F. Data reduction based on bi-directional point cloud slicing for reverse engineering. *Key Eng. Mater.* **2010**, *437*, 492–496. [[CrossRef](#)]
6. Budak, I.; Hodolic, J.; Sokovic, M. Development of a programme system for data-point pre-processing in Reverse Engineering. *J. Mater. Process. Technol.* **2005**, *162*, 730–735. [[CrossRef](#)]

7. Yan, R.J.; Wu, J.; Lee, J.Y.; Khan, A.M.; Han, C.S.; Kayacan, E.; Chen, I.M. A novel method for 3D reconstruction: Division and merging of overlapping B-spline surfaces. *Comput. Aided Des.* **2016**, *81*, 14–23. [[CrossRef](#)]
8. Pal, P.; Ballav, R. Object shape reconstruction through NURBS surface interpolation. *Int. J. Prod. Res.* **2007**, *45*, 287–307. [[CrossRef](#)]
9. Cali, M.; Ambu, R. Advanced 3D Photogrammetric Surface Reconstruction of Extensive Objects by UAV Camera Image Acquisition. *Sensors* **2018**, *18*, 2815. [[CrossRef](#)]
10. Zanetti, E.; Aldieri, A.; Terzini, M.; Cali, M.; Franceschini, G.; Bignardi, C. Additively manufactured custom load-bearing implantable devices. *Australas. Med. J.* **2017**, *10*. [[CrossRef](#)]
11. Cavas-Martinez, F.; Fernandez-Pacheco, D.G.; Canavate, F.J.F.; Velazquez-Blazquez, J.S.; Bolarin, J.M.; Alio, J.L. Study of Morpho-Geometric Variables to Improve the Diagnosis in Keratoconus with Mild Visual Limitation. *Symmetry* **2018**, *10*, 306. [[CrossRef](#)]
12. Manavella, V.; Romano, F.; Garrone, F.; Terzini, M.; Bignardi, C.; Aimetti, M. A novel image processing technique for 3D volumetric analysis of severely resorbed alveolar sockets with CBCT. *Minerva Stomatol.* **2017**, *66*, 81–89. [[CrossRef](#)] [[PubMed](#)]
13. Aldieri, A.; Terzini, M.; Osella, G.; Priola, A.M.; Angeli, A.; Veltri, A.; Audenino, A.L.; Bignardi, C. Osteoporotic Hip Fracture Prediction: Is T-Score-Based Criterion Enough? A Hip Structural Analysis-Based Model. *J. Biomech. Eng. Trans. ASME* **2018**, *140*, 111004. [[CrossRef](#)] [[PubMed](#)]
14. Jia, Z.Y.; Lu, X.H.; Yang, J.Y. Self-learning fuzzy control of scan tracking measurement in copying manufacture. *Trans. Inst. Meas. Control* **2010**, *32*, 307–318. [[CrossRef](#)]
15. Wang, Y.Q.; Tao, Y.; Nie, B.; Liu, H.B. Optimal design of motion control for scan tracking measurement: A CMAC approach. *Measurement* **2013**, *46*, 384–392. [[CrossRef](#)]
16. Li, W.L.; Zhou, L.P.; Yan, S.J. A case study of blade inspection based on optical scanning method. *Int. J. Prod. Res.* **2015**, *53*, 2165–2178. [[CrossRef](#)]
17. Khameneifar, F.; Feng, H.Y. Extracting sectional contours from scanned point clouds via adaptive surface projection. *Int. J. Prod. Res.* **2017**, *55*, 4466–4480. [[CrossRef](#)]
18. Budak, I.; Sokovic, M.; Barisic, B. Accuracy improvement of point data reduction with sampling-based methods by Fuzzy logic-based decision-making. *Measurement* **2011**, *44*, 1188–1200. [[CrossRef](#)]
19. Shi, B.Q.; Liang, J.; Liu, Q. Adaptive simplification of point cloud using k-means clustering. *Comput. Aided Des.* **2011**, *43*, 910–922. [[CrossRef](#)]
20. Feng, C.; Taguchi, Y. FasTFit: A fast T-spline fitting algorithm. *Comput. Aided Des.* **2017**, *92*, 11–21. [[CrossRef](#)]
21. Meng, X.L.; He, W.T.; Liu, J.Y. An investigation of the high efficiency estimation approach of the large-scale scattered point cloud normal vector. *Appl. Sci.* **2018**, *8*, 454. [[CrossRef](#)]
22. Song, H.; Feng, H.Y. A progressive point cloud simplification algorithm with preserved sharp edge data. *Int. J. Adv. Manuf. Technol.* **2009**, *45*, 583–592. [[CrossRef](#)]
23. Chen, L.C.; Hoang, D.C.; Lin, H.I.; Nguyen, T.H. Innovative methodology for multi-view point cloud registration in robotic 3D object scanning and reconstruction. *Appl. Sci.* **2016**, *6*, 132. [[CrossRef](#)]
24. Macher, H.; Landes, T.; Grussenmeyer, P. From point clouds to building information models: 3D semi-automatic reconstruction of indoors of existing buildings. *Appl. Sci.* **2017**, *7*, 1030. [[CrossRef](#)]
25. Han, H.; Han, X.; Sun, F.; Huang, C. Point cloud simplification with preserved edge based on normal vector. *Optik* **2015**, *126*, 2157–2162. [[CrossRef](#)]
26. Wang, Y.Q.; Tao, Y.; Zhang, H.J.; Sun, S.S. A simple point cloud data reduction method based on Akima spline interpolation for digital copying manufacture. *Int. J. Adv. Manuf. Technol.* **2013**, *69*, 2149–2159. [[CrossRef](#)]
27. Arpaia, P.; Buzio, M.; Inglese, V. A two-domain real-time algorithm for optimal data reduction: A case study on accelerator magnet measurements. *Meas. Sci. Technol.* **2010**, *21*. [[CrossRef](#)]
28. Wang, D.; He, C.; Li, X.; Peng, J. Progressive point set surface compression based on planar reflective symmetry analysis. *Comput. Aided Des.* **2015**, *58*, 34–42. [[CrossRef](#)]
29. Lee, K.H.; Woo, H.; Suk, T. Data reduction methods for reverse engineering. *Int. J. Adv. Manuf. Technol.* **2001**, *17*, 735–743. [[CrossRef](#)]
30. Ma, X.; Cripps, R.J. Shape preserving data reduction for 3D surface points. *Comput. Aided Des.* **2011**, *43*, 902–909. [[CrossRef](#)]
31. Smith, J.; Petrova, G.; Schaefer, S. Progressive encoding and compression of surfaces generated from point cloud data. *Comput. Graph.* **2012**, *36*, 341–348. [[CrossRef](#)]

32. Morell, V.; Orts, S.; Cazorla, M.; Garcia-Rodriguez, J. Geometric 3D point cloud compression. *Pattern Recognit. Lett.* **2014**, *50*, 55–62. [[CrossRef](#)]
33. Lu, J.C.; Yang, J.K.; Mu, L.C. Automatic tracing measurement and close data collection system of the free-form surfaces. *J. Dalian Univ. Technol.* **1986**, *24*, 55–59. (In Chinese)
34. ElKott, D.F.; Veldhuis, S.C. Isoparametric line sampling for the inspection planning of sculptured surfaces. *Comput. Aided Des.* **2005**, *37*, 189–200. [[CrossRef](#)]
35. Wozniak, A.; Balazinski, A.; Mayer, R. Application of fuzzy knowledge base for corrected measured point determination in coordinate metrology. In Proceedings of the Annual Meeting of the North American Fuzzy Information Processing Society, San Diego, CA, USA, 24–27 June 2007.
36. Jia, Z.Y.; Lu, X.H.; Wang, W.; Yang, J.Y. Data sampling and processing for contact free-form surface scan-tracking measurement. *Int. J. Adv. Manuf. Technol.* **2010**, *46*, 237–251. [[CrossRef](#)]
37. Tao, Y.; Li, Y.; Wang, Y.Q.; Ma, Y.Y. On-line point cloud data extraction algorithm for spatial scanning measurement of irregular surface in copying manufacture. *Int. J. Adv. Manuf. Technol.* **2016**, *87*, 1891–1905. [[CrossRef](#)]
38. Li, R.J.; Fan, K.C.; Huang, Q.X.; Zhou, H.; Gong, E.M. A long-stroke 3D contact scanning probe for micro/nano coordinate measuring machine. *Precis. Eng.* **2016**, *43*, 220–229. [[CrossRef](#)]
39. Wang, Y.Q.; Liu, H.B.; Tao, Y.; Jia, Z.Y. Influence of incident angle on distance detection accuracy of point laser probe with charge-coupled device: Prediction and calibration. *Opt. Eng.* **2012**, *51*, 083606. [[CrossRef](#)]
40. Valkenburg, R.J.; McIvor, A.M. Accurate 3D measurement using a structured light system. *Image Vis. Comput.* **1998**, *16*, 99–110. [[CrossRef](#)]
41. Carmignato, S. Accuracy of industrial computed tomography measurements: Experimental results from an international comparison. *CIRP Ann. Manuf. Technol.* **2012**, *61*, 491–494. [[CrossRef](#)]
42. Lamberty, A.; Schimmel, H.; Pauwels, J. The study of the stability of reference materials by isochronous measurements. *Anal. Bioanal. Chem.* **1998**, *360*, 359–361. [[CrossRef](#)]
43. Li, W.D.; Zhou, H.X.; Hong, W. A Hermite inter/extrapolation scheme for MoM matrices over a frequency band. *IEEE Antennas Wirel. Propag. Lett.* **2009**, *8*, 782–785. [[CrossRef](#)]



© 2018 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).

Article

# Novel Boundary Edge Detection for Accurate 3D Surface Profilometry Using Digital Image Correlation

Liang-Chia Chen <sup>1,\*</sup> and Ching-Wen Liang <sup>2</sup> 

<sup>1</sup> Department of Mechanical Engineering, National Taiwan University, No. 1, Sec. 4, Roosevelt Rd., Taipei 10617, Taiwan

<sup>2</sup> National Chung-Shan Institute of Science and Technology, No. 481, 6th Neighborhood, Sec. Jia'an, Zhongzheng Rd., Longtan Dist., Taoyuan City 32546, Taiwan; maghuang@ms5.hinet.net

\* Correspondence: lchen@ntu.edu.tw

Received: 27 September 2018; Accepted: 30 November 2018; Published: 7 December 2018



**Abstract:** Digital image correlation (DIC) has emerged as a popular full-field surface profiling technique for analyzing both in-plane and out-of-plane dynamic structures. However, conventional DIC-based surface 3D profilometry often yields erroneous contours along surface edges. Boundary edge detection remains one of the key issues in DIC because a discontinuous surface edge cannot be detected due to optical diffraction and height ambiguity. To resolve the ambiguity of edge measurement in optical surface profilometry, this study develops a novel edge detection approach that incorporates a new algorithm using both the boundary subset and corner subset for accurate edge reconstruction. A pre-calibrated gauge block and a circle target were reconstructed to prove the feasibility of the proposed approach. Experiments on industrial objects with various surface reflective characteristics were also conducted. The results showed that the developed method achieved a 15-fold improvement in detection accuracy, with measurement error controlled within 1%.

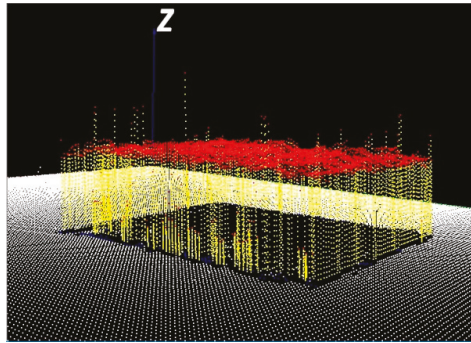
**Keywords:** digital image correlation (DIC); edge detection; random speckle images; surface profilometry; automated optical inspection (AOI)

## 1. Introduction

Nowadays, optical surface profilometry has a broad range of applications, especially for the measurement of critical dimensions and the process automation of industrial workpieces. Three-dimensional surface profilometry is a crucial technology for the quality inspection of various manufacturing processes. However, one of the significant challenges in digital image correlation (DIC)-based profilometry is surface edge detection. Figure 1 shows the depth image of a square box with four surface line edges detected using traditional DIC. As can be seen, the measurement and reconstruction of the edges are far from accurate [1], indicating that the current DIC-based approach cannot accurately detect and reconstruct 3D surface edges. Such inaccuracy also suggests that pattern cross-correlation in DIC is not effective in determining the geometric information of surface edges. Conventional DIC-based surface profilometry often suffers poor accuracy due to noisy data along surface edges.

The commonly used structured light projection techniques include phase-shifting, structured-code, random-speckle, and digital image correlation (DIC) detection approaches. This article focuses on the 3D DIC-based random-speckle processing method, which is an efficient full-field on-shot measurement technique, and can provide absolute depth information. To achieve single-exposure full-field surface profilometry, the 3D DIC-based random-speckle methodology projects random laser speckles on the tested surface and evaluates absolute depth through correlating the captured image with pre-calibrated reference images [2,3]. In such a process, a set of image blocks with unique

deformed random-speckle patterns is acquired to correlate mathematically with the images in the pre-calibrated database. In random-speckle profilometry, small lens apertures are often employed to obtain high depth-of-field (DOF) measurements. However, both depth resolution and accuracy are often undermined due to increased depth range.



**Figure 1.** Depth image of a square box with four line edges detected using the traditional Digital image correlation (DIC) method.

For precise 3D absolute measurement, a fundamental problem is the random variation of light intensity in each small sub-area of the projected random-speckle image without a repeating pattern. Current studies on DIC detection accuracy aim at the optimization of measuring parameters, enhanced quality of random-speckle patterns, improvement in speckle-generating methods, and speckle image-processing algorithms. Measuring parameters including the size, shape, and contrast of laser speckles, as well as the correlation subset size, are closely related to speckle quality [4,5]. In general, a larger speckle size and correlation subset can obtain a higher contrast in speckle quality which can diminish the measured noise but increase the calculation time in DIC. Some speckle quality indexes were developed to determine the quality of a speckle pattern and the optimal subset size for accurate and efficient DIC operation [6,7]. In 2013, Crammond pointed out potential limitations on contrast-type criteria that fail to provide a reliable indication for the randomness of the speckle pattern [8]. Speckle size distribution and speckle shapes are regarded as important local unique features for achieving a steady DIC. For better accuracy of 3D DIC-based measurements, different strategies such as ground-glass laser emission, white-paint spray, or computer simulation have been developed to generate random image features to enhance the uniqueness of speckle patterns [9–11]. Moreover, some speckle image-processing algorithms have been proposed to optimize the signal-to-noise ratio in cross-correlation. Some approaches, utilizing manifold reference image sets or through implementing correlation peak detection, overcome discontinuous displacement with a split subset [12,13]. Other methods use either adaptive binarization or reliable-point matching to enhance the image contrast to improving measurement accuracy [14]. Furthermore, a feature speckle was extracted by considering statistical binary characteristics, and a graphical model-based scheme was developed for refining disparity maps to enhance measurement accuracy [15]. The accurate reconstruction of sharp surface edges has been achieved using an innovative edge detection algorithm with structured illumination microscopy (SIM) [16]. Nevertheless, 3D surface boundary edge detection remains a major challenge for DIC-based surface measurement. An initial concept was developed to improve the accuracy of the critical dimension by detecting surface edges [17]. However, the full algorithm for accurate edge detection has not yet been developed, nor has the feasibility of current measurement approaches in industrial applications been carefully evaluated. In view of this, a new speckle image-processing algorithm that uses a boundary subset and a corner subset for accurate reconstruction of surface boundary edges was developed, and its applicability to measuring industrial objects was assessed.

## 2. Methodology of the Random-Speckle Optical Detection and Processing Algorithm

### 2.1. Principle of Digital Imaging Correlation (DIC)

The laser triangulation principle with random-speckle projection is commonly used for the depth detection of 3D tested surface. Optical elements for generating laser random-speckles include an infrared laser, micro-lens arrays, and a diffuser. Figure 2 shows the optical configuration of the 3D surface profilometry, in which a near-infrared (NIR) laser is used for random-speckle projection. An imaging sensor is employed to detect the shifted phase corresponding to the difference in optical paths formed by the detected surface depth.

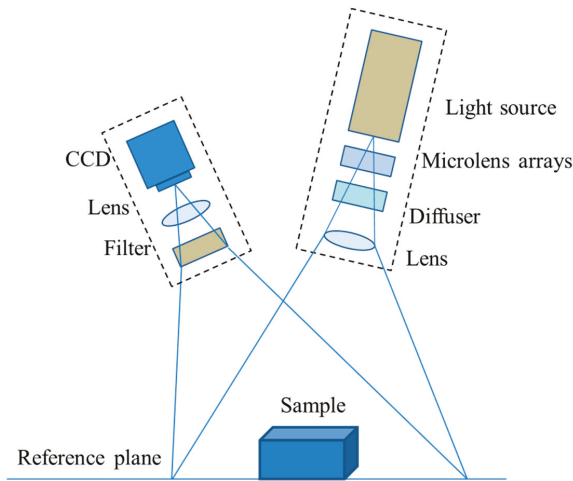


Figure 2. Optical configuration for 3D surface profilometry.

To generate random speckles of various intensities without a repeating pattern, the NIR laser was used as a light source to reduce interference from environmental light. A large grit ground glass served as a diffuser to produce high-contrast grey-level patterns, while a low dark current charge-coupled device (CCD) camera captured the correct pixel intensity.

The DIC principle involves cross-correlating the surface depth detected using a deformed speckle image with the corresponding depth of the reference image in the stored database. Each random-speckle image is unique in its database; hence, the deformed pattern with a specific depth can be matched with its reference image. The surface depth of each image block is thus transformed into the surface depth of the object by triangulation transformation.

### 2.2. Normalized Cross-Correlation (NCC) for Edge Detection

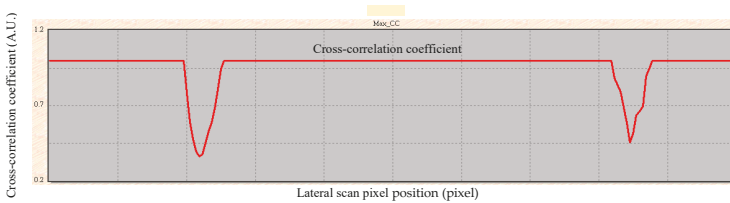
The normalized cross-correlation coefficient can be defined in Equation (1) [17,18] and is commonly taken as a matching index for the correlation between the reference image subset and the measured image subset.

$$C(X, Y) = \frac{\sum_{x,y} T'(x, y, H) \times R'(x, y, t)}{\sqrt{\sum_{x,y} T'(x, y, H)^2 \times R'(x, y, t)^2}} \quad (1)$$

where  $T'(x, y, H) = T(x, y, H) - \frac{\sum_{x,y} T(x, y, H)}{w \times h}$ ,  $R'(x, y, t) = R(x, y, t) - \frac{\sum_{x,y} R(x, y, t)}{w \times h}$ ,  $H$  is the corresponding height of the reference image,  $t$  is the duration of time after the reference image is calibrated,  $w$  is the width of the subset, and  $h$  is the height of the subset.

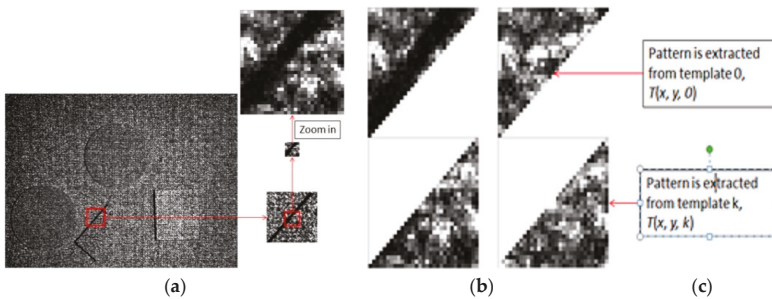
Obviously, the subset size is an important parameter for best cross-correlation in terms of both measurement accuracy and efficiency. The optimal size depends strongly on the image contrast uniqueness of the speckle pattern. In general, the more unique the pattern is, the smaller the subset size required to obtain matching accuracy. Therefore, the uniqueness of the speckle pattern is first determined through cross-correlation between subset images; and a suitable subset window size is established by evaluating the uniqueness of the subset image at various locations of the reference speckle image.

Traditional cross-correlation in DIC involves matching two entire image subsets in both the reference and measured sets. However, this is adequate only for a continuous surface with no contour jump. When there exists a discontinuous surface edge, subset pattern matching between the reference template and the measured image can no longer succeed because the subset random-speckle pattern, under such circumstances, is projected on both sides of the detected surface jump (or edge) with a step-height difference between the projected images. Figure 3 shows how a continuous surface edge is measured by detecting the peak of the cross-correlation coefficient curve along the scanning axis. As can be seen, a lateral scan along the surface edge reveals a decline of the curve to a local minimum, indicating a decrease in cross-correlation between the reference subset image and the measured subset image, followed by a rise back to its original level.



**Figure 3.** Cross-correlation coefficient curve along the lateral scan axis on a rectangular gauge block with two surface edges.

When measuring a geometrically discontinuous surface edge, the measured subset image pattern on the detected edge is affected by edge light diffraction. The measured image pattern comprises two sub-patterns represented by the two split surfaces on either side of the measured edge. Figure 4 shows an example of the above scenario. The measured image pattern marked by a red square comprises the top-left triangular region, the light diffracted region, and the bottom-right triangular region. More importantly, the two triangular subset images correspond to their own reference heights, which represent the measured top-left and bottom-right surface contour regions, respectively.



**Figure 4.** Measured subset image pattern formed by a surface contour edge: (a) measured image of multi objects; (b) a red square indicating measured subset image pattern along a titling contour edge, which comprises the top-left triangular region, the light diffracted region, and the bottom-right triangular region; and (c) two triangular subset images corresponding to their own reference heights.



Next, the edge detection accuracy of two typical subsets was evaluated. Figure 5 shows two typical subsets formed when the template scans along a surface contour edge. The blue square represents a subset image acquired from a continuous surface contour, while the red square shows a subset image scanned on a discontinuous surface with an edge.

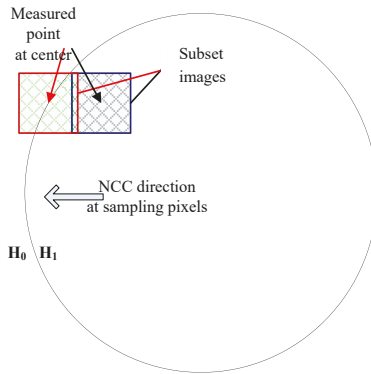


Figure 5. Typical subsets formed when the template scans on a surface with an edge.

Figure 6 shows the two normalized cross-correlation coefficient curves acquired when both blue and red subset images are matched with their corresponding reference templates. It is apparent that the red subset image scanned past an edge did not possess the clear peak that is essential for accurate depth detection. To resolve this problem, two new subsets, namely boundary and corner subsets, were proposed as described in the following section.

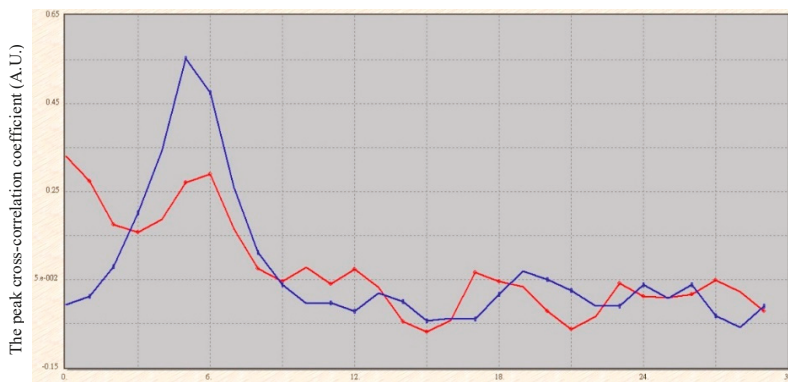


Figure 6. Two normalized cross-correlation coefficient curves acquired when both the blue and red subset images were matched with their corresponding reference templates.

To achieve precise contour measurements for an object with discontinuous surface features, a novel 3D edge detection approach was proposed. The related procedures are presented in the flowchart shown in Figure 7. Algorithm 1 describes the steps and conditions required. First, a digital image-normalized cross-correlation was performed on the full-field detected speckle image to generate an initial 3D point cloud. In Step 2, object segmentation was applied to the initial 3D point cloud for the detection of object surface boundaries and contour edges. Following this, the mean and standard deviation of the correlation coefficient of object points were calculated and used as indexes for edge searching. Furthermore, the edge point was estimated using the geometry relationship of

each measured point as the detection center with its eight neighboring points. In Step 5, the boundary search point was defined as a non-edge point located adjacent to an edge point. The subset type of search NCC was determined by the geometry relationship of the inspection point and its eight neighboring points. A new edge point was detected and updated when the correlation coefficient of the search portion exceeded the pre-defined edge point threshold. The algorithm is detailed in Sections 2.3–2.5.

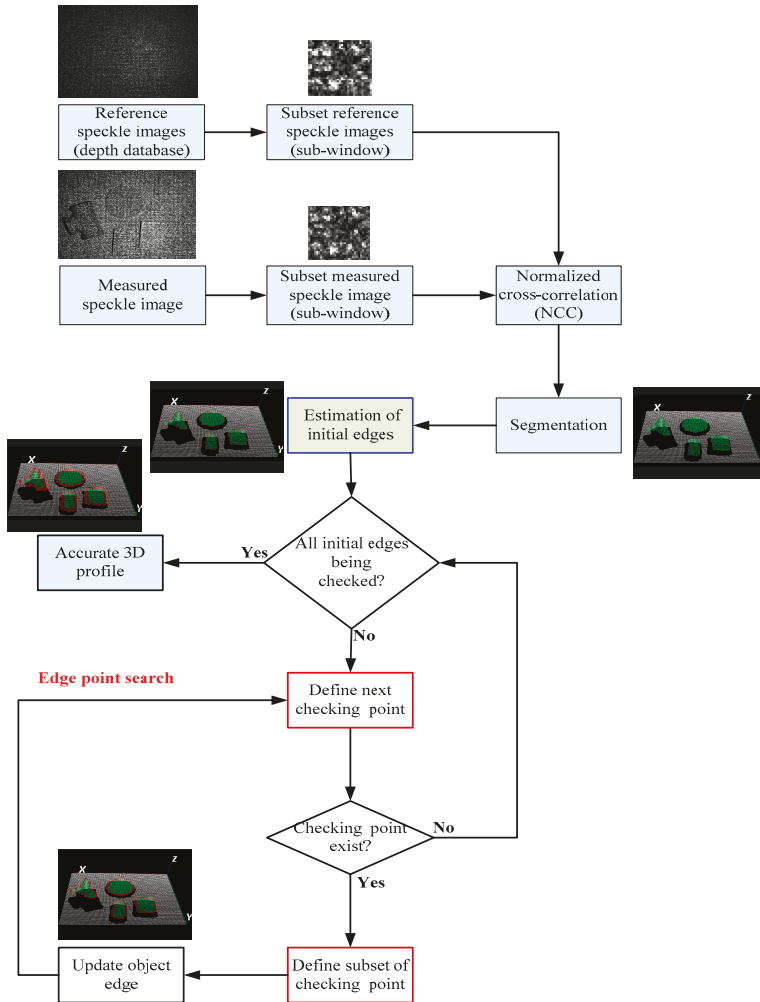


Figure 7. Flowchart of the developed approach.

**Algorithm 1**

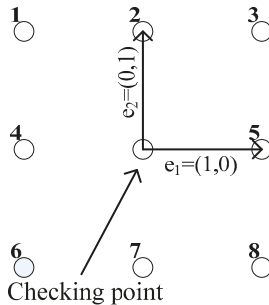
Input: Standard-height template database  $T$  and measured speckle image  $R$ .

Output: 3D profiles of objects

- 1 Implement normalized cross-correlation digital image correlation (NCC DIC) to generate point clouds.
- 2 Segment point clouds into individual measured objects.
- 3 Calculate the mean and standard deviation of the correlation coefficient of the measured point and use the mean as the threshold in the following search.
- 4 Estimate the edge boundary of each measured point using the geometry relationship of the point with its eight neighboring points.
- 5 Edge point search:
  - 5.1 Define the next inspection point according to the relationship of the point with its neighbors.
  - 5.2 Define the subset of inspection points according to the relationships of the points with their neighbors.
  - 5.3 Perform NCC between the subset of inspection points and the corresponding standard-height templates.
  - 5.4 Update the edge as an inspection point if its NCC result is larger than the current one.
  - 6 Stop the edge point search when there is no more inspection points.

**2.3. Initial Edge Point**

First, viewpoint-independent 3D object segmentation was performed to separate point clouds into individual objects and obtain the initial edges of each object [19]. The inspection point was defined as the origin of a Cartesian coordinate system with two vectors of unit length pointing along the x and y axis to its eight neighboring pixels, as shown in Figure 8.



**Figure 8.** The Cartesian coordinate system of inspection points.

The vector comprises the standard basis vectors  $e_1$  and  $e_2$ , as shown in Equation (2).

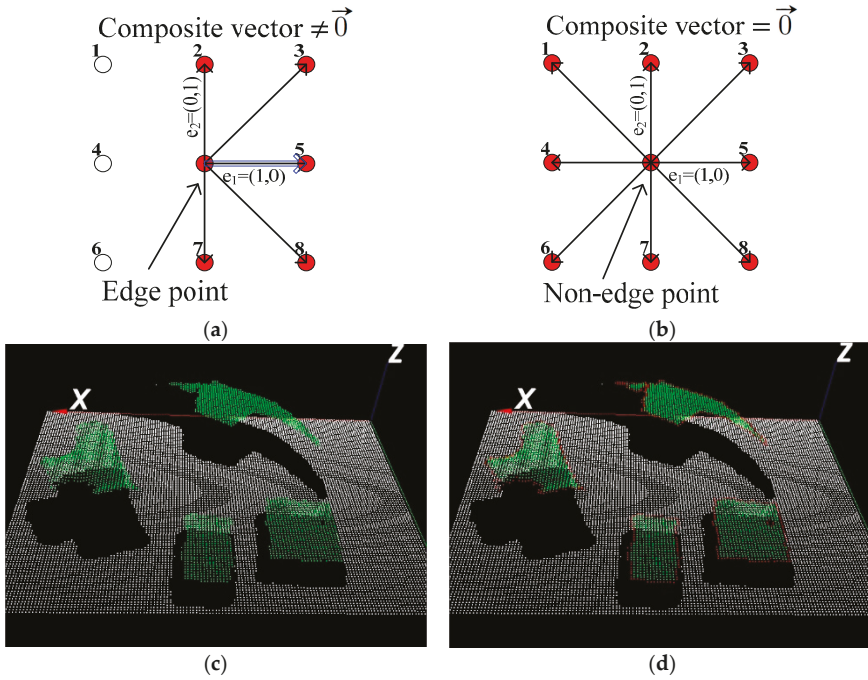
$$v_n = a_{1n}e_1 + a_{2n}e_2 \tag{2}$$

Where  $a_{1n}$  and  $a_{2n}$  are vector components of  $v_n$  on the basis vectors.

Then, the composite vector of an inspection point is defined by Equation (3).

$$\text{Composite vector} = \sum_{n=1}^8 a_{1n}e_1 + \sum_{n=1}^8 a_{2n}e_2 \tag{3}$$

There exists a vector from the inspection point to each of its neighboring points, which are marked as red spots in Figure 9. An inspection point in this step was defined as an initial edge point if its composite vector is not a zero vector (shown as Figure 9a). In contrast, the composite vector of a non-edge point is a zero vector (shown as Figure 9b). The initial edge link was generated by enclosing all the neighboring initial edge points.



**Figure 9.** Estimation of initial edge point: (a) edge point, (b) non-edge point, (c) an example of the measured point cloud, and (d) an initial edge link established for the example in (c).

2.4. Search Strategy for Accurate Edge Points

In view of the fact that the speckle pattern measured along a contour edge does not match the image in the reference template database, two new types of speckle subsets, namely boundary and corner subsets, were proposed to search the accurate edge points. The boundary subset had a condition that the split line was either of a horizontal or vertical boundary. Figure 10a shows the difference in edge detection between the proposed boundary subset and the traditional central subset when an inspection point was located on the contour edge. The overlapping region shown as a blue rectangle is the difference between the surface edge and the center of the boundary subset. Similarly, the corner subset had a condition that the split line is of a tilting boundary. Figure 10b shows the comparison between the corner subset and the traditional subset when an inspection point was located on the contour edge. The overlapping region shown as a blue square represents the potential difference in object edge detection. When the inspection point was scanned toward the eight directions shown in Figure 10c, eight speckle subsets could be formed either by the boundary subset or corner subset. More speckle subsets could be formed when more scanning directions were taken, for higher accuracy in the surface edge detection. In addition, the higher the scanning resolution, the more accurate the reconstruction of the object contour.

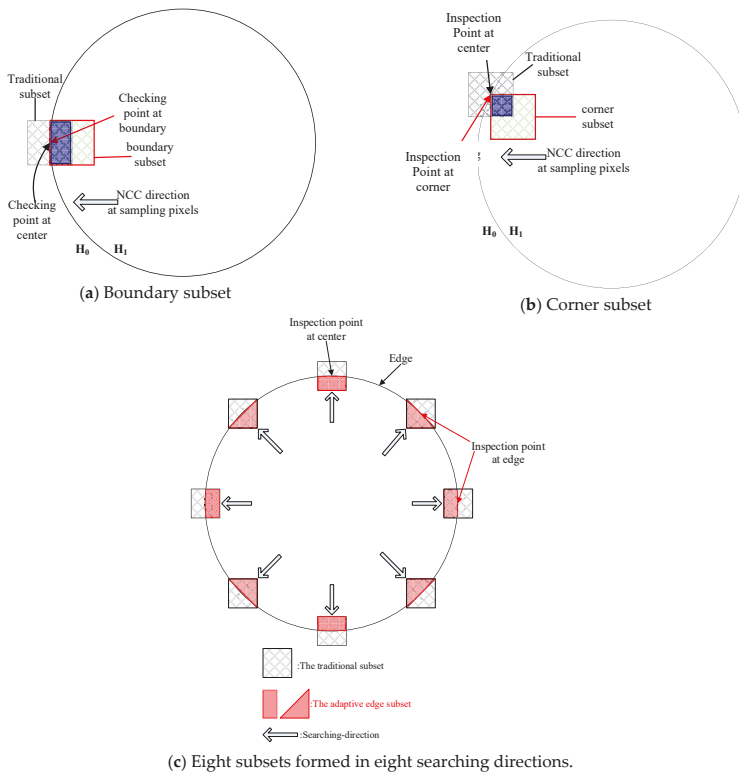


Figure 10. Proposed speckle subsets for contour edge detection.

The search direction of the edge point was specified to be opposite to its composite vector, as shown in Figure 11a. When an inspection point was set along the search direction, the composite vector of the inspection point was calculated to determine the subset type for the edge point search. Figure 11b shows the boundary speckle subset to be defined for a horizontal or vertical search direction.

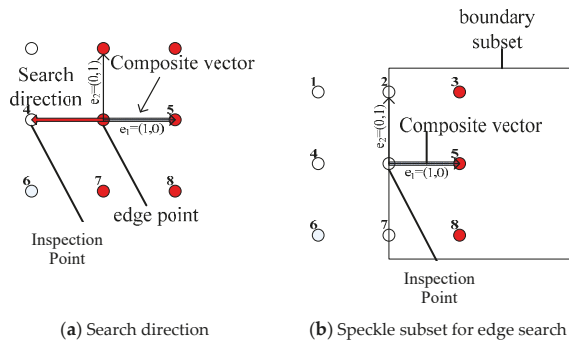


Figure 11. Boundary subset for horizontal or vertical edge search.

On the other hand, the corner subset was defined for a titling search direction. Figure 12a shows an inspection point set in the horizontal direction and another in the vertical direction. These two inspection points can define a corner subset for the edge search, as shown in Figure 12b.

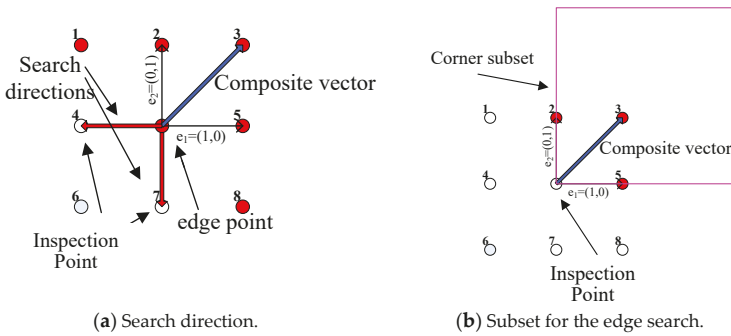


Figure 12. Corner subset for a titling edge search.

2.5. Criterion for Edge Detection

To ensure accurate edge detection, a threshold should be defined for the traditional NCC to filter low S/N coefficients from the point cloud. With a high-pass threshold for point filtration, the initial point cloud of the DIC measurement can provide a good basis for contour edge detection. As defined in Equation (4), an average correlation coefficient obtained by matching the reference image subset with its neighboring subset can generally provide a reliable threshold.

$$T_{ncc} = \frac{1}{c \times (r - 1)} \sum_1^c \sum_{N=1}^r \frac{\sum_{x,y} T_N(x, y) T_{N+1}(x + s, y + s)}{\sqrt{\sum_{x,y} T_N(x, y)^2 T_{N+1}(x + s, y + s)^2}}, \text{ for } N < r \tag{4}$$

where  $c$  is the number of subsets in a column;  $r$  is the number of subsets in a row, and  $s$  is the sample pixel.

When performing NCC with the boundary or corner subset, the threshold required for reliable edge detection should be further determined by the NCC of the inspection subset. In the proposed approach, the variance of the normalized cross-correlation coefficients along the search direction was taken as the quality-filtering threshold. The correlation coefficient curves obtained from the edge detection process using the traditional subset and the proposed search subset are shown in Figure 13a,b, respectively. The size of the partial search subset was either  $\frac{p}{n} \times n$  or  $n \times \frac{p}{n}$ , depending on the search direction, where  $p$  is the number of sampling pixels in DIC. As can be seen, the normalized cross-correlation coefficient (NCCC) of the proposed search subset generally had a much higher sensitivity or slope than that of the traditional subset, because the subset speckle images of the proposed approach better matched their corresponding reference template images.

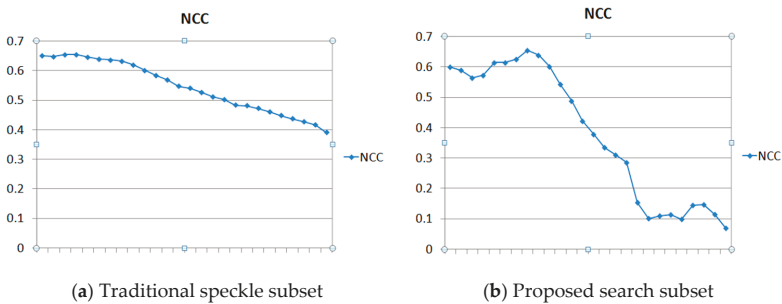
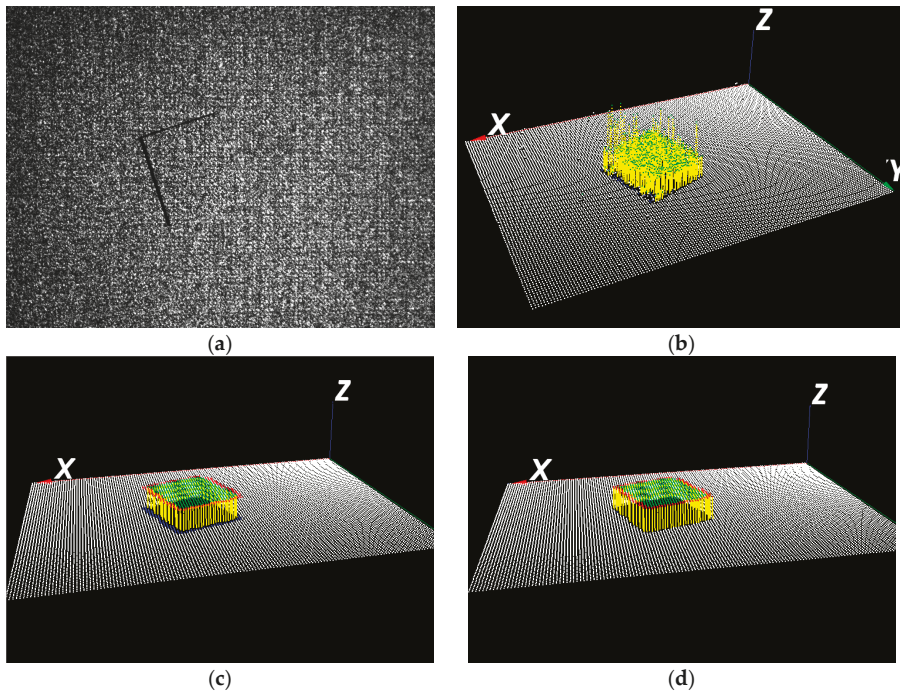


Figure 13. Normalized cross-correlation coefficient (NCCC) curves.

### 3. Experimental Results and Accuracy Analyses

The proposed approach was developed using C++ on a personal computer with 3.2 Hz i5-4570 CPU and 4GB RAM. The 3D optical probe was incorporated with a three-axis scanning stage via Ethernet to detect speckle images. Three experiments with different setups were conducted using the probe and the developed approach to evaluate their effectiveness. First, to verify the accuracy of the developed approach, the reconstruction of a pre-calibrated gauge block with a step height of 9.000 mm and a coordinate measuring machine (CMM) calibrated circle target with a thickness of 5.000 mm was performed. Figure 14 shows the reconstruction results of the pre-calibrated gauge block. Repeated tests yielded improvements in the width-measuring error, standard deviation, and measurements of 1.06%, 315  $\mu\text{m}$  and 22-fold, respectively; compared with the results obtained by the traditional DIC. Figure 14 also shows the improvement in accuracy obtained via 3D profiles and the corresponding measured speckle images with 2D edge detection of each step. The experimental results obtained verified the superiority of the developed approach to the traditional DIC for measuring surface contours, especially for edge measurement and reconstruction.



**Figure 14.** Measurement results obtained using the proposed approach for reconstruction of the pre-calibrated gauge block: (a) speckle image; (b) 3D result measured using traditional DIC; (c) initial edge reconstruction using the proposed approach, in which some edge positions are still missing from its original object shape; and (d) 3D map reconstructed using the proposed edge detection method.

Figure 15 shows the reconstruction results of the pre-calibrated circle target. As can be seen, the measurement error was kept within 0.4% of the original diameter while the standard deviation was maintained at 267  $\mu\text{m}$ . In comparison with traditional DIC, the developed approach achieved an 18-fold improvement in diameter measurement, indicating the superiority of the proposed method for critical dimension measurement. The comparison of different measurement results is shown in Table 1.

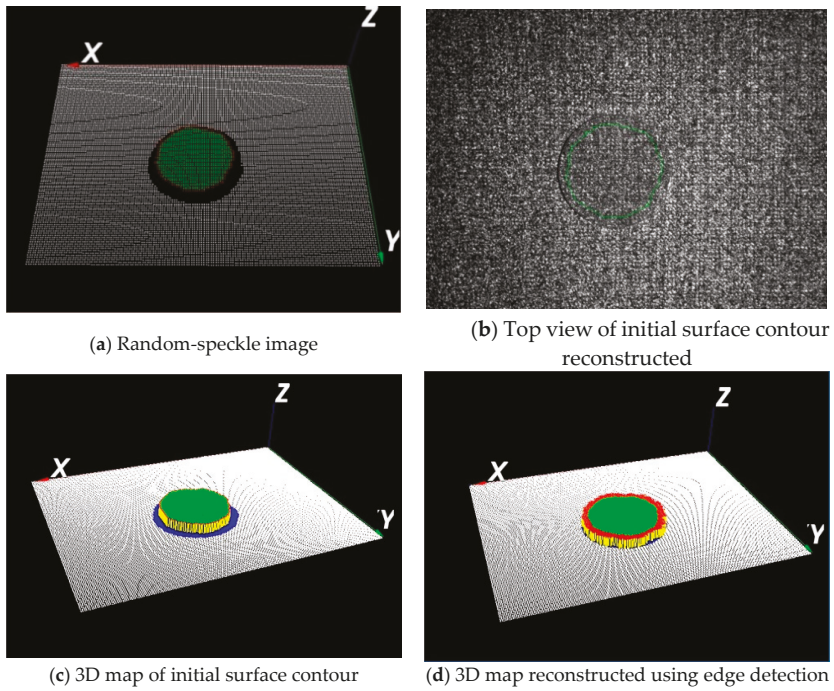


Figure 15. Pre-calibrated circle target reconstructed with edge detection.

Table 1. Comparison of measurement accuracy and improvement achieved by coordinate measuring machine (CMM), traditional digital image correlation (DIC) and the developed approach.

Method	CMM (A)	Traditional DIC (B)	Developed Approach			
	Mean Diameter or Width (mm)	Mean Diameter/Width (mm)	Mean Diameter or Width (mm) (C)	Standard Deviation (µm)	Measurement Error (%) (D)	Measurement Improvement (%) (E)
Circular plate (diameter)	45.473	41.513	45.253	267	0.4%	1800%
Gauge block (width)	29.98	22.97	29.662	315	1.06%	2204%

Note:  $D = \frac{|C-A|}{A}$ ,  $E = \frac{C-B}{A-B}$ .

Edge detection was also performed on industrial workpieces to validate the feasibility of the developed method. Figure 16 shows the surface profile reconstruction of a 3D-printed hammer using the proposed approach, while Figure 17 compares the 3D map of a molding part obtained using both traditional DIC and the developed method. Experimental results and analyses revealed that traditional DIC has difficulty providing reliable edge information. The critical dimension could not be reconstructed accurately and conventional 2D image edge-detection technique had to be employed in conjunction to accomplish the measurement. The developed method achieved a 15-fold improvement in measurement, with error controlled within 1%.



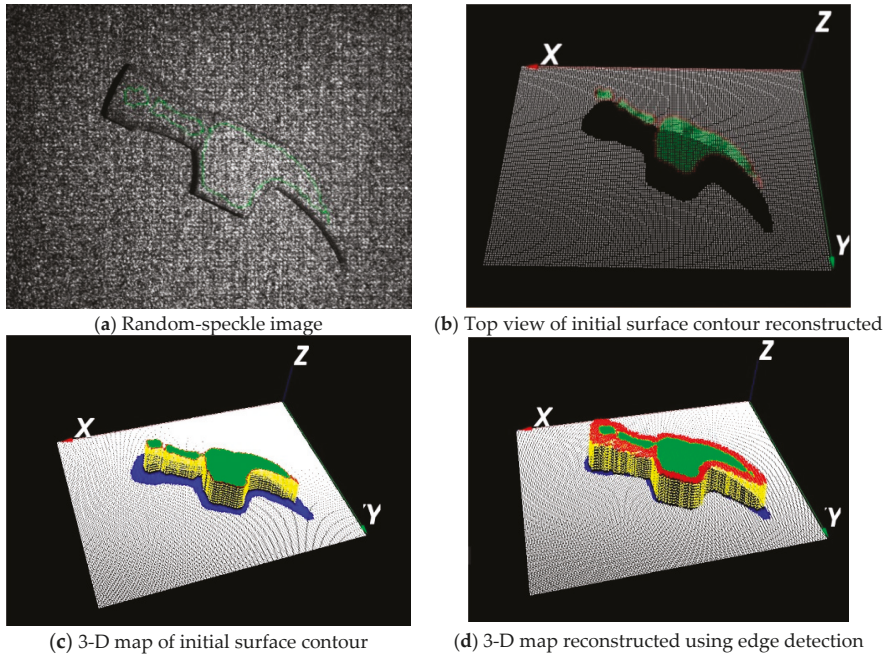


Figure 16. Reconstruction of a 3D-printed hammer using the proposed approach.

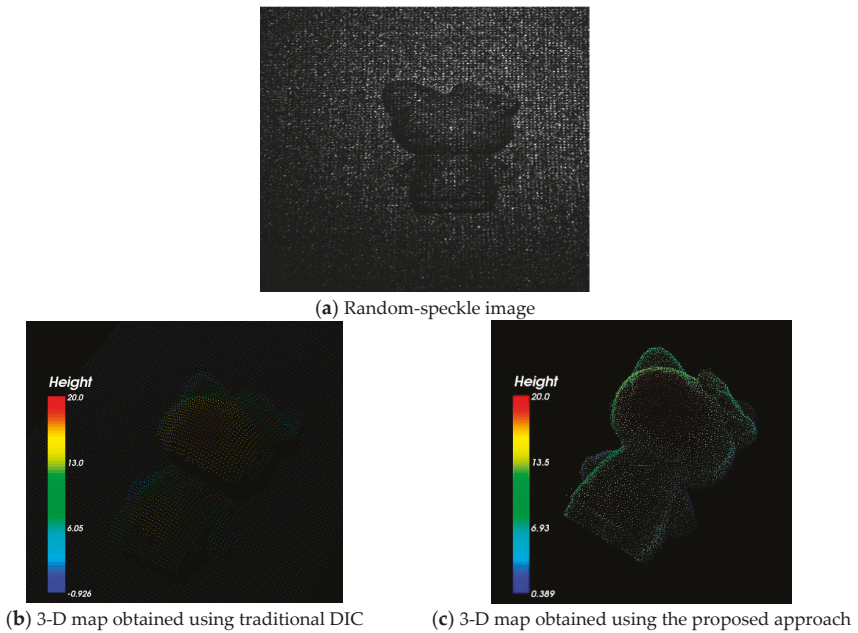
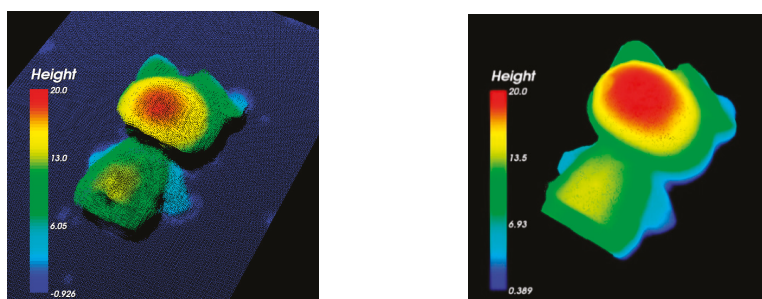


Figure 17. Cont.



(d) 3-D depth map obtained using traditional DIC (e) 3-D depth map obtained using the proposed method

**Figure 17.** Comparison of an industrial molding part reconstructed using DIC and the proposed approach.

#### 4. Conclusions

In this research, a new surface edge detection approach was developed to achieve accurate surface edge measurement using various speckle DIC-based techniques. The developed algorithm has its basis in edge detection subsets and the variance of the normalized cross-correlation coefficients. Preliminary experimental results confirmed that the maximum measured error of the selected critical dimension was significantly reduced and controlled within 1% of the measuring range or size. The precision of the measurement, as one standard deviation, was also kept below 1.05% of the measuring range. Moreover, this study also verified that the subset size and speckle pattern contrast are crucial parameters influencing measurement precision.

**Author Contributions:** Conceptualization, L.-C.C.; Data curation, C.-W.L.; Formal analysis, L.-C.C.; Investigation, L.-C.C. and C.-W.L.; Methodology, L.-C.C.; Resources, L.-C.C.; Software, C.-W.L.; Supervision, L.-C.C.; Validation, L.-C.C.; Writing – original draft, C.-W.L.; Writing – review & editing, L.-C.C.

**Conflicts of Interest:** The authors declare no conflicts of interest.

#### References

- Andersen, M.R.; Jensen, T.; Lisouski, P.; Mortensen, A.K.; Hansen, M.K.; Gregersen, T.; Ahrendt, P. *Kinect Depth Sensor Evaluation for Computer Vision Applications*; Technical Report; Aarhus University: Aarhus, Denmark, 2012.
- McNeil, S.R.; Sutton, M.A.; Miao, Z.; Ma, J. Measurement of surface profile using digital image correlation. *Exp. Mech.* **1997**, *37*, 13–20. [[CrossRef](#)]
- Sjödahl, M.; Synnergren, P. Measurement of shape by using projected random patterns and temporal digital speckle photography. *Appl. Opt.* **1999**, *38*, 1990–1997. [[CrossRef](#)] [[PubMed](#)]
- Lecompte, D.; Smits, A.; Bossuyt, S. Quality assessment of speckle patterns for digital image correlation. *Opt. Lasers Eng.* **2006**, *44*, 1132–1145. [[CrossRef](#)]
- Schreier, H.W.; Sutton, M.A. Systematic errors in digital image correlation due to under matched subset shape functions. *Exp. Mech.* **2002**, *42*, 303–310. [[CrossRef](#)]
- Pan, B.; Lu, Z.; Xie, H. Mean intensity gradient: An effective global parameter for quality assessment of the speckle patterns used in digital image correlation. *Opt. Lasers Eng.* **2010**, *48*, 469–477. [[CrossRef](#)]
- Pan, B.; Xie, H.; Wang, Z.; Qian, K.; Wang, Z. Study on subset size selection in digital image correlation for speckle patterns. *Opt. Express.* **2008**, *16*, 7037–7048. [[CrossRef](#)]
- Crammond, G.; Boyd, S.W.; Dulieu-Barton, J.M. Speckle pattern quality assessment for digital image correlation. *Opt. Lasers Eng.* **2013**, *51*, 1368–1378. [[CrossRef](#)]
- García, J.; Zalevsky, Z. Three-dimensional mapping and range measurement by means of projected speckle patterns. *Appl. Opt.* **2008**, *47*, 3032–3040. [[CrossRef](#)] [[PubMed](#)]
- Pan, B.; Qian, K.; Xie, H.; Asundi, A. Two-dimensional digital image correlation for in-plane displacement and strain measurement: A review. *Meas. Sci. Technol.* **2009**, *20*, 062001. [[CrossRef](#)]



11. Stoilov, G.; Kavardzhikov, V.; Pashkouleva, D. A comparative study of random patterns for digital image correlation. *J. Theor. Appl. Mech. Pol.* **2012**, *42*, 55–66. [[CrossRef](#)]
12. Poissant, J.; Barthelat, F. A novel “subset splitting” procedure for digital image correlation on discontinuous displacement fields. *Exp Mech.* **2010**, *50*, 353–364. [[CrossRef](#)]
13. Dai, H.; Su, X. Shape measurement by digital speckle temporal sequence correlation with digital light projector. *Opt. Eng.* **2001**, *40*, 793–800. [[CrossRef](#)]
14. Wang, G.; Yin, X.; Pei, X.; Shi, C. Depth estimation for speckle projection system using progressive reliable points growing matching. *Appl. Opt.* **2013**, *52*, 516–524. [[CrossRef](#)] [[PubMed](#)]
15. Yin, X.; Wang, G.; Shi, C.; Liao, Q. Efficient active depth sensing by laser speckle projection system. *Opt. Eng.* **2014**, *53*, 013105. [[CrossRef](#)]
16. Le, M.-T.; Chen, L.-C.; Lin, C.-J. Reconstruction of accurate 3-D surfaces with sharp edges using digital structured light projection and multi-dimensional image fusion. *Opt. Lasers Eng.* **2017**, *96*, 17–34. [[CrossRef](#)]
17. Chen, L.-C.; Liang, C.-W.; Tseng, H.-Y.; Lin, S.-T. Accurate 3-D surface profilometry using novel boundary edge detection on digital image correlation. *Appl. Mech. Mater.* **2017**, *870*, 295–302. [[CrossRef](#)]
18. Sutton, M.A.; Orteu, J.J.; Schreier, H.W. *Image Correlation for Shape, Motion and Deformation Measurements: Basic Concepts, Theory and Applications*; Springer: New York, NY, USA, 2009.
19. Chen, L.-C.; Nguyen, T.-H.; Lin, S.-T. Viewpoint-independent 3-D object segmentation for randomly stacked objects using optical object detection. *Meas. Sci. Technol.* **2015**, *26*, 105202. [[CrossRef](#)]



© 2018 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).

Article

# Uncertainty Evaluation for Measurements of Pitch Deviation and Out-of-Flatness of Planar Scale Gratings by a Fizeau Interferometer in Littrow Configuration

Xin Xiong <sup>1</sup>, Yuki Shimizu <sup>1,\*</sup>, Xiuguo Chen <sup>1,2</sup> , Hiraku Matsukuma <sup>1</sup>  and Wei Gao <sup>1</sup>

<sup>1</sup> Precision Nanometrology Laboratory, Department of Finemechanics, Tohoku University, Sendai 980-8579, Japan; xinxiong@nano.mech.tohoku.ac.jp (X.X.); xiuguochen@nano.mech.tohoku.ac.jp (X.C.); hiraku.matsukuma@nano.mech.tohoku.ac.jp (H.M.); gaowei@cc.mech.tohoku.ac.jp (W.G.)

<sup>2</sup> State Key Laboratory for Digital Manufacturing Equipment and Technology, Huazhong University of Science and Technology, Wuhan 430074, China

\* Correspondence: yuki.shimizu@nano.mech.tohoku.ac.jp; Tel.: +81-22-795-6950

Received: 25 October 2018; Accepted: 5 December 2018; Published: 7 December 2018



**Abstract:** Form errors of a planar scale grating, such as pitch deviations and out-of-flatness, are major contributors to the final measurement uncertainty of an interferential scanning-type planar encoder. Following the previous work, in which a method has been proposed to evaluate both the out-of-flatness and the pitch deviations of a planar scale grating by a Fizeau interferometer in Littrow configuration, uncertainty analysis on this method is performed in this paper. Theoretical equations are derived to make quantitative uncertainty analysis while taking possible error factors into account. To overcome the drawbacks of a traditional uncertainty matrix approach, a new procedure is proposed to evaluate the uncertainty in the PV (peak-to-valley) deviation of a surface form, so as to assure the quality of measurement. Experiments are finally conducted to demonstrate the feasibility of proposed uncertainty evaluation method.

**Keywords:** scale grating; Fizeau interferometer; optical encoder; pitch deviation; out-of-flatness; uncertainty

## 1. Introduction

An optical encoder, which is mainly composed of an optical sensor head and a scale grating, is one of the most commonly used sensors for precision positioning in production engineering [1], due to the advantages of non-contact measurement, a high resolution, and a wide bandwidth [2]. In the case of incremental type one-axis encoders, a diffraction grating having periodic line pattern structures is employed as a scale for measurement of relative translational displacement between the scale grating and an optical sensor head [3]. With the employment of a planar scale grating having two-dimensional grating pattern structures [4], the dimension of displacement/position measurement can be expanded to two or more [5,6]. Among the components in a multi-axis planar encoder system [5], a planar scale grating having periodical structures along the X- and Y-directions is one of the key components that influence the accuracy of measurement, since the period of interference signal in the optical head, which will be interpolated and used to calculate the displacement, is directly influenced by the pitch of the scale grating, in principle [4]. Moreover, the Z-directional out-of-flatness of the scale grating could also be a source of measurement uncertainty in the planar encoder system. The X- and Y-directional pitch deviations, as well as the out-of-flatness of the scale grating, are thus required to be evaluated and calibrated over the entire surface of a scale grating for high precision measurement applications [7].

Line scale comparator [8,9] and scanning probe microscopes (SPMs) [10,11] are available as conventional solutions to evaluate the pitch deviations of a planar scale grating. The line scale comparator has long been used in the enterprises that manufacture linear encoders and national standard institutes for testing one-axis linear scale gratings [8,9]. However, it is too expensive to construct a measurement system for two-dimensional scale grating that is based on the line scale comparator. Another drawback of the line scale comparator is that the out-of-flatness of a planar scale grating cannot be measured in principle. Meanwhile, SPMs can measure both the grating pitch deviations and the out-of-flatness of a planar scale grating. However, on the other hand, a limited scanning area and a measurement throughput of SPMs [11] prevent them being applied for the evaluation of planar scale grating, especially for the condition of laboratory-level, it may take hours and days to conduct the measurement of a large size planar scale grating over its whole area.

In responding to the background described above, the authors' group has proposed a method to simultaneously evaluate both the *X*- and *Y*-directional pitch deviations and the *Z*-directional out-of-flatness of a planar scale grating by utilizing a Fizeau interferometer in the previous work [7]. In the propose method, wavefront of the zeroth-order diffracted beam and wavefronts of positive and negative first-order diffracted beams from a scale grating obtained in Littrow configurations were utilized to perform a non-contact evaluation of a planar scale grating over the entire area. The feasibility of proposed method has been verified in experiments [7]. In addition, the proposed method has been improved so that the self-calibration of the Fizeau interferometer and the planar scale grating can be conducted simultaneously while considering the out-of-flatness of the reference optical flat and the change of the coordinate system in the Littrow configuration [12]. As the first step of the research, a preliminary verification of the improved method has been performed successfully through the simulation in both the noise-free and noisy cases [12]. The key advantage of this method is that the measurement using the Fizeau interferometer can be carried out without significant investments in time and/or capital to assess the out-of-flatness and pitch deviations of the planar scale grating accurately compared with other measurement techniques. Nevertheless, further experimental demonstration as well as the uncertainty analysis of the proposed method have not been conducted yet, and they remain to be addressed.

In this paper, as the second step of the research, experimental verification of the feasibility of the improved method [12] has been conducted as well as the uncertainty analysis when considering the possible error factors influencing the measurement uncertainties, such as the form error of the reference surface in a Fizeau interferometer or the inclination error of a scale grating in measurement. Theoretical equations are derived to evaluate the influence of the error factors on the final uncertainties in the measured *Z*-directional out-of-flatness as well as the *X*- and *Y*-directional pitch deviations of a scale grating. Moreover, when considering the drawbacks of the traditional uncertainty matrix approach [13], which is intellectually satisfying but difficult to handle or specify the surface form in a straightforward way, a method aiming to evaluate the uncertainty by using one of the simplest but most commonly used peak-to-valley (P-V) deviation of the surface form is proposed. Experiments are also conducted to demonstrate the feasibility of the proposed method.

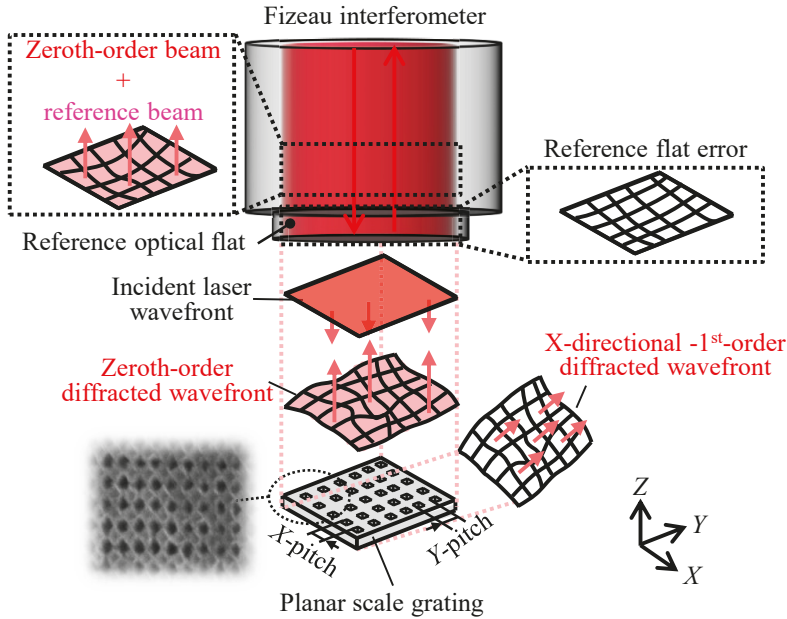
## 2. Measurement Principle

Figure 1 demonstrates the measurement of the *Z*-directional out-of-flatness of a planar scale grating  $e_Z(x,y)$  using the typical function of the Fizeau interferometer. Assume that the field-of-view (FOV) of the Fizeau interferometer is larger than the size of the grating employed for measurement. As can be seen in the figure, the wavefront of the zeroth-order diffracted beam from the grating superimposed on the wavefront of the reflected beam from the reference optical flat. Consequently, the zeroth-order phase output  $I_0(x,y)$  from the interferometer can be expressed by the following equation [7]:

$$I_0(x,y) = \frac{4\pi[e_Z(x,y) - e_R(x,y)]}{\lambda} \quad (1)$$

where  $\lambda$  represents the light wavelength of the laser source and  $e_R(x,y)$  represents the out-of-flatness of the reference optical flat in the Fizeau interferometer. The out-of-flatness of the grating  $e_Z(x,y)$  can be obtained by the following equation according to Equation (1):

$$e_Z(x,y) = \frac{\lambda}{4\pi} I_0(x,y) + e_R(x,y) \tag{2}$$



**Figure 1.** Measurement of zeroth-order diffracted wavefront using Fizeau interferometer (positive first-order diffracted wavefront is not shown for the sake of simplicity).

As can be seen in Equation (2), the result of  $e_Z(x,y)$  contains the out-of-flatness of the reference flat  $e_R(x,y)$ . Since the calibration of the absolute flatness of the reference flat is quite complicated in practice [14], in this paper, the mean (average or expectation) of  $e_R(x,y)$  is taken to be zero and its uncertainty being the specified peak-to-valley (PV) value, which is typically less than  $\lambda/20$  over the whole field of view (FOV) of a commercial Fizeau interferometer.

For the evaluation of X-directional pitch deviation  $e_X(x,y)$  and Y-directional pitch deviation  $e_Y(x,y)$  of the planar scale grating, measurement of the wavefronts of the X and Y-directional positive and negative first-order diffracted beams in Littrow configurations is needed. Figure 2 shows an example of the Littrow configuration for the measurement of the wavefront of X-directional positive first-order diffracted beam, where the grating is tilted clockwise, so that the X-directional positive first-order diffracted beam is back-reflected directly into the direction of the incident laser beam and superimposed with the reference beam in the Fizeau interferometer. Assume that the actual inclination angle in measurement  $\theta_1$  can be represented by  $\theta_1 = \theta/2 + \varepsilon_1$ , where  $\theta = \arcsin(\lambda/g)$  is the first-order diffraction angle,  $\varepsilon_1$  is the misalignment of  $\theta$ , and  $g$  is the nominal pitch of the grating along the X-direction. The X-directional pitch deviation  $e_X(x,y)$  of the planar scale grating causes a phase shift in the wavefront of the diffracted beam [7]. As depicted in Figure 2, the coordinate system has changed after setting the grating in Littrow configuration. Since the measurement results should correspond

to the original grating coordinate system, the X-directional positive first-order phase output  $I_{X+1}(x,y)$  that was obtained in the Fizeau measurement can be described, as follows:

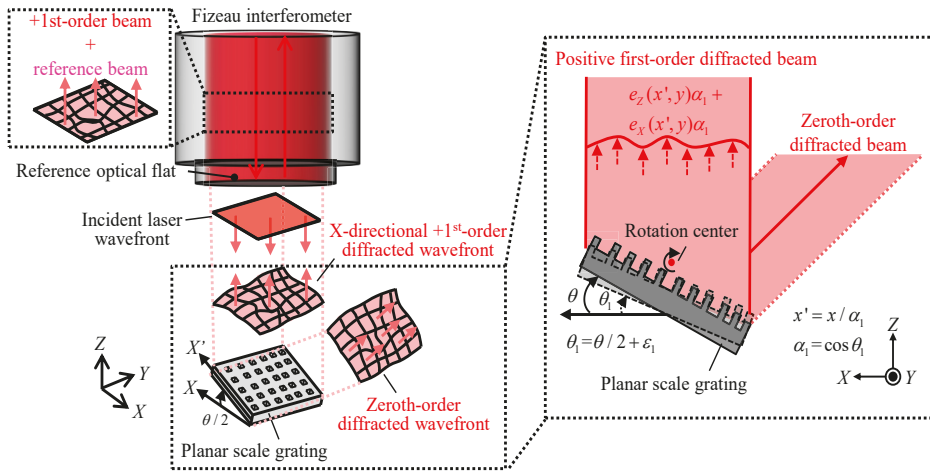
$$I_{X+1}(x,y) = 2\pi \frac{e_X(x/\alpha_1,y)}{g} + \frac{4\pi}{\lambda} [\alpha_1 e_Z(x/\alpha_1,y) - e_R(x,y)] \quad (3)$$

where  $\alpha_1 = \cos\theta_1$ . In the same manner, the wavefront of the X-directional negative first-order diffracted beam can be measured by tilting the scale grating counterclockwise with an opposite inclination angle, the X-directional negative first-order phase output  $I_{X-1}(x,y)$  can be obtained, as follows:

$$I_{X-1}(x,y) = -2\pi \frac{e_X(x/\alpha_2,y)}{g} + \frac{4\pi}{\lambda} [\alpha_2 e_Z(x/\alpha_2,y) - e_R(x,y)] \quad (4)$$

where  $\alpha_2 = \cos\theta_2$ . By using the error  $\varepsilon_2$  in the alignment of  $\theta_2$  in measurement of the negative first-order diffracted beam,  $\theta_2$  can be represented as  $\theta_2 = -\theta/2 + \varepsilon_2$ . According to Equations (3) and (4), the following equation can be obtained:

$$e_X(x/\alpha_1,y) + e_X(x/\alpha_2,y) = \frac{g}{2\pi} [I_{X+1}(x,y) - I_{X-1}(x,y)] - \frac{2g}{\lambda} [\alpha_1 e_Z(x/\alpha_1,y) - \alpha_2 e_Z(x/\alpha_2,y)] \quad (5)$$



**Figure 2.** Measurement of X-directional positive first-order diffracted wavefront with inclination error (negative first-order diffracted wavefront is not shown for the sake of simplicity).

In the same manner, the Y-directional pitch deviation  $e_Y(x,y)$  of the planar scale grating can be obtained as follows by measuring the Y-directional positive and negative first-order phase outputs  $I_{Y+1}(x,y)$  and  $I_{Y-1}(x,y)$ , respectively:

$$I_{Y+1}(x,y) = 2\pi \frac{e_Y(x,y/\alpha'_1)}{g} + \frac{4\pi}{\lambda} [\alpha'_1 e_Z(x,y/\alpha'_1) - e_R(x,y)], \quad (6)$$

$$I_{Y-1}(x,y) = -2\pi \frac{e_Y(x,y/\alpha'_2)}{g} + \frac{4\pi}{\lambda} [\alpha'_2 e_Z(x,y/\alpha'_2) - e_R(x,y)], \quad (7)$$

where  $\alpha'_1 = \cos\theta'_1 = \cos(\theta/2 + \varepsilon'_1)$  and  $\alpha'_2 = \cos\theta'_2 = \cos(\theta/2 + \varepsilon'_2)$ .  $\varepsilon'_1$  and  $\varepsilon'_2$  represent the misalignments of  $\theta$  in the measurement of the wavefronts of the Y-directional positive and negative first-order diffracted beams, respectively. Note here that the nominal pitches of the planar scale grating

along the X- and Y-directions are assumed to be the same for the sake of simplicity. According to Equations (6) and (7), the following equation can be obtained:

$$e_Y(x, y/\alpha'_1) + e_Y(x, y/\alpha'_2) = \frac{g}{2\pi} [I_{Y+1}(x, y) - I_{Y-1}(x, y)] - \frac{2g}{\lambda} [\alpha'_1 e_Z(x, y/\alpha'_1) - \alpha'_2 e_Z(x, y/\alpha'_2)] \quad (8)$$

In the following section of this paper, how to further obtain the mean values and uncertainties of  $e_Z(x, y)$ ,  $e_X(x, y)$ , and  $e_Y(x, y)$  according to Equations (2), (5), and (8) are presented. Since the evaluations of the mean value and uncertainty of  $e_Y(x, y)$  are similar to those of  $e_X(x, y)$ , only the latter is focused in the following for the sake of clarity.

### 3. Uncertainty Evaluation

Major sources of uncertainty in the evaluation of out-of-flatness  $e_Z(x, y)$  and pitch deviation  $e_X(x, y)$  of a planar scale grating by a Fizeau interferometer primarily arise from uncertainties in the measured phase outputs  $I_0(x, y)$  and  $I_{X\pm 1}(x, y)$ . In addition, out-of-flatness  $e_R(x, y)$  of the reference flat in a Fizeau interferometer and uncertainties in the nominal pitch  $g$  of the planar scale grating under test, as well as the inclination angles  $\theta_1$  and  $\theta_2$ , could also be major sources of uncertainty. According to the GUM (guide to the expression of uncertainty in measurement) [13], the uncertainties in  $I_0(x, y)$  and  $I_{X\pm 1}(x, y)$  attribute to the Type A evaluation, while those in  $e_R(x, y)$ ,  $g$ ,  $\theta_1$  and  $\theta_2$  attribute to the Type B evaluation. Through the uncertainty analysis, The  $E(\cdot)$ ,  $\sigma(\cdot)$  and  $\sigma^2(\cdot)$  are used to represent the mean, standard deviation, and variance of a random variable, respectively. The expanded uncertainty of a random variable is given by  $u(\cdot) = k\sigma(\cdot)$ , where  $k$  denotes the coverage factor;  $k = 2$  produces a confidence interval of approximately 95%; and,  $k = 3$  produces an interval with a confidence of approximately 99%.

As mentioned in Section 2, in this paper, the mean and  $3\sigma$  uncertainty of  $e_R(x, y)$  are treated to be zero and the specified PV value, respectively. The uncertainty in the inclination angles  $\theta_1$  and  $\theta_2$  contains two components; one is the uncertainty induced by the deviation of the nominal grating pitch with respect to its true value, and another is induced by the misalignments in measurement. In the uncertainty analysis, the means of  $\theta_1$  and  $\theta_2$  are taken to be  $\theta/2$  and  $-\theta/2$ , respectively, and the uncertainties that are induced by the misalignments are assumed to be identical; namely,  $\sigma(\varepsilon_1) = \sigma(\varepsilon_2)$ . Due to the independence between  $\theta$  and  $\varepsilon_1$  ( $\varepsilon_2$ ), the means of  $\alpha_1$  and  $\alpha_2$  will be  $\cos(\theta/2)$ , which is denoted as  $\alpha$  in the following ( $\alpha = \cos(\theta/2)$ ). By taking the means for both sides of Equations (2) and (5), the following equations can be obtained:

$$E[e_Z(x, y)] = \frac{\lambda}{4\pi} E[I_0(x, y)], \quad (9)$$

$$E[e_X(x/\alpha, y)] = \frac{g}{4\pi} \{E[I_{X+1}(x, y)] - E[I_{X-1}(x, y)]\}, \quad (10)$$

It should be pointed out that here has  $E[\alpha_1 e_Z(x/\alpha_1, y)] = E[\alpha_2 e_Z(x/\alpha_2, y)]$  mathematically. In practical experiments, the second term of the right side of Equation (5) could be minimized by making interference fringes of the Fizeau interferometer minimum though adjusting the tilt stage on which the scale grating is fixed. Although the influence from the out-of-flatness of the substrate surface cannot completely be eliminated by the differential operation, when considering the holographic grating with a good surface quality, the effect on the evaluation of the pitch deviation can be ignored by conducting the operation mentioned above.

A straightforward approach to evaluate the uncertainty in the obtained  $e_Z(x, y)$  and  $e_X(x, y)$  is to generate a matrix, called the uncertainty matrix [15], which has the same size to  $e_Z(x, y)$  or  $e_X(x, y)$  and it contains the corresponding uncertainty of  $e_Z(x, y)$  or  $e_X(x, y)$  at every pixel position  $(x_i, y_j)$  ( $i = 1, 2, \dots$ ,



$M; j = 1, 2, \dots, N$ ). Taking the propagation of errors into consideration, the variances  $\sigma^2[e_Z(x,y)]$  and  $\sigma^2[e_X(x/\alpha,y)]$  can be represented, as follows:

$$\sigma^2[e_Z(x,y)] = \left(\frac{\lambda}{4\pi}\right)^2 \sigma^2[I_0(x,y)] + \sigma^2[e_R(x,y)], \tag{11}$$

$$\begin{aligned} \sigma^2[e_X(x/\alpha,y)] &= \left(\frac{g}{4\pi}\right)^2 \{ \sigma^2[I_{X+1}(x,y)] + \sigma^2[I_{X-1}(x,y)] \} + \left\{ \frac{E[e_X(x/\alpha,y)]}{g} \right\}^2 \sigma^2(g) \\ &+ 2 \left[ \frac{g}{\lambda} e_Z(x/\alpha,y) + \frac{g}{\lambda} \alpha \frac{\partial e_Z(x/\alpha,y)}{\partial \alpha} \right]^2 \sigma^2(\alpha) + 2 \left(\frac{g}{\lambda} \alpha\right)^2 \sigma^2[e_Z(x/\alpha,y)], \end{aligned} \tag{12}$$

where the variance  $\sigma^2(\alpha)$  is evaluated by the following equation:

$$\sigma^2(\alpha) = \sin^2\left(\frac{\theta}{2}\right) \left\{ \left[ \frac{1}{2} \frac{1}{\sqrt{1 - (\lambda/g)^2}} \frac{\lambda}{g^2} \right]^2 \sigma^2(g) + \sigma^2(\varepsilon) \right\}. \tag{13}$$

The first and second terms in curly brackets of the right side of Equation (13) correspond to the uncertainties that are induced by the deviation of the nominal grating pitch to its true value and the misalignment in measurement, respectively. It should be noted from Equations (11) and (12) that, for  $e_Z(x,y)$ , the corresponding uncertainty matrix can be obtained, while for  $e_X(x,y)$  it cannot, because of the change in the coordinate system after tilting the scale gating. Moreover, as shown in Equation (12), even for  $e_X(x/\alpha,y)$ , the corresponding uncertainty matrix cannot be obtained either due to the involvement of the evaluation of the unknown  $e_Z(x/\alpha,y)$ ,  $\partial e_Z(x/\alpha,y)/\partial \alpha$ , and  $\sigma^2[e_Z(x/\alpha,y)]$ . On the other hand, it should be pointed out that the uncertainty matrix approach has a number of drawbacks in practice. As stated in [16], the information density that is contained in the uncertainty matrix is too high for users who only want a single number. In addition, the uncertainty matrix cannot easily be used for the decision of either acceptance or rejection. Evaluations of measurement uncertainty should be expressed in a manner that is informative and actionable [17]. Therefore, in this paper, a PV value is adopted to evaluate the quality of the obtained form errors. Although a PV value also has inadequacies in charactering surfaces, it is still widely recognized that a PV value remains one of the most commonly used specifications of surface form errors [18]. Next, the details are presented in evaluating a PV value and its uncertainty for the characterization of  $e_Z(x,y)$  and  $e_X(x,y)$ .

In this paper, the form errors  $e_Z(x,y)$  and  $e_X(x,y)$  are assumed can be characterized in terms of polynomials. This assumption is reasonable, since, as stated in [19], the low-frequency errors can be completely modeled by the set of polynomials and are of the main concern in the application. Moreover, modeling the form errors in terms of polynomials is also favorable to evaluate the associated PV values since the estimated PV values will not be biased by the high-frequency noises contained in the raw data. Without losing generality, the  $n$ -degree bivariate polynomials are used to fit the discrete form errors  $(x_i, y_j, E[e_Z(x_i,y_j)])$  and  $(x_i, y_j, E[e_X(x_i/\alpha,y_j)])$  obtained at each pixel position  $(x_i,y_j)$ , according to Equations (9) and (10), as follows:

$$e_Z(x,y) = [1, x, x^2, \dots, x^n] \begin{bmatrix} A_{00} & A_{01} & A_{02} & \cdots & A_{0,n} \\ A_{10} & A_{11} & \cdots & & A_{1,n-1} \\ A_{20} & \cdots & A_{2,n-2} & & \\ \vdots & & & & \\ A_{n0} & & & & 0 \end{bmatrix} \begin{bmatrix} 1 \\ y \\ y^2 \\ \vdots \\ y^n \end{bmatrix} = \mathbf{x}^T \mathbf{A} \mathbf{y}, \tag{14}$$

$$e_X(x/\alpha, y) = [1, x, x^2, \dots, x^n] \begin{bmatrix} B'_{00} & B'_{01} & B'_{02} & \dots & B'_{0,n} \\ B'_{10} & B'_{11} & \dots & B'_{1,n-1} & \\ B'_{20} & \dots & B'_{2,n-2} & & \\ \vdots & & & & \\ B'_{n0} & & & & 0 \end{bmatrix} \begin{bmatrix} 1 \\ y \\ y^2 \\ \vdots \\ y^n \end{bmatrix} = \mathbf{x}^T \mathbf{B}' \mathbf{y}, \quad (15)$$

where the matrices **A** and **B'** contain the undetermined coefficients of the fitted bivariate polynomials. As noted from Equations (14) and (15), here the set of *n*-degree bivariate polynomials are used to fit the form errors of the scale grating. In addition, the highest degree for *x* and *y* coordinates is not necessary to be identical, but it depends on the actual fitting performance. With Equations (14) and (15), the  $e_Z(x/\alpha, y)$  and  $e_X(x, y)$  can be obtained, as follows:

$$e_Z(x/\alpha, y) = \mathbf{x}^T \mathbf{A}' \mathbf{y}, \quad (16)$$

$$e_X(x, y) = \mathbf{x}^T \mathbf{B} \mathbf{y}, \quad (17)$$

where

$$\mathbf{A}' = \begin{bmatrix} 1 & & & & \\ & 1/\alpha & & & \\ & & 1/\alpha^2 & & \\ & & & \ddots & \\ & & & & 1/\alpha^n \end{bmatrix} \begin{bmatrix} A_{00} & A_{01} & A_{02} & \dots & A_{0,n} \\ A_{10} & A_{11} & \dots & A_{1,n-1} & \\ A_{20} & \dots & A_{2,n-2} & & \\ \vdots & & & & \\ A_{n0} & & & & 0 \end{bmatrix}, \quad (18)$$

$$\mathbf{B} = \begin{bmatrix} 1 & & & & \\ & \alpha & & & \\ & & \alpha^2 & & \\ & & & \ddots & \\ & & & & \alpha^n \end{bmatrix} \begin{bmatrix} B'_{00} & B'_{01} & B'_{02} & \dots & B'_{0,n} \\ B'_{10} & B'_{11} & \dots & B'_{1,n-1} & \\ B'_{20} & \dots & B'_{2,n-2} & & \\ \vdots & & & & \\ B'_{n0} & & & & 0 \end{bmatrix} \quad (19)$$

According to Equations (14), (16), and (18), once the matrix **A** is obtained, the  $e_Z(x/\alpha, y)$ ,  $\partial e_Z(x/\alpha, y) / \partial \alpha$ ,  $\sigma^2[e_Z(x/\alpha, y)]$  can also be calculated successively, and finally  $\sigma^2[e_X(x/\alpha, y)]$  by Equation (12),  $e_Z(x/\alpha, y)$  by Equation (16),  $e_X(x, y)$  by Equations (17) and (19), and also  $\sigma^2[e_X(x, y)]$ . With the obtained  $e_Z(x, y)$ ,  $\sigma^2[e_Z(x, y)]$ ,  $e_X(x, y)$ , and  $\sigma^2[e_X(x, y)]$ , evaluation of the PV values and the associated uncertainty for both  $e_Z(x, y)$  and  $e_X(x, y)$  can also be conducted. Since the procedures for obtaining the matrices **A** and **B'** are similar, only the determination of the matrix **A** is presented in the following.

Taking the vectorization operator on both sides of Equation (14) leads to

$$\text{vec}[e_Z(x, y)] = (\mathbf{y}^T \otimes \mathbf{x}^T) \text{vec}(\mathbf{A}), \quad (20)$$

where  $\text{vec}(\cdot)$  signifies the vectorization operator that converts a matrix into a column vector, and the symbol  $\otimes$  represents the Kronecker product. Discarding the zero elements in  $\text{vec}(\mathbf{A})$  and denoting the processed  $\text{vec}(\mathbf{A})$  as a vector  $\mathbf{a} = [a_0, a_1, a_2, \dots, a_{K-1}]^T$  ( $K = (n + 1)(n + 2)/2$ ) and then choosing new basis functions  $\Xi_k$  ( $k = 0, 1, 2, \dots, K - 1$ ), where  $\Xi_k \{1, x, y, x^2, xy, y^2, \dots, xy^{n-1}, y^n\}$ , the Equation (20) can be rewritten, as follows:

$$\eta(\xi) = a_0 \Xi_0(\xi) + a_1 \Xi_1(\xi) + a_2 \Xi_2(\xi) + \dots + a_{K-1} \Xi_{K-1}(\xi), \quad (21)$$

where the variable  $\xi$  represents the value of the terms  $x, y, x^2, xy, y^2, \dots, xy^{n-1}, y^n$  at any pixel position  $(x_i, y_j)$  ( $i = 1, 2, \dots, M; j = 1, 2, \dots, N$ ), and  $\eta(\xi)$  represents the value of  $\text{vec}[e_Z(x, y)]$  at the corresponding

pixel position. To determine the elements in vector  $\mathbf{a}$ , the general linear least-squares method [20] is used by defining a merit function, as follows:

$$\chi^2 = \sum_{l=1}^{MN} \left[ \frac{\eta_l - \sum_{k=0}^{K-1} a_k \Xi_k(\xi_l)}{\sigma(\eta_l)} \right]^2, \tag{22}$$

where  $\eta_l$  corresponds to the value  $E[e_Z(x_i, y_j)]$  given in Equation (9) and  $\sigma(\eta_l)$  is corresponding standard deviation  $\sigma[e_Z(x_i, y_j)]$ , as obtained by Equation (11). In addition, an  $MN \times K$  matrix  $\mathbf{D}$ , which is also called the design matrix of the fitting problem, is defined with elements given as follows:

$$D_{lk} = \frac{\Xi_k(\xi_l)}{\sigma(\eta_l)}. \tag{23}$$

Also, define a vector  $\mathbf{b}$  of length  $MN$  with elements, as follows:

$$b_l = \frac{\eta_l}{\sigma(\eta_l)}. \tag{24}$$

The minimum of Equation (22) occurs where the derivative of  $\chi^2$  with respect to all  $K$  parameters  $a_k$  vanish. This condition yields the following matrix equation:

$$(\mathbf{D}^T \mathbf{D}) \mathbf{a} = \mathbf{D}^T \mathbf{b}. \tag{25}$$

According to Equation (25), the least-squares estimation of vector  $\mathbf{a}$  can be obtained, as follows:

$$\hat{\mathbf{a}} = (\mathbf{D}^T \mathbf{D})^{-1} \mathbf{D}^T \mathbf{b} = \mathbf{D}^+ \mathbf{b}, \tag{26}$$

where  $\mathbf{D}^+ = (\mathbf{D}^T \mathbf{D})^{-1} \mathbf{D}^T$  is the Moore-Penrose pseudo-inverse of  $\mathbf{D}$ . Let's denote the matrix  $(\mathbf{D}^T \mathbf{D})^{-1}$  as  $\mathbf{C}$ , namely,  $\mathbf{C} = (\mathbf{D}^T \mathbf{D})^{-1}$ , which is also called the covariance matrix and it has a close relation with the uncertainties of the estimated parameters  $a_k$  ( $k = 0, 1, 2, \dots, K - 1$ ) by Equation (26). Specifically, the variance in  $a_k$  can be evaluated by

$$\sigma^2(\hat{a}_k) = C_{kk}, \tag{27}$$

where  $C_{kk}$  is the  $k$ -th diagonal element of the matrix  $\mathbf{C}$ . According to Equations (26) and (27), the least-square estimation  $A_{pq}$  ( $p, q = 0, 1, 2, \dots, n$ ) of all elements in matrix  $\mathbf{A}$  as well as their uncertainties  $u(A_{pq})$  can be finally obtained.

With the estimated elements  $A_{pq}$  and their uncertainties  $u(A_{pq})$ , the uncertainty of  $e_Z(x, y)$  at any pixel position  $(x_i, y_j)$  can be further evaluated. Specifically, according to Equation (14), there is  $e_Z(x_i, y_j) = \mathbf{x}_i^T \mathbf{A} \mathbf{y}_j$ , which can be regarded as a linear function of the coefficients  $A_{pq}$ . The evaluation of the upper and lower bounds of  $e_Z(x_i, y_j)$  can be transformed into a simple linear programming problem [21], described as follows:

$$\begin{cases} \max e_Z(x_i, y_j) = \mathbf{x}_i^T \mathbf{A} \mathbf{y}_j \\ \text{s.t. } A_{pq} \in [\hat{A}_{pq} - u(\hat{A}_{pq}), \hat{A}_{pq} + u(\hat{A}_{pq})], \\ \quad p, q = 0, 1, 2, \dots, n \end{cases}, \tag{28a}$$

$$\begin{cases} \min e_Z(x_i, y_j) = \mathbf{x}_i^T \mathbf{A} \mathbf{y}_j \\ \text{s.t. } A_{pq} \in [\hat{A}_{pq} - u(\hat{A}_{pq}), \hat{A}_{pq} + u(\hat{A}_{pq})], \\ \quad p, q = 0, 1, 2, \dots, n \end{cases}, \tag{28b}$$

where Equations (28a) and (28b) yield the upper and lower bounds of  $e_Z(x_i, y_j)$ , respectively. The uncertainty of  $e_Z(x, y)$  at the pixel position of  $(x_i, y_j)$  is then evaluated by

$$u[e_Z(x_i, y_j)] = \frac{\overline{e_Z}(x_i, y_j) - e_Z(x_i, y_j)}{2}. \tag{29}$$

In addition, according to Equation (18), the least-squares estimation  $A'_{pq}$  of elements in matrix  $A'$  can be calculated, as follows:

$$\hat{A}'_{pq} = \frac{1}{\alpha^p} \hat{A}_{pq}. \tag{30}$$

It follows that the uncertainty in  $A'_{pq}$  can be evaluated by the following equation:

$$u^2(\hat{A}'_{pq}) = \left(\frac{p}{\alpha^{p+1}} \hat{A}_{pq}\right)^2 u^2(\alpha) + \left(\frac{1}{\alpha^p}\right)^2 u^2(\hat{A}_{pq}). \tag{31}$$

With  $A'_{pq}$  and  $u(A_{pq})$ , the values of  $e_Z(x/\alpha, y)$  and  $\partial e_Z(x/\alpha, y)/\partial \alpha$  can be obtained at any pixel position of  $(x_i, y_j)$ . The uncertainty of  $e_Z(x/\alpha, y)$  at any pixel position  $(x_i, y_j)$  can be also evaluated in a similar manner to Equation (28). Likewise, the value of  $e_X(x, y)$  and its uncertainty  $u[e_X(x, y)]$  at any pixel position  $(x_i, y_j)$  can be calculated as well.

The PV value of  $e_Z(x, y)$  in the region  $\Omega$  of interest can be evaluated by

$$PV[e_Z(x, y)] = \max[e_Z(x, y)] - \min[e_Z(x, y)], \quad (x, y) \in \Omega, \tag{32}$$

where  $\max(\cdot)$  and  $\min(\cdot)$  represent the maximum and minimum of a function. The peak and valley values of  $e_Z(x, y)$ ,  $\max[e_Z(x, y)]$ , and  $\min[e_Z(x, y)]$  can be readily calculated according to the obtained expression of  $e_Z(x, y)$ . The uncertainty in the calculated PV of  $e_Z(x, y)$  can then be evaluated, as follows:

$$u^2\{PV[e_Z(x, y)]\} = u_p^2[e_Z(x, y)] + u_v^2[e_Z(x, y)], \tag{33}$$

where  $u_p[e_Z(x, y)]$  and  $u_v[e_Z(x, y)]$  represent the uncertainties of  $e_Z(x, y)$  at the peak and valley positions, respectively, which can be directly obtained by substituting the coordinates of the peak and valley positions into Equation (29).

Other approaches to evaluate the uncertainty in PV could simply double the maximum value of the uncertainty matrix or applying the Monte Carlo (MC) simulation. However, the former method obviously overestimates the uncertainty in the PV, except coincidence. The MC method could estimate the PV uncertainty by collecting the PV value in each trail assuming that all of the significant errors are included and the probability distributions of each error source is known. Nevertheless, it is burdensome to run the simulation when routine experiment is conducted in a similar condition and the noise could exist explicitly in the evaluated results. Therefore, in this paper, the MC method is not employed, but just the part of error form is analyzed directly to evaluate the uncertainty in the peak and valley separately, and then combine them together. Since the PV of each error form is usually larger than the uncertainty in this experiment, this method is thus a reasonable way to obtain the uncertainty in PV.

As can be noted from the above description, the success of the proposed evaluation method lies on the fitting performance of the polynomials fitted to  $e_Z(x, y)$  and  $e_X(x/\alpha, y)$ . There are two measures that are commonly used to evaluate the goodness-of-fit in statistics. One is the coefficient of determination  $R^2$ , while another is the reduced chi-square  $\chi^2_v$ , which for the polynomial fitted to  $e_Z(x, y)$  are defined, as follows:

$$R^2 = 1 - \frac{\sum_{l=1}^{MN} \left[ \eta_l - \sum_{k=0}^{K-1} \hat{a}_k \Xi_k(\xi_l) \right]^2}{\sum_{l=1}^{MN} \left[ \eta_l - \sum_{l=1}^{MN} \eta_l / MN \right]^2}, \tag{34a}$$

$$\chi^2_v = \frac{\chi^2_{\min}}{v} = \frac{1}{MN - K} \sum_{l=1}^{MN} \left[ \frac{\eta_l - \sum_{k=0}^{K-1} \hat{a}_k \Xi_k(\xi_l)}{\sigma(\eta_l)} \right]^2, \tag{34b}$$

where  $v = MN \times K$  signifies the degree-of-freedom in the least-square regression. Typically,  $0 < R^2 < 1$ , and the larger the  $R^2$  is, the better the polynomial model fits the data. For the reduced chi-square,  $\chi_v^2 = 1$  indicates that the model properly fits the data, while  $\chi_v^2 > 1$  indicates that the fit does not fully capture the data (or that the variance  $\sigma^2(\eta_l)$  is underestimated) and  $\chi_v^2 < 1$  indicates that the model overly fits the data (either the model improperly fits the noise or the variance  $\sigma^2(\eta_l)$  is overestimated). According to the value of  $\chi_v^2$ , the variance of  $a_k$  that is represented in Equation (27) is usually rescaled as follows in order to pass the  $\chi^2$ -test:

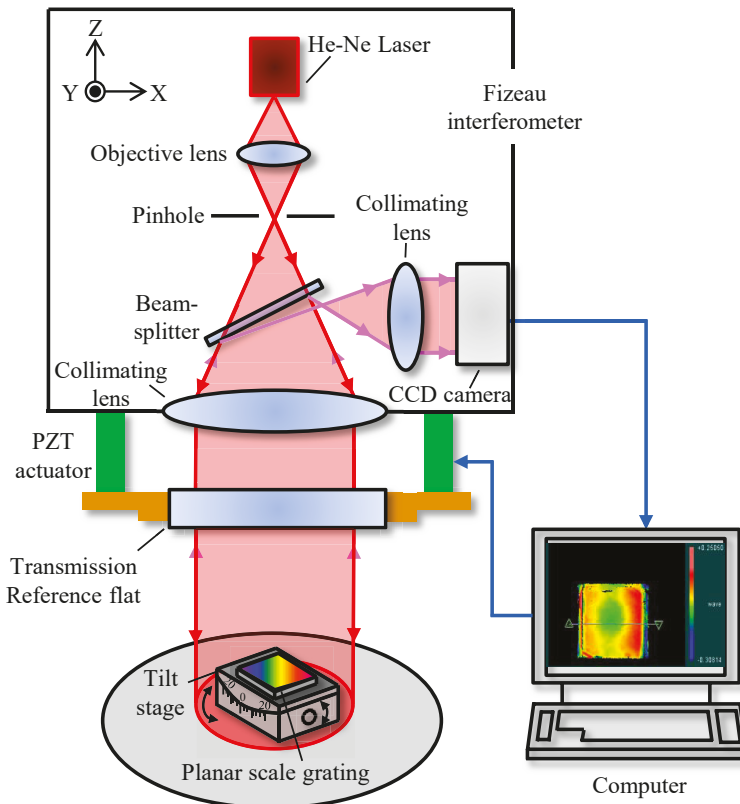
$$\sigma_r^2(\hat{a}_k) = \chi_v^2 C_{kk} \tag{35}$$

In addition, as illustrated in Equation (26), the least-squares estimation of vector  $\mathbf{a}$  can be obtained by calculating the Moore-Penrose pseudo-inverse of the design matrix  $\mathbf{D}$ . It should be pointed out that this solving approach is susceptible to round-off error in practice and it will lead to errors in both the solution  $\mathbf{a}$  and its uncertainty, especially when the variance  $\sigma^2(\eta_l)$  (corresponds to the variance  $\sigma^2(e_z(x_i, y_j))$ ) is not uniform at different pixel positions. To relieve the round-off problem, the solution by use of the singular value decomposition is recommended [22].

## 4. Experiments

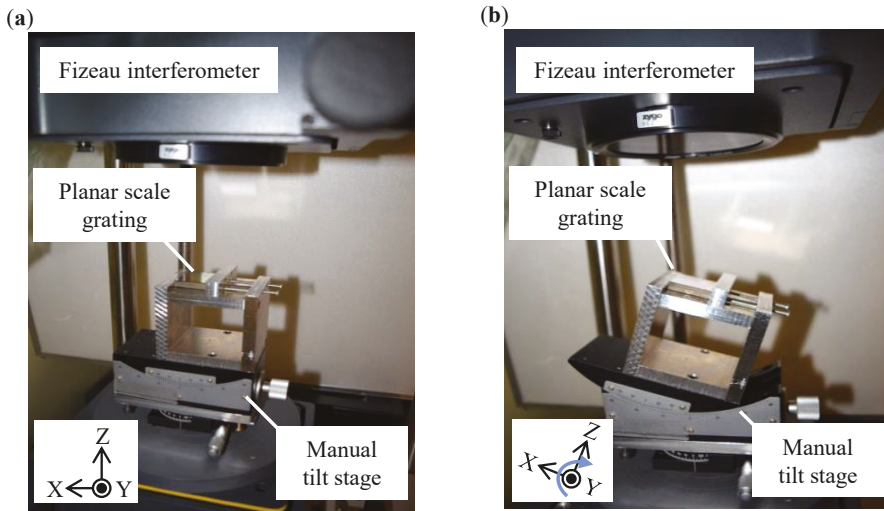
### 4.1. Experimental Setup

A commercial Fizeau interferometer (Verifire™, Zygo Corp, Middlefield, CT, USA) having a laser source with a wavelength of 632.8 nm was used in experiments. A field-of-view (FOV) of the Fizeau interferometer was 100 mm in diameter. Vertical and horizontal resolutions of the interferometer were 0.05 nm and 300 μm, respectively. A nominal PV value of the reference optical flat in the interferometer was less than  $\lambda/20$  over the whole FOV. Figure 3 shows the experimental setup with the Fizeau interferometer. In the experiments, a reflective-type planar scale grating in a size of 35 mm × 30 mm that was fabricated by interference lithography was tested. This holographic planar scale grating had a nominal grating pitch of 1 μm along the X- and Y-directions. The scale grating was first mounted on a manual tilt stage and then together mounted on the sample stage in the Fizeau interferometer. Phase shifts in the wavefronts of diffracted beams caused by the out-of-flatness and pitch deviations of the scale grating can directly be measured by the Fizeau interferometer by setting the scale grating in the Littrow configuration. Although measuring higher-order diffracted beams would be helpful to enhance the measurement sensitivity, it requires a larger rotation angle, resulting in degradations of the quality of measured wavefront and the lateral resolution along the direction perpendicular to the grating grooves. Therefore, only the first-order diffracted beams were measured in the following experiments.



**Figure 3.** Experiment setup with a Fizeau interferometer. PZT: piezoelectric ceramic transducer; CCD: charge-coupled Device.

The zeroth-order diffracted wavefront was firstly measured to evaluate the out-of-flatness of the planar scale grating. After that, the scale grating was rotated clockwise and counterclockwise about the Y-axis by using the manual tilt stage to measure the X-directional positive and negative first-order diffracted beams, respectively. The inclination angle of the planar scale grating was adjusted by the manual tilt stage. With the visual feedback of the fringe pattern from the monitor in the Fizeau interferometer system, the number of interference fringes can be made to be minimum while adjusting the tilt stage. In the same manner, the Y-directional positive and negative first-order diffracted beams were also measured by rotating the tilt stage 90 degrees about the Z-axis. Figure 4a,b show photographs of the setup for measurement of the zeroth-order diffracted beam and the X-directional positive and negative first-order diffracted beams, respectively. Totally, it spent about 10 min to complete the experiments procedure, including the adjustment of the tilt stage and the alignment process of the Fizeau interferometer to minimize the number of fringes. Once the positive and negative first-order diffracted beams in both the X- and Y-directions are measured, the uncertainty evaluation can be conducted based on the models that are proposed in the previous section.



**Figure 4.** Photographs of measurement of the wavefronts of the diffracted beams. (a) 0th-order diffracted beam; (b) X-directional +1st-order diffracted beam.

#### 4.2. Results and Discussions

The form errors of the tested grating were first fitted by the set of  $n$ -degree bivariate polynomials, as proposed in Section 3. The highest degree  $n$  of the  $x$  and  $y$  coordinates in the respective model was set to be identical, which was 5 for the fitting of the  $Z$ -directional out-of-flatness error map and was 4 for the fitting of the  $X$  and  $Y$ -directional pitch deviations error maps for the sake of clarity. The regression surface may be too sharp or wiggly in the higher degree case. In addition, the higher degree polynomials will have problems with numerical accuracy. Therefore, the highest degree of the  $n$ -degree bivariate polynomials for each fitting model was selected not to exceed 5. Before the least square approximation, the  $x$  and  $y$  coordinates were normalized by the  $Z$ -score method [23] for a better fitting performance. In addition, the calculated variances from Equations (11) and (12) were used as the weights in the weighted least squared fitting process for fitting the corresponding error forms to make the approximation more precise. The fitting performance characterized by the proposed criterion in Section 3 was summarized in Table 1. As can be seen in the table, the values of  $R^2$  and  $\chi_v$  of each model indicate that the set of polynomials well captured the error forms in the three directions. The small root-mean-square error (RMSE) values also show that the regression surface closely fits the measured form errors without the loss of precision. In the following, the comprehensive comparisons of the calibration results and polynomials fitting results of each directional form error are obtained.

**Table 1.** Fitting performance of the polynomial models for the form errors of the grating. RMSE: root-mean-square error.

Form Error Type	$R^2$	$\chi_v$	RMSE	Polynomials Degree
$e_z(x,y)$	0.9844	0.9463	2.1599	5
$e_x(x',y')$	0.9320	0.2905	2.7534	4
$e_y(x,y')$	0.8608	0.1522	4.5063	4

An area of 20 mm × 20 mm over the scale grating located near the center of FOV was extracted for the evaluations. Figure 5a shows the out-of-flatness of the planar scale grating that was obtained in the experiment. The PV value was evaluated to be 85.80 nm. Figure 5b shows the polynomials fitting result of the measured out-of-flatness, as shown in Figure 5a. The PV value was evaluated to be 75.16 nm.

To take a closer look at the fitting detail, the central cross-sections of the two results were extracted and compared with each other. Figure 6a,b show the X- and Y-directional central cross-sections of the out-of-flatness shown in Figure 5. The differences in the PV values between the measured result and the fitted results of the X- and Y-directional cross-sections were evaluated to be 13.71 nm and 6.64 nm, respectively. The X'-directional pitch deviation was also evaluated. Figure 7 shows the measured X'-directional deviation ( $e_X(x',y)$ ) obtained by analyzing the phase outputs of the positive and negative first-order diffracted beams that were measured by the commercial Fizeau interferometer. Figure 8 shows the X- and Y-directional central cross-sections of the pitch deviation that are shown in Figure 7. The PV of the X'-directional pitch deviation from the measured and fitted results were evaluated to be 53.06 nm and 39.89 nm, respectively.

In the same manner, the Y'-directional deviation ( $e_Y(x,y')$ ) was also evaluated as shown in Figure 9. The PV of the Y'-directional pitch deviation from the measured and fitted results were evaluated to be 79.68 nm and 68.52 nm, respectively, as shown in Figure 9. It should be noted that the PV values of the central XY axes cross-sections from the X'- and Y'-directional deviations were almost within 10 nm, as shown in Figures 8 and 10; when considering the existence of the high frequency components in the original measurement data, the actual fitting error may be even smaller. Therefore, the fitted results could express the measured forms well with the existence of noise even though the PV value was small, which is acceptable for the precision requirement of the proposed evaluation methods.

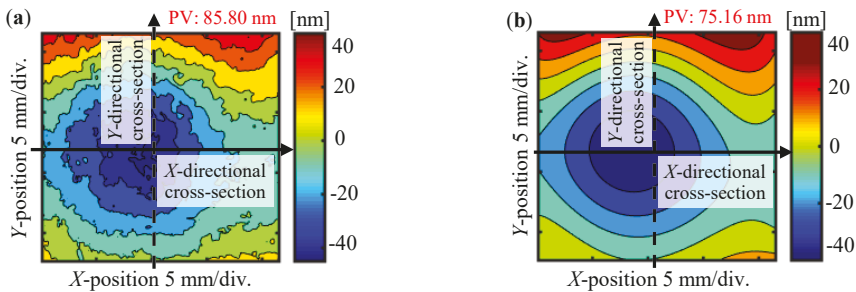


Figure 5. Measured and fitted results of the Z-directional out-of-flatness. (a) Measured Z-directional out-of-flatness; (b) Z-directional out-of-flatness fitted by a polynomial. PV: peak-to-valley.

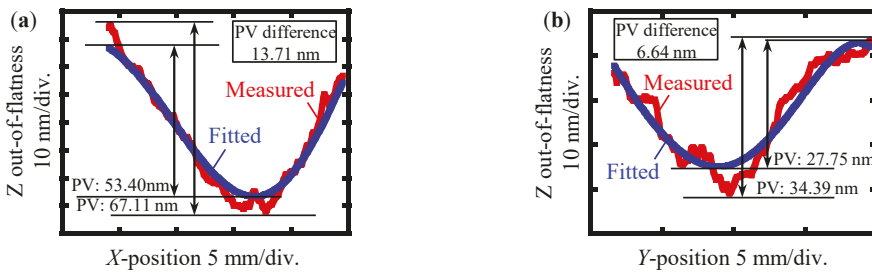
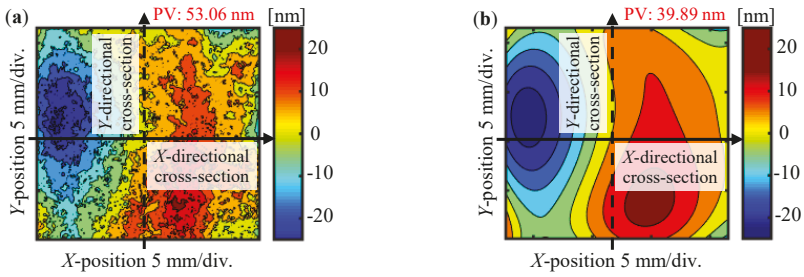
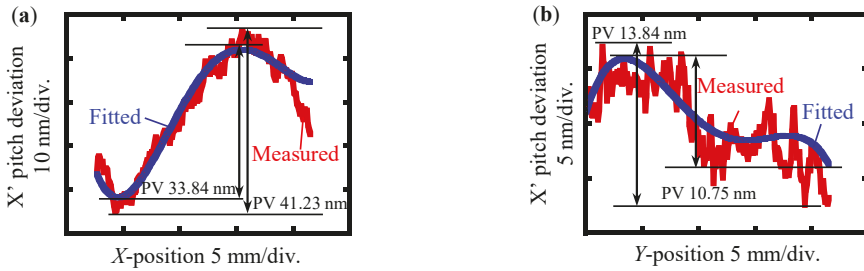


Figure 6. Central cross-sections of the Z-directional out-of-flatness shown in Figure 5. (a) Central cross-section along the X-direction; (b) Central cross-section along the Y-direction.

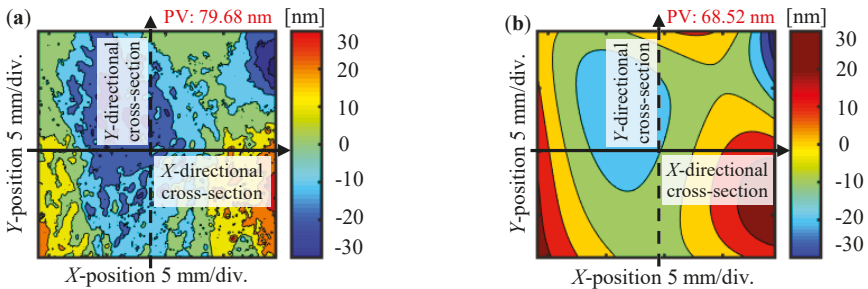




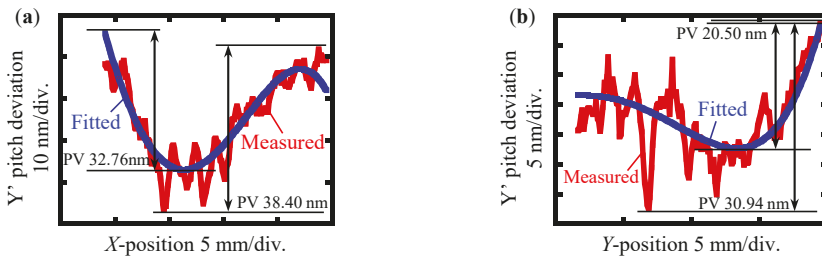
**Figure 7.** Measured and fitted results of the  $X'$ -directional pitch deviation. (a) Measured  $X'$ -directional pitch deviation (b)  $X'$ -directional pitch deviation fitted by a polynomial.



**Figure 8.** Central cross-sections of the  $X'$ -directional pitch deviation shown in Figure 7. (a) Central cross-section along the X-direction; (b) Central cross-section along the Y-direction.

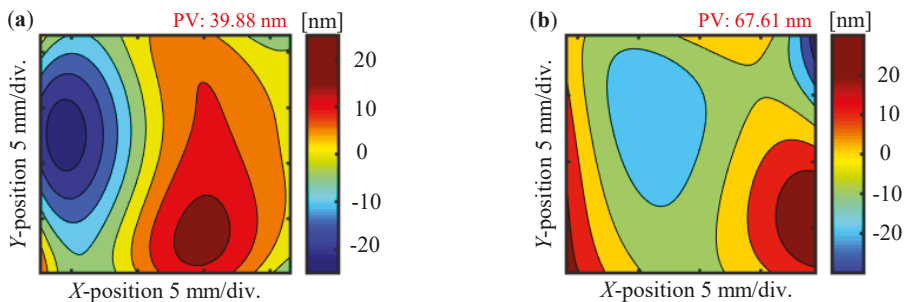


**Figure 9.** Measured and fitted results of the  $Y'$ -directional pitch deviation. (a) Measured  $Y'$ -directional pitch deviation (b)  $Y'$ -directional pitch deviation fitted by a polynomial.



**Figure 10.** Central cross-sections of the  $Y'$ -directional pitch deviation shown in Figure 9. (a) Central cross-section along the X-direction; (b) Central cross-section along the Y-direction.

The error maps of the X- and Y- directional pitch deviations were finally obtained based on the fitted results of  $e_X(x',y)$  and  $e_Y(x,y')$ , as shown in Figure 11a,b. The PV values of the fitted X- and Y-directional pitch deviations were evaluated to be 39.88 nm and 67.61 nm, respectively. It should be noted that the differences of the PV values in the fitted pitch deviation errors between the original X-, Y-directional results shown in Figure 11 and the tilted  $X',Y'$ -directional results shown in Figures 7b and 9b were evaluated to be 0.01 nm and 0.91 nm, respectively, which were negligibly small, since the required Littrow angle was small and the newly calculated regression coefficients were similar to the original coefficients according to Equation (30). In addition, the analyzed area has slightly shifted after changing the coordinate system in the Littrow configuration back to the original grating coordinate system. It could be readily explained by the cosine impact from the inclination angle, which would induce a shift of the analyzed area with a distance determined by the cosine value of the inclination angle. However, since the influence of the inclination angle is already considered in the regression coefficients calculation of the polynomial fitting and the uncertainty evaluation process for the new area according to Equations (30) and (31), the finally obtained PV values and the PV uncertainties will thus not be overestimated after changing the coordinate system.



**Figure 11.** Measurement and fitted results of the X- and Y-directional pitch deviations. (a) X-directional pitch deviation of polynomial fitted results; (b) Y-directional pitch deviation of polynomial fitted results.

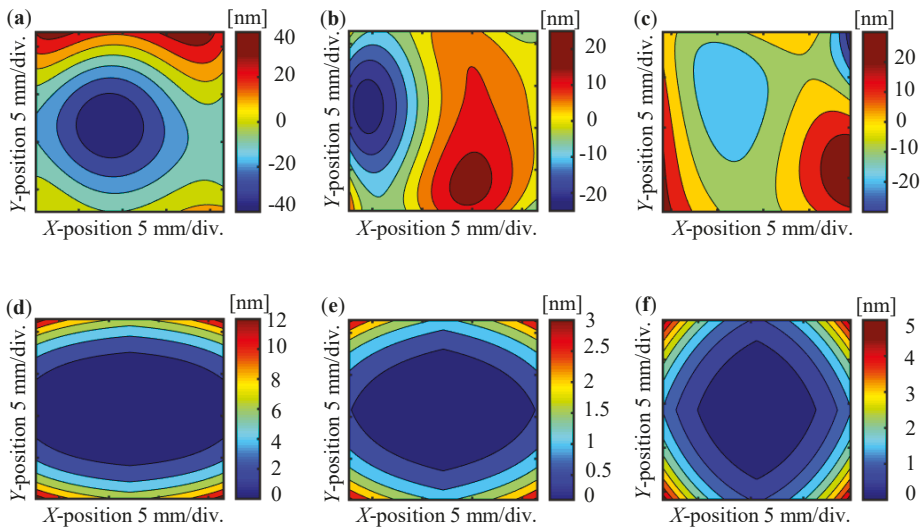
Table 2 shows the uncertainty budget. Since the complete model is quite complex, only the major uncertainty sources were taken into consideration, as discussed in Section 3. During the period of experiment, the environment effects, such as the air pressure, relative humidity, and temperature, could induce a drift in air refractive index as well as the laser wavelength. However, the refractive index drift and laser wavelength fluctuation in the whole measurement can be small under a well-controlled experimental environment with a limited testing time, which could be roughly estimated to be several picometers [24] and the caused combined errors are negligibly small. Therefore, the environmental effects and laser wavelength fluctuation are not problems in this experiment and its uncertainty is omitted in Equation (12). In addition, the angular displacement errors of the scale grating about the X- and Y-axes during the experiment were small, thus the misalignment angles due to the rotation error with respect to the X- and Y-axes do not contribute significantly to the overall combined uncertainty. Moreover, since this paper has focused on the form errors and PV uncertainties evaluation of the planar scale grating, the rotational error about the Z-axis is not a problem either in this case; the phase shifts information in the diffracted wavefronts will not change and the rotation error about the Z-axis is believed to be very small with careful adjustments.

Table 2. Uncertainty budget.

Uncertainty Source	Symbol	Type	Standard Uncertainty
Z-directional zeroth-order phase outputs	$u[I_0(x,y)]$	A	0.0825 rad
Out-of-flatness of reference flat	$u[e_R(x,y)]$	B	2.11 nm
Regression coefficients matrix	$u(A)$	A	Multiple values
PV uncertainty of $e_Z(x,y)$ ( $k = 2$ )	$u\{PV [e_Z(x,y)]\}$		4.12 nm
X-directional positive first-order phase outputs	$u[I_{X+1}(x,y)]$	A	0.0761 rad
X-directional negative first-order phase outputs	$u[I_{X-1}(x,y)]$	A	0.0347 rad
Grating pitch	$u(g)$	B	100 nm
Mechanical adjustment error of the tilt stage	$u(\epsilon_{1,2})$	B	0.1°
Regression coefficients matrix	$u(B)$	A	Multiple values
PV uncertainty of $e_X(x,y)$ ( $k = 2$ )	$u\{PV [e_X(x,y)]\}$		1.60 nm
Y-directional positive first-order phase outputs	$u[I_{Y+1}(x,y)]$	A	0.8420 rad
Y-directional negative first-order phase outputs	$u[I_{Y-1}(x,y)]$	A	0.5013 rad
Grating pitch	$u(g)$	B	100 nm
Mechanical adjustment error of the tilt stage	$u(\epsilon'_{1,2})$	B	0.1°
Regression coefficients matrix	$u(C)$	A	Multiple values
PV uncertainty of $e_Y(x,y)$ ( $k = 2$ )	$u\{PV [e_Y(x,y)]\}$		10.72 nm

Since the tested location of the planar scale grating was near the center of the FOV of the Fizeau interferometer, the PV value of the reference optical flat was expected to be far better than  $\lambda/20$  in the experiment. In the uncertainty analysis, the standard uncertainty of the out-of-flatness of the reference optical flat is thereby taken to be  $\sigma[e_R(x,y)] = \lambda/300$ , according to the size of the analyzed region of the grating under test. The standard uncertainty of the grating pitch was taken as  $u(g) = 100$  nm, which was obtained by the specification sheet of the grating provided by the grating manufacturer, in addition, the uncertainty was taken as the same in the X- and Y-directions approximately, since the evaluated values were roughly the same. The standard uncertainty induced by the mechanical adjustment errors of the manual tilt stage was taken as  $u(\epsilon) = 0.1^\circ$ , according to the specification of the tilt stage. The rectangular distribution was assumed for the mechanical adjustment error according to the scale interval of the vernier of the manual tilt stage. For the sake of simplicity, the uncertainty components were considered to be the same in the X- and Y-directions. The mean values as well as the uncertainties of the measured phase outputs  $I_0(x,y)$ ,  $I_{X\pm 1}(x,y)$  and  $I_{Y\pm 1}(x,y)$  were estimated by the data from six attempts. The standard uncertainty of each pixel in the analyzed area was calculated followed by the Type A standard uncertainty evaluation approach. The maximum value was adopted as the repeatability ( $1\sigma$ ) of the corresponding phase output. The uncertainty of the regression coefficients matrices contains multiple values and the detailed results are presented in Appendix A.

Figure 12 shows the obtained error maps as well as the corresponding uncertainty maps of the planar scale grating over the analyzed area. Figure 12a–c show the error maps of the Z-directional out-of-flatness, X- and Y-directional pitch deviation of the planar scale grating, respectively. Figure 12d–f present the corresponding standard uncertainty maps of the form errors according to Equations (28) and (29). The finally evaluated PV uncertainties in the above three errors were 4.21 nm, 1.60 nm, and 10.72 nm (with a confidence interval of approximately 95%), respectively. The uncertainty maps indicate that the regression models fit closely to the original error forms, which yield a small uncertainty value in each pixel. On the other hand, the estimated PV uncertainties are believed to coincide with the results using another approach with the uncertainty matrix. By adding and subtracting the uncertainty matrix to the calibration data set, the standard uncertainty of the peak ( $u_p$ ) and valley ( $u_v$ ) of the form errors expected to lie could be obtained separately. The PV uncertainties could be obtained in exactly the same manner as Equation (33), which would be similar to the results when using the proposed PV uncertainty evaluation method in the observed directions. This part of work will be conducted as future work.



**Figure 12.** Error and uncertainty maps of the planar scale grating. (a) Error map of the Z-directional out-of-flatness; (b) Error map of the X-directional pitch deviation; (c) Error map of the Y-directional pitch deviation; (d) Uncertainty map of the Z-directional out-of-flatness; (e) Uncertainty map of the X-directional pitch deviation; and, (f) Uncertainty map of the Y-directional pitch deviation.

## 5. Conclusions

The Z-directional out-of-flatness as well as the X- and Y-directional pitch deviations of a reflective-type XY planar grating have been evaluated by analyzing the wavefronts of diffracted beams that were obtained by a Fizeau interferometer in the Littrow configuration. The experiment has been conducted to verify the proposed method by evaluating the planar scale grating with a nominal pitch of 1  $\mu\text{m}$  along both the X- and Y-directions by a commercial Fizeau interferometer having a field-of-view (FOV) of 100 mm equipped with a laser source with a wavelength of 632.8 nm. The X-, Y-, and Z-directional error maps over an area of 20 mm  $\times$  20 mm of the scale grating located near the center of the FOV have been obtained, and the small difference between the measured results and the fitted results have demonstrated the validity of the proposed measurement method. It should be noted that the proposed method can also be applied to measure the form errors of various types of the reflective-type diffraction gratings with a planar surface. In addition, intensive uncertainty analyses have been performed, while relative theoretical equations have been derived by taking the possible error factors, such as the error in the reference flat and the errors in the grating inclination angle, into consideration. The X-, Y-, and Z- directional uncertainty maps as well as the PV uncertainty of each error forms have been obtained by the proposed evaluation procedure. As a result, the PV uncertainties have been evaluated to be 4.12 nm, 1.60 nm, and 10.72 nm for the Z-directional out-of-flatness, X- and Y-directional pitch deviations of the planar scale grating, respectively. Future research will focus on how to reduce the uncertainties in the measurement results.

**Author Contributions:** Conceptualization, W.G., X.C. and Y.S.; Methodology, X.C. and X.X.; Software, X.X.; Validation, X.C. and X.X.; Formal Analysis, X.C., X.X.; Investigation, X.C. and X.X.; Resources, X.X. and X.C.; Data Curation, X.X., H.M. and X.C.; Writing-Original Draft Preparation, X.X., X.C. and Y.S.; Writing-Review & Editing, W.G. and Y.S.; Visualization, W.G. and Y.S.; Supervision, W.G.; Project Administration, W.G.; Funding Acquisition, W.G. and Y.S.

**Funding:** This work is supported by Japan Society for the Promotion of Science (JSPS).

**Conflicts of Interest:** The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, and in the decision to publish the results.

### Appendix A

According to Equations (14) and (17), there is  $e_z(x,y) = x^T \mathbf{A}y$ ,  $e_x(x,y) = x^T \mathbf{B}y$  and, in the same manner,  $e_y(x,y) = x^T \mathbf{C}y$ . The least-square estimation  $A_{pq}, B_{pq}, C_{pq}$  ( $p,q = 0, 1, 2, \dots, n$ ) of all the elements in the corresponding regression coefficients matrices  $A, B, C$  as well as their uncertainties  $u(A_{pq}), u(B_{pq}), u(C_{pq})$  can thus be obtained according to Equations (26) and (27). Finally, the uncertainty of the regression coefficients matrices  $u(A), u(B), u(C)$  can be obtained as follows:

$$u(A) = \begin{bmatrix} 0.0294 & 0.0569 & 0.0441 & 0.0636 & 0.0158 & 0.0182 \\ 0.0569 & 0.0364 & 0.0504 & 0.0137 & 0.0158 & 0 \\ 0.0441 & 0.0504 & 0.0135 & 0.0154 & 0 & 0 \\ 0.0635 & 0.0137 & 0.0154 & 0 & 0 & 0 \\ 0.0157 & 0.0158 & 0 & 0 & 0 & 0 \\ 0.0181 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

$$u(B) = \begin{bmatrix} 0.0375 & 0.0374 & 0.0563 & 0.0175 & 0.0201 \\ 0.0696 & 0.0439 & 0.0146 & 0.0206 & 0 \\ 0.1332 & 0.0921 & 0.0229 & 0 & 0 \\ 0.0447 & 0.0278 & 0 & 0 & 0 \\ 0.0697 & 0 & 0 & 0 & 0 \end{bmatrix}$$

$$u(C) = \begin{bmatrix} 0.0647 & 0.0645 & 0.0970 & 0.0301 & 0.0346 \\ 0.0987 & 0.0755 & 0.0668 & 0.0437 & 0 \\ 0.1109 & 0.0674 & 0.0845 & 0 & 0 \\ 0.1116 & 0.0609 & 0 & 0 & 0 \\ 0.0397 & 0 & 0 & 0 & 0 \end{bmatrix}$$

### References

1. Gao, W. *Precision Nanometrology—Sensors and Measuring Systems for Nanomanufacturing*; Springer: London, UK, 2010.
2. Gao, W.; Kim, S.W.; Bosse, H.; Haitjema, H.; Chen, Y.L.; Lu, X.D.; Knapp, W.; Weckenmann, A.; Estler, W.T.; Kunzmann, H. Measurement technologies for precision positioning. *CIRP Ann. Manuf. Technol.* **2015**, *64*, 773–796. [CrossRef]
3. Teimel, A. Technology and applications of grating interferometers in high-precision measurement. *Precis. Eng.* **1992**, *14*, 147–154. [CrossRef]
4. Gao, W.; Araki, T.; Kiyono, S.; Okazaki, Y.; Yamanaka, M. Precision nano-fabrication and evaluation of a large area sinusoidal grid surface for a surface encoder. *Precis. Eng.* **2003**, *27*, 289–298. [CrossRef]
5. Li, X.; Gao, W.; Muto, H.; Shimizu, Y.; Ito, S.; Dian, S. A six-degree-of-freedom surface encoder for precision positioning of a planar motion stage. *Precis. Eng.* **2013**, *37*, 771–781. [CrossRef]
6. HEIDENHAIN Catalogue. *Two-Coordinate Encoders*; 2016, p. 281. Available online: [https://www.heidenhain.de/fileadmin/pdb/media/img/350457-2L\\_General-Catalog\\_en.pdf](https://www.heidenhain.de/fileadmin/pdb/media/img/350457-2L_General-Catalog_en.pdf) (accessed on 20 September 2018).
7. Gao, W.; Kimura, A. A fast evaluation method for pitch deviation and out-of-flatness of a planar scale grating. *CIRP Ann. Manuf. Technol.* **2010**, *59*, 505–508. [CrossRef]
8. Beers, J.S.; Penzes, W.B. NIST Length Scale Interferometer Measurement Assurance. *NISTIR* **1992**, *4998*, 1–28.
9. Sawabe, M.; Maeda, F.; Yamaryo, Y.; Simomura, T.; Saruki, Y.; Kubo, T.; Sakai, H.; Aoyagi, S. A New Vacuum Interferometric Comparator for Calibrating the Fine Linear Encoders and Scales. *Precis. Eng.* **2004**, *28*, 320–328. [CrossRef]
10. Hocken, R.J.; Trumper, D.L.; Wanga, C. Dynamics and Control of the UNCC/MIT Sub-atomic Measuring Machine. *Ann. CIRP* **2001**, *50*, 373–376. [CrossRef]

11. Weckenmann, A.; Hoffmann, J. Long Range 3D Scanning Tunnelling Microscopy. *Ann. CIRP* **2007**, *56*, 525–528. [[CrossRef](#)]
12. Chen, X.; Shimizu, Y.; Xiong, X.; Chen, Y.L.; Gao, W. Self-calibration of Fizeau interferometer and planar scale gratings in Littrow setup. *Opt. Express* **2017**, *25*, 21567–21582. [[CrossRef](#)] [[PubMed](#)]
13. BIPM; IEC; IFCC; ILAC; ISO; IUPAC; IUPAP; OIML. *Evaluation of Measurement Data—Guide to the Expression of Uncertainty in Measurement (GUM)*; International Bureau of Weight and Measures: Sèvres, France, 2008.
14. Vannoni, M.; Molesini, G. Absolute planarity with three-flat test: An iterative approach with Zernike polynomials. *Opt. Express* **2008**, *16*, 340–354. [[CrossRef](#)] [[PubMed](#)]
15. Schmitz, T.L.; Davies, A.; Evans, C.J.; Parks, R.E. Silicon Wafer Thickness Variation Measurements Using the National Institute of Standards and Technology Infrared Interferometer. *Opt. Eng.* **2003**, *42*, 2281–2290. [[CrossRef](#)]
16. Evans, C.J. Uncertainty evaluation for measurements of peak-to-valley surface form errors. *CIRP Ann. Manuf. Technol.* **2008**, *57*, 509–512. [[CrossRef](#)]
17. Possolo, A.; Lye, H.K. Concepts and tools for the evaluation of measurement uncertainty. *Rev. Sci. Instrum.* **2017**, *88*, 011301. [[CrossRef](#)] [[PubMed](#)]
18. ISO. Part 5: Surface form tolerances. In *Optics and Photonics. Preparation of Drawings for Optical Elements and Systems*; BSI: London, UK, 2007.
19. Fritz, B.S. Absolute calibration of an optical flat. *Opt. Eng.* **1984**, *23*, 379–383. [[CrossRef](#)]
20. Aitken, A.C. IV. —On least squares and linear combination of observations. *Proc. R. Soc. Edinb.* **1936**, *55*, 42–48. [[CrossRef](#)]
21. Zoutendijk, G. *Methods of Feasible Directions: A Study in Linear and Non-Linear Programming*; Elsevier: Amsterdam, The Netherlands, 1960.
22. Golub, G.H.; Reinsch, C. Singular value decomposition and least squares solutions. *Numer. Math.* **1970**, *14*, 403–420. [[CrossRef](#)]
23. Kreyszig, E. *Advanced Engineering Mathematics*; John Wiley & Sons: New York, NY, USA, 2010.
24. Buhr, E.; Michaelis, W.; Diener, A.; Mirandé, W. Multi-wavelength VIS/UV optical diffractometer for high-accuracy calibration of nano-scale pitch standards. *Meas. Sci. Technol.* **2007**, *18*, 667. [[CrossRef](#)]



© 2018 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).

Article

# Design and Testing of a Compact Optical Prism Module for Multi-Degree-of-Freedom Grating Interferometry Application

Xinghui Li <sup>1,\*</sup>, Yaping Shi <sup>1,†</sup>, Xiang Xiao <sup>1</sup>, Qian Zhou <sup>1</sup>, Guanhao Wu <sup>1,2</sup>, Haiou Lu <sup>1</sup> and Kai Ni <sup>1,\*</sup>

- <sup>1</sup> Division of Advanced Manufacturing, Graduate School at Shenzhen, Tsinghua University, Tsinghua Campus, Xili University Town, Shenzhen 518055, China; cdshiyaping@gmail.com (Y.S.); xiaoxiang\_x\_x@126.com (X.X.); zhou.qian@sz.tsinghua.edu.cn (Q.Z.); guanhaowu@tsinghua.edu.cn (G.W.); lho16@mail.tsinghua.edu.cn (H.L.)
- <sup>2</sup> Department of Precision Instrument, Tsinghua University, Haidian District, Beijing 100084, China
- \* Correspondence: li.xinghui@sz.tsinghua.edu.cn (X.L.); ni.kai@sz.tsinghua.edu.cn (K.N.)
- † These authors contributed equally to this work.

Received: 24 October 2018; Accepted: 1 December 2018; Published: 4 December 2018



**Abstract:** In this research, a key optical component for multi-degree-of-freedom (MDOF) surface encoder was designed, fabricated and evaluated. In a MDOF grating interferometry system, there are four diffraction beams from a two-axis scale grating and reference grating, respectively. For further modulation, these beams will propagate more than 100 mm, which makes paralleling these beams necessary. In previous research, collimation lens, separate prisms and a home fabricated diffraction device by combining four separate one-axis line gratings in a glass substrate have been demonstrated. However, large power loss and assembly complicity makes these techniques less competitive. For solving this problem, this research proposed a new lens module, which is an improved type prism, quadrangular frustum pyramid (QFP) prism. The prism is designed in such a way that these four reflected beams from the grating are symmetrically incident into the prism through the upper surface, total reflected on the inner sides of the prism, and then parallel getting through the bottom surface. A prism that allows an incident beam diameter of 1 mm and four paralleling beams with a 10 mm distance between the two diffraction beams along one direction was designed, fabricated and tested. Testing results based on an entire grating interferometry system verified that the proposal in this research is greatly effective in beam paralleling in terms of less power loss and high paralleling and greatly reduces the assembly complicity, which will eventually be beneficial for grating interferometry application.

**Keywords:** surface encoder; multi-degree-of-freedom; interferometry; grating; prism

## 1. Introduction

With the development of industrial automation and intelligence, the popularization and demand of multi-axis computer numerical control (CNC) equipment in industrial production are increasing. Therefore, the multi-degree-of-freedom (MDOF) measurement technology is a key basic technology. For instance, in a MDOF CNC machine tool, the current common measurement scheme is to equip each motion axis with a separate single-axis measuring device, including linear grating ruler, linear photoelectric encoder, and so forth [1–3]. The installation of these measuring devices is usually stacked along the motion axis. This brings about the abbe error problem in the assembly process, which will have an adverse effect on the measurement accuracy [4,5].

In order to eliminate the abbe error in multi-axis measurement and improve the integration of the measurement system, measurement devices with MDOF measurement capabilities have been paid more attention. Among them, the technique schemes of two-dimensional grating rulers and the MDOF grating rulers are of great application potential [6–9]. These MDOF grating rulers, generally with usage of a two-axis planar scale grating, are able to measure not only the translational motions but the small order rotational error motions, and it can be also applied to six-DOF measurement. Compared with the traditional single-axis measuring devices, the MDOF grating ruler has a higher level integration and the abbe error could be avoided. Thus, the reliability and precision of the multi-axis precision stage and measurement system can be improved effectively [10–12].

As illustrated in these prototype MDOF grating rulers, grating diffraction beams interferometry is the key technology, which means that the interference is generated by using four beams of  $\pm 1$  order diffracted light of a two-dimensional grating so that to obtain the displacement distance by detecting the interference signal [6–8]. Compared with the one-dimensional interferometric grating ruler, the two-dimensional grating ruler is equivalent to integrating two sets of one-dimensional grating rulers whose working direction is orthogonal to each other, however this makes complexity of the two-dimensional grating ruler light path increase greatly. The complexity is mainly manifested in the need of the interference of four beams both from the reference grating and the scale grating parallel or perpendicular to the incident light in the system in order to produce an ideal interference signal. This means that the above eight diffracted beams need to be finely deflected, which determines the quality of the interference signals. In previous research, there are three proposed methods, collimation lens represented by Reference [6], separate prisms represented by Reference [5], and a home fabricated diffraction device by combining four separate one-axis line gratings in a glass substrate [13]. These methods meet the demand of the light path design and can be manufactured under the lab conditions. However, due to the complex structure and alignment, loss of signal intensity and the low sensitivity, these solutions are difficult to meet the condition of mass production.

Therefore, in this paper, focusing on the demand of mass production and engineering prototype, facing these challenges in current technologies, we design a new lens module for this kind of grating rulers, a single prism called a quadrangular frustum pyramid, enabling simultaneously deflecting the four diffraction beams without any power loss and easy alignment. To test the lens module, a grating ruler system is designed, constructed and experimental verification are introduced in this manuscript.

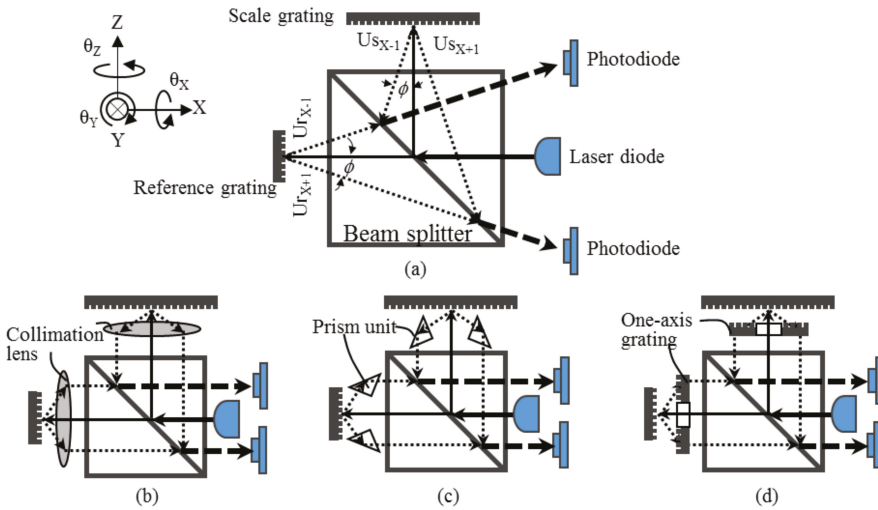
## 2. Design of the Integrated Optical Prism

Figure 1a is the schematic diagram of the principle of two-dimensional grating ruler [5]. When a laser beam is irradiated vertically on a two-dimensional grating, the  $\pm 1$  order diffractive light will be generated along the two grating lines of the two-dimensional grating. The total diffraction light is four beams, denoted as  $U_{X+1}$ ,  $U_{X-1}$ ,  $U_{Y+1}$ ,  $U_{Y-1}$  ( $U_{Y+1}$ ,  $U_{Y-1}$  are not shown in this Figure for clarity). The angle between the diffraction light and the normal line of the grating plane is equal to the diffraction angle of the grating. As shown in Figure 1a, the beam emitted by the laser source is vertically irradiated on the reference grating and the scale grating after the beam is divided by the beam splitter. The interference of  $\pm 1$  order diffractive beams from the scale grating and reference grating occurs after the beams combining in the splitter. The interference fringes are formed and can be detected by photoelectric detectors for its light intensity, and finally four interference signals are obtained. After that, the measured displacement can be solved according to the detected interference signal through the corresponding data processing algorithm.

As illustrated in previous research [5,6,13], further phase delay and amplitude division of these signals are required so that the direct current components of the interference signals can be eliminated and the motion direction can be distinguished. Thus, more optics are added into this optical layout and this results the light path increasing to more than 100 mm [14]. This makes paralleling these diffraction beams necessary. As mentioned in the first section of this paper, Figure 1b–d illustrate these three present optical designs. As shown in Figure 1b, a collimation lens was employed. The diffraction

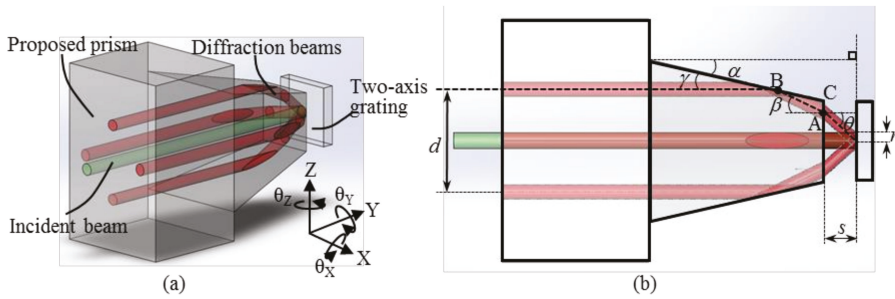


beams were refracted. This is suitable for large grating period, generally a 10 micrometer order. Under such a condition, the diffraction angle from gratings are relatively small, and the collimation effect can be ensured [6]. However, a large grating period will result in a relatively poor resolution. A high resolution requires a small grating period. Correspondingly a large diffraction angle makes using the collimation difficult. The second trial is that of the triangular prism refraction method, as shown in Figure 1c. Similarly, a triangular prism can be replaced with a mirror. Both methods need the diffractive light deflection mirror to be precisely calibrated so that the reference and the scale grating diffractive light will be able to interfere after transmission to the detector. It was found in the trial production of the actual test prototype [5,12], the process of assembly and adjusting the diffractive light deflection prism is extremely complicated and difficult for the mechanical assembly stress and the influence of such factors as poor precision stability of the optical path. In order to avoid the problem that diffractive light deflection prisms need to be adjusted separately, a structure using transmission grating for diffraction light deflection is proposed, as shown in Figure 1d [13,15,16]. In this structure, the transmission grating is actually four one-dimensional transmission gratings combined. This scheme greatly reduces the difficulty of assembling and debugging optical circuits by employing the integrating transmission grating, which is used as a diffractive light deflection device. However, limiting the transmission grating diffraction efficiency, the plan will greatly reduce the intensity of the diffractive light. And it is not conducive to improve the signal-to-noise ratio (SNR) in subsequent signal processing, which leads to reduce the measuring accuracy of the grating ruler system.



**Figure 1.** (a) Principle of a two-grating encoder; (b) diagram of employing a collimation lens to parallel diffracted beams; (c) diagram of employing prisms to parallel the diffracted beams; (d) diagram of employing gratings to parallel the diffracted beams.

In order to overcome the above problems, an integrated grating diffractive light deflection prism design is proposed. The design takes the advantages of prism deflection that does not reduce the beam intensity and transmission grating high integration, providing a high precision, high brightness and easy to assemble and debug solution for reference grating and scale grating diffractive light deflection. The effect of the integrated prism and the deflection of four  $\pm 1$  order diffractive beams is shown in Figure 2. The key parameters in the design are the integration of the prism angle, performance parameters for the diffraction light horizontal/vertical emergent spacing, and variable parameter to the top of the prism and grating spacing and the beam radius.



**Figure 2.** (a) The 3D structure diagram of the integrated quadrangular frustum pyramid prism and the optical route; (b) the optical configuration of the deflection diffractive beams in the design.

The design and calculation of parameters are described in Figure 2b. The grating diffraction angle is denoted as  $\theta$ . The refractive angle of diffraction light is denoted as  $\beta$ . The reflection angle of diffraction light is denoted as  $\gamma$ . The angle between the normal side of the prism and the top plane is denoted as  $\alpha$ . The design requires that the diffraction light emitted is parallel to the incident light, then  $\alpha$  and  $\gamma$  should satisfy Equation (1).

$$\alpha = \gamma. \tag{1}$$

In triangle ABC, there is a geometric relationship expressed by:

$$\gamma + (90^\circ - \beta) + (90^\circ + \alpha) = 180^\circ. \tag{2}$$

Diffraction angle  $\theta$  and refractive angle  $\beta$  of diffraction light meet the refractive law of Equation (3), where  $n$  is the refractive index of optical materials used for prism:

$$\frac{\sin \theta}{\sin \beta} = n. \tag{3}$$

Grating diffraction angle  $\theta$  is determined by grating distance  $d$  and incident light wavelength  $\lambda$ :

$$\theta = \arcsin \frac{\lambda}{d}. \tag{4}$$

$$\alpha = \frac{\beta}{2}. \tag{5}$$

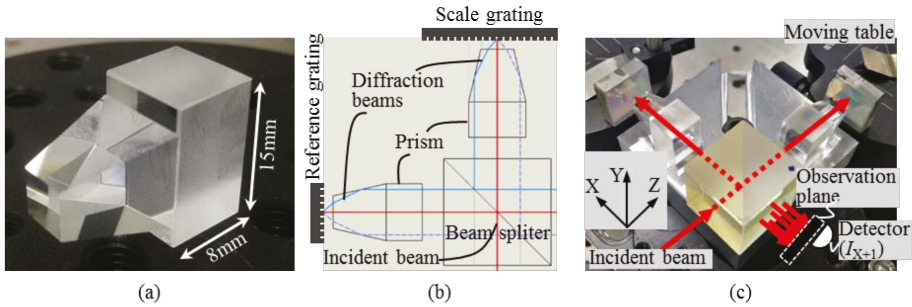
$$\alpha = \frac{\beta}{2} = \frac{1}{2} \arcsin \frac{\sin \theta}{n} = \frac{1}{2} \arcsin \frac{\lambda}{nd}. \tag{6}$$

In our design, the optical material of a prism is H-K9Lglass (Union Optic Incorporated, Wuhan, China), and the wavelength  $\lambda = 660$  nm, refractive index  $n = 1.5128$ , grating pitch  $d = 1000$  nm. Then, the key parameter of the integrated prism  $\alpha$  can be calculated:

$$\alpha = \frac{1}{2} \arcsin \frac{\lambda}{nd} = \frac{1}{2} \arcsin \frac{660}{1.5138 \times 1000} = 12.924^\circ. \tag{7}$$

### 3. Results

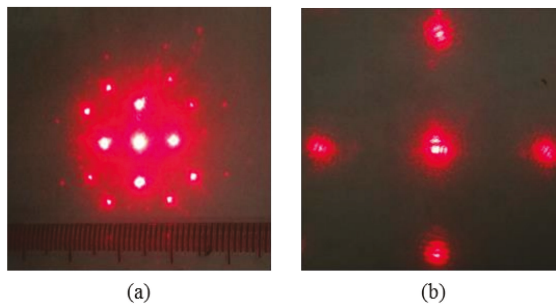
Based on the above geometric optical design, a batch of proposed QFP prisms were fabricated by a lens manufacturing company (Union Optic Incorporated, Wuhan, China) and the parameters were measured and given as shown in Figure 3a. It should be noted that a cuboid base was added onto the bottom of the QFP prism so that the prism can be mounted onto the plate base of the grating ruler, which also benefits the prism manufacture process.



**Figure 3.** (a) The picture of the fabricated deflection prism; (b) the light path of the deflection of the diffractive beams; (c) the grating interferometry testing system with the fabricated deflection prism.

In order to test the effectiveness of the QFP prism, a fundamental grating ruler optical path using interferometric measuring principle is constructed, as shown in Figure 3b,c. This optical layout is the same as the optical path principle shown in Figure 1c. The incident beam with a calibrated wavelength of 660 nm was divided into two beams and are projected to the reference grating and the scale grating after passing through the QFP prisms, respectively. The scale grating is mounted onto a moving table and the reference grating is stable.

First of all, performance of the QFP prism in diffraction beams propagation directions modulation was evaluated by observing the beam spots on the screen. The observation as shown in Figure 3c was placed at a distance about 150 mm, a similar distance with that of a real grating ruler. Figure 4a illustrated the beam spots on the screen. It can be seen that the four diffraction beams from the reference grating coincided well with those from the scale grating. The distance between these four diffraction beams from the central beam (reflective beam, also called zero-order diffraction beam) are consistent with the designed values. These results verified the geometry accuracy of the fabrication process. It should be noted that there are many undemand spots on the screen which are caused by the non-orthogonal diffraction beams from the grating and can be easily blocked by setting a physical holes aperture.



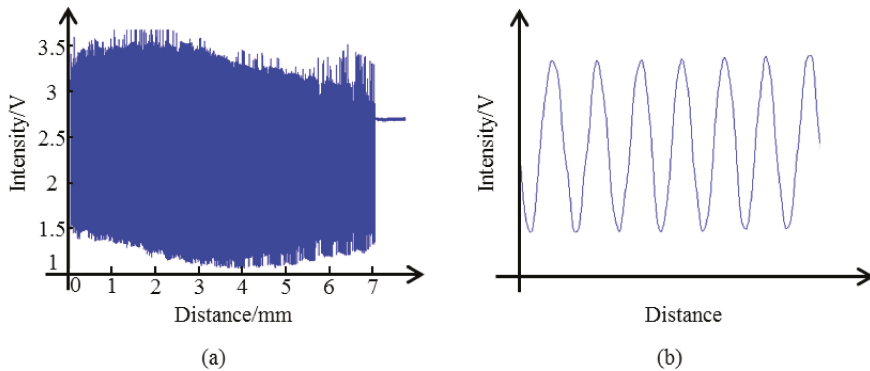
**Figure 4.** (a) The diffraction beams spots on the observation screen; (b) the interference fringe on the observation screen.

In order to further verify the effectiveness of the proposal in terms of optical property, such as polarization, the interference signal was tested. As shown in Figure 4b, clear diffraction stripe structure can be observed in all four interference spots and these represent that they have almost same order of interference intensity, while in previous separated prism layouts this was always the most challenging task.

The results can preliminarily prove that the QFP prisms based interferometric grating ruler optical configuration is feasible and effective. To further test the optical path, the scale grating was

driven by the micron motion platform (PI-M112.1 DG, Physik Instrumente (PI) GmbH & Co. KG, Karlsruhe, Germany), as shown in Figure 3c.

As the principle of grating interferometry, the movement of the scale grating brings a periodic change of the interference signal. In this demonstration, two interference signals,  $I_{X+1}$  and  $I_{X-1}$ , were recorded by a photodiode. The interference signals are given in Figure 5. It can be seen that the visibility of the interference signal almost remains stable, which is essential for the grating ruler application. The direct current components in the interference signal can be eliminated optically. The results show that the given QFP prisms based optical layout can form a good interference signal, the signal is in good sinusoidal shape under constant speed drive and can be used for bit shifting calculation.



**Figure 5.** (a) The interference signal of collected data; (b) the detail interference signal of one channel ( $I_{X+1}$ ) during moving.

The motion velocity along the direction X and direction Y of the measured grating was 0.025 mm/s, the motion stroke was 7 mm, and the sampling rate was 1000 Hz. A total of four repeated experiments were conducted. The calculation results of bit shifting of each test are shown in Table 1. As can be seen from the data in the table, the repeatability of the bit transfer calculation is good, and the relative error is no more than 2.5%. The measuring error can be reduced by calibrating the measure system for the machining error in fabricating the grating and the QFP prism and assembly error. In each group, the result of  $X_{+1}$  is slight less than  $X_{-1}$ , because the rear surface of the QFP prism is not fully parallel to the grating surface, which is inevitable and will have no or little influence on the result after calibrating the system and further processing the subdivision signals.

**Table 1.** The result of the displacement solution ( $\mu\text{m}$ ).

Testing Group	Result of $X_{+1}$	Result of $X_{-1}$
1	6829	6889
2	6843	6890
3	6877	6891
4	6874	6890

#### 4. Discussion

First of all, the geometrical parameters of the fabricated QFP prisms were evaluated before they were constructed into the optical testing system. The parameters variation is less than 0.1 mm, which proves this structure design and fabrication technology effective. About cost, the price of one QFP prim is about 60 USD when the customized number is no large than 10 in a Chinese optics factory

and this can greatly reduce to be 15 USD when this product number is larger than 1000. The reduction of cost of this QFP prism will be beneficial for the grating ruler application.

Second, because the scope of this research is focused on the design, fabrication, construction and evaluation of this proposed QFP prism, the systematical algorithm for motion determination that was illustrated in References [6–16] was not introduced or not employed to precisely calculate the displacement. Furthermore, because of the slight misalignment of the motion stage and the scale grating, crosstalk will inevitably be involved. The systematical misalignment could be previously calibrated with a compensation factor used [12].

Lastly, it should be also noted that the proposal in this research requires the grating pitch and the light source wavelength to be time-stable or a small order variation within the designed tolerance. Improvement of the tolerance generally can be achieved by enlarging the beam spot, yet will correspondingly enlarge the size of the QFP prism.

## 5. Conclusions

This research proposed, fabricated and tested a new optics, the QFP prism for simultaneously paralleling multiple diffraction beams for grating interferometry application. To get a compact and effective QFP prism we calculated the angle  $\alpha$ , the key parameter of the prism, according to the wavelength, the reflective index and grating patch, and designed a QFP prism with  $\alpha = 12.924^\circ$ . The fabricated QFP prism was tested in a grating ruler system to evaluate the prism performance in terms of shape parameters and optical functionality. The testing results preliminary proved that the design and fabrication process are feasible. The less complexity in optical configuration and assembly procedure with usage of the QFP prism will be beneficial for mass production of this kind of grating rulers.

## 6. Patents

A prism for diffraction beams propagation directions modulation, Xinghui Li, Kai Ni, Xiang Xiao, Qian Zhou, Huanhuan Wang, Lijiang Zeng, Weihai Yuan, Xiao Su and Xiaohao Wang, Application No. CN201611217716.2, Open No. CN106646907A, May 10, 2017.

**Author Contributions:** Conceptualization, X.L. and X.X., methodology, X.X., X.L.; validation, Y.S., X.X., and X.L.; resources, Q.Z., K.N., X.L.; data curation, X.X., Y.S., H.L.; writing—original draft preparation, X.L., X.X.; writing—review and editing, X.L., Y.S.; visualization, X.L., X.X.; supervision, Q.Z.; project administration, Q.Z.; funding acquisition, K.N., Q.Z.

**Funding:** This research is supported by the National Natural Science Foundation of China under Project No. 51427805, the Shenzhen fundamental research funding with Grant No. JCYJ20170817160808432 and Grant No. JCYJ20160531195459678, the Youth funding of Shenzhen Graduate of Tsinghua University with Grant No. QN20180003, and the National key research and development program under Grant No. 2016YFF0100704.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Tomita, Y.; Kojima, E.; Kawachi, S.; Koyanagawa, Y.; Otsuku, S. Development and applications of Sumitomo precision stage technologies for FPD process. *Int. J. Jpn. Soc. Precis. Eng.* **2012**, *72*, 117–121. [[CrossRef](#)]
2. *Surface-Motor Driven Stage Series*; Mechatronics Division, Sumitomo Heavy Industries Ltd.: Tokyo, Japan, 2018.
3. Gao, W.; Kim, S.W.; Bosse, H.; Haitjema, H.; Chena, Y.L.; Lu, X.D.; Knapp, W.; Weckenmann, A.; Estler, W.T.; Kunzmann, H. Measurement technologies for precision positioning. *CIRP Ann.* **2015**, *64*, 773–796. [[CrossRef](#)]
4. Kimura, A.; Hosono, K.; WooJae, K.; Shimizu, Y.; Gao, W.; Zeng, L.J. A two-degree-of-freedom linear encoder with a mosaic scale grating. *Int. J. Nanomanuf.* **2011**, *7*, 73–91. [[CrossRef](#)]
5. Gao, W.; Kimura, A. A three-axis displacement sensor with nanometric resolution. *CIRP Ann.* **2007**, *56*, 529–532. [[CrossRef](#)]
6. Bryan, J. The Abbe principle revisited: An updated interpretation. *Precis. Eng.* **1979**, *1*, 129–132. [[CrossRef](#)]
7. Schwenke, H.; Knapp, W.; Haitjema, H.; Weckenmann, A.; Schmitt, R.; Delbressine, F. Geometric error measurement and compensation of machines—An update. *CIRP Ann.* **2008**, *57*, 660–675. [[CrossRef](#)]

8. Caley, A.J.; Waddie, A.J.; Taghizadeh, M.R. Two-dimensional displacement sensing using a cross diffraction grating scheme. *J. Opt. A Pure Appl. Opt.* **2003**, *6*, 106.
9. Kimura, A.; Gao, W.; Arai, Y.; Zeng, L.J. Design and construction of a two-degree-of-freedom linear encoder for nanometric measurement of stage position and straightness. *Precis. Eng.* **2010**, *34*, 145–155. [[CrossRef](#)]
10. Li, X.H.; Gao, W.; Muto, H.S.; Shimizu, Y.; Ito, S.; Dian, S. A six-degree-of-freedom surface encoder for precision positioning of a planar motion stage. *Precis. Eng.* **2013**, *37*, 771–781. [[CrossRef](#)]
11. Wang, W.N.; Sun, J.H. A modified moiré Interferometer for three-dimensional displacement measurement. *Chin. Opt. Lett.* **2004**, *2*, 396–398.
12. Kimura, A.; Gao, W.; Kim, W.; Hosono, K.; Shimizu, Y.; Shi, L.; Zeng, L.J. A sub-nanometric three-axis surface encoder with short-period planar gratings for stage motion measurement. *Precis. Eng.* **2012**, *36*, 576–585. [[CrossRef](#)]
13. Li, X.H.; Shimizu, Y.; Ito, T.; Cai, Y.D.; Ito, S.; Gao, W. Measurement of six-degree-of-freedom planar motions by using a multi-probe surface encoder. *Opt. Eng.* **2014**, *53*, 122405. [[CrossRef](#)]
14. Gao, W.; Araki, T.; Kiyono, S.; Okazaki, Y.; Yamanaka, M. Precision nano-fabrication and evaluation of a large area sinusoidal grid surface for a surface encoder. *Precis. Eng.* **2003**, *27*, 289–298. [[CrossRef](#)]
15. Gao, W.; Saito, Y.; Muto, H.; Arai, Y.; Shimizu, Y. A three-axis autocollimator for detection of angular error motions of a precision stage. *CIRP Ann.* **2011**, *60*, 515–518. [[CrossRef](#)]
16. Li, X.H.; Wang, H.H.; Ni, K.; Zhou, Q.; Mao, X.Y.; Zeng, L.J.; Wang, X.H.; Xiao, X. Two-probe optical encoder for absolute positioning of precision stages by using an improved scale grating. *Opt. Express* **2016**, *24*, 21378–21391. [[CrossRef](#)] [[PubMed](#)]



© 2018 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).

Article

# Obtaining Vital Distances Using Wearable Inertial Measurement Unit for Real-Time, Biomechanical Feedback Training in Hammer-Throw

Ye Wang <sup>1</sup>, Hua Li <sup>1</sup>, Bingjun Wan <sup>2</sup>, Xiang Zhang <sup>3</sup> and Gongbing Shan <sup>2,3,4,\*</sup>

- <sup>1</sup> Department of Mathematics & Computer Science, University of Lethbridge, Lethbridge, AB T1K 3M4, Canada; ye.wang3@uleth.ca (Y.W.); hua.li@uleth.ca (H.L.)
  - <sup>2</sup> School of Physical Education, Shaanxi Normal University, Xian 710119, China; bingjunw55@snnu.edu.cn
  - <sup>3</sup> Department of Physical Education, Xinzhou Teachers' University, Shanxi 034000, China; xiangzhang@xztc.edu.cn
  - <sup>4</sup> Biomechanics Lab, Faculty of Arts & Science, University of Lethbridge, Lethbridge, AB T1K 3M4, Canada
- \* Correspondence: g.shan@uleth.ca; Tel.: 1-403-329-2683

Received: 29 October 2018; Accepted: 30 November 2018; Published: 3 December 2018



**Abstract:** The hammer throw is one of the regular track and field competitions, but unlike other events, it has not seen a new world record for over three decades. The standstill may be caused by the lack of scientifically based training. In our previous work, we have developed a wireless/wearable device for the wire tension measurement in order to develop real-time biomechanical feedback training. In this paper, we show the improvement of our wearable system by adding two sensors for tracking of two vital vertical distances. The paper describes the details related to the development of turning an inertial measurement unit into a tracking device for the dynamic distances. Our preliminary data has shown that the dynamic data of the hip and wrist could be used for revealing the coordination between the upper and the lower limbs during a throw. In conjunction with wearable wire-tension measurement, various motor control patterns employed for hammer throwing could be demystified. Such real-time information could be valuable for hammer-throw learning and optimization. Further studies are required to verify the potentials of the wearable system for its efficiency and effectiveness in coaching practice.

**Keywords:** IMU; dynamic tracking; limbs' coordination; motor control pattern; motor learning

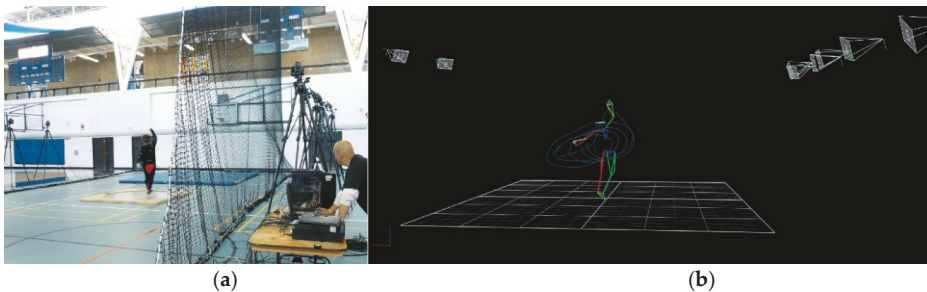
## 1. Introduction

Optimization of any sport skill requires re-organization of limbs coordination responsible for governing the movement performance [1]. This type of motor learning can be enhanced through a number of methods that are utilized in research and application settings alike. In general, verbal feedback of coaches in real-time is commonly used as a preliminary means of instilling motor learning [1,2]. Due to the rapidity and complexity of some sport skills as well as invisibility of some parameters (e.g., force), the real-time feedback of coaches is often a subjective guess based on experience. For increasing the reliability of feedback in training, biomechanical means are used to supplement the verbal instructions [3–6]. The hammer throw is such a sport skill that needs a combination of a coach's experience and biomechanical feedback in elite sport training to facilitate motor learning and optimize outcomes.

Men's hammer throw has been part of Olympics track and field competitions since 1900, but unlike other events, the hammer throw has not seen a new world record since 1986 [7]. This standstill may be caused by the lack of scientifically based training. While extensive 3D motion analysis technologies do supply highly trustworthy information for human motor skill quantification [8–11], due to their

drawbacks, the analysis and feedback has traditionally occurred offline after completion of a given testing session, i.e., it is post-measurement feedback, rather than real-time [11–14]. The drawbacks of a 3D motion capture system include complicated operation, high cost, long calibration and setup procedures, time-consuming course on data collection, processing, analysis as well as the movement constraints induced by dozens of capture markers attached on a subject's body [15]. These drawbacks have hindered the use of 3D motion capture systems in sport training practice. As a consequence, research has been initiated to develop real-time biomechanical feedback devices for hammer throw training, beginning with wire-tension measurement [12].

Most recently, a pilot study [16] using 3D motion capture technology (Figure 1) found that the timely displacements of hip and wrist may be used to reveal the upper and lower limbs' coordination when analyzing hammer throw. The pilot study has shown that the timely change of vertical displacements of hip and wrist are closely related to the turning speed, the ratio of one-leg/two-leg support (power generation), and hammer velocity change during the skill performance. Therefore, obtaining the dynamic distance data of these two anatomical landmarks would be vital for real-time feedback training.



**Figure 1.** The 3D motion capture of hammer throw. (a) The set-up of the data collection; (b) a sample of the 3D data.

The results of the two studies would suggest that a combination of the wire-tension measurement and the dynamic vertical-displacements of hip and wrist could have great potential for substituting 3D motion capture technology in the skill analysis of the hammer throw. Since real-time wire-tension is developed (i.e., already wearable) [12], developing wearables for tracking hip and wrist movements would realize the real-time biomechanical feedback learning/training in the hammer throw. Encouraged by the results of the studies, we are aiming at developing a practical wearable device for pursuing real-time training. This paper will highlight our approach of using the inertial measurement unit (IMUs) as a practical approach for the development of a wearable system for biomechanical feedback training in the hammer throw.

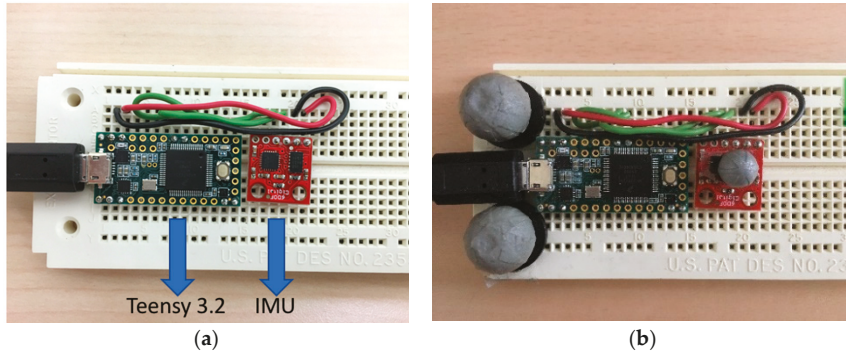
## 2. Materials and Methods

### 2.1. Hardware Configuration

The constitution of the hardware in our system is straightforward. Intuitively shown in Figure 2a, a six degree of freedom (6DoF) IMU [17] and a Teensy 3.2 board [18] were connected with each other. We built these on a breadboard as a testing device with an Arduino Mega board in our previous work for the wire-tension measurement [12]. The 6DoF means there is a tri-axial accelerometer and a tri-axial gyroscope, which can return the acceleration and the angular speed, respectively, on the X, Y, and Z axis of a coordinate system. In other words, 6DoF can be described as the freedom of movement of a rigid body in three-dimensional space, which refers to the following: Forward/back (on X axis), left/right (on Y axis), up/down (on Z axis), roll (around X axis), pitch (around Y axis), and yaw (around Z axis). In our particular case, we do not need the magnetometer, which can be potentially combined with the



accelerometer and the gyroscope to construct a 9DoF IMU. In our application, the IMU is designed as a combo board, which has the accelerometer, ADXL345, and the gyroscope, TG3200. The Teensy 3.2 board is a breadboard-friendly microcontroller, which can be programmed in Arduino IDE (Integrated Development Environment). Compared to several Arduino boards, the current one is smaller than the Arduino UNO and Mega boards, and it has its own USB (Universal Serial Bus) port while Arduino Mini does not have one, which makes it relatively easier to be programmed.



**Figure 2.** (a) Teensy 3.2 and Inertial Measurement Unit (IMU) connected on a breadboard; (b) Teensy 3.2 and IMU connected on a breadboard with motion capture markers.

In addition, as shown in Figure 2b, we attached three motion capture markers (two are 9 mm in diameter and one is 5 mm in diameter) on the sensor device for constructing a capture model for 3D motion capture using a 10-camera VICON MX40 motion capture system (VICON Motion Systems, Oxford Metrics Ltd., Oxford, England). The motion capture rate was set at 200 frames/s. Calibration residuals were determined in accordance with VICON's guidelines and yielded positional data accurate within 1 mm. The VICON data was used to help us develop a tracking algorithm of the IMUs in the vertical direction.

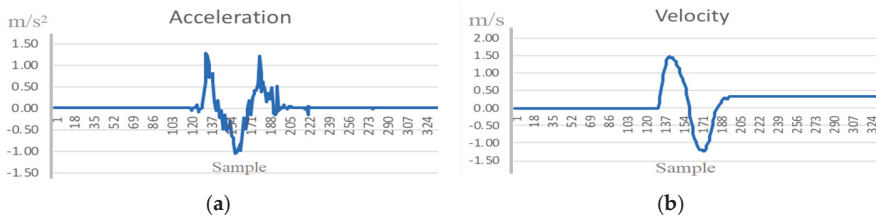
## 2.2. Methodology

First, we needed to configure the IMU in an Arduino sketch program by applying the Wire library [19], which allowed communication with I2C (Inter-Integrated circuit) devices. This configuration program was uploaded to the Teensy 3.2 microcontroller. It was used to set up the accelerometer and the gyroscope. We set the data format register of the accelerometer to  $0 \times 09$ , which could set the acceleration range from  $-4 \text{ g}$  ( $1 \text{ g} = 9.8 \text{ m/s}^2$ ) to  $+4 \text{ g}$ . According to the datasheet of ADXL345 [20], it sets the device to a full resolution mode, where the output resolution increases with the g range set by the range bits to maintain a  $3.9 \text{ mg/LSB}$  scale factor. This setting information is useful for converting the unit of the raw output data to g. We set the power control register to  $0 \times 08$  to make the accelerometer in a measurement mode. Should we want to have the minimum power consumption, we could set this register to  $0 \times 00$  to make it in a standby mode. Similarly, we configured the ITG3200 gyroscope's settings in this program according to its datasheet [21]. The range of rotation speed was set from  $-2000 \text{ degree/second (dps)}$  to  $+2000 \text{ degree/second}$ , which was a full-scale range. The sensitivity was  $14.375 \text{ LSB}$ , which was also useful for converting the unit of the raw output data to dps. The low pass filter bandwidth was set to  $98 \text{ Hz}$ , and its sampling rate was set to  $100 \text{ Hz}$ . In addition, we implemented the functions for reading and outputting the acceleration values relative to the earth and the angular speed values in the Arduino sketch program.

The next step was to do some calibration work. We applied the sensitivity values,  $256 \text{ LSB}$  and  $14.375 \text{ LSB}$ , to convert the units of the raw output of accelerometer and the raw output of gyroscope to

g and dps, respectively. These values can be found in the datasheets of the sensors. We implemented a function of calculating the offsets (chip specific) of the first 200 samples in the Arduino sketch program.

Finally, we needed to apply an algorithm to predict the orientation. Kalman-based filters have been widely used in orientation estimation [22]. In the beginning, we tried to implement a complementary Kalman-based algorithm. However, the result of the orientation estimation was bad due to a drifting error that kept occurring in calculating the velocities. Figure 3 displays the acceleration data obtained from the IMU sensor and the corresponding velocity data, which was calculated by the complementary Kalman-based filter. Obviously, the velocity data cannot come back to zero at the end of the test, which is known as a drifting error. Then, we started looking for some other ways to get the accurate orientation. Madgwick’s algorithm [23] is also known as Madgwick’s MARG (magnetic, angular rate, and gravity) filter or AHRS (attitude and heading reference systems) algorithm. As Madgwick mentions in his work, Kalman-based filters are difficult to implement because they may require sampling rates far exceeding the subject bandwidth. Our sensor device has a fairly low sampling rate, which is only 50 Hz. This is probably one of the reasons that we experienced the velocity drifting error. As Madgwick claims in his study, his algorithm can be effective even at lower sampling rates, like 10 Hz. In addition, Madgwick compares the performance of his algorithm with a Kalman-based filter in his study, and the results indicate his algorithm has a slightly better accuracy. Therefore, we decided to apply Madgwick’s algorithm. We tried two versions of his filters to predict the orientation in MATLAB R2017a.



**Figure 3.** (a) The acceleration data obtained from the IMU sensor; (b) the corresponding velocity data, which was calculated by the complementary Kalman-based filter.

### 2.3. Our Prototype and In-Field Test

After the development of the wearables for vertical distance monitoring, we built our prototype for the real-time biomechanical feedback training of the hammer throw. The prototype integrated the two newly developed distance sensors into our Arduino Mega board developed in our previous work for wire-tension monitoring during the hammer throw [12], increasing the device capacity to monitor three key variables in real time. The two new distance sensors were buried into a self-made belt and a self-made armband, attached to the waist and right wrist, respectively.

The new device was tested in the field. A varsity-level athlete (male, 25 years, 81 kg, 1.75 m with seven years training experience) tried out the real-time feedback device. Our wearable device permitted considerable freedom of movement for the subject with negligible influence on his performance. Taking advantage of this, we placed no restrictions on the subject’s movements during the in-field test to preserve his normal “control style”. Four trials were done.

## 3. Results and Discussion

The result of the test by applying Madgwick’s filter is satisfied. During a test, we moved our system device up and down three times. As shown in Figure 4, we got relatively accurate feedback of the 3D positioning data. Indeed, the Madgwick algorithm eliminates the drifting error from integrating the velocities. The three different curves stand for the changing distances over time on the X axis, Y axis, and Z axis in a 3D space. The dynamic distance on the Z axis (blue lines) shows exactly three times up and down of the device. The range of vertical movements is ~0.33 m for the first moving-up

and down, ~29 cm for the second circle, and ~32 cm for the last one, respectively. The next step is to validate the accuracy of the device.

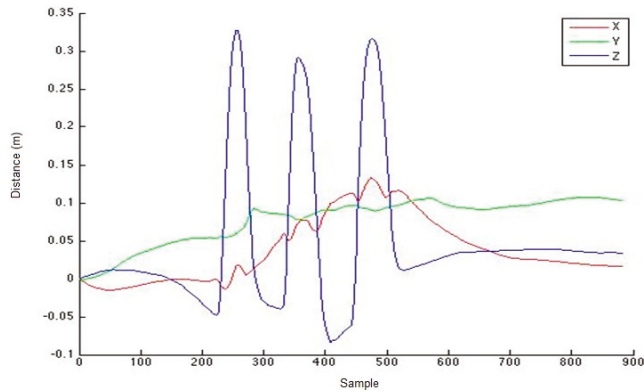


Figure 4. The 3D positioning data obtained by the IMU with the Madgwick filter.

It is well known that 3D motion capture technology provides accurate and objective analysis of a variety of human motor skills [14,24–27]. Therefore, we employed the synchronized data collection of the IMU and 3D motion capture (Figure 1b) for validating and improving the accuracy of the IMU device. There were eight synchronized tests performed to obtain thousands of data for our validation. Since we aimed to gain the dynamic vertical distance, the validation of the Z axis was selected. Figure 5 shows a typical test data. The synchronized data demonstrate a matching vertical excursion over time between the IMU data and the accurate 3D motion capture data. The results suggest that our device works principally.

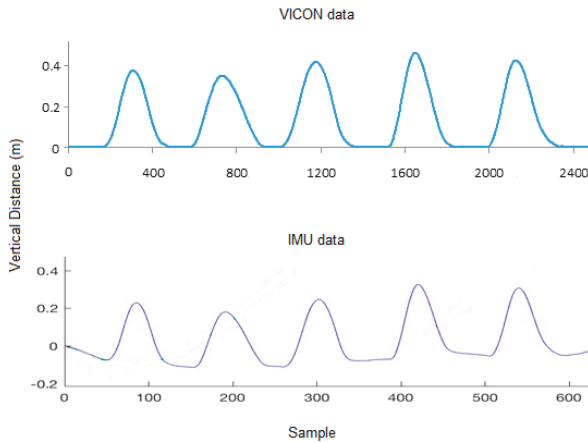
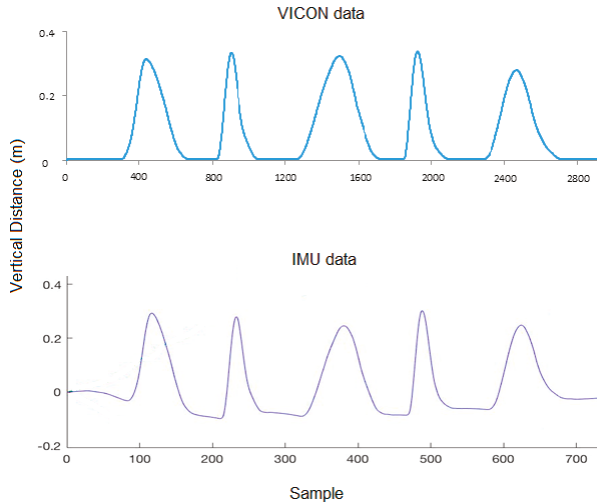


Figure 5. A synchronized test data obtained from 3D motion capture (VICON data, top, sampling rate 200 Hz) and our IMU device without calibration (IMU data, bottom, sampling rate 50 Hz).

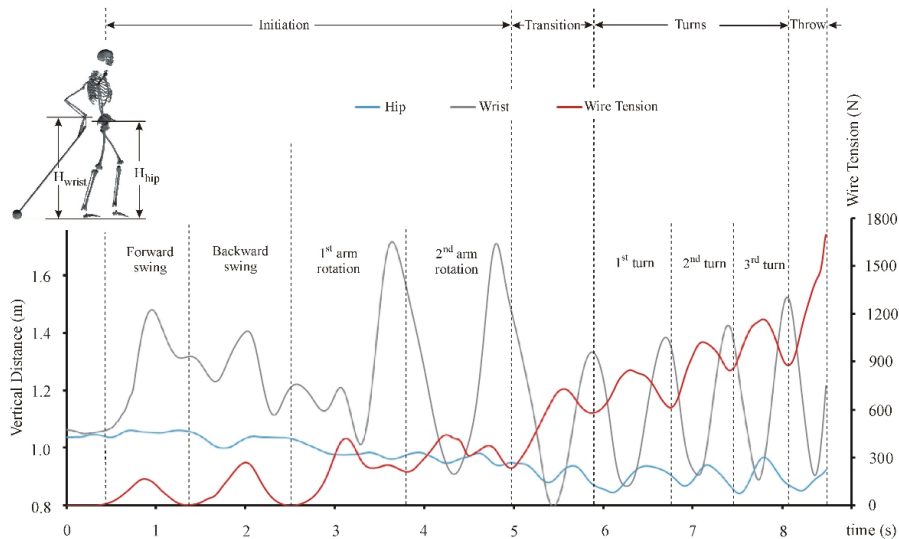
A magnitude comparison shows that the excursion of the VICON data was larger than that of the IMU data (Figure 5). A timely comparison between the synchronized data of all trials revealed that the two excursions ran in a quasi-parallel way, which suggested that we could apply a factor for re-calibrating the IMU device to improve the accuracy of the IMU data. After the quantitative comparison between the two excursions of all trials, a re-calibration factor of 1.31 was determined.

After the simple re-calibration, a renewed synchronized measurement was done and the result is shown in Figure 6. This time, the average data error of our IMU data decreases to under 6%, which is accurate enough for sport skills analysis using the biomechanical modeling method [28–32].



**Figure 6.** A renewed synchronized test data obtained from 3D motion capture (VICON data, top, sampling rate 200 Hz) and our IMU device after calibration (IMU data, bottom, sampling rate 50 Hz).

Finally, it should be noted that our device needs an initial value for its application. As shown in Figure 6, the device will start at zero regardless of its actual vertical position. Therefore, for its application in the hammer throw, an accurate feedback needs the initial heights of the hip and wrist ( $H_{hip}$  and  $H_{wrist}$ ) as shown in Figure 7.



**Figure 7.** The upper and lower limb coordination (i.e., motor control pattern) revealed by the vertical distances of hip and wrist as well as the wire tension during a hammer throw by a college-level athlete.

The in-field test on the college-level athlete using our prototype device confirms the potential of using wire-tension and IMUs in real-time feedback training (Figure 7). In practice, the motor control of the hammer throw could be divided into four phases: Initiation, transition, turns, and throw. The goal of the initiation phase is to launch the circulation of the hammer around the body. It commonly consists of a forward and backward swing of the hammer (i.e., to set the hammer to motion) and two over-head arm rotations (i.e., to set the hammer into rotation). The transition phase aims to switch the body from standing posture to the first body rotation, building a rotating system of the body and the hammer. The phase of turns accelerates the rotating system of the body and the hammer to their highest circulation. The final phase is the throwing. Our data has revealed the following motor control information: (1) During the transition phase, the upper and lower limbs' controls are transferring from an unclear coordination pattern to a quasi-out-of-phase coordination in the phase of turns (Figure 7). (2) The transition phase helps the power generation (i.e., wire tension) become in phase (quasi) with the hips' up-and-down movement, indicating the hammer acceleration depends on the timely flexion/extension of lower limbs. (3) Finally, the characteristic of quasi-out-of-phase between the arm control and wire tension finishes in the transition phase. Would such characteristics appear at different levels of athletes? How can the real-time feedback (i.e., wearable devices) be helpful in optimization of individual hammer-throw skills? Are there additional potentials of wearables in the learning and training of the hammer throw? Future studies are needed to answer the above application questions.

#### 4. Conclusions

In conclusion, we used IMUs to build a wearable sensing system to determine the dynamic vertical distances of the hip and wrist during hammer throws. The dynamic data could play a vital role in skill optimization, as they could be used to reveal the coordination between upper and lower limbs. In conjunction with wearable wire-tension measurement, various motor control patterns during the hammer throw could be demystified. Hence, such a wearable system could realize the real-time biomechanical feedback training for the hammer throw. Such an approach has a great potential to become a coach-friendly tool for effectively learning and/or training in practice. In short, our device could make three potential contributions to hammer throw learning and/or training: (1) Making scientific monitoring from a lab-based environment to in field, (2) simplifying a scientific quantification from using a complicated motion capture system to easily-applied wearables, and (3) transferring the biomechanical feedback training from a post-measurement one to a real-time one. Further studies are required to verify the potentials.

**Author Contributions:** Y.W. designed, prototyped, programmed the wearable system, and tested its performance; B.W., X.Z. and G.S. analyzed and interpreted the data; G.S. and H.L. proposed the architecture and improved the design; Y.W., H.L. and G.S. prepared the draft; all authors contributed to the revisions and proof reading of the article.

**Funding:** This research was funded by National Sciences and Engineering Research Council of Canada (NSERC), grant number RGPIN-2014-03648.

**Conflicts of Interest:** The authors declare no conflict of interest. The founding sponsors had no role in the collection, analyses, or interpretation of data; in the writing of the manuscript, and in the decision to publish the results.

#### References

1. Magill, R.A. *Motor Learning Concepts and Applications*, 6th ed.; McGraw-Hill: Boston, MA, USA, 2001.
2. Schmidt, R.; Lee, T. *Motor Learning and Performance: From Principles to Application*, 5th ed.; Human Kinetics: Windsor, ON, USA, 2013; p. 336.
3. Shan, G.; Westerhoff, P. Full body kinematic characteristics of the maximal instep Soccer kick by male soccer players and parameters related to kick quality. *Sports Biomech.* **2005**, *4*, 59–72. [[CrossRef](#)] [[PubMed](#)]

4. Zhang, X.; Shan, G. Where do golf driver swings go wrong?—Factors Influencing Driver Swing Consistency. *Scand. J. Med. Sci. Sports* **2014**, *24*, 749–757. [CrossRef] [PubMed]
5. Yu, D.; Yu, Y.; Wilde, B.; Shan, G. Biomechanical characteristics of the axe kick in Tae Kwon-Do. *Arch. Budo* **2012**, *8*, 213–218. [CrossRef]
6. Shan, G.; Visentin, P.; Zhang, X.; Hao, W.; Yu, D. Bicycle kick in soccer: Is the virtuosity systematically entrainable? *Sci. Bull.* **2015**, *60*, 819–821. [CrossRef]
7. IAAF. Hammer Throw. Available online: <https://www.iaaf.org/disciplines/throws/hammer-throw> (accessed on 11 May 2018).
8. Shan, G.; Zhang, X. From 2D leg kinematics to 3D full-body biomechanics—the past, present and future of scientific analysis of maximal instep kick in soccer. *Sports Med. Arthrosc. Rehabil. Ther. Technol.* **2011**, *3*, 23. [CrossRef] [PubMed]
9. Wan, B.; Shan, G. Biomechanical modeling as a practical tool for predicting injury risk related to repetitive muscle lengthening during learning and training of human complex motor skills. *SpringerPlus* **2016**, *5*, 441. [CrossRef] [PubMed]
10. Shan, G.; Daniels, D.; Wang, C.; Wutzke, C.; Lemire, G. Biomechanical analysis of maximal instep kick by female soccer players. *J. Hum. Mov. Stud.* **2005**, *49*, 149–168.
11. Shan, G. Influences of Gender and Experience on the Maximal Instep Soccer Kick. *Eur. J. Sport Sci.* **2009**, *9*, 107–114. [CrossRef]
12. Wang, Y.; Wan, B.; Li, H.; Shan, G. A wireless sensor system for a biofeedback training of hammer throwers. *SpringerPlus* **2016**, *5*, 1395. [CrossRef]
13. Aminian, K.; Najafi, B. Capturing human motion using body-fixed sensors: Outdoor measurement and clinical applications. *Comput. Animat. Virtual Worlds* **2004**, *15*, 79–94. [CrossRef]
14. Shan, G.; Zhang, X.; Li, X.; Hao, W.; Witte, K. Quantification of Golfer-club Interaction and Club-type's Affect on Dynamic Balance during a Golf Swing. *Int. J. Perform. Anal. Sport* **2011**, *11*, 417–426. [CrossRef]
15. Shan, G.; Yuan, J.; Hao, W.; Gu, M.; Zhang, X. Regression Equations related to the Quality Evaluation of Soccer Maximal Instep Kick for Males and Females. *Kinesiology* **2012**, *44*, 139–147.
16. Wan, B.; Shan, G.; Wang, Y.; Zhang, X.; Li, H. 3D Quantification of Key Parameters for Developing Wearables of Biomechanical Feedback Training in Hammer Throw. *Unpubl. Artic.* **2018**.
17. Sparkfun. SparkFun 6 Degrees of Freedom IMU Digital Combo Board—ITG3200/ADXL345. Available online: <https://www.sparkfun.com/products/retired/10121> (accessed on 16 August 2017).
18. Sparkfun. Teensy 3.2. Available online: <https://www.sparkfun.com/products/13736> (accessed on 16 August 2017).
19. Arduino. Wire Library. Available online: <https://www.arduino.cc/en/Reference/Wire> (accessed on 11 July 2015).
20. AnalogDevices. Digital Accelerometer. Available online: <https://www.sparkfun.com/datasheets/Sensors/Accelerometer/ADXL345.pdf> (accessed on 11 August 2017).
21. InvenSense. ITG-3200 Product Specification Revision 1.4. Available online: <https://www.sparkfun.com/datasheets/Sensors/Gyro/PS-ITG-3200-00-01.4.pdf> (accessed on 11 August 2017).
22. Won, S.-H.; Melek, W.; Golnaraghi, F. Position and orientation estimation using Kalman filtering and particle diltering with one IMU and one position sensor. In Proceedings of the 34th Annual Conference of IEEE Industrial Electronics, Orlando, FL, USA, 10–13 November 2008; pp. 3006–3010.
23. Madgwick, S. An efficient orientation filter for inertial and inertial/magnetic sensor arrays. *Rep. x-io Univ. Bristol (UK)* **2010**, *25*, 113–118.
24. Shan, G.B.; Visentin, P. A quantitative three-dimensional analysis of arm kinematics in violin performance. *Med. Probl. Perform. Artist.* **2003**, *18*, 3–10.
25. Shan, G. Biomechanical Know-how of Fascinating Soccer-kicking Skills—3D, Full-body Demystification of Maximal Instep Kick, Bicycle kick & Side Volley. In Proceedings of the 8th International Scientific Conference on Kinesiology, Zagreb, Opatija, Croatia, 10–14 May 2017; pp. 133–135.
26. Visentin, P.; Li, S.; Tardif, G.; Shan, G. Unraveling mysteries of personal performance style; biomechanics of left-hand position changes (shifting) in violin performance. *PeerJ* **2015**, *3*, e1299. [CrossRef] [PubMed]
27. Shan, G.; Zhang, X.; Meng, M.; Wilde, B. A Biomechanical Study for Developing Wearable-Sensor System to Prevent Hip Fractures among Seniors. *Appl. Sci.* **2017**, *7*, 771. [CrossRef]

28. Shan, G.; Bohn, C. Anthropometrical data and coefficients of regression related to gender and race. *Appl. Ergon.* **2003**, *34*, 327–337. [[CrossRef](#)]
29. Shan, G.; Bohn, C.; Sust, M.; Nicol, K. How can dynamic rigid-body modeling be helpful in motor learning?—Learning performance using dynamic modeling. *Kinesiology* **2004**, *36*, 182–191.
30. Shan, G.; Sust, M.; Simard, S.; Bohn, C.; Nicol, K. How can dynamic rigid-body modeling be helpful in motor learning?—Diagnosing performance using dynamic modeling. *Kinesiology* **2004**, *36*, 5–14.
31. Ballreich, R.; Baumann, W. *Grundlagen der Biomechanik des Sports (The Basics of Biomechanics in Sports)*; Enke Verlag: Stuttgart, Germany, 1996.
32. Shan, G.; Zhang, X.; Wan, B.; Yu, D.; Wilde, B.; Visentin, P. Biomechanics of Coaching Maximal Instep Soccer Kick for Practitioners. *Interdiscip. Sci. Rev.* **2018**. [[CrossRef](#)]



© 2018 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).

Article

# Calculating the Effective Center Wavelength for Heterodyne Interferometry of an Optical Frequency Comb

Shilin Xiong<sup>1</sup>, Yue Wang<sup>1</sup>, Yawen Cai<sup>2</sup>, Jiuli Liu<sup>2</sup>, Jie Liu<sup>2</sup> and Guanhao Wu<sup>1,\*</sup>

<sup>1</sup> State Key Laboratory of Precision Measurement Technology and Instruments, Department of Precision Instrument, Tsinghua University, Beijing 100084, China; xiongshilin0808@163.com (S.X.); wy310275@163.com (Y.W.)

<sup>2</sup> Institute of Spacecraft System Engineering, China Academy of Space Technology, Beijing 100094, China; ywcai02@gmail.com (Y.C.); justin\_2009@163.com (J.L.); liujie@cast.cn (J.L.)

\* Correspondence: guanhaowu@mail.tsinghua.edu.cn; Tel.: +86-010-6279-8763

Received: 11 November 2018; Accepted: 29 November 2018; Published: 3 December 2018



**Abstract:** Heterodyne interferometry based on an optical frequency comb (OFC) is a powerful tool for distance measurement. In this paper, a method to calculate the effective center wavelength of wide spectrum heterodyne interference signal was explored through both simulation and experiment. Results showed that the effective center wavelength is a function of the spectra of the two interfered beams and time-delay of the two overlapped pulses. If the product of the spectra from two arms is symmetric, the effective center wavelength does not change with time-delay of the two pulses. The relative difference between the simulation and experiment was less than 0.06%.

**Keywords:** optical frequency comb; heterodyne interferometry; center wavelength

## 1. Introduction

An optical frequency comb (OFC) emits an evenly spaced ultra-short pulse train with a broad spectrum consisting of discrete, narrow lines with uniform mode-spacing [1,2]. The absolute frequency of each mode can be expressed as:

$$f_m = m f_{\text{rep}} + f_{\text{ceo}}, \quad (1)$$

where  $f_{\text{rep}}$  is the repetition rate,  $f_{\text{ceo}}$  stands for the carrier-envelope-offset frequency, and  $m$  is the mode order. When  $f_{\text{rep}}$  and  $f_{\text{ceo}}$  are stabilized referencing a frequency standard, OFC becomes an ultra-precise ruler in the space, time, and frequency domain [3–5]. Therefore, it is useful for absolute distance measurement.

In the past decade, numerous methods based on OFC have been proposed to measure absolute distance with high precision. These methods can be categorized into several groups according to the measurement principle and include: using the inter-mode beat signals of the comb [6,7], applying dispersive interferometry [8–13], using the pulse separation distance as a ruler [14–19], and the dual-comb method [20–25]. In order to suppress the effect of intensity noise, heterodyne interferometry has been introduced into the OFC distance measurement system [26]. The OFC heterodyne interferometry displays excellent results in temporal coherence interferometry, two-color correction of the refractive index of air [27,28], and pulse-to-pulse alignment [29]. Therefore, it is a powerful tool for distance measurement.

In a traditional laser heterodyne interferometer, the accuracy of the wavelength is very important for the distance measurement. However, in an OFC heterodyne interferometer, the light source has a wide optical spectrum; the spectral width of an optical frequency comb is normally tens of nanometers.

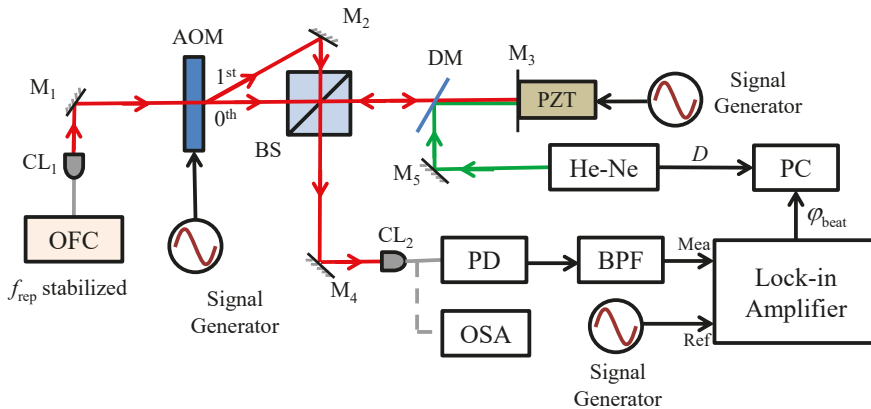


Additionally, the optical spectra of the beams in the two arms of the interferometer can be different, and thus, it is ambiguous to determine the center wavelength. Determining the effective center wavelength of the interference signal is a basic but essential question for the OFC heterodyne interferometry which has not yet been investigated in detail.

In order to resolve this problem, this paper introduced a method for calculating the effective center wavelength of the OFC heterodyne interference signal in an equal-arm interferometer included in a commercial interferometer. The method was then verified by simulation and experimental results.

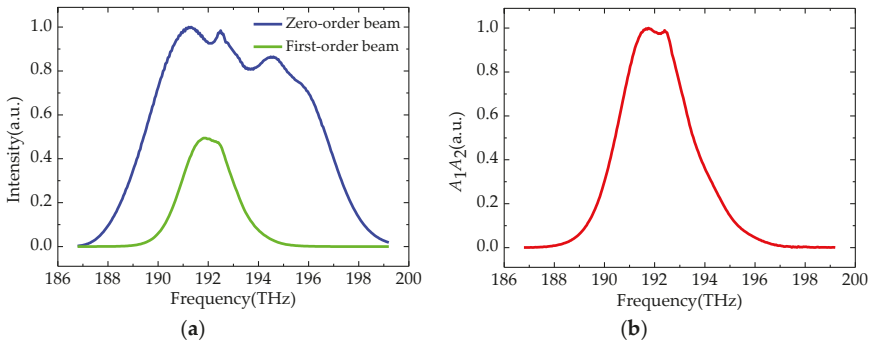
## 2. Methods and Experiments

A schematic of the OFC heterodyne interferometer used in this study is provided in Figure 1. The light source (OFC) was a homemade mode-locked erbium-doped fiber femtosecond laser. The central wavelength of the OFC was 1560 nm, the full width at half maximum (FWHM) of the spectrum was 55 nm, the output power was approximately 8 mW, and the repetition frequency was stabilized to a frequency synthesizer (78 MHz, 33250A, Agilent, Santa Clara, CA, USA), which was referenced to an atomic clock (5071A, Symmetricom, San Jose, CA, USA).



**Figure 1.** Schematic of the optical frequency comb (OFC) heterodyne interference system. Gray lines: optical fibers; red and green lines: optical paths in free space; black lines: electrical connections. OFC: optical frequency comb; CL: collimating lens; M: mirror; AOM: acoustic optical modulator; BS: beam splitter; DM: dichroic mirror; PZT: piezo-electric transducer; OSA: optical spectrum analyzer; PD: photodetector; BPF: bandpass filter; He-Ne: commercial interferometer; PC: computer.

The output laser beam from the OFC passed through an acousto-optic modulator (AOM, MGAS80-A1, AA Opto Electronic, Orsay, France) driven by a sinusoidal signal at a constant frequency  $f_{\text{AOM}} = 80$  MHz. After the AOM, the zero-order beam travels along the original direction with the original frequency, while the first-order beam spreads in another direction due to optical diffraction, and its frequency is shifted by  $f_{\text{AOM}}$ . The zero-order beam travels through a beam splitter (BS) and arrives at mirror  $M_3$ , then reflects back to BS and overlaps with the first-order beam which is adjusted by mirror  $M_2$ . Two beams were coupled into an optical fiber together by a collimating lens ( $CL_2$ ). The optical lengths of the two beams were set to be equal. The optical spectra of the two beams were then individually measured by an optical spectrum analyzer (OSA, AQ6370C, Yokogawa Electric, Musashino, Tokyo, Japan). The first-order beam is referred to as the reference arm, while the zero-order beam is called the measurement arm. Figure 2a illustrates the optical spectrum of the two arms. Note that the optical spectra of the two beams are different, due to the diffraction of the AOM.



**Figure 2.** (a) Optical spectrum of two beams; (b) Product of electric amplitude of two arms.

Because the frequency of the first-order beam was shifted by  $f_{AOM}$ , a heterodyne interference signal was detected by the photodetector (PD, Model 1811, New Focus, CA, USA), and the frequency of the heterodyne interference signal was  $f_{beat} = f_{AOM} - f_{rep} = 2$  MHz. The beat signal was extracted by a bandpass filter and then sent into a lock-in amplifier (SR844, Stanford Research System, Sunnyvale, CA, USA) to measure the phase compared with the reference signal.

The beam of the commercial interferometer overlapped with the beam of OFC at the dichroic mirror (DM). The target mirror  $M_3$  was driven by a piezo-electric transducer (PZT). The OFC heterodyne interferometer and the commercial interferometer were both used to measure the displacement of  $M_3$  simultaneously, see Section 3 for details. According to the comparison results, the effective center wavelength of the OFC heterodyne interference signal can then be estimated.

In the following section, a theoretical method to calculate the effective center wavelength of the OFC heterodyne interferometer was introduced. The complex amplitude of the electric field of two beams can be expressed as:

$$E_1(t) = \sum_m A_1 e^{i[2\pi f_m(t-t_1) + \beta_m]}, \quad (2)$$

$$E_2(t) = \sum_m A_2 e^{i[2\pi(f_m + f_{beat})(t-t_2) + \beta_m]}, \quad (3)$$

where  $A_1(f_m) \propto \sqrt{P_1(f_m)}$ ,  $A_2(f_m) \propto \sqrt{P_2(f_m)}$ ,  $P_1$ , and  $P_2$  are the optical spectral intensity of the two beams as shown in Figure 2a. The imaginary unit is  $i$ ,  $t_1$ , and  $t_2$  are the time delay of the two beams, and  $\beta_m$  is the initial phase of the  $m^{\text{th}}$  mode. The interference signal intensity is:

$$I = |E_1 + E_2|^2 = (E_1 + E_2) \cdot (E_1 + E_2)^*. \quad (4)$$

After a bandpass filter, the beat signal of OFC heterodyne interferometry is:

$$I_{beat} = \sum_m 2A_1 A_2 \cos[2\pi f_{beat}(t - t_2) + 2\pi f_m(t_1 - t_2)]. \quad (5)$$

Equation (5) is equivalent to this form:

$$I_{beat} = 2a \cos[2\pi f_{beat}(t - t_2) + \varphi_{beat}], \quad (6)$$

where  $\varphi_{beat}$  is the phase measured by the lock-in amplifier, and:

$$a \cos \varphi_{beat} = \sum_m A_1(f_m) A_2(f_m) \cos 2\pi f_m \tau, \quad (7)$$

$$a \sin \varphi_{beat} = \sum_m A_1(f_m) A_2(f_m) \sin 2\pi f_m \tau, \quad (8)$$

where  $\tau = t_1 - t_2$  is the relative time delay of the two beams. According to Equations (7) and (8):

$$\tan \varphi_{\text{beat}} = \frac{\sum_m A_1(f_m)A_2(f_m) \sin 2\pi f_m \tau}{\sum_m A_1(f_m)A_2(f_m) \cos 2\pi f_m \tau} \tag{9}$$

then  $\varphi_{\text{beat}}$  is a function of relative time delay  $\tau$ :

$$\varphi_{\text{beat}} = \varphi(\tau). \tag{10}$$

For the continuous-wave laser heterodyne interferometer,  $\varphi(\tau) = 2\pi f_c \tau$ , thus, the effective center frequency for OFC heterodyne interferometry is:

$$f_c = \frac{1}{2\pi} \varphi'(\tau). \tag{11}$$

The effective center wavelength can then be calculated using  $\lambda_c = c/f_c$ , where  $c$  is the velocity of light in the vacuum.

Figure 2b illustrates the product of the electric amplitude of the two beams, which is the function of the optical frequency. If the product is symmetrical for the frequency at  $f_{m0}$ , Equation (9) can be simplified as:

$$\tan \varphi_{\text{beat}} = \frac{\sin 2\pi f_{m0} \tau}{\cos 2\pi f_{m0} \tau} = \tan 2\pi f_{m0} \tau, \tag{12}$$

therefore,  $\varphi_{\text{beat}} = 2\pi f_{m0} \tau$ , meaning that the effective center wavelength does not change with the relative time delay of two beams. In this study, however, the product was not symmetric, thus, the effective center wavelength was calculated by simulation.

In this simulation, the parameters of the system were the same as the experimental setup, including the electric amplitude of the two beams and repetition frequency of OFC. The carrier-envelope-offset frequency was regarded as zero. The phase  $\varphi_{\text{beat}}$  of the heterodyne interference signal was calculated at different relative time delay according to Equation (9), and numerical differentiation methods from Equations (10) and (11) were used to calculate the effective center wavelength at different relative time delay.

### 3. Results

To calculate the effective center wavelength of the OFC heterodyne interferometer, a square wave signal generated from a signal generator controlled the displacements of the PZT. The frequency of the square wave signal is 10 Hz, with a peak-to-peak value of 0.4 V. The sensitivity of PZT to light path variations was about 6.4  $\mu\text{m}/\text{V}$ , and the positions of  $M_3$  were monitored precisely by the commercial interferometer.

Figure 3a shows the phase change of OFC heterodyne interferometer and the commercial interferometers while tuning the displacements of PZT.

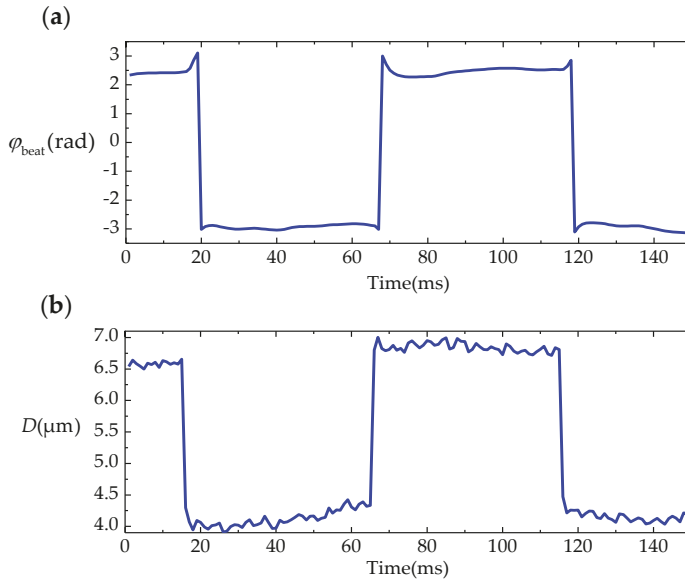
The effective center wavelength at this position is calculated by:

$$\lambda_c = \frac{2\pi n \Delta D}{\Delta \varphi}, \tag{13}$$

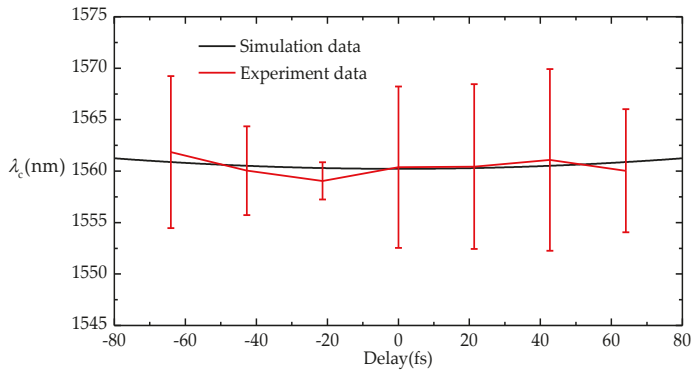
where  $n$  is the refractive index of air. Since the displacement was larger than the wavelength, it was necessary to unwrap the phase when calculating the effective center wavelength. Environmental parameters were recorded throughout the experiment in order to calculate the air refractive index for both commercial interferometer and OFC heterodyne interferometer.

To evaluate the stability of the effective center wavelength while the relative position of two overlapped pulses change, the bias voltage of the square wave was altered to shift the equilibrium position of PZT. The bias voltage was changed by six steps at 1 V interval (corresponding to 6.4  $\mu\text{m}$ ).

At each position, the effective center wavelength was measured 10 times. The average value and standard deviation are presented in Figure 4. The simulation results are also provided for comparison.



**Figure 3.** (a)  $\varphi_{\text{beat}}$  change with time measured by lock-in amplifier; (b) displacement change with time measured by commercial interferometer.



**Figure 4.** Effective center wavelength  $\lambda_c$  with different pulse delay. The data was measured by scanning the mirror  $M_3$  through modulating the voltage of PZT. Black line: results of simulation; red line: results of experiments. Error bars represents  $\pm \sigma$  (standard deviation) for 10 measurements.

When the pulse delay changed from  $-80$  fs to  $80$  fs, the effective center wavelength changed approximately  $1.02$  nm according to the simulation results. This is because the product of the electric amplitude of two beams was not symmetric, and the function of phase  $\varphi_{\text{beat}}$  change over the relative time delay was not a linear function. The biggest relative difference between the simulation and experiment was approximate to  $0.06\%$ , which was predominantly caused by the random error of the commercial interferometer and the phase drifting caused by the AOM. The optical frequency comb had an especially wide spectrum; its coherence length was very short. When calculating the effective center wavelength for heterodyne interferometry, the displacement was only about  $2.5 \mu\text{m}$ . The repeatability

of commercial interferometer was only about 15 nm for the short displacement measurement, which was the main error source in the evaluation results shown in Figure 4. Note that the two beams in this experiment come out of the same optical fiber, thus, the initial phase  $\beta_m$  of two beams was equal. However, if the beams would have come from different optical fiber, the initial phase  $\beta_m$  of the two beams would be different because of the chirp of optical fiber. This case requires further investigation for calculating the effective center wavelength.

#### 4. Conclusions

A method was proposed to calculate the effective center wavelength for an OFC heterodyne interferometer based on the spectra of two interfered beams. An experimental setup was established to verify the theoretical formula. The results show that the theoretical analysis corresponds well with the experimental results, and illustrate that the proposed model is reasonable and effective. This method is an important tool for OFC heterodyne interferometry that can be used for ranging or pulse alignment and other applications.

**Author Contributions:** G.W. and S.X. conceived and designed the experiments; Y.C., J.L. (Jiuli Liu), and J.L. (Jie Liu) discussed the experimental scheme; S.X. and Y.W. performed the experiments; S.X. analyzed the data; S.X. and G.W. wrote the paper. All authors have read and approved the final manuscript.

**Funding:** This research was funded by the National Key Research and Development Project (2016YFF0101804), the National Natural Science Foundation of China (61575105, 61611140125), the Beijing Natural Science Foundation (3182011), and Shenzhen fundamental research funding (Grant No. JCYJ20170412171535171).

**Conflicts of Interest:** The authors declare no conflict of interest.

#### References

1. Cundiff, S.T.; Ye, J. Colloquium: Femtosecond optical frequency combs. *Rev. Mod. Phys.* **2003**, *75*, 325–342. [[CrossRef](#)]
2. Kim, S.-W. Combs rule. *Nat. Photonics* **2009**, *3*, 313–314. [[CrossRef](#)]
3. Cundiff, S.T. Phase stabilization of ultrashort optical pulses. *J. Phys. D-Appl. Phys.* **2002**, *35*, R43–R59. [[CrossRef](#)]
4. Wu, X.; Yang, L.; Zhang, H.; Yang, H.; Wei, H.; Li, Y. Hybrid mode-locked Er-fiber oscillator with a wide repetition rate stabilization range. *Appl. Opt.* **2015**, *54*, 1681–1687. [[CrossRef](#)]
5. Nakajima, Y.; Minoshima, K. Highly stabilized optical frequency comb interferometer with a long fiber-based reference path towards arbitrary distance measurement. *Opt. Express* **2015**, *23*, 25979–25987. [[CrossRef](#)]
6. Minoshima, K.; Matsumoto, H. High-accuracy measurement of 240-m distance in an optical tunnel by use of a compact femtosecond laser. *Appl. Opt.* **2000**, *39*, 5512–5517. [[CrossRef](#)]
7. Bitou, Y.; Schibli, T.R.; Minoshima, K. Accurate wide-range displacement measurement using tunable diode laser and optical frequency comb generator. *Opt. Express* **2006**, *14*, 644–654. [[CrossRef](#)]
8. Dandliker, R.; Salvade, Y.; Zimmermann, E. Distance measurement by multiple-wavelength interferometry. *J. Opt.-Nouv. Rev. D Opt.* **1998**, *29*, 105–114. [[CrossRef](#)]
9. Joo, K.N.; Kim, S.W. Absolute distance measurement by dispersive interferometry using a femtosecond pulse laser. *Opt. Express* **2006**, *14*, 5954–5960. [[CrossRef](#)]
10. Salvade, Y.; Schuhler, N.; Leveque, S.; Le Floch, S. High-accuracy absolute distance measurement using frequency comb referenced multiwavelength source. *Appl. Opt.* **2008**, *47*, 2715–2720. [[CrossRef](#)]
11. Falaggis, K.; Towers, D.P.; Towers, C.E. Multiwavelength interferometry: Extended range metrology. *Opt. Lett.* **2009**, *34*, 950–952. [[CrossRef](#)]
12. Cui, M.; Zeitouny, M.G.; Bhattacharya, N.; van den Berg, S.A.; Urbach, H.P. Long distance measurement with femtosecond pulses using a dispersive interferometer. *Opt. Express* **2011**, *19*, 6549–6562. [[CrossRef](#)]
13. van den Berg, S.A.; Persijn, S.T.; Kok, G.J.P.; Zeitouny, M.G.; Bhattacharya, N. Many-Wavelength Interferometry with Thousands of Lasers for Absolute Distance Measurement. *Phys. Rev. Lett.* **2012**, *108*. [[CrossRef](#)] [[PubMed](#)]

14. Schuhler, N.; Salvade, Y.; Leveque, S.; Daendliker, R.; Holzwarth, R. Frequency-comb-referenced two-wavelength source for absolute distance measurement. *Opt. Lett.* **2006**, *31*, 3101–3103. [[CrossRef](#)] [[PubMed](#)]
15. Cui, M.; Schouten, R.N.; Bhattacharya, N.; van den Berg, S.A. Experimental demonstration of distance measurement with a femtosecond frequency comb laser. *J. Eur. Opt. Soc.-Rapid Publ.* **2008**, *3*. [[CrossRef](#)]
16. Joo, K.-N.; Kim, Y.; Kim, S.-W. Distance measurements by combined method based on a femtosecond pulse laser. *Opt. Express* **2008**, *16*, 19799–19806. [[CrossRef](#)] [[PubMed](#)]
17. Balling, P.; Kren, P.; Masika, P.; van den Berg, S.A. Femtosecond frequency comb based distance measurement in air. *Opt. Express* **2009**, *17*, 9300–9313. [[CrossRef](#)]
18. Cui, M.; Zeitouny, M.G.; Bhattacharya, N.; van den Berg, S.A.; Urbach, H.P.; Braat, J.J.M. High-accuracy long-distance measurements in air with a frequency comb laser. *Opt. Lett.* **2009**, *34*, 1982–1984. [[CrossRef](#)]
19. Lee, J.; Kim, Y.-J.; Lee, K.; Lee, S.; Kim, S.-W. Time-of-flight measurement with femtosecond light pulses. *Nat. Photonics* **2010**, *4*, 716–720. [[CrossRef](#)]
20. Liu, T.-A.; Newbury, N.R.; Coddington, I. Sub-micron absolute distance measurements in sub-millisecond times with dual free-running femtosecond Er fiber-lasers. *Opt. Express* **2011**, *19*, 18501–18509. [[CrossRef](#)]
21. Lee, J.; Han, S.; Lee, K.; Bae, E.; Kim, S.; Lee, S.; Kim, S.-W.; Kim, Y.-J. Absolute distance measurement by dual-comb interferometry with adjustable synthetic wavelength. *Meas. Sci. Technol.* **2013**, *24*. [[CrossRef](#)]
22. Hebert, N.B.; Boudreau, S.; Genest, J.; Deschenes, J.-D. Coherent dual-comb interferometry with quasi-integer-ratio repetition rates. *Opt. Express* **2014**, *22*, 29152–29160. [[CrossRef](#)] [[PubMed](#)]
23. Wu, G.; Zhou, Q.; Shen, L.; Ni, K.; Zeng, X.; Li, Y. Experimental optimization of the repetition rate difference in dual-comb ranging system. *Appl. Phys. Express* **2014**, *7*. [[CrossRef](#)]
24. Zhang, H.; Wei, H.; Wu, X.; Yang, H.; Li, Y. Absolute distance measurement by dual-comb nonlinear asynchronous optical sampling. *Opt. Express* **2014**, *22*, 6597–6604. [[CrossRef](#)] [[PubMed](#)]
25. Wu, G.; Xiong, S.; Ni, K.; Zhu, Z.; Zhou, Q. Parameter optimization of a dual-comb ranging system by using a numerical simulation method. *Opt. Express* **2015**, *23*, 32044–32053. [[CrossRef](#)]
26. Wu, G.H.; Liao, L.; Xiong, S.L.; Li, G.Y.; Cai, Z.J.; Zhu, Z.B. Synthetic wavelength interferometry of an optical frequency comb for absolute distance measurement. *Sci. Rep.* **2018**, *8*, 7. [[CrossRef](#)] [[PubMed](#)]
27. Wu, G.; Arai, K.; Takahashi, M.; Inaba, H.; Minoshima, K. High-accuracy correction of air refractive index by using two-color heterodyne interferometry of optical frequency combs. *Meas. Sci. Technol.* **2013**, *24*, 015203. [[CrossRef](#)]
28. Wu, G.; Takahashi, M.; Arai, K.; Inaba, H.; Minoshima, K. Extremely high-accuracy correction of air refractive index using two-colour optical frequency combs. *Sci. Rep.* **2013**, *3*, 01894. [[CrossRef](#)]
29. Wu, G.; Takahashi, M.; Inaba, H.; Minoshima, K. Pulse-to-pulse alignment technique based on synthetic-wavelength interferometry of optical frequency combs for distance measurement. *Opt. Lett.* **2013**, *38*, 2140–2143. [[CrossRef](#)]



© 2018 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).

Article

# Selection of Optimal Path Control Algorithms for Probe Heads Used on Five-Axis Measuring Systems

Adam Gąska \*, Piotr Gąska, Maciej Gruza and Jerzy Sładek

Laboratory of Coordinate Metrology, Cracow University of Technology, 31-155 Kraków, Poland; pjgaska@gmail.com (P.G.); gruzam@interia.pl (M.G.); sladek@mech.pk.edu.pl (J.S.)

\* Correspondence: agaska@mech.pk.edu.pl; Tel.: +48-12-374-32-38

Received: 30 October 2018; Accepted: 22 November 2018; Published: 2 December 2018



**Featured Application:** Results of this work may be directly applied by all users of probe heads with the possibility of continuous indexation for the reduction of measurement errors by the selection of optimal path control algorithms.

**Abstract:** The utilization of rotational movements of a probing system during points measurements contributes to the reduction of measurement duration and increases measurement repeatability. However, knowledge on such behavior and accuracy of probing systems is still unsatisfactory. Machines combined with articulating probing systems that have the ability of continuous indexation become redundant systems, which means that the same points can be measured using almost infinite mutual configurations of the machine and probe stylus orientations. Therefore, the proper selection of inspection path planning method becomes one of the main factors affecting the accuracy of the measurement. It is possible to assess the impact of this factor on the accuracy of the measurement by comparing the results of the measurements of gauge elements, which are done using different path controlling algorithms. After that, the best method for basic measuring tasks can be chosen in order to reduce measurement errors. Measurements of the multi-feature check gauge, using the default method for path planning and those chosen on the basis of described experiments, indicates that the improvement of accuracy may reach several microns. Results presented in this paper can be directly transferred to similar systems and measuring tasks, which are commonly met in industrial and scientific practice.

**Keywords:** five-axis system; CMM; dimensional measurements; inspection planning; accuracy

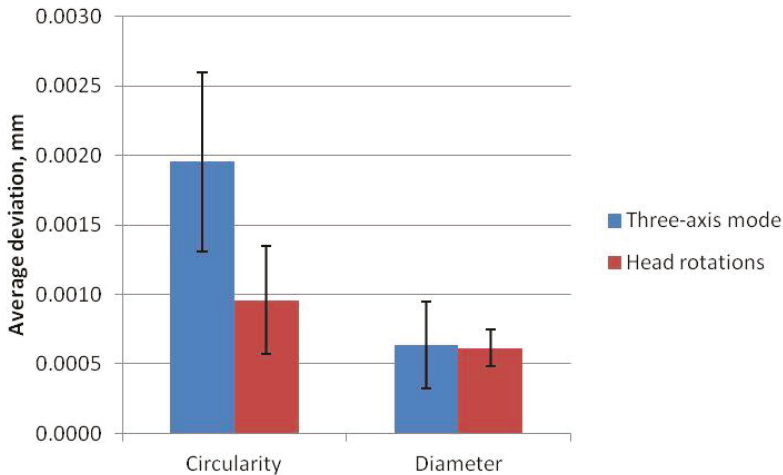
## 1. Introduction

### 1.1. State of Art in Modern Coordinate Measurement Metrology

Changes that can be observed in industry over the last years are commonly termed as a fourth industrial revolution. They are aimed at tightening the connection between individual components of the production process in order to reduce the manufacturing time and improve the quality of produced goods [1,2]. This trend is present in all industry branches connected with the production process, including quality control, especially coordinate metrology. Currently, the efforts of producers of measuring systems are focused mainly on the development of contactless techniques such as computed tomography (CT) or 3D scanners, often combined with industrial robots [3–5]. Their advantages are well known, most of all their high measurement speed and, in case of CT, their possibility of part interior inspection. However, the accuracy of mentioned systems still cannot be compared to classic tactile coordinate measuring machines (CMMs), which for decades have been the key element of modern, automated quality control.

Classic CMMs are still developed and enriched with solutions that improve their accuracy and efficiency. Utilization of the articulating probing systems that have the ability of continuous indexation (described in details in sub Section 1.2) fits in with this trend. A machine equipped with such a device becomes the five-axis system, capable of carrying out measurements with rotational movements of the probing system. In case of measurements of rotational features, it may cause a significant reduction of measurement duration and, for some measuring tasks, even accuracy improvement.

Figure 1 shows results of the experiment that involved reference ring measurements performed in two different modes on the same five-axis measuring system; once using only transitional moves of a CMM during probing, and once utilizing only the rotational movements of an articulating probing system. In both cases the applied measurement strategy was the same: The ring was measured in 32 points, distributed evenly across half of the reference ring’s height, and, for both tested modes, measurements were performed 10 times. As can be seen, measurements performed with the rotation of the probing system reduced values of errors of form deviation measurements by nearly half, in relation to traditional three-axis measurements, and also contributed to the improvement in measurement repeatability.



**Figure 1.** The results of the experiment showing the difference between measurements done using only the classic, three-axis coordinate measuring machine (CMM) and measurements done using measuring head movements. Error bars present the range of individual measurement results.

Additionally, it should be mentioned that the time needed to perform measurements with rotations of the probing system was two times shorter than in the case of the three-axis measurement. That results show that articulating probing systems are an interesting addendum to the possibilities of classic CMMs and they should be developed in the future. However, there are still many questions arising around this subject, with questions about articulating probing systems’ accuracy and factors that may affect such probing systems’ performance, being the most important ones. The software path control method, software correction, and algorithms used during measurement points determination are crucial for the proper operation of such measuring systems. This paper focuses on the first of mentioned factors, whereby we show the changes in measurement results depending on chosen path control method, and present how proper measurement strategy selection can contribute to the improvement of performance of the five-axis measuring system.

Considerations presented in this paper are topically related to the research on computer-aided inspection planning and the optimization of measurement strategy. There are a lot of papers focused on the optimization of the strategy of measurements performed on CMMs that use indexing probe



heads ([6–8] to cite few of them), but there are no publications dealing with this topic for probe heads that have the possibility of continuous articulation.

### 1.2. Five-Axis Coordinate Measuring System

Articulating probing systems have been utilized in CMM measurements for many years. They allow the ability to change the orientation of the stylus using two rotations around mutually perpendicular axes. Two configurations are used in their construction. In the first approach, the axis running through the stylus coincides with the probing system's vertical axis of rotation (axial adjustment); whereas in the second solution, the axis running through the stylus is parallel to the vertical axis of rotation of the probing system, and offset from it at a specified distance (side adjustment). The first solution is regarded as more accurate; however, on the contrary to the second approach, it has a limited range for rotation about the horizontal axis of a probe. Another possible classification of articulating probing systems can be found in [9], dividing them depending on the qualification process. For some probing systems the experimental qualification is needed for all orientations used during measurement; however, in the second option qualification is done only in several positions and then interpolated for any orientation. Despite the used solution, articulating probing systems working with classic CMMs have a fixed orientation during coordinate measurements, which is kept by a locking mechanism, or by the servo control system. The stylus orientation can be changed before or after point probing, but actual measurements are done only using translational moves of machine elements. Additionally, it should be noted that usually the producers specify the step of the angular increment for possible rotations, for example  $2.5^\circ$ ,  $7.5^\circ$ , or  $15^\circ$ . The reason for the requirement of the proper operation of the articulating probing systems is the high repeatability of angular positioning. Otherwise, any change in orientation would entail the requalification of the probe head.

In five-axis measuring systems the articulating probing system becomes a crucial element of the whole device. Such a probe can remain fixed in a chosen orientation, but its rotary motions can also be utilized during measurements. In that case, the data set needed for point coordinate determination is expanded by information about the angular position of the probe head, which is given by angular encoders. The probe heads utilize axial adjustment, with two orthogonal axes of rotation. The range of possible rotations for the vertical axis of revolution, which will be called B axis in the rest of the paper, is  $-180^\circ$  to  $180^\circ$ , and for the horizontal axis, which will be called A axis in the rest of the paper, the range from  $-115^\circ$  to  $115^\circ$ . The probing system is oriented vertically (along the machine quill) when A and B angles are set to  $0^\circ$ . Both measuring and touch trigger probes can work as a part of the five-axis measuring system, and both kinds use the same kinematics. One of the main differences between articulated probe heads used in three-axis CMMs and those utilized in five-axis measurements is the continuous indexation of the head. As the articulated probes use rotary moves for measurements, it must be possible to set the probe orientation freely within the working range of the probe. This is realizable thanks to the qualification process designed especially for such probing systems. It is divided into two steps: Firstly, the geometrical calibration of the head (as it is named by producers of such devices) is performed, then the probe is qualified. Both procedures are similar and based on calibration sphere measurements. However, the first procedure is done with short styli, to minimize bending and gravitational effects and, in turn, to accurately assess the geometrical errors of the head mechanism, whereas the second procedure is done for stylus and stylus tip configuration, which are chosen for measurements. The steps comprised in the procedures includes: The determination of calibration of ball position in the three-axis mode for the probe directed vertically, with B angle set to  $0^\circ$ , and then the second time with B angle set to  $180^\circ$ ; the determination of probing parameters, such as probing speed and maximal acceleration (this step is done without standard measurements); and the determination of the interpolated probe map, based on sphere measurements using head rotational movements, performed for different angular orientations of the probe head and for different cross-sections of the standard. Both parts of the qualification process take a considerable amount of time (more than 10 min for the Renishaw PH20 probe head), but they do not have to be repeated, in the case of the probe, until

a new probe configuration is needed, and in the case of the head geometrical calibration, until the probe head would be disassembled from the machine quill.

The problem of the articulating probing system’s accuracy was studied mostly for systems with fixed indexation, retained during probing. Described aspects of this problem include angular positioning repeatability [10], general accuracy [11,12], and probing system modelling [13]. Fewer studies have been focused on a probing system with continuous indexation, mostly done in the Laboratory of Coordinate Metrology (LCM), among them research has been done about such probing system’s hysteresis and the influence of stylus length [14] and the angular position of the probing system on measurement accuracy [15]. Conducted experiments revealed the existence of additional problems worth further examination; one of them is the issue of selecting the appropriate control algorithm for the measurement path for measurements utilizing the rotational movements of the probing system. This problem will be investigated in the following chapters of this paper.

**2. Path Control Algorithms of the Articulating Probing System**

The software responsible for path control and the determination of measured points coordinates plays a crucial role in the proper performance of the five-axis coordinate system. Most metrological programs allow the five-axis coordinate system to be used as a classic CMM equipped with a articulating probe, which retains a fixed orientation during point probing. Currently the most popular software, which allows the ability to fully use the rotational moves capabilities of a probing system, is Modus software by Renishaw. The programming environment is based on dimensional measuring interface specification (DMIS), but it is also possible to combine it with external applications in order to enrich the software capabilities. To control the probing system’s performance, Modus uses an option called “tilt” and “advance”, which defines the orientation of the probe during measurements. Tilt defines the rotation of the probe away from the plane, perpendicular to the measured surface normal vector [16]; and advance specifies the rotation around surface normal, while the negative value of this parameter means that the head is behind the tip [16]. As can be seen, the definitions are not exactly clear. However, the code created in the DMIS format uses, directly, the values of the A and B angles during measurement, or utilizes the Euler angles notation when the orientation is given in the part coordinate system (PCS). During the measuring path determination, Modus visualizes the location of the measured points, and the probe’s stylus orientation corresponding to them, by drawing a leader line from each measured point to the probing system’s central point [16]. Depending on the geometrical feature chosen for measurement, four different methods can be applied to manipulate the probe orientation: Guide point (GP), guide rail (GR), guided curve (GC), and fixed tilt and advance (FTAA). All mentioned methods are shown in Figure 2.

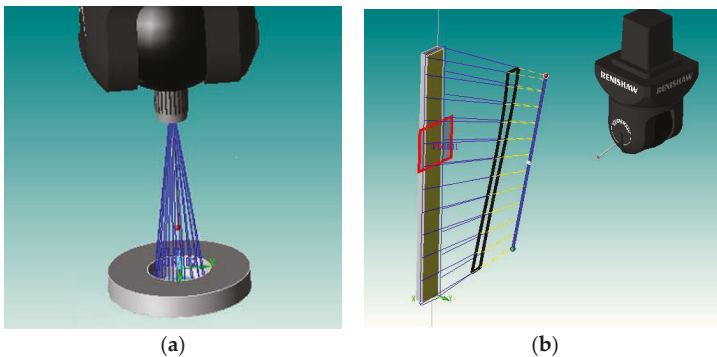
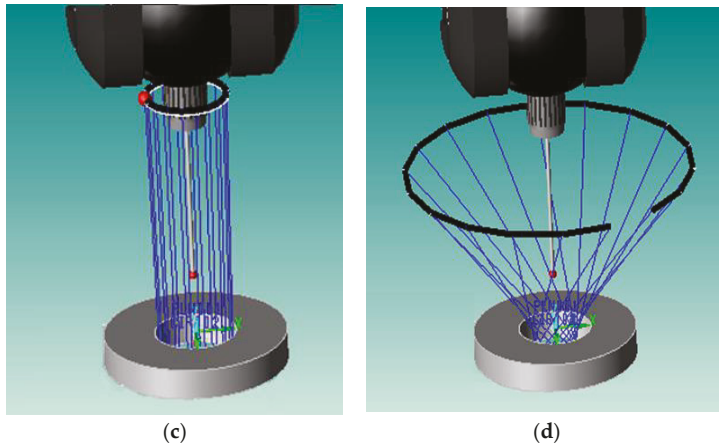


Figure 2. Cont.



**Figure 2.** Path control algorithms used during five-axis measurements: (a) Guide point (GP); (b) guide rail (GR); (c) guided curve (GC); and (d) fixed tilt and advance (FTAA).

Guide point is the option used for minimizing machine motion during probing. The centroid, which determines the position of the center point of the probing system, is calculated from points included in the inspection path. It is offset along the normal direction to the plane in which the measured points are defined. The user can also specify the position to be taken by the head during measurements. The guide rail method can be used to define the path for the center point of the probing system during measurement. It consists of two segments between three reference points, of which the position can be defined by the user. By default, the position of the starting and ending points of the rail is calculated as offset, in normal direction to the plane, on which the measured points lie at the two extreme points of the inspection path. The offset has the same length as the utilized stylus. The guided curve is calculated, taking into account all points in the inspection path, by projecting them along the normal direction to the plane, on which the inspection path lies. Fixed tilt and advance can be used to ensure that each measured point would be inspected using the same tilt and advance values. In the case of the measurement of rotational features, it results with big changes in the probing system’s center point position during measurements. Table 1 shows which method can be applied for the measurements of different geometrical elements.

**Table 1.** Path control algorithms that can be applied for measurement of specified features.

Feature	Guide Point (GP)	Guide Rail (GR)	Guided Curve (GC)	Fixed Tilt and Advance (FTAA)	Comments
Point	+	—	—	—	
Line	—	+	+	—	
Plane	+	+	—	—	
Circle	+	—	+	+	
Cylinder	+	—	+	+	
Cone	—	—	—	—	“Tilt and advance” option unavailable
Sphere	—	—	—	—	“Tilt and advance” option unavailable
Surface	—	—	—	—	“Tilt and advance” option unavailable

### 3. Experiment and Results

All measurements described in this article were performed on a five-axis measuring system, which was part of the LCM’s equipment. It consisted of a CMM with a moving bridge, Zeiss WMM 850S (ZEISS International, Oberkochen, Germany), and the articulating probing system, PH20 by Renishaw

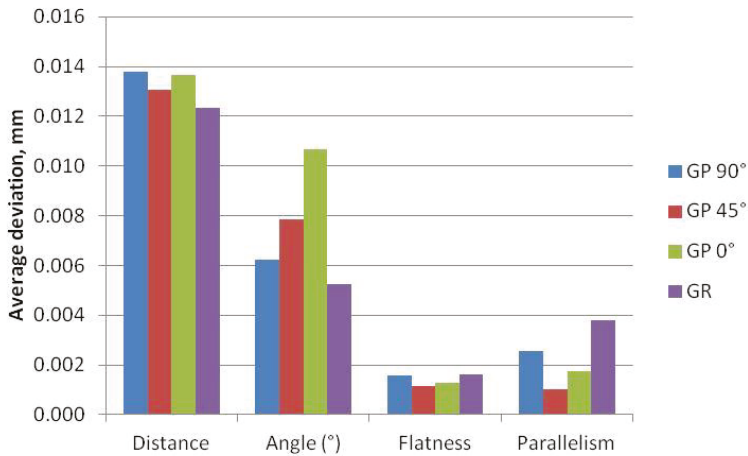
(Wotton-under-Edge, UK), working with the TP20 STD probe. The whole system was located in an air-conditioned room, with constant ambient conditions that were monitored. The temperature during experiments varied between 19.8 to 20.4 °C. The experiments undertaken at the LCM consisted of two parts: The first part was conducted to assess the influence of the different methods used for path control on the measurement accuracy. It involved multiple measurements of standard objects; gauge block of 200 mm length and gauge ring of 28 mm diameter (Figure 3).



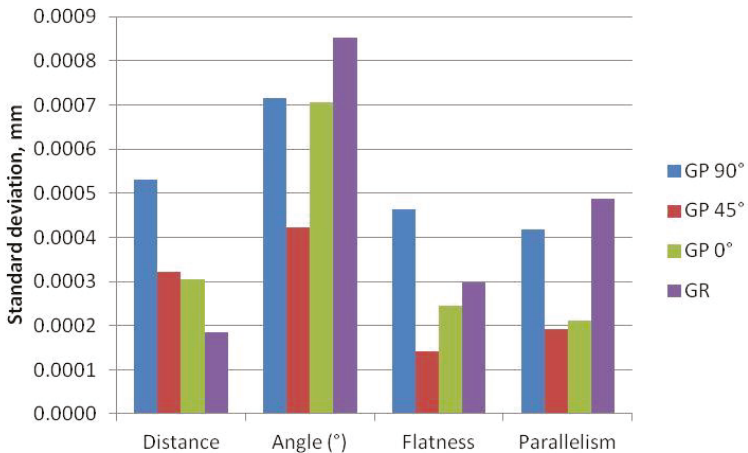
**Figure 3.** Photographs of the Zeiss WMM 850S (right side), the Renishaw PH20 measuring head (upper left), and the mounting of gauges during measurements (lower left).

The first step of the gauge block measurements was the determination of the local coordinate system, which was defined using three mutually perpendicular planes of gauge (the measurement of the planes that define the part coordinate system were performed using the standard three-axis mode, so only transitional movements of the machine were used during probing). After this step, the distance, angle between planes, parallelism, and flatness deviations were measured using different methods for inspection planning. Two methods were included in this part of the experiment, guide point and guide rail, as only these two were available for measurement of planes. Additionally, three different orientations of the guide point were utilized, marked as 0°, 45°, and 90° (denoted respectively as GP 0°, GP 45°, and GP 90° on Figures 4 and 5) with names corresponding to values around which A angle used during probing oscillated. Measurements performed with all mentioned methods were repeated 10 times and if it was possible in all cases the measurements were done with the same distribution of measured points. The results of measurements are presented in Figure 4 which shows average deviations from nominal values (taken from calibration certificate of gauge block) and in Figure 5 which gives information about standard deviation for each measurement task.

The most noticeable differences between the results obtained for the different path planning methods during gauge block measurements can be seen for the measurements of the angle and parallelism deviation. The smallest difference was obtained for measurements of flatness. In most cases the lowest values of standard deviation can be observed for the path control algorithms that also gave the lowest values of deviations.



**Figure 4.** The average deviations for the four chosen tasks, measured on the gauge block using different path control algorithms.



**Figure 5.** The standard deviations for the four chosen tasks, measured on the gauge block using different path control algorithms.

The next part of the experiment involved measurements of the ring gauge using all possible methods of inspection path planning. The local coordinate system was created in the center of the ring, with the ring’s axis determining the spatial alignment of the coordinate system. Again, the measurement needed for the coordinate system definition was performed in the three-axis mode of the measuring system. The ring was measured across half of its height, in 16 evenly-distributed points. Three available methods were tested: Guide point, guide curve, and fixed tilt and advance. In the guide curve option it was possible to specify the offset value for each point of the determined guide curve, in the direction from regarded point to the probing system’s central point. The negative value of this parameter moves all points of the curve closer to the central point of the probing system, and the positive one moves them away. Three cases were tested during experiment: Guide curve with offset set to 0 mm, −5 mm, and −10 mm (denoted as GC 0, GC −5, and GC −10, respectively). Additionally, the fixed tilt and advance method was tested for different values of tilt and advance: With a fixed tilt value of 30° and three advance values, 10°, 20°, and 30° (denoted as FTAA 30°/10°,

FTAA 30°/20°, and FTAA 30°/30°, respectively); and with a fixed advance value of 30° and three changing tilt values, 10°, 20°, and 30° (denoted as FTAA 10°/30°, FTAA 20°/30°, and FTAA 30°/30°, respectively). Measurements of the diameter and circularity deviation were repeated 10 times for all described methods and cases. Results of this part of the experiment are presented in Figure 6, which shows average deviations from nominal values (taken from the calibration certificate of the standard ring), and Figure 7, which gives information about the standard deviation for each measurement task.

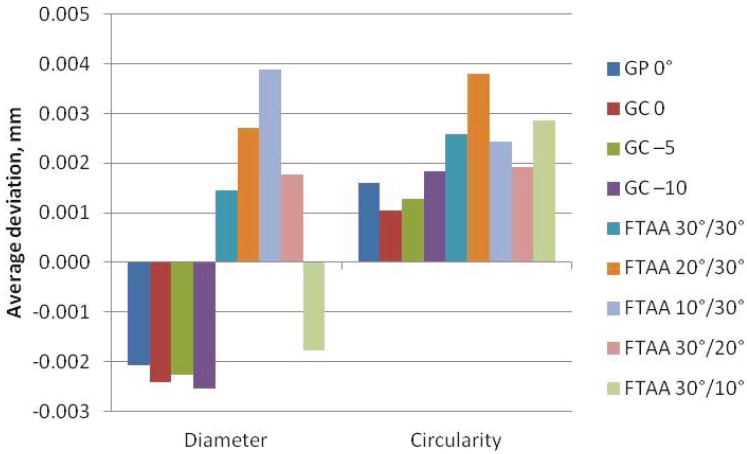


Figure 6. The average deviations for the two chosen tasks, measured on the ring standard using nine different path control algorithms.

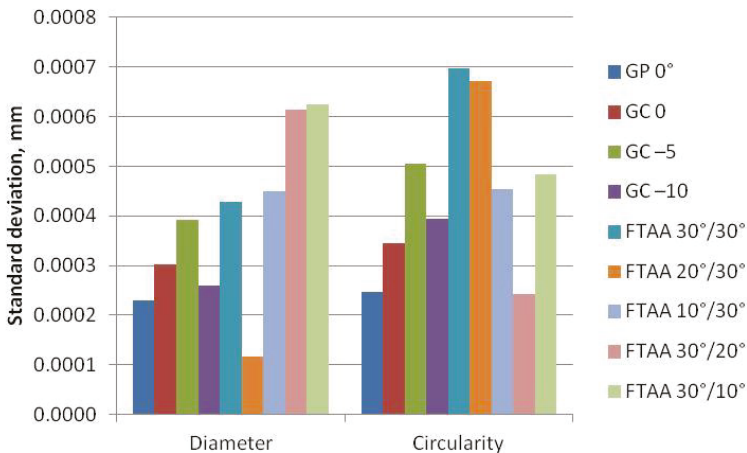


Figure 7. The standard deviations for the two chosen tasks, measured on the ring standard using nine different path control algorithms.

In the case of the standard ring measurements, the range of deviation produced by the different path control algorithms used for the determination of the diameter was equal to 6.3 μm, whereas for the measurement of circularity the deviation was 2.8 μm. What should be noticed, is that a slight change in parameter values for tilt and advance may cause significant changes in the results of both the diameter and circularity deviation measurements.

After completing the first part of the experiments, which involved gauges measurements, the obtained results (presented in Figures 4–7) were analyzed in order to find the inspection path planning methods that gave the best results for the measurement tasks included in the experiments. The main criterion for the selection of the most suitable method was the minimization of average deviation from nominal value. Additionally, the standard deviation was also taken into account during selection, as a secondary criterion. The result of this analysis is presented in Table 2.

**Table 2.** The best inspection path planning method to be used for basic measuring tasks.

Measuring Task	Best Inspection Path Planning Method
Flatness	Guide point (stylus orientation 45°)
Distance between planes	Guide rail
Parallelism	Guide point (stylus orientation 45°)
Angle between planes	Guide rail
Circularity	Guided curve (with offset set to 0 mm)
Circle diameter	Fixed tilt and advance (tilt 30°; advance 30°)

The second part of the experiments involved measurements of chosen dimensions and relations of the multi-feature check (MFC) gauge, and determination of the relationship of the results produced by the two methods used for inspection path control: the default option (suggested by the Modus software) and the option selected on the basis of results obtained in the first part of experiments. The multi-feature check standard allows for measurements of different geometrical features and the evaluation of the diversity of geometrical dimensions and tolerances, such as distances and angles between elements, perpendicularity, parallelism, flatness, cylindricity deviations, etc. The main element forming the standard’s shape was a cylinder with a length of 200 mm (between front planes) and an outer diameter of 100 mm. The local coordinate system of the measured object was defined with the measurement of the external cylinder (the axis of the cylinder determines z-axis of the coordinate system), with the measurement of the plane cut into the side of the cylinder (of which the normal vector defines the x-axis of the coordinate system), and with the measurement of the front plane of the cylinder, which was measured in order to define the origin of the coordinate system. Next the results were compared with the actual values of the controlled dimensions taken from the MFC’s calibration certificate. Calibration of the multi-feature check was performed on the Leitz PMM machine, which was located at LCM. The machine maximum permissible error Equation (1) was:

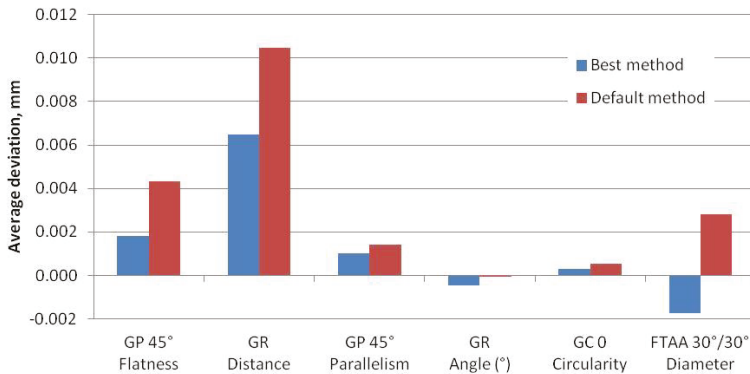
$$E_{L,MPE} = 0.8 + 2.5 \times L/1000 \mu\text{m} \tag{1}$$

where *L* is the measured length given in mm.

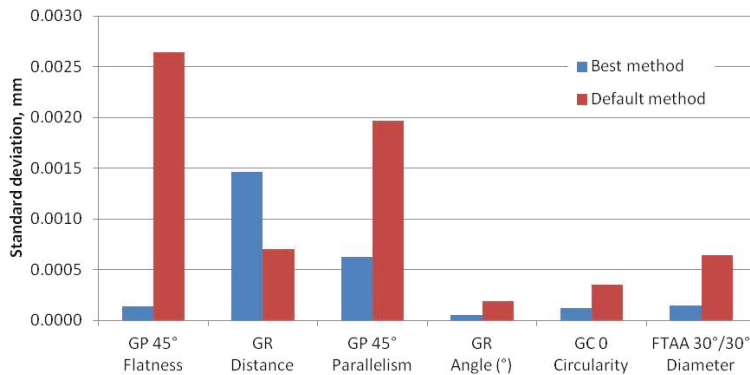
It was located in an air-conditioned room with a thermal stability of ±0.05 °C. The LCM is laboratory accredited, by the Polish Centre for Accreditation, for calibration of artifacts using the coordinate measuring technique.

The following dimensions were checked (the default method for all measured features is guide point and the best method is given in brackets after each task described): The distance between the front planes of the gauge (GR); the angle between the front and side planes (GR); the parallelism deviation between the front planes of the gauge (GP 45°); the flatness of the front plane (GP 45°); and the circularity (GC 0) and diameter of the circle (FTAA 30°/30°).

Figures 8 and 9 show the results obtained during this part of the experiment.



**Figure 8.** The average deviations for the measurements of the multi-feature check using the default and the best path control algorithms.



**Figure 9.** The standard deviations for the measurements of the multi-feature check using the default and the best path control algorithms.

#### 4. Discussion

The results of the investigations presented in this paper confirmed that it is possible to reduce the measurement errors (an error is taken as the difference between the measurement result and the reference value, represented as a result of the measured workpiece calibration) using a relevant path control algorithm for the probe head used on a five-axis coordinate measuring system. In the case of some measuring tasks (e.g., distance between planes and flatness deviation), the reduction of error is obvious and reaches up to 4 μm; however, in the case of other (e.g., parallelism deviation and circularity deviation) it is not that significant, but the tendency to reduce the measurement error is still observed if the proper path control algorithm is used. When analyzing the values of standard deviations associated to each measuring task, it may also be observed that their values are smaller for measurements performed using the optimal path control algorithm (except one case, which is the measurement of distance between planes).

It is hard to give one exact reason why there are such big differences in results obtained using different path control algorithms. As shown, the results of the previous research undertaken by the authors in [15] showed that the accuracy of the measurements performed using the probe heads used on five-axis CMMs is strongly dependent on the orientation (the A and B angles) of the probe head during point measurement. In the majority of the presented path control algorithms, with the change of the algorithm’s control parameters (e.g., tilt, advance, and offset), the probe orientation during measurement is also changing, so this may be the main cause of the differences in results. The first



analysis confirmed this assumption; however, more detailed studies are now underway in order to unequivocally prove it. What is most important is that, as was mentioned above, it is possible to choose the best path control algorithm for each measuring task and thus reduce the measurement error.

The next observation that can be made is that it is not as important to select the relevant path control algorithm for the type of measuring feature, but rather for the type of measuring task. Therefore, for some measuring tasks, which consist of measurements of the same features (e.g., distance between two planes and parallelism deviation between them—both of these tasks consist of measurement of the same planes) but also aim to determine the different relations between them, it may occur that the measurement errors for each task are smaller if both features are measured using different path control algorithms.

The considerations presented in this paper provides guidance on the selection of the best possible path control algorithm for chosen measuring tasks (Table 2). Use of them may contribute to the general reduction of measurement error for measurements performed on five-axis systems, and the improvement of awareness of its users regarding how to run the measurements on them. This, in turn, could cause a decrease in the number of faulty classified parts during the assessment of the conformity with their geometrical specifications.

The future works on this subject will involve the preparation of complete guidelines on the selection of the best path control algorithm for chosen metrological software (preferably Modus software, which is dedicated to be used on five-axis coordinate measuring systems). It will give guidance on the selection of proper algorithms during the measurement of all measuring tasks that are able to be evaluated using this software, including all possible combinations used for solving considered tasks. In future, the proposed guidelines could also be used by the companies responsible for the development of metrological software, in order to set the default measurements of the path control algorithms that will give the lowest measurement errors.

**Author Contributions:** All authors conceived and designed the experiments; A.G, P.G., and M.G. performed the experiments; M.G., A.G., and P.G. analyzed the results; and P.G., A.G., and M.G. wrote the paper.

**Funding:** This research was funded by the National Science Centre, Poland, as part of a project grant no.: 2015/17/D/ST8/01280.

**Conflicts of Interest:** The authors declare no conflicts of interest.

## References

1. Majchrowski, R.; Grzelka, M.; Wieczorowski, M.; Sadowski, L.; Gapiński, B. Large area concrete surface topography measurements using optical 3D scanner. *Metrol. Meas. Sys.* **2015**, *22*, 565–576. [[CrossRef](#)]
2. Kohut, P.; Kurc, K.; Szybicki, D.; Cioch, W.; Brudzik, R. Vision-based motion analysis and deflection measurement of a robot's crawler unit. *J. Vibroeng.* **2015**, *17*, 4112–4122.
3. Gapiński, B.; Wieczorowski, M.; Grzelka, M.; Alonso, P.A.; Tomé, A.B. The application of micro computed tomography to assess quality of parts manufactured by means of rapid prototyping. *Polimery* **2017**, *62*, 53–59. [[CrossRef](#)]
4. Kohut, P.; Holak, K.; Martowicz, A. An uncertainty propagation in developed vision based measurement system aided by numerical and experimental tests. *J. Theo. Appl. Mech.* **2012**, *50*, 1049–1061.
5. Delčev, S.; Gučević, J.; Ogrizović, V. Necessity of involvement of calibration laboratories in proficiency testing schemes. In Proceedings of the 11th IMEKO TC14 International Symposium on Measurement and Quality Control, Krakow/Kielce, Poland, 11–13 September 2013.
6. Zhao, F.; Xu, X.; Xie, S.Q. Computer-aided inspection planning—the state of the art. *Comp. Indust.* **2009**, *60*, 453–466. [[CrossRef](#)]
7. Lin, Y.-J.; Murugappan, P. New algorithm for CAD-directed CMM dimensional inspection. *Int. J. Adv. Manuf. Technol.* **2000**, *16*, 107–112. [[CrossRef](#)]
8. Hwang, C.Y.; Tsai, C.Y.; Chang, C.A. Efficient inspection planning for coordinate measuring machines. *Int. J. Adv. Manuf. Technol.* **2004**, *23*, 732–742. [[CrossRef](#)]
9. *ISO 10360-5:2010*; International Organization for Standardization (ISO): Geneva, Switzerland, 2010.

10. Woźniak, A.; Jankowski, M. New method of testing of the repeatability of CMM articulating heads. *Int. J. Adv. Manuf. Technol.* **2011**, *56*, 677–682. [[CrossRef](#)]
11. Míguez, P.C.; King, T.; Abackerli, A. A review on methods for probe performance verification. *Measurement* **1998**, *23*, 15–23. [[CrossRef](#)]
12. Ratajczyk, E.; Woźniak, A. *Coordinate Measuring Systems (In Polish: Współrzędnościowe Systemy Pomiarowe)*; Warsaw University of Technology Publishing House: Warsaw, Poland, 2016.
13. Woźniak, A.; Dobosz, M. Metrological feasibilities of CMM touch trigger probes. Part I: 3D theoretical model of probe pretravel. *Measurement* **2003**, *34*, 273–286. [[CrossRef](#)]
14. Gaśka, P.; Gaśka, A.; Gruza, M. Challenges for modeling of five-axis coordinate measuring systems. *Appl. Sci.* **2017**, *7*, 803. [[CrossRef](#)]
15. Gaśka, A.; Gaśka, P.; Gruza, M. Simulation model for correction and modeling of probe head errors in five-axis coordinate systems. *Appl. Sci.* **2016**, *5*, 144. [[CrossRef](#)]
16. *MODUSTM v1.4 SP1 User Guide*; Renishaw: Wotton-under-Edge, UK, 2011.



© 2018 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).

Article

# Demodulation Technique Based on Laser Interference for Weak Photo-Acoustic Signals on Water Surface

Xiaolin Zhang \*, Hongjie Mao and Wenyan Tang

Institute of Precision Instruments, Harbin Institute of Technology, Harbin 150001, China; xiaomaolv2hao@163.com (H.M.); tangwy@hit.edu.cn (W.T.)

\* Correspondence: zhangxiaolin@hit.edu.cn; Tel.: +86-139-3626-1257

Received: 30 October 2018; Accepted: 27 November 2018; Published: 29 November 2018



**Abstract:** To detect underwater sound-generating targets, a water surface acoustic wave laser interference and signal demodulation technique is proposed in this paper. The underlying principle of this technique involves casting a laser beam onto the water surface disturbed by an underwater acoustic source and creating interference between lights reflected by the surface and reference lights. A data acquisition and processing system was employed to obtain water surface acoustic wave information from the interference signals by means of demodulation, thus allowing detection of the underwater target. For the purpose of this study, an interference detection platform was set up in an optical dark chamber. High-frequency water surface fluctuations were introduced in the reference optical path as the phase generated carriers to create laser interference signals in two different paths, which received demodulation based on an improved arc tangent demodulation algorithm and characteristic ratio algorithm, respectively, in view of their different frequencies. Water surface wave information was then derived from such low-frequency and high-frequency signals. According to test results, in the frequency range of 200 Hz–10 kHz, the frequency detection accuracy was better than 1 Hz. The amplitude measurements exhibited high repeatability, with a standard deviation lower than 2.5 nm. The theory proposed in this paper is therefore experimentally verified with good results.

**Keywords:** optical interference; phase generated carrier; phase demodulation; water surface acoustic waves

## 1. Introduction

When an acoustic wave originating from an underwater object reaches the water/air interface, it will generate an elastic surface wave that transmits transversely along the interface in the medium's superficial layer, with a penetration depth of one wavelength approximately [1]. It has been discovered through studies that if an underwater acoustic source causes water surface vibration, only those vibrations sharing the same frequency as the acoustic source are significant in amplitude [2]. Hence, detection of the water surface acoustic waves (WSAWs) resulting from acoustic waves provides an indirect way of obtaining sound-generating information of an underwater acoustic source.

Currently, sonar remains the main technology for underwater sound field detection, but the detection coverage is limited by low-speed movement of vehicles. Researchers have therefore been highly interested in laser-based WSAW detection techniques [3,4]. One of the earliest studies with this regard was reported in the paper of Lee M. S. [5], which illustrated the mechanism of WSAWs stemming from underwater acoustic signals. In his experiments, he also obtained optical signals with the same frequency as their acoustic source. Since then, a number of techniques have been brought forward by other researchers, including those based on laser scattering, laser diffraction, luminous flux, and laser interference [6,7], all contributing to direct processing of light intensity signals. By using a laser sonar technique measuring Doppler vibrations, Antonelli L. T. et al. [8,9] successfully

detected WSAWs, and later completed single-point detection at different locations of dynamic water surface. Miao Runcai et al. [10] conducted laser interference measurements of acoustic waves on a low-frequency liquid surface and obtained modulation interference patterns. Information like surface acoustic wavelength, frequency, and amplitude was then demodulated from such interference patterns. Great efforts have been made by our project team in the research of underwater target detection through aerial laser interference and effective extraction of complicated underwater sound field signals, and so far, concrete results have been achieved under laboratory conditions (in the optical dark chamber with a high SNR). Detection of underwater acoustic sources in the frequency range of 1 kHz–18 kHz has been made possible by us with techniques such as spectrum analysis, wavelet transform analysis, and local turning point data demodulation [11–13]. However, all the above techniques fail to cover the whole acoustic wave frequency range, and are unable to provide information on variations of acoustic source signal density. In contrast, the new interference demodulation method proposed in this paper enables acquisition of underwater acoustic source vibration information and tracking of varying vibration density based on detection of WSAW. The effectiveness of the method has been proven by us through tests.

## 2. Materials and Methods

### 2.1. General Concept

An interference system is used for water surface detection, and information on water surface waves is contained in the phase of the detection signals. The disturbances have two components: natural disturbances of the water surface and WSAW. By expressing the water surface disturbances with simple harmonic vibration superimposition, we depict the interference signals in the following equation:

$$U(t) = A \cos\left\{\frac{4\pi}{\lambda} \left[\sum_i A_i \cos(\omega_i t + \theta_i) + A_s \cos(\omega_s t + \theta_s)\right] + \Omega\right\} \quad (1)$$

where  $A$  is the system gain of the interference signals;  $\lambda$  is the laser wavelength; and  $\Omega$  is the fixed phase caused by the interference optical path.  $A_i$ ,  $\omega_i$ , and  $\theta_i$  represent amplitude, angular frequency, and initial phase of the natural disturbance waves, respectively, and  $i$  represents number of harmonic waves. Similarly,  $A_s$ ,  $\omega_s$ , and  $\theta_s$  are amplitude, angular frequency, and initial phase of the WSAW, respectively. In Appendix A, the above equation receives a trigonometric function expansion and a Bessel function simplification in turn, followed by a simulation at an underwater acoustic source frequency of 1 kHz.

The resulting signal spectrum of Table 1 is shown in Figure 1. The dense spectral lines in the low-frequency zone are attributed to natural fluctuations of water surface and various low-frequency disturbance waves. Concentrated frequency bands exist around 1 kHz, i.e., the acoustic source frequency, and exhibit a symmetric distribution. Spectral lines also appear around multiples of 1 kHz such as 2 kHz and 3 kHz, although with negligible amplitude. In reality, the wide low-frequency bands imply superimposition of low frequency acoustic source signals and low-frequency natural water surface fluctuation signals, making it impossible to perform spectrum analysis of the underwater low-frequency signals directly. A new phase demodulation method is therefore required to obtain the vibration information.

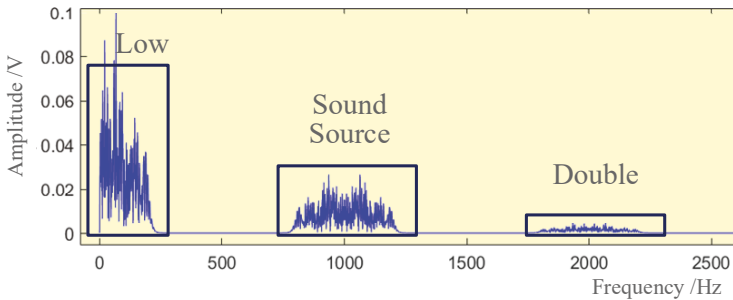
Two demodulation algorithms were used in our study for interpretation of signals acquired in the course of underwater acoustic source detection, as shown in Figure 2. The signals provided by the photodetector underwent a DC filtering process, and a spectrum analysis was then performed on the interference signals. Different demodulation schemes were adopted depending on the frequency. For underwater high-frequency acoustic source signals above 2 kHz, frequency bands independent of the low-frequency ones were created in the spectrum, with a symmetrical distribution around the underwater sound-generating source frequency. After determination of the underwater source

frequency by locating the distribution center, a characteristic ratio method was used to obtain the WSAW amplitude. For low-frequency signals below 2 kHz, given their aliasing with low-frequency bands, we were not able to derive further information from the spectrum. To address this, an improved arc tangent function algorithm was employed to demodulate the phase generated carriers (PGCs) signals, thus characterizing phase changes caused by WSAWs, and obtaining amplitude and frequency of the WSAWs.

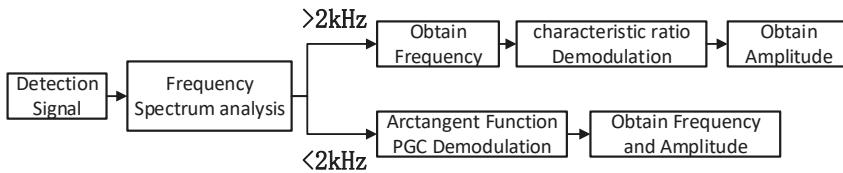
**Table 1.** Simulation parameter settings.

	Amplitude/nm	Frequency/Hz	Initial Phases/rad
Disturbances (superimposition)	1500	3	$0.8\pi$
	3000	4	$0.2\pi$
	1000	5	$1.2\pi$
WSAW	30	1000	$0.5\pi$
Laser wavelength $\lambda$ /nm			632.8
System gain of the signals $A/V$			1
The fixed system phase $\Omega$ /rad			10,000
SNR of system/dB			20

WSAW = water surface acoustic wave; SNR = signal to noise ratio.



**Figure 1.** Spectral distribution of 1 kHz simulated interference signals.



**Figure 2.** Principle diagram of demodulation algorithm for underwater acoustic source detection.

2.2. Characteristic Ratio Demodulation Algorithm

The spatial phase modulation and time phase modulation provided by water surface waves for incident laser are analogous. Spatial modulation depth can be estimated from the light intensity distribution of diffraction. For laser interference detection, we can also estimate the laser time phase modulation depth caused by water surface waves from the spectrum of the interference signals [14]. To achieve this, Bessel function was used to expand interference signals, and  $P(\omega)$  was used to represent the amplitude of the frequency of  $\omega$  in the spectrum. The characteristic ratio  $R$  was then defined as:

$$R = \sqrt{\frac{\sum_m P[(2m + 1)\omega_i]}{\sum_m P[(2m + 1)\omega_i + \omega_s]} \cdot \frac{\sum_m P(2m\omega_i)}{\sum_m P(2m\omega_i + \omega_s)}} = \frac{J_0(\frac{2\pi}{\lambda} A_s)}{J_1(\frac{2\pi}{\lambda} A_s)} \tag{2}$$

where  $(2m + 1)\omega_i$  and  $2m\omega_i$  were odd and even frequency doubling of low frequency components;  $J_0, J_1$  are Bessel functions. Since the amplitude  $A_i$  of water surface disturbance waves was significantly higher than WSAW amplitude  $A_s$  caused by the underwater acoustic source, according to the nature of Bessel function of the first kind,  $R$  can be rewritten as:

$$R = \frac{\sum_m P(m\omega_i)}{\sum_m P(m\omega_i + \omega_s)} \tag{3}$$

It can be seen from the above equation that without aliasing of the low-frequency bands with the signals under measurement, the characteristic ratio  $R$  can be determined by calculating the ratio of the sum of low-frequency band amplitudes to that of the frequency component amplitudes after a frequency shift by  $\omega_s$ . With the characteristic ratio  $R$  thus obtained, WSAW amplitude  $A_s$  can be resolved with the Bessel function. Since the analytic solution of  $A_s$  is hard to derive in practice, we estimate the numerical solution of  $A_s$  from the look-up table which is shown in Figure 3.

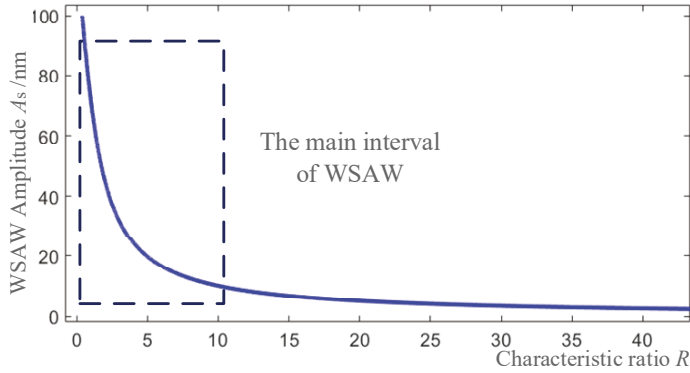


Figure 3. The function image of  $A_s(R)$ .

### 2.3. PGC-Based Improved Arc Tangent Demodulation Algorithm

When the signal frequency under test is lower than 2 kHz, PGCs are used to detect the underwater acoustic source. The test system built for this purpose is shown in Figure 4. A high-frequency underwater acoustic source was placed on the reference optical path, and the high-frequency water surface waves excited by it were introduced into the system as the carriers. This introduction of carriers can avoid the use of high-price optical phase modulators and are easy to realize; meanwhile, the interference effect is better because the reference light has a similar light intensity with the measured light. Furthermore, there are also environmental disturbances on the water surface of the reference pool, which can be reduced with measurement pool disturbances and the anti-interference ability of the system can be improved.

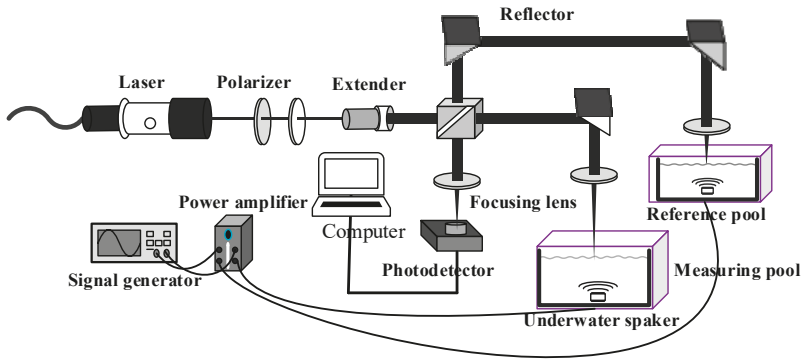


Figure 4. Test system diagram.

A high-frequency signal was added to the reference arm, its amplitude, frequency, and initial phase being represented by  $A_c$ ,  $\omega_c$ , and  $\theta_c$ , respectively.

$$\begin{cases} H = \frac{2\pi}{\lambda} A_c \\ \theta(t) = \frac{2\pi}{\lambda} [\sum_i A_i \cos(\omega_i t + \theta_i) + A_s \cos(\omega_s t + \theta_s)] \end{cases} \quad (4)$$

Assuming that the reference optical path carriers have a modulation depth of  $H$ , and that the phase caused by water surface disturbance and WSAWs under measurement is  $\theta(t)$ , we achieved the following interference signals of the PGCs:

$$U(t) = A \cos[H \cos(\omega_c t + \theta_c) + \theta(t) + \Omega] \quad (5)$$

Based on Equation (5), a demodulation was performed with the configuration given in Figure 5. This process starts with a sine/cosine mixing of the carrier frequency, followed by low-pass filtering that generated three mixing signals:

$$\begin{cases} U_{1s} = -AJ_1(H) \sin(\theta_c) \sin[\theta(t)] \\ U_{1c} = -AJ_1(H) \cos(\theta_c) \sin[\theta(t)] \\ U_{2c} = -AJ_2(H) \cos(2\theta_c) \cos[\theta(t)] \end{cases} \quad (6)$$

It can be seen from the above equation that  $U_{1s}$  and  $U_{1c}$  signals only differ in amplitude. Their amplitude ratio is equal to the tangent of the initial phase  $\theta_c$  of the carriers. Initial phase of the carriers can be derived from these two signals, which can be written as:

$$\theta_c = \arctan\left(\frac{U_{1s}}{U_{1c}}\right) \quad (7)$$

With the PGC demodulation technique, an arc tangent function demodulation method [15] was used to achieve electronic mixing, filtering, and phase division of the signals expressed by Equation (6). On this basis, the water surface fluctuations under measurement can be obtained by means of demodulation and expressed as follows:

$$\theta(t) = \arctan\left[\frac{J_2(H) \cos(2\theta_c)}{J_1(H) \cos(\theta_c)} \cdot \frac{U_{1c}}{U_{2c}}\right] \quad (8)$$

A combination of the carrier modulation depth  $H$  obtained from the spectrum and the carrier initial phase  $\theta_c$  in Equation (6) gives complete information on demodulated water surface

fluctuations, which, after elimination of low-frequency disturbances through high-pass filtering, provides WSAW information.

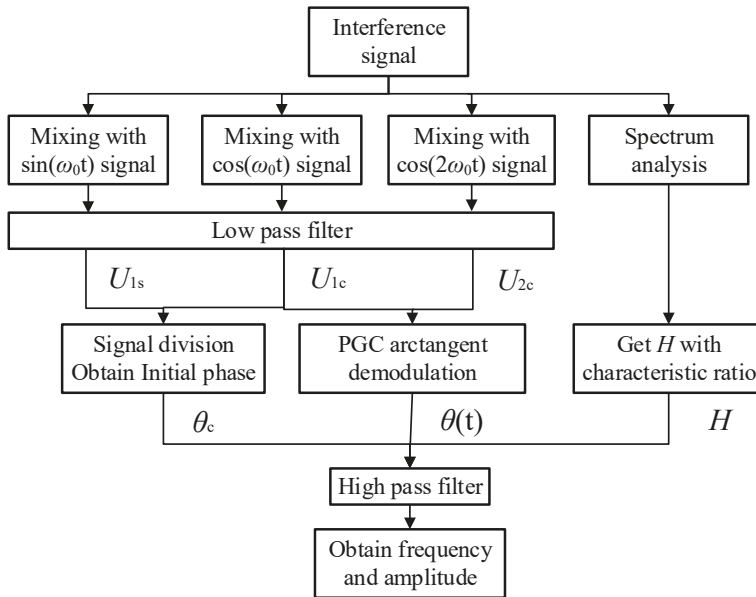


Figure 5. Phase generated carriers (PGCs) interference signal demodulation flow.

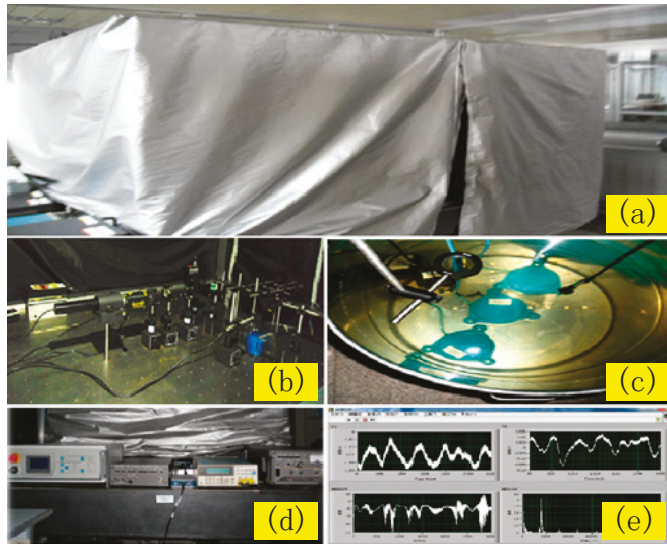
### 3. Results

In our study, a test platform was set-up with separate optical elements in an optical dark chamber. The optical dark chamber, interference optical path, and other test devices used in the study are shown in Figure 6. During the test, the signal generator created signals on two paths. The 10 kHz high-frequency signals on the first path acted on the reference optical path to generate high-frequency carriers of the same frequency, while the signals on the second path caused the speaker in the water pool to generate waves as the underwater target to be measured. The main equipment in our experiment is shown in Table 2.

Table 2. Performance index of main equipment.

Equipment	Parameter	Value
He-Ne Laser	Wavelength/nm	632.8
	Power/mW	0.6–1.4
	Power stability	±0.1%
Underwater speaker	Rated power/W	15
	Frequency response /Hz	80–20,000
Data acquisition card	Sampling rate/kS/s	100
	Resolution/bit	16

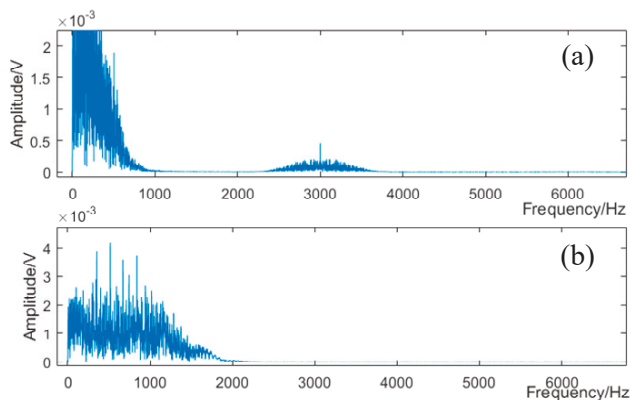




**Figure 6.** (a) Optical dark chamber; (b) interference optical path; (c) test water pool (height 100 cm, diameter 80 cm); (d) signal amplifier, power amplifier, and laser controller; (e) PC measurement screen.

### 3.1. Frequency Measurement Results

Taking the interference signals of 3 kHz and 1 kHz underwater acoustic sources as an example, the spectrograms are shown in Figure 7. It can be seen from this figure that the low-frequency band caused by the natural water surface fluctuations is about 1 kHz in width. The 1 kHz underwater acoustic source signals experienced aliasing with low-frequency bands in Figure 6b, which coincides with theoretical analysis.



**Figure 7.** Spectrum aliasing phenomenon: (a) spectrum of 3 kHz source; (b) spectrum of 1 kHz source.

Based on the previously mentioned demodulation theory, we first grouped the detected signals in the interference signal spectrum into high-frequency and low-frequency categories. Low-frequency signals are handled with improved PGC demodulation technique. The phase demodulation signals have exactly the same frequency as the signal generator. For high-frequency signals, the central points of the high-frequency components were extracted for frequency measurement, as shown in Figure 8.

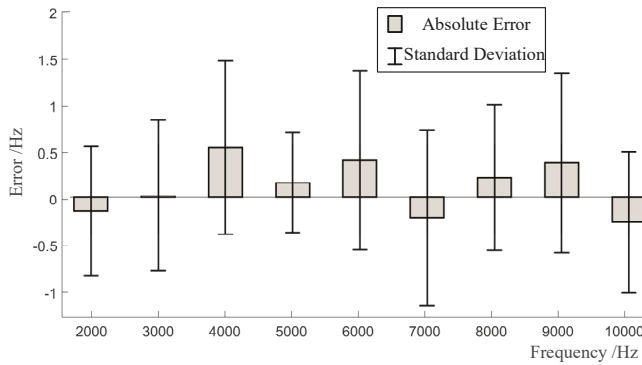


Figure 8. Errors based on extraction of central frequency points in the spectrum.

With the central frequency extraction method, detection tests were performed on underwater acoustic sources in the frequency range between 2 kHz and 10 kHz. For each frequency, the measurement was repeated five times. It can be seen from the results given in the above table that for signals above 2 kHz, the frequency measurements were generally accurate and repeatable, with a measurement error and a standard deviation not exceeding 1 Hz.

### 3.2. Amplitude Measurement Results

The spectrum of interference signals was analyzed first. The amplitude was derived directly with characteristic ratio method at high frequencies. In the case of low frequencies, high-frequency carriers were added to the reference optical path, and a signal generator was used to create 10 kHz high-frequency signals, which were fed to the reference speaker in the pool to excite carrier waves. This was followed by demodulation with improved PGC demodulation method. The test procedure covers all frequency bands between 200 Hz and 10 kHz, each corresponding to five repeated measurements. Some results are given in Table 3. For 500 Hz low-frequency signals, the modulation depth  $H$  was first estimated from the spectrum, and then the initial phase was obtained with arc tangent method, finally leading to the amplitude demodulation results. For 3-kHz high-frequency signals, after extraction of low-frequency bandwidth, the sum of spectrum amplitudes was calculated. Next a ratio of this sum to that of spectrum amplitudes after a left frequency shift by  $\omega_s$  was calculated, which is known as the characteristic ratio  $R$ . In the last step, the water surface acoustic wave amplitude was determined with the look-up table  $A_s(R)$ .

Table 3. Some amplitude measurement results.

$f/\text{Hz}$	$H (A_c = 39.28 \text{ nm})$	$\theta_c/\text{rad}$	$A_s/\text{nm}$	$\sigma/\text{nm}$	
500	0.39	0.628	48.59	1.63	
		1.309	50.03		
		0.211	51.77		
		0.785	53.42		
		1.553	51.22		
3000	Sum of Low/V	Sum of high/V	$R$	$A_s/\text{nm}$	1.31
	3.157	1.044	3.023	31.64	
	2.684	0.861	3.117	30.78	
	2.551	0.854	2.986	32.00	
	4.003	1.408	2.843	33.44	
	3.996	1.453	2.750	34.44	

The standard deviations of amplitudes measured at different frequency bands are given in Figure 9. The figure implies a lower standard deviation of the demodulated signals at high frequencies. For low-frequency signals, due to errors arising from estimation of carrier characteristic ratio and PGC demodulation, the repeatability was lower, but still within the limit of 2.5 nm. In sum, the amplitude measurements in our study provide satisfactory repeatability.

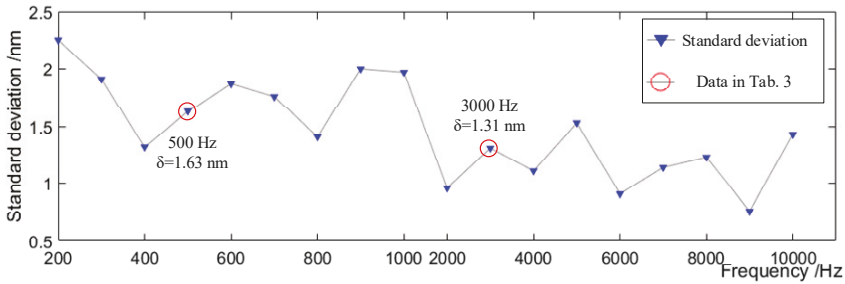


Figure 9. Standard deviation of amplitude measurement.

### 3.3. Results of Tracking Demodulation of Amplitude Modulation Signals

The signal generator creates 1 kHz signals with an amplitude modulated by 1 Hz sine waves. The modulation results at a water surface vibration sampling rate of 100 kS/s are given in Figure 8. Apparent periodical changes of the amplitude can be seen from the time domain signals after demodulation, and the correct 1 kHz demodulation result can be found in the spectrum analysis as well. The envelope of the time domain signals was extracted and fitted as a curve, resulting in the amplitude curve shown in Figure 10. The fitting curve equation is expressed as  $A_s = -26.39 \sin(2\pi \times 1.03t - 3.59) + 65.28$ . The fact that the fitting curve frequency 1.03 Hz exhibits the same as the amplitude-modulating frequency 1 Hz of the signal generator within a small tolerance proves the ability of the method used in our study to accurately track signal amplitude changes.

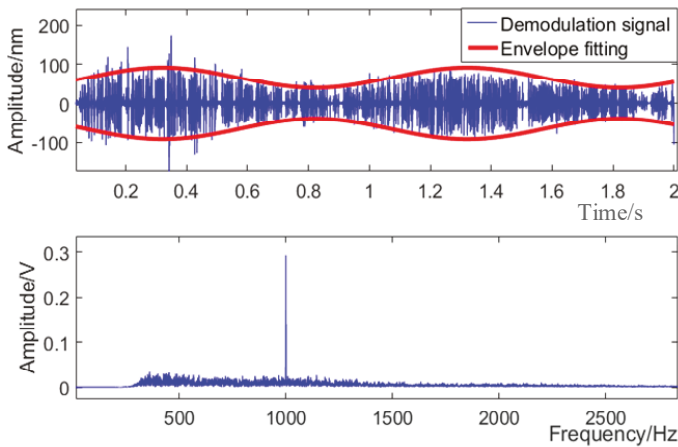


Figure 10. Demodulation result and fitting curve of amplitude modulation signal.

## 4. Discussion

To address spectrum aliasing that occurs during underwater acoustic source detection with laser interference, a demodulation method capable of measuring multiple frequency bands is proposed in this paper. Our method achieves lower frequency detection and tracking of change signals,

which cannot be realized by previous research. In our approach, the spectrum of interference signals was analyzed, and different solutions were chosen depending on frequencies: for frequencies higher than 2 kHz, a characteristic ratio method was adopted, while for those lower than 2 kHz, a PGC-based improved arc tangent function demodulation method was used. In the former, the central point of a high-frequency band was used as the measurement frequency of the detection signals, and characteristic ratio  $R$  was calculated before the amplitude was obtained in a lookup table. The PGC-based demodulation method enables demodulation with an improved arc tangent function, which gives initial phase and modulation depth of high-frequency carriers. After low-frequency filtering, amplitude and frequency of WSAW can be extracted in the optical dark chamber (SNR about 14 dB). The detection approach adopted in this paper applies to the frequency range between 200 Hz and 10 kHz while providing a frequency measurement accuracy better than 1 Hz. It also features high repeatability for amplitude measurement, with a standard deviation not exceeding 2.5 nm. As demodulated signals are generated through demodulation, time-varying signals may be measured as well. The results of amplitude modulated signal tests show that this approach is able to track signal changes.

Certain restrictions still exist in the approach adopted in this paper. The lower limit of frequency detection is dependent on filter performance and water surface disturbance waves, due to which the demodulation effect for signals below 200 Hz is not significant. On the other end of the spectrum, the upper limit of frequency detection is related to the high-frequency carriers, and should be limitless in theory. As an underwater speaker was used in our study to excite WSAWs as carriers, the frequency response attribute of the speaker also plays a role. Frequency shift devices will therefore be incorporated in subsequent research, and will hopefully enable carriers with a frequency much higher than that of acoustic waves. Furthermore, the next experiment is to consider the impact of natural light in the natural environment, perhaps considering optical filters and other devices to improve system SNR.

**Author Contributions:** Conceived the Method and Wrote the Paper, Z.-X.L.; Designed and Performed the Experiments, M.-H.J.; Edited the Manuscript, T.-W.Y.

**Funding:** This study was funded by the National Natural Science Foundation of China (61108073); Shanghai Aerospace Science and Technology Innovation Foundation (SAST2015029); Postdoctoral Researchers Settled in Heilongjiang Research Foundation (AUGA4120006016).

**Conflicts of Interest:** The authors declare no conflict of interest.

## Appendix A

The interference signal we set in the article is:

$$U(t) = A \cos\left\{\frac{4\pi}{\lambda}\left[\sum_i A_i \cos(\omega_i t + \theta_i) + A_s \cos(\omega_s t + \theta_s)\right] + \Phi\right\} \tag{A1}$$

In order to simplify the derivation, we set  $k = 2\pi/\lambda$ ,  $\theta_i = 0$ ,  $\theta_s = 0$ , therefore:

$$U(t) = A \cos\{2k[A_i \sin(\omega_i t) + A_s \sin(\omega_s t)] + \Phi\} \tag{A2}$$

We first use the trigonometric function to expand it:

$$\begin{aligned} U(t) &= A \cos \Phi \{ \cos[2kA_i \sin(\omega_i t)] \cdot \cos[2kA_s \sin(\omega_s t)] \} \\ &= -A \cos \Phi \{ \sin[2kA_i \sin(\omega_i t)] \cdot \sin[2kA_s \sin(\omega_s t)] \} \\ &= -A \sin \Phi \{ \sin[2kA_i \sin(\omega_i t)] \cdot \cos[2kA_s \sin(\omega_s t)] \} \\ &= -A \sin \Phi \{ \cos[2kA_i \sin(\omega_i t)] \cdot \sin[2kA_s \sin(\omega_s t)] \} \end{aligned} \tag{A3}$$

Combining the expansion formula of Bessel function:

$$\begin{cases} \cos[x \sin(\omega t)] = J_0(x) + 2 \sum_{k=1}^{\infty} J_{2k}(x) \cos(2k\omega t) \\ \sin[x \sin(\omega t)] = 2 \sum_{k=0}^{\infty} J_{2k+1}(x) \sin[(2k+1)\omega t] \end{cases} \quad (A4)$$

The  $J_k(x)$  is the first-class Bessel function value of  $k$  order of  $x$ :

$$J_k(x) = \sum_{m=0}^{\infty} \frac{(-1)^m}{m! \Gamma(m+k+1)} \left(\frac{x}{2}\right)^{2m+k} \quad (A5)$$

We further simplify the interference signal by using the Bessel function. We can clearly see the frequency components and their magnitude of the interference signal:

$$\begin{aligned} U(t) &= AJ_0(2kA_s)J_0(2kA_i) \cos \Phi \\ &+ 2AJ_0(2kA_s) \cos \Phi \sum_{m=1}^{\infty} J_{2m}(2kA_i) \cos(2m\omega_i t) \\ &- 2AJ_0(2kA_s) \sin \Phi \sum_{m=0}^{\infty} J_{2m+1}(2kA_i) \sin[(2m+1)\omega_i t] \\ &+ 2AJ_0(2kA_i) \cos \Phi \sum_{n=1}^{\infty} J_{2n}(2kA_s) \cos(2n\omega_s t) \\ &- 2AJ_0(2kA_i) \sin \Phi \sum_{n=0}^{\infty} J_{2n+1}(2kA_s) \sin[(2n+1)\omega_s t] \\ &+ 4A \cos \Phi \left[ \sum_{m=1}^{\infty} J_{2m}(2kA_s) \cos(2m\omega_i t) \right] \cdot \left[ \sum_{n=1}^{\infty} J_{2n}(2kA_i) \cos(2n\omega_s t) \right] \\ &- 4A \cos \Phi \left\{ \sum_{m=0}^{\infty} J_{2m+1}(2kA_i) \sin[(2m+1)\omega_i t] \right\} \cdot \left\{ \sum_{n=0}^{\infty} J_{2n+1}(2kA_s) \sin[(2n+1)\omega_s t] \right\} \\ &- 4A \sin \Phi \left[ \sum_{m=1}^{\infty} J_{2m}(2kA_i) \cos(2m\omega_i t) \right] \cdot \left\{ \sum_{n=0}^{\infty} J_{2n+1}(2kA_s) \sin[(2n+1)\omega_s t] \right\} \\ &- 4A \sin \Phi \left\{ \sum_{m=0}^{\infty} J_{2m+1}(2kA_i) \sin[(2m+1)\omega_i t] \right\} \cdot \left[ \sum_{n=1}^{\infty} J_{2n}(2kA_s) \cos(2n\omega_s t) \right] \end{aligned} \quad (A6)$$

There is no detailed explanation in this paper, but we describe the interference signal with simulation, which is consistent with the theory from the simulation result.

**References**

1. Hirao, M.; Fukuoka, H.; Hori, K. Acoustoelastic Effect of Rayleigh Surface Wave in Isotropic Material. *J. Appl. Mech.* **1981**, *48*, 119–124. [CrossRef]
2. Dai, Z.H.; Sun, J.Z.; Sui, P.F. Theoretical Study on the Water Surface Transversal Mini-wave due to the Underwater Sound Field. *J. Natl. Univ. Def. Technol.* **2004**, *26*, 95–98.
3. Blackmon, F.; Estes, L.; Fain, G. Linear optoacoustic underwater communication. *Appl. Opt.* **2005**, *44*, 3833–3845. [CrossRef] [PubMed]
4. Minami, Y. Surface tension measurement of liquid metal with inelastic light-scattering spectroscopy of a thermally excited capillary wave. *Appl. Phys. B* **2014**, *117*, 969–972. [CrossRef]
5. Lee, M.S.; Bourgeois, B.S.; Hsieh, S.T.; Martinez, A.B.; Hsu, L.; Hickman, G.D. A laser sensing scheme for detection of underwater acoustic signals. In Proceedings of the Conference Proceedings '88, IEEE Southeastcon, Knoxville, TN, USA, 10–13 April 1988; pp. 253–257.
6. Farrant, D.; Burke, J.; Dickinson, L.; Fairman, P.; Wendoloski, J. Opto-acoustic underwater remote sensing (OAUERS)—an optical sonar? In Proceedings of the OCEANS'10 IEEE SYDNEY, Sydney, NSW, Australia, 24–27 May 2010; pp. 1–7.
7. Bruno, F.; Lagudi, A.; Ritacco, G.; Muzzupappa, M.; Guida, R. Opto-Acoustic Data Fusion for Supporting the Guidance of Remotely Operated Underwater Vehicles (ROVs). *ISPRS Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **2015**, *XL-5/W5*, 47–53. [CrossRef]

8. Antonelli, L.T.; Kenneth, M.W.; Andrew, A. Laser interrogation of air-water interface for in-water sound detection initial feasibility tests. *J. Acoust. Soc. Am.* **1999**, *106*, 2298. [[CrossRef](#)]
9. Antonelli, L.T.; Blackmon, F.; Meier, L.I. Acousto-optic localization using a dynamic, Spatio-temporal array. In Proceedings of the Proceedings of OCEANS 2005 MTS/IEEE, Washington, DC, USA, 17–23 September 2005; Volume 1, pp. 1–8.
10. Miao, R.C.; Wang, Y.M.; Meng, F.; Ma, J. Optical measurement of the liquid surface wave amplitude with different intensities of underwater acoustic signal. *Opt. Commun.* **2014**, *313*, 285–289. [[CrossRef](#)]
11. Zhang, L.S.; Zhang, X.L.; Tang, W.Y. Detection of water surface capillary wave by analysis of turning-point local signal data using a laser interferometer. *Chin. Opt. Lett.* **2017**, *15*, 36–41.
12. Zhang, X.L.; Tang, W.Y.; Sun, H.Y. Frequency identification of underwater acoustic signals based on Morlet wavelet. *J. Optoelectron. Laser* **2010**, *21*, 1839–1841.
13. Zhang, L.S.; Zhang, X.L.; Tang, W.Y. Detection of water surface acoustic wave using laser interference based on wavelet ridge. *J. Optoelectron. Laser* **2015**, *26*, 103–107.
14. Zhang, L.S.; Zhang, X.L.; Tang, W.Y. Amplitude measurement of weak sinusoidal water surface acoustic wave using laser interferometer. *Chin. Opt. Lett.* **2015**, *13*, 33–37.
15. Li, S.W.; Shao, S.Y.; Mei, H.P.; Hao, Q.; Tao, R. Interference Phase Generated Carrier Demodulation Algorithm Based on Synchronous Carrier restoration Method. *Chin. J. Lasers* **2016**, *43*, 1204004.



© 2018 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).

Article

# A High Precision Capacitive Linear Displacement Sensor with Time-Grating That Provides Absolute Positioning Capability Based on a Vernier-Type Structure

Xiaokang Liu <sup>1</sup>, Hui Zhang <sup>2</sup>, Kai Peng <sup>1,\*</sup>, Qifu Tang <sup>1</sup> and Ziran Chen <sup>1</sup>

<sup>1</sup> Engineering Research Center of Mechanical Testing Technology and Equipment (Ministry of Education), Chongqing University of Technology, Chongqing 400054, China; lxx@cqut.edu.cn (X.L.); tqf@cqut.edu.cn (Q.T.); czr@cqut.edu.cn (Z.C.)

<sup>2</sup> School of Electrical and Electronic Engineering, Chongqing University of Technology, Chongqing 400054, China; ZHandLCL@126.com

\* Correspondence: pkgogo1987@163.com; Tel.: +86-023-6256-3387

Received: 31 October 2018; Accepted: 26 November 2018; Published: 28 November 2018



**Abstract:** Nanometer-scale measurement devices with high accuracy and absolute long-range positioning capability are increasingly demanded in the field of computer numerical control machining. To meet this demand, the present report proposes a capacitive absolute linear displacement sensor with time-grating that employs a vernier-type structure based on a previously proposed single-row capacitive sensing structure. The novel proposed vernier-type absolute time-grating (VATG) sensor employs two capacitor rows, each with an equivalent measurement range. The first capacitor row is designed with  $n$  periods to realize fine measurement, while the second capacitor row is designed with  $n - 1$  periods, and the phase difference between the second row and the first row is employed to obtain absolute positioning information. A prototype VATG sensor with a total measurement range of 600 mm and  $n = 150$  is fabricated using printed circuit board manufacturing technology, and its measurement performance is evaluated experimentally. Harmonic analysis demonstrates that the measurement error mainly consists of first-harmonic error, which is mostly caused by signal crosstalk. Accordingly, an optimized prototype VATG sensor is fabricated by adding a shielding layer between the two capacitor rows and designing a differential induction structure. Experimental results demonstrate that the measurement error of the optimized prototype sensor is  $\pm 1.25 \mu\text{m}$  over the full 600 mm range and  $\pm 0.25 \mu\text{m}$  over a single 4 mm period.

**Keywords:** capacitive linear displacement sensor; vernier-type absolute structure; differential sensing structure; time-grating

## 1. Introduction

Absolute position displacement measurement sensors play an irreplaceable role in the field of high-precision measurements and as a component in the closed loop feedback control of computer numerical control (CNC) machine tool systems. The rapid development of high-precision positioning technology in CNC machine tool systems has generated an increasing demand for robust absolute displacement measurement sensors with high positioning accuracy and excellent repetitive positioning precision in harsh environments, such as under the conditions of electromagnetic interference, mechanical vibration, extreme temperature variation, and dust contamination [1–7]. Meanwhile, many types of absolute grating displacement measurement sensors with high precision

and high resolution have been developed, including capacitive grating sensors [1,2], magnetic grating sensors [8], and optical grating sensors [9].

Absolute optical grating (AOG) sensors are based on an evaluation of the intensity of incident light from a known illumination source, and represent an efficient means for conducting closed loop position control in CNC machine tool systems [10–13]. The measurement error of AOG sensors has been reduced to values as small as  $\pm 0.275 \mu\text{m}$  over a 10 mm range and  $\pm 5 \mu\text{m}$  over a full 3040 mm range. However, mechanical vibration, temperature variation, and dust contamination affect the light intensity and other characteristics of the illumination source incident on the sensor, all of which detract from the measurement performance of AOG sensors. It has a weak anti-interference capability, particularly in the harsh environments associated with many CNC machining applications [13,14].

In contrast to AOG sensors, both magnetic and capacitive grating sensors have high tolerance for mechanical vibration and contamination due to dust and oil and, therefore, are commonly employed in servo drive systems [2,8,15]. The measurement error of absolute magnetic grating sensors can achieve values as low as  $\pm 5 \mu\text{m}$  over a 1000 mm range and can obtain a measurement resolution of  $0.75 \mu\text{m}$  [15–17]. However, the large size of the magnetic head and poor resistance to electromagnetic interference restrict the use of such sensors for high accuracy applications. In comparison, capacitive grating sensors provide a fuller range of environmental advantages, which make them ideally suited to harsh machining environments [2,18–21]. Moreover, the measurement error of absolute capacitive grating sensors can achieve values as low as  $\pm 0.3 \mu\text{m}$  over a 6 mm range, with an error of less than 0.01% over the full measurement range [2]. However, the measurement range of absolute capacitive grating sensors is limited, which restricts their use in practical applications [2].

The present work addresses the limitations of absolute capacitive grating sensors by developing a novel capacitive absolute linear displacement sensor that employs a vernier-type structure with time-grating based on a previously proposed single-row capacitive sensing structure [22]. The proposed vernier-type absolute time-grating (VATG) sensor employs two capacitor rows. According to the design of vernier calipers, the auxiliary scale, or vernier scale, includes  $n$  gradations over the total caliper length, while the main scale includes  $n - 1$  gradations over the same total caliper length. Similarly, one capacitor row of the VATG sensor is composed of  $n$  periods over the total sensor length and serves as a fine measurement scale to ensure high measurement resolution, which is analogous to the auxiliary scale of vernier calipers. The other row of the VATG sensor is composed of  $n - 1$  periods over the total sensor length as a coarse measurement scale used to find the phase difference of two rows to realize absolute positioning, which is analogous to the main scale of vernier calipers. A prototype VATG sensor is fabricated using printed circuit board (PCB) manufacturing technology with a capacitor row composed of 150 periods to realize fine measurement, and the second capacitor row composed of 149 periods to realize course measurement. Based on an analysis of the measurement performance of the prototype VATG sensor, the design is optimized by increasing the distance between the two rows and adding a shielding layer between them to reduce the influence of signal crosstalk, and a differential induction structure is adopted to eliminate common-mode interference, which significantly improves the measurement performance of the modified prototype VATG sensor. The remainder of this paper is organized as follows. Section 2 introduces the structures and measurement principles of the single-row capacitive sensor and the proposed VATG sensor. Section 3 presents an analysis of the limit of error of the VATG sensor. The experimental results and optimization methodology are presented and discussed in Section 4. Finally, Section 5 presents the main conclusions of the work.

## 2. Sensor Structure and Measurement Principle

### 2.1. Single-Row Sensor Measurement Principle

The structure of the single-row capacitive displacement sensor is illustrated in Figure 1a. The sensing structure consists of independent plate capacitor arrays, where the sequence of four adjacent rectangular



excitation electrodes labeled  $S+$ ,  $C+$ ,  $S-$ , and  $C-$  are fixed in stationary positions to provide a small period of  $W$ . Four orthogonal sinusoidal AC excitation signals of amplitude  $A$  and angular frequency  $\omega$ , denoted as  $U_{S+} = A_m \sin(\omega t)$ ,  $U_{C+} = A_m \cos(\omega t)$ ,  $U_{S-} = -A_m \sin(\omega t)$ , and  $U_{C-} = -A_m \cos(\omega t)$  are respectively applied to the electrodes labeled  $S+$ ,  $C+$ ,  $S-$ , and  $C-$ . Here, the length of the excitation electrodes in the  $y$  direction is  $H$  and the width is  $(W/4 - l)$ , where  $l$  is the interval between two adjacent excitation electrodes in the  $x$  direction. The stationary excitation electrodes are overlaid by an array of sinusoidal-shaped induction electrodes of length  $H$  and maximum width  $W/2$  that are movable in the  $x$  direction and have relative positions fixed with a periodicity of  $W$ , where the interval between two adjacent induction electrodes in the  $x$  direction is  $W/2$ . As illustrated in Figure 1b, the moveable induction electrodes are aligned parallel with the stationary excitation electrodes with a distance of separation  $d_0$  in the  $z$  direction. The sequence of four excitation electrodes labeled  $S+$ ,  $C+$ ,  $S-$ , and  $C-$  and one induction electrode thereby form four planar capacitors with capacitance values denoted as  $C_1$ ,  $C_3$ ,  $C_2$ , and  $C_4$ , respectively. In addition, the values of  $C_1$  and  $C_3$ , and  $C_2$  and  $C_4$  vary oppositely as the induction electrode moves along the  $x$  axis, as illustrated by the equivalent circuit shown in Figure 1d. Here,  $R$  represents load resistance. According to the superposition theorem of the equivalent circuit, the output  $U_0$  is obtained as

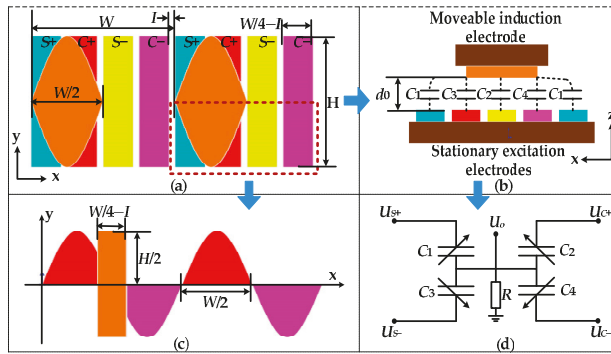
$$U_0 = U_{S+} \frac{Z_{L1}}{Z_{L1} + Z_1} + U_{S-} \frac{Z_{L3}}{Z_{L3} + Z_3} + U_{C+} \frac{Z_{L2}}{Z_{L2} + Z_2} + U_{C-} \frac{Z_{L4}}{Z_{L4} + Z_4}, \tag{1}$$

where  $Z_{Lj} = -1 / (\sum_{i=1}^4 \frac{1}{C_i} + \frac{1}{R})$  is equivalent capacitive-reactance for  $i \neq j, j = 1, 2, 3, 4$ , when  $S+$ ,  $C+$ ,  $S-$ , and  $C-$  work in dependence;  $Z_1, Z_2, Z_3$  and  $Z_4$  are the capacitive-reactance of  $C_1, C_2, C_3$ , and  $C_4$ , changing with the movement of the moveable induction electrodes owing to the variation of the value of  $C_1, C_2, C_3$ , and  $C_4$ . Thus, the final output can be given as follows [22]:

$$U_0 = \frac{U_{S+} \Delta S_{S+}}{2S_0} + \frac{U_{S-} \Delta S_{S-}}{2S_0} + \frac{U_{C+} \Delta S_{C+}}{2S_0} + \frac{U_{C-} \Delta S_{C-}}{2S_0}, \tag{2}$$

where the value  $S_0 = (HW/\pi) \cos(\pi l/W)$  represents the maximum area of overlap between an excitation electrode and an induction electrode, and  $\Delta S_{S+}, \Delta S_{C+}, \Delta S_{S-}$ , and  $\Delta S_{C-}$  are the efficient variations in the overlapping area between the  $S+, C+, S-$ , and  $C-$  excitation electrodes and a single induction electrode, respectively, when  $d_0$  and the dielectric coefficient  $\epsilon_0$  remain constant. However, Figure 1a,c illustrates the case where the sinusoidal-shaped induction electrode moving along the  $x$  axis with respect to the stationary rectangular excitation electrodes can be equivalent to a rectangular induction electrode moving along the  $x$  axis with respect to a set of four stationary half-sinusoidal-shaped excitation electrodes if the single-row sensing structure is divided into two halves along the  $x$  axis. Here, the latter case can be regarded as an uncertain limit integral of a cosine or sine function [3,21–24]. Therefore, the effective area variation  $\Delta S$  can be formulated according to the displacement  $x$  along the  $x$  direction as follows [22]:

$$\begin{aligned} \Delta S_{S+} &= \frac{WH}{\pi} (1 - \cos \frac{2\pi x}{W}), \quad \Delta S_{S-} = \frac{WH}{\pi} (1 + \cos \frac{2\pi x}{W}), \\ \Delta S_{C+} &= \frac{WH}{\pi} (1 + \sin \frac{2\pi x}{W}), \quad \Delta S_{C-} = \frac{WH}{\pi} (1 - \sin \frac{2\pi x}{W}). \end{aligned} \tag{3}$$



**Figure 1.** Schematic diagrams of a single-row capacitive displacement sensing structure and its measurement principle. Top view (a) and side view (b) schematic diagrams of the single-row sensing structure. (c) Schematic diagram presents a geometric representation of the effective area variation of the selected region in (a). (d) Equivalent capacitor circuit based on (b).

Thus, the final output is given as follows:

$$U_o = A \sin\left(\omega t - \frac{2\pi x}{W}\right), \tag{4}$$

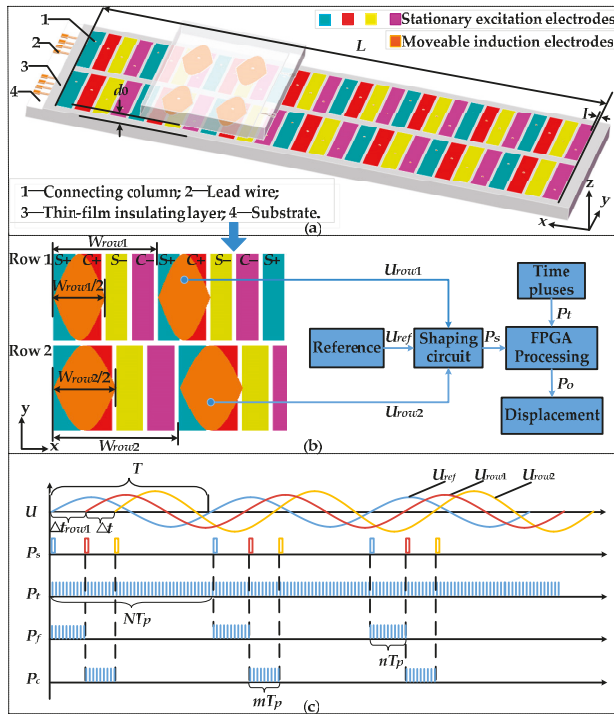
where  $A$  is the coupling coefficient, and  $U_o$  is a travelling wave signal whose frequency and amplitude are determined by the excitation signals  $U_{S+}$ ,  $U_{C+}$ ,  $U_{S-}$ , and  $U_{C-}$ , but its phase varies proportionately with respect to  $x$ , and changes with a periodicity of  $W$ . Because the travelling wave signal represents a relationship between the spatial displacement and a time standard, we express the output as follows:

$$U_o = A \sin\left(\frac{2\pi}{T}t - \frac{2\pi x}{W}\right), \tag{5}$$

where  $T$  is the time period of the travelling wave signal determined by the excitation signals. Accordingly, a single-row capacitive structure cannot obtain absolute position measurements without first knowing the starting position. However, absolute position measurements can be obtained by adopting a double-row capacitive structure.

### 2.2. Absolute Measurement Principle

The double-row capacitive structure of the VATG sensor is illustrated in Figure 2a, where stationary excitation electrodes serve as the stator, and moveable induction electrodes serve as the mover. Here, the stator consists of two-row stationary excitation electrodes, and each row occupies an equivalent overall length of  $L$ . In the same way, the mover consists of two-row corresponding inductive electrodes, and each row occupies an equivalent length of  $L/12$ . Meanwhile, a multilayer thin-film sensing structure is employed, where the electrodes and the lead wires are separated by a thin-film insulating layer, and the electrodes and lead wires are connected by a connecting column.



**Figure 2.** (a) Schematic diagram of a double-row Vernier-type capacitive structure for conducting absolute position measurements with time-grating. (b) Schematic diagram illustrating the absolute position measurement principle of double-row capacitive structure. (c) High-frequency clock pulse interpolation.

As shown in Figure 2b, the VATG sensor includes two single-row capacitive sensing structures denoted as row 1 and row 2 with different periodicities of  $W_{row1}$  and  $W_{row2}$ , respectively. Because the value of  $I$  is equivalent for both rows, the widths of the excitation electrodes in row 1 and row 2 are  $(W_{row1}/4 - I)$  and  $(W_{row2}/4 - I)$ , respectively. In addition,  $L/W_{row1}$  and  $L/W_{row2}$  are both integers, and  $L/W_{row1} - L/W_{row2} = 1$ . According to (5), the corresponding output voltage signals  $U_{row1}$  and  $U_{row2}$  of row 1 and row 2 can be given as

$$U_{row1} = A_{row1} \sin\left(\frac{2\pi}{T}t - \frac{2\pi x}{W_{row1}}\right), U_{row2} = A_{row2} \sin\left(\frac{2\pi}{T}t - \frac{2\pi x}{W_{row2}}\right), \quad (6)$$

where  $A_{row1}$  and  $A_{row2}$  are coupling coefficients of row 1 and row 2, respectively. We then define a reference signal  $U_{ref}$  with an equivalent frequency as  $U_{row1}$  and  $U_{row2}$ . Accordingly, we can define the time difference  $\Delta t_{row1}$  along the time axis as representing the phase difference between  $U_{row1}$  and  $U_{ref}$ , and the time difference  $\Delta t$  as representing the phase difference between  $U_{row1}$  and  $U_{row2}$  [25]. As shown in Figure 2b,  $U_{row1}$ ,  $U_{row2}$ , and  $U_{ref}$  are routed to a shaping circuit to convert them into square wave signals. The respective rising edges of the square wave signals, which are collectively denoted as  $P_s$  for convenience, are then routed to a phase-comparing circuit constructed using a field-programmable gate array (FPGA). High-frequency clock pulses  $P_t$  with a period  $T_p$  are also input into the phase-comparing circuit as the time standard to interpolate  $\Delta t_{row1}$  and  $\Delta t$ . The interpolation process is illustrated in Figure 2c, where the rising edge of the converted  $U_{ref}$  initiates the counting of pulses  $P_f$  until the rising edge of the converted  $U_{row1}$  is encountered. Then, pulses  $P_c$  are counted until the next rising edge

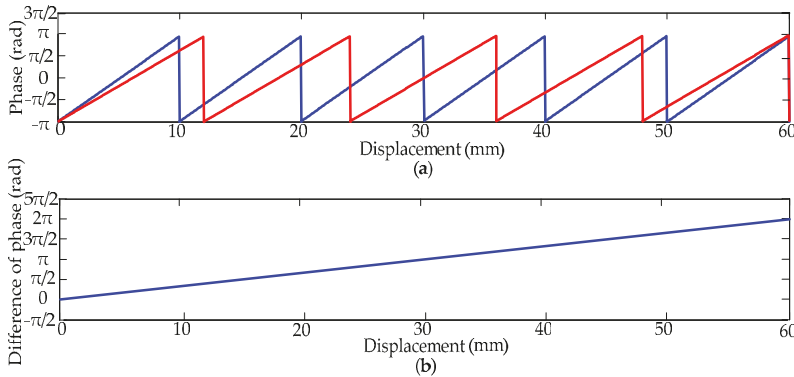
of the converted  $U_{ref}$  is encountered after a total period  $T$ . Accordingly, if  $n$  pulses  $P_f$  are counted,  $\Delta t_{row1} = nT_p$  to realize the fine measurement of row 1, and, if  $m$  pulses  $P_c$  are counted,  $\Delta t = mT_p$  to realize the coarse measurement of row 2. We also note from Figure 2c that  $T = NT_p$ , and  $mT_p$  can be employed to calculate the number of periods,  $M$ , that the induction electrode array has moved relative to the stationary excitation electrodes as follows:

$$M = \text{int}((\frac{mT_p}{NT_p}L)/W), \tag{7}$$

where the operator  $\text{int}(\cdot)$  returns the integer value of its argument. Therefore, the relationship between  $x$  and  $\Delta t$  can be defined as follows:

$$x = W(M + \frac{\Delta t}{T}) = W(\text{int}((\frac{m}{N}L)/W) + \frac{n}{N}), \tag{8}$$

Figure 3 presents the phase curves of  $U_{row1}$  and  $U_{row2}$  within their periods  $W_{row1}$  and  $W_{row2}$  with respect to displacement  $x$  obtained for  $L = 60$  mm,  $W_{row1} = 10$  mm,  $W_{row2} = 12$  mm. According to Figure 3a, the phases of the two signals are distributed in a sawtooth form according to the value of  $x$  in each period. We also note that the phase curve of  $U_{row1}$  becomes increasingly advanced relative to that of  $U_{row2}$  with increasing  $x$ . Additionally, until  $x = 600$  mm, the phase of row 2 differs by exactly a period from that of row 1. From Figure 3b, we note that the phase difference arising between the two signals with increasing  $x$  is monotonic and very nearly linear. As such, positioning can be realized according to the phase difference.



**Figure 3.** (a) Phase of outputs of row 1 ( $U_{row1}$ ) and row 2 ( $U_{row2}$ ) within their periods  $W_{row1}$  and  $W_{row2}$ , respectively, versus the displacement of the induction electrode array. (b) Phase difference with respect to the displacement of the induction electrode array.

### 3. Analysis of the Limit of Error

As discussed, row 1 and row 2 of the VATG sensor correspond to the auxiliary ruler and the main ruler of a vernier caliper, respectively. In a similar manner, we can denote the magnitude of a single scale division on row 2 of the VATG sensor (corresponding to the main ruler of a vernier caliper) as  $W_{row2}$  and the magnitude of a single scale division on row 1 (corresponding to the auxiliary ruler of a vernier caliper) as  $W_{row1}$ . Therefore, the limit of error of the VATG sensor is given as follows.

$$E_{limit} = W_{row2} - W_{row1}, \tag{9}$$

This relationship is illustrated in Figure 4. According to the measurement principle of the VATG sensor, the realization of absolute position measurement is defined completely according to  $E_{limit}$ , that is,  $E < E_{limit}$ , where  $E$  is the error in absolute position measurements.

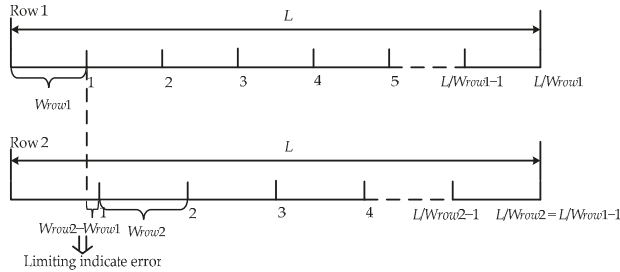
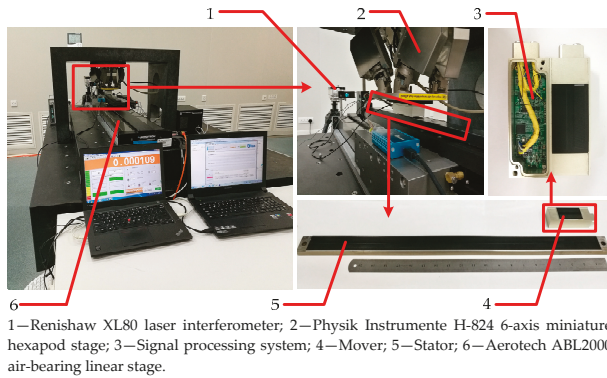


Figure 4. Schematic illustrating the determination of the limit of error for the proposed VATG sensor.

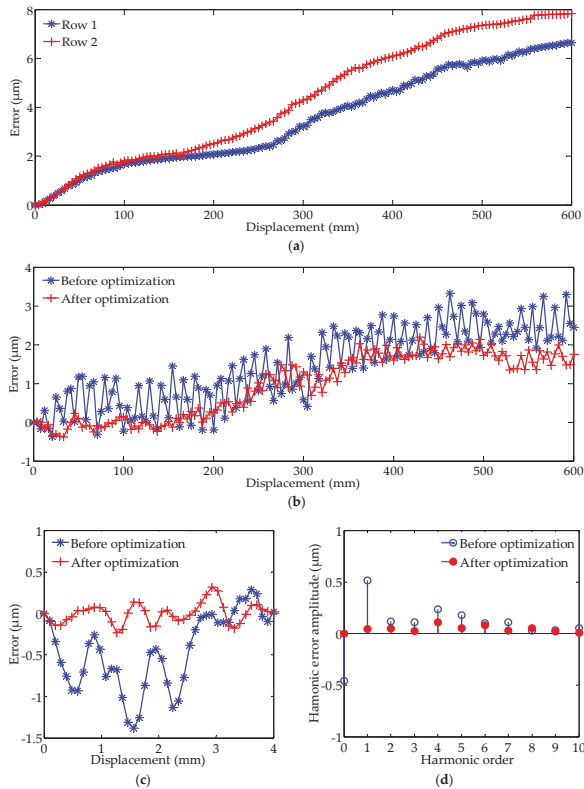
#### 4. Results and Discussion

A prototype VATG sensor like that illustrated in Figure 2a was fabricated using PCB manufacturing technology with row 1 composed of  $n = 150$  periods and row 2 composed of 149 periods, where the accuracy of PCB manufacturing technology can reach  $10 \mu\text{m}$  [23]. The parameters of the sensor were  $L = 600 \text{ mm}$ ,  $W_{row1} = L/n = 600/150 = 4 \text{ mm}$ ,  $W_{row2} = L/(n - 1) = (600/149) \approx 4.02684564 \text{ mm}$ , and  $E_{limit} = W_{row2} - W_{row1} \approx 26.845 \mu\text{m}$ . The measurement performance of the VATG sensor was evaluated experimentally. The experimental setup is shown in Figure 5. It consists of an Aerotech ABL2000 air-bearing linear stage with a positioning accuracy of  $\pm 1.5 \times 10^{-6}$  over a 120 mm traveling range, a Physik Instrumente H-824 6-axis miniature hexapod stage with a repeatability of  $\pm 0.1 \mu\text{m}$ , a Renishaw XL80 laser interferometer with a  $\pm 0.5 \text{ ppm}$  precision as a measurement standard, a signal processing system, and a sensor loaded with shells. Here, the shell of the stator is 655 mm long, 37 mm wide, and 15 mm high, while that of the mover is 96 mm long, 37 mm wide, and 26 mm high. The stator was fixed on the air-bearing linear stage. The mover was mounted on the six-axis miniature hexapod stage for adjusting the relative spatial position between the excitation and induction electrodes with sub-micrometer precision, and maintaining a gap width  $d_0 = 0.8 \text{ mm}$ . As such, the stator was actually translated during the experiments, and the mover was held stationary. The data was acquired and processed by the signal processing system, and the results were compared with the data obtained from the laser interferometer.

The displacement measurement errors of the two rows of the prototype VATG sensor individually are shown in Figure 6a as a function of the displacement at 145 points along the entire 600 mm range of the sensor. Here, the error is observed to increase linearly with increasing displacement, and the minimum error is around  $6.5 \mu\text{m}$ . This type of error is characteristic of all sensors using a repeated periodic structure and that employ an incremental measurement method. However, the double-row capacitive structure can eliminate measurement error accumulated over a small period, as indicated by the measurement error of the prototype VATG sensor over the full range shown by the curve in Figure 6b obtained before optimization. Compared with the measurement errors of the single capacitive rows in Figure 6a, the measurement error has been reduced by  $3 \mu\text{m}$  at least, and achieves a measurement error of  $3.5 \mu\text{m}$  at most over the full displacement range, which is much less than the  $E_{limit}$  value of  $26.845 \mu\text{m}$ . Therefore, the prototype VATG sensor can realize absolute displacement measurements.



**Figure 5.** Experimental setup employed for evaluating the measurement performance of the prototype VATG sensors manufactured using printed circuit board technology.



**Figure 6.** (a) Measurement errors of two rows of the initial prototype VATG sensor over the full measurement range. (b) Measurement errors of the VATG sensor obtained over the full measurement range before optimization and after optimization. (c) Measurement errors obtained over a single 4 mm period before optimization and after optimization. (d) Frequency spectra of measurement errors for a single 4 mm period before optimization and after optimization.

The measurement performance of the prototype VATG sensor was further evaluated by plotting the measurement error with respect to displacement taken at 42 points over a single 4 mm period. The results presented in Figure 6c before optimization demonstrate that the sensor exhibits a periodic measurement error of  $\pm 0.75 \mu\text{m}$  over the single period. Figure 6d presents the results of harmonic analysis based on the fast Fourier transform of the periodic measurement error data obtained before optimization. The figure indicates that the primary harmonic component includes the first-order harmonic error component with an amplitude greater than  $0.5 \mu\text{m}$  with much smaller fourth-order and fifth-order harmonic component contributions and relatively minor contributions from the remaining components. Here, the first-order harmonic error component is faintest because the experimental environment is controlled by high-precision instruments. However, physical non-uniformities in the four excitation electrode arrays, as well as non-uniformities in the electric fields of the electrodes, can introduce first-order harmonic error [24,25]. Moreover, non-uniformities in the electric fields of the electrodes can be exacerbated by signal crosstalk among the electrodes of two rows. That is to say, a certain quantity of excitation signals (briefly summarized as two types of excitation signals, sinusoidal and cosine signals) are introduced into the inductive output signal. To analyze the effect of crosstalk among electrodes on the measurement accuracy of the VATG sensor, we take row 1 as an example and consider the following mathematical model of a capacitive time-grating sensor.

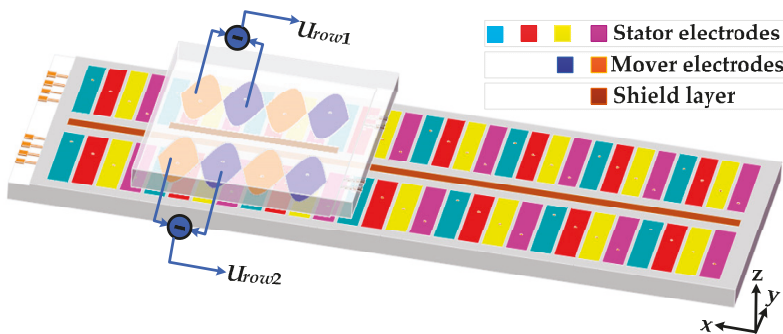
$$\begin{aligned}
 U_{row1} &= A_{row1} \sin\left(\omega t - \frac{2\pi x}{W}\right) + \delta_1 \sin(\omega t) + \delta_2 \cos(\omega t) \\
 &= A_{row1} \sin\left(\omega t - \frac{2\pi x}{W}\right) + \sqrt{\delta_1^2 + \delta_2^2} \sin\left(\omega t + \arctan \frac{\delta_2}{\delta_1}\right),
 \end{aligned}
 \tag{10}$$

where  $\delta_1$  and  $\delta_2$  are the relational coefficients of crosstalk signals caused by a certain quantity of excitation signals. Thus, signal crosstalk can result in first-order harmonic error. To address this issue, the VATG sensor design was optimized by increasing the distance between the two rows and adding a shielding layer between them to reduce the influence of signal crosstalk among the electrodes of two rows. In addition, we adopted a differential induction structure to reduce common mode disturbance. To analyze the effect of the differential induction structure for reducing common mode disturbance, the following mathematical model can be established.

$$\begin{aligned}
 U_{row1} &= \left[ A_{row1} \sin\left(\omega t - \frac{2\pi x}{W}\right) + \delta'_1 \sin(\omega t) + \delta'_2 \cos(\omega t) \right] \\
 &\quad - \left[ A_{row1} \sin\left(\omega t - \frac{2\pi x}{W} + \pi\right) + \delta'_1 \sin(\omega t) + \delta'_2 \cos(\omega t) \right] \\
 &= 2A_{row1} \sin\left(\omega t - \frac{2\pi x}{W}\right),
 \end{aligned}
 \tag{11}$$

where  $\delta'_1$  and  $\delta'_2$  are the relational coefficients of common mode disturbance caused by physical non-uniformities in the four excitation electrode arrays and by non-uniformities in the electric fields of the electrodes. These changes are illustrated in Figure 7.

A new prototype VATG sensor employing the optimized structure illustrated in Figure 7 was then fabricated, and its measurement performance was evaluated using the experimental setup shown in Figure 5. The experimental results obtained after optimization in Figure 6b–d demonstrate that the optimized sensor structure provides greatly enhanced measurement performance. Figure 6b indicates that the measurement error obtained over the full range of the optimized sensor was reduced to no more than  $2.5 \mu\text{m}$ . Moreover, Figure 6c indicates that the error obtained over a 4 mm period was reduced to no more than  $\pm 0.25 \mu\text{m}$ . Finally, Figure 6d indicates that the first-order harmonic error was greatly reduced from greater than  $0.5 \mu\text{m}$  to about  $0.05 \mu\text{m}$ .



**Figure 7.** Schematic diagram illustrating the optimized structure of the VATG sensor, which includes increasing the distance between the two rows and adding a shielding layer between them to reduce the influence of signal crosstalk. In addition, a differential induction structure is adopted to reduce common mode disturbance.

## 5. Conclusions

The present work addressed the limitations of absolute capacitive grating sensors by developing a novel VATG sensor that employs a double-row capacitive structure based on a previously proposed single-row capacitive sensing structure. One capacitor row is composed of  $n$  periods over the total sensor length and serves as a fine measurement scale analogous to the auxiliary scale of vernier calipers. The other row is composed of  $n - 1$  periods over the total sensor length as a coarse measurement scale analogous to the main scale of vernier calipers. A prototype VATG sensor with a total length of 600 mm was fabricated using PCB manufacturing technology with a fine measurement capacitor row composed of 150 periods and coarse measurement capacitor row composed of 149 periods. Based on an analysis of the measurement performance of the prototype VATG sensor, the sensor design was optimized by increasing the distance between the two rows and adding a shielding layer between them to reduce the influence of signal crosstalk, and a differential induction structure was adopted to reduce common mode disturbance, which significantly improved the measurement performance of the sensor. Experimental results demonstrated that the measurement error of the optimized prototype sensor was  $\pm 1.25 \mu\text{m}$  over the full 600 mm range,  $\pm 0.25 \mu\text{m}$  over a single 4 mm period, and the first-order harmonic error was very small.

**Author Contributions:** Conceptualization, X.L.; Writing and preparation of original draft, H.Z.; Methodology, K.P.; Validation, Q.T. and Z.C.

**Funding:** This work was funded by the National Natural Science Foundation of China (Grant No. 51435002), by the National Natural Science Foundation of China (Grant No. 51605062), by the University Innovation Team of Chongqing (Grant No. CXTDX201601029), and by the Science and Technology Research Program of Chongqing Municipal Education Commission (Grant No. KJQN201801127).

**Acknowledgments:** The authors would like to thank Engineering Research Center of Mechanical Testing Technology and Equipment (Ministry of Education), for equipment support services.

**Conflicts of Interest:** The authors have no conflicts of interest to declare.

## References

1. Kai, T.; Akiyama, M.; Nakamura, Y.; Wakui, S. Effects of Back Electromotive Force in an Absolute Displacement Sensor and Improvement Using Current-feedback-type Driver. In *Proceedings of the 2011 International Conference on Advanced Mechatronic Systems*; IEEE: Zhengzhou, China, August 2011; p. 78.
2. Bai, Y.; Lu, Y.; Hu, P.; Wang, G.; Xu, J.; Zeng, T.; Li, Z.; Zhang, Z.; Tan, J. Absolute Position Sensing Based on a Robust Differential Capacitive Sensor with a Grounded Shield Window. *Sensors* **2016**, *16*, 680. [[CrossRef](#)] [[PubMed](#)]




3. Liu, X.; Peng, K.; Chen, Z.; Pu, H.; Yu, Z. A New Capacitive Displacement Sensor with Nanometer Accuracy and Long Range. *IEEE Sens. J.* **2016**, *16*, 2306–2316.
4. Fleming, A.J. A Review of Nanometer Resolution Position Sensors: Operation and performance. *Sens. Actuators A Phys.* **2013**, *190*, 106–126. [[CrossRef](#)]
5. Lin, C.Y.; Lee, C.H. Remote Servo Tuning System for Multi-Axis CNC Machine Tools Using a Virtual Machine Tool Approach. *Appl. Sci.* **2017**, *7*, 776. [[CrossRef](#)]
6. Li, K.; Zhang, Y.; Wei, S.; Yue, H. Evolutionary Algorithm-Based Friction Feedforward Compensation for a Pneumatic Rotary Actuator Servo System. *Appl. Sci.* **2018**, *8*, 1623. [[CrossRef](#)]
7. Bahn, W.; Nam, J.H.; Lee, S.H.; Cho, D.I.D. Digital Optoelectrical Pulse Method for Vernier-Type Rotary Encoders. *IEEE Trans. Instrum. Meas.* **2016**, *65*, 431–440. [[CrossRef](#)]
8. Zhang, Z.; Ni, F.; Dong, Y.; Guo, C.; Jin, M.; Liu, H. A Novel Absolute Magnetic Rotary Sensor. *IEEE Trans. Ind. Electron.* **2015**, *62*, 4408–4419. [[CrossRef](#)]
9. Cai, N.; Xie, W.; Peng, H.; Wang, H.; Yang, Z.; Chen, X. A Novel Error Compensation Method for an Absolute Optical Encoder Based on Empirical Mode Decomposition. *Mech. Syst. Signal Process.* **2017**, *88*, 81–88. [[CrossRef](#)]
10. Girao, P.M.B.S.; Postolache, O.A.; Faria, J.A.B.; Pereira, J.M.C.D. An Overview and a Contribution to the Optical Measurement of Linear Displacement. *IEEE Sens. J.* **2001**, *1*, 322–331. [[CrossRef](#)]
11. Fleming, W.J. Overview of Automotive Sensors. *IEEE Sens. J.* **2002**, *1*, 296–308. [[CrossRef](#)]
12. Xiaokang, L.; Hongji, P.; Fangyan, Z.; Jiqin, F.; Zhicheng, Y. Research on electric field distribution and error characteristics of the nanometer time grating displacement sensor. *Chin. J. Sci. Instrum.* **2013**, *34*, 2257–2264.
13. Friedland, I.; Gurwich, I.; Brandes, A. absolute optical sensors, positioners, optical linear sensor, linear displacement, sensor volt-displacement characteristic, absolute position sensor, sensor scale factor. *Sens. Transducers J.* **2009**, *100*, 125–136.
14. Gage, S.; Evans, D.; Hodapp, M.; Sorensen, H.; Jamison, D.; Krause, B. *Optoelectronics/Fiber-Optics Applications Manual*, 2nd ed.; McGraw-Hill Book Co.: New York, UY, USA, 1981.
15. Hao, S.H.; Liu, J.Z.; Liu, Y.; Hao, M.H.; Song, B.Y. Design of a New Kind of Absolute Magnetic Grid Displacement Sensor. *High Volt. Eng.* **2009**, *35*, 2120–2125.
16. Jiang, K.D. Magnetic flux responsive gate with numerical control system. *Machinery* **2012**, *50*, 39–40.
17. Hao, S.H.; Liu, Y.; Song, B.Y.; Hao, M.H. A Novel Absolute Displacement Detecting Magnetic Sensor. *J. Harbin Inst. Technol.* **2011**, *18*, 81–85.
18. Smith, P.T.; Vallance, R.R.; Marsh, E.R. Correcting Capacitive Displacement Measurements in Metrology Applications with Cylindrical Artifacts. *Precis. Eng.* **2005**, *29*, 324–335. [[CrossRef](#)]
19. Yu, H.; Zhang, L.; Shen, M. Novel Capacitive Displacement Sensor Based on Interlocking Stator Electrodes with Sequential Commutating Excitation. *Sens. Actuators A Phys.* **2015**, *230*, 94–101. [[CrossRef](#)]
20. Ahn, H.J.; Park, J.H.; Um, C.Y.; Han, D.C. A Disk-type Capacitive Sensor for Five-dimensional Motion Measurements. *Meas. Sci. Technol.* **2008**, *19*, 045202. [[CrossRef](#)]
21. Chen, Z.; Pu, H.; Liu, X.; Peng, D.; Yu, Z. A Time-Grating Sensor for Displacement Measurement with Long Range and Nanometer Accuracy. *IEEE Trans. Instrum. Meas.* **2015**, *64*, 3105–3115. [[CrossRef](#)]
22. Yu, Z.; Peng, K.; Liu, X.; Pu, H.; Chen, Z. A New Capacitive Long-range Displacement Nanometer Sensor with Differential Sensing Structure Based on Time-Grating. *Meas. Sci. Technol.* **2018**, *29*, 054009. [[CrossRef](#)]
23. Peng, K.; Yu, Z.; Liu, X.; Chen, Z.; Pu, H. Features of Capacitive Displacement Sensing that Provide High-accuracy Measurements with Reduced Manufacturing Precision. *IEEE Trans. Ind. Electron.* **2017**, *64*, 7377–7386. [[CrossRef](#)]
24. Pu, H.; Liu, H.; Liu, X.; Peng, K.; Yu, Z. A Novel Capacitive Absolute Positioning Sensor Based on Time Grating with Nanometer Resolution. *Mech. Syst. Signal Process.* **2018**, *104*, 705–715. [[CrossRef](#)]
25. Peng, K.; Liu, X.; Chen, Z.; Yu, Z.; Pu, H. Sensing Mechanism and Error Analysis of a Capacitive Long-range Displacement Nanometer Sensor Based on Time Grating. *IEEE Sens. J.* **2017**, *17*, 1596–1607. [[CrossRef](#)]



© 2018 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).

# Metrology Data-Based Simulation of Freeform Optics

Ingo Sieber <sup>1,\*</sup> , Allen Y. Yi <sup>2</sup> and Ulrich Gengenbach <sup>1</sup>

- <sup>1</sup> Karlsruhe Institute of Technology, Institute for Automation and Applied Informatics, Hermann-von-Helmholtz-Platz 1, 76344 Eggenstein-Leopoldshafen, Germany; ulrich.gengenbach@kit.edu
- <sup>2</sup> Department of Integrated Systems Engineering, The Ohio State University, 1971 Neil Ave, Columbus, OH 43210, USA; yi.71@osu.edu
- \* Correspondence: ingo.sieber@kit.edu; Tel.: +49-721-608-25746

Received: 31 October 2018; Accepted: 19 November 2018; Published: 22 November 2018



**Abstract:** This paper describes the approach to use measurement data to enhance the simulation model for designing freeform optics. Design for manufacturing of freeform optics is still challenging, since the classical tolerancing procedures cannot be applied. In the case of spherical optics manufacturing, tolerances are more or less isotropic, and this relationship is lost in case of freeform surfaces. Hence, an accurate performance prediction of the manufactured optics cannot be made. To make the modeling approach as accurate as possible, integration of measured surface data of fabricated freeform optics in the modeling environment is proposed. This approach enables performance prediction of the real manufactured freeform surfaces as well as optimization of the manufacturing process. In our case study this approach is used on the design of an Alvarez-optics manufactured using a microinjection molding ( $\mu\text{IM}$ ) process. The parameters of the  $\mu\text{IM}$  process are optimized on the basis of simulation analysis resulting in optics, with a performance very close to the nominal design. Measurement of the freeform surfaces is conducted using a tactile surface measurement tool.

**Keywords:** systems design; simulation; form measurements; 3D measurements

## 1. Introduction

With the advancement of modern precision optical fabrication technologies, freeform optical components can be manufactured with a sufficiently high accuracy [1,2]. Under this condition, a spherical surface is not anymore the only choice in current optical design. Due to their enlarged degrees of freedom and strong possibilities of aberration correction, optical freeform surfaces are increasingly used in different fields of application, including vehicle lighting [3,4] or ophthalmic applications [5–8] as well as beam expanders [9]. From the geometrical viewpoint, an optical freeform surface has non-rotationally symmetric features, while from the aspect of fabrication and design, an optical freeform surface is regarded as an optical surface that leverages a third independent axis during the fabrication process to form the optical surface with as-designed non-symmetric features [10]. Both viewpoints reveal the basic feature of the non-rotational symmetry of optical freeform surfaces, which is a challenge for the optical community. Surface shapes of freeform optics range from weak deviations from spherical surface to surface shapes without any symmetries [10], e.g., Alvarez-optics [11] or optics using mutual rotation of two helically formed surfaces for tuning of the focal length [12–19].

Since deviations in position and shape of freeform surfaces cannot be expressed by traditional classification of rotational symmetric surfaces, the classical tolerancing procedures cannot be used in designing and manufacturing freeform optics [20], hence a performance estimation of real manufactured optical subsystems is not possible using the conventional approach of a tolerance

analysis [21]. Standardization bodies address these problems by adapting the existing standards to freeform optics [22]. In [23], the option to use a fit of the surface form deviation to determine surface form errors of real manufactured surfaces is explicitly mentioned.

Quality of optical surfaces can be evaluated and characterized using a variety of measures: Root mean square (RMS) errors as well as peak to valley (PV) errors are used to describe low-spatial frequency (LSF) errors, power-spectra density analysis can be used to analyze mid-spatial frequency (MSF) errors like spokes or ripples [24,25]. Based on these figures the manufacturing accuracy can be characterized but no conclusions can be drawn regarding the optical performance in the case where freeform optics are used in a manufactured subsystem. To be able to use optical simulation for performance prediction of a manufactured system, the optical simulation model has to be enhanced by means of measurement data of the manufactured freeform surfaces [26,27].

In addition, the performance prediction based on a model enhancement by means of metrology data can be used to optimize manufacturing parameters. In particular, this might be useful to improve the performance of freeform optics designed-for-manufacture by low-cost mass fabrication methods such as replication technologies (like  $\mu\text{IM}$ ) [28]. As a result, mass fabrication methods can compete with highly precise and more expensive single-part manufacturing processes, such as ultraprecision diamond machining.

We propose the use of measured surface data of fabricated freeform optics in the modeling environment to make the modeling approach as accurate as possible. Due to the non-symmetric shape of optical freeform surfaces, the manufacturing tolerances are usually not isotropic as it is assumed for spherical optics. For this reason, the analyzation of the measured surface data of the real, manufactured optics is of great interest. This approach allows predicting the optical performance of the real manufactured freeform surfaces.

In this paper, a method is demonstrated to use metrology data to enhance the model to be able to predict the optical performance of the manufactured system. The article is organized as follows: Section 2 describes the manufacture of the freeform optics by  $\mu\text{IM}$ . In Section 3, analysis and model enhancement by means of metrology data from the manufactured freeform optics are demonstrated. The paper closes with conclusions in Section 4.

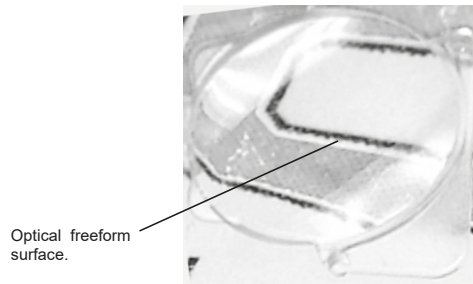
## 2. Manufacture of the Freeform Optics

The manufacture of the freeform optics was conducted using a  $\mu\text{IM}$ -replication process. The initial Alvarez-surface is described by Equation (1).

$$z(x, y) = 0.01958x^2y + 0.00629y^3, \quad (1)$$

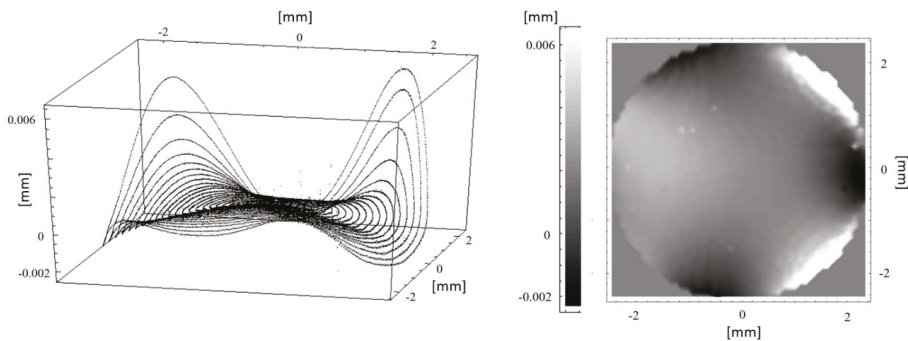
$\mu\text{IM}$  has the potential to fabricate optical components with low cycle times, usually ranging from seconds to minutes. In the first step, a molding tool, i.e., a negative form of the component, must be produced. In mold design, shrinkage compensation and placement of runners and vents have to be considered. Manufacturing parts in small dimensions frequently leads to a freezing of the melt due to the disadvantageous surface to volume ratio, thus representing a particular challenge in  $\mu\text{IM}$ . Simulating the mold flow can prevent this problem: information about the impact of different process parameters can be gathered and used to optimize the molding process.

A set of injection molding simulation experiments was carried out to form a basis to design the mold and to derive the process parameters [29]. On the basis of these simulation experiments, molding tools were fabricated by which the freeform optics was molded. Figure 1 depicts a photograph of a freeform lens manufactured by the  $\mu\text{IM}$  process.



**Figure 1.** Freeform lens manufactured by the  $\mu$ IM process.

An atomic scale profiler was used to measure the manufactured surfaces. The profiler used was a Panasonic UA3P with a resolution of  $\pm 0.1 \mu\text{m}$  at an inclination angle of  $70^\circ$ . The measurement range was adjusted to  $\varnothing 5 \text{ mm}$ . Measurements were conducted on concentric circles, the spatial frequency was set to 48 cyc/aperture, resulting in a radial distance of 0.1 mm. One measurement consists of 50,000 measurement points. The measurement was targeted to detect the surface form deviation error, which is an LSF error. To analyze MSF errors linear scanning of the sample would be the preferred procedure. One advantage of conducting the measurement on radial paths is avoidance of a device oscillation. Figure 2 depicts the residuals in lens sag based on a reduced data set on a  $48 \times 48$  grid (2304 measurement points), plotted along the measurement path (left) and as a contour plot (right). High residues are found at the coordinates near  $\{2 \text{ mm}, 2 \text{ mm}\}$  and  $\{2 \text{ mm}, -2 \text{ mm}\}$ . Otherwise, the surface looks notably smooth. Some non-uniformly distributed pits can be observed, interpreted as artifacts caused by the tactile measuring process.



**Figure 2.** Surface residuals with regard to the nominal surface: plotted on the concentric measuring path (left), contour plot (right).

### 3. Analysis and Results

Analysis of the metrology data results in an averaged RMS value of  $0.85 \mu\text{m}$  and an averaged PV value of  $8.1 \mu\text{m}$ . Using RMS and PV residuals to evaluate the optical surface quality, it is obvious, that the smaller an RMS value is, the better the nominal geometry of the surface is resembled. The same holds for PV: the higher the PV value the larger the maximum deviation between measured and nominal surface. However, neither of the two values describe the optical performance due to geometrical deviations. It is not possible to derive how this surface deviation will affect the system's optical performance. The absolute PV value of  $8.1 \mu\text{m}$  seems to be very small for a 5 mm clear aperture. The same is true for the RMS error below  $1 \mu\text{m}$ . In the event of spherical optics, no one would expect the optical performance to decrease significantly due to these small values. However, using freeform optics an estimation of the impact of geometrical deviations becomes less intuitive.

To analyze the effect of geometric deviations due to the manufacturing processes, the simulation model has to be enhanced by measured surface data. This approach enables simulation and performance analysis of the real manufactured geometry. The surface fit on the measuring points was conducted by means of a polynomial equation, since the nominal surface is also described by a polynomial. The use of higher order polynomial terms is due to the complexity of freeform optics. Surface generation using fourth order polynomials achieved better results than an approximation of the surface based on spline interpolation [20]. Optical surfaces were often fitted by means of Zernike polynomials. Zernike polynomials offer the advantage to describe directly low frequency errors and hence, can be used as direct measure of optical aberrations. Since the individual Alvarez lens part has no meaningful optical effect, optical aberrations cannot be assigned to the individual surfaces directly. Hence, the advantage of using Zernike polynomials does not exist. Alvarez optics generate its optical effect only in conjunction with a second element. Determination of form deviations of freeform surfaces by means of a surface fit is conform to the modified ISO-10110 standard [23].

Table 1 depicts the functions generated by means of the surface fit, the RMS error between the manufactured surface and the fitted function, and the mean error of the surface fit. The mean error as well as the RMS error are small enough to allow a representation of the manufactured optics by the fitted surface function given in Table 1. The simulation model now is enhanced by means of the fitted surface function. A performance simulation based on the enhanced model allows for a comparison of the optical performance with the originally designed nominal freeform surfaces. To quantify the optical performance, the modulation transfer function (MTF) was applied as criterion. The MTF is an accepted criterion to quantify the optical imaging quality of lenses or lens systems [29].

**Table 1.** Fitted function and its parameters. RMS and mean error describe the fit accuracy with respect to the measured data.

Surface Functionz			
$(x,y) = a_1 x^4 + a_2 x^3 y + a_3 x^2 y^2 + a_4 xy^3 + a_5 y^4 + a_6 x^3 + a_7 x^2 y + a_8 xy^2 + a_9 y^3 + a_{10} x^2 + a_{11} xy + a_{12} y^2 + a_{13} x + a_{14} y$			
Injection-Molded Surface	Parameter	RMS [ $\mu\text{m}$ ]	Mean Error [ $1 \times 10^{-16} \mu\text{m}$ ]
	$a_1 = 1.139 \times 10^{-4}$ $a_5 = -8.416 \times 10^{-5}$ $a_7 = 0.01832$ $a_8 = -9.148 \times 10^{-5}$ $a_9 = 6.636 \times 10^{-3}$	0.105	-1.28
	$a_{10} = 1.335 \times 10^{-4}$ $a_{11} = 2.688 \times 10^{-4}$ $a_{12} = 1.184 \times 10^{-4}$ $a_{14} = 2.915 \times 10^{-3}$		

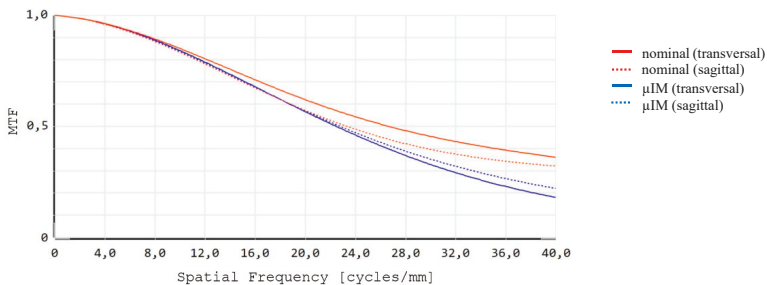
The performance of the manufactured Alvarez-optics was tested in an existing setup in context of the development of the Artificial Accommodation System [8]. The AAS features a modular concept, where housing as well as optical surfaces integrated in the housing is used for different varifocal principles. Originally, the housing is designed for an optical subsystem using a lens moveable along the optical axis to adjust the focal length. Hence, the rear-side optics is a negative aspheric lens. Combining this rear-side optics with an Alvarez-optics, an aspherical inset lens has to be designed to add a positive refraction power to the negative lens of the rear side. The optical performance of the assembly containing the Alvarez-optics suffers from the adaptation of the existing setup. This strategy leads to a degradation of the imaging quality, which can be observed by relatively poor values of the MTF. Nevertheless, this setup is suitable to evaluate our approach to enhance the optical simulation model by means of measurement surface data. The MTF of the central field was calculated using the nominal surface of the Alvarez-optics and the fitted function corresponding to the  $\mu\text{IM}$  process. Five different settings of the varifocal optics were used in the calculation corresponding to different states of refraction power, respectively. Table 2 depicts the MTF of the fitted functions with respect to the nominal values and the central field. The optical subsystem used suffers from a dependency of the imaging quality (depicted by the MTF value) on the adjustment of focal length. This effect can be observed in the nominal design as well as for the manufactured optics. The injection-molded optics shows a non-acceptable decline in performance in particular at the infinite focus length.

**Table 2.** Modulation transfer function (MTF) of the manufactured optics compared with the nominal design (central field, evaluated at 40 cyc/mm).

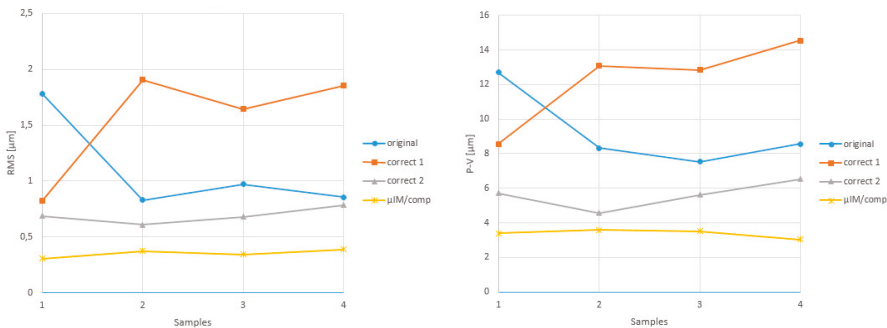
Focal Length [m]	MTF Nominal	MTF Manufactured	Relative MTF [%]
$\infty$	0.34	0.2	59
4	0.49	0.43	88
0.667	0.66	0.64	97
0.364	0.63	0.54	86
0.263	0.55	0.43	78

Figure 3 shows the graph of the MTF. Again, the central field is displayed in an adjustment of an infinite focal length. The splitting of the curves in two parts is due to the non-rotationally symmetric shape of the freeform optics. The transversal part is shown as solid line; the sagittal part is represented as dashed line. The red curves describe the MTF of the nominal design, the blue curves represent the MTF of the simulation based on the measurement data of the manufactured Alvarez-optics. The reduced MTF of the manufactured optics is caused by geometrical part deformation occurring in the  $\mu\text{IM}$  process. An uneven refractive index distribution may also contribute to the poor MTF. An optimization of the  $\mu\text{IM}$  process was conducted to compensate for these two limiting factors. Optimization of the molding process lead to different settings of the process parameters and is described in detail in [28]. Figure 4 presents RMS and PV values of intermediate states of the optimization for four samples, respectively. The initial setting is depicted as “original”. Several corrections led eventually to the final process settings ( $\mu\text{IM}/\text{comp}$ ) featuring a smooth characteristic over all four samples. Based on the  $\mu\text{IM}/\text{comp}$ -process, modified Alvarez-optics were manufactured.

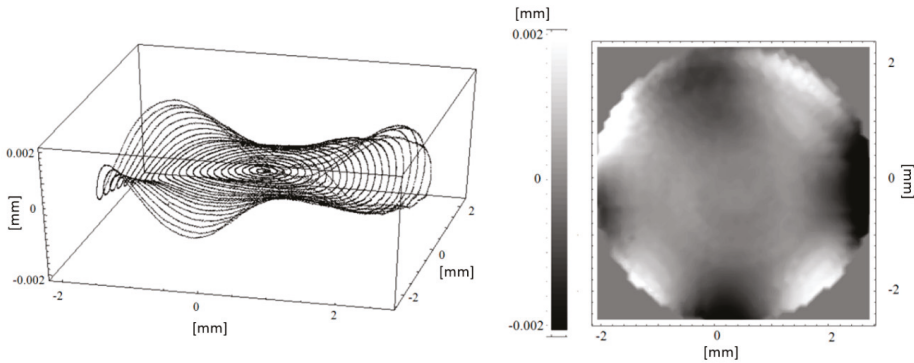
Figure 5 shows the surface residuals for the compensated Alvarez-lens parts manufactured by means of the  $\mu\text{IM}/\text{comp}$ -process.



**Figure 3.** MTF of the manufactured optics compared with the nominal design.

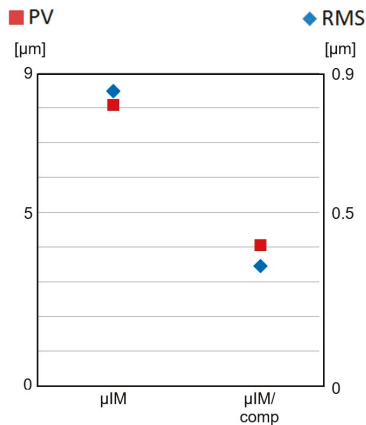


**Figure 4.** RMS and PV-values of intermediate states of the process optimization for four samples, respectively. “original” means the initial setting.



**Figure 5.** Surface residuals of the freeform optics manufactured using the  $\mu\text{IM}/\text{comp}$ -process (data set on a reduced  $48 \times 48$  grid). **Left:** plotted on the concentric measuring path; **right:** Contour plot.

Noticeable is the smooth surface. The improvement of the surface form deviation can be seen clearly by comparing Figures 2 and 5. Of course, this aspect can also be seen regarding the PV and RMS errors, resulting in  $4.1 \mu\text{m}$  and  $0.34 \mu\text{m}$ , respectively. Figure 6 shows the PV and RMS value of the Alvarez-surface manufactured by the two microinjection molding processes. The improvement of the compensated surface manufactured by means of the  $\mu\text{IM}/\text{comp}$ -process is obvious.



**Figure 6.** PV and RMS values of the individual optics manufactured. Left:  $\mu\text{IM}$ ; right:  $\mu\text{IM}/\text{comp}$ .

To see if the optical performance of the freeform optics manufactured by the  $\mu\text{IM}/\text{comp}$ -process meets the criterion to preserve the MTF of the nominal Alvarez-optics, the measured surfaces were used to simulate the performance of the compensated Alvarez-optics. The approach used is the same as described in case of the  $\mu\text{IM}$  process. Table 3 depicts the function used in the enhanced model.

Table 4 shows the resulting, simulated MTF based on the surface representation shown in Table 3 compared again with the nominal surface.

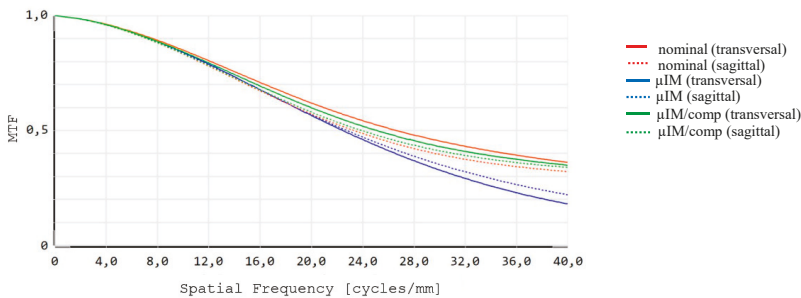
**Table 3.** Function of the fitted point clouds as well as the RMS and mean errors showing the accuracy of the fit.

Surface Functionz				
$(x, y) = a_1 x^4 + a_2 x^3 y + a_3 x^2 y^2 + a_4 xy^3 + a_5 y^4 + a_6 x^3 + a_7 x^2 y + a_8 xy^2 + a_9 y^3 + a_{10} x^2 + a_{11} xy + a_{12} y^2 + a_{13} x + a_{14} y$				
Surface	Parameter		RMS [ $\mu\text{m}$ ]	Mean Error [ $1 \times 10^{-16} \mu\text{m}$ ]
$\mu\text{IM}/\text{comp}$	$a_1 = -3.182 \times 10^{-6}$ $a_2 = -3.180 \times 10^{-5}$ $a_3 = -1.482 \times 10^{-4}$ $a_5 = 5.137 \times 10^{-5}$ $a_6 = -1.911 \times 10^{-5}$ $a_7 = 0.0195$	$a_8 = 6.796 \times 10^{-5}$ $a_9 = 6.396 \times 10^{-3}$ $a_{10} = 1.335 \times 10^{-4}$ $a_{11} = 1.235 \times 10^{-4}$ $a_{12} = 8.923 \times 10^{-5}$ $a_{13} = -1.675 \times 10^{-5}$ $a_{14} = 1.235 \times 10^{-4}$	0.147	-0.85

**Table 4.** MTF of the manufactured optics compared with the nominal design (central field, evaluated at 40 cyc/mm).

Focal Length [m]	MTF Nominal	MTF $\mu\text{IM}/\text{Comp}$	Relative MTF $\mu\text{IM}/\text{Comp}$ [%]
$\infty$	0.34	0.34	100
4	0.49	0.47	98
0.667	0.66	0.65	98
0.364	0.63	0.63	101
0.263	0.55	0.55	100

The results presented in Table 4 show that the optics manufactured using the  $\mu\text{IM}/\text{comp}$ -process performs very closely to the nominal value. The deviations in performance are as small as 2% in maximum. The effect of improvement of the MTF by means of the  $\mu\text{IM}/\text{comp}$ -process can also be observed in Figure 7 showing the graphs of the simulated MTFs for the different molding processes ( $\mu\text{IM}$  and  $\mu\text{IM}/\text{comp}$ ) in comparison with the nominal optics. The graphs are depicted for the central field and an allocation of infinite focal length.



**Figure 7.** MTF derived with the compensated injection-molded freeform optics included (green) in comparison with Figure 3.

It is of no doubt that the values of RMS and PV are very useful to describe surface form deviations and are of importance in the characterization of manufacturing quality. This research was designed to show that PV and RMS errors alone are not sufficient to enable statements of the optical performance of freeform optics. In the example shown even surface accuracies of  $8.1 \mu\text{m}$  PV and  $0.85 \mu\text{m}$  RMS do not meet the performance requirements for a clear aperture of 5.0 mm. This was shown in case of the uncompensated microinjection-molded freeform optics (see Figure 3). Using metrology data to enhance the model offers the possibility to predict reasonably the expected optical performance of a manufactured freeform optics.



#### 4. Conclusions

Freeform optics can be seen as an enabling technology since they allow completely new designs of optical systems. Classical optical components have a rotational symmetric shape. The tolerances are therefore describable by simple parameters, e.g., radius. Since freeform optics are lacking symmetry, such simple tolerance parameters cannot be found. This makes the estimation of shape tolerances nearly impossible as well as the estimation of the impact of manufacturing tolerances on the optical performance. Doubtless, the use of PV and RMS data is very important to characterize and analyze the manufactured optics with respect to quality and accuracy of the manufacturing processes. However, a performance prediction of a fabricated optical system using freeform optics renders to be impossible if only based on PV and RMS data. For performance prediction of the manufactured freeform optics in an optical system, we propose the use of measurement data of the manufactured parts to enhance the optical model. The enhanced model will better resemble the real system hence simulation results can also be used to improve process optimization.

The approach was shown using Alvarez-optics as freeform surfaces. As manufacturing process, a  $\mu\text{IM}$  process was used. It turned out, that freeform surfaces with seemingly insignificant PV and RMS values could cause severe effects on the optical performance. This was shown for a freeform optics with a 5 mm clear aperture by way of simulation of an enhanced model. Optimization of the  $\mu\text{IM}$  process leads to improved optics with a maximum deviation of only 2% from the nominal MTF values. These results show that enhancing the simulation model by means of metrology data will allow for performance prediction of the optical system. This approach can also be used to evaluate and optimize process parameters to improve the parts manufactured.

**Author Contributions:** I.S. developed the conceptualization of the design approach as well as the underlying methodology. Furthermore, he performed the optical simulation and formal analysis as well as the validation. He also wrote the original draft. A.Y.Y. conceived and designed the advanced microinjection molding process, did investigations with respect to the manufacturing parameters, manufactured the Alvarez-optics, and acquired funding. U.G. conceived the system integration of the optics, acquired funding, administrated the project, and contributed the respective text passages as well as reviewing and editing of the complete draft.

**Funding:** German Federal Ministry of Education and Research (BMBF) (16SV5472K).

**Acknowledgments:** The work presented in this paper was conducted under the KD OptiMi2 programme (grant no. 16SV5472K), subtopic B "Fabrication of housing shells & measurement data" funded by the German Federal Ministry of Education and Research. The authors would also like to acknowledge Likai Li from the Ohio State University for his support in manufacturing of the freeform optics and Erik Beckert and Ralf Steinkopf from Fraunhofer-Institut für Angewandte Optik und Feinmechanik (IOF) for surface measurement. Furthermore, we acknowledge support by Deutsche Forschungsgemeinschaft and Open Access Publishing Fund of Karlsruhe Institute of Technology.

**Conflicts of Interest:** The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

#### References

1. Fang, F.Z.; Zhang, X.D.; Weckenmann, A.; Zhang, G.X.; Evans, C. Manufacturing and measurement of freeform optics. *CIRP Ann.–Manuf. Technol.* **2013**, *62*, 823–846. [CrossRef]
2. Owen, J.D.; Davies, M.A.; Schmidt, D.; Urruti, E.H. On the ultra-precision diamond machining of chalcogenide glass. *CIRP Ann.–Manuf. Technol.* **2015**, *64*, 113–116. [CrossRef]
3. Chen, F.; Wang, K.; Qin, Z.; Wu, D.; Luo, X.; Liu, S. Design method of high-efficient LED headlamp lens. *Opt. Express* **2010**, *18*, 20926–20938. [CrossRef] [PubMed]
4. Hsieh, C.; Li, Y.; Hung, C. Modular design of the LED vehicle projector headlamp system. *Appl. Opt.* **2013**, *52*, 5221–5229. [CrossRef] [PubMed]
5. Focus on Vision. Available online: <http://www.focus-on-vision.org/en/adjustable-glasses> (accessed on 26 September 2018).
6. Barbero, S.; Rubinstein, J. Power-adjustable spherocylindrical refractor comprising two lenses. *Opt. Eng.* **2013**, *52*, 063002. [CrossRef]


7. Simonov, A.N.; Vdovin, G.; Rombach, M.C. Cubic optical elements for an accommodative intraocular lens. *Opt. Express* **2006**, *14*, 7757–7775. [[CrossRef](#)] [[PubMed](#)]
8. Sieber, I.; Martin, T.; Gengenbach, U. Robust Design of an Optical Micromachine for an Ophthalmic Application. *Micromachines* **2016**, *7*, 85. [[CrossRef](#)] [[PubMed](#)]
9. Duerr, F.; Thienpont, H. Analytic design of a zoom XY-beam expander with freeform optical surfaces. *Opt. Express* **2015**, *23*, 30438–30447. [[CrossRef](#)] [[PubMed](#)]
10. Ye, J.; Chen, L.; Li, X.; Yuan, Q.; Gao, Z. Review of optical freeform surface representation technique and its application. *Opt. Eng.* **2017**, *56*, 110901. [[CrossRef](#)]
11. Alvarez, L.W.; Humphrey, W.E. Variable-Power Lens and System. U.S. Patent 3,507,5651, 21 April 1970.
12. Bernet, S.; Ritsch-Martel, M. Adjustable refractive power from diffractive moiré elements. *Appl. Opt.* **2008**, *47*, 3722–3730. [[CrossRef](#)] [[PubMed](#)]
13. Bernet, S.; Harm, W.; Ritsch-Martel, M. Demonstration of focus-tunable diffractive Moiré-lenses. *Opt. Express* **2013**, *21*, 6955–6966. [[CrossRef](#)] [[PubMed](#)]
14. Harm, W.; Bernet, S.; Ritsch-Martel, M.; Harder, I.; Lindlein, N. Adjustable diffractive spiral phase plate. *Opt. Express* **2015**, *23*, 413–421. [[CrossRef](#)] [[PubMed](#)]
15. Busch, S.F.; Balzer, J.C.; Bastian, G.; Town, G.E.; Koch, M. Extending the Alvarez-Lens Concept to Arbitrary Optical Devices: Tunable Gratings, Lenses, and Spiral Phase Plates. *IEEE Trans. Terahertz Sci. Technol.* **2017**, *7*, 320–325. [[CrossRef](#)]
16. Grewe, A.; Fesser, P.; Sinzinger, S. Diffractive array optics tuned by rotation. *Appl. Opt.* **2017**, *56*, A89–A96. [[CrossRef](#)]
17. Sieber, I.; Martin, T.; Stiller, P. Tunable refraction power by mutual rotation of helical lens parts. *Proc. SPIE* **2017**, *10375*, 103750L.
18. Sieber, I.; Stiller, P.; Moser, D.; Gengenbach, U. Design-for-manufacture of a varifocal rotation optics. *Proc. SPIE* **2018**, *10690*, 106901B. [[CrossRef](#)]
19. Sieber, I.; Stiller, P.; Gengenbach, U. Design studies of varifocal rotation optics. *Opt. Eng.* **2018**.
20. Sieber, I.; Li, L.; Gengenbach, U.; Beckert, E.; Steinkopf, R.; Yi, A. Optical performance simulation of free-form optics for an eye implant based on a measurement data enhanced model. *Appl. Opt.* **2016**, *55*, 6671–6679. [[CrossRef](#)] [[PubMed](#)]
21. Sieber, I. Comprehensive modeling and simulation of micro-optical subsystems. *Proc. SPIE* **2012**, *8550*, 855002.
22. ISO 10110, Optics and Photonics—Preparation of Drawings for Optical Elements and Systems. Available online: [www.iso.org](http://www.iso.org) (accessed on 26 September 2018).
23. Youngworth, R.N.; Kiontke, S.R.; Aikens, D.M. Implementing ISO standard-compliant freeform component drawings. *Opt. Eng.* **2016**, *55*, 071205.
24. Filhaber, J. Artificial Large optics: Mid-spatial-frequency errors: The hidden culprit of poor optical performance. *LaserFocusWorld*, 8 August 2013.
25. Youngworth, R.N.; Stone, B.D. Simple estimates for the effects of mid-spatial-frequency surface errors on image quality. *Appl. Opt.* **2000**, *39*, 2198–2209. [[CrossRef](#)] [[PubMed](#)]
26. Sieber, I.; Rübenaich, O. Integration of measurement data in the comprehensive modelling approach. *Proc. SPIE* **2013**, *8884*, 88841F.
27. Li, L.; Raasch, T.W.; Sieber, I.; Beckert, E.; Steinkopf, R.; Gengenbach, U.; Yi, A.Y. Fabrication of Microinjection Molded Miniature Freeform Alvarez Lenses. *Appl. Opt.* **2014**, *53*, 4248–4255. [[CrossRef](#)] [[PubMed](#)]
28. Li, L.; Sieber, I.; Beckert, E.; Steinkopf, R.; Yi, A.Y. Design and fabrication of microinjection molded miniature freeform alvarez lenses. In Proceedings of the ASPE/ASPEN Summer Topical Meeting, Kohala Coast, HI, USA, 26–27 June 2014; pp. 13–16.
29. Smith, G.H. *Practical Computer-Aided Lens Design*; Willmann-Bell: Richmond, VA, USA, 1998; pp. 157–170, ISBN 0-943396-57-3.



© 2018 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).

Article

# Experimental Analyses on Multiscale Structural and Mechanical Properties of $\epsilon$ -Si/GeSi/C-Si Materials

Wei Qiu <sup>1,2,\*</sup>, Lu-Lu Ma <sup>1</sup>, Hong-Tao Wang <sup>3</sup>, Ren-Rong Liang <sup>4</sup>, Yu-Cheng Zhao <sup>1</sup> and Yun-Shen Zhou <sup>2,\*</sup>

<sup>1</sup> Tianjin Key Laboratory of Modern Engineering Mechanics, Department of Mechanics, Tianjin University, Tianjin 300072, China; ll\_ma@tju.edu.cn (L.-L.M.); daniel\_l\_q@hotmail.com (Y.-C.Z.)

<sup>2</sup> Department of Electrical and Computer Engineering, University of Nebraska-Lincoln, Lincoln, NE 68588-0511, USA

<sup>3</sup> Institute of Applied Mechanics, Zhejiang University, Hangzhou 310027, China; htw@zju.edu.cn

<sup>4</sup> Institute of Microelectronics, Tsinghua University, Beijing 100084, China; liangrr@tsinghua.edu.cn

\* Correspondence: qiuwei@tju.edu.cn (W.Q.); yunshen.zhou5@gmail.com (Y.-S.Z.);  
Tel.: +86-139-2046-2608 (W.Q.)

Received: 30 October 2018; Accepted: 19 November 2018; Published: 22 November 2018



**Abstract:** Strained silicon ( $\epsilon$ -Si) is a promising material that could extend Moore's law by enhancing electron mobility. A  $\epsilon$ -Si material is usually composed of multiscale, multilayer heterostructures, where the strained-silicon film or strap is tens-of-nanometers thick, and its buffer layers are of the micrometer scale. The structural properties determine the electrical performance and reliability of  $\epsilon$ -Si-based devices. Inhomogeneous residual stress is induced during the preparation, which induces  $\epsilon$ -Si structure failure. In this work, biaxial strained-silicon films that contain graded and relaxed germanium-silicon buffer layers were prepared on monocrystalline silicon wafers through reduced-pressure chemical-vapor epitaxy. The layer components and thicknesses were measured using energy-dispersive spectroscopy and scanning-electron microscopy. Crystal and lattice characters were observed by using high-resolution transmission-electron microscopy and micro-Raman spectroscopy. The residual stress distribution along cross-sections of the  $\epsilon$ -Si multilayer structures was examined by using micro-Raman mapping. The experimental results showed that, with a gradual increase in germanium concentration, the increasing residual stress was suppressed owing to dislocation networks and dislocation loops inside the buffer layers, which favored the practical application.

**Keywords:** strained silicon ( $\epsilon$ -Si); multiscale; structural property; mechanical property; micro-Raman spectroscopy; cross-section; dislocation

## 1. Introduction

Strained silicon ( $\epsilon$ -Si) is regarded as a promising material that could extend Moore's law [1]. Because introduced strain leads to an increased carrier mobility and reduced resistance [2], the performance of  $\epsilon$ -Si-based devices is improved significantly [3]. In recent years, with continuous developments in  $\epsilon$ -Si technology, semiconductor chips have become smaller, faster, and more energy efficient [4]. Strained-silicon technology has been applied widely in core chips inside mobile phones, personal computers, laptops, tablets and even televisions [5].

The  $\epsilon$ -Si structure complexity determines the difficulty of the manufacturing process. In general,  $\epsilon$ -Si is achieved by manufacturing a set of multiscale, multilayer heterogeneous two- or three-dimensional structures [6]. The manufacture of such a complex structure requires a strict and well-designed control of process parameters to ensure that each processing procedure step

may yield a material/structure that meets the design requirements of geometrical dimensions and structural-mechanical properties [7,8]. Furthermore, the design, manufacture, performance evaluation and control technologies for microelectronics with non-uniform and local-singular stress distribution is regarded as the future of the microelectronics industry [9].

The (equal biaxial or uniaxial) strain state and its magnitude determine the optical/electrical properties of  $\epsilon$ -Si-based devices [10]. The strain state and magnitude of the  $\epsilon$ -Si depend heavily on the buffer-layer properties (such as the element composition, thickness, lattice quality, defects and residual stress) and interfacial properties (such as mismatch and shear stress) [11]. Size miniaturization of a  $\epsilon$ -Si-based device leads to increased defect densities and residual stress, which results in a degraded optoelectronic performance and structure reliability of the entire integrated device [12].

Experimental studies of multilayer heterogeneous structures, such as strained-silicon, face two major challenges, namely, “multiscale” and “multiproperty coupling”. A strained-silicon layer is usually several to tens-of-nanometers thick, whereas the buffer-layer thickness tends to be of the micrometer scale, and the substrate is typically several hundreds of microns thick. Unique key physical and mechanical behaviors that interact exist at each spatial scale. Cross-scale interactions between the multiscale physical-mechanical properties increase the complexity of experimental studies of strain silicon.

A number of methods at different spatial scales have been applied to experimental studies on multilayer heterogeneous structures [13–16]. The profiler has been used to measure the average residual stress from the mesoscale [17]. By using (scanning- and transmission-) electron microscopy, the structural morphology and dislocation distribution in the microscale have been determined [18]. At the same scale, nano-indentation has been used to measure the modulus [19], whereas Raman microscopy has been applied to characterize the lattice quality and residual stress [20–22]. High-resolution transmission-electron microscopy (TEM) has been used to analyze the lattice structure and local strain of materials at the nanoscale [23].

However, most widely used experimental methods can measure or characterize only one or several material parameters at their respectively applicable scales. These techniques do not provide sufficient understanding of the multiscale structure properties, such as strained silicon, and it is difficult to discover the interaction between the different structural and mechanical properties at different spatial scales. Therefore, there is a need for collaborative analysis using methods at different spatial scales to investigate the interaction mechanisms of multiproperty coupling.

In this work, a biaxial strained-silicon material with germanium-silicon as a buffer layer was prepared and studied experimentally. The structural properties, including the components and thickness of each layer at the microscale and the crystal and lattice characteristics and quality at the nanoscale, was characterized by energy-dispersive spectroscopy (EDS), scanning-electron microscopy (SEM), micro-Raman spectroscopy (MRS) and TEM. The residual stress distribution at different depths was analyzed by using micro-Raman mapping on the sample cross-section.

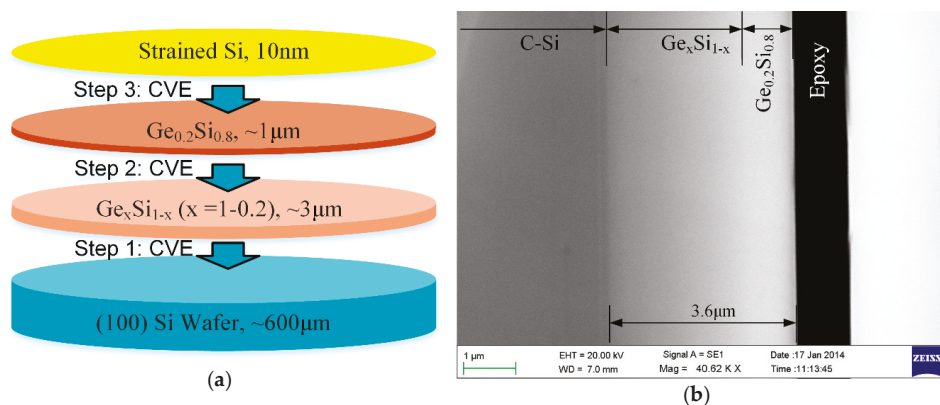
## 2. Materials and Methods

The  $\epsilon$ -Si/GeSi/C-Si materials were prepared through reduced pressure chemical vapor epitaxy (CVE) [24].  $\text{SiH}_2\text{Cl}_2$  was used as the Si precursor,  $\text{GeCl}_4$  was the Ge precursor, and  $\text{H}_2$  was the carrier gas. The monocrystalline Si (100) wafer was pre-treated using an  $\text{H}_2\text{SO}_4:\text{H}_2\text{O}_2 = 3:1$  solution for 10 min to remove contaminants and then dipped in 10% HF solution to remove the oxide layer. The CVE growth temperature was 900 °C, and the pressure was  $1.33 \times 10^4$  Pa. In the beginning, the Ge component increased linearly to form a graded germanium-silicon ( $\text{Ge}_x\text{Si}_{1-x}$ ,  $x = 0 \rightarrow 0.2$ ) buffer layer of  $\sim 3$   $\mu\text{m}$  thickness. The growth rate slowed and the Ge concentration was fixed at 20% to form a  $\text{Ge}_{0.2}\text{Si}_{0.8}$  buffer layer of  $\sim 0.8$   $\mu\text{m}$  thickness. A 10 nm strained-silicon cap layer was deposited at 750 °C and  $1.33 \times 10^4$  Pa. Figure 1a provides a schematic structure of the prepared  $\epsilon$ -Si/GeSi/C-Si material.

The cross-sectional samples were prepared as follows. A Si-based  $\epsilon$ -Si/GeSi wafer was cut to approximately 10 mm  $\times$  10 mm square pieces. Two wafers were pasted by using epoxy resin with

the  $\epsilon$ -Si films face-to-face. Then, two C-Si pieces, cut from another (100) silicon wafer with a 430  $\mu\text{m}$  thickness were pasted outside the two  $\epsilon$ -Si/GeSi wafers, respectively. Then, this piece group was cut to some thin slices about a  $\sim 1$  mm thickness. The two cross-section surfaces of each thin slice were polished by using the chemical-mechanical polishing technique (CMP). Samples for TEM analysis were thinned further by using sandpaper (3M, Maplewood, Minnesota, USA), lapping film (Allied High Tech Products, Rancho Dominguez, CA) and a precise ion-polishing machine (Gatan, Pleasanton, CA, USA).

A ZEISS EVO MA15 system and a JEOL 2100 system were used for the SEM and TEM tests, respectively. Raman experiments for the cross-sectional samples were conducted by using a laser confocal micro-Raman system (InVia Reflex, Renishaw, Gloucestershire, UK) with a 532 nm laser exciting light. A  $50\times$  Leica objective (N.A. 0.8) was used, and hence, the spot size of the incident laser was about 1  $\mu\text{m}$  in diameter. Besides, Raman tests on the untreated surface of the strained-silicon samples used the same system, but with a 325-nm laser and a  $40\times$  Leica ultraviolet (UV) objective (N.A. 0.5). The sampling spot of the incident laser was about 5  $\mu\text{m}$  in diameter.



**Figure 1.** (a) Schematic structure of prepared  $\epsilon$ -Si/GeSi/C-Si material, (b) scanning-electron microscopy (SEM) image of multilayer vicinity of sample cross-section. CVE, chemical vapor epitaxy.

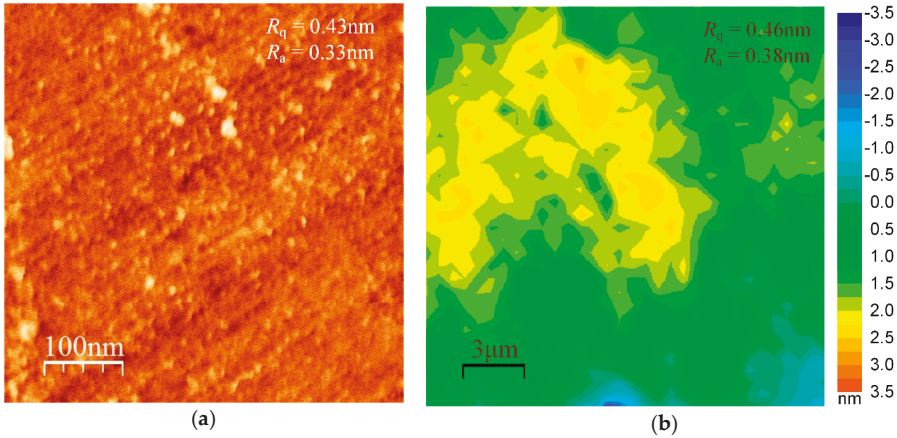
### 3. Results and Discussion

#### 3.1. Microscale Structural Properties

Figure 1b shows an SEM image of the multilayer vicinity of the cross-sectional sample. Because of differences in the electrical resistance of each layer, distinctive monocrystalline silicon,  $\text{Ge}_x\text{Si}_{1-x}$ ,  $\text{Ge}_{0.2}\text{Si}_{0.8}$  and epoxy resin layers were observed. A narrow black strip of  $\sim 1$   $\mu\text{m}$  thickness is the adhesive layer of epoxy resin that adheres the two multilayer wafers. The region with uniform and relatively deeper greyscale, at the left of the image and far from the adhesive layer, is the C-Si substrate. From the C-Si substrate to the wafer edge ( $\sim 3.6$   $\mu\text{m}$  thickness), the grayscale becomes brighter gradually, which is the graded and relaxed buffer layers. The C-Si/ $\text{Ge}_x\text{Si}_{1-x}$  interface is clear and straight, but that of the  $\text{Ge}_x\text{Si}_{1-x}/\text{Ge}_{0.2}\text{Si}_{0.8}$  interface is barely distinguishable visible in this SEM image. The strained-silicon layer ( $\sim 10$  nm thick) is not recognized in the SEM image, because of the resolution limitation of the SEM system.

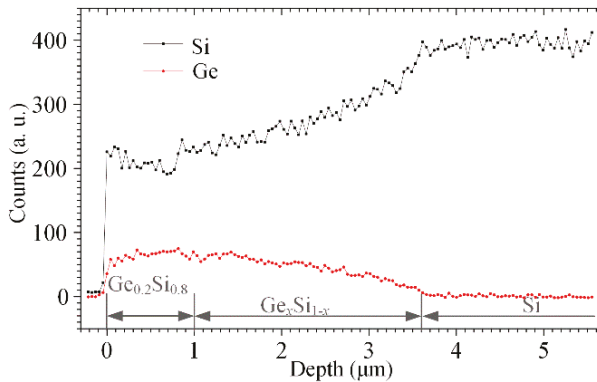
The SEM image also shows that there is no visible fluctuation on the surface of the sample cross-section. The cross-sectional surface roughness was measured by using an atomic-force microscope (AFM, Dimension Icon, Bruker, Billerica, Massachusetts, USA) and a white-light interferometer (WLI, NewView 8300, Zygo, Middlefield, Connecticut, USA). Figure 2a gives the height data of a random 500 nm  $\times$  500 nm area on the cross-sectional surface detected by AFM. The root-mean-square deviation of the height data ( $R_q$ ) is 0.43 nm, and the arithmetical mean deviation of the height data ( $R_a$ ) is 0.33 nm.

Figure 2b gives the height data of a random  $21 \mu\text{m} \times 21 \mu\text{m}$  area on the cross-sectional surface detected by using a white-light interferometer with a spatial resolution of  $0.34 \mu\text{m}$  in the X/Y direction and  $0.01 \text{ nm}$  in the Z direction (root-mean-square repeatability).  $R_q$  and  $R_a$  are essential surface-roughness parameters. The experimental results showed that the surface roughness of the sample cross-section was less than  $0.5 \text{ nm}$ , regardless of the nano- or microscale area.



**Figure 2.** The surface roughness of cross-sectional surface measured by (a) atomic-force microscope (AFM) and (b) white-light interferometer.

The elemental distribution along the depth direction of the sample cross-section is shown in Figure 3 as determined by EDS (Oxford Instruments, Oxford, UK) with a  $42\text{-nm}$  step length. When the EDS sampling point on the cross-section surface approaches the wafer edge, the elemental silicon content exhibits an approximately linear decrease and that of elemental germanium shows a linear increase. Such linear variations of the elemental components that are sustained inside a region of  $\sim 2.6 \mu\text{m}$  in thickness, corresponding to the graded buffer layer (via the  $\text{Ge}_x\text{Si}_{1-x}$  layer).



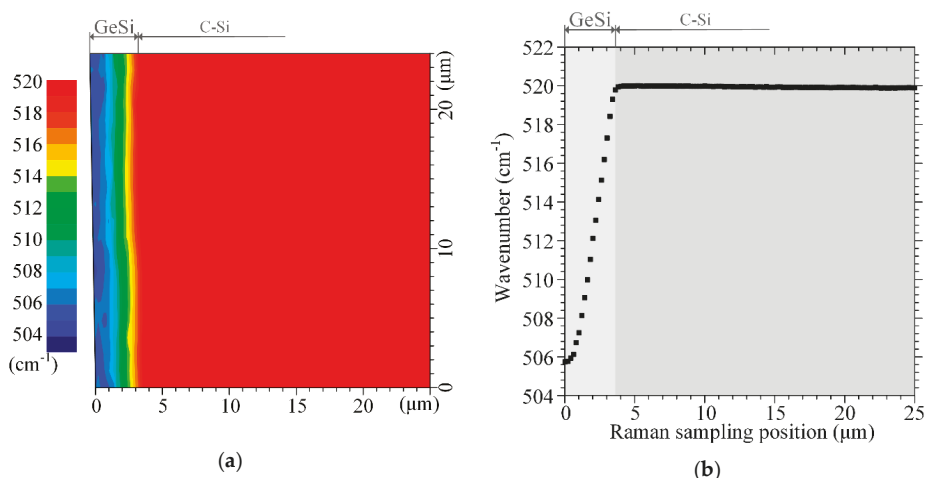
**Figure 3.** Elemental distribution along the depth direction achieved by using energy-dispersive spectroscopy (EDS).

The content of elemental silicon increased slightly and gradually until the sampling point reached the wafer edge. The opposite behavior resulted in the content of elemental germanium. The thickness of this region was  $\sim 1 \mu\text{m}$ , which corresponds to the relaxed buffer layer (viz. the  $\text{Ge}_{0.2}\text{Si}_{0.8}$  layer). In this layer, the actual distributions of elemental Si and Ge did not agree with, but were not distinct

from, the design shown in Figure 1. The EDS result showed that it was difficult to recognize the interface between the two buffer layers, and the exact thickness of the two buffer layers. The sudden drop of the Si component at a 1  $\mu\text{m}$  depth may have been caused by parameter alternation during the ultra-high-vacuum chemical-vapor epitaxy, which helps to locate the  $\text{Ge}_x\text{Si}_{1-x}/\text{Ge}_{0.2}\text{Si}_{0.8}$  interface.

Both elements decreased to zero because the EDS sampling point entered the epoxy resin layer of the sample cross-section. The EDS data did not show the signal from the  $\epsilon$ -Si layer, which should be because the  $\epsilon$ -Si layer thickness ( $\sim 10$  nm) was less than the EDS sampling step length.

Figure 4a shows the wavenumber image of the Si-Si band ( $\sim 520$   $\text{cm}^{-1}$ ) that was obtained by Raman mapping at a  $25 \mu\text{m} \times 24 \mu\text{m}$  region near the multilayer interfaces on the sample cross-section. The step lengths for Raman mapping along the parallel and vertical direction to the interfaces were  $1.2 \mu\text{m}$  and  $0.2 \mu\text{m}$ , respectively. Figure 4b shows the mean values of the wavenumbers at different locations along the depth direction based on Figure 4a. Figure 4a,b shows that the wavenumber at  $\sim 520$   $\text{cm}^{-1}$  changes slightly inside the Si substrate, acutely in the buffer layers, and continuously at the interfaces. When the scanning spot originated from the Si substrate into the buffer layers, the wavenumber showed a sharp and linear decrease with a speed of  $-4.97 \pm 0.07$   $\text{cm}^{-1}/\mu\text{m}$ . However, this rapidly decreasing trend stopped within a region of  $\sim 1 \mu\text{m}$  next to the epoxy layer (compared with the top surface of the strained-silicon wafer), where the wavenumber decreased gently with a speed of  $-0.83 \pm 0.28$   $\text{cm}^{-1}/\mu\text{m}$ . Hence, from the wavenumber image, the thickness of the rapidly decreased wavenumber region is  $\sim 2.6$ – $2.8 \mu\text{m}$ , and that of the gently decreased wavenumber region is  $\sim 0.8$ – $1 \mu\text{m}$ . The former region is the  $\text{Ge}_x\text{Si}_{1-x}$  layer and the latter is the  $\text{Ge}_{0.2}\text{Si}_{0.8}$  layer. The strain distribution can be achieved by using the wavenumber results, as discussed below.

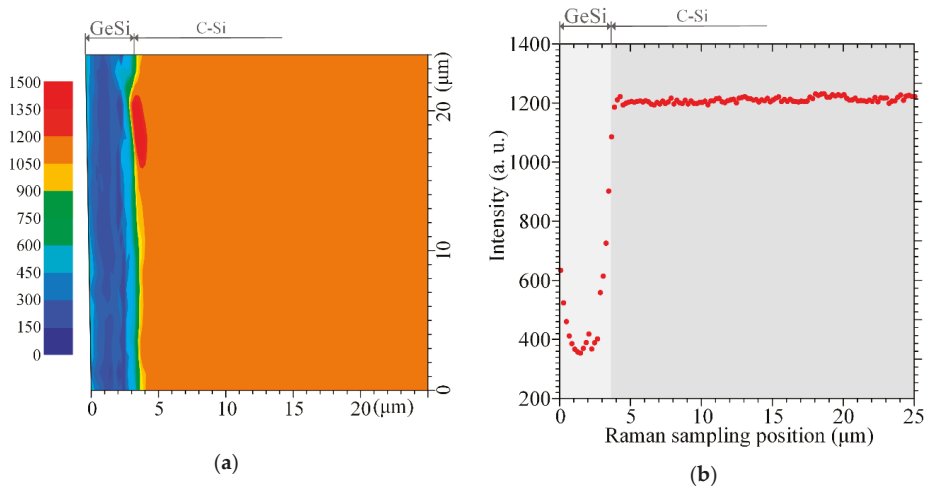


**Figure 4.** (a) Si-Si band wavenumber image obtained by Raman mapping on the sample cross-section, (b) wavenumber distribution along the depth direction.

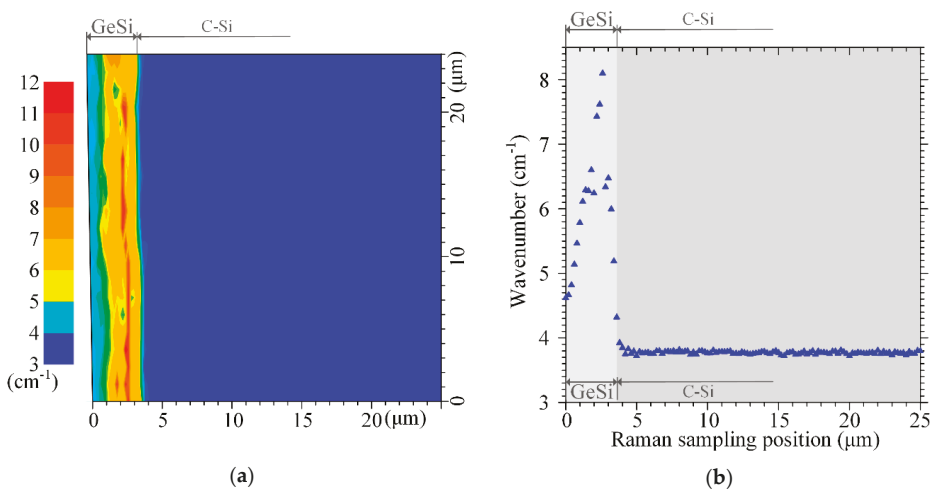
Figure 5a shows the Raman intensity image of the same region in Figure 4a. Figure 5b shows the mean values at different locations along the depth direction based on Figure 5a. In Figure 5, when the sampling spot for Raman mapping from the Si substrate crossed the buffer layer, the Raman intensity of the Si-Si band decreased first and then increased. The Raman-intensity-decreased region in Figure 5 is almost the same as that of the wavenumber rapid-decreased region in Figure 4, so the Raman intensity-increased region to the wavenumber gently-decreased region. This result indicates that the quantity of elemental Si in the monocrystalline stage decreased with a continuous enhancement of elemental Ge in the graded-buffer ( $\text{Ge}_x\text{Si}_{1-x}$ ) layer, and then improved gradually

inside the relaxed-buffer ( $\text{Ge}_{0.2}\text{Si}_{0.8}$ ) layer even though the elemental content of the materials did not change any more.

Figure 6a shows the full-width-at-half-maximum (FWHM) image of the same region of Figure 4a. Figure 6b gives the mean values at different locations along the depth direction based on Figure 6a. In Figure 6, the FWHM of the Si-Si band broadened first and then narrowed when the sampling spot of Raman mapping from the Si substrate crossed the buffer layer. The FWHM broadened region in Figure 6 is different from (smaller than) the wavenumber rapid-decreased region in Figure 4 and the Raman-intensity-decreased region in Figure 5. Therefore, with a continuous enhancement of elemental Ge, the average lattice quality of the Si element in the monocrystalline stage in the graded-buffer ( $\text{Ge}_x\text{Si}_{1-x}$ ) layer decreased to an extreme and then improved continuously. This tendency was retained in the relaxed-buffer ( $\text{Ge}_{0.2}\text{Si}_{0.8}$ ) layer.



**Figure 5.** (a) Si-Si band Raman-intensity image obtained by Raman mapping on the sample cross-section, (b) Raman-intensity distribution along the depth direction.

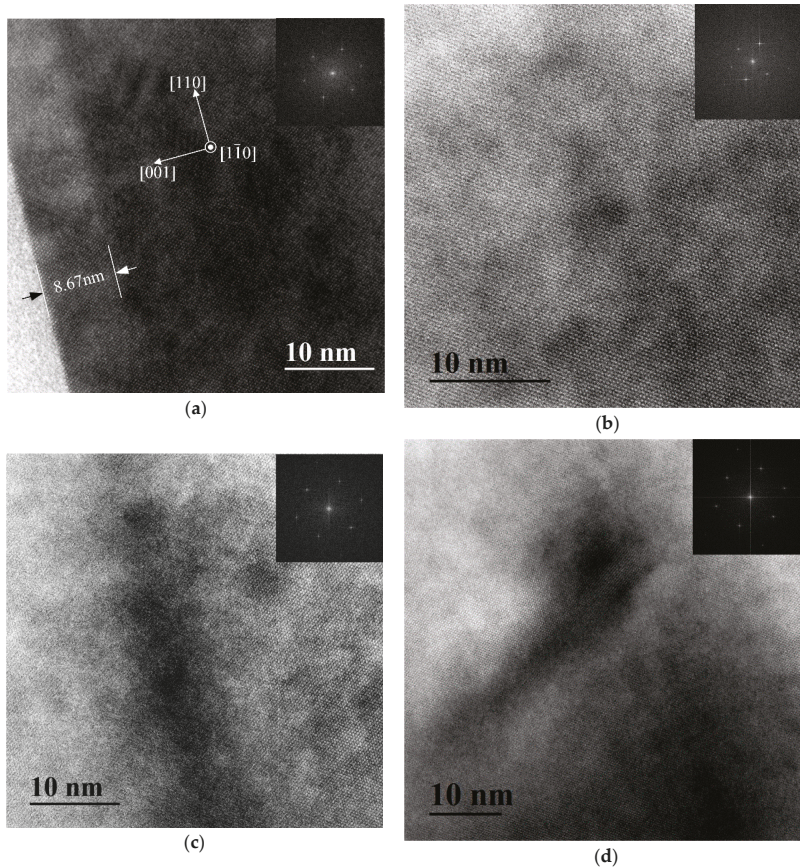


**Figure 6.** (a) Si-Si band full-width-at-half-maximum (FWHM) image obtained by Raman mapping on the sample cross-section, (b) FWHM distribution along the depth direction.

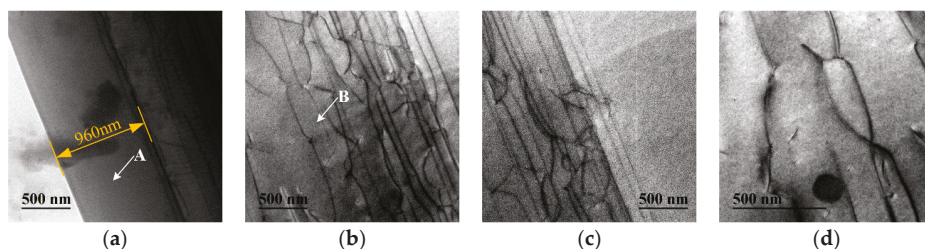


### 3.2. Nanoscale Structural Properties

The structural properties at the nanoscale were investigated by using TEM. Figures 7 and 8 give high- and low-resolution TEM images of the sample cross-section, respectively. Figure 7a shows a high-resolution TEM image, including the  $\epsilon$ -Si and  $\text{Ge}_{0.2}\text{Si}_{0.8}$  layers. Three regions exist, each of which has a uniform contrast grade, and which corresponds to a dissimilar layer. The left lower region, which is brightest, is an epoxy resin layer of the sample that serves as the adhesive for the two wafers. The image shows no regular lattice structure in this region. The region adjacent to the epoxy is the  $\epsilon$ -Si layer. The largest region in the image is the  $\text{Ge}_{0.2}\text{Si}_{0.8}$  layer. Inside the  $\epsilon$ -Si layer, the lattice structure is complete and clear, and shows a regular mesh of the monocrystalline state. The growth surface is in the same atomic sheet. No visible defects exist in this layer. Moreover, the interface of the  $\epsilon$ -Si/ $\text{Ge}_{0.2}\text{Si}_{0.8}$  layers appears clear, straight and almost atomically flat, where the lattices match well. The  $\epsilon$ -Si layer thickness is relatively uniform at  $\sim 8.67$  nm. Figure 7 also shows that the cross-section is a (100) crystal plane.



**Figure 7.** High-resolution transmission-electron microscopy (TEM) images of sample cross-section (a) at the  $\epsilon$ -Si and  $\text{Ge}_{0.2}\text{Si}_{0.8}$  layers, (b) inside the  $\text{Ge}_{0.2}\text{Si}_{0.8}$  layer, (c) at the  $\text{Ge}_{0.2}\text{Si}_{0.8}/\text{Ge}_x\text{Si}_{1-x}$  interface and (d) inside the  $\text{Ge}_x\text{Si}_{1-x}$  layer.



**Figure 8.** Low-resolution TEM images of sample cross-section (a) around  $\text{Ge}_{0.2}\text{Si}_{0.8}$  layer, (b,c) at  $\text{Ge}_x\text{Si}_{1-x}/\text{C-Si}$  interface and (d) inside  $\text{Ge}_x\text{Si}_{1-x}$  layer with a dislocation loop.

Figure 8a shows a low-resolution TEM image around the  $\text{Ge}_{0.2}\text{Si}_{0.8}$  layer, and Figure 7b shows a high-resolution enlarged image of spot A in Figure 8a, which is inside the  $\text{Ge}_{0.2}\text{Si}_{0.8}$  layer. In the low-resolution TEM image, no obvious defect exists inside the  $\text{Ge}_{0.2}\text{Si}_{0.8}$  layer. The interface between the  $\text{Ge}_x\text{Si}_{1-x}$  and  $\text{Ge}_{0.2}\text{Si}_{0.8}$  layers is clearly recognizable, which shows that the  $\text{Ge}_{0.2}\text{Si}_{0.8}$  layer thickness is uniform, at  $\sim 0.96 \mu\text{m}$  ( $\sim 1 \mu\text{m}$ ). According to the SEM, EDS and Raman results, as shown in Figures 1 and 3–7, the total thickness of the two GeSi buffer layers is  $3.6 \mu\text{m}$ , and hence, that of the graded layer ( $\text{Ge}_x\text{Si}_{1-x}$ ) is  $\sim 2.6 \mu\text{m}$ .

In the high-resolution TEM images Figure 7a,b, the lattice phenomena in the  $\text{Ge}_{0.2}\text{Si}_{0.8}$  layer are analogous to those in the  $\epsilon\text{-Si}$  layer. The two layers have the same lattice structure and dimensions and can only be distinguished according to differences in the contrast grades. The lattice maintained a good and similar crystal structure, which indicates that the stress in the  $\epsilon\text{-Si}$  and  $\text{Ge}_{0.2}\text{Si}_{0.8}$  layers is not released by defects or dislocations.

Figure 8b,c presents low-resolution TEM images around the  $\text{Ge}_x\text{Si}_{1-x}$  layer. Figure 7c,d presents high-resolution enlarged images at the  $\text{Ge}_x\text{Si}_{1-x}/\text{Ge}_{0.2}\text{Si}_{0.8}$  interface and at spot B in Figure 8b, respectively, where spot B is inside the  $\text{Ge}_x\text{Si}_{1-x}$  layer. In the low-resolution TEM image in Figure 8b,c, the interface between the  $\text{Ge}_x\text{Si}_{1-x}$  layer and C-Si substrate appears clear and straight. Numbers of dislocations exist in the  $\text{Ge}_x\text{Si}_{1-x}$  layer, especially close to the  $\text{Ge}_x\text{Si}_{1-x}/\text{Si}$  interface. Some appear as dislocation loops, such as the one in Figure 8d. Most dislocations connect and form dislocation networks inside the layer. Some vertical threading dislocation lines reach the  $\text{Ge}_x\text{Si}_{1-x}/\text{C-Si}$  interface.

Dislocation lines can only be terminated on the inner surface of the interface or on the free crystal surface, but they may not stop inside the crystal and appear in the form of isolated sections. Therefore, inside the crystal, any dislocation must be closed to form a dislocation ring or be connected to form a dislocation network, or move toward a surface/edge to form a threading dislocation.

All TEM images of the  $\epsilon\text{-Si}$  layer,  $\text{Ge}_{0.2}\text{Si}_{0.8}$  layer and Si substrate of the sample are clean and flat, without visible defects/dislocations. In contrast, dislocations are concentrated in the  $\text{Ge}_x\text{Si}_{1-x}$  layer with the highest dislocation intensity near the  $\text{Ge}_x\text{Si}_{1-x}/\text{C-Si}$  interface. The dislocations first arose in the beginning of the  $\text{Ge}_x\text{Si}_{1-x}$  epitaxial growth. With an increase in thickness during the epitaxial growth, some (existing and/or new born) dislocation lines connected with others, to form dislocation networks, and remained within the crystal. Some connected and became dislocation loops, or they continued extending to the sample edges. Because of the existence of dislocation in the  $\text{Ge}_x\text{Si}_{1-x}$ , especially near the  $\text{Ge}_x\text{Si}_{1-x}/\text{C-Si}$  interface, the average lattice quality of monocrystalline silicon deteriorated rapidly in the beginning of the graded buffer layer, which leads the FWHM of the Raman peak to broaden (as shown in Figure 6). With the gradual reduction in dislocations, the lattice quality improved.

Each type of dislocation helps to relax internal/residual stresses that are induced by the change in lattice dimension [25,26]. However, the threading dislocation is harmful to the structural reliability and performance of semiconductor devices, so it is necessary to avoid the appearance of threading dislocations. By using the manufacturing process to produce a strained silicon, dislocation

lines mainly form rings or networks inside the crystal structure, which avoids the generation of threading dislocations.

### 3.3. Distribution of Residual Stress with Depth

The residual stress inside the samples was analyzed, using MRS, which was regarded as an effective method for residual stress analysis of any semiconductor material, such as silicon. The basic Raman-mechanical model for crystal silicon (C-Si) follows the secular equation of the lattice dynamics [21,27]. According to existing investigations, the Raman-mechanical relationship of C-Si depends on the crystal plane and the stress state for measurement. In this work, the UV Raman tests were performed on the surface of the strained-silicon film, whose crystal plane was (100). Figure 9a shows a typical spectrum from the UV Raman tests. The 325-nm UV laser can only penetrate the crystal silicon to a depth of ~10 nm, which is similar to the thickness of the strained silicon. Hence, the wavenumber detected by the 325-nm laser achieves a mean wavenumber of 511.98 cm<sup>-1</sup> based on results from 66 different sampling spots, which originated solely from the ε-Si layer and were characterized by the average strain and stress of this layer. The residual stress in the ε-Si layer can be treated as a biaxial stress state, with stress components  $\sigma_x = \sigma_y = \sigma$  and strain components  $\epsilon_x = \epsilon_y = \epsilon$ , where  $x$  and  $y$  denote the in-plane directions. According to the Raman-mechanical relationship of C-Si with the (100) crystal plane [28],

$$\begin{aligned} \epsilon &= -1.204 \times 10^{-3} \Delta w \\ \sigma &= -217.28 \Delta w \text{ (MPa)} \end{aligned} \quad (1)$$

According to Equation (1),  $\epsilon = -1.204 \times 10^{-3} \times (511.98 - 519.78) = 9.39\% \approx 1\%$  and  $\sigma = -217.28 \times (511.98 - 519.78) \approx 1.69 \text{ GPa}$ , where wavenumber 519.78 cm<sup>-1</sup> was calibrated on a silicon wafer of the same production batch with the ε-Si/GeSi/C-Si samples and on the same day of the measurements above. The results showed that the real in-plane strain in the ε-Si layer was ~1%.

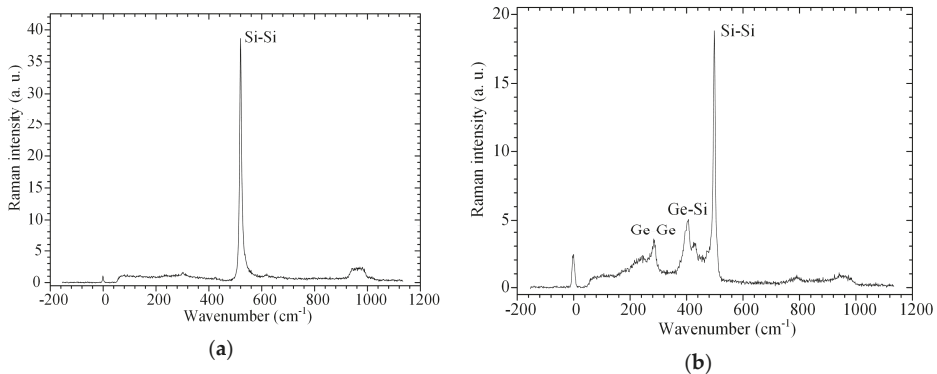


Figure 9. Raman spectra of (a) strained silicon (surface) and (b) germanium-silicon (cross-section).

For the Raman measurement on the cross-section with a (110) crystal plane, the applicable Raman-mechanical relationship of (110) C-Si was given as Equation (2), where the residual stress is treated as a uniaxial stress state, because the normal component on the cross-section surface (denoted as  $\sigma_x$ ) was released during the sample preparation,  $\sigma_x = 0$ ,  $\sigma_y = \sigma$  and strain components  $\epsilon_x = \epsilon_y = \epsilon$ , where  $x$  and  $y$  denote the in-plane directions [28].

$$\begin{aligned} \epsilon &= -2.58 \times 10^{-3} \Delta w \\ \sigma &= -434.58 \Delta w \text{ (MPa)} \end{aligned} \quad (2)$$

The cross-sectional stress analyses for the Raman data from the sampling spot inside the C-Si layer can use the Raman-mechanical relationship shown as Equation (2), whereas those from the GeSi

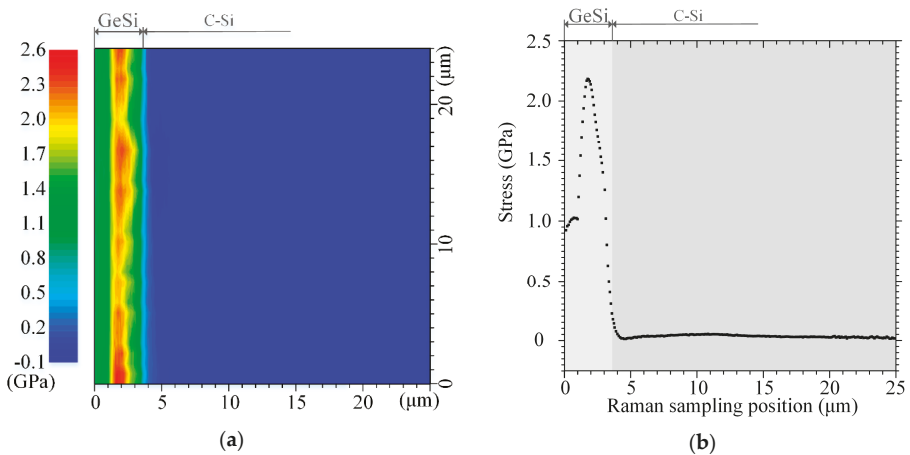
layers cannot use Equation (2) directly. Figure 9b shows a typical Raman spectrum of the germanium silicon alloy, which includes three major peaks: A Si-Si band at  $\sim 500\text{ cm}^{-1}$ , a Ge-Si band at  $\sim 400\text{ cm}^{-1}$  and a Ge-Ge band at  $\sim 300\text{ cm}^{-1}$ . If elemental Si dominated the Ge-Si alloy, just like the buffer layers in this work, the Raman peak of the Si-Si band would be strong and sufficiently narrow (shown in Figures 5 and 6) for mechanical measurement. However, Equation (2) is not applicable directly for the Si-Si band of germanium silicon alloy. As shown in Figure 4, the Raman wavenumber decreases sharply inside the GeSi layers, whose increment is up to  $14\text{ cm}^{-1}$ , denoting more than 6 GPa by using Equation (2). This is inconsistent with the reality of C-Si or GeSi strength. In fact, due to the influence of phonon confinement, the change in the wavenumber of the Si-Si band in the Raman spectrum of germanium silicon is not only induced by the strain, but also by the influence of phonon confinement. Tsang [29] presented an empirical relationship for the wavenumber  $w_{\text{Si-Si}}$ , the Ge content  $x$  and the strain  $\epsilon$ , which is regarded as suitable for the germanium-silicon alloy with  $0 < x < 0.25$ . Meanwhile, the elastic parameters of the germanium-silicon alloy followed the modulus-content linear model proposed by Kasper et al. [30].

$$C_{ij} = xC_{ij,\text{Ge}} + (1 - x)C_{ij,\text{Si}}, \quad i, j = 1 - 6, \tag{3}$$

where  $C_{ij}$ ,  $C_{ij,\text{Ge}}$  and  $C_{ij,\text{Si}}$  were components of the elastic coefficient tensors of the Ge-Si alloy, germanium and silicon, respectively. Hence the Raman-mechanical relationship suitable for the stress analysis inside the GeSi layers on the (110) cross-section in this work is given by Equation (4) [11].

$$\sigma = (100x - 435)(\Delta w + 62x) \text{ (MPa)}. \tag{4}$$

The residual stress distribution along the deepness direction of this multilayer structure is given as Figure 10, where the stress inside the C-Si substrate was achieved by using Equation (2), and that inside the germanium-silicon buffer layers was achieved by using Equation (4). Figure 10 shows that there exists serious residual stress inside the buffer layers. The residual stress increases rapidly in the beginning of the epitaxial growth. The mean stress gradient during the first  $2\text{ }\mu\text{m}$  of the graded buffer layer was  $777.5\text{ MPa}/\mu\text{m}$ . With further growth of the  $\text{Ge}_x\text{Si}_{1-x}$  layer, the residual stress stopped increasing and then began to decrease, which reduced the stress gradients and even turned them negative, until the beginning of the relaxed buffer layer. In the  $\text{Ge}_{0.2}\text{Si}_{0.8}$  layer, the residual stress was stable compared with that in the  $\text{Ge}_x\text{Si}_{1-x}$  layer. The mean residual stress in the relaxed buffer layer was  $\sim 1.01\text{ GPa}$ .



**Figure 10.** (a) Distribution image of residual stress on sample cross-section, (b) mean residual stress along the depth direction.

#### 4. Discussion

The stress distribution achieved by using Raman mapping showed that the residual stress did not increase continuously with the gradual rise in germanium content. The diameter of elemental germanium is larger than that of the silicon, and so is the lattice size of the Ge crystal. In the germanium–silicon layer prepared by reduced-pressure CVE, most silicon maintained its monocrystalline stage according to the Raman FWHM results (shown in Figure 7) and high-resolution TEM image (shown in Figure 7b,d). Therefore, some strain, and hence residual stress, was introduced into the buffer layer. The maximum residual tensile stress introduced by doping Ge in Si was 2.2 GPa, which did not induce any visible break or crack in each layer or interface, because of the birth and propagation of the dislocation, which was coincident with the GeSi alloy growth. Dislocation networks and dislocation loops helped to release the residual stress induced by the change in lattice dimension during the growth of a graded buffer layer where the germanium content increased gradually, which was positive for the reliability and stability of the  $\epsilon$ -Si structure during its application.

In the  $\text{Ge}_x\text{Si}_{1-x}$  layer, the germanium component started from zero at the  $\text{Ge}_x\text{Si}_{1-x}/\text{C-Si}$  interface and increased with a linear gradient (shown in Figure 3a), which resulted in a gradual change in lattice constant, avoided a sudden difference in lattice constant at the interface and reduced the mismatch strain and dislocation nucleation density compared with other growth methods. When the epitaxial  $\text{Ge}_x\text{Si}_{1-x}$  layer grew accompanied by a linear increase in Ge content, the mismatch strain relaxed gradually, and slowed down the extent of dislocation lines and the formation of dislocation networks. The dislocation lines stopped when the strain energy was released, which made the dislocation networks limited inside the  $\text{Ge}_x\text{Si}_{1-x}$  layer and unable to extend into the  $\text{Ge}_{0.2}\text{Si}_{0.8}$  and  $\epsilon$ -Si. Thus, the growth of the gradient-graded component can be effective for confining dislocations within the buffer layer.

As shown by the Raman results, the residual stress remained stable and the lattice quality improved inside the relaxed buffer layer. No obvious dislocation existed in this layer in the TEM images. These results proved that it was helpful to reduce the negative effect of germanium, and to improve the liability and performance of the microdevice based on strained-silicon technology.

#### 5. Conclusions

In this work, multiscale structural and mechanical properties of a  $\epsilon$ -Si/GeSi/C-Si material were investigated by several experimental methods. Based on sample cross-sections, the components and thickness of each layer were measured by EDS and SEM and the crystal and lattice characteristics and quality were observed by TEM and MRS. The residual stress distribution inside this multilayer material was achieved through micro-Raman mapping. The experimental results showed that the residual stress introduced during the heteroepitaxial growth was effectively controlled inside the buffer layers within a safe magnitude. The twin (graded plus relaxed) GeSi alloy structure provided a stable and reliable buffer for the manufacture of a strained-silicon film/ribbon in microelectronic devices.

**Author Contributions:** Conceptualization, W.Q. and Y.-S.Z.; methodology, W.Q., H.-T.W. and Y.-C.Z.; validation, L.-L.M.; resources, R.-R.L.; data curation, L.-L.M.; writing—original draft preparation, W.Q.; writing—review and editing, W.Q. and Y.-S.Z.

**Funding:** This research was funded by the National Key Research and Development Program of China (2018YFB0703500) and the National Natural Science Foundation of China (Grant Nos. 11827802, 11772223 and 11772227).

**Acknowledgments:** We thank Yongfeng Lu from the University of Nebraska-Lincoln for providing the experimental facilities, Dawei Li from the University of Nebraska-Lincoln for the AFM and WLI tests, and Laura Kuhar, from Liwen Bianji, Edanz Group China ([www.liwenbianji.cn/ac](http://www.liwenbianji.cn/ac)), for editing the English text of a draft of this manuscript.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

- Roldan, R.; Castellanos-Gomez, A.; Cappelluti, E.; Guinea, F. Strain engineering in semiconducting two-dimensional crystals. *J. Phys. Condens. Matter* **2015**, *27*, 313201. [[CrossRef](#)] [[PubMed](#)]
- Cao, T.; Wang, D.; Geng, D.S.; Liu, L.M.; Zhao, J. A strain or electric field induced direct bandgap in ultrathin silicon film and its application in photovoltaics or photocatalysis. *Phys. Chem. Chem. Phys.* **2016**, *18*, 7156–7162. [[CrossRef](#)] [[PubMed](#)]
- Saran Yalamarthy, A.; Senesky, D.G. Strain- and temperature-induced effects in AlGaIn/GaN high electron mobility transistors. *Semicond. Sci. Technol.* **2016**, *31*, 035024. [[CrossRef](#)]
- Dzuba, J.; Vanko, G.; Babchenko, O.; Lalinský, T.; Horvát, F.; Szarvas, M.; Kováč, T.; Hučko, B. In Strain induced response of AlGaIn/GaN high electron mobility transistor located on cantilever and membrane. *Int. Conf. Adv. Semicond. Devices Microsyst.* **2017**, 227–230. [[CrossRef](#)]
- Shen, T.; Penumatcha, A.V.; Appenzeller, J. Strain Engineering for Transition Metal Dichalcogenides Based Field Effect Transistors. *ACS Nano* **2016**, *10*, 4712. [[CrossRef](#)] [[PubMed](#)]
- Wen, H.; Borlaug, D.; Wang, H.; Ji, Y.; Jalali, B. Engineering Strain in Silicon Using SIMOX 3-D Sculpting. *IEEE Photonics J.* **2016**, *8*, 1–9. [[CrossRef](#)]
- Tian, P.; Edwards, P.R.; Wallace, M.J.; Martin, R.W.; McKendry, J.J.D.; Gu, E.; Dawson, M.D.; Qiu, Z.-J.; Jia, C.; Chen, Z.; et al. Characteristics of GaN-based light emitting diodes with different thicknesses of buffer layer grown by HVPE and MOCVD. *J. Phys. D Appl. Phys.* **2017**, *50*, 075101. [[CrossRef](#)]
- Kaganer, V.M.; Jenichen, B.; Ramsteiner, M.; Jahn, U.; Hauswald, C.; Grosse, F.; Fernández-Garrido, S.; Brandt, O. Quantitative evaluation of the broadening of x-ray diffraction, Raman, and photoluminescence lines by dislocation-induced strain in heteroepitaxial GaN films. *J. Phys. D Appl. Phys.* **2015**, *48*, 385105. [[CrossRef](#)]
- Lin, T.; Wang, F.; Cheng, C.-H.; Chen, S.; Feng, Z.C.; Lin, G.-R. Strain-related recombination mechanisms in polar InGaIn/GaN MQWs on amorphous SixC1-x buffers. *Opt. Mater. Express* **2018**, *8*, 1100. [[CrossRef](#)]
- Guo, Q.L.; Di, Z.F.; Lagally, M.G.; Mei, Y.F. Strain engineering and mechanical assembly of silicon/germanium nanomembranes. *Mater. Sci. Eng. R* **2018**, *128*, 1–31. [[CrossRef](#)]
- Qiu, W.; Cheng, C.L.; Liang, R.R.; Zhao, C.W.; Lei, Z.K.; Zhao, Y.C.; Ma, L.L.; Xu, J.; Fang, H.J.; Kang, Y.L. Measurement of residual stress in a multilayer semiconductor heterostructure by micro-Raman spectroscopy. *Acta Mech. Sin.* **2016**, *32*, 805–812. [[CrossRef](#)]
- Lockwood, D.J.; Wu, X.; Baribeau, J.M.; Mala, S.A.; Wang, X.; Tsybeskov, L. Si/SiGe Heterointerfaces in One-, Two-, and Three-Dimensional Nanostructures: Their Impact on SiGe Light Emission. *Front. Mater.* **2016**, *3*. [[CrossRef](#)]
- Cen, H.; Kang, Y.L.; Lei, Z.K.; Qin, Q.H.; Qiu, W. Micromechanics analysis of Kevlar-29 aramid fiber and epoxy resin microdroplet composite by Micro-Raman spectroscopy. *Compos. Struct.* **2006**, *75*, 532–538. [[CrossRef](#)]
- Wang, M.; Hu, X.F.; Wu, X.P. Internal microstructure evolution of aluminum foams under compression. *Mater. Res. Bull.* **2006**, *41*, 1949–1958. [[CrossRef](#)]
- Li, X.; Peng, Y. Investigation of capillary adhesion between the microcantilever and the substrate with electronic speckle pattern interferometry. *Appl. Phys. Lett.* **2006**, *89*, 234104. [[CrossRef](#)]
- Li, X.D.; Su, D.C.; Zhang, Z. A novel technique of microforce sensing and loading. *Sens. Actuators A Phys.* **2009**, *153*, 13–23. [[CrossRef](#)]
- Stoney, G.G. The Tension of Metallic Films Deposited by Electrolysis. *Proc. R. Soc. Lond.* **1909**, *82*, 40–43. [[CrossRef](#)]
- Van, A.S.; Batenburg, K.J.; Rossell, M.D.; Erni, R.; Van, T.G. Three-dimensional atomic imaging of crystalline nanoparticles. *Nature* **2011**, *470*, 374–377. [[CrossRef](#)]
- Nawaz, A.; Mao, W.G.; Lu, C.; Shen, Y.G. Mechanical properties, stress distributions and nanoscale deformation mechanisms in single crystal 6H-SiC by nano-indentation. *J. Alloys Compd.* **2017**, *708*, 1046–1053. [[CrossRef](#)]
- Qiu, W.; Li, Q.; Lei, Z.K.; Qin, Q.H.; Deng, W.L.; Kang, Y.L. The use of a carbon nanotube sensor for measuring strain by micro-Raman spectroscopy. *Carbon* **2013**, *53*, 161–168. [[CrossRef](#)]
- Kang, Y.L.; Qiu, Y.; Lei, Z.K.; Hu, M. An application of Raman spectroscopy on the measurement of residual stress in porous silicon. *Opt. Laser Eng.* **2005**, *43*, 847–855. [[CrossRef](#)]

22. Qiu, W.; Kang, Y.L. Mechanical behavior study of microdevice and nanomaterials by Raman spectroscopy: A review. *Sci. Bull.* **2014**, *59*, 2811–2824. [[CrossRef](#)]
23. Zhao, C.; Wen, S.; Hou, Q.; Qiu, W.; Xing, Y.; Su, S.; Cheng, B. Strain status of epitaxial Ge film on a Si (001) substrate. *J. Phys. Chem. Solids* **2016**, *90*, 87–92. [[CrossRef](#)]
24. Liang, R.; Zhang, T.; Wang, J.; Xu, J. High-Resolution Light-Emitting Diode Array Based on an Ordered ZnO Nanowire/SiGe Heterojunction. *IEEE Trans. Nanotechnol.* **2016**, *15*, 539–548. [[CrossRef](#)]
25. Shubhakar, K.; Bosman, M.; Neucheva, O.A.; Loke, Y.C.; Raghavan, N.; Thamankar, R.; Ranjan, A.; O'Shea, S.J.; Pey, K.L. An SEM/STM based nanoprobe and TEM study of breakdown locations in HfO<sub>2</sub>/SiO<sub>x</sub> dielectric stacks for failure analysis. *Microelectron. Reliab.* **2015**, *55*, 1450–1455. [[CrossRef](#)]
26. Zhao, C.W.; Xing, Y.M.; Zhou, C.E.; Bai, P.C. Experimental examination of displacement and strain fields in an edge dislocation core. *Acta Mater.* **2008**, *56*, 2570–2575. [[CrossRef](#)]
27. De Wolf, I. Relation between Raman frequency and triaxial stress in Si for surface and cross-sectional experiments in microelectronics components. *J. Appl. Phys.* **2015**, *118*, 053101. [[CrossRef](#)]
28. Qiu, W.; Ma, L.L.; Li, Q.; Xing, H.D.; Cheng, C.L.; Huang, G.Y. A general metrology of stress on crystalline silicon with random crystal plane by using micro-Raman spectroscopy. *Acta Mech. Sin.* **2018**. [[CrossRef](#)]
29. Tsang, J.C.; Mooney, P.M.; Dacol, F.; Chu, J.O. Measurements of alloy composition and strain in thin Ge<sub>x</sub>Si<sub>1-x</sub> layers. *J. Appl. Phys.* **1994**, *75*, 8098–8108. [[CrossRef](#)]
30. Dismukes, J.P.; Ekstrom, L.; Paff, R.J. Lattice Parameter and Density in Germanium-Silicon Alloys. *J. Phys. Chem.* **1964**, *68*, 3021–3027. [[CrossRef](#)]



© 2018 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).

Article

# A Real-Time Measurement Method of Air Refractive Index Based on Special Material Etalon

Guo-Ying Ren <sup>1,2,\*</sup>, Xing-Hua Qu <sup>1</sup> and Shuang Ding <sup>1</sup>

<sup>1</sup> State Key Laboratory of Precision Measurement Technology and Instruments, Tianjin University, Tianjin 300072, China; quxinghua@tju.edu.cn (X.-H.Q.); dingshuang0626@163.com (S.D.)

<sup>2</sup> National Institute of Metrology, Beijing 100029, China

\* Correspondence: rengy@nim.ac.cn

Received: 29 October 2018; Accepted: 16 November 2018; Published: 21 November 2018



**Abstract:** In the precise displacement measurement based on laser interferometry, the measurement technology for the refractive index of air is widely used to improve the measurement accuracy. However, the existing measurement method of the refractive index of air based on direct measurement is not easy to realize in practical work because of its complex measurement principle and the huge volume of the measurement device; while the measurement accuracy and speed based on the indirect method cannot adapt to the real-time, fast and accurate measurement requirements of industrially changing environments, resulting in distortion of the results. In this study, a measurement method of the refractive index of air based on a special material etalon is proposed. The method enables rapid and direct measurement of the air refractive index when the environment changes and it is given the realization process. Finally, the experimental results show that the deviation between this method and the modified Edlen formula is about  $2.5 \times 10^{-7}$ , and that this method can quickly reflect the changes of the environment, which prove the correctness of this method and its ability manage rapid environmental responses. This method is worth popularizing in industrial measurement.

**Keywords:** metrology; precision measurement; air refractive index; wavelength correction

## 1. Introduction

During the displacement measurement based on the principle of laser interferometry, the refractive index of air in the optical path is often changed because of the change of ambient atmosphere, but it is difficult to measure the refractive index of air rapidly and accurately because of the limitation of the conditions or measurement schemes [1]. According to the principle of laser interference:

$$L = \frac{K \lambda_0}{2n} \quad (1)$$

where  $L$  is the length of the measurement,  $K$  is the interference level,  $n$  is the refractive index of air, and  $\lambda_0$  is the vacuum wavelength.

It is known that the variation ( $\Delta L$ ) of the displacement measurement  $L$ , caused by the variation ( $\Delta n$ ) of air refractive index  $n$ , is as follows:

$$\Delta L = \frac{K \lambda_0}{2 n^2} \Delta n \quad (2)$$

Therefore, for precise measurement, the change in the air refractive index  $n$ , caused by the change in environmental factors and the inaccuracy of its measurement, become one of the main reasons for the limitation on the improvement of the accuracy of the measurement results. How to measure the



air refractive index accurately and quickly in order to compensate for the limitation is the primary consideration in the scheme design of measurement.

At present, there are two ways to obtain a measurement of the air refractive index: The direct method and the indirect method. The direct method is to calculate the air refractive index by directly comparing the optical path difference between the air cavity and the same-length vacuum cavity with the refractive index measuring instrument or laser interference technology [2,3], which can reach a  $10^{-8}$  level of measurement accuracy. However, its design and fabrication is complex and the vacuum required by the experiment is very high. At the same time, the atmospheric pressure makes the vacuum system easy to deform and re-introduce the error source, so it is seldom used in engineering. The indirect method is to measure the parameters of the air composition through various sensors, such as ambient temperature, humidity, air pressure and so on. The refractive index value is calculated in order to obtain the air refractive index formula [4–12]. The important characteristic of the indirect measurement method is that it is convenient and easy to use. The measurement accuracy is  $10^{-7}$  level but because of the delay of the sensor itself, the air refractive index value calculated after the measurement is not the true value at that time. The measurement system accuracy is also low and is prone to produce errors in fact. Some errors are up to a  $10^{-6}$  level in some measurement systems.

In order to synthesize the advantages of the above two methods and avoid their disadvantages, this paper presents a new measurement method for the refractive index of air based on a special material etalon.

## 2. Theory

### 2.1. Principle of Measurement

From Equation (2), it can be derived that:

$$\Delta n = n_1 - n_0 = \Delta L \frac{n_0}{L} \tag{3}$$

where  $n_0$  is the initial measurement value of the air refractive index and  $n_1$  is the current measurement value of the air refractive index.

So, the current value  $n_1$  of refractive index of air can be calculated as:

$$n_1 = n_0 + \Delta n = n_0 + \Delta L \frac{n_0}{L} = n_0 \cdot \left( 1 + \frac{\Delta L}{L} \right) \tag{4}$$

According to this principle, a type of zero-expansion quartz material with the known coefficient of thermal expansion and its value close to zero, was selected to design and manufacture a standard etalon work-piece (as shown in Figure 1) with length  $L$ . It was placed in the measuring light path in order to monitor the variation of the wavelength of the light path caused by the fluctuation of the environment during the measurement.

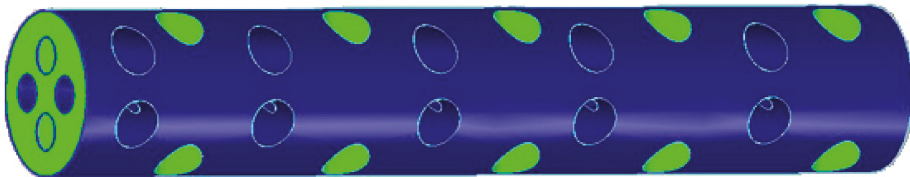
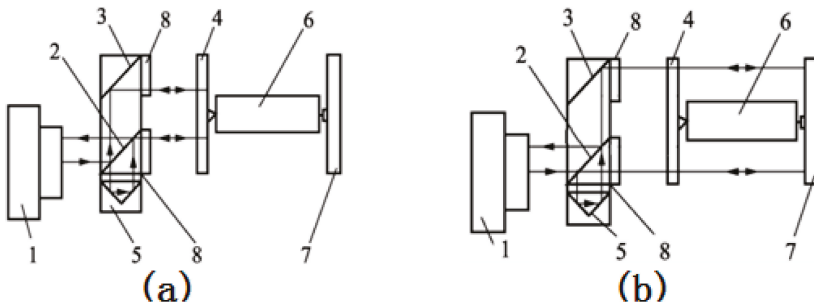


Figure 1. Standard etalon work-piece for measuring the refractive index of air.

The front-end face of the standard etalon was coated with reflective film as a reference arm, and the back end of the standard etalon was coated with reflective film as a measuring arm. The layout of the reference light path and the measured light path based on the etalon are shown in Figure 2 [13].



**Figure 2.** The layout of light path based on the etalon. (a) Description of the layout of the reference light path based on the etalon. (b) Description of the layout of the measurement light path based on the etalon.

In Figure 2a, the beam is emitted from laser one, reflected from the reflector mirror two to beam bender mirror three, passed through the quarter wave plate eight, which is incident to measuring mirror four fixed on the front face of the tested piece six, returned from four and passed through two, reflected from two to four again through cube corner mirror five, and returned from four to form a reference beam through two. The measurement optical path in Figure 2b can be analyzed similarly.

As shown in Figure 2, the measured beam and the reference beam pass through the same optical path in the interferometer. They are designed based on the concept of the common light path and symmetrical distribution of the interferometer. So, the measurement system has a heat balance function, that is to say, when the size of the interferometer varies with the temperature effect, the influence on the length measurements of the two light paths are cancelled out one another.

### 2.2. Measurement Model

In order to obtain an accurate initial  $n_0$ , the indirect measurement method of the air refractive index should be used first. Therefore, some high-sensitivity temperature, pressure and humidity sensors, and CO<sub>2</sub> water vapor sensors were arranged in the measuring optical path in order to measure the air temperature, pressure and humidity, and then the initial refractive index  $n_0$  was calculated and obtained by the refractive index correction formula.

Because the conventional Edlen formula for calculating the air refractive index does not include the modified item of the CO<sub>2</sub> error [14,15], the improved Edlen formula is as follows:

$$(n - 1)_s = 10^{-8} \cdot [A + B/(130 - \sigma^2) + C/(38.9 - \sigma^2)] \tag{5}$$

$$(n - 1)_x = [1 + 0.540 \cdot (x - 0.0003)] \cdot (n - 1)_s \tag{6}$$

$$n_{Tp} = 1 + P \cdot (n - 1)_x / D \cdot [1 + 10^{-8} \cdot (E - F \cdot T) \cdot P] / (1 + G \cdot T) \tag{7}$$

$$Y = -J \cdot [1 - (T/273.16)^{-1.5}] + K \cdot [1 - (T/273.16)^{-1.25}] \tag{8}$$

$$n_{Tpf} - n_{Tp} = -(f/100 \cdot 611.657 \cdot e^Y) \cdot (292.75/T) \cdot [3.7345 - 0.0401 \cdot \sigma^2] \cdot 10^{-10} \tag{9}$$

Here, the constants can be defined as follows [11]:

$A = 8342.54$ ;  $B = 2,406,147$ ;  $C = 15,998$ ;  $D = 96,095.43$ ;  $E = 0.601$ ;  $F = 0.00972$ ;  $G = 0.003661$ ;  $J = 13.928169$ ;  $K = 34.7078238$ ;  $\sigma = 1/\lambda$ ;  $T = t + 273.15$ .

Here  $(n - 1)_s$  is the refractive index of air in the standard state,  $\sigma$  is the wave number in the vacuum, expressed in the unit of  $\text{um}^{-1}$ ,  $(n - 1)_x$  is the refractive index of air with the molar content of carbon dioxide, where the unit is  $(\times 10^{-6})$ ,  $n_{Tp}$  is the refractive index of air under the conditions of temperature, air pressure and carbon dioxide, and  $n_{Tpf}$  is the refractive index of air under the conditions of temperature, humidity, air pressure and carbon dioxide.

After obtaining the initial value of the air refractive index, the etalon can be utilized to calculate the real-time wavelength compensation. The compensation factor  $C$  is:

$$C = \frac{10^6}{n_{Tpf} + 10^6} \tag{10}$$

Suppose that the number  $P_n$  of preset wavelengths at the moment of measurement is:

$$P_n = \frac{LRf}{C\lambda} \tag{11}$$

where  $L$  is the effective length between the two mirrors of the etalon of the air refractive index,  $R$  is the electrical resolution of the interferometer, and  $f$  is the fold factor of the laser interference in the light path.

The value of the compensation  $C_n$  is therefore:

$$C_n = \frac{LRf}{(P_A + P_n)\lambda} \tag{12}$$

where  $P_A$  is cumulative count.

The dead-path count  $C_{Dn}$  is:

$$C_{Dn} = \frac{DRf}{C\lambda} \tag{13}$$

where  $D$  is the dead-path length in the interferometer optical path.

Then the initial actual displacement  $P_{s0}$  after compensation is:

$$P_{s0} = \frac{C_n(P_A + C_{Dn})\lambda}{Rf} \tag{14}$$

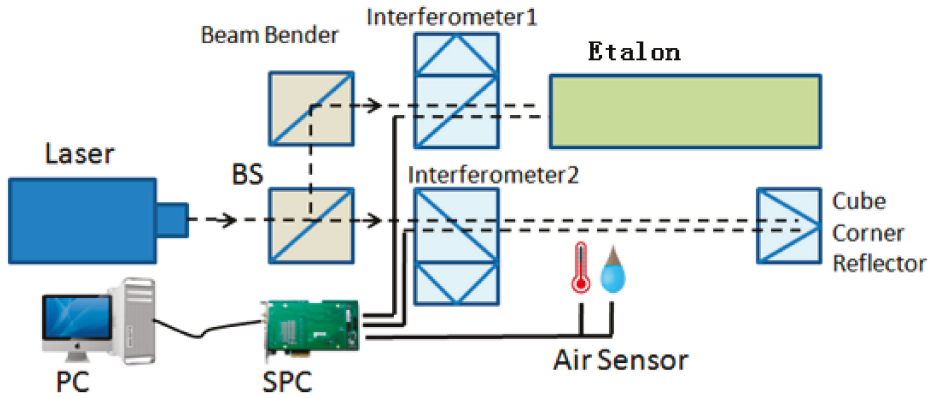
Finally, the current value  $n_1$  of the reflective index of air is:

$$n_1 = \left( 1 + \frac{\Delta L}{P_{s0}} \right) \tag{15}$$

### 3. Experiments and Results

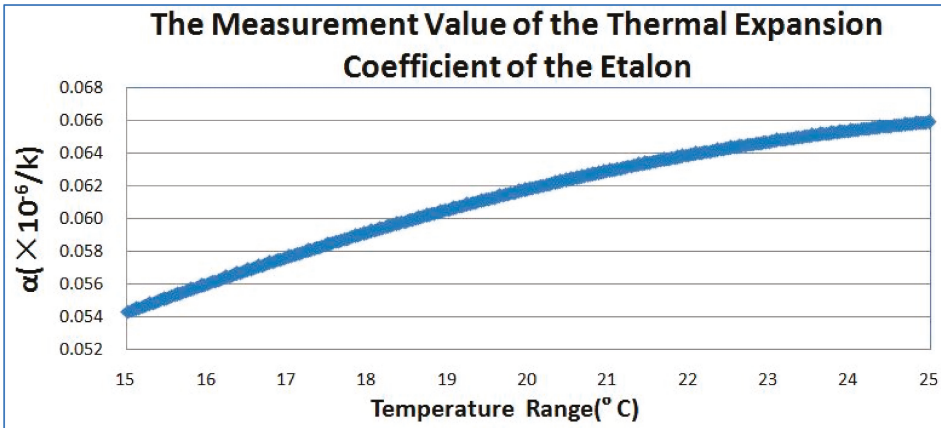
In order to verify the correctness of the above formulae of the air refractive index based on the etalon, the verification scheme of measurement of the optical path of the air refractive index was constructed as follows (see Figure 3).

The light from the laser passes through the beam splitter (BS) mirror and divides into two paths. One path passes through the beam bender mirror and the first interferometer to the air refractive index standard (etalon). The returned light passes through the first interferometer and returns to the electronic signal subdivision circuit (SPC). The processed signal data is collected by the computer. The other light path passes through the second interferometer to the cube corner reflector. It is reflected back to the second interferometer and back to the SPC. The processed signal data is collected by the computer. At the same time, the environmental sensors are arranged on a second path in order to calculate the initial value  $n_0$  of the refractive index of the air. In particular, these two light paths can be placed close to each other, so it can be thought that the two light paths have the same environmental conditions and the same refractive index of the air.



**Figure 3.** The verification scheme for the measurement optical path of the air refractive index. BS: beam splitter, SPC: signal subdivision circuit, PC: personal computer.

In the experiment, a Renishaw RLE10 was selected as the laser source, which has a laser signal with 64 times electronic subdivision. The actual measurement value of the thermal expansion coefficient of the Etalon used in the air refractive index standard tool is  $6.1 \times 10^{-8}/K$  (see Figure 4). The variation of the thermal expansion coefficient of the Etalon in the range of  $20 \text{ }^\circ\text{C} \pm 5 \text{ }^\circ\text{C}$  is not more than  $1.0 \times 10^{-8}/K$ , which guarantees the outstanding thermal stability of the length.



**Figure 4.** The measurement value of the thermal expansion coefficient of the Etalon.

The diagram comparing the variation between the measured values of the air refractive index obtained from the Etalon in the first light path and the calculated value of the air refractive index from the ambient atmosphere sensors in the second path are shown in the following figure (see Figure 5).

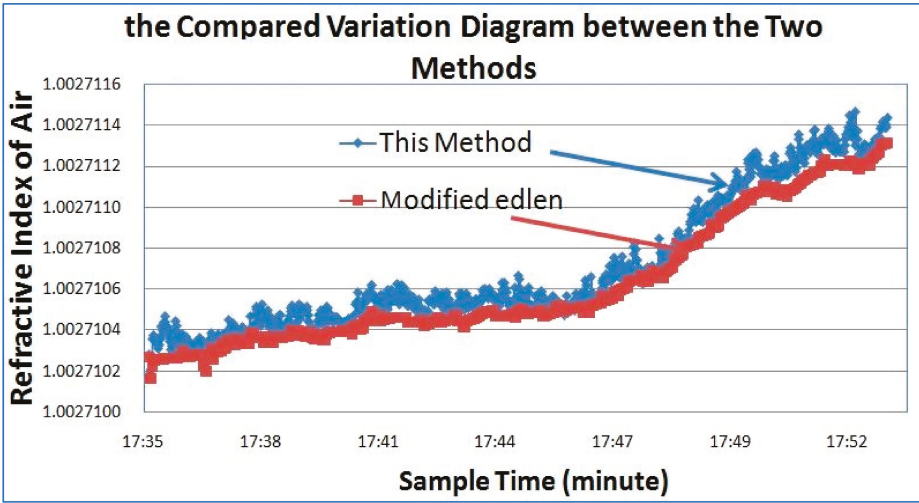


Figure 5. The compared variation diagram between the two methods.

It can be seen from Figure 5 that the change in the air refractive index calculated by the modified Edlen formula is relatively gentle because of the hysteresis of the sensors, which cannot fully reflect the effect of air fluctuation on the actual air refractive index. However, the method proposed in this paper can obtain the real-time changes in the air refractive index easily and quickly, and perform the further modification and compensation in the subsequent length measurement. From Figure 6, it can be seen that the maximum difference between the two measurement methods is up to  $2.5 \times 10^{-7}$ .

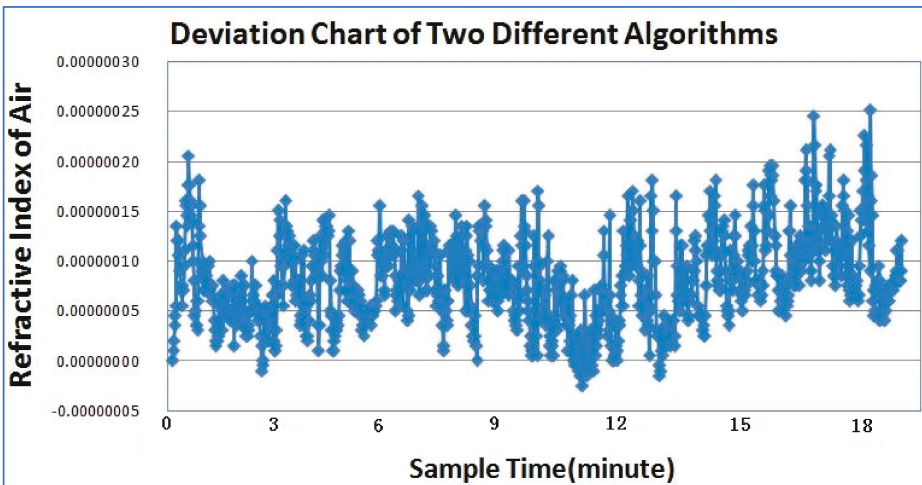


Figure 6. The maximum difference between the two measurements.

#### 4. Conclusions

From the above analysis it can be seen that the requirement for length measurement accuracy becomes higher and higher in the precision measurement. However, the current measurement methods for the air refractive index based on Edlen formula cannot provide real-time measurement results because of the ambient atmosphere sensors it requires. In this paper, a new method for air refractive

index measurement is proposed, which can rapidly measure a changing air refractive index in real time, and provides real-time wavelength correction data for high-precision length measurement. Finally, the correctness of the method is proven by the experimental results. In future we will attempt to analyze and describe the high-frequency noise problem which might have been present in the described method, although the noise may have been very small.

There is hope that this measurement method can provide a reference for others who need to improve the accuracy of laser interferometry in the precision manufacturing industry.

**Author Contributions:** G.-Y.R. conceived the method, improved the experiments and wrote the paper; X.-H.Q. conducted the experiments; S.D. curated the data.

**Funding:** This research was funded by the National Key Research and Development Program of China, grant number: 2018YFF0212702.

**Acknowledgments:** This study was supported by the National Key Research and Development Program of China (Grant No.: 2018YFF0212702). The authors would like to thank the other members of the research team for their contributions to this study.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Minoshima, K.; Arai, K.; Inaba, H. High-accuracy self-correction of refractive index of air using two-color interferometry of optical frequency combs. *Opt. Express* **2011**, *19*, 26095–26105. [CrossRef] [PubMed]
2. Li, D.G.; Zhang, G.X. Measurement and compensation for refractive index of air using a pre-evacuated airtight sealed cavity. *Opt. Precis. Eng.* **2001**, *9*, 80–84.
3. Li, D.G.; Zhang, G.X. Measuring refractive index of air by using a pre-evacuated airtight sealed cavity and its accuracy analysis. *Opt. Tech.* **2000**, *26*, 499–501.
4. Ciddor, P.E. Refractive index of air: New equations for the visible and near infrared. *Appl. Opt.* **1996**, *35*, 1566–1573. [CrossRef] [PubMed]
5. Edlén, B. The refractive index of air. *Metrologia* **1966**, *2*, 71–80. [CrossRef]
6. Birch, K.P.; Downs, M.J. Correction to the Updated Edlén Equation for the Refractive Index of Air. *Metrologia* **2005**, *31*, 315–316. [CrossRef]
7. Bonsch, G.; Potulski, E. Measurement of the refractive index of air and comparison with modified Edlén's formulae. *Metrologia* **2003**, *35*, 133–139. [CrossRef]
8. Giacomo, P. Equation for the Determination of the Density of Moist Air (1981). *Metrologia* **2005**, *18*, 33–40. [CrossRef]
9. Beers, J.; Doiron, T. Short Communication: Verification of Revised Water Vapor Correction to the Index of Refraction of Air. *Metrologia* **1992**, *29*, 315–316. [CrossRef]
10. Birch, K.P.; Reinboth, F.; Ward, R.E.; Wilkening, G. The Effect of Variations in the Refractive Index of Industrial Air upon the Uncertainty of Precision Length Measurement. *Metrologia* **1993**, *30*, 7–14. [CrossRef]
11. Stone, J.A.; Zimmerman, J.H. Index of Refraction of Air. Available online: <https://emtoolbox.nist.gov/Wavelength/Documentation.asp#IndexofRefractionofAir> (accessed on 29 October 2018).
12. Cui, J.J.; Liu, X.B.; Kang, Y.H.; Zhang, H.; Li, J.S. Calibration on Measurement System for the Refractive Index of Air Based on Edlén Formula. *Acta Metrol. Sin.* **2014**, *35*, 210–215.
13. Zhang, X.D.; Su, Y.C.; Ye, X.Y. Thermal Expansion Coefficient Measurement with Precision Interferometer. *Acta Metrol. Sin.* **2012**, *33*, 1–4.
14. Ren, G.Y. Precision Measurement for Feature Length of Cylinder Acoustic Resonator Cavity. *Acta Metrol. Sin.* **2012**, *33*, 212–214.
15. Ren, G.Y.; Wang, W.N.; Wang, Z.J. Uncertainty Analysis for Measurement of Step Gauges Based on Coordinate Measuring Machine. *Acta Metrol. Sin.* **2008**, *29* (Suppl. S1), 151–153.



© 2018 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).

Article

# An Orthogonal Type Two-Axis Lloyd's Mirror for Holographic Fabrication of Two-Dimensional Planar Scale Gratings with Large Area

Xinghui Li <sup>1,\*</sup>, Haiou Lu <sup>1,†</sup>, Qian Zhou <sup>1,\*</sup>, Guan hao Wu <sup>1,2</sup>, Kai Ni <sup>1</sup> and Xiaohao Wang <sup>1</sup>

<sup>1</sup> Division of Advanced Manufacturing, Graduate School at Shenzhen, Tsinghua University, Tsinghua Campus, Xili University Town, Shenzhen 518055, China; lho16@mails.tsinghua.edu.cn (H.L.); guanhaowu@tsinghua.edu.cn (G.W.); ni.kai@sz.tsinghua.edu.cn (K.N.); wang.xiaohao@sz.tsinghua.edu.cn (X.W.)

<sup>2</sup> Department of Precision Instrument, Tsinghua University, Haidian District, Beijing 100084, China

\* Correspondence: li.xinghui@sz.tsinghua.edu.cn (X.L.); zhou.qian@sz.tsinghua.edu.cn (Q.Z.)

† These authors contributed equally to this work.

Received: 30 October 2018; Accepted: 15 November 2018; Published: 19 November 2018



**Abstract:** In this paper, an orthogonal type two-axis Lloyd's mirror interference lithography technique was employed to fabricate two-dimensional planar scale gratings for surface encoder application. The two-axis Lloyd's mirror interferometer is composed of a substrate and two reflective mirrors (X- and Y-mirrors), which are placed edge by edge perpendicularly. An expanded and collimated beam was divided into three beams by this interferometer, a direct beam and two reflected beams, projected onto the substrate, X- and Y-mirrors, respectively. The unexpected beam sections having twice reflected off the mirrors were blocked by a filter. The remaining two reflected beams interfered with the direct beam on the substrate, generating perpendicularly cross patterns thus forming two-dimensional scale gratings. However, the two reflected beams undesirably interfere with each other and generate a grating pattern along 45-degree direction against the two orthogonal direction, which influence the pattern uniformity. Though an undesired grating pattern can be eliminated by polarization modulation with introduction of waveplates, spatial configuration of waveplates inevitably downsized the eventual grating, which is a key parameter for grating interferometry application. For solving this problem, theoretical and experimental study was carefully carried out to evaluate the fabrication quality with and without polarization modulation. Two-dimensional scale gratings with a 1  $\mu\text{m}$  period in X- and Y-directions were achieved by using the constructed experiment system with a 442 nm He-Cd laser source. Atomic force microscopy (AFM) images and the result of diffraction performances demonstrated that the orthogonal type two-axis Lloyd's mirror interferometer can stand a small order undesired interference, that is, a degree of orthogonality between two reflected beams, denoted by  $\gamma$ , no larger than a nominal value of 0.1.

**Keywords:** interference lithography; two-axis planar scale grating; Lloyd's mirror; surface encoder

## 1. Introduction

Planar encoders play a key role in precision positioning of linear stages due to their high resolution, high robustness, and relatively low cost [1–9]. The key component in the planar encoder is a two-dimensional (2D) planar scale grating, on which typically surface relief square holes or sinusoidal hills with a period spacing from submicron to several microns are arrayed along two orthogonal directions.

Laser interference lithography (LIL), engine machining, and imprinting are three typical methods for fabrication of the scale grating. Compared with engine machining or imprinting, LIL is more

convenient and cost-effective because only a coherent laser source is required and neither sophisticated precision positioning system nor expensive photolithography equipment and mask preparation are required. For fabrication of a 2D planar grating, according to the exposure times, the LIL process can be separated into two categories, double exposure and single exposure. For the two-beam two-exposure, it requires twice exposure procedures and the second time exposure requires rotation of the substrate by 90 degrees [5,10]. The exposure can be based on an amplitude-division type two-beam LIL, called two beam two exposures (TBTE) or a wavefront-division type Lloyd's mirror type LIL, called one beam two exposures (OBTE) [11]. TBTE is good at fabrication of gratings with a relatively large area up to 100 mm × 100 mm, while OBTE can be easily handled with a relatively simple and stable optical system. However, a big challenge existing in the method of the two-exposure process is that the grating structures generated in the first exposure will be influenced by the background light in the second exposure. Thus the depths of the grating structures will be different in two orthogonal directions. As a result, diffraction efficiencies of both the negative and positive diffraction beams will not be consistent. This is a critical problem for the encoder system since it directly affects the accuracy and the detecting limit of the interference signals. Therefore, efforts for solving this problem using a single exposure for 2D scale grating have been done these years. There are two trends for this motivation, multibeam LIL and one beam LIL. Multibeam LIL includes, but is not limited to, triple-beam and quadruple-beam interference technique. This technology can achieve the goal of 2D grating patterns with a single exposure. Some researchers invented a novel nanofabrication technique, called four-beam interference lithography, by using a two-dimensional grating [12]. Nevertheless, it cannot be neglected that the multibeam interference lithography has complicated optical structure easily to be disturbed by external environment. Furthermore, the optical path difference (OPD) is also difficult to be manipulated.

In contrast, one beam LIL, represented by two-axis Lloyd's mirror interferometer and a corner-like reflector interferometer, allows fabricating a crossed 2D grating with a single exposure [13–18]. For one beam LIL, one more proposal called the two-axis Lloyd's mirror interferometer with polarization modulation technique, proposed in 2014, has the advantages of large grating area, easy handling, and stable grating patterns [13]. However, optimization of the theoretical inclination in grating shape requires sophisticated in situ detection system for exposure and development. The inclination will influence the diffraction efficiency in a certain degree when it is directly used as a photoresist grating, though this influence might be reduced when a further etched process. On the other hand, the corner-like reflector interferometer, also called orthogonal type two-axis Lloyd's mirror interferometer, also proposed in 2014, has the advantages on grating shape in depth direction, although the complicated interference patterns on the grating substrate require deep modification for application as a 2D scale grating [14]. For modification of the grating patterns, Yuki and Chen took polarization modulation technique and results verified this technology has a great potential for 2D scale grating fabrication. But it should be noted that, introduction of the polarization modulation both reduce the grating area and bring a new challenge in the system stability [16–18].

Thus, investigation of the orthogonal type two-axis Lloyd's mirror interferometer based LIL without polarization modulation for a large grating area is carried out in this research. We constructed an orthogonal type two-axis Lloyd's mirror interferometer and fabricated 2D scale grating under different polarization state. Design details, construction, experimental setup, and evaluation are presented in this paper.

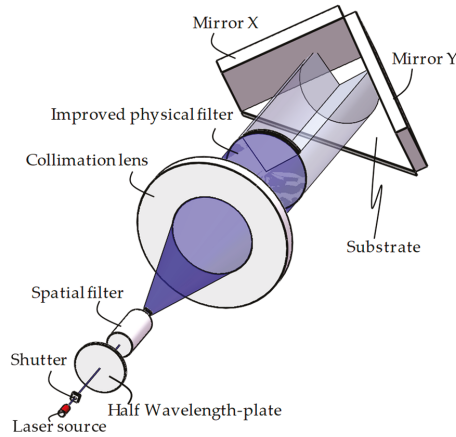
## 2. Methods

### 2.1. Optical Setup of the Exposure System

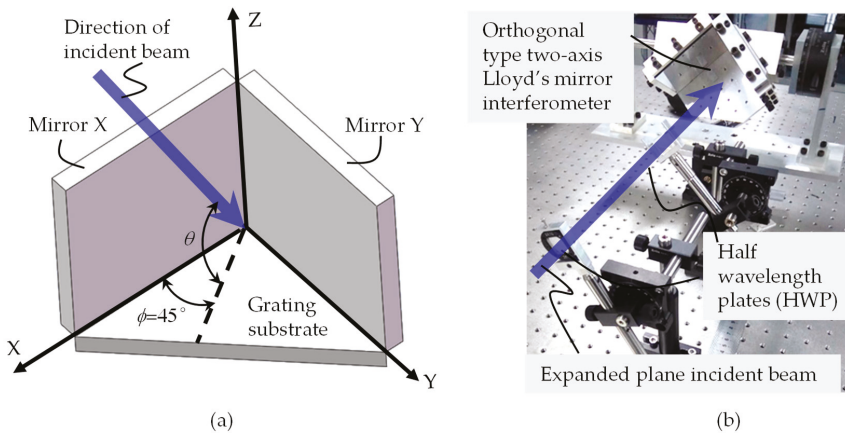
An orthogonal two-axis Lloyd's mirror interferometer was designed and assembled, as shown in Figure 1, which mainly includes a light source, beam shaping unit, polarization adjustment unit, and two-axis Lloyd's mirror unit. For the light source, with consideration of laser source quality (stability of power, polarization, and wavelength), photoresist sensitivity against laser wavelength, and optical



alignment low-complexity related to laser visibility, we employed a commercial, well-used He-Cd laser of the KIMMON KOHA company (Tokyo, Japan) with the wavelength of 442 nm and power output of 180 mW. Two optical elements are contained in the beam-shaping unit: the spatial filter element and collimating lens element. The two-axis Lloyd’s mirror unit, as shown in Figure 2a, consists of three elements: X-mirror, Y-mirror, and the jig of substrate. Besides, the three elements are placed perpendicularly to each other. In addition, the two-axis Lloyd’s mirror unit possess the function of rotation so that it’s easy to manipulate the grating period. We designed the azimuth angle  $\varphi$  equals to  $45^\circ$ , and the angle  $\theta$  ( $0-90^\circ$ ) called the incident angle in this paper.



**Figure 1.** The optical configuration of the orthogonal type two-axis Lloyd’s mirror interferometer-based laser interference lithography (LIL) system.

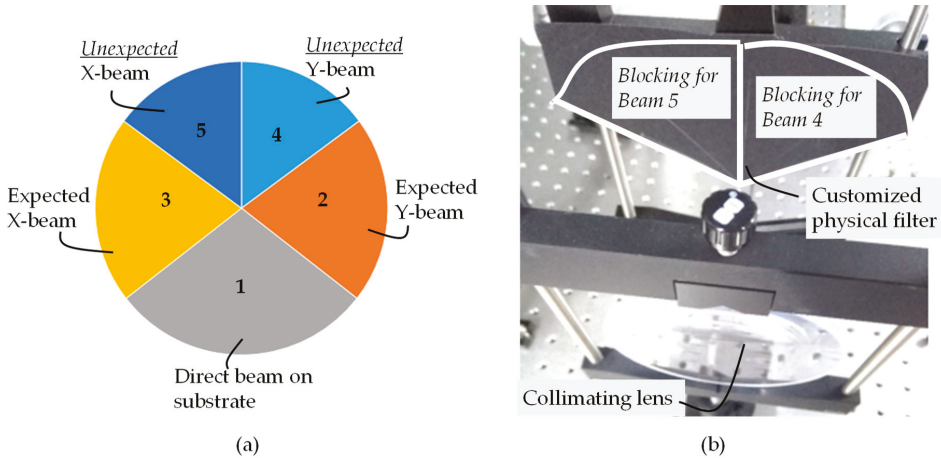


**Figure 2.** The global coordinate system of exposure system. (a) Two-axis Lloyd’s mirror unit and definition of incident angle  $\theta$  and azimuth angle  $\varphi$ . (b) Photograph of half-wavelength plates (HWPs) setup.

We selected the mirror with the coating of enhanced aluminum for the purpose of adjusting arbitrarily the grating pitch. The mirror with the dielectric interface, by contrast, has a limited angle of incidence, usually within a range of 0 to  $45^\circ$ . Two half-wavelength plates were employed in the polarization adjustment unit that controlled the polarization state of corresponding incidence beams, as shown in Figure 2b. Section 2.2 introduces the mechanism of polarization adjusting.

2.2. Interference Theory of Orthogonal Two-Axis Lloyd’s Mirror Interferometer

According to the analysis of the interferometer, the collimated beam was divided into five sections, as shown in the Figure 3a, called beam 1, 2, 3, 4, and 5. The beam directly projecting onto the substrate is beam 1; the beam propagating onto the substrate after being reflected by the X-mirror is beam 2; the beam propagating onto the substrate after being reflected by the Y-mirror is beam 3; the beams propagating onto the substrate after being reflected by the Y-mirror firstly and reflected by the X-mirrors secondly is beam 4; and the beams propagating onto the substrate after being reflected by the X-mirror firstly and reflected by the Y-mirrors secondly is beam 5.



**Figure 3.** Optical configurations of orthogonal two-axis Lloyd’s mirror interferometer. (a) The sections of collimated beam (from the direction of the incident beam). (b) Improved physical filter.

First of all, however, it should be noted that only beams 1, 2, and 3 are required for fabrication of the desired 2D gratings. Beams 4 and 5 will introduce unexpected grating patterns [14]. Thus, a physical filter is stuck in the optical path as shown in Figure 3b, a similar manner as that illustrated in present methods [16–18]. It should be noted that the incident plane wavefront beam still follows a Gaussian distribution, which will influence the grating shape uniformity among the whole grating surface.

Theoretically, assuming that only three—beams 1, 2, and 3—have uniform intensity distribution at the wavefront, the electric field of a laser beam can be expressed as follows.

$$E_i = E_i \cdot e_i \exp(jk_i \cdot r + \gamma_i) \quad (i = 1, 2, 3, 4, 5), \tag{1}$$

where  $r$  is the position vector,  $E_i$  is the real electric field amplitude,  $e_i$  is a unit vector in the polarization direction of the laser beam,  $\gamma_i$  is the initial phase, and  $k_i$  is the wave vector. The total electric field of needed incident beams can be expressed by

$$e(r) = \sum_{i=1}^3 E_i \exp(jk_i \cdot r + \gamma_i) \quad (i = 1, 2, 3). \tag{2}$$

The intensity distribution of interference field upon the substrate can be written as follows,

$$I(r) = \sum_{i=1}^3 |E_i|^2 + 2 \sum_{m=2}^3 \sum_{n<m} E_m E_n e_m e_n \cos\{(k_n - k_m) \cdot r + \varphi_n - \varphi_m\}. \tag{3}$$

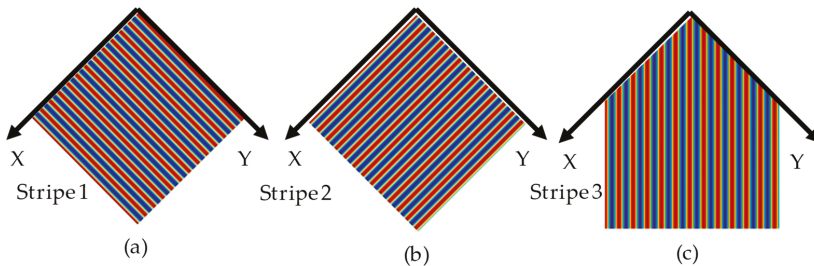
According to Equation (3), the periods of the interference fringes to be generated by the interference between the two of those beams can be expressed as follows,

$$g_{nm} = \frac{2\pi}{|k_n - k_m|} \quad (n < m \leq 3). \tag{4}$$

Assume  $\varphi = 45^\circ$  and the mirrors and substrate are orthogonal with each other, the X and Y-directional fringe periods and, respectively, can be expressed as follows,

$$g_x = g_y = \frac{\lambda}{\sqrt{2}\cos\theta}. \tag{5}$$

Figure 4 summarizes the line interference fringes to be generated by the interference between the two of those beams. As we can be seen in Figure 4, the interference fringes generated by beams 1 and 2, and beams 1 and 3 are parallel to Y-axis and X-axis, respectively, as shown in Figure 4a,b. Obviously, these two interference fringes are perpendicular to each other. The interference fringes generated by beam 2 and beam 3 are  $45^\circ$  with respect to X-axis as shown in Figure 4c. The interference fringe is enough for fabricating a 2D-orthogonal grating with the interference fringes generated by beams 1 and 2 and beams 1 and 3. Consequently, it is necessary to eliminate the interference of beam 2 and beam 3.



**Figure 4.** Interference fringe patterns generated by two of beam 1, 2, and 3. (a) beams 1 and 2; (b) beams 1 and 3; and (c) beams 2 and 3.

### 2.3. Optimum Combination of Initial Polarization

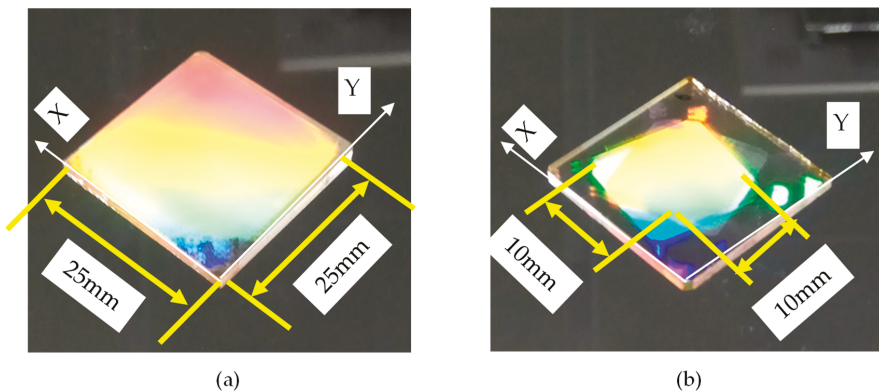
Figure 4 clearly illustrates the grating stripe direction. In previous research [13,16–18], effects are paid on elimination/reduction of Stripe 3 by using polarization modulation and experimental results verified that these proposals are effective in the reduction of Stripe 3 despite downsizing of the grating surface. But after deep investigation and careful calculation of these grating stripes with consideration of the exposure system configuration, it can be found that the influence of the Stripe 3 varies when not only the polarization but also the incident angle  $\theta$  are different. As that proposed in [16], denoting the degree of orthogonality between beams 2 and 3  $\gamma_{23}$ , and Stripe 3 might be neglected when  $\gamma_{23}$  is lower than a threshold or an experimentally obtained tolerance.

We employed the combination of three angles ( $\alpha_1$ ,  $\alpha_2$ , and  $\alpha_3$ ) to represent the combination of the initial orientation of polarization. Counter-clockwise and clockwise rotation were recorded as positive and negative angle, respectively. With the strict method to modulate the initial polarization status based on the theory of three-dimensional polarization ray-tracing calculus [16,19], the lowest  $\gamma_{23}$  of approximately 0.048 can be achieved when the incidence angle  $\theta = 71.79^\circ$  corresponding to a grating period of  $1 \mu\text{m}$  under a polarization status of polarization modulated three beams (direct beam, X-beam, and Y-beam: s,  $-33.5^\circ$ ,  $33.5^\circ$ , respectively). On the other hand, under such an incident angle, without polarization modulation, i.e., three beams polarization status (s, s, and s), the values  $\gamma_{23}$  of are 0.06 and 0.08 for metal-type mirror and dielectric-type mirror, respectively. It can be seen that this difference is not significant. It can be expected that, even without polarization, the Stripe 3 will not influence the grating uniformity so much. These findings are fundamental to this research.

Unfortunately, however, it should be noted that on the premise of the incident beam with linear polarization states, there does not exist a combination of  $\alpha_2$  and  $\alpha_3$  that could completely prevent the interference of beam 2 and beam 3, i.e.,  $\gamma_{23} = 0$ . For a near zero  $\gamma_{23}$ , polarization modulation of the incident beam might be required, this will be discussed in the future.

### 3. Experiment and Results

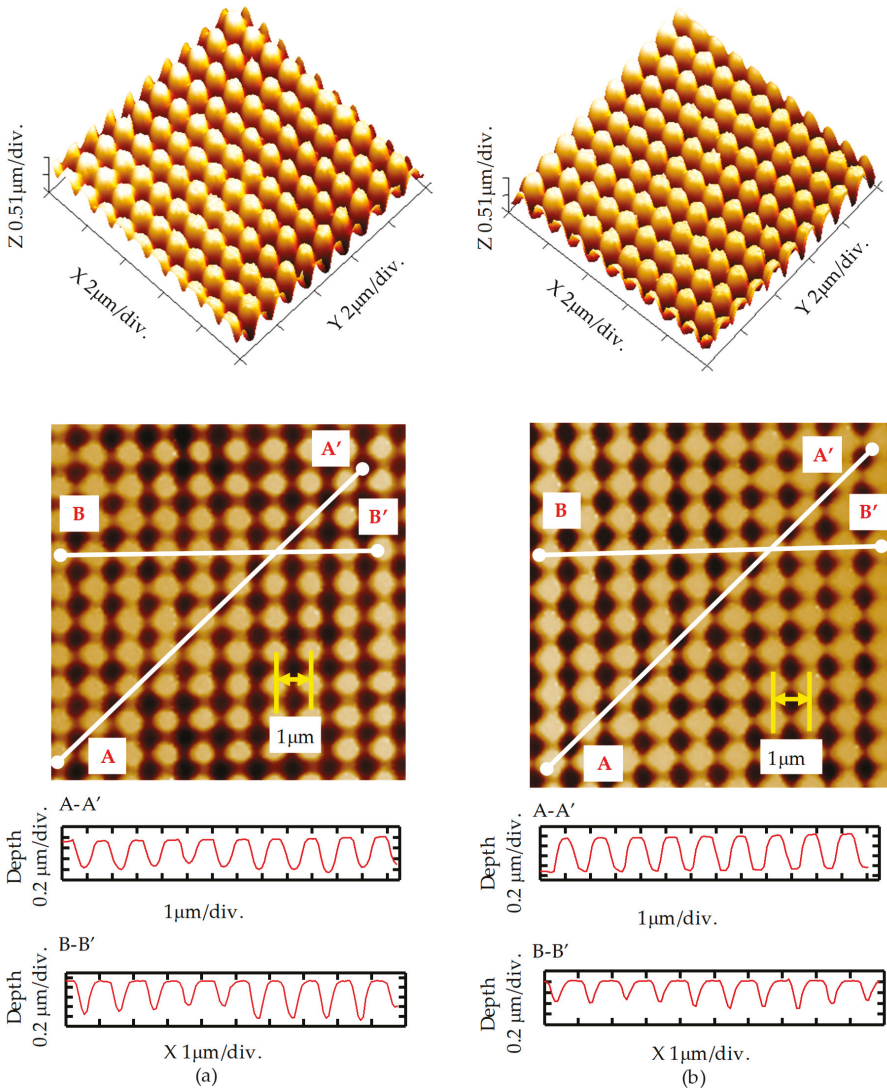
According to the abovementioned theories about polarization modulation, we used half-wavelength plates to realize the optimal combination of initial polarization status. We used a linear polarizer to measure the initial orientation of polarization of the laser source, and use a half-wavelength plate to modulate it with 0 degrees; we marked it as s. We employed two HWPS, which is half-wavelength plates, to alter the orientation of polarization of beams 2 and 3. However, the basic requirement is that optical elements after the collimation lens are as few as possible, to keep a good wavefront and a relatively large effective exposure area. To a certain extent, if the degree of orthogonality is small enough as mentioned above, we can ignore the interference between beam 2 and beam 3. Therefore, experiments with half-wavelength plates were carried out to verify the effects of polarization modulation. Meanwhile, for the purpose of obtaining maximum fabrication area and simplification of the optical system, we conducted a bold attempt that abandoned employing the two half-wavelength plates. It is worth noting that the reflected beam had different polarization status with the change of incident angle according to Fresnel's theories. For the sake of simplification, all the experiments were carried out under the constant condition of incident angle  $\theta = 71.8^\circ$ . The fabricated 2D gratings could have the same grating pitch at two orthogonal directions ( $g_{12} = g_{13} = 1 \mu\text{m}$ ). Figure 5 shows the fabricated gratings without and with polarization, i.e., under polarization state of (s, s, s) and (s,  $-33.5^\circ$ ,  $33.5^\circ$ ), respectively. It can be seen that the area of 2D grating fabricated without polarization status, i.e., initial polarization of the laser source (s, s, s) is larger than the modulated polarization status (s,  $-33.5^\circ$ ,  $33.5^\circ$ ). The area of the 2D grating fabricated with the initial polarization status (s, s, s) and (s,  $-33.5^\circ$ ,  $33.5^\circ$ ) are  $25 \text{ mm} \times 25 \text{ mm}$  and  $10 \text{ mm} \times 10 \text{ mm}$ , respectively. The enlarged grating area will be beneficial for grating interferometry application.



**Figure 5.** Photographs of the fabricated 2D gratings under the initial polarization status of (a) (s, s, s) under conditions of exposure time 30 s, development time 8 s and (b) (s,  $-33.5^\circ$ ,  $33.5^\circ$ ) under exposure time 30 s, development time 8 s.

The fabricated grating shapes were then evaluated by using an atomic force microscope (AFM). The AFM (Bruker Dimension Icon) worked on a peak force tapping mode with a height sensor data density of  $128 \times 128$  of an AFM image with a size of  $10 \mu\text{m} \times 10 \mu\text{m}$ . The grating surface material is photoresist, and AFM tip material is SiC and the shape aspect ratio is 1. The peak force setpoint and amplitude are 0.05 and 300 nm, respectively. Figure 6a,b show the images of the fabricated 2D

gratings under the initial polarization status of (s, s, s) and (s,  $-33.5^\circ$ ,  $33.5^\circ$ ), respectively. First of all, it can be seen that the grating shapes under these two polarization status are all regularly uniform, a near pillar shape, and no obvious elliptical shape as illustrated in [13] occurred. More than ten samples have been fabricated and results were highly consistent. Furthermore, the grating periods were evaluated as shown in Figure 6, B-B'. The grating periods are all consistent with the designed value of 1  $\mu\text{m}$ . It should be noted that, the period consistence is more importance than the absolute grating period value. From the A-A' sectional data, we can see that the grating depth are about 500 nm, (s, s, s) is slightly lower than that of the (s,  $-33.5^\circ$ ,  $33.5^\circ$ ), which are also consistent with the grating shape demand.



**Figure 6.** Atomic force microscopy (AFM) images and grating shape details of the fabricated 2D gratings under the initial polarization status of (a) (s, s, s) and (b) modulated polarization status (s,  $-33.5^\circ$ ,  $33.5^\circ$ ).

From Figures 5 and 6, it can be concluded that the grating shape and area of these two polarization status are all consistent with the preliminarily designed value. However, before it can be employed to the grating interferometry system, grating diffraction efficiency that influencing the inference signals are required to be evaluated. An experimental setup, as shown in Figure 7a, was constructed and the measurement point distribution is shown in Figure 7b. As for the real interferometry system, the incident beam was polarized to be circular and pass through the fabricated grating. After the diffraction of the grating, four first-order diffracted lights appear on the observation screen. The power of each diffracted light can be read by the power meter, which was then divided by the output power of the laser source, thus we can get the diffraction efficiency. The measurement results were summarized by Tables 1 and 2. Each diffraction efficiency is larger than 6%, and all the sum efficiency of all the first diffraction beam are larger than 50%. For the uniformity of the +1 and the -1 diffraction beams along the X- and Y-directions, the initial polarization is better than the modulated polarization, this may be because the intensity was redistributed by the two QWPs. The diffraction difference between the X- and the Y-direction is mainly due to the unideal intensity uniformity within the plane wavefront.

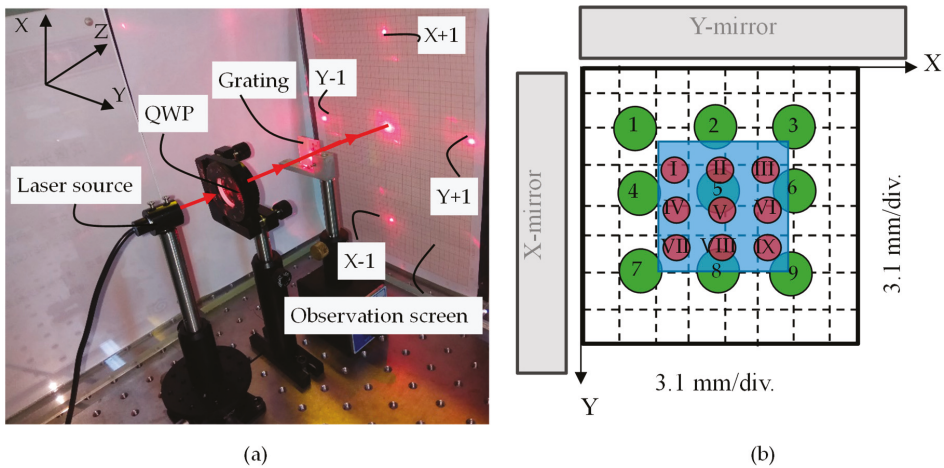


Figure 7. Diffraction efficiency testing setup (a) and measurement points on the grating surface (b).

Table 1. Diffraction efficiency testing result of the grating fabricated without polarization.

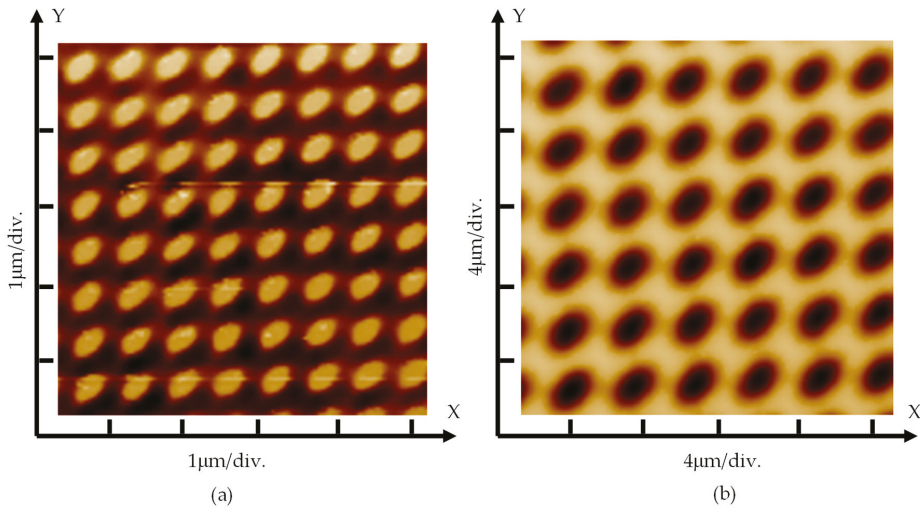
Points	1	2	3	4	5	6	7	8	9
X + 1	0.21	0.18	0.17	0.22	0.19	0.19	0.22	0.18	0.18
X - 1	0.19	0.16	0.16	0.21	0.18	0.18	0.21	0.17	0.17
Y + 1	0.13	0.14	0.12	0.09	0.13	0.13	0.07	0.12	0.13
Y - 1	0.13	0.14	0.13	0.08	0.13	0.13	0.06	0.13	0.13

Table 2. Diffraction efficiency testing result of the grating fabricated with polarization.

Points	I	II	III	IV	V	VI	VII	VIII	IX
X + 1	0.06	0.05	0.06	0.09	0.06	0.09	0.14	0.13	0.12
X - 1	0.12	0.08	0.10	0.17	0.11	0.14	0.22	0.19	0.18
Y + 1	0.22	0.23	0.22	0.12	0.20	0.17	0.07	0.13	0.14
Y - 1	0.18	0.22	0.19	0.07	0.17	0.13	0.01	0.09	0.12

For comparison, we conducted experiments at the incident angle  $\theta$  of  $60^\circ$  and  $85^\circ$  without polarization modulation, i.e., (s, s, s), corresponding to the grating periods of 625 nm and 3586 nm, respectively. Under such polarization and incident angle, the  $\gamma_{23}$  is larger than that of  $\theta$  of  $71.8^\circ$ .

Figure 8 shows the AFM images. This result reveals that the groove of 2D gratings deviated from square and was more inclined to elliptical shape, which cannot meet the interferometry application. Therefore, it is necessary for fabricating 2D grating to operate HWP's at  $60^\circ$  and  $85^\circ$ , as well as other angles. The specific values of modulation need to be calculated based on the three-dimensional polarization ray-tracing calculus [20].



**Figure 8.** AFM images of fabricated grating without polarization under different incident angle  $\theta$  shown in Figure 2: (a)  $60^\circ$  and (b)  $85^\circ$ .

#### 4. Discussion and Conclusions

We have designed and assembled the orthogonal type two-axis Lloyd's mirror interferometer as the exposure system. Apart from removing of the unexpected beams 4 and 5, the grating shape uniformity is theoretically evaluated in terms of an undesired interference between beams 2 and 3 by an index of degree of orthogonality.

Experimental results reveal that the difference is not significant for the fabricated 2D gratings under the original polarization status (s, s, s) and the optimal combination polarization status (s,  $-33.5^\circ$ ,  $33.5^\circ$ ) under a specific incident angle corresponding to a  $1\ \mu\text{m}$  grating period in both X and Y directions. The degree of orthogonality of beams 2 and 3 that fall in the same amounts level in the above two cases is the main reason. The degree of orthogonality of beam 2 and 3 are 0.07 and 0.048, respectively. The fabrication result under polarization was consistent with that shown in previous research, where similar initial polarization states (s,  $-45^\circ$ ,  $45^\circ$ ) were employed [18]. However, the grating area would be greatly enlarged without polarization as shown in Figure 5a compared with all the simulations and experiments shown in [18]; the optical system complexity was greatly reduced compared with the present technique, TBTE [5,10].

It is also confirmed that, under the condition of incident linear polarized light, the combination of initial polarization status of beams 2 and 3 corresponding to orthogonal polarization state in the orthogonal two-axis Lloyd's interferometer does not exist. The microstructures of the fabricated gratings are all consistent with the designed value. However the diffraction efficiency uniformity, both in a relative large area and the four diffraction beams on a point, require improvement.

In summary, polarization modulation technique of the LIL system can bring a good grating shape despite grating area downsizing, while, without polarization, i.e., usage of the initial polarization under some specific angle that can reach a similar grating performance, with a reduced system complexity and an enlarge grating area.

**Author Contributions:** Conceptualization, X.L.; Methodology, X.L. and H.L.; Software, H.L.; Validation, X.L., Q.Z. and K.N.; Resources, Q.Z., G.W., and X.W.; Data Curation, H.L.; Writing—original draft preparation, X.L., H.L.; Writing—review and editing, X.L.; Visualization, X.L.; Supervision, X.L.; Project Administration, Q.Z.; Funding Acquisition, G.W. and K.N.

**Funding:** This research was supported by the National Natural Science Foundation of China under Project No. 51427805, the Youth Funding of Shenzhen Graduate of Tsinghua University with Grant No. QN20180003, Shenzhen Fundamental Research Funding Grant No. JCYJ20170817160808432 and Grant No. JCYJ20160531195459678, and the National Key Research and Development Program under Grant No. 2016YFF0100704.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. *Two-Coordinate Incremental Encoder PP 281R*; KGM Series Catalogue; Heidenhain GmbH: Traunreut, Germany, 2018.
2. Gao, W. *Precision Nanometrology: Sensors and Measuring Systems for Nonmanufacturing*; Springer: London, UK, 2010; ISBN 78-1-84996-253-7.
3. Gao, W.; Araki, T.; Kiyono, S.; Okazaki, Y.; Yamanaka, M. Precision nano-fabrication and evaluation of a large area sinusoidal grid surface for a surface encoder. *Precis. Eng.* **2003**, *27*, 289–298. [[CrossRef](#)]
4. Gao, W.; Kimura, A. A three-axis displacement sensor with nanometric resolution. *Ann. CIRP* **2007**, *56*, 529–532. [[CrossRef](#)]
5. Kimura, A.; Gao, W.; Kim, W.J.; Hosono, K.; Shimizu, Y.; Shi, L.; Zeng, L.J. A sub-nanometric three-axis surface encoder with short-period planar gratings for stage motion measurement. *Precis. Eng.* **2012**, *36*, 576–585. [[CrossRef](#)]
6. Lee, C.; Kim, G.; Lee, S. Design and construction of a single unit multi-function optical encoder for a six-degree-of-freedom motion error measurement in an ultraprecision linear stage. *Meas. Sci. Technol.* **2011**, *22*, 105901. [[CrossRef](#)]
7. Huang, H.; Liu, C.; Jywe, W.; Wang, M.; Fang, T. Development of a three-degree-of-freedom laser linear encoder for error measurement of a high precision stage. *Rev. Sci. Instrum.* **2007**, *78*, 066103. [[CrossRef](#)] [[PubMed](#)]
8. Kao, C.; Lu, S.; Shen, H.; Fan, K. Diffractive laser encoder with a grating in Littrow Configuration. *Jpn. J. Appl. Phys.* **2008**, *47*, 1833–1837. [[CrossRef](#)]
9. Li, X.H.; Gao, W.; Muto, H.; Shimizu, Y.; Ito, S.; Dian, S.Y. A six-degree-of-freedom surface encoder for precision positioning of a planar motion stage. *Precis. Eng.* **2013**, *37*, 771–781. [[CrossRef](#)]
10. Byun, I.J.; Kim, J.W. Cost-effective Laser interference lithography using a 405 nm AlInGaN semiconductor laser. *J. Micromech. Microeng.* **2010**, *20*, 055024. [[CrossRef](#)]
11. Brueck, S.R.J. Optical and Interferometric Lithography-Nanotechnology Enablers. *Proc. IEEE* **2005**, *93*, 1704–1721. [[CrossRef](#)]
12. Chua, J.K.; Murkeshan, V.M. Patterning of Two-dimensional Nanoscale Features Using Grating-based Multiple Beams Interference Lithography. *Phys. Scr.* **2009**, *80*, 015401. [[CrossRef](#)]
13. Vala, M.; Homola, J. Flexible method based on four-beam interference lithography for fabrication of large areas of perfectly periodic plasmonic arrays. *Opt. Express* **2014**, *22*, 18778–18789. [[CrossRef](#)] [[PubMed](#)]
14. Boor, J.D.; Geyer, N.; Gösele, U.; Schmidt, V. Three-beam interference lithography: Upgrading a Lloyd’s interferometer for single-exposure hexagonal patterning. *Opt. Lett.* **2009**, *34*, 1783–1785. [[CrossRef](#)] [[PubMed](#)]
15. Li, X.H.; Gao, W.; Shimizu, Y.; Ito, S. A two-axis Lloyd’s mirror interferometer for fabrication of two-dimensional diffraction gratings. *Manuf. Technol.* **2014**, *63*, 461–464. [[CrossRef](#)]
16. Zhou, H.Y.; Zeng, L.J. Method to fabricate orthogonal crossed gratings based on a dual Lloyd’s mirror interferometer. *Opt. Commun.* **2016**, *360*, 68–72. [[CrossRef](#)]
17. Shimizu, Y.; Aihara, R.; Ren, Z.; Chen, Y.L.; Ito, S.; Gao, W. Influences of misalignment errors of optical components in an orthogonal two-axis Lloyd’s mirror interferometer. *Opt. Express* **2016**, *24*, 27521–27535. [[CrossRef](#)] [[PubMed](#)]
18. Chen, X.G.; Ren, Z.W.; Shimizu, Y.; Chen, Y.L.; Gao, W. Optimal polarization modulation for orthogonal two-axis Lloyd’s mirror interference lithography. *Opt. Express* **2017**, *25*, 22237–22252. [[CrossRef](#)] [[PubMed](#)]



19. Yun, G.; Crabtree, K.; Chipman, R.A. Three-dimensional polarization ray-tracing calculus I: Definition and diattenuation. *Appl. Opt.* **2011**, *50*, 2855–2865. [[CrossRef](#)] [[PubMed](#)]
20. Chen, X.G.; Shimizu, Y.; Chen, C.; Chen, Y.L.; Gao, W. Generalized method for probing ideal initial polarization states in multi-beam Lloyd's mirror interference lithography of 2D scale gratings. *J. Vac. Sci. Technol. B* **2017**, *36*, 021601. [[CrossRef](#)]



© 2018 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).

Article

# Simultaneous Measurement Method and Error Analysis of the Six Degrees-of-Freedom Motion Errors of a Rotary Axis

Chuanchen Bao, Qibo Feng <sup>\*</sup> and Jiakun Li

MoE Key Laboratory of Luminescence and Optical Information, Beijing Jiaotong University, Beijing 100044, China; 15118423@bjtu.edu.cn (C.B.); jkli@bjtu.edu.cn (J.L.)

\* Correspondence: qbfeng@bjtu.edu.cn; Tel.: +86-10-51685558

Received: 28 September 2018; Accepted: 7 November 2018; Published: 13 November 2018



**Abstract:** Error measurement of a rotary axis is the key to error compensation and to improving motion accuracy. However, only a few instruments can measure all the motion errors of a rotary axis. In this paper, a device based on laser collimation and laser interferometry was introduced for simultaneous measurement of all six degrees-of-freedom motion errors of a rotary axis. Synchronous rotation of the target and reference rotary axes was achieved by developing a proportional–integral–derivative algorithm. An error model for the measuring device was established using a homogeneous transformation matrix. The influences of installation errors, manufacturing errors, and error crosstalk were studied in detail, and compensation methods for them were proposed. After compensation, the repeatability of axial and radial motion errors was significantly improved. The repeatability values of angular positioning error and of tilt motion error around the  $y$  axis and  $x$  axis were 28.0", 2.8", and 3.9". The repeatability values of translational motion errors were less than 2.8  $\mu\text{m}$ . The comparison experiments show that the comparison errors of angular positioning error and tilt motion error around the  $y$  axis were 2.3" and 2.9", respectively. These results demonstrate the effectiveness of our method and the error compensation model.

**Keywords:** simultaneous measurement; six degrees-of-freedom errors; rotary axis; error model

## 1. Introduction

Rotary axes are widely applied in computer numerical control machining, robotics, aerospace, and other fields. Rotary axes with inadequate motion accuracy directly reduce the working accuracy of the related precision machines. The first and most important step to increase the accuracy of the axes and decrease the influence of the errors is to quickly and accurately measure the motion errors of the rotary axes.

A complete description of the motion errors of a rotary axis requires six physical quantities [1]: the angular positioning error ( $E_{CC}$ ), the tilt motion error around the  $y$  axis ( $E_{BC}$ ), the tilt motion error around the  $x$  axis ( $E_{AC}$ ), the radial motion error along the  $x$  axis ( $E_{XC}$ ), the radial motion error along the  $y$  axis ( $E_{YC}$ ), and the axial motion error ( $E_{ZC}$ ). In order to make the symbols convenient for the subsequent calculation of the coordinate transformation, we use the simpler symbols  $\varepsilon_z$ ,  $\varepsilon_y$ ,  $\varepsilon_x$ ,  $\delta_x$ ,  $\delta_y$ , and  $\delta_z$  instead of  $E_{CC}$ ,  $E_{BC}$ ,  $E_{AC}$ ,  $E_{XC}$ ,  $E_{YC}$ , and  $E_{ZC}$  in the calculation. The real motion trajectory of the axis can be reconstructed using these six motion errors, and the application precision of the related mechanical equipment can be improved by compensating for the errors. Therefore, the key problem to be solved is how to obtain the six degrees-of-freedom (DOF) motion errors of the axis with high precision and high efficiency. At present, the simultaneous measurement methods of multi-DOF

motion errors are mainly based on laser properties [2–4], material profile [5], trajectory analysis [6], and other factors [7].

Chen’s apparatus, using three lasers, is a rare method that can measure the six DOF motion errors of a rotary axis simultaneously and directly [8]. While its structure is complicated, H. Schwenke used the laser tracker [9] to study the measurement of the six DOF errors. Ibaraki used the R-test [10,11] method to identify the six errors. However, most of these indirect measurement methods need to determine each error using complex function-fitting algorithms. Unlike these existing methods, our method can measure the six DOF errors directly at the same time. It also has the advantages of easy installation and high measurement efficiency. Two important methods—a servo-tracking strategy and error model analysis—are used to realize and improve the functions of the measuring device.

In this paper, a device for the simultaneous measurement of the six DOF errors of a rotary axis is introduced, and a method to improve the measurement accuracy is proposed based on our previous research for five DOF errors [2,12]. The automatic measurement is achieved by using the servo-tracking strategy. It is also used to overcome the difficulty of the beam needing to be uninterrupted during the interferometric process. A complete model for measuring all six DOF motion errors is proposed. The factors influencing the measurement result are analyzed, and a compensation method is proposed. In Section 2, the measurement principle of our device is briefly introduced. The servo-tracking strategy of the reference rotary axis is described in detail. An error model is established using the homogeneous transformation matrix (HTM) [13,14]. In Section 3, we describe a series of experiments, which were carried out to simultaneously measure the six DOF errors of a rotary axis. The repeatability and comparison results were obtained after compensation and demonstrate the effectiveness of our method and the error compensation model.

## 2. Principle and Model

### 2.1. Measurement Principle

The measuring device includes a laser and fiber coupling unit, a measurement unit, an error-sensitive unit, and a reference rotary unit [2]. The measurement unit and the error-sensitive unit constitute the main optical path. The optical path is divided into a collimation measurement part and an interferometry measurement part. The collimation measurement part consists of two quadrant detectors (QDs) and two position-sensitive detectors (PSDs). As shown in Figure 1, QD1 receives the light reflected back from a retro-reflector (RR3) to detect the radial motion error along the  $y$  axis and the axial motion error of the target rotary axis. QD2 is similar to QD1, and the difference between the two QDs in the vertical direction corresponds to the tilt motion error around the  $x$  axis. PSD1 receives the light reflected from a mirror (a beam-splitting film, BS3) and detects the angular error around the  $z$  axis and tilt motion error around the  $y$  axis. The target rotary axis rotates a nominal angle ( $\theta_{\text{nominal}}$ ), and the reference rotary axis rotates in the opposite direction ( $\theta_{\text{encode}}$ ). The angular positioning error is determined by  $Y_{\text{PSD1}}$  (the  $y$  direction of PSD1),  $\theta_{\text{encode}}$ , and  $\theta_{\text{nominal}}$ . PSD2 detects the laser beam drift [2]. The interference part consists of the reflected light by RR1 (as the reference light) and the reflected light by RR3 (as the signal light). Detector D1 records the changes in the interference signals, which correspond to the radial motion error along the  $x$  axis. Combined with collimation and interferometry, the measuring device is capable of the simultaneous measurement of all six of the DOF motion errors of a rotary axis.

To ensure the continuity of the interference signal, the reference rotary axis needs to rotate synchronously with the target rotary axis. The proportional–integral–derivative (PID) algorithm was developed to achieve the servo-tracking strategy, and this process is shown in Figure 2. The start and end thresholds are based on the horizontal data of QD1. After the target rotary axis starts to rotate, the light incident to QD1 will deviate from its initial position. When the light reaches the start threshold, the reference rotary axis starts to rotate in the opposite direction. According to the data in the horizontal direction of QD1, the PID algorithm continuously outputs the speed commands to the

reference axis. By constantly adjusting the speed, the reference axis rotates at a relatively stable speed with the target axis. When approaching the measuring position, the target axis rapidly decelerates to a stop. When the QD data reaches the end threshold, the computer stops the reference axis, and the servo-tracking process is completed. In order to eliminate errors introduced by bi-directional rotation, the reference axis is set to rotate in only one direction. Thus, if the speed calculated using PID is negative, the speed command will be zero.

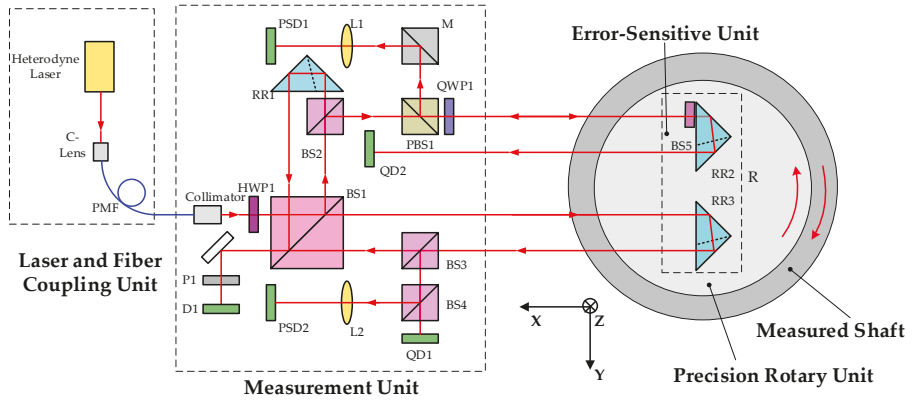


Figure 1. Measurement principle.

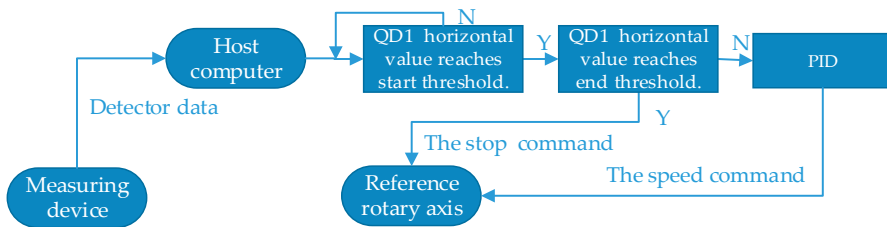


Figure 2. Schematic diagram of the tracking process.

In order to shorten the control delay of the host computer, separate threads for the data acquisition from the detectors and PID algorithm were opened. By changing the end threshold, the stop position of the spot on QD1 is controlled; thus, the linear range of the QD detector can be fully utilized. The proportion coefficient of PID determines the stable state of the motion. An excessively small proportion coefficient will cause interference signal interruption, while an excessively large proportion coefficient will cause unstable movement of the reference axis, which is not suitable for measurement. The differential coefficient is sensitive to a sudden change. Thus, it can cope with excessive acceleration when the target axis starts or stops. Through the servo-tracking strategy, the device not only realizes interferometry but also reduces the time of a full-circle measurement from about 30 min [12] to less than 15 min.

## 2.2. Error Model Establishment and Analysis

Various errors will be inevitably introduced in the measurement process. The data from the detectors (QD1, QD2, PSD1, PSD2, and D1) are affected by the six DOF motion errors, installation errors, manufacturing errors, etc. An error model was established to improve the accuracy of our measuring device. The error sources and influences on the measurement results were analyzed using the model, and the corresponding compensation methods were studied.

The model was established in a three-dimensional space. The measuring device mainly includes the target rotary axis A, the reference rotary axis B, the error-sensitive unit C, and the measurement unit. The coordinate systems “1” to “4”, fixed to the corresponding parts and moving with them, were established, respectively, as shown in Figure 3. The world coordinate system “0” was established, and the initial state of coordinate system, “1”, coincides with the coordinate system “0”. The HTM  $T_n^m$  between the adjacent coordinate systems was established. This matrix can be used to simulate the motion of rotary axes and to realize the coordinate transformation [12].  $T_n^m$  represents the transformation matrix from the coordinate system  $m$  to the coordinate system  $n$ . We assume that the coordinate system “1” is obtained after a series of motions of the coordinate system “0”. The order of the motions is as follows: first translate  $\delta_x$ ,  $\delta_y$ , and  $\delta_z$ ; then rotate  $\varepsilon_x$ ,  $\varepsilon_y$ , and  $(\theta + \varepsilon_z)$ . This process defines the six DOF errors. The coefficient matrix  $LC^4$  of the output beam equation is defined in the coordinate system “4”, and the reflection matrix  $R$  is defined in the coordinate system “3” [12]. According to Equation (1), the coefficient matrix  $LoutC^4$  of the reflected light can be calculated using the HTMs and the reflection matrix. Based on the structure of the measuring device and the ray tracing principle, the expressions of the six DOF motion errors of the rotary axis can be obtained.

$$LoutC^4 = LC^4 T_4^0 T_0^1 T_1^2 T_2^3 R T_3^2 T_2^1 T_1^0 T_0^4 \tag{1}$$

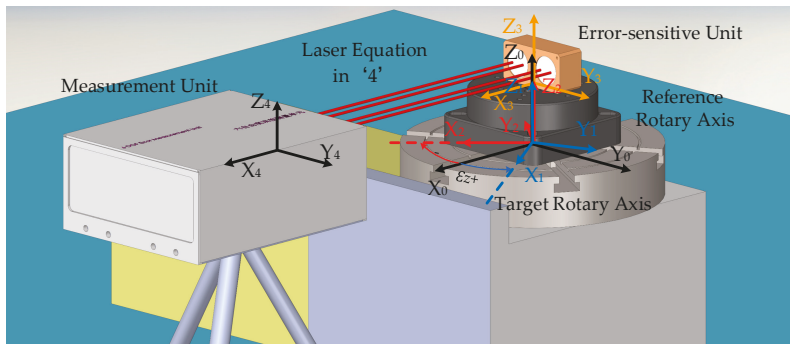


Figure 3. Schematic diagram of the experimental setup and coordinate systems.

In the error model, angular errors are generally expressed in  $\varepsilon$ , and translational errors are expressed in  $\delta$ . Different subscripts are used as a distinction. The specific variables and their naming rules are shown in Table 1.

Table 1. Error variables and the naming rules.

Error Variables	Naming Rules
Motion error	The subscript is the name of the coordinate axis (i.e., $\varepsilon_z$ ).
Installation error	The name of the axis and two lowercase letters ( $a$ , $b$ , and $c$ ) are used as the subscript (i.e., $\varepsilon_{yab}$ denotes the angular error around the $y$ axis between A and B). The subscript of the measurement unit is $L$ (i.e., $\varepsilon_{yL}$ ).
Manufacturing error	The subscript is the name of the axis and the letter $c$ (i.e., $\varepsilon_{yc}$ ).
Laser beam drift error	The subscript is the name of the axis and the letter $t$ (i.e., $\varepsilon_{yt}$ ).

The data in the horizontal and vertical directions of the detectors are represented by  $Y$  and  $Z$ , respectively. PSDs and QDs are distinguished by the subscript. For example,  $Y_{PSD1}$  is the data in the horizontal direction of PSD1. ( $\theta = 0$ ) is added, which represents the initial value of the measurement. For example,  $Y_{PSD1} (\theta = 0)$  is the initial value in the vertical direction of PSD1. The naming of other

parameters, such as structural parameters and optical parameters of the measuring device, is shown in Table 2.

**Table 2.** The naming of parameters.

Parameters	Naming
Height of the reference axis	$H_a$
Vertex coordinate of the retro-reflector (in “3”)	$(OC_{1x}, OC_{1y}, OC_{1z})$
Vertex distance of two retro-reflectors	$D = OC_{2y} - OC_{1y}$
Center point coordinate of the bottom edge of the measurement unit’s front surface (in “0”)	$(P_x, P_y, P_z)$
The first laser output position coordinates on the measurement unit (in “4”)	$(P_{1x}, P_{1y}, P_{1z})$
Center coordinates of the QD1 photosensitive surface (in “4”)	$(QDC_{1x}^4, QDC_{1y}^4, QDC_{1z}^4)$
Center coordinates of the PSD1 photosensitive surface (in “4”)	$(PSDC_{1x}^4, PSDC_{1y}^4, PSDC_{1z}^4)$
Focal length of the lens	$f$
Refractive index of the retro-reflector glass	$n$

According to the principle of laser collimation, all the DOF motion errors can be obtained—except for the radial motion error along the  $x$  axis—with the two types of photoelectric detectors [12].

The angular positioning error is calculated using the formula:

$$\epsilon_z = -(Y_{PSD1}(\theta = 0) - Y_{PSD1})/2f + \theta_{encode} - \theta_{nominal}, \tag{2}$$

where  $Y_{PSD1}(\theta = 0) = -f(-2\epsilon_{zca} + 2\epsilon_{zL})$ .

The tilt motion error around the  $y$  axis is calculated using the formula:

$$\epsilon_y = -(Z_{PSD1} - Z_{PSD1}(\theta = 0))/2f - \epsilon_{xab} \sin \theta - \epsilon_{yab} \cos \theta + \epsilon_{yab}, \tag{3}$$

where  $Z_{PSD1}(\theta = 0) = -2f(\epsilon_{yca} + \epsilon_{yab} - \epsilon_{yL})$ .

The tilt motion error around the  $x$  axis is calculated using the formula:

$$\epsilon_x = -(\Delta - \Delta(\theta = 0))/2D - \epsilon_{xab} \cos \theta + \epsilon_{yab} \sin \theta + \epsilon_{xab}, \tag{4}$$

where  $\Delta = Z_{QD1} - Z_{QD2}$  and

$$\Delta(\theta = 0) = (Z_{QD1} - Z_{QD2})(\theta = 0) = 2D(-\epsilon_{xca} - \epsilon_{xab} + \epsilon_{xL}) + (-nQDC_{1x}^4 + nQDC_{2x}^4)\epsilon_{yc} + (D - P_{1y} + P_{2y})\epsilon_{xc}.$$

According to Equation (2), the only influence on the angular positioning error is constant installation errors, which can be eliminated by subtracting the initial value in the first position of the rotary axis. As described in Equations (3) and (4), the two tilt motion errors are affected by the constant installation errors. The influence of the trigonometric form on the tilt motion errors is from the angle of the coaxiality deviation between the reference and target axes ( $\epsilon_{xab}$  and  $\epsilon_{yab}$ ). This influence can be compensated for by function fitting. In addition, the manufacturing error ( $\epsilon_{xc}$  and  $\epsilon_{yc}$ ) of the retro-reflector has a constant influence on the tilt error around the  $x$  axis, which can be eliminated by subtracting the initial value.

The radial motion error along the  $y$  axis can be calculated using the equation:

$$\begin{aligned} \delta_y &= Y_{QD}/2 - Y_{PSD}OC_{1x}/2nf - \Delta(OC_{1z} + H_a)/2D - \delta_{xab} \sin \theta - \delta_{yab} \cos \theta + \delta_{yab} \\ &- [Y_{QD}/2 - Y_{PSD}OC_{1x}/2nf - \Delta(OC_{1z} + H_a)/2D](\theta = 0) \\ &+ [\epsilon_{zt} - \epsilon_{zt}(\theta = 0)](-\frac{OC_x}{n} + 2P_x + QDC_x) \\ &+ \frac{(OC_{1z} + H_a)}{2D}(-QDC_{1x} + QDC_{2x})[-\epsilon_{yt} + \epsilon_{yt}(\theta = 0)] \end{aligned} \tag{5}$$

where

$$\begin{aligned} & [Y_{QD} - OC_{1x}Y_{PSD}/nf - \Delta(OC_{1z} + H_a)/D](\theta = 0) \\ & = 2\delta_{yca} + 2\varepsilon_{zL}P_x - 2\varepsilon_{xL}P_z + 2\delta_{yab} + 2H_a\varepsilon_{xca} + (-OC_{1z} - H_a - P_z + P_{1z})\varepsilon_{xc} \\ & - \varepsilon_{yc}(-nQDC_{1x}^4 + nQDC_{2x}^4)(OC_{1z} + H_a)/D \\ & + (nP_x - OC_{1x} + nQDC_{1x}^4)\varepsilon_{zc} - (D - P_{1y} + P_{2y})\varepsilon_{xc}(OC_{1z} + H_a)/D \end{aligned}$$

The axial motion error can be calculated using the equation:

$$\delta_z = Z_{QD}/2 - OC_{1x}Z_{PSD}/2nf + OC_{1y}\Delta/2D - [Z_{QD}/2 - OC_{1x}Z_{PSD}/2nf + OC_{1y}\Delta/2D](\theta = 0) - \frac{\varepsilon_{yt} - \varepsilon_{yt}(\theta=0)}{2} \left( -\frac{OC_y}{D}QDC_{1x} + \frac{OC_y}{D}QDC_{2x} - QDC_x - 2P_x + \frac{OC_x}{n} \right) \tag{6}$$

where

$$\begin{aligned} & [Z_{QD}/2 - OC_{1x}Z_{PSD}/2nf + OC_{1y}\Delta/2D](\theta = 0) = 2\delta_{zca} + 2\delta_{zab} + 2\varepsilon_{xL}P_y - 2\varepsilon_{yL}P_x \\ & + (-nP_x - nQDC_{1x}^4 + OC_{1x})\varepsilon_{yc} + (-P_y + OC_{1y} - P_{1y})\varepsilon_{xc} \\ & + \varepsilon_{yc}(-nQDC_{1x}^4 + nQDC_{2x}^4)OC_{1y}/D + (D - P_{1y} + P_{2y})\varepsilon_{xc}OC_{1y}/D \end{aligned}$$

As described in Equations (5) and (6), the manufacturing errors and some of the installation errors have constant influences on the radial and axial motion errors. These influences can be eliminated by subtracting the initial value. The measurement result of the radial motion error along the *y* axis is affected by the offset of the coaxiality deviation of the two rotary axes. This effect—the trigonometric terms in equations—can be compensated for using function fitting. In Equation (5),  $Y_{PSD}/2f$  is related to the angular positioning error, and  $\Delta/2D$  is related to the tilt motion error around the *x* axis. Thus, these two motion errors have influences on the measurement of the radial motion error along the *y* axis, which is the error crosstalk. Similarly, in Equation (6), two tilt motion errors have influences on the measurement of the axial motion error. All of the error crosstalk can be compensated for by the above model.

To measure the radial motion error along the *x* axis, the interference principle is used. Therefore, accurate analysis of the total optical path of the interference part is the key to describing the radial motion error along the *x* axis. The total optical path is divided into three parts: the optical path of the output light ( $L_o$ , yellow), the optical path in RR3 ( $L_i$ , red), and the optical path of the reflected light ( $L_r$ , blue), as shown in Figure 4. According to the equations of the output and reflected light, the incident and emitted positions of the lights on the measurement unit or the retro-reflector can be obtained, and the total optical path of the interference can be calculated.

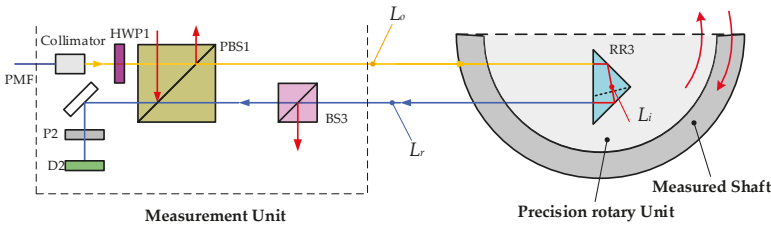


Figure 4. Optical path of the interference.

The optical path in RR3 is affected by the angular changes of the incident light, and the influences are second- and higher-order items, as shown in Equation (7).  $h$  is the height of the retro-reflector and  $i$  is the angle of incidence.

$$L_i = 2nh \left( 1 - \sin^2 i/n^2 \right)^{-1/2} \tag{7}$$

The angular motion errors have larger influences on the optical paths of the output light and reflected light. The influences include the first-order items. After ignoring the second- and higher-order items, the optical paths of the output light  $L_o$  and the reflected light  $L_r$  are:

$$\begin{aligned} L_o &= -B^3 + (P_y + P_{1y})(\varepsilon_{zca} - \varepsilon_{zL} + \varepsilon_z) \\ &\quad - (H_a - P_z - P_{1z})(-\varepsilon_{yca} - \varepsilon_y + \varepsilon_{yL} - \varepsilon_{yab} \cos \theta - \varepsilon_{xab} \sin \theta) \\ L_r &= -B^3 - (P_y + P_{1y} - 2OC_y)(\varepsilon_{zca} - \varepsilon_{zL} + \varepsilon_z) \\ &\quad - (P_z + P_{1z} - H_a - 2OC_z)(-\varepsilon_{yca} - \varepsilon_y + \varepsilon_{yL} - \varepsilon_{xab} \sin \theta - \varepsilon_{yab} \cos \theta) \end{aligned} \tag{8}$$

where

$$B^3 = \delta_{xca} + \delta_x + \delta_{xab} \cos \theta - \delta_{yab} \sin \theta + H_a(\varepsilon_{yab} \cos \theta + \varepsilon_{xab} \sin \theta + \varepsilon_y - \varepsilon_{yL}) - P_x - P_y \varepsilon_{zL} + P_z \varepsilon_{yL} + L_s$$

The total optical path can be expressed as:

$$\begin{aligned} L &= L_o + L_i + L_r = \\ &\quad -2(\delta_{xca} + \delta_x + \delta_{xab} \cos \theta - \delta_{yab} \sin \theta) + OC_y Y_{PSD} / f + (H_a + OC_z) Z_{PSD} / f \\ &\quad -2(-P_y \varepsilon_{zL} + P_z \varepsilon_{yL} - H_a \varepsilon_{yca}) - 2(-P_x + L_s - nOC_x) + OC_y \varepsilon_{zt} - (H_a + OC_z) \varepsilon_{yt} \end{aligned} \tag{9}$$

Therefore, the radial motion error along the  $x$  axis is:

$$\begin{aligned} \delta_x &= -(L - L(\theta = 0)) / 2 + OC_y (Y_{PSD} - Y_{PSD}(\theta = 0)) / 2f \\ &\quad + (H_a + OC_z)(Z_{PSD} - Z_{PSD}(\theta = 0)) / 2f - (\delta_{xab} \cos \theta - \delta_{yab} \sin \theta - \delta_{xab}) \\ &\quad + OC_y (\varepsilon_{zt} - \varepsilon_{zt}(\theta = 0)) / 2 - (H_a + OC_z) (\varepsilon_{yt} - \varepsilon_{yt}(\theta = 0)) / 2 \end{aligned} \tag{10}$$

where

$$\begin{aligned} L(\theta = 0) &= -2(\delta_{xca} + \delta_{xab}) + OC_y Y_{PSD}(\theta = 0) / f + (H_a + OC_z) Z_{PSD}(\theta = 0) / f \\ &\quad -2(-P_y \varepsilon_{zL} + P_z \varepsilon_{yL} - H_a \varepsilon_{yca}) - 2(-P_x + L_s - nOC_x) + OC_y \varepsilon_{zt}(\theta = 0) - (H_a + OC_z) \varepsilon_{yt}(\theta = 0) \end{aligned} \tag{11}$$

$L_s$  is the distance between the laser emission surface and the front surface of the measurement unit. The constant influence of some installation errors can be eliminated by subtracting the initial value in the first position of the rotary axis. The offset of the coaxiality deviation ( $\delta_{xab}$  and  $\delta_{yab}$ ) has an influence on the measurement of the radial motion error. The trigonometric items can be compensated for by function fitting. In Equation (10), two angular motion errors also influence the measurement of the radial motion error. This kind of error crosstalk can be compensated for by using Equation (10).

The complete theoretical analysis shows that the measurement results of the six DOF motion errors are affected by several errors. The compensation methods are given in the above theoretical model, which can effectively compensate for error crosstalk, installation errors, and manufacturing errors and thus improve measurement accuracy.

### 3. Experiment Results

#### 3.1. Experiment Conditions

Under laboratory conditions, the six DOF motion errors of a rotary axis were measured using our measuring device, as shown in Figure 5. The performance of the device was studied through experiments. The target rotary axis was the SKQ-12200 produced by KEOLEA. Its angular positioning error was 40", and the repeatability value was 20". The reference rotary axis was the ANT95-360-R produced by Aerotech. Its angular positioning error was 10", and the unidirectional repeatability value was 0.5". The repeatability value of the tilt motion errors was 3"; the repeatability values of the axial and radial motion errors were 0.5 and 1 μm, respectively. The measuring interval was 30°, and the rotation speed of the target rotary axis was about 0.55°/s. The experimental temperature was fixed



at  $25 \pm 1$  °C. The rotational velocity of the reference axis fluctuated around the velocity of the target rotary axis. The following errors were about 250".

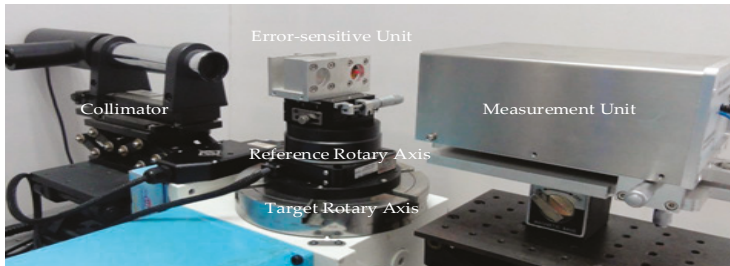


Figure 5. Measurement device.

### 3.2. Repeatability Experiment

The full-circle measurement of the six DOF motion errors of the target axis was repeated nine times. At the same position, half of the fluctuation range of the nine measurements was defined as the repeatable value of this position. The maximum value of all the 13 positions was selected as the repeatability value of each motion error measured by the device.

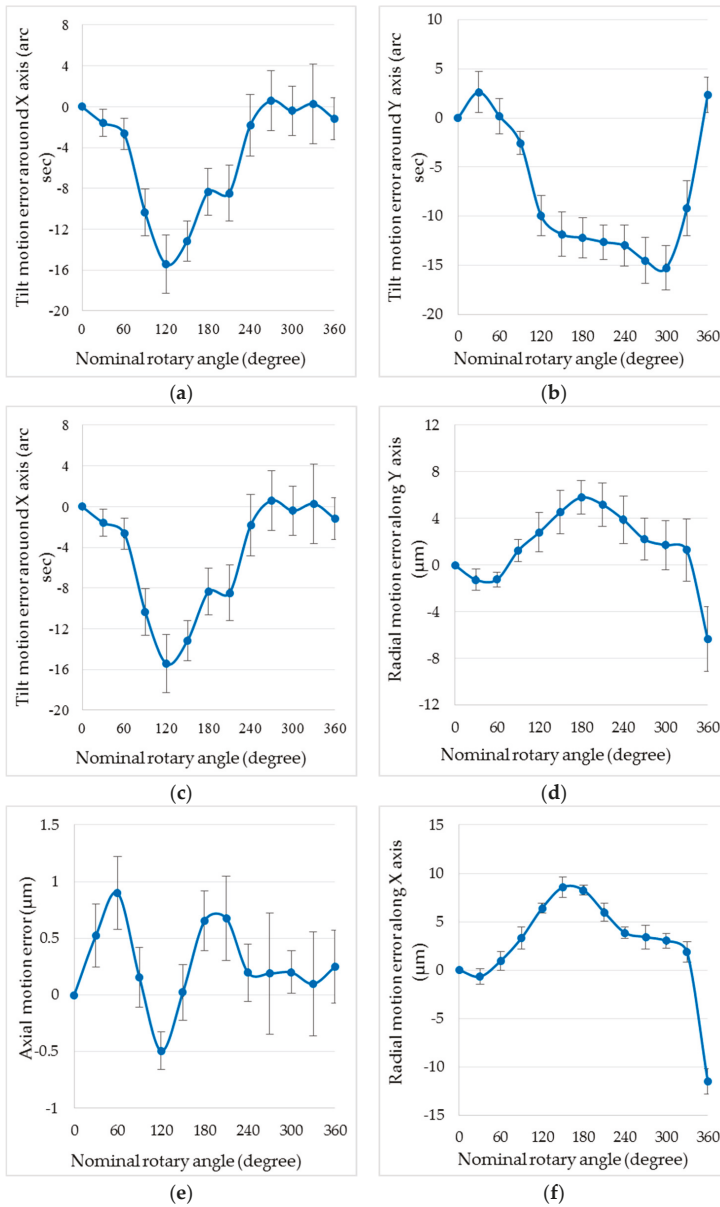
The original data consist of various errors, and the real motion errors of the target axis can be obtained through error compensation. First, the initial values of the first position were subtracted from all the original data. Thus, the constant influence of errors, such as installation errors and manufacturing errors, could be eliminated. Then, the error crosstalk was compensated for according to the above equations. Finally, the installation errors due to the coaxiality deviations of the two axes were compensated for by function fitting. The final results of the repeatability experiment are shown in Figure 6.

The repeatability value of the angular positioning error was 28.0". The repeatability values of tilt motion errors around the *y* axis and *x* axis were 2.8" and 3.9", respectively. The repeatability value of the radial motion error along the *y* axis was 2.8 μm. The repeatability value of axial motion error was 0.5 μm. The radial motion error along the *x* axis was 1.3 μm. As shown in Figure 6, the results at 0 and 360 degrees are not the same. There are mainly two reasons for this. One is that our device has repeatability values, which result in the noncoincidence of the measuring points. Our experimental results show that the maximum repeatability value of the angular motion error was 28.0". The other is that the measured axis itself has a large repeatability value. The positions of 0 degrees and 360 degrees are the same in theory, but they are not in fact. Due to the characteristics of the measured axis, some motion errors of the measured axis near the 0 point vary greatly. Thus, even a small angular positioning error will lead to a large change in motion error.

The error model helps us not only obtain error data much closer to the real value but also compensate for the influence of error crosstalk on the repeatability measurement. As shown in Table 3, the repeatability values of translational motion errors were significantly improved by removing the fluctuations introduced by other angular motion errors.

Table 3. Repeatability changes before and after crosstalk compensation.

Motion Errors	Repeatability after Subtracting the Initial Value	Repeatability after Compensating for the Crosstalk	Percentage Decline
Radial error along the <i>y</i> axis	4.1 μm	2.8 μm	32%
Axial error	0.8 μm	0.5 μm	38%
Radial error along the <i>x</i> axis	4.5 μm	1.3 μm	71%



**Figure 6.** The repeatability of six DOF motion errors: the (a) angular positioning error, (b) tilt motion error around the  $y$  axis, (c) tilt motion error around the  $x$  axis, (d) radial motion error along the  $y$  axis, (e) axial motion error, and (f) radial motion error along the  $x$  axis.

### 3.3. Comparison Experiment

Restricted by the experimental conditions, only the angular positioning error and the tilt error around the  $y$  axis measured by our device were compared to the results measured by a standard instrument. An autocollimator AC300 (the standard instrument was produced by AcroBeam) measured the target axis simultaneously with the measuring device. The max error of the standard instrument

is  $\pm 0.05''$  in the range of  $\pm 300''$  and the uncertainty of indication error is  $0.05''$ . The instrument was calibrated by the National Institute of Metrology in China. At the same position of measurement, the difference between our device and the standard instrument was considered the comparison error at this measurement position''.

There are six DOF installation errors of the autocollimator and six installation errors of the mirror, which are introduced in the comparison experiment. Because of the measurement principle of the autocollimator, the translation errors and the tilt errors around the  $x$  axis have little effect on the measurement and can be ignored. Only four angular installation errors have an influence on the measurement results. The influences are constant and can be compensated for by subtracting the value of the first measurement point. As shown in Figure 7, the comparison error of the angular positioning error was  $2.3''$ , and the comparison error of tilt motion error around the  $y$  axis was  $2.9''$ .

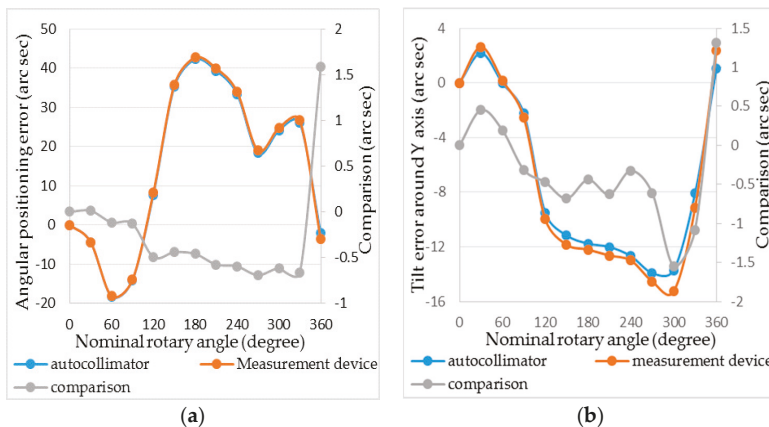


Figure 7. Comparison experiment results for the (a) angular positioning error and the (b) tilt motion error around the  $y$  axis.

The experiment results show that the repeatability of the device for measuring all motion errors except for the angular positioning error is good, as shown in Figure 6. We believe that the large repeatability value of the angular positioning error is due to the low positioning accuracy of the target rotary axis, which was verified by the comparison experiments. As shown in Figure 7, the comparison errors of both the angular positioning error and the tilt motion error around the  $y$  axis were small. Moreover, the validity of the model compensation was also verified by the experiment. The three repeatability values of the measuring device were significantly improved after compensation, as shown in Table 3.

#### 4. Conclusions

In this paper, we introduced a device for simultaneously and directly measuring all six of the DOF motion errors of a rotary axis. This device is based on laser collimation and laser interferometry and has the advantages of easy installation, high measurement efficiency, and simple data processing. A PID algorithm was developed to realize the servo-tracking strategy of the reference rotary axis. The six DOF motion errors, including the radial motion error along the  $x$  axis, were measured automatically and simultaneously. The measurement efficiency was greatly improved. The whole measurement time was less than 15 min. The error model was established by using HTMs. The influences of error crosstalk, installation error, and manufacturing error were clarified, and the corresponding compensation methods were given. The repeatability values of the three translational motion errors were obviously improved after compensation. The results of the repeatability and comparison experiments demonstrate the effectiveness of our method and the error compensation

model. The relatively poor repeatability of the angular positioning error was due to the low accuracy of the target rotary axis. In general, the accuracy of measuring devices should be 3 to 10 times the accuracy of the measured equipment. Thus, according to the measurement data, our device can only meet the measurement requirements of lower-precision rotary axes. The repeatability of the measuring device will be further improved so that the device can be used in more fields. The analysis method and model for error compensation also provide a reference for improving the repeatability of other measuring instruments. In the future, the measurement time will be further shortened by optimizing the PID algorithm. Comparison experiments of other DOF motion errors will also be carried out.

**Author Contributions:** All works in relation to this paper were accomplished by all the authors' efforts. Conceptualization was done by Q.F., J.L., and C.B. The experiment was carried out by C.B. and J.L. C.B., J.L., and Q.F. wrote the paper.

**Funding:** This research was funded by the National Natural Science Foundation of China (NSFC) (51527806) and the Fundamental Research Funds for the Central Universities (2016RC019).

**Conflicts of Interest:** The authors declare no conflict of interest.

## References



1. International Organization for Standardization. *ISO 230-7:2015—Test Code for Machine Tools—Part 7: Geometric Accuracy of Axes of Rotation*; International Organization for Standardization: Geneva, Switzerland, 2015.
2. Li, J.; Feng, Q.; Bao, C.; Zhao, Y. Method for simultaneous measurement of five DOF motion errors of a rotary axis using a single-mode fiber-coupled laser. *Opt. Express* **2018**, *26*, 2535–2545. [[CrossRef](#)] [[PubMed](#)]
3. Chen, C.J.; Lin, P.D. High-accuracy small-angle measurement of the positioning error of a rotary table by using multiple-reflection optoelectronic methodology. *Opt. Eng.* **2007**, *46*, 113604. [[CrossRef](#)]
4. Park, S.; Hoang, T.; Yang, S. A new optical measurement system for determining the geometrical errors of rotary axis of a 5-axis miniaturized machine tool. *J. Mech. Sci. Technol.* **2010**, *24*, 175–179. [[CrossRef](#)]
5. Wang, Q.; Miller, J.; Freyberg, A.V.; Steffens, N.; Fischer, A.; Goch, G. Error mapping of rotary tables in 4-axis measuring devices using a ball plate artifact. *CIRP Ann. Manuf. Technol.* **2018**, *67*, 559–562. [[CrossRef](#)]
6. Li, J.; Xie, F.; Liu, X.; Dong, Z.; Song, Z.; Li, W. A geometric error identification method for the swiveling axes of five-axis machine tools by static R-test. *Int. J. Adv. Manuf. Technol.* **2017**, *89*, 3393–3405. [[CrossRef](#)]
7. Schwenke, H.; Knapp, W.; Haitjema, H.; Weckenmann, A.; Schmitt, R.; Delbressine, F. Geometric error measurement and compensation of machines—An update. *CIRP Ann. Manuf. Technol.* **2008**, *57*, 660–675. [[CrossRef](#)]
8. Chen, C.J.; Lin, P.D.; Jywel, W.Y. An optoelectronic measurement system for measuring 6-degree-of-freedom motion error of rotary parts. *Opt. Express* **2007**, *15*, 14601–14617. [[CrossRef](#)] [[PubMed](#)]
9. Schwenke, H.; Schmitt, R.; Jatzkowski, P.; Warmann, C. On-the-fly calibration of linear and rotary axes of machine tools and CMMs using a tracking interferometer. *CIRP Ann. Manuf. Technol.* **2009**, *58*, 477–480. [[CrossRef](#)]
10. Hong, C.; Ibaraki, S. Non-contact R-test with laser displacement sensors for error calibration of five-axis machine tools. *Precis. Eng.* **2013**, *37*, 159–171. [[CrossRef](#)]
11. Ibaraki, S.; Oyama, C.; Otsubo, H. Construction of an error map of rotary axes on a five-axis machining center by static R-test. *Int. J. Mach. Tools Manuf.* **2011**, *51*, 190–200. [[CrossRef](#)]
12. Bao, C.; Li, J.; Feng, Q.; Zhang, B. Error-compensation model for simultaneous measurement of five degrees of freedom motion errors of a rotary axis. *Meas. Sci. Technol.* **2018**, *29*, 075004. [[CrossRef](#)]
13. Nojehdeh, M.V.; Arezoo, B. Functional accuracy investigation of work-holding rotary axes in five-axis CNC machine tools. *Int. J. Mach. Tools Manuf.* **2016**, *111*, 17–30. [[CrossRef](#)]
14. Huang, N.; Bi, Q.; Wang, Y. Identification of two different geometric error definitions for the rotary axis of the 5-axis machine tools. *Int. J. Mach. Tools Manuf.* **2015**, *91*, 109–114. [[CrossRef](#)]



© 2018 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).

Article

# Development of a Compact Three-Degree-of-Freedom Laser Measurement System with Self-Wavelength Correction for Displacement Feedback of a Nanopositioning Stage

Yindi Cai <sup>1</sup> , Zhifeng Lou <sup>1</sup>, Siying Ling <sup>1</sup>, Bo-syun Liao <sup>2</sup> and Kuang-chao Fan <sup>1,2,\*</sup> 

<sup>1</sup> School of Mechanical Engineering, Dalian University of Technology, Dalian 116023, China; caiyd@dlut.edu.cn (Y.C.); louzf@dlut.edu.cn (Z.L.); lingsy@dlut.edu.cn (S.L.)

<sup>2</sup> Department of Mechanical Engineering, National Taiwan University, Taipei 10617, Taiwan; b95502050@ntu.edu.tw

\* Correspondence: fan@ntu.edu.tw

Received: 29 October 2018; Accepted: 7 November 2018; Published: 10 November 2018



**Abstract:** This paper presents a miniature three-degree-of-freedom laser measurement (3DOFLM) system for displacement feedback and error compensation of a nanopositioning stage. The 3DOFLM system is composed of a miniature Michelson interferometer (MMI) kit, a wavelength corrector kit, and a miniature autocollimator kit. A low-cost laser diode is employed as the laser source. The motion of the stage can cause an optical path difference in the MMI kit so as to produce interference fringes. The interference signals with a phase interval of  $90^\circ$  due to the phase control are detected by four photodetectors. The wavelength corrector kit, based on the grating diffraction principle and the autocollimation principle, provides real-time correction of the laser diode wavelength, which is the length unit of the MMI kit. The miniature autocollimator kit based on the autocollimation principle is employed to measure angular errors and compensate induced Abbe error of the moving table. The developed 3DOFLM system was constructed with dimensions of 80 mm ( $x$ )  $\times$  90 mm ( $y$ )  $\times$  20 mm ( $z$ ) so that it could be embedded into the nanopositioning stage. A series of calibration and comparison experiments were carried out to test the performance of this system.

**Keywords:** laser diode; interferometer; wavelength corrector; angular error; nanopositioning stage

## 1. Introduction

Long-stroke and nanopositioning planar stages with nanometer-scale resolution and accuracy are essential in various precision systems, such as electronic device manufacturing equipment and scanning-type measuring instruments [1,2]. Stacked-type planar stages are conventionally constructed with two linear stages by stacking one on top of the other [3,4]. However, the stacked-type planar stages suffer from Abbe error, because the reference axis cannot be in line with the functional axis in the working space; thus, the stacked-type planar stages are not proper from the viewpoints of both positioning accuracy and positioning speed. Gantry stages [5], consisting of two  $x$ -axes and a  $y$ -axis, provide high capacity and high stiffness. Their travel accuracy, however, is significantly influenced by the parallelism of two  $x$ -axes and the Abbe error is also large due to the large Abbe offset.

Compared with the stacked-type planar stage and the gantry stage, coplanar stages can eliminate the Abbe error by allowing the  $x$ - and  $y$ -motions along a common plane [6–9]. The coplanar stage needs to be equipped with a high-precision position sensor for closed-loop control of each axis' motion. Linear encoders are regarded as more stable than laser interferometers for ultra-precision displacement measurement, as the mechanical grid is less sensitive to the ambient condition than

the laser wavelength [10]. However, due to the geometric design of linear encoders, it is difficult to follow the Abbe principle when being installed into the coplanar stage as the feedback sensor. Commercial laser interferometers, made of He–Ne lasers featuring a long measurement range, fast speed, and nanometric resolution, are widely used for the calibration tool of linear stages. However, their sizes are too bulky to be embedded into the small nanopositioning stage as a feedback sensor. A linear displacement grating interferometer (LDGI), which employs the polarization technique to obtain low-noised sinusoidal waveforms, was developed to detect the  $x$ - and  $y$ -motions [6,7]. However, the stage motion has inevitable geometrical angular errors that would cause pulse counting errors of the LDGI. A novel miniature laser diode interferometer based on the Michelson interferometer principle was, thus, developed for precision displacement feedback in each axis of a nanopositioning stage by the authors' group [11]. However, although the laser diode has the advantage of small size, it suffers from low stability in output power and wavelength, which is the length unit of laser interferometer, thereby limiting its applications in length measurements. In addition, the novel miniature laser diode interferometer cannot measure angular errors of the moving stage. It is known that angular errors induce Abbe errors in displacement measurements, which influence the displacement feedback accuracy of the laser diode interferometer. In order to achieve a high-precision closed-loop displacement feedback control of the nanopositioning stage, it is essential to ensure the accuracy of the laser diode wavelength and compensate for the angular-error-induced Abbe error.

There are many methods for correcting the laser wavelength in room conditions. An air sensor, which is used to detect the current temperature, relative humidity, and atmospheric pressure to correct the refractive index of air through an empirical equation with a reference wavelength in vacuum, is commonly required in commercial laser interferometers. The empirical equation is usually taken from the Edlen equation [12] or its modified equations [13,14]. However, on the one hand, the empirical equation was obtained based on the He–Ne laser, which has different properties from the laser diode. On the other hand, the empirical method is an indirect measurement, which is subject to the sensitivity and accuracy of the air sensor. Therefore, the correction of laser diode wavelength cannot directly adopt existing empirical equations.

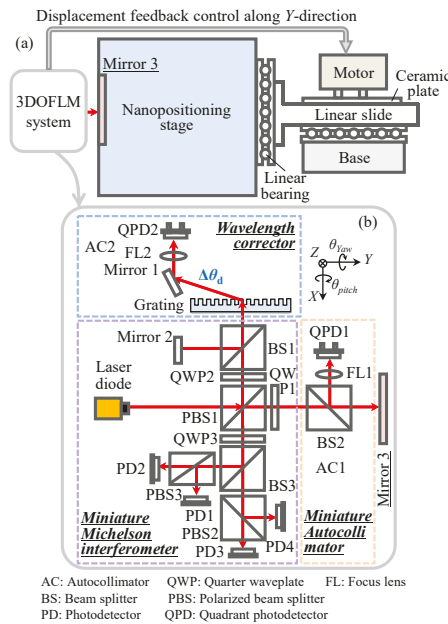
Compared with the indirect measurement method, the direct measurement method can obtain real-time laser wavelength. An interferometer comparator can be used for laser wavelength measurements by analyzing the relationship between the phase [15,16] or the period [17,18] of interference fringes and the laser wavelength. The interference fringes are obtained when an optical path difference between a reference path and a measurement path is generated in the interferometer. The common measurement systems are Michelson interferometers [19,20] and Mach–Zehnder interferometers [21] with mirrors or cube corners. However, in the interferometer comparator, a reference laser is necessary, and the measurement accuracy depends significantly on the laser coherence. Optical beating methods with a certain optical path difference can be employed to make a highly accurate measurement of the laser wavelength [22,23]. Optical beating methods are based on frequency comparisons with a laser stabilized by means of the saturated absorption method [22] or with a comb-generator [23]. Wavelength can also be measured based on specific wavelength-dependent material properties, such as specialties of optical fiber [24], polarization of light [25], and diffraction gratings [16,26]. The nature of simple structure and portability of this kind of wavelength measurement method predestinates it to be applied in different kinds of commercial systems, especially gratings, in which laser wavelength can be calculated by detecting the variation of diffraction angle with a charge-coupled device (CCD) [16]. However, CCD cannot provide high-precision wavelength measurements due to its low resolution and accuracy.

In this study, a laser-diode-based miniature Michelson interferometer (MMI) with self-wavelength correction based on the grating diffraction principle and the autocollimation principle is proposed. Combining the MMI with a dual-angle sensor based on the autocollimation principle, a compact three-degree-of-freedom laser measurement (3DOFLM) system is developed, which is employed as the displacement feedback sensor, as well as pitch and yaw angular sensors, of a nanopositioning

stage. The optical layout of the 3DOFLM system is designed in such a way that the displacement and two angular motions of the nanopositioning stage can be detected simultaneously. The laser wavelength can be corrected in real time during the displacement feedback. A prototype system was constructed. The principle of the 3DOFLM system design and its performance tests are presented in the following sections.

## 2. Principle of the Miniature Michelson Interferometer (MMI)

Figure 1a shows the configuration of a nanopositioning  $y$ -stage. The linear bearing between the linear slide and the moving table is to allow the installation of another  $x$ -stage to form a coplanar stage. For the sake of clarity, only the displacement feedback control along the  $y$ -direction is plotted. The ultrasonic motor is employed to actuate the linear slide through a friction force on the ceramic plate. Thus, the moving table can be moved by the linear slide in push-pull mode. The three-degree-of-freedom laser measurement (3DOFLM) system, consisting of a miniature Michelson interferometer (MMI) kit for displacement measurement and feedback, a wavelength correction kit for the MMI kit, and a miniature autocollimator kit for angular error measurement, is mounted onto the baseplate of the stage and on the opposite side of the ultrasonic motor. The laser beam of the 3DOFLM system is in line with the moving axis at the center of the stage. With this kind of arrangement, the Abbe principle is totally observed. The optical design of the 3DOFLM system is shown in Figure 1b. A low-cost laser diode is employed as the laser source. The miniature autocollimator kit is employed to detect the pitch and yaw angular errors for the purpose of Abbe error compensation of the stage. The displacement of the stage detected by the MMI kit is used as a feedback signal to the ultrasonic motor. The laser diode wavelength of the MMI kit is corrected in real time by the wavelength corrector kit. The detailed principles of angular error measurement and displacement measurement with and without wavelength correction are introduced below.



**Figure 1.** (a) Top view of configuration of the nanopositioning  $y$ -stage; (b) optical design of the three-degree-of-freedom laser measurement (3DOFLM) system ( $xy$  view).

### 2.1. Angular Error Measurement of the Nanopositioning Stage

The principle of angular error measurement is based on the autocollimation principle. A collimated laser beam emitted from a laser diode is divided into a transmitted measurement beam (P-polarized beam) and a reflected reference beam (S-polarized beam) by a polarizing beam splitter (PBS1), as shown in Figure 1. The measurement beam reflected by a moving mirror (mirror 3) is split into two beams by a beam splitter (BS2). The reflected beam is used to measure pitch and yaw errors of the nanopositioning stage by an autocollimator (AC1), comprising a focus lens (FL1) and a four-quadrant photodetector (QPD1).

Figure 2 shows the optical layout of the miniature autocollimator kit. BS2 is mounted on an angle mirror mount for easy angle adjustment. QPD1 is positioned on the focal point of FL1. When mirror 3 is rotated with a pitch angle or a yaw angle, the position of the focused spot on QPD1 is moved along the *x*- or *y*-directions, accordingly. The relationship between the angle and laser spot movement can be expressed by the following equations:

$$\theta_{pitch} = \frac{\Delta x}{2f} \text{ and } \theta_{yaw} = \frac{\Delta y}{2f}, \tag{1}$$

where  $\theta_{pitch}$  and  $\theta_{yaw}$  are the pitch and yaw errors,  $\Delta x$  and  $\Delta y$  are the corresponding spot shifts on QPD1, and  $f$  is the focal length of FL1.

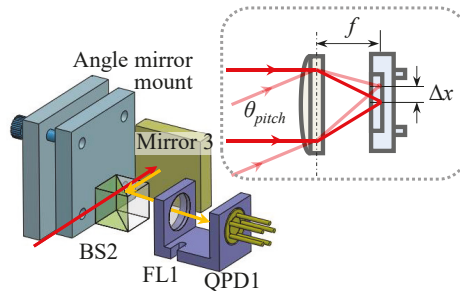


Figure 2. Optical layout of the miniature autocollimator kit.

### 2.2. Displacement Measurement of the Nanopositioning Stage

Due to the space limit of the nanopositioning stage, a miniature Michelson interferometer (MMI) kit is developed for displacement feedback, as shown in Figure 3a. A collimated laser beam emitted from a laser diode is divided into a measurement beam and a reference beam by PBS1, as shown in Figure 1. The reference beam passes through a quarter waveplate (QWP2) and is reflected back by a fixed mirror (mirror 2). The measurement beam passes through QWP1 and is reflected back by the moving mirror (mirror 3). It should be noted that the polarization state of the measurement and reference beams are changed by 90° after passing QWP2 and QWP1 twice, which prevents the reflected beams from entering back into the laser diode. Those two reflected beams are combined at PBS1 and converted into left and right circularly polarized beams after passing through QWP3. The interference signals with a phase shift of 90° due to the phase control by two PBSs (PBS2 and PBS3) and a BS (BS3) are detected by four photodetectors (PD1–PD4), as shown in Figure 3b. The motion of the moving mirror 3 causes an optical path difference between the two reflected beams so as to produce interference. Analyzed by the Jones matrix, the intensity of each photodetector can be expressed as follows:

$$I_{PD1} = I_0(1 - \cos(\frac{2\pi \cdot \Delta y}{\lambda})), \tag{2}$$

$$I_{PD2} = I_0(1 + \cos(\frac{2\pi \cdot \Delta y}{\lambda})), \tag{3}$$



$$I_{PD3} = I_0(1 + \sin(\frac{2\pi \cdot \Delta y}{\lambda})), \tag{4}$$

$$I_{PD4} = I_0(1 - \sin(\frac{2\pi \cdot \Delta y}{\lambda})), \tag{5}$$

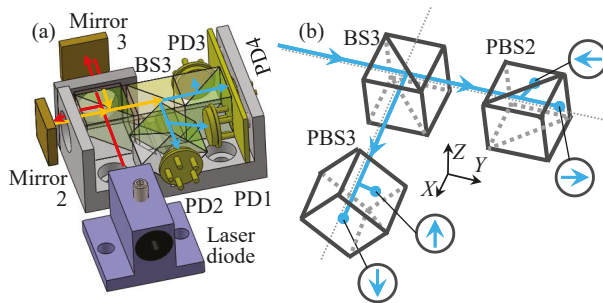
where  $I_0$  and  $\lambda$  are the intensity and the wavelength of the laser beam, respectively, and  $\Delta y$  is the optical path difference of two reflected beams. The signal direct current (DC) offset can be eliminated by subtracting Equation (2) from Equation (3) for the cosine term, and Equation (5) from Equation (4) for the sine term. The phase change can be obtained by

$$\Delta\Phi = \tan^{-1}(\frac{I_{PD1} - I_{PD2}}{I_{PD3} - I_{PD4}}) = \frac{2\pi \cdot \Delta y}{\lambda}. \tag{6}$$

Based on the optical design of the MMI kit, when mirror 3 has a motion of  $d$ , the optical path difference of the two reflected beams  $\Delta z$  will be  $2d$ . Thus, the displacement  $d$  of the moving mirror 3 can be calculated by

$$d = \frac{\lambda}{2}(N + \frac{\Phi_{initial} - \Phi_{final}}{2\pi}), \tag{7}$$

where  $N$  is the number of sinusoidal waves, and  $\Phi_{initial}$  and  $\Phi_{final}$  are the exact phases of initial and final incomplete wave cycles, respectively. It can be seen from Equation (7) that the interference waveform has the period of one-half laser wavelength. Thus, the accuracy of the laser wavelength is significantly important in the MMI kit.



**Figure 3.** Optical layout of the miniature Michelson interferometer (MMI) kit: (a) three-dimensional (3D) image; (b) geometric configuration of generating four interference fringes with 90° phase shift.

### 2.3. Correction of Laser Diode Wavelength in the MMI Kit

The wavelength corrector kit is based on the grating diffraction principle and the autocollimation principle, as shown in Figure 4. The divided reference beam is transmitted through BS1 and impinged into a grating. A group of diffraction beams can be observed. Here, the zeroth (0th) order and positive first (+1st) order diffraction beams are picked up for calculating laser wavelength. Based on the diffraction grating equation, the following equation can be obtained:

$$d(\sin \theta_i + \sin \theta_d) = \lambda_n, \tag{8}$$

where  $d$  is the grating pitch,  $\theta_i$  is the incidence angle,  $\lambda_n$  is the nominal laser wavelength, and  $\theta_d$  is the +1st order diffraction angle with respect to  $\lambda_n$ . As seen from Equation (8), the wavelength can be calculated if the +1st order diffraction angle can be measured at known incidence angle and grating pitch. This is also the measurement principle of typical spectrometers [26]. In the designed wavelength corrector, the laser beam is set to normally project onto the grating; thus, the incidence angle is 0.

However, if the nominal wavelength is varied by  $\Delta\lambda$ , the diffraction angle will have a drift. Thus, Equation (8) can be modified to

$$\lambda_n + \Delta\lambda = d \sin(\theta_d + \Delta\theta_d), \tag{9}$$

where  $\Delta\theta_d$  is the drift of +1st order diffraction angle due to  $\Delta\lambda$ .

From Equation (9), it is seen that, if  $\Delta\theta_d$  can be measured in real time,  $\Delta\lambda$  can be immediately calculated, and the real-time laser wavelength can then be automatically self-corrected without the need of the air sensor and the Edlen equation. This is the significance of the proposed self-wavelength correction.

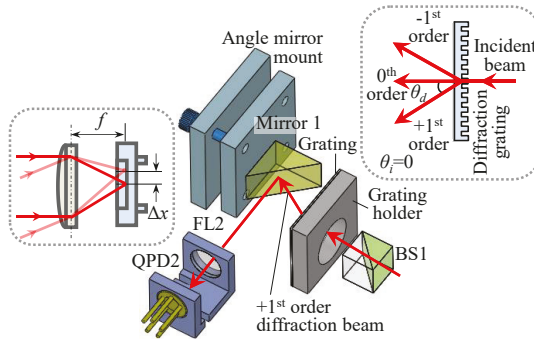


Figure 4. Optical layout of the wavelength corrector kit.

### 3. Experiments and Discussions

Figure 5 shows a photo of the constructed 3DOFLM system. Its size, including the electric cables and the base plate for mounting the optical components, is 80 mm ( $x$ )  $\times$  90 mm ( $y$ )  $\times$  20 mm ( $z$ ), which is much smaller than commercial He–Ne laser interferometers. The U-shaped holder was made to fix both FL and QPD in one place for better assembly accuracy. It should be noted that, due to the space limit of the nanopositioning stage, the dead path of the 3DOFLM system, which is the distance between the MMI kit and the null point of the measurement position, was designed to be 36 mm. Therefore, the error due to the dead path is negligible [27].

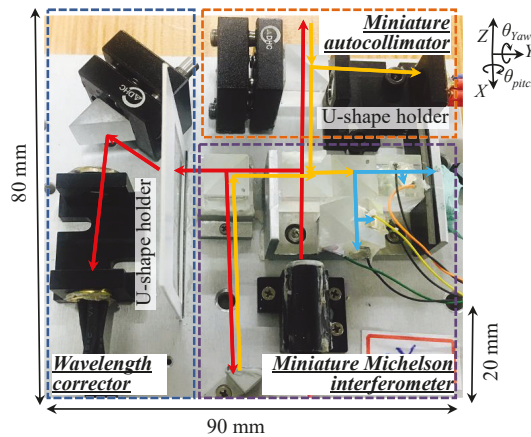


Figure 5. Photo of the constructed 3DOFLM system.

In order to achieve accurate displacement of the nanopositioning stage, it is essential to investigate the performance of the designed 3DOFLM system. It should be noted that all experiments were carried out in a non-environmental controlled open laboratory in which disturbances caused by temperature change and humidity variation cannot be avoided.

### 3.1. Calibration and Comparison Experiments of Angular Error Measurement

Firstly, the basic performances of the miniature autocollimator kit, including stability and measurement accuracy, were tested. A commercial low-cost collimated laser diode with a nominal wavelength of 635 nm and a divergence angle less than 0.3 mrad was adopted as the laser source. A high-precision QPD (SPOT-4DMI, OSI Optoelectronics, California, USA) with an active area of 0.25 mm per element, an element gap of 13  $\mu\text{m}$ , and a responsivity of 0.65 A/W was chosen as the angle detector. The sampling frequency was set to be 10 Hz, and a Butterworth filter was used to reduce the noise level. Figure 6 shows the stability of the angular errors over a duration of 1 h. As seen, the stability of the output signals for measuring the pitch angle error and yaw angle error were approximately 0.55 and 0.35 arcsec, respectively. It should be noted that, although the laser beams for measuring the pitch and yaw angle errors were from an identical laser diode, due to the non-uniformity of the output intensity distribution of the laser diode, the stability in pitch and yaw angle error measurements were not similar, but inversely correlated in Figure 6. The reason for the larger pitch variation could be due to the mechanical vibration. It can be seen that the developed miniature autocollimator kit satisfied stability for angular error measurements with sub-arcsecond precision.

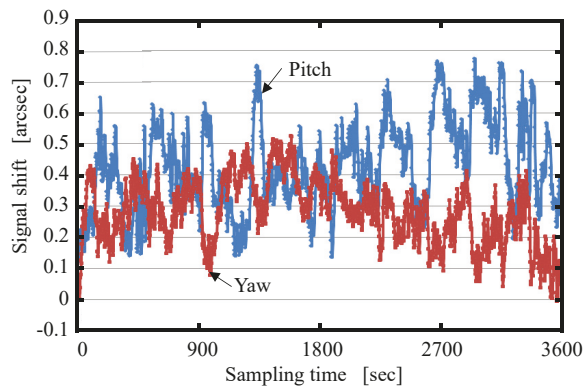


Figure 6. Stability of output signals of the miniature autocollimator kit.

Prior to the testing the nanopositioning stage's angular error, the miniature autocollimator kit was calibrated by a commercial laser interferometer (HP5529A, Keysight Technologies, California, USA), which has an angle measurement accuracy of  $\pm 0.2\%$  of displayed value and  $\pm 0.05$  arcsec per meter of distance traveled. The experiments were repeated five times, and the calibration range was set to be  $\pm 40$  arcsec. The comparative results are shown in Figure 7a,b. It can be seen that the residuals of pitch angle and yaw angle were all within  $\pm 0.3$  arcsec.

The 3DOFLM system was then mounted on the designed nanopositioning stage [6] to measure the angular error of the stage. The HP5529A interferometer was also adopted for comparison. The laser beams of both the 3DOFLM and HP5529A systems were adjusted in line with the motion axis of the stage. Figure 8a,b show the measurement results. It is seen that the output of the 3DOFLM system was in a good agreement with that of the HP5529A system. The maximum residual of both angles was within  $\pm 0.5$  arcsec in the measurement distance of 20 mm. It is verified that the performance of the designed 3DOFLM system was acceptable for sub-arcsecond precision angular error measurement.

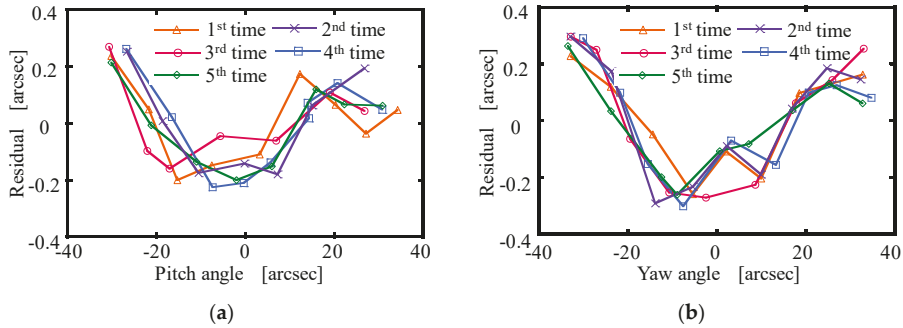


Figure 7. Calibration angle results of the miniature autocollimator kit: (a) pitch; (b) yaw.

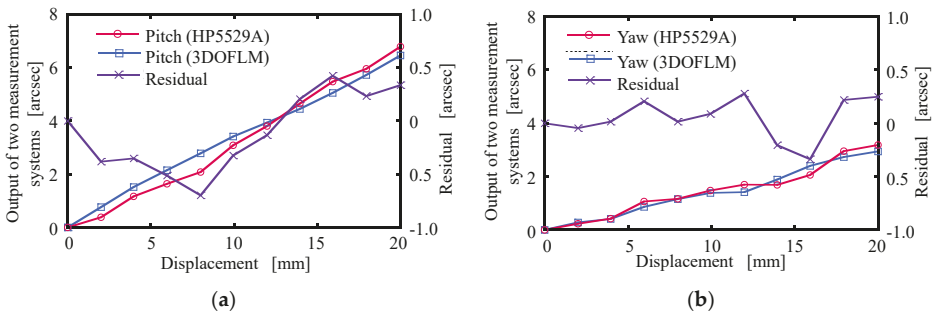


Figure 8. Comparison angle results of measured angular error of the nanopositioning stage by the HP5529A and the 3DOFLM system: (a) pitch; (b) yaw.

3.2. Displacement Measurement without Wavelength Correction

The displacement measurement range of the MMI kit was firstly confirmed using the HP5529A interferometer, which has a measurement range of 15 m. In experiments, the 3DOFLM and HP5529A systems were located on both sides of a linear stage (MSR 100, PARKER, Charlotte, USA), which was moved by a commanded step of 2 to 36 mm long. Figure 9 shows the bi-directional displacement measurement results of the MMI kit and the HP5529A system, determined in duplicate. It can be seen that both outputs were in good agreement. Output sinusoidal signals in the form of a Lissajous circle were found to be very good and stable during motion, as shown in Figure 10. The measuring range of the MMI kit was, thus, confirmed to be at least 36 mm, which is regarded as a long-stroke for a nanopositioning stage.

The displacement measurement accuracy of the 3DOFLM system was then calibrated by a commercial nanopositioning and nanomeasuring machine (NMM-1, SIOS Meßtechnik GmbH, Ilmenau, Germany), which has a positioning range of 25 mm in the  $x$ - and  $y$ -directions with sub-nanometer positioning accuracy. The misalignment error of the MMI kit and NMM-1 will induce a cosine error in the displacement measurement, which cannot be ignored in the nanometer range. Therefore, particular effort was put into aligning the laser beam of the MMI kit with the moving axis of the NMM-1. In this alignment experiment, a QPD was adopted to detect the two lateral motions of the moving table of the NMM-1 relative to the laser beam of the MMI kit. An angle mirror mount was applied to adjust the laser beam so that the beam spot was always close to the center of the QPD during the motion of the NMM-1. Figure 11 shows the finally adjusted straightness error of the NMM-1 in the  $x$ - and  $y$ -directions. The maximum straightness error was evaluated to be 1.272  $\mu\text{m}$ , and it corresponded to a cosine error of 0.04 nm on the displacement measurement in the range of 20 mm. Such a cosine

error was small enough compared to the displacement of NMM-1 and was deemed negligible for the subsequent experiments.

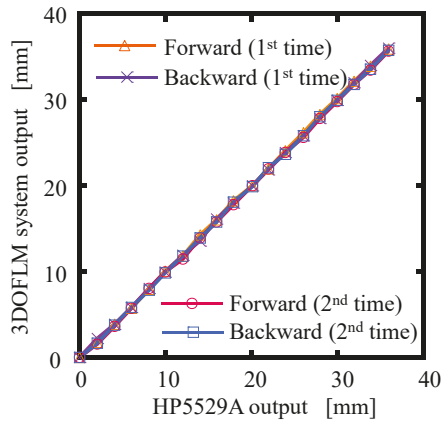


Figure 9. Confirmation of measurement range of the 3DOFLM system.

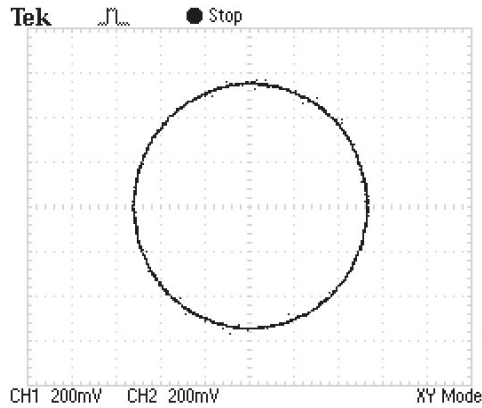


Figure 10. Output Lissajous circle of the 3DOFLM system.

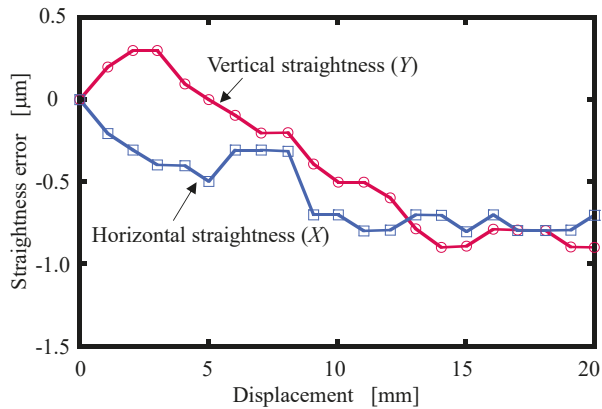
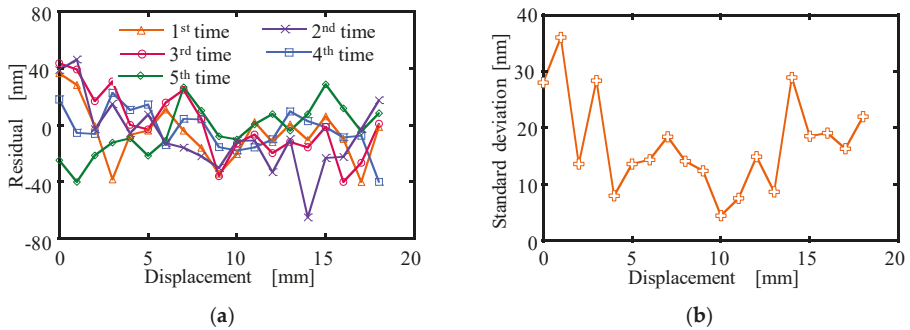


Figure 11. Straightness error of the commercial nanopositioning and nanomesuring machine (NMM-1).

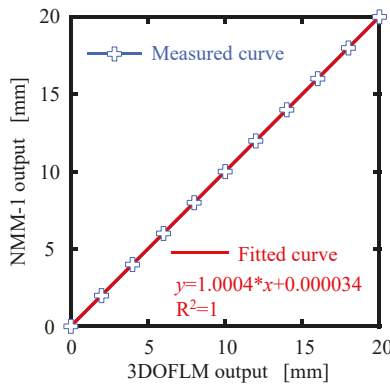
The calibration experiments without wavelength correction were then carried out with the nominal wavelength of the laser diode (i.e., 635 nm) given by the manufacturer. The displacement residuals between the 3DOFLM system and the NMM-1 system were evaluated within  $\pm 45$  nm up to a travel of 18 mm, as shown in Figure 12. The standard deviation of the displacement residuals for five measurements was estimated between 4.49 and 35.98 nm. The residual error between the 3DOFLM and NMM systems was deemed to be mainly caused by the incorrect laser diode wavelength of the MMI kit. Therefore, it is essential to correct the laser diode wavelength for better measurement accuracy.



**Figure 12.** Comparison results: (a) residual of measured displacement by the NMM-1 and 3DOFLM systems without wavelength correction; (b) standard deviation of five measurements.

### 3.3. Displacement Measurement with Wavelength Correction

Figure 13 shows simultaneous displacement measurement results of the NMM-1 system and the MMI kit with a step interval of 2 mm in the measurement distance of 20 mm. The actual calibrated wavelength (referred to as  $\lambda_{cal}$ ) was estimated by multiplying the nominal laser wavelength of 635 nm by the slope  $k_p$  of the fitted curve of the results in Figure 13.



**Figure 13.** Comparison results of measured displacement by the NMM-1 and 3DOFLM systems without wavelength correction.

In the wavelength corrector kit, a transmission diffraction grating (Edmund optics, Barrington, USA) with a grating pitch of 1  $\mu$ m was selected. The optical components used in AC2 were the same as that in AC1. The performance investigations of wavelength correction were firstly conducted. Figure 14 shows the stability of AC2, which was employed to measure the drift of +1st order diffraction angle. As seen from Figure 14, the maximum angle drift was measured as less than 1 arcsec. The laser wavelength variation caused by the variation of AC2 was evaluated to be 0.0037 nm by Equation (9).

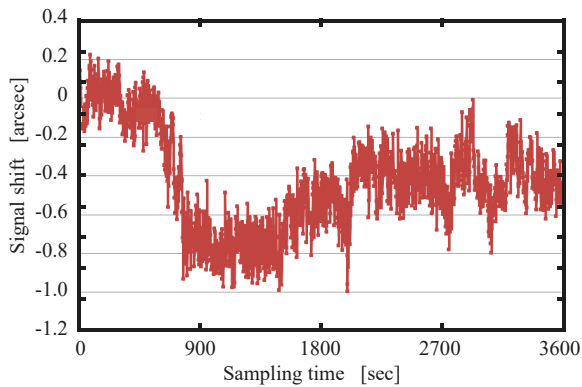


Figure 14. Stability of output signals of the autocollimator (AC2) in the wavelength corrector kit.

The angular measurement accuracy of AC2 was calibrated by the HP5529A system and the results are shown in Figure 15. It can be seen that the angle residual for the range of  $\pm 100$  arcsec was within  $\pm 0.5$  arcsec. It is verified that AC2 was acceptable for sub-arcsecond precision angular error measurement.

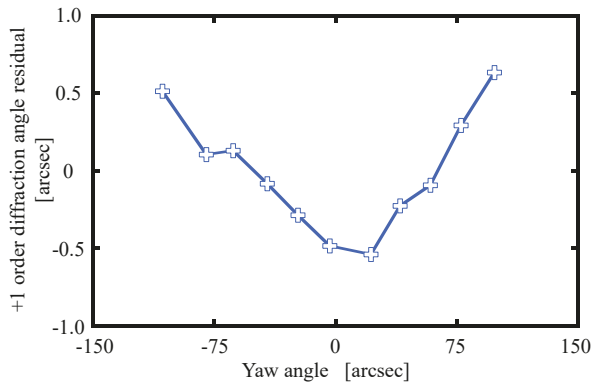
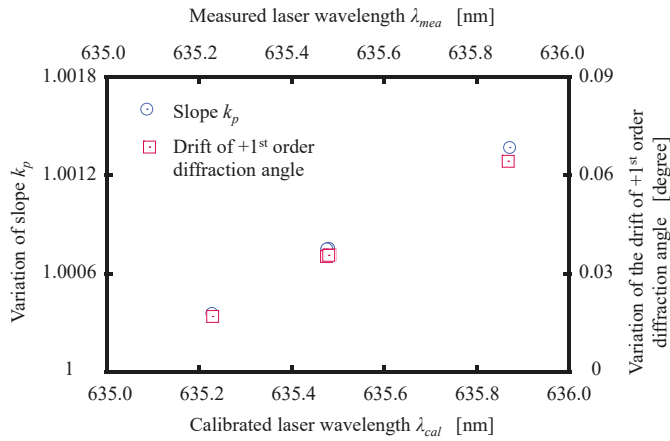


Figure 15. Calibration results of AC2 in the wavelength corrector kit.

The real-time laser diode wavelength was then evaluated based on Equation (9) according to the output of AC2. The wavelength measured by the self-wavelength correction is referred to as  $\lambda_{mea}$ . The calibrated wavelength ( $\lambda_{cal}$ ) was obtained from the ratio of the actual displacement of NMM-1 system and the measured total phase change of the MMI kit, as given by Equation (7). Figure 16 shows the relationship between the calibrated laser wavelength  $\lambda_{cal}$  and the variation of slope  $k_p$  obtained from the displacement measurement results of the NMM-1 system and the MMI kit. The relationship between the measured laser wavelength  $\lambda_{mea}$  and the variation of drift +1st order diffraction angle detected by AC2 was also plotted in Figure 16. For the sake of comparison,  $\lambda_{cal}$  and  $\lambda_{mea}$  are listed in Table 1. It is clearly seen that an order of  $10^{-6}$  accuracy was achieved in measured wavelength by the designed wavelength correction. It means that, if the 3DOFLM system measures a distance of 20 mm, the maximum measurement error, on average, would be estimated as about 20 nm.

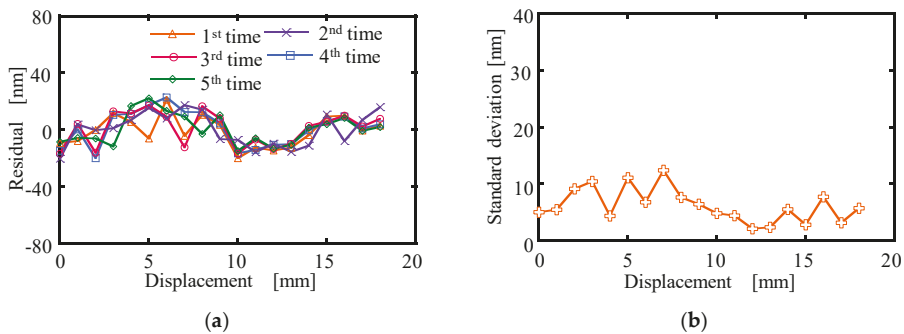


**Figure 16.** Relationship between  $\lambda_{cal}$  and the variation of slope  $k_{cal}$ , and between  $\lambda_{cal}$  and the variation of the drift of positive first-order diffraction angle.

**Table 1.** Calibrated and measured laser wavelength.

Times	$\lambda_{cal}$ (nm)	$\lambda_{mea}$ (nm)	$\frac{\lambda_{cal}-\lambda_{mea}}{\lambda_{cal}}$
1	635.2275	635.2294	$-2.90 \times 10^{-6}$
2	635.4802	635.4750	$8.18 \times 10^{-6}$
3	635.4761	635.4824	$-9.91 \times 10^{-6}$
4	635.4764	635.4803	$-6.13 \times 10^{-6}$
5	635.8709	635.8676	$5.18 \times 10^{-6}$
Average	635.5069	635.5062	$-1.16 \times 10^{-6}$

From the above experiments, it can be realized that a special feature of this 3DOFLM system is that it only needs to calibrate and store the nominal wavelength and the corresponding initial diffraction angle once using a reference. With these stored parameters, the system can automatically correct the wavelength by itself for the remaining measurements. The comparison results between the NMM-1 and 3DOFLM systems with wavelength correction are given in Figure 17. It can be seen that, as expected, the measurement error of the 3DOFLM system falls within  $\pm 25$  nm, which was much smaller than that without wavelength correction in Figure 12. The corresponding standard deviation of the measurement results was evaluated to be between 2.2 and 13.5 nm. The effectiveness of the wavelength corrector kit was, thus, verified.



**Figure 17.** Comparison results: (a) residual of measured displacement by the NMM-1 and 3DOFLM systems with wavelength correction; (b) standard deviation of five measurements.



### 3.4. Error Analysis

The measurement error of the 3DOFLM system was significantly reduced after correcting the laser diode wavelength. However, the measurement accuracy of the 3DOFLM system is still a little lower than commercial laser interferometers. The reason could be that the measurement accuracy of the wavelength corrector is influenced by the linearity error of the Lissagous circle, the incident beam angle and its drift, thermal deformation of the grating pitch, etc., which were not considered in this study. As a report for the first step of the research, this paper focuses on the design and construction of the 3DOFLM system. The improvement of wavelength measurement accuracy will be carried out in the next step of the research for the proposal of a practical application of the nanopositioning stage.

## 4. Conclusions

In this paper, a low-cost and compact three-degree-of-freedom laser measurement (3DOFLM) system with real-time wavelength correction was proposed for displacement feedback, angular error measurement, and Abbe error compensation of a nanopositioning stage. The method of wavelength correction does not rely on the Edlen equation, which was adapted to the He-Ne laser and was an off-line process. The principle of the 3DOFLM system was expressed. It only needs to calibrate and store the nominal wavelength and the corresponding initial diffraction angle once using a reference. With these stored parameters, the system can automatically correct the wavelength by itself in the remaining measurements. A prototype was constructed and tested. It was verified by experiments that, compared with the HP5529A interferometer, the angular error measurement achieved an accuracy of  $\pm 0.5$  arcsec in a travel distance of 20 mm. The displacement measurement achieved a range up to 36 mm, and the accuracy was improved from  $\pm 45$  nm without wavelength correction to  $\pm 25$  nm with wavelength correction. The feasibility of the designed 3DOFLM system was demonstrated. In order to further improve the real-time measured wavelength, considerations of the influences of linearity error of the Lissagous circle, incident beam angle and its drift, and thermal deformation of the grating pitch on the wavelength measurement accuracy will be carried out in future works.

**Author Contributions:** Y.C. wrote the main concepts of the manuscript and contributed the formal analysis; B.L. implemented numerical experiments; Z.L. and S.L. collaboratively discussed the results; K.F. conceived the experiments and checked the English writing.

**Acknowledgments:** The authors are grateful for the financial support provided by the National Key Research and Development Program of China (2017YFF0204800), the Fundamental Research Funds for the Central Universities (No. DUT16TD20) and the National Natural Science Foundation of China (U1508211, 51621064).

**Conflicts of Interest:** The authors declare no conflicts of interest.

## References

1. Buice, E.; Otten, D.; Yanga, R.; Smith, S.; Hocken, R.; Trumper, D. Design Evaluation of a Single-Axis Precision Controlled Positioning Stage. *Precis. Eng.* **2009**, *33*, 418–424. [[CrossRef](#)]
2. Schwenke, H.; Knapp, W.; Haitjema, H.; Weckenmann, A.; Schmitt, R.; Delbressine, F. Geometric error measurement and compensation of machines—An update. *CIRP Ann. Manuf. Technol.* **2008**, *57*, 660–675. [[CrossRef](#)]
3. Li, Y.; Xu, Q. A novel piezoactuated XY stage with parallel, decoupled, and stacked flexure structure for micro-/nanopositioning. *IEEE Trans. Ind. Electron.* **2011**, *58*, 3601–3615. [[CrossRef](#)]
4. Polit, S.; Dong, J. Development of a high-bandwidth XY nanopositioning stage for high-rate Micro-/Nanomanufacturing. *IEEE-ASME Trans. Mechatron.* **2011**, *16*, 724–733. [[CrossRef](#)]
5. Herreros, I.G.; Kestelyn, X.; Coleman, R.; Barre, P.J. Model-based decoupling control method for dual-drive gantry stages: A case study with experimental validations. *Control Eng. Pract.* **2013**, *21*, 298–307. [[CrossRef](#)]
6. Cheng, F.; Fan, K.C. Linear diffraction grating interferometer with high alignment tolerance and high accuracy. *Appl. Opt.* **2011**, *50*, 4550–4556. [[CrossRef](#)] [[PubMed](#)]
7. Wang, H.Y.; Fan, K.C.; Ye, J.K.; Lin, C.H. A Long-Stroke Nanopositioning Control System of the Coplanar Stage. *IEEE/ASME Trans. Mechatron.* **2014**, *19*, 348–356. [[CrossRef](#)]

8. Li, X.; Gao, W.; Muto, H.; Shimizu, Y.; Ito, S.; Dian, S. A six-degree-of-freedom surface encoder for precision positioning of a planar motion stage. *Precis. Eng.* **2013**, *37*, 771–781. [[CrossRef](#)]
9. Gao, W. *Precision Nanometrology: Sensors and Measuring Systems for Nanomanufacturing*; Springer: London, UK, 2010.
10. Otsuka, J.; Ichikawa, S.; Masuda, T.; Suzuki, K. Development of a small ultraprecision positioning device with 5 nm resolution. *Meas. Sci. Technol.* **2005**, *16*, 2186–2192. [[CrossRef](#)]
11. Fan, K.C.; Zhou, H.; Li, R. Development of a High Precision Coplanar Stage by using Miniature Michelson Interferometer. *Int. J. Precis. Eng. Manuf.* **2014**, *15*, 2251–2256. [[CrossRef](#)]
12. Edlen, B. The Refractive Index of Air. *Metrologia* **1996**, *2*, 71–80. [[CrossRef](#)]
13. Birch, K.P.; Downs, M.J. An Updated Edlen Equation for the Refractive Index of Air. *Metrologia* **1993**, *30*, 155–162. [[CrossRef](#)]
14. Birch, K.P.; Downs, M.J. Correction to the Updated Edlén Equation for the Refractive Index of Air. *Metrologia* **1994**, *31*, 315–316. [[CrossRef](#)]
15. Shi, Z.; Boyd, R.W.; Camacho, R.M.; Vudyasetu, P.K.; Howell, J.C. Slow-light Fourier transform interferometer. *Phys. Rev. Lett.* **2007**, *99*, 240801. [[CrossRef](#)] [[PubMed](#)]
16. Dobosz, M.; Kożuchowski, M. Overview of the laser-wavelength measurement methods. *Opt. Lasers Eng.* **2017**, *98*, 107–117. [[CrossRef](#)]
17. Sun, J.; Zhang, J.; Zhang, X.; Lin, H.; Feng, X. Length determination of a fixed-path cylindrical resonator with DHE dual wavelength laser interference method. *Int. J. Thermophys.* **2011**, *32*, 1330–1338. [[CrossRef](#)]
18. Dubov, M.; Klishevich, G.; Sergienko, V.; Obukhov, I.; Turchin, A.; Chernomorets, M. An astigmatic wavelength meter. *Instrum. Exp. Tech.* **2001**, *44*, 537–543. [[CrossRef](#)]
19. Yan, L.; Chen, B.; Yang, W.; Wei, R.; Zhao, S. A novel laser wavelength meter based on the measurement of synthetic wavelength. *Rev. Sci. Instrum.* **2010**, *81*, 115104. [[CrossRef](#)] [[PubMed](#)]
20. Solomakha, D.; Toropov, A. Laser wavelength measurements (review). *Sov. J. Quantum Electron* **1977**, *7*, 929–942. [[CrossRef](#)]
21. Fox, P.; Scholten, R.; Walkiewicz, M.; Drullinger, R. A reliable, compact, and low-cost Michelson wavemeter for laser wavelength measurement. *Am. J. Phys.* **1999**, *67*, 624–630. [[CrossRef](#)]
22. Balling, P. Absolute frequency measurement of wavelength standard at 1542 nm: Acetylene stabilized DFB laser. *Opt. Express* **2005**, *13*, 9196–9201. [[CrossRef](#)] [[PubMed](#)]
23. Zhang, J.; Lu, Z.H.; Wang, Y.H.; Liu, T.; Stejskal, A.; Zhao, Y.N. Exact frequency comb mode number determination in precision optical frequency measurements. *Laser Phys.* **2007**, *17*, 1025–1028. [[CrossRef](#)]
24. Cooper, D.; Smith, P. Simple and highly sensitive method for wavelength measurement of low-power time-multiplexed signals using optical amplifiers. *J. Lighthwave Technol.* **2003**, *21*, 1612–1620. [[CrossRef](#)]
25. Kundikova, N.; Lonschakova, A. Wavelength measurement by polarization method. *Proc. SPIE* **2005**, *6024*, 1–8.
26. Saitoh, T.; Nakamura, K.; Takahashi, Y.; Miyagi, K. Optical spectrum analyzer utilizing MEMS scanning mirror. *Photonics Technol. Lett.* **2006**, *18*, 767–769. [[CrossRef](#)]
27. Samir, M. *Introduction to Precision Machine Design and Error Assessment*; CRC Press: Boca Raton, FL, USA, 2013.



© 2018 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).

Article

# Sub-Pixel Chessboard Corner Localization for Camera Calibration and Pose Estimation

Tianlong Yang, Qiancheng Zhao \*, Xian Wang and Quan Zhou

College of Mechanical and Electrical Engineering, Hunan University of Science and Technology, Taoyuan Rd, Xiangtan 411201, China; eastlife0108@163.com (T.Y.); 15111388435@163.com (X.W.); zhouquan1103@126.com (Q.Z.)

\* Correspondence: qczhao@163.com; Tel.: +86-131-0732-6852

Received: 30 September 2018; Accepted: 26 October 2018; Published: 1 November 2018



**Abstract:** This work describes a novel approach to localize sub-pixel chessboard corners for camera calibration and pose estimation. An ideally continuous chessboard corner model is established, as a function of corner coordinates, rotation and shear angles, gain and offset of grayscale, and blurring strength. The ideal model is evaluated by a low-cost and high-similarity approximation for sub-pixel localization, and by performing a nonlinear fit to input image. A self-checking technique is also proposed by investigating qualities of the model fits, for ensuring the reliability of addressing perspective-n-point problem. The proposed method is verified by experiments, and results show that it can share a high performance. It is also implemented and examined in a common vision system, which demonstrates that it is suitable for on-site use.

**Keywords:** chessboard corner; camera calibration; pose estimation; sub-pixel localization

## 1. Introduction

Computer vision is an interdisciplinary field that deals with how computers can be made for gaining high-level understanding from digital images or videos. From the perspective of engineering, it seeks to automate assignments that the human eyes can do. As a sub-domain of computer vision, visual measurement is employed for some applications involving dimensional survey tasks, and it always utilizes one or more cameras with exactly known intrinsic parameters for addressing perspective-n-point problem [1] and, therefore, camera calibration is pivotal for ensuring the system accuracy [2].

Most camera calibration approaches require a certain number of correspondences between world and image frames, which are also known as control points, and they usually are called “targets” in photogrammetry [3]. These approaches are performed with planar or non-planar targets with exactly known geometries. After the targets are photographed, their corresponding image points need to be localized for solving intrinsic and extrinsic parameters based on bundle adjustment or other optimization models [4]. As a result, the accuracy of camera calibration is largely dependent on the localization of image points, and usually evaluated by re-projection errors [5].

Circular dots and chessboards are the most common target types. Without a loss of generality, projecting the center of a circle yields an image point that is not necessary to be the center of a pattern projected from the circle, unless the pattern is still circular. Contrarily, the corner of a chessboard is scarcely subjected to projective transformations. For that reason, chessboards are more convenient for achieving targets in visual measurements [6]. In addition, since Zhang [2] proposed a flexible calibration approach employing a planar rig with chessboard patterns, this approach has been cited more than ten thousand times, and made chessboards the most frequently used targets for camera calibration.

Chessboard corners can be detected at pixel level using conventional detectors, such as Harris [7] and Kanade-Lucas-Tomasi (KLT) [8]. These detectors usually extract a set of redundant points from one corner, due to user-defined parameters. Some modified means [9–11] are contributed to make arbitrary points converge into nearby corners, but their operational accuracy is still at pixel level. In many application scenarios [12–14], however, pixel resolution is not yet accurate enough and, therefore, mathematical techniques, such as interpolations or approximations, are used to localize sub-pixel corners. In this paper, a novel approach for sub-pixel corner localization is proposed, and experiments are conducted to verify the new approach.

## **2. Related Work**

During the last two decades, a quantity of sub-pixel localization approaches have been proposed for ensuring the accuracy of camera calibration and pose estimation, which can be roughly divided into three categories discussed in this section.

### *2.1. Approaches Based on Image Gradient*

Sroba [15] shares a sub-pixel localization technique based on the observation that a vector from a corner to any part of its adjacent area is perpendicular to the image gradient of the corner. Points from the adjacent area are used to apply some mathematical treatments for solving a location iteratively; this location is taken as the new center of the adjacent area, until the center stays within a set threshold. Bok [16] adopts a sub-pixel finder algorithm based on Harris detector. From the given initial corner locations, the algorithm iteratively updates the individual corner locations to the largest gradient values using patch-based structure tensor calculation. The algorithm calculates the structure tensor by directly interpolating the gradients, instead of first interpolating the image and second computing the gradients for reducing computational costs. The first mentioned category has been implemented by a toolbox [17] and a library function [18] and, therefore, frequently employed in many application scenarios. These approaches can achieve high efficiency, but they are sensitive to image noise, and often lead to unstable results for on-site use.

### *2.2. Approaches Based on Grayscale Symmetry*

Chu [19] introduces a sub-pixel detector using a round template under image physical coordinates. The round template is employed to pass through a dilated image, and corners are ultimately determined by calculating the centroid of redundant points based on the symmetry of chessboard patterns. Zhao [20] proposes a method based on the property that the symmetry of a square region is more significant when the central pixel of it is closer to a corner. Symmetric factors of all pixels in a selected area are calculated for obtaining a weighted sub-pixel corner position. These approaches can achieve better results when detecting blurred or overexposed images. However, they need a bigger region of interest (ROI), and come at a higher computation cost due to the determination of symmetric factors using a sliding template or window and, therefore, they are subject to certain constraints of lens distortions and expensive for real-time use.

### *2.3. Approaches Based on Polynomial Fitting*

Lucchese [21] performs a least-squares fit of a quadratic polynomial to a low-pass version of the input image. The approach obtains saddle points from the polynomial coefficients and is invariant to affine transformations. Chen [22] computes intermediate values in the Harris-corner-like detection phase to obtain a second-order Taylor expansion of input image, saddle points are found based on a corner model restricted to orthogonal corners. Mallon [23] proposes an edge-based nonlinear corner localizer. The localizer performs a least-squares fit of a parametric edge model to an edge version of input image. Placht [24] develops a modified strategy inspired by the method mentioned in [21]. A corner is refined to sub-pixel accuracy by filtering the adjacent region around it using a 2-D cone filter for an intensity surface amenable to fitting a quadratic polynomial. On the premise of selecting a

reasonably small neighborhood, these approaches can yield a suitable approximation in the presence of nonlinear distortions and projective transformations, due to the affine invariance [25]. However, they require a filtered version of the input image for polynomial fitting, because their corner models are not accurate enough for direct processing, which still need to be optimized for improving reliability and efficiency of sub-pixel localization.

### 3. Methodology

In this section, an accurate model of chessboard images is established for localizing sub-pixel corners; the methodology is based on polynomial fitting and without being dependent on image filtering.

#### 3.1. Ideally Continuous Corner Model

It is intuitive to imagine that a corner is located at a junction of two edges, and has the smallest radius of curvature; pixels around it appear as a high change of brightness in all directions. As represented in Figure 1, a square region  $C$  with a center  $o = [0, 0]$  and an area of  $(2r + 1)^2$  pixels is observed to analyze a chessboard image for the following description. Geometrically, a straight line  $L$  passing through  $o$  can be given as

$$L : \chi(\omega, u, v) = u \sin \omega - v \cos \omega = 0, \text{ with } \{\omega, u, v\} \in \mathbb{R}, \tag{1}$$

where  $\omega$  is the angle of inclination. Using the sign function  $\text{sgn}(x)$  yields an ideal edge  $E$  related to  $L$  via

$$E(\omega, u, v) = \text{sgn}[\chi(\omega, u, v)]. \tag{2}$$

An ideally continuous chessboard image, with a gray value +1 in the white and -1 in the black regions, is then defined:

$$C_i(u, v) = E(\alpha, u, v)E(\beta, u, v), \quad 0 \leq \alpha < \pi, \quad \alpha < \beta < \pi + \alpha, \tag{3}$$

where  $\alpha$  and  $\beta$  are the angles similar to  $\omega$ , and determine two edges  $E_1$  and  $E_2$ . It is worth mentioning that Equation (3) is subject to a reasonably small  $r$ . Otherwise,  $E_1$  and  $E_2$  may be re-defined by two curve functions for a suitable approximation of lens distortions.

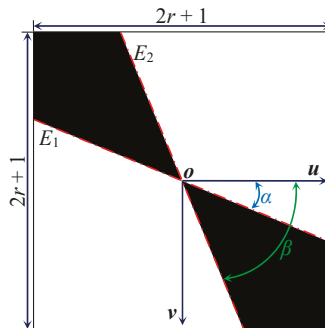


Figure 1. Definition of a chessboard image.

In actual imaging, however,  $C_i$  is inevitably blurred by the lens of a vision system. A point input, represented as a single pixel in  $C_i$ , will be reproduced as a spread region in a blurred image  $C_f$ . For practical purposes, the blurring response described by the point spread function (PSF) is always

approximated by a radio-symmetrical Gaussian kernel [26]. Similarly,  $C_f$  is modeled by convolving  $C_i$  with a 2-D Gaussian filter:

$$C_f(u, v) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{u^2 + v^2}{2\sigma^2}\right) \otimes C_i(u, v), \tag{4}$$

with  $\sigma$  effectively denoting the blurring strength.

Distinctly, the gray level of  $C_f$  is not in the same range as that of a digital image in the common use. Under the assumption that the vision system has a linear response to the light intensity within a reasonable range,  $C_f$  can be transformed by

$$C_s(u, v) = \lambda C_f(u, v) + \kappa, \tag{5}$$

with  $\kappa$  and  $\lambda$  related to the maximum and minimum gray values  $g_{\max}$  and  $g_{\min}$  of  $C_s$  via

$$\kappa = \frac{g_{\max} + g_{\min}}{2}, \quad \lambda = \frac{g_{\max} - g_{\min}}{2}. \tag{6}$$

### 3.2. Sub-Pixel Corner Localization

According to the existing techniques discussed in Section 2.3, since a real region  $R$  is detected with a known corner position  $c_p = [u_p, v_p]$  at pixel level, the ideal model,  $C_s$ , most similar to  $R$  (the highest PSNR), can be found by determining

$$\operatorname{argmin}_{\mu, v, \alpha, \beta, \lambda, \kappa, \sigma} \left[ \sum_{i=-r}^{+r} \sum_{j=-r}^{+r} (\varepsilon_{i,j})^2 \right], \quad \text{with } \varepsilon_{i,j} = C_s(i + \mu, j + v) - R(i + u_p, j + v_p), \tag{7}$$

where  $\mu$  and  $v$  form a vector  $d$  from the ideal corner position to the center of  $C_s$ . It is evident that the closed expression of  $C_f$  is required to address the above optimization by common means, e.g., the Gauss–Newton method. Despite the fact that Equation (4) cannot be directly analyzed by anti-derivatives, it is approximately evaluated by separating the Gaussian kernel and using integration by parts, given as

$$C_f(u, v) = \operatorname{erf}\left[\frac{\chi(\alpha, u, v)}{\sqrt{2}\sigma}\right] \operatorname{erf}\left[\frac{\chi(\beta, u, v)}{\sqrt{2}\sigma}\right] + \Delta(u, v), \tag{8}$$

$$\Delta(u, v) \approx \left(1 - \frac{4\theta_1}{\pi}\right) \left\{ 1 - \operatorname{erf}^2 \left[ \sqrt{\frac{\delta_1 \chi^2(\theta_2, u, v) + \delta_2 \chi^2\left(\frac{\pi}{2} + \theta_2, u, v\right)}{2\sigma^2}} \right] \right\}, \quad \text{with} \tag{9}$$

$$\begin{cases} \theta_1 = \frac{\beta - \alpha}{2} \\ \theta_2 = \frac{\beta + \alpha}{2} \end{cases}, \quad \begin{pmatrix} \delta_1 & \delta_2 \end{pmatrix} = \begin{cases} \begin{pmatrix} 1 & \tan \theta_1 \\ \cot \theta_1 & 1 \end{pmatrix} & \text{if } \theta_1 < \frac{\pi}{4} \\ \begin{pmatrix} 1 & \tan \theta_1 \\ \cot \theta_1 & 1 \end{pmatrix} & \text{otherwise} \end{cases},$$

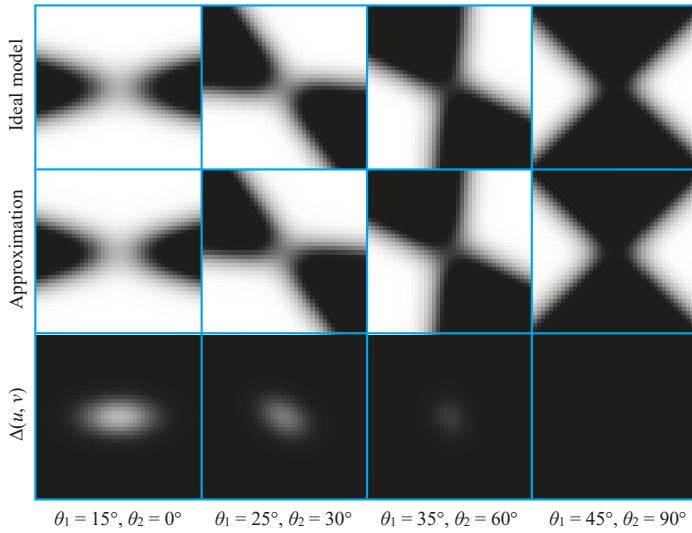
where  $\Delta(u, v)$  is the integral remainder term,  $\operatorname{erf}(x)$  denotes the Gaussian error function, and  $\theta_1$  and  $\theta_2$  are also known as the angles of shear and rotation in image plane. The corner model, approximated in Equations (8) and (9), share a high similarity with the ideal one in Equation (4), due to an effective estimation and compensation of  $\Delta(u, v)$  (Figure 2).

However, there is still a lack of the closed form for the Gaussian error function; an accurate but expensive way is replacing it by piecewise polynomials [27]. Considering that most applications utilize 8-bit gray images (256 gray-levels), this replacement should be a balance between the computational accuracy and efficiency. Alternatively, a low-cost approximation  $\tanh(\rho x)$  is used, and leads to an acceptable result by selecting a suitable value for the coefficient  $\rho$  (Figure 3).

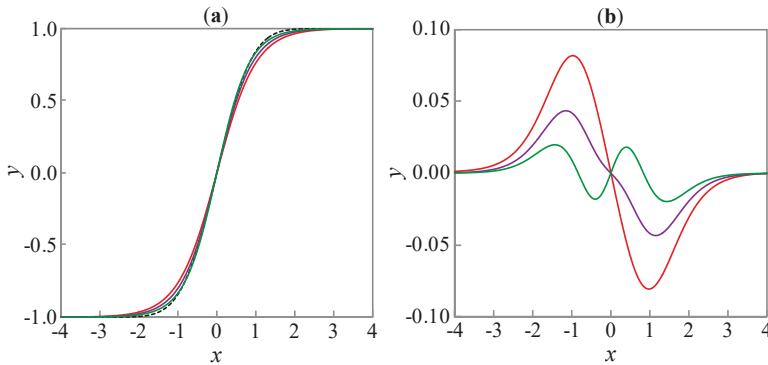
Finally, Equation (7) can be achieved using a linear optimization in iterations.  $\mu$  and  $v$  are initialized to 0 and  $\sigma$  to 1.  $\alpha$  and  $\beta$  are initialized based on edge extraction [28].  $\kappa$  and  $\lambda$  are initialized using the gray values in the black and white areas close to  $c_p$ . Generally, about 14 pixels are suitable

for  $r$  with an overall consideration of the lens distortions, image noise, and computational efficiency. After sufficient iterations for the system convergence, sub-pixel corner  $c_s$  can be calculated from

$$c_s = c_p - d. \tag{10}$$



**Figure 2.** Gray images simulated by the ideal model, approximated model, and  $\Delta(u, v)$ , with different values of  $\theta_1$  and  $\theta_2$ . The image size is  $41 \times 41$  pixels,  $\sigma = 4$ .



**Figure 3.** Function curve plots of (a):  $y = \text{erf}(x)$  (dashed),  $y = \tanh(\rho x)$  (solid), and (b):  $y = \tanh(\rho x) - \text{erf}(x)$ . The coefficient  $\rho$  is set to 1.0 (red), 1.1 (purple), and 1.2 (green).

### 3.3. Self-Checking for Perspective-n-Point

Resulting from a maximum likelihood estimation of  $C_s$ , the residual  $\epsilon_{i,j}$  can be used to evaluate the quality of model fit. Let  $\acute{E}$  be similar to the root-mean-square error (RMSE), and expressed as

$$I_o = \frac{1}{2r + 1} \sqrt{\sum_{i=-r}^{+r} \sum_{j=-r}^{+r} (\epsilon_{i,j})^2}. \tag{11}$$

The factor  $\acute{E}$  is slightly and unavoidably affected by the lens distortions and image noise, but remarkable when there is a great deal of light pollution (or imbalanced illumination) brought into a chessboard image, which often happens in on-site applications. Thus, it can also be considered as a metric to reflect the reliability of sub-pixel localization. In order to achieve a self-checking technique for perspective-n-point, boxplot analysis is more appropriate than conventional means, e.g., 3-sigma rule [29]. The boxplot distinguishes outliers using quantiles (Figure 4), rather than depending on a prior knowledge about the distribution of actual dataset and, therefore, it has a higher flexibility.

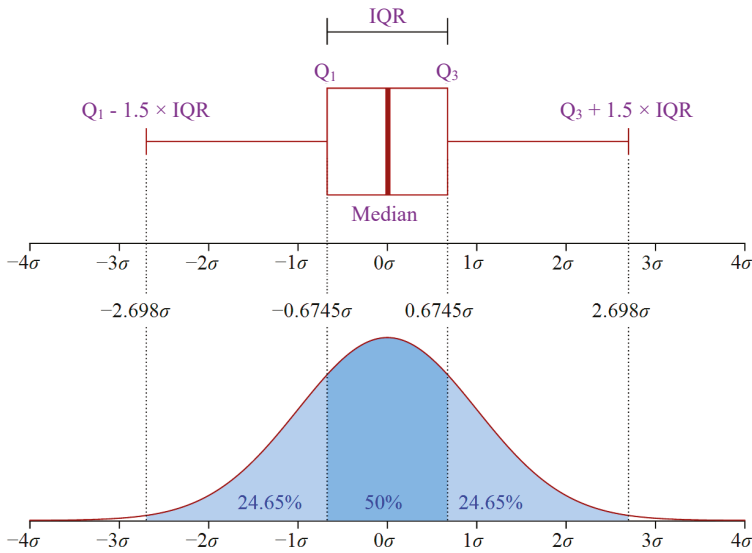


Figure 4. Boxplot with respect to a probability density function of  $N(0, \sigma^2)$ .

It is assumed that a chessboard image is detected with an array of sub-pixel corners (the array size is  $M \times N$ ). Using Equation (11) gives a corresponding metric  $\acute{E}_{m,n}$  for the corner  $c_{m,n} = [u_{m,n} \ v_{m,n}]$ ,  $m \in \{1, \dots, M\}$ ,  $n \in \{1, \dots, N\}$ . As illustrated in Figure 4, since the quantiles  $Q_1$  and  $Q_3$  are obtained by investigating all the metrics, the factor  $w_{m,n}$ , standing for the reliability of  $c_{m,n}$ , is then determined:

$$w_{m,n} = \begin{cases} 1 & \text{if } Io_{m,n} \in [2.5Q_1 - 1.5Q_3 \quad 2.5Q_3 - 1.5Q_1] \\ 0 & \text{otherwise} \end{cases} \quad (12)$$

Perspective-n-point is the problem of estimating a 3-D rotation  $r$  and a translation  $t$  of a calibrated camera, with respect to the world frame. Since the chessboard mentioned above is defined in the world frame accurately, 3-D points in it and their corresponding image points follow a pin-hole model for the camera [2]:

$$s_{m,n} \mathbf{n}_{m,n} = r \mathbf{q}_{m,n} + \mathbf{t}, \text{ with } \mathbf{n}_{m,n} = \begin{bmatrix} \frac{u_{m,n} - u_0}{f_x} & \frac{v_{m,n} - v_0}{f_y} & 1 \end{bmatrix}^T, \quad (13)$$

where  $\mathbf{q}_{m,n}$  is the  $(m, n)^{\text{th}}$  3-D point and  $s_{m,n}$  the corresponding scale factor.  $f_x$  and  $f_y$  are the scaled focal lengths,  $[u_0 \ v_0]$  is the principal point. An optimal solution of  $s_{m,n}$  is related to the given  $r$  and  $t$  via

$$s_{m,n} = (\mathbf{n}_{m,n}^T \mathbf{n}_{m,n})^{-1} \mathbf{n}_{m,n}^T (r \mathbf{q}_{m,n} + \mathbf{t}). \quad (14)$$



Considering  $w_{m,n}$  as the penalty factor in association with the above estimator yields

$$\operatorname{argmin}_{r,t} \left\{ \sum_{m=1}^M \sum_{n=1}^N \left\| w_{m,n} \left[ \left( n_{m,n}^T n_{m,n} \right)^{-1} n_{m,n} n_{m,n}^T - I \right] (rq_{m,n} + t) \right\|^2 \right\}. \quad (15)$$

Rodrigues parameters, instead of Euler rotations, are recommended for simplifying the above optimization [30]. It is worth being pointed out that sub-pixel corners obtained according to Section 3.2 cannot be directly taken as image points, which need to be corrected beforehand, due to lens distortions [2].

#### 4. Evaluation

In this section, experiments on synthetic and real datasets are conducted to verify the proposed method with three references detailed in literatures [16,20,24]. In order to decrease the influence on localization result due to different parameter settings, for both the proposed and referenced methods, each chessboard corner is detected with the same initial pixel coordinates, and refined from the same local neighborhood with a square size of  $31 \times 31$  pixels.

##### 4.1. Synthetic Data

In order to acquire synthetic chessboard image, a pin-hole camera is simulated with the properties:  $[f_x, f_y] = [7000, 7000]$ ,  $[u_0, v_0] = [1296, 972]$ . The image resolution is set to  $2592 \times 1944$ . A single chessboard pattern with 20 mm cell size in both directions is projected to the image plane. Since optical paths are reversible, an ideal projection from the pattern center can be found and defined as ground truth. Gaussian blur with the window parameter  $\sigma_f$  and Gaussian noise with 0 mean and standard deviation  $\sigma_n$  are added to make the image similar in appearance to a real one (Figure 5). For each given  $\sigma_f$  and  $\sigma_n$ , 100 independent trials are performed, with other simulation parameters varied and limited in their ranges (Table 1), under the premise of ensuring faultless projections.

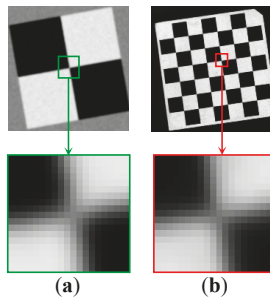


Figure 5. Chessboard images captured by (a) simulation and (b) real device.

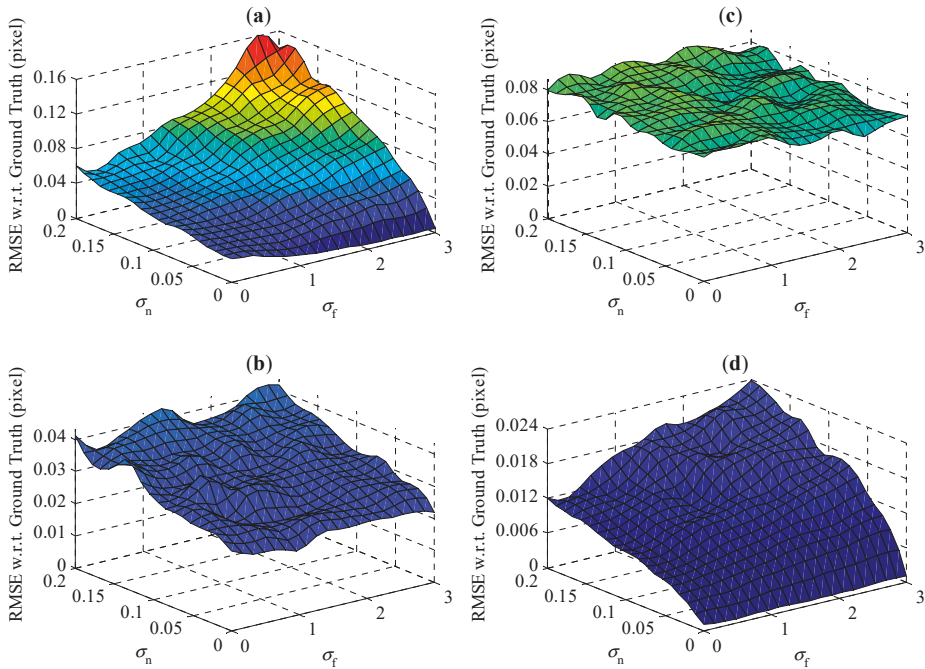
Table 1. Range set of simulation parameters.  $yaw$ ,  $pitch$ , and  $roll$  are the Euler angles related to  $r$ .  $t_x$ ,  $t_y$ , and  $t_z$  are the dimensional elements of  $t$ .

$g_{\max}$	$g_{\min}$	$yaw, pitch, roll$	$t_x$ (mm)	$t_y$ (mm)	$t_z$ (mm)
[191, 255]	[0, 63]	$[-\pi/4, \pi/4]$	[-40, 40]	[-30, 30]	[950, 1050]

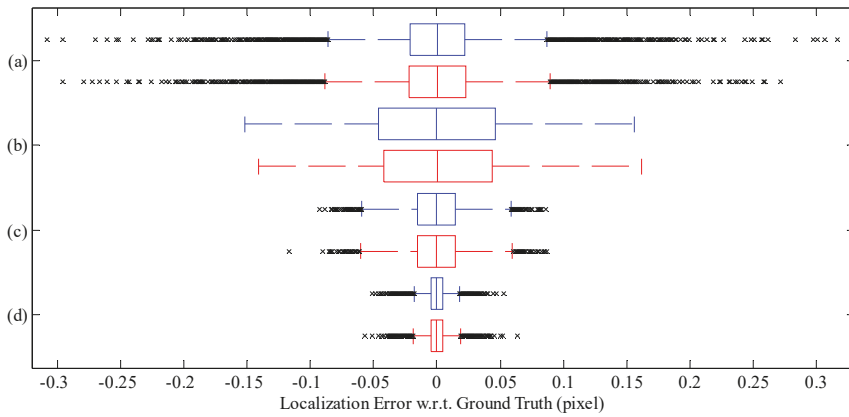
Figure 6 depicts the RMS error of sub-pixel localization as a function of  $\sigma_f$  and  $\sigma_n$ . The proposed technique performs significantly better than the referenced ones. Although it results in a higher error due to the increase of  $\sigma_f$  and  $\sigma_n$ , the performance drop is not as pronounced as for the others. Concretely, for the poorest image quality ( $\sigma_f = 3$ ,  $\sigma_n = 0.2$ ), the result shows that the errors are about 0.154, 0.041, 0.077, and 0.024 pixels for [16,20,24], and the proposed technique, respectively. Remarkably,

Placht et al. [24] yields a stable, but significant, error in the presence of the change of  $\sigma_f$  and  $\sigma_n$  for taking filtered images as inputs. That is to say, it not only eliminates noise distinctly, but also leads to an extra uncertainty of sub-pixel localization.

In addition, sub-pixel localization errors from all trials (the total number is 40,000) are gathered for an overall evaluation represented by boxplots. As shown in Figure 7, for the proposed and referenced methods, interquartile ranges (IQRs) are highly symmetrical about medians pretty close to zero. In detail, the IQRs are about 0.18, 0.13, 0.32, and 0.04 pixels in both directions for [16,20,24], and the proposed method, respectively. The smaller IQR reflects the better performance of sub-pixel localization. Again, using filtered images as inputs lead to a particular outcome, that there are no outliers to be distinguished with the largest IQR for [24].

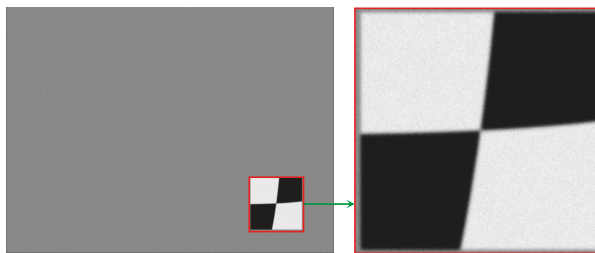


**Figure 6.** Localization error with respect to blur strength  $\sigma_f$  and noise level  $\sigma_n$  for (a) [16], (b) [20], (c) [24], and (d) the proposed technique.



**Figure 7.** Boxplots of the errors between localized and standard values in  $u$  (red boxes) and  $v$  (blue boxes) directions, regarding (a) [16], (b) [24], (c) [20], and (d) the proposed method.

The above simulation relies on the assumption that edges defining a corner are completely straight in the observation area, or region of interest, where the corner is going to be found. However, it is well known that lenses inevitably have distortions. To obtain maximum allowable distortions for the method, another simulation is conducted, with the fixed blur strength and noise level ( $\sigma_f = 1.5$ ,  $\sigma_n = 0.1$ ), and the first order radial distortion with the degree  $k_1$  is added to the image (Figure 8). Again, for each given  $k_1$ , 100 independent trials are performed, with other simulation parameters varied and limited in their ranges (Table 1), except for  $[t_x, t_y, t_z]$  set to [115, 80, 1000], for ensuring the projections farther away from the principal point.



**Figure 8.** Chessboard image captured by simulation with the first order radial distortion,  $k_1 = -5$ .

Figure 9 depicts the RMS error of sub-pixel localization as a function of  $k_1$ . The highest errors are 0.089 pixels for [16], 0.046 pixels for [20], 0.123 pixels for [24], and 0.037 pixels for the proposed method. Again, the proposed method performs significantly better than the referenced ones when  $k_1$  varies from  $-5$  to  $5$ . Different from [20] and the proposed method, Bok et al. and Placht et al. [16,24] show a distinct variability due to the limitation of their methodologies; the blur strength and noise level in the simulation have greater impact on the localization result than the distortion. For practical applications, however, cameras with the coefficient  $k_1$  larger than  $5$  are lesser used in photogrammetry because the pinhole model is no longer applicable for them. Therefore, for calibrating a camera for common use, the proposed method can be effectively performed without any pretreatment.

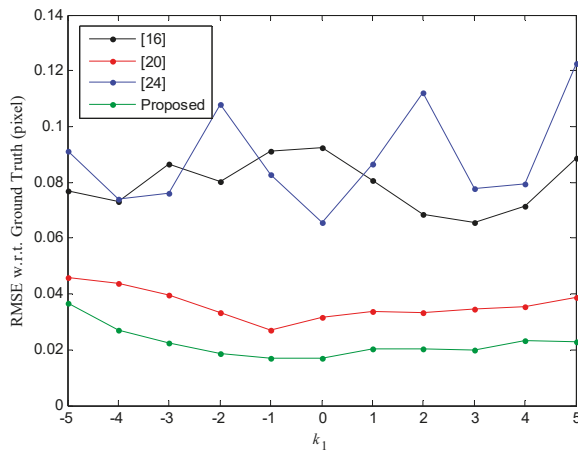


Figure 9. RMS error between localized and standard values as a function of  $k_1$  for four different approaches.

4.2. Real Data

In contrast to simulations, real data experiments cannot directly evaluate the accuracy of sub-pixel localization via the observed corner coordinates, due to their undetermined ground truth data. An alternative and indirect way is examining it based on camera calibration technique. Figure 10 shows that a camera (JPLY, G1GD05C) with 16 mm lens and  $2592 \times 1944$  image resolution is employed for conducting a camera calibration experiment based on a coordinate measuring machine (CMM) (Brown & Sharpe, Global Image 7107) with a single chessboard pattern ( $20 \times 20$  mm cell size) mounted on the end of its probe. 3-D control points are achieved by programmatically driving the probe to a set of specially designed positions, and provided with a dimensional error of less than 0.003 mm in both directions. For each position, the chessboard pattern is recorded by the camera for capturing a corresponding corner. Since all corners are located at the sub-pixel level, the camera can be calibrated based on bundle adjustment [4,6].

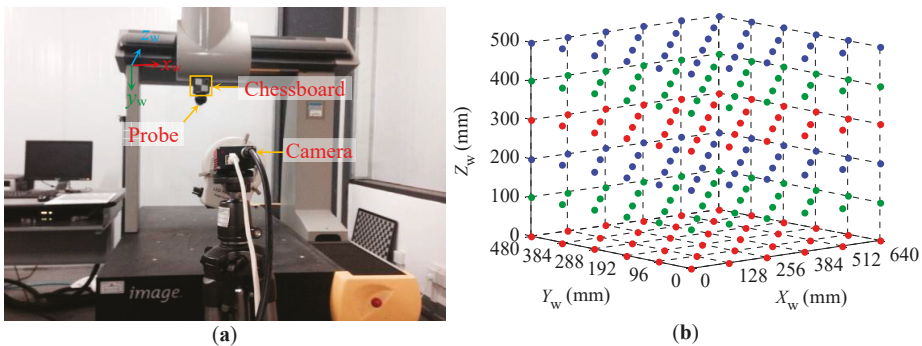


Figure 10. Camera calibration using (a) single chessboard and CMM for achieving (b) 3-D control points with specially designed positions.

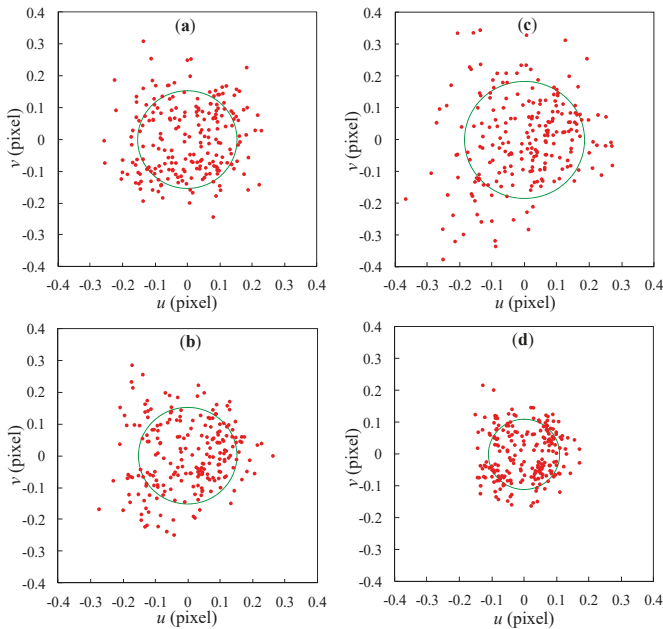
Table 2 lists the result of intrinsic parameters calibrated from the corners based on four different approaches. According to the definition of radial distortion coefficients detailed in [2], for [16,20,24] and the proposed method, the maximum distortions evaluated using the image point furthest from the principal point are 22.78, 25.92, 29.33, and 24.54 pixels in the radial direction, respectively. Among them,

the contributions of  $k_2$  are 2.03 pixels for [16], 6.06 pixels for [20], 10.06 pixels for [24], and 4.62 pixels for the proposed method. Therefore,  $k_2$  has a much smaller influence on the pixel offsets than  $k_1$ . Or rather, the estimator of  $k_2$  is more sensitive to noise in the corner coordinates. In spite of the fact that the result cannot intuitively demonstrate the performance of each approach, it is pivotal for the following investigations.

**Table 2.** Calibration result for four different approaches.  $k_1$  and  $k_2$  denote the 1st and 2nd order radial distortion coefficients.

	$[u_0, v_0]$	$[f_x, f_y]$	$[k_1, k_2]$
[16]	[1344.42, 937.37]	[7296.06, 7298.85]	[0.2332083, 0.4313581]
[20]	[1342.78, 936.71]	[7298.67, 7302.12]	[0.2237942, 1.2905459]
[24]	[1340.02, 937.04]	[7297.59, 7300.76]	[0.2118124, 2.2750773]
Proposed	[1343.70, 936.62]	[7299.13, 7302.50]	[0.2253345, 0.9926340]

Figure 11 represents four scatter plots of re-projection errors. For general examinations, the maximum and mean re-projection errors for the proposed method are 0.22 pixels and 0.11 pixels, evidently less than 0.32 pixels and 0.15 pixels for [16], 0.29 pixels and 0.15 pixels for [20], 0.39 pixels and 0.18 pixels for [24]. From the standpoint of addressing perspective-n-point problem, the re-projection errors, assessing the validity of calibration, are subjected to some optical indications, e.g., image and lens resolutions, and integrated with certain methodologies, including calibration model, target geometry, and sub-pixel localization. The mentioned experiment employs a robust model with stereo points establishing correspondences between world and image frames accurately and, therefore, the lower re-projection errors not only reflect the better solution of perspective-n-point, but also testify the higher accuracy of sub-pixel localization. Therefore, the corners obtained using the proposed technique are better suited for camera calibration.



**Figure 11.** Scatter plots of re-projection errors (red dots) for (a) [16], (b) [20], (c) [24], and (d) proposed method. In each sub-figure, green circle is rendered with a radius equal to the mean re-projection error.

In order to alternatively examine the proposed technique, different measurements on displacement and attitude are carried out using the CMM and camera mentioned above. Firstly, for displacement measurement, a target ( $6 \times 6$  grid of points,  $20 \times 20$  mm cell size) fixed on the end of the probe is moved with the guide and imaged by the camera placed in front of the CMM, for measuring a distance  $d$  between two different positions as an evaluation factor (Figure 12). Secondly, for attitude measurement, two targets,  $T_1$  and  $T_2$ , mounted on the base with the same grid and cell as that of the above measurement, are imaged by the camera (Figure 13). Among the three axis vectors, only the one in the z direction can be perfectly measured using the probe (Renishaw, SP600), by scanning the pattern plane of each target, due to a restriction that makes it hard to capture 3-D coordinates of a corner accurately, by means of contact measurements. Thus, the included angle  $\theta$  between two normal vectors is adopted as another evaluation factor more suitably. Fifteen independent trials are performed to localize sub-pixel corners, employing both the proposed and referenced methods, and estimate the camera poses from their respective intrinsic parameters listed in Table 2. The metrics  $d$  and  $\theta$  are then computed for investigating discrepant deviations with respect to the CMM data.

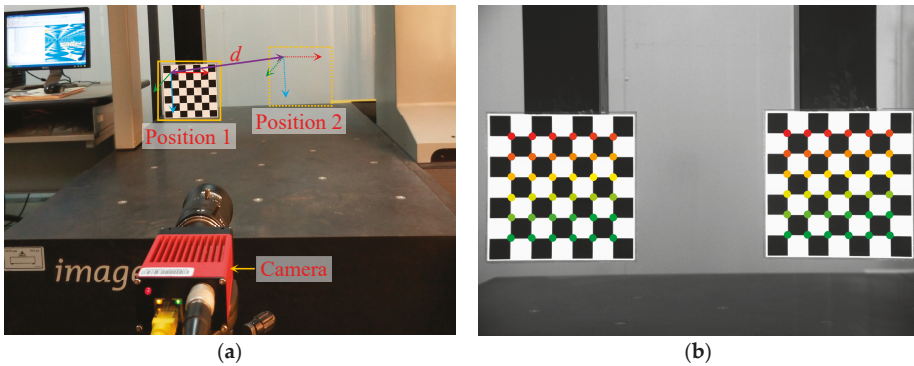


Figure 12. Experiment for measuring displacement. (a) Determining  $d$  via CMM and camera. (b) A merged image of two positions with located corners.



Figure 13. Experiment for measuring attitude. (a) Determining  $\theta$  via CMM and camera. (b) One shot in 15 trials with located corners.

Figure 14 presents the results from the above measurements. Under the premise that the CMM provides baselines with a higher accuracy, the RMS errors of  $d$  and  $\theta$  are 0.032 mm and  $0.010^\circ$  for [16], 0.021 mm and  $0.009^\circ$  for [20], 0.037 mm and  $0.013^\circ$  for [24], and 0.014 mm and  $0.006^\circ$  for the proposed approach. Although there are many estimable and inestimable influences during the experiments,

the results are mainly dependent on the accuracies of intrinsic parameters and corner coordinates, and essentially subject to the performance of each sub-pixel localization method because the camera is also calibrated from the respective corner set. From a synthetic point of view, exact values of  $d$  and  $\theta$  are derived from reliable estimations of camera poses predetermined by accurate corner coordinates. As an apparent outcome of the comparison, the proposed technique presents a higher performance than others.

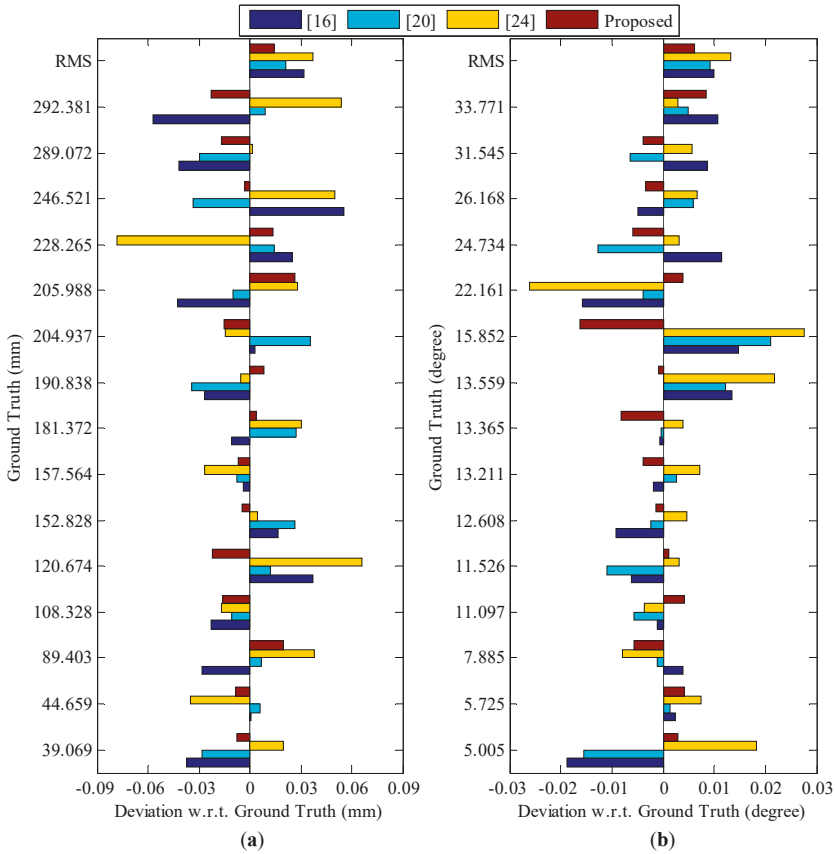


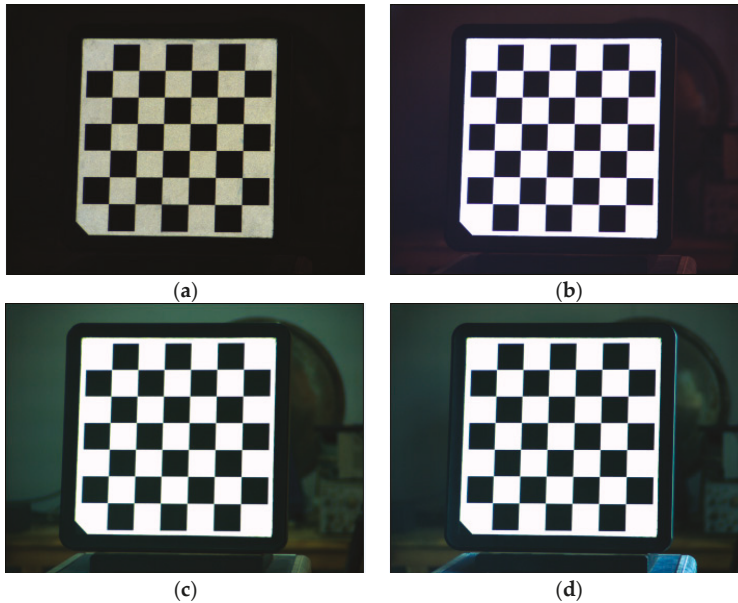
Figure 14. Measurement results of (a) displacement  $d$  and (b) attitude angle  $\theta$ .

As shown in Figure 15, in order to test the proposed approach in terms of its robustness to real-world data gathering, four images of a stationary chessboard are captured by the mentioned camera under underexposed, overexposed, indoor light interfered, and outdoor light interfered scenarios. For each corner in a  $6 \times 6$  array, its maximin deviation between different scenarios is computed and gathered for an overall evaluation.

Table 3 lists the overall evaluation result for four different approaches. The RMS deviations are 0.419 pixels for [16], 0.287 pixels for [20], 0.396 pixels for [24], and 0.241 pixels for the proposed approach. Considering the fact that the relative pose between the target and camera is stationary, the variability of each detected corner is mainly subject to the robustness of corner localization in the presence of the ambient light changes. The smallest RMS deviation proves that the proposed approach has higher interference immunity, resulting from a more robust corner model.

**Table 3.** Overall evaluation result for four different approaches.

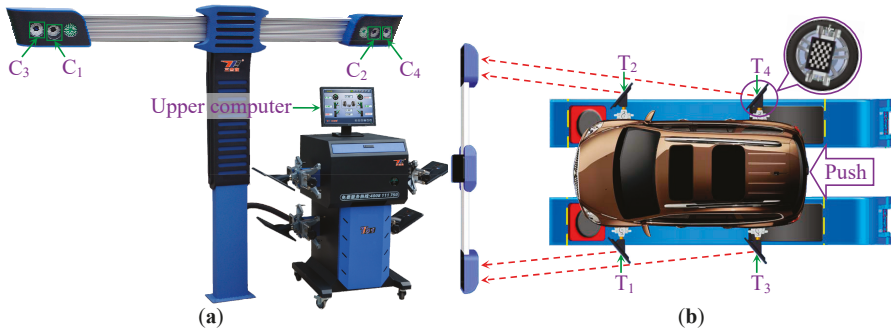
	[16]	[20]	[24]	Proposed
RMSD (pixel)	0.419	0.287	0.396	0.241

**Figure 15.** Four images of a stationary chessboard captured under (a) underexposed, (b) overexposed, (c) indoor light interfered, and (d) outdoor light interfered scenarios.

#### 4.3. Practical Application

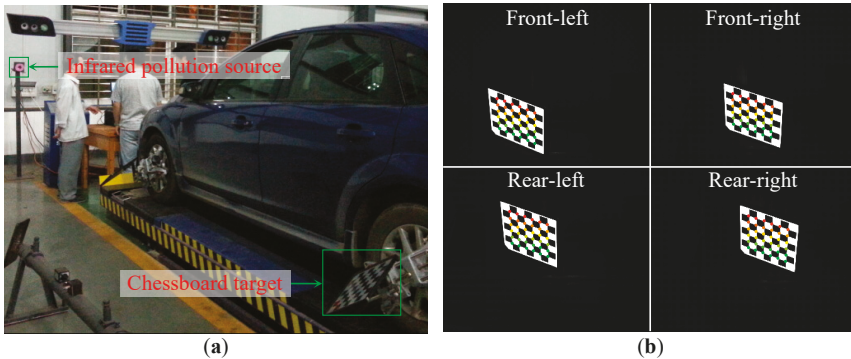
The proposed approach is implemented in a visual measurement system called 3D four-wheel aligner (3Excel, T50). The system, designed for aligning four automobile wheels, mainly consists of an upper computer and four cameras and chessboard targets (Figure 16). Each camera is equipped with infrared filter and illuminant, for ensuring a high immunity to the complicated imaging conditions at customer sites. During an initial operation, the automobile under test is driven up to a certain distance by external force. Meanwhile, the cameras  $C_1$  to  $C_4$  are triggered in synchronous mode to capture image sequences of the targets  $T_1$  to  $T_4$  mounted on the front-left, front-right, rear-left, and rear-right wheels, respectively. For each image sequence, sub-pixel corners are detected for estimating a wheel attitude with respect to the corresponding camera; two alignment parameters *toe-in/toe-out* and *camber* are then determined by decomposing angles of the wheel attitude unified in a global frame defined by the bodywork. During a real-time alignment, the parameters are dynamically calculated from continuous estimations of the wheel attitude changes with respect to their initial values.





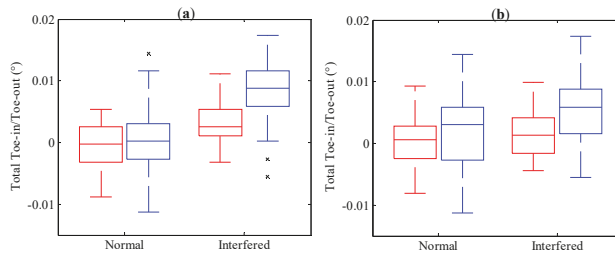
**Figure 16.** 3D four-wheel alignment. (a) System composition. (b) Initial operation.

As demonstrated in Figure 17, an automobile (Ford Focus) in healthy condition is used for on-site alignment. The introduced aligner can capture chessboard images with black backgrounds due to the usage of infrared filters and illuminants. After finishing the initial operation, the alignment is carried out and divided into two periods: one performs normally, and the other is interfered by infrared pollution sources. During each period, the alignment parameters are incessantly computed based on both the proposed and built-in techniques, until the number of their recorded values reaches 120.



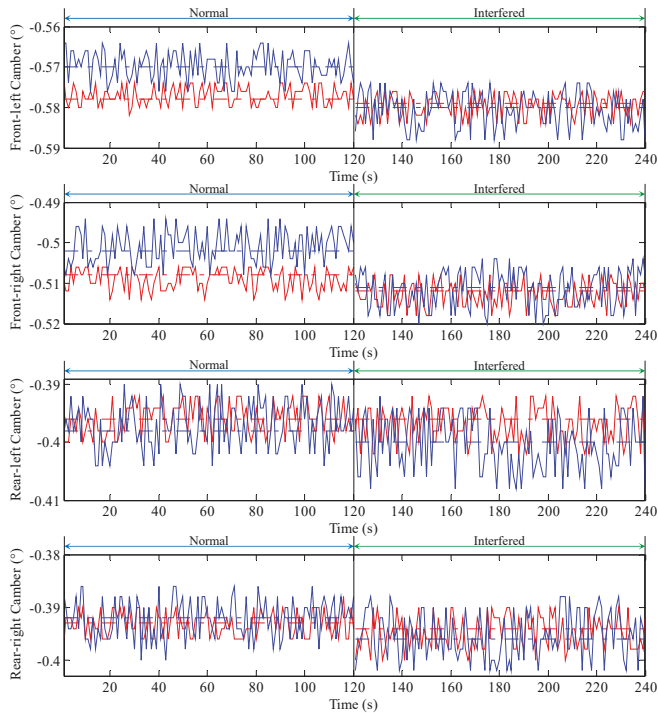
**Figure 17.** On-site aligning experiment. (a) Using infrared light as pollution source. (b) One shot in image sequence with located corners for each wheel position.

Figure 18 shows two boxplots of total *toe-in/toe-out* for the front and rear wheel-sets. This parameter, called *toe-in* for positive and *toe-out* for negative values, is defined for investigating the symmetry of each wheel-set about the geometric centerline (or thrust line). For both normal and interfered periods, the proposed method results in a median closer to zero and IQR of minor scope, compared with the built-in algorithm. Considering the fact that two total values should be pretty small because of the healthy condition of the automobile, the boxplots prove that the proposed method shows better central tendency, due to the accurate corner localization. When comparing the medians of the proposed method during two periods, the deviations between them are about  $0.003^\circ$  and  $0.002^\circ$  for the front and rear wheel-sets, significantly less than that of the built-in algorithm ( $0.009^\circ$  and  $0.003^\circ$ ), which also shows that the proposed method has a higher interference immunity resulting from the self-checking technique.



**Figure 18.** Boxplots of (a) front and (b) rear toe-in/toe-out values for the proposed (red boxes) and built-in (blue boxes) techniques. Regarding normal and interfered periods.

Figure 19 depicts four curve plots of *camber* as a function of time stamp for the front-left, front-right, rear-left, and rear-right wheel positions, which are divided into two parts, according to two different periods of the alignment. This parameter is defined for measuring the inclination of a wheel with respect to vertical line of the bodywork. Different from *toe-in/toe-out*, it is separately investigated using the wheel attitude, and weakly restricted to the absolute symmetry about its baseline for the corresponding wheel-set and, therefore, a total value makes poor sense for the evaluation. However, when observing the median change between two periods of each front wheel, there is a strong comparison that the difference is less than  $0.003^\circ$  for the proposed method, and more than  $0.008^\circ$  for the built-in algorithm.



**Figure 19.** Camber as a function of time stamp for each wheel position. Solid curves (dashed lines) in red and blue denote function values (medians) for the proposed and built-in techniques, respectively. Regarding normal and interfered periods.

It should be remarked that both methods yield median changes of the rear positions smaller than that of the front ones. This can be found from both Figures 18 and 19, and especially for the built-in technique. There is a logical explanation, as follows: the distances from the front and rear wheels to the infrared pollution source are about 1.5 m and 3.9 m, respectively. The energy of interference is in a state of decay when the distances become larger and, therefore, has no pivotal influence on corner localization and pose estimation for the rear wheels. That is to say, when there is a lack of robust localization technique, a direct way to improve system accuracy is enhancing image quality. Or rather, perspective-n-point is prone to errors if there are outliers in the set of point correspondences. Thus, the self-checking technique can be used in conjunction with existing solutions to make the final solution for the camera pose more robust to outliers.

#### 4.4. Computational Efficiency Test

One hundred real images are obtained using the camera and  $6 \times 6$  chessboard mentioned in Section 4.2, for testing computational efficiency of the proposed method. An optimized dynamic library of it is implemented in C++ code (available online: <https://pan.baidu.com/s/1PgRl3qG8HDi49f8n8Jwe3Q>), to make a more objective analysis in terms of processing time compared with two mature functions “findChessboardcorners” and “cornerSubPix” built-in OpenCV. The test is run in VS2010 installed on a desktop computer (CPU: Intel Core i7-6700; RAM: DDR4-2133 16GB; HDD: 1TB). All the images are preloaded in the RAM for an undifferentiated access performance, instead of an unstable reading speed of the HDD.

Table 4 lists the result of the processing time for three different algorithms. Although “cornerSubPix” runs two times as fast as the proposed method due to a low-cost computation based on image gradient, it is not essential for real-time detection, because sub-pixel corners are refined from their pixel coordinates located by expensive pretreatments. There is a common view that “findChessboardcorners” has high performance for rough detection. When comparing with two sub-pixel algorithms, however, it costs 41,715 ms, almost 32 and 14 times longer than that of “cornerSubPix” and the proposed algorithm. Therefore, the present efficiency bottleneck is the pixel detection, not the sub-pixel refinement. What can be expected is that this bottleneck is not unbreakable; some state-of-the-art techniques, such as CUDA and multithread computing, are powerful for addressing this kind of problem.

**Table 4.** Processing time of detecting 100 chessboard images for three different algorithms.

	findChessboardcorners	cornerSubPix	Proposed
Time (ms)	41,715	1296	3051

## 5. Summary

In this work, a new approach is proposed to localize chessboard corners at sub-pixel level. The proposed approach is based on an ideal chessboard model, established as a function of corner coordinates, rotation and shear angles, gain and offset of grayscale, and blurring strength. In order to localize the sub-pixel corner using a nonlinear fit to input image directly, the ideal chessboard model is approximated by a low-cost and high-similarity expression in the closed form. In order to ensuring the reliability of perspective-n-point, a self-checking technique for pose estimation is proposed by investigating qualities of model fits. The proposed approach has the following superiorities: (1) the methodology is effective without being dependent on image filtering employed as the pretreatment in the references; (2) the approximated corner model is more accurate than that in the references and has a high performance; (3) the self-checking technique, in association with existing solutions, is powerful for on-site use.

**Author Contributions:** Conceptualization, T.Y. and Q.Z. (Qiancheng Zhao); Methodology, T.Y.; Software, X.W.; Writing—review and editing, Q.Z. (Quan Zhou) and T.Y.

**Funding:** This work was supported by the National Nature Science Foundation of China [No: 51405154, 51275169]; the Hunan Provincial Natural Science Foundation of China [No: 2015JJ5012]; and the Hunan Provincial Innovation Foundation for Postgraduate of China [No: CX2016B545].

**Acknowledgments:** The authors are very grateful to Shenzhen 3Excel Tech. Co., Ltd. for providing experimental resources and technical supports.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Penate-Sanchez, A.; Andrade-Cetto, J.; Moreno-Noguer, F. Exhaustive Linearization for Robust Camera Pose and Focal Length Estimation. *IEEE Trans. Pattern Anal. Mach. Intell.* **2013**, *35*, 2387–2400. [CrossRef] [PubMed]
2. Zhang, Z. A flexible new technique for camera calibration. *IEEE Trans. Pattern Anal. Mach. Intell.* **2000**, *22*, 1330–1334. [CrossRef]
3. Sužiedelytė-Visockienė, J. Accuracy analysis of measuring close-range image points using manual and stereo modes. *Geodesy Cartogr.* **2013**, *39*, 18–22. [CrossRef]
4. Bundle Adjustment. Available online: [http://en.wikipedia.org/wiki/Bundle\\_adjustment/](http://en.wikipedia.org/wiki/Bundle_adjustment/) (accessed on 5 April 2018).
5. Lepetit, V.; Moreno-Noguer, F.; Fua, P. EPnP: An Accurate O(n) Solution to the PnP Problem. *Int. J. Comput. Vis.* **2009**, *81*, 155–166. [CrossRef]
6. Wang, Y.; Yuan, F.; Jiang, H.; Hu, Y. Novel camera calibration based on cooperative target in attitude measurement. *Optik* **2016**, *127*, 10457–10466. [CrossRef]
7. Harris, C.; Stephens, M. A combined corner and edge detector. In Proceedings of the Alvey Vision Conference, Manchester, UK, 31 August–2 September 1988; pp. 147–151.
8. KLT: An Implementation of the Kanade-Lucas-Tomasi Feature Tracker. Available online: <http://cecas.clemson.edu/~stb/klf/> (accessed on 26 October 2018).
9. Escalera, A.D.; Armingo, J.M. Automatic chessboard detection for intrinsic and extrinsic camera parameter calibration. *Sensors* **2010**, *10*, 2027–2044. [CrossRef] [PubMed]
10. Tu, D.; Zhang, Y. Auto-detection of chessboard corners based on grey-level difference. *Opt. Precis. Eng.* **2011**, *19*, 1360–1365.
11. Liu, Y.; Liu, S.; Cao, Y.; Wang, Z. Automatic chessboard corner detection method. *IET Image Process* **2016**, *10*, 16–23. [CrossRef]
12. Yang, S.; Scherer, S.A.; Yi, X.; Zell, A. Multi-camera visual SLAM for autonomous navigation of micro aerial vehicles. *Robot. Auton. Syst.* **2017**, *93*, 116–134. [CrossRef]
13. Song, L.; Wang, M.; Lu, L.; Huan, H. High precision camera calibration in vision measurement. *Opt. Laser Technol.* **2007**, *39*, 143–1420. [CrossRef]
14. Zhang, T.; Liu, J.; Liu, S.; Tang, C.; Jin, P. A 3D reconstruction method for pipeline inspection based on multi-vision. *Measurement* **2017**, *98*, 35–48. [CrossRef]
15. Sroba, L.; Ravas, R.; Grman, J. The Influence of Sub-pixel Corner Detection to Determine the Camera Displacement. *Procedia Eng.* **2015**, *100*, 834–840. [CrossRef]
16. Bok, Y.; Ha, H.; Kweon, I.S. Automated checkerboard detection and indexing using circular boundaries. *Pattern Recognit. Lett.* **2016**, *71*, 66–72. [CrossRef]
17. Camera Calibration Toolbox for Matlab. Available online: [http://www.vision.caltech.edu/bouguetj/calib\\_doc/](http://www.vision.caltech.edu/bouguetj/calib_doc/) (accessed on 1 September 2018).
18. Camera Calibration and 3D Reconstruction. Available online: <http://docs.opencv.org/2.4/modules/imgproc/doc/> (accessed on 5 September 2018).
19. Chu, J.; Lu, A.G.; Wang, L. Chessboard corner detection under image physical coordinates. *Opt. Laser Technol.* **2013**, *48*, 599–605. [CrossRef]
20. Zhao, Q.; Chen, Z.; Yang, T.; Zhao, Y. Detection of sub-pixel chessboard corners based on gray symmetry factor. In Proceedings of the SPIE Ninth International Symposium on Precision Engineering Measurement and Instrumentation, Changsha, China, 8–11 August 2014; Volume 9446, p. 94464S.

21. Lucchese, L.; Mitra, S.K. Using saddle points for sub-pixel feature detection in camera calibration targets. In Proceedings of the Asia-Pacific Conference on Circuits and Systems, Denpasar, Indonesia, 28–31 October 2002; Volume 2, pp. 191–195.
22. Chen, D.; Zhang, G. A New Sub-Pixel Detector for X-Corners in Camera Calibration Targets. In Proceedings of the 13th International Conference in Central Europe on Computer Graphics, Visualization and Computer Vision, Plzen, Czech Republic, 31 January–4 February 2005.
23. Mallon, J.; Whelan, P.F. Which pattern? biasing aspects of planar calibration patterns and detection methods. *Pattern Recognit. Lett.* **2007**, *28*, 921–930. [[CrossRef](#)]
24. Placht, S.; Fürsattel, P.; Mengue, E.A.; Hofmann, H.; Schaller, C.; Balda, M.; Angelopoulou, E. ROCHADE: Robust Checkerboard Advanced Detection for Camera Calibration. In Proceedings of the European Conference on Computer Vision, Zurich, Switzerland, 6–12 September 2014; Fleet, D., Pajdla, T., Schiele, B., Tuytelaars, T., Eds.; Springer: Cham, Switzerland, 2014.
25. Alturki, A.S.; Loomis, J.S. X-Corner Detection for Camera Calibration Using Saddle Points. In Proceedings of the International Conference on Image Analysis and Processing, Boston, MA, USA, 25–26 April 2016.
26. Wang, C.; Sun, T.; Wang, T.; Miao, X.; Wang, R. Multi-PSF fusion in image restoration of range-gated systems. *Opt. Laser Technol.* **2018**, *103*, 219–225. [[CrossRef](#)]
27. Chang, S.H.; Cosman, P.C.; Milstein, L.B. Chernoff-Type Bounds for the Gaussian Error Function. *IEEE Trans. Commun.* **2011**, *59*, 2939–2944. [[CrossRef](#)]
28. Wang, Z.; Wu, W. Recognition and location of the internal corners of planar checkerboard calibration pattern image. *Appl. Math. Comput.* **2007**, *185*, 894–906. [[CrossRef](#)]
29. Hubert, M.; Vandervieren, E. An adjusted boxplot for skewed distribution. *Comput. Stat. Data Anal.* **2008**, *52*, 5186–5201. [[CrossRef](#)]
30. Yang, T.; Zhao, Q.; Wang, X.; Huang, D. Accurate calibration approach for non-overlapping multi-camera system. *Opt. Laser Technol.* **2018**. [[CrossRef](#)]



© 2018 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).

Article

# Measurement of High Numerical Aperture Cylindrical Surface with Iterative Stitching Algorithm

Dingfu Chen <sup>1</sup>, Junzheng Peng <sup>2</sup>, Sergiy Valyukh <sup>3</sup>, Anand Asundi <sup>4</sup> and Yingjie Yu <sup>1,\*</sup>

<sup>1</sup> Lab of Applied Optics and Metrology, Department of Precision Mechanical Engineering, Shanghai University, Shanghai 200072, China; eeyorechen@shu.edu.cn

<sup>2</sup> Guangdong Provincial Key Laboratory of Optical Fiber Sensing and Communications, Department of Optoelectronic Engineering, Jinan University, Guangzhou 510632, China; junzpeng@jnu.edu.cn

<sup>3</sup> Department of Physics, Chemistry and Biology (IFM), Linköping University, SE-581 83 Linköping, Sweden; sergiy.valyukh@liu.se

<sup>4</sup> Centre for Optical and Laser Engineering, School of Mechanical and Aerospace Engineering, Nanyang Technological University, Singapore 639798, Singapore; anand.asundi@pmail.ntu.edu.sg

\* Correspondence: yingjieyu@staff.shu.edu.cn; Tel.: +86-021-6613-0822

Received: 22 September 2018; Accepted: 23 October 2018; Published: 29 October 2018



**Featured Application:** The proposed method is used to test the contour of a cylindrical lens and improve the performance of optical systems.

**Abstract:** There are some limitations in null test measurements in stitching interferometry. In order to meet the null test conditions, the moving distance between the sub-apertures often deviates from the theoretical preset distance, which leads to a position deviation of sub-apertures when measured. To overcome this problem, an algorithm for data processing is proposed in this paper. An optimal estimation of the deviation between sub-apertures is used to update their positions, and then a new overlapped region is obtained and again optimized. This process is repeated until the algorithm converges to an acceptable tolerance, and finally exact stitching is realized. A cylindrical lens was taken as an object for experimental examination of the proposed method. The obtained results demonstrate the validity, reliability, and feasibility of our iterative stitching algorithm.

**Keywords:** null test measurement; stitching interferometry; cylindrical surface; iterative algorithm

## 1. Introduction

The requirements for optical instruments are constantly increasing. Many optical instruments need one-dimensional shaping of the light source realized by cylindrical lenses. Cylindrical lenses have been widely used in high intensity laser systems and spectroscopic and interferometric devices. Their manufacturing processes, which include cutting, grinding, and polishing, are more complex and difficult than for spherical lenses [1,2]. Therefore, quality control and characterization are required.

Tactile measurements are usually recorded with a coordinate measuring machine or a cylindricity measuring instrument. This technology is mature as has been in use for a long time. These methods are highly accurate and provide various error compensation, but only have a small number of sampling points in a certain direction [3]. Besides, it is necessary to be extraordinarily careful to avoid damaging the precision instrument.

Non-contact measurements are realized with optical interferometric techniques known for their high accuracy and dynamic measurement capability coupled with non-destructiveness. At present, plane and spherical interferometers, like that developed by Zygo Co. (Middlefield, CT, USA), are relatively accurate. However, there are some difficulties in using interferometry to measure large optical surfaces or high numerical apertures due to the problems in manufacturing reference

optics for these purposes. The sub-aperture stitching technique [4–10] has to be applied in this case. In J Peng et al. [11,12], stitching interferometry of high numerical aperture cylindrical optics without using a fringe-nulling routine was proposed. An outgoing plane wave from an interferometer is converted into a cylindrical wave by using a Computer Generated Hologram (CGH) [13,14]. Although this work simplified the measurement procedure without meeting the null test condition [8], the described method cannot totally remove high-order errors. J Peng et al. [15] introduced the first order cylindrical coordinate transformation approximation method to measure the complete 360° cylinder surface. The wavefront of the light carrying information on the large aperture was divided into several sub-parts, and the whole surface profile of the metal cylindrical shaft was obtained by the stitching algorithm. However, the sub-aperture stitching technique, based on the null test method, produces errors when the axes of the lens, the wavefront, and the rotation stage do not coincide with each other. A multi-dimensional adjustment device was used to satisfy the null test condition by adjusting the position of the axis of rotation and to obtain the interferogram with a minimal number of fringes in the measured part (sub-aperture) of the lens. The actual errors of the turntable movements can cause miss-arrangement of the sub-apertures. In order to overcome this problem, we propose an iterative algorithm to correct the mismatch and to establish an interferometric stitching measurement system for cylindrical lens examination, where a CGH cylinder null is used as the cylindrical wave converter. We obtain a more accurate stitching result of the surface of the cylindrical lens.

The remainder of the paper is organized as follows. In Section 2, we introduce the principles of the stitching algorithm that have been used in the past. The interferometric system for cylindrical measurement is described in this section. Considering the adjustment errors caused by the null test, we describe a cylinder stitching model based on the iterative algorithm, which can constantly adjust the position of the sub-aperture in the global coordination. In Section 3, we compare the results of the simulations obtained using the proposed iterative algorithm with the traditional algorithms. The results of the proposed method show good performance, with the errors we added to the sub-apertures being almost eliminated. Section 4 is devoted to experimental measurements of a cylindrical lens, and the obtained data are processed with the iterative algorithm. Conclusions are drawn in the final section.

## 2. Principles

### 2.1. Experimental Setup and Error Analysis

In the null test, when the tested and reference surfaces have the same profile, the interferogram is a uniform zero fringe pattern. A non-zero fringe pattern indicates differences between the two surfaces. In order to minimize the misalignment aberrations during the null test, the sub-apertures must be adjusted so that the interferograms contain minimal fringes. In this paper, a CGH was chosen as a null corrector to measure the cylindrical lens.

Figure 1 presents the cylindrical interferometric stitching system. Firstly, the plane wave from the interferometer is converted into a cylindrical wave by the CGH and projected on the cylindrical lens to be tested. Then, after interaction with the cylindrical lens surface, the reflected light returns to the CGH, where the cylindrical wavefront is converted into the plane wavefront and directed back into the interferometer. After, the discrepancy between the measured and reference cylindrical surfaces are obtained by analyzing the recorded interferograms. In order to measure cylindrical lenses with large apertures, it was necessary to have a CGH with a small  $f$ /number (the ratio of the system's focal length to the diameter of the entrance pupil, usually greater than 1). The CCH used in our experiment had a  $f$ /number of 3, which enables observation and detection within 20° and, as a result, prevents obtaining the whole surface profile of a cylindrical lens by scanning once.

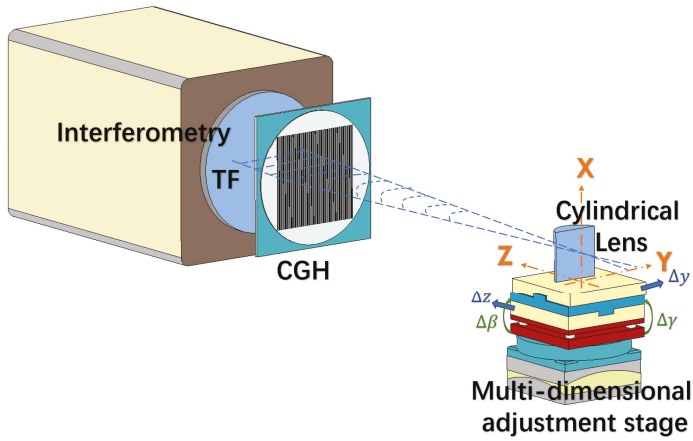


Figure 1. Schematic view of the cylindrical interference system.

The problem mentioned above is usually solved with the sub-aperture stitching technique [4–10]. Such an approach is also helpful for characterization of optical components with high numerical apertures. Multiple scanning of a sample requires preliminary planning of dividing the whole aperture into sub-apertures at different angles of observation. To ensure the whole tested surface is covered during the measurements and to control the stitching process, the adjacent segments from the sub-apertures must have overlapping regions. The reconstructed profiles of these overlapping regions have to be identical to smoothly and continuously supplement the information obtained from the non-overlapping regions.

In the measuring approach, we set the global Cartesian coordinate system so that the X-axis coincides with the focus of the CGH, the optical axis of the tested sample, and the axis of rotation of the rotatable stage (Figure 2a). Z represents the direction of defocus and Y is the horizontal direction. Since the actual positions of the sub-apertures are unknown in the traditional algorithm [5], we assumed that the areas covered by the sub-apertures were uniformly allocated on the tested surface. To obtain uniform allocations of the overlapped and non-overlapped regions, the rotation axis of the multi-dimensional adjustment device platform is adjusted with the optical axis of the cylinder and the focus of the CGH (Figure 2b).

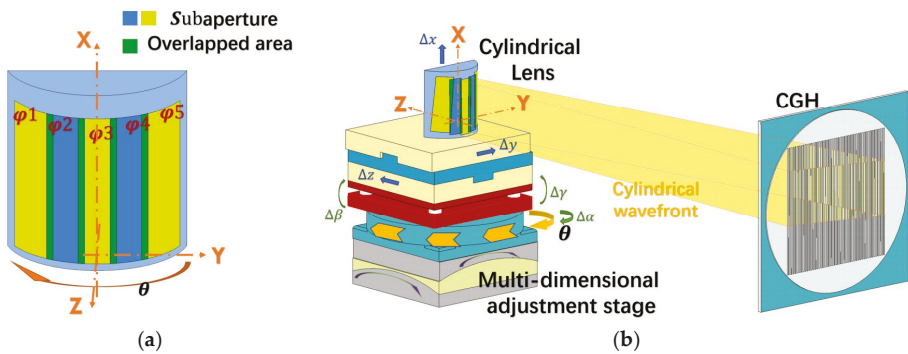


Figure 2. (a) Schematic diagram of sub-aperture layout; (b) Schematic view of a multi-dimensional adjustment stage.



However, in real situations, a deviation may exist in the adjustment. As a result, the mentioned three axes are usually mismatched. Sub-aperture data acquisition in the rotating process may introduce relative position errors, including: (1) turntable radial error caused by inaccuracies in  $\Delta y$  and  $\Delta z$ , (2) the turntable axial runout, caused by  $\Delta x$ , (3) the rotating error of the turntable shaft caused by the change in  $\Delta\beta$  and  $\Delta\gamma$  (Figure 2b), and (4) the turntable angle error caused by the change in  $\Delta\alpha$  (Figure 2b). The real coordinates on the surface of the cylindrical lens,  $\rho_2$ ,  $\Theta_2$ , and  $X_2$ , can be expressed through the real coordinates,  $\rho_1$ ,  $\Theta_1$ , and  $X_1$ , respectively, according to the following formula [16]:

$$\begin{bmatrix} \rho_2 \\ \Theta_2 \\ X_2 \end{bmatrix} = \begin{bmatrix} \rho_1 \\ \Theta_1 \\ X_1 \end{bmatrix} + \begin{bmatrix} \cos \Theta_1 & \sin \Theta_1 & 0 & -X_1 \sin \Theta_1 & X_1 \cos \Theta_1 & 0 \\ -\sin \Theta_1 / \rho_1 & \cos \Theta_1 / \rho_1 & 0 & -X_1 \cos \Theta_1 / \rho_1 & -X_1 \sin \Theta_1 / \rho_1 & 1 \\ 0 & 0 & 1 & \rho_1 \sin \Theta_1 & -\rho_1 \cos \Theta_1 & 0 \end{bmatrix} \begin{bmatrix} \Delta z \\ \Delta y \\ \Delta x \\ \Delta \gamma \\ \Delta \beta \\ \Delta \alpha \end{bmatrix} \quad (1)$$

Since the measured phase  $\varphi$  is associated with the lens surface, the radius can be replaced by the phase in this equation.

At present, the minimum resolution of a commercial precision turntable rotating stage is less than 1'. Such a low deviation cannot be registered by the camera in our experimental setup. Therefore,  $\Delta\alpha$  and  $\Delta x$  can be ignored.

### 2.2. Stitching Algorithm for the Cylindrical Lens

Suppose that the entire cylindrical surface can be totally covered by  $N$  sub-apertures forming the  $N - 1$  overlapping regions as shown in Figure 2a. Due to the non-ideality of the rotating stage and the position errors, each measurement after a rotation requires preliminary tuning for minimization of the fringes in the interferogram. This, in turn, forced us to describe the sample position during measurement with a correction based on the errors  $\Delta z_i$ ,  $\Delta y_i$ ,  $\Delta \gamma_i$ , and  $\Delta \beta_i$ .

Let the whole interferogram of the lens consist of  $M \times K$  pixels. The results of the measurement of each sub-aperture can be represented by the matrix  $M \times K$ , which includes the measured sub-aperture and the zero values. In particular, the data for the first sub-aperture characterized by  $\varphi_1$  (Figure 2) are:

$$\Phi_1 = \begin{bmatrix} \varphi_{1,1} & \varphi_{1,2} & \cdots & \varphi_{1,L} & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 & 0 \\ \varphi_{2,1} & \varphi_{2,2} & \cdots & \varphi_{2,L} & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \varphi_{M,1} & \varphi_{M,2} & \cdots & \varphi_{M,L} & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 & 0 \end{bmatrix}$$

whereas the matrix characterizing the second sub-aperture is:

$$\Phi_2 = \begin{bmatrix} 0 & \cdots & 0 & \varphi_{2,1} & \varphi_{2,2} & \cdots & \varphi_{2,L} & 0 & 0 & \cdots & 0 & 0 & 0 \\ 0 & \cdots & 0 & \varphi_{2,1} & \varphi_{2,2} & \cdots & \varphi_{2,L} & 0 & 0 & \cdots & 0 & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & \cdots & 0 & \varphi_{2,M,1} & \varphi_{2,M,2} & \cdots & \varphi_{2,M,L} & 0 & 0 & \cdots & 0 & 0 & 0 \end{bmatrix}$$

The rows in the matrixes correspond to the data obtained along the  $X$  axis and the columns correspond to the data along the  $\Theta$  axis. Due to the overlapping region between the two measurements, the exact positions of the measured values in  $\Phi_2$  can be determined only after comparing the maps of the first and second measurements and finding the exact overlapping region between them.

Since each element of the matrixes can be considered a function of  $X$  and  $\Theta$ , the relative position error of adjacent sub-apertures can be eliminated using the correlation of the overlapping regions.



as the right array division by dividing each element of the sum of the corrected sub-apertures by the corresponding element of the sum of mask:

$$\hat{\varphi}(X, \Theta) = \sum_{i=1}^N (\Phi_i + \Delta z_i \cos \Theta + \Delta y_i \sin \Theta - \Delta \gamma_i X \sin \Theta + \Delta \beta_i X \cos \Theta) / \sum_{i=1}^N \text{mask}_i \quad (9)$$

where  $\text{mask}_i$  is a matrix with  $M \times K$  elements that equals 1 or 0. Positions of the nonzero elements coincide with the positions of the measured values  $\varphi_{iM,L}$  in  $\Phi_i$ . For example, the  $\text{mask}_1$  is:

$$\text{mask}_1 = \begin{bmatrix} 1 & 1 & \dots & 1 & 0 & \dots & 0 & 0 & 0 & \dots & 0 & 0 & 0 \\ 1 & 1 & \dots & 1 & 0 & \dots & 0 & 0 & 0 & \dots & 0 & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 1 & 1 & \dots & 1 & 0 & \dots & 0 & 0 & 0 & \dots & 0 & 0 & 0 \end{bmatrix}.$$

### 2.3. Iterative Algorithm for the Cylindrical Sub-Aperture

The surface profile of the high high-numerical-aperture (NA) cylindrical lens obtained by stitching the collected phase data with the approach described in Section 2.2 may not immediately have a seamless result. One of the reasons for this is that the restriction by the first order of the Taylor expansion in Equation (1). Besides, the cylindrical lens is not a closed surface, and unlike the cylindrical shaft, it lacks some constraints in the circumference direction. To produce better stitching, an iterative algorithm that enables us to make constant adjustments to the position of each sub-aperture is preferred. This can be realized by applying the least square method for finding the position errors  $R_i$  according to Equation (8). Then, the position errors  $R_i$  are substituted into the Rigid Body Transformation to correct the mismatch of the sub-apertures. After correcting the sub-apertures, the new corrected position errors  $R_i$  are calculated. The loop with the iterations enables us to get more reliable results. The specific steps of the algorithm are as follows:

Step 1. Assume that the measuring aperture is arranged in space according to the preset coordinates. Then, the initial values of the cumulative errors have to be set as  $Err_i^0 = (\Delta z_i^0, \Delta y_i^0, \Delta \gamma_i^0, \Delta \beta_i^0) = (0, 0, 0, 0)$ . The initial phase values  $\varphi_i^0 = \varphi_i$  are the phase data measured by the interferometer. The superscript denotes a current iteration number, and the subscript  $i$  denotes the number of sub-apertures.

Step 2. Substitute  $\theta$ ,  $X$ , and  $\varphi_i^{(t)}$  into Equation (8) to obtain the least squares solution  $(\Delta z_i^{(t+1)}, \Delta y_i^{(t+1)}, \Delta \gamma_i^{(t+1)}, \Delta \beta_i^{(t+1)}) = R_i^{(t+1)}$ , which is the error vector in the cylindrical coordination system. The superscript  $t$  denotes the iteration number,  $t = 0, 1, 2, \dots$

Step 3. Find a new corrected  $Err_i^{(t+1)}$  according to the equation:

$$Err_i^{(t+1)} = Err_i^{(t)} + R_i^{(t+1)} \quad (10)$$

Step 4. Change  $\varphi_i^{(t)}$  into Cartesian coordinate sand substitute it into the transformation matrix, which includes the coefficient  $R_i^t$ . After that, change the coordinate back into the cylindrical coordination:

$$y_i^{(t)} = \varphi_i^{(t)} \times \cos(\theta_i^{(t)}); z_i^{(t)} = \varphi_i^{(t)} \times \sin(\theta_i^{(t)}); x_i^{(t)} = X_i^{(t+1)} \quad (11)$$

$$\begin{bmatrix} y_i^{(t+1)} & z_i^{(t+1)} & x_i^{(t+1)} & I \end{bmatrix} = \begin{bmatrix} y_i^{(t)} & z_i^{(t)} & x_i^{(t)} & I \end{bmatrix} * T_1 T_2 T_3 T_4 \quad (12)$$

where  $T_1, T_2, T_3$ , and  $T_4$  are the transformation matrixes taking  $R_i^{(t+1)}$  into account, and  $I$  is the unity matrix.

$$\varphi_i^{(t+1)} = \sqrt{y_i^{(t+1)2} + z_i^{(t+1)2}}; X_i^{(t+1)} = x_i^{(t+1)}; \theta_i^{(t+1)} = \arctan\left(\frac{z^{(t+1)}}{y^{(t+1)}}\right) \quad (13)$$

Step 5. Repeat the steps described above until  $R_i^{(t+1)}$  is less than the threshold or reaches the iteration number (usually  $10^{-3} \mu\text{m}$  or 3 iterations according to practical experience). Then we can obtain a more reliable result for  $\varphi_i^{(t+1)}$ .

### 3. Simulation

In order to estimate the accuracy of the proposed method, we carried out a stitching interferometry simulation for a high-NA cylindrical surface using MATLAB (MathWorks. Inc, Natick, MA, USA). First, a cylindrical surface 50 mm high and a radius of 26.4 mm was built (Figure 3a). The angular aperture equaled  $102^\circ$ . Then, the simulated surface was divided into eight equal fragments, each of which was observed under a sub-aperture of about  $18^\circ$ . The rotation angle of each fragment was  $12^\circ$ , and the overlapping region was  $6^\circ$ . Each fragment was associated with its own position error as shown in Table 1. The fourth fragment (sub-aperture) was chosen as the benchmark; therefore, its error was zero. Besides, because the errors  $\Delta x$  and  $\Delta \alpha$  were negligible, they were set zero. The errors given in the table correspond to the actual situation in the measurement.

Table 1. Errors of the sub-apertures.

Error	1	2	3	4	5	6	7	8
$\Delta z (\mu\text{m})$	1.2	0.6	2.3	0	-1.9	-2.1	1.1	1
$\Delta y (\mu\text{m})$	0.8	1.4	1.5	0	2.6	1.4	-0.5	-0.3
$\Delta \gamma (")$	-3.1	-3.7	-6.4	0	-5.6	-6.6	-2.9	-1.7
$\Delta \beta (")$	-3.5	-2.3	3.7	0	-7.0	-3.9	-6.4	-5.6
$\Delta x (\mu\text{m})$	0	0	0	0	0	0	0	0
$\Delta \alpha (")$	0	0	0	0	0	0	0	0

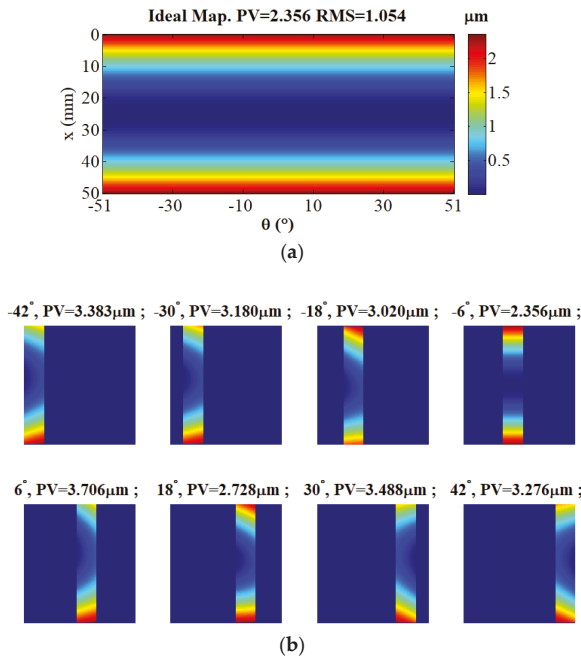
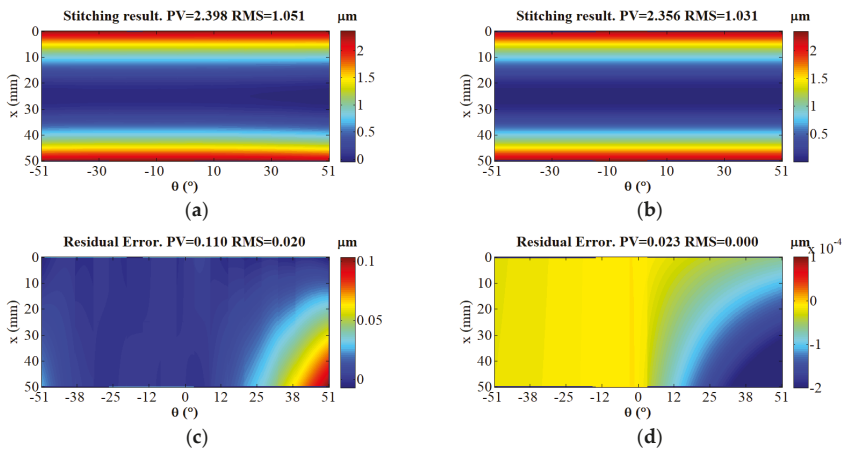


Figure 3. (a) The simulation of the proposed method. An ideal cylindrical surface was built and (b) eight sub-apertures were segmented to cover the whole surface and each sub-aperture was attached with some position errors.

After obtaining the sub-aperture with different position errors (Figure 3b), it was necessary to stitch them using iterations if necessary (Figure 4a,b). Subtracting the obtained data (Figure 4a,b) from the simulated data (Figure 3a) enabled us to see that the peaks-and-valleys (PV) value and the root mean square (RMS) value of the stitching result without using the iterative method are larger than those obtained using the iterative method. The PV value of the residual error (Figure 3d) was  $0.023 \mu\text{m}$  and the RMS value was  $0.000 \mu\text{m}$ , which means that the sub-apertures were correctly stitched. Because the curve of the lens was not closed, some constraints in the circumference direction were lacking. So, the errors produced by the traditional method cannot be eliminated totally, especially when there is an error near the boundary. Through iterations, this situation will improve. The final value of the residual error is affected by the size of the pixel.



**Figure 4.** (a) The stitching result without using the iteration algorithm. (b) The result obtained with the proposed iterative method, (c) the residual error between the origin map and traditional method, and (d) the residual error between the origin map and iterative method.

#### 4. Experimental Demonstration

Experimental measurements were carried out to verify the practicality of the proposed algorithm. The experimental arrangement consisted of a Fizeau interferometer (Zygo GPI/XP 4, Zygo, Middlefield, CT, USA), a CGH cylinder null (H80F3C, Diffraction international, Minnetonka, MN, USA) and a multi-dimensional adjustment stage (Figure 5). A CGH cylinder null, 80 mm high  $\times$  80 mm wide, was used as the cylindrical wavefront converter of the interference system. The maximum measurable aperture angle was nearly  $18.9^\circ$  (the treatment of the image borders led to a measured angle less than 20 degrees). The cylindrical lens was 53 mm high  $\times$  50.8 mm wide and curvature radius was 26.4 mm. For obtaining the whole surface of the cylindrical lens, we collected 13 sub-apertures and set the rotation angle to  $12^\circ$ . We did not strictly align the lens, so after each rotation, the fringe of interferogram was too dense to be analyzed. So, we manually tilted and leveled the sub-aperture to obtain an interferogram that was near to non-fringe. Then, the remaining position errors were calculated through the iterations.

Figure 6 shows the phase maps and interferograms of the acquired sub-aperture. In order to meet the condition of the null test, we ensured that the fringe was small enough to make the ideal cylindrical wavefront and cylindrical lens coincide. The remaining fringes were caused by the contour deviation and remaining position errors.

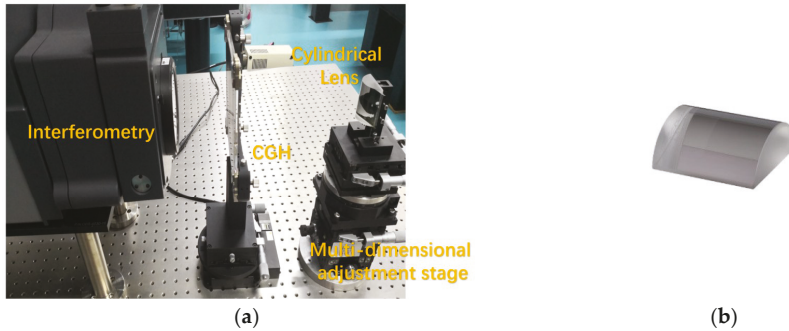


Figure 5. (a) Photograph of the experimental system and (b) the experimental sample.

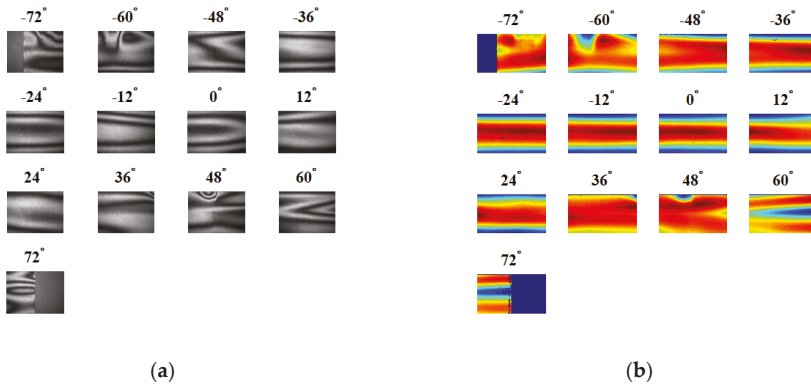


Figure 6. (a) Measured interferograms of different sub-apertures, fringe maps of sub-apertures and (b) the phase maps of sub-apertures.

The stitching result map without the iterations is depicted in Figure 7a, where the stitching marks can be seen. Accordingly, the result map after three iterations (Figure 7b) was seamless, the PV value of the stitching result was 1.003  $\mu\text{m}$ , and the RMS value was 0.174  $\mu\text{m}$ . To evaluate the effectiveness of the proposed stitching method, we observed the variation in the mismatch map within the overlapping regions, as shown in Figure 7c,d. The variation was determined by calculating the radial deviations of the corresponding points in the overlapping region, and can be thought of as the residual noise following stitching process. The PV value of the residual error without iterations was 0.619  $\mu\text{m}$  and the RMS value was 0.029  $\mu\text{m}$ . The RMS value obtained with the proposed method was 0.005  $\mu\text{m}$ . This confirms that the proposed stitching model can eliminate the discrepancies among the overlapping regions. The invalid points in Figure 7a were caused by the spurious fringe in the central sub-aperture. Due to the interpolation process in the coordinate transformation, there was no invalid point in the middle part of the stitching result shown in Figure 7b.

In order to evaluate the reliability of the proposed algorithm, we carried out 10 iterations of the experimental data and compared the coefficients of the sub-aperture adjustment for each iteration of the two sub-apertures. The experimental data (Figure 8) showed that after three iterations, the *y*-direction and *z*-direction errors were stabilized within a very small range. From the graph, for the whole experimental system and its operation, the rotation errors around the *Y* and *Z* axes were relatively small.

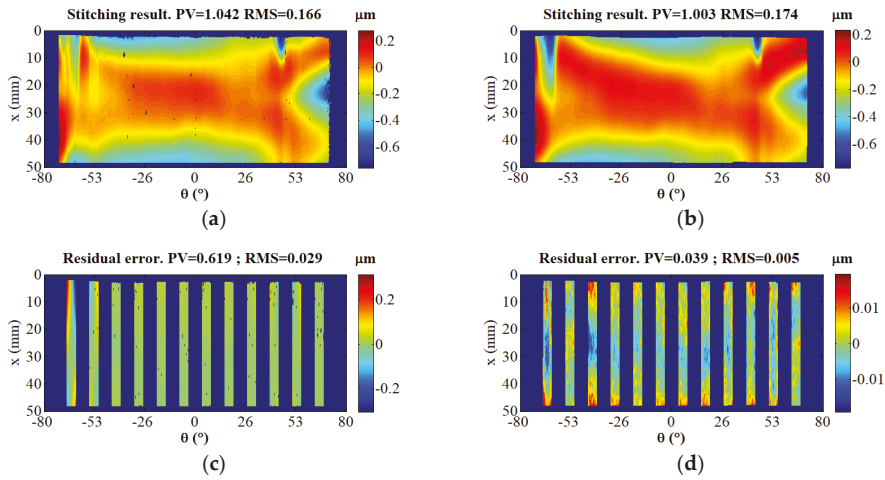


Figure 7. (a,b) Stitching result without iteration and the mismatch map; (c,d) comparison with the proposed method with three iterations.

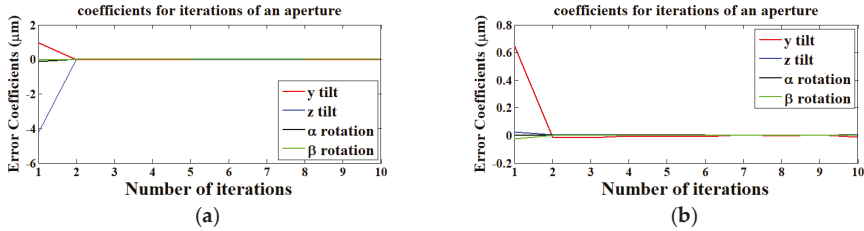


Figure 8. (a,b) The change of the relative motion errors of two sub-apertures in 10 iterations.

Table 2 summarizes the position error  $Err_i^k$  of three iterations and Figure 9 shows an interferogram of  $Err_i^k$ . As the ninth sub-aperture was chosen as the benchmark, the  $24^\circ$  interferogram is not shown. Compared with the number of fringes in the pattern in Figure 6a, the value of the position error is near the magnitude of the contour of the lens. The position errors of the first few sub-apertures are quite large, which is consistent with the result in Figure 7c. Because the smallest reading value of the tilt stage was  $1^\circ$  and the resolution was  $18''$ , the precision angles of manual adjustment could not be obtained. However, through the simulation according to the fringe number before manual adjustment, the range of the manually tilt angle was about  $\pm 10'$ , and the horizontally motion was about  $\pm 200 \mu\text{m}$ .

Table 2. The adjustment errors of each sub-aperture for three iterations.

$Err_i^{(3)}$	1	2	3	4	5	6	7	8
$\Delta z$ ( $\mu\text{m}$ )	1.2137	-0.5741	0.8844	0.7786	0.7942	0.7232	0.5038	0.4332
$\Delta y$ ( $\mu\text{m}$ )	-4.4061	2.7769	-0.1865	0.0128	0.0205	0.0574	0.1129	0.0850
$\Delta\gamma$ ( $''$ )	-14	-16	-0.7	-0.5	-0.27	0	0.2	0.3
$\Delta\beta$ ( $''$ )	-3.51	-2.27	3.71	0	-7.01	-3.92	-6.39	-5.57
$Err_i^{(3)}$	9	10	11	12	13			
$\Delta z$ ( $\mu\text{m}$ )	0	0.4166	-0.6006	0.5141	0.5827			
$\Delta y$ ( $\mu\text{m}$ )	0	0.1540	-0.5952	0.3576	0.3520			
$\Delta\gamma$ ( $''$ )	0	-1.6	-4.4	-0.3	-0.3			
$\Delta\beta$ ( $''$ )	0	-4.1	4.6	-0.8	-0.9			

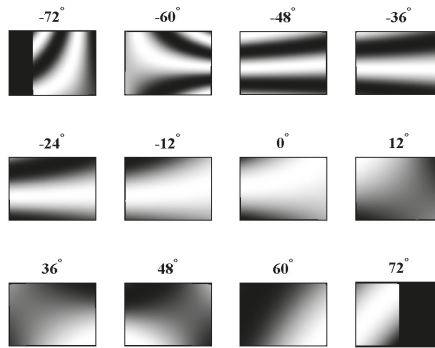


Figure 9. The interferogram of translation error  $Err_i^k$ .

Finally, to verify the robustness of the algorithm, the lens was measured eight times, with or without the iteration to compare the residual errors (Figure 10). The PV value and the RMS value of the residual error solved by iteration method were, as a rule, kept within a stable range, which further verifies the high reliability of the measurement.

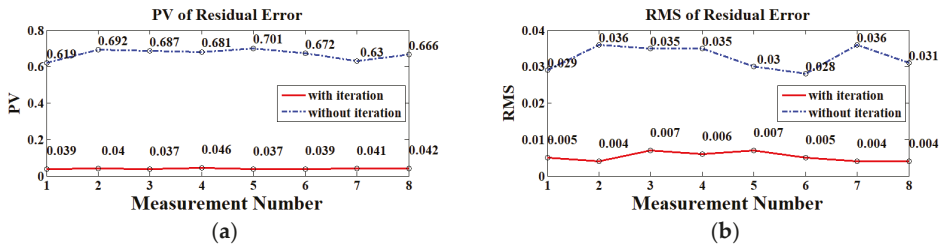


Figure 10. (a) Data comparison of eight groups of the peaks-and-valleys (PV) value of residual error; (b) the root mean square (RMS) value of the residual error.

From the experiment, the stitching result obtained by our algorithm has advantages in comparison with the traditional technique. Among the next practical steps for improving the proposed method include slight modifications and adaptation for characterization of complex structured materials and surfaces [17,18] as well as for inline quality control in a product line [19], which is often highly required in mass production of optical components.

### 5. Conclusions

An iterative stitching algorithm was proposed in this study for data processing of interferometric measurements of cylindrical surfaces. The technique was demonstrated with an example by determining the surface topography of a cylindrical lens with high NA. By constantly calculating the deviation of the measured overlapping region and then adjusting the position of each sub-aperture, the obtained experimental results had fewer residual errors than when using the traditional method. Possible registered miss-arrangements during measurement were corrected in the data processing. The advantages of this new method enable achieving higher reasonability and reliability than the traditional techniques. The calculation results quickly converge, but the improvement in accuracy is not obvious with increasing numbers of iterations. In addition, due to the high accuracy of the experimental device, a higher number of iterations was not required.

**Author Contributions:** Writing—Original Draft Preparation and Software: D.C.; Formal Analysis, Investigation and Methodology: J.P.; Project Administration and Supervision: Y.Y. and A.A.; Supervision: A.A.; Editing & Review and editing: Y.Y. and S.V.



**Funding:** The authors are grateful for support from the following research grants: (1) the National Natural Science Foundation of China (NSFC) (No. 51775326); (2) National Science and Technology Major Project (No.2016YFF0101905); (3) Natural Science Foundation of China (NSFC) (No. 61605126).

**Acknowledgments:** Dingfu Chen thanks the 'Strategic Partners Scholarship Fund' program of Shanghai University for the support.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Fang, F.; Zhang, X.; Hu, X. Cylindrical coordinate machining of optical freeform surfaces. *Opt. Express* **2008**, *16*, 7323–7329. [[CrossRef](#)] [[PubMed](#)]
2. Zhang, X.; Zeng, Z.; Liu, X.; Fang, F. Compensation strategy for machining optical freeform surfaces by the combined on- and off-machine measurement. *Opt. Express* **2015**, *23*, 24800–24810. [[CrossRef](#)] [[PubMed](#)]
3. Ramaswami, H.; Kanagaraj, S.; Anand, S. An inspection advisor for form error in cylindrical features. *Int. J. Adv. Manuf. Technol.* **2009**, *40*, 128–143. [[CrossRef](#)]
4. Chen, M.; Cheng, W.; Wang, C.W.W. Multi aperture overlap scanning technique for large-aperture test. *Proc. SPIE* **1991**, *1553*, 626–635. [[CrossRef](#)]
5. Otsubo, M.; Okada, K.; Tsujiuchi, J. Measurement of large plane surface shapes by connecting small aperture interferograms. *Opt. Eng.* **1994**, *33*, 608–613. [[CrossRef](#)]
6. Peng, J.; Chen, D.; Guo, H.; Zhong, J.; Yu, Y. Variable optical null based on yawing CGH for measuring steep acylindrical surface. *Opt. Express* **2018**, *16*, 20306–20318. [[CrossRef](#)] [[PubMed](#)]
7. Wen, Y.; Cheng, H.; Tam, H.Y.; Zhou, D. Modified stitching algorithm for annular subaperture stitching interferometry for aspheric surface. *Appl. Opt.* **2013**, *52*, 5686–5694. [[CrossRef](#)] [[PubMed](#)]
8. Tricard, M.; Kulawiec, A.; Bauer, M. Subaperture stitching interferometry of high-departure aspheres by incorporating a variable optical null. *CIRP Ann. Manuf. Technol.* **2010**, *59*, 547–550. [[CrossRef](#)]
9. Hagino, T.; Yokoyama, Y.; Kuriyama, Y.; Haitjema, H. Sphericity measurement using stitching interferometry. *Key Eng. Mater.* **2012**, *523*, 883–888. [[CrossRef](#)]
10. Fan, Y.; Struik, K.; Mulders, P.; Velzel, C. Stitching interferometry for the measurement of aspheric surfaces. *CIRP Ann. Manuf. Technol.* **1997**, *46*, 459–462. [[CrossRef](#)]
11. Peng, J.; Wang, Q.; Peng, X.; Yu, Y. Stitching interferometry of high numerical aperture cylindrical optics without using a fringe-nulling routine. *J. Opt. Soc. Am. A* **2015**, *32*, 1964–1972. [[CrossRef](#)] [[PubMed](#)]
12. Peng, J.; Xu, H.; Yu, Y. Stitching interferometry for cylindrical optics with large angular aperture. *Meas. Sci. Technol.* **2015**, *26*, 025204. [[CrossRef](#)]
13. Mac Govern, A.J.; Wyant, J.C. Computer Generated Holograms for Testing Optical Elements. *Appl. Opt.* **1971**, *10*, 619–624. [[CrossRef](#)] [[PubMed](#)]
14. Asfour, J.M.; Poleshchuk, A.G. Asphere testing with a Fizeau interferometer based on a combined computer-generated hologram. *J. Opt. Soc. Am. A* **2006**, *23*, 172–178. [[CrossRef](#)]
15. Peng, J.; Yu, Y.; Chen, D. Stitching interferometry of full cylinder by use of the first-order approximation of cylindrical coordinate transformation. *Opt. Express* **2017**, *25*. [[CrossRef](#)] [[PubMed](#)]
16. Guo, H.; Chen, M. Multiview connection technique for 360-deg 3D measurement. *Opt. Eng.* **2003**, *42*, 900–905. [[CrossRef](#)]
17. Valyukh, S.; Osterman, J.; Valyukh, I.; Skarp, K. Characterization of flexible reflective liquid crystal cells. *J. SID* **2005**, *13*, 501–506. [[CrossRef](#)]
18. Valyukh, S.; Valyukh, I.; Skarp, K. Spectrophotometric determination of reflective liquid crystal cell parameters. *J. Appl. Phys.* **2006**, *99*, 053102. [[CrossRef](#)]
19. Valyukh, S.; Sorokin, S.; Chigrinov, V. Inline quality control of liquid crystal cells. *IEEE/OSA J. Disp. Technol.* **2015**, *11*, 1042–1047. [[CrossRef](#)]



© 2018 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).

Article

# Error Correction for FSI-Based System without Cooperative Target Using an Adaptive Filtering Method and a Phase-Matching Mosaic Algorithm

Xing-Ting Xiong , Xing-Hua Qu and Fu-Min Zhang \*

State Key Laboratory of Precision Measurement Technology and Instruments, Tianjin University, Tianjin 300072, China; tjuxiong@163.com (X.-T.X.); quxinghua@tju.edu.cn (X.-H.Q.)

\* Correspondence: zhangfumin@tju.edu.cn

Received: 1 September 2018; Accepted: 14 October 2018; Published: 17 October 2018

**Featured Application:** Absolute distance measurement and surface profiling.

**Abstract:** In our frequency scanning interferometry-based (FSI-based) absolute distance measurement system, a frequency sampling method is used to eliminate the influence of laser tuning nonlinearity. However, because the external cavity laser (ECL) has been used for five years, factors such as the mode hopping of the ECL and the low signal-to-noise ratio (SNR) in a non-cooperative target measurement bring new problems, including erroneous sampling points, phase jumps, and interfering signals. This article analyzes the impacts of the erroneous sampling points and interfering signals on the accuracy of measurement, and then proposes an adaptive filtering method to eliminate the influence. In addition, a phase-matching mosaic algorithm is used to eliminate the phase jump, and a segmentation mosaic algorithm is used to improve the data processing speed. The result of the simulation proves the efficiency of our method. In experiments, the measured target was located at eight different positions on a precise guide rail, and the incident angle was 12 degrees. The maximum deviation of the measured results between the FSI-based system and the He-Ne interferometer was 9.6  $\mu\text{m}$ , and the maximum mean square error of our method was 2.4  $\mu\text{m}$ , which approached the Cramer-Rao lower bound (CRLB) of 0.8  $\mu\text{m}$ .

**Keywords:** frequency scanning interferometry; adaptive filtering method; mosaic algorithm

## 1. Introduction

Absolute distance measurement systems, which measure several tens of meters with microns uncertainties, are of significant interest in the field of metrology [1–4]. With the improvement of the external cavity laser (ECL), systems using frequency scanning interferometry FSI and the ECL play an important role in these fields: absolute distance measurement [5–8], imperfection detection [9], optical coherence tomography (OCT) [10], and three-dimensional surface profiling [11]. In our paper, attention is focused on absolute distance measurement.

An FSI-based absolute distance measurement system sends out a frequency-modulated signal to an object and receives the backscattering signal. The output is a low-frequency cosine wave in time domain that is affected by the tuning nonlinearity of the ECL [12,13]. Hence, in order to obtain precision in the order of microns, it is important for the ECL to provide linear frequency scanning over a broad tuning bandwidth [14]. However, because of the hysteresis and creep of the piezoelectric actuator (PZT), the ECL often shows tuning nonlinearity in practical situations. To eliminate the influence of the nonlinear tuning, many methods had been proposed [15–19].

In our paper, we sample the measurement interference signal at equidistant optical frequency points (EIOFs, namely sampling points) of the auxiliary interferometer with a long single mode

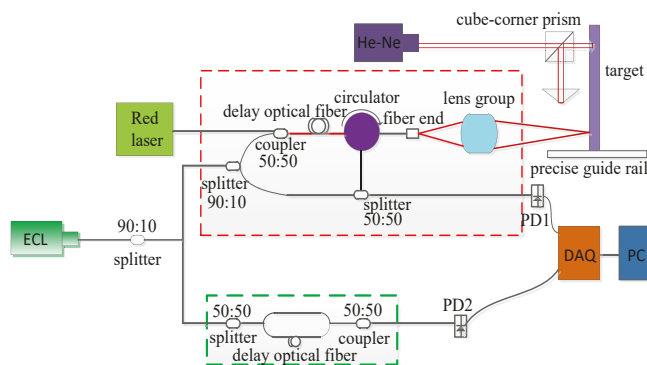
polarization-maintaining fiber to eliminate the influence of tuning nonlinearity. This method is a common and effective method, but it creates a dispersion mismatch problem that results in target peak broadening. Fortunately, many dispersion compensation methods had been proposed and proved efficiently [20–22]. In our experiments without cooperate targets, we use an ECL that has been used for five years, so the SNR is low, and mode-hopping signals often occur. These bring new problems: erroneous sampling points and interfering signals. Our paper analyzes the influences of erroneous sampling points and interfering signals, and then an adaptive filtering method is proposed to eliminate the influences completely. In addition, a phase-matching mosaic algorithm and a segmentation mosaic algorithm are adopted to eliminate the phase jump and improve the data processing speed, respectively.

The article is organized as follows: Section 2 reviews the principle of the FSI-based system and discusses the measurement errors caused by erroneous sampling points and interfering signals. Then, an adaptive filtering method, a phase-matching mosaic algorithm, and a segmentation mosaic algorithm are proposed to eliminate the erroneous sampling points and phase jump and reduce the data processing time. In Sections 3 and 4, the results of the simulation and experiment prove the efficiency of the methods that were proposed in Section 2. Finally, a brief conclusion is given in Section 5.

## 2. Theory

### 2.1. The Principle of FSI Using Frequency Sampling Method

As is shown in Figure 1, the ECL has been used for five years and is connected with two interferometers. One is the auxiliary interferometer (as shown in the green box), whose output is used as the sampling signal to correct the tuning nonlinearity. The other is the measurement interferometer (as shown in the red box), whose output contains the information of measured distance. First, 10% power of the ECL goes into the auxiliary interferometer and the output is received by photo-detector 2 (PD2). Meanwhile, the remaining 90% of the power of the ECL goes into the measurement interferometer. Then, making use of a 90:10 optical splitter, 10% is used as the reference signal and 90% is used as the measurement light. Photo-detector 1 (PD1) obtains the output of the measurement interferometer. Finally, PD1 and PD2 convert the optical signals to the electric signals, and send the signals to the data acquisition card (DAQ). The He-Ne interferometer is used to verify the accuracy of our system.



**Figure 1.** Schematic diagram of our frequency scanning interferometry (FSI)-based system and the experimental setup.

The output of the measurement interferometer can be expressed as:

$$I_m(f) = A_m \cos(2\pi f \tau_m) = A_m \cos(2\pi(f_0 + \Delta f)\tau_m) \tag{1}$$

where  $A_m$  is the amplitude of the measurement interference signal,  $f_0$  is the initial frequency of the ECL,  $\Delta f$  is the variation of the ECL instantaneous frequency, and  $\tau_m$  is the time delay of the measurement interferometer.

Similarly, the ideal output of the auxiliary interferometer can be expressed as:

$$I_a(f) = A_a \cos(2\pi f \tau_a) = A_a \cos(2\pi(f_0 + \Delta f)\tau_a) \tag{2}$$

where  $A_a$  is the amplitude of the auxiliary interference signal, and  $\tau_a$  is the time delay of the auxiliary interferometer.

To eliminate the influence of the ECL tuning nonlinearity, taking the output of the auxiliary interferometer as the sampling signal yields:

$$2\pi\Delta f(k)\tau_a = \pi k, k = 1, 2, \dots, N \tag{3}$$

where  $\Delta f(k)$  is the variation of instantaneous frequency,  $k$  is the sampling point index, and  $N$  is the number of the ideal sampling point. So, the output of the measurement interferometer can be rewritten as:

$$I_m(k) = A_m \cos\left(2\pi f_0 \tau_m + 2\pi \frac{\tau_m}{2\tau_a} k\right) \tag{4}$$

Equation (4) by fast Fourier transform (FFT) yields:

$$\frac{\tau_m}{2\tau_a} = \frac{P}{N} \tag{5}$$

where  $P$  is the abscissa index of the peak point in the frequency spectrum of the measurement interferometer. When the dispersion is considered, according to John [23], it is easy to know:

$$\tau_a = \frac{n_{air} L_m}{c} (1 - 2\pi\alpha\beta v_g t) \tag{6}$$

where  $n_{air}$  is the refractive index of air,  $L_m$  is the measured distance,  $c$  is the speed of light in the vacuum,  $n_f$  is the refractive index of the single mode fiber,  $L_a$  is the length of the single mode fiber in the auxiliary interferometer,  $\beta$  is the group velocity dispersion of the single-mode optical fiber,  $\alpha$  is the tuning rate of the ECL,  $v_g$  is the group velocity, and  $t$  is the time.

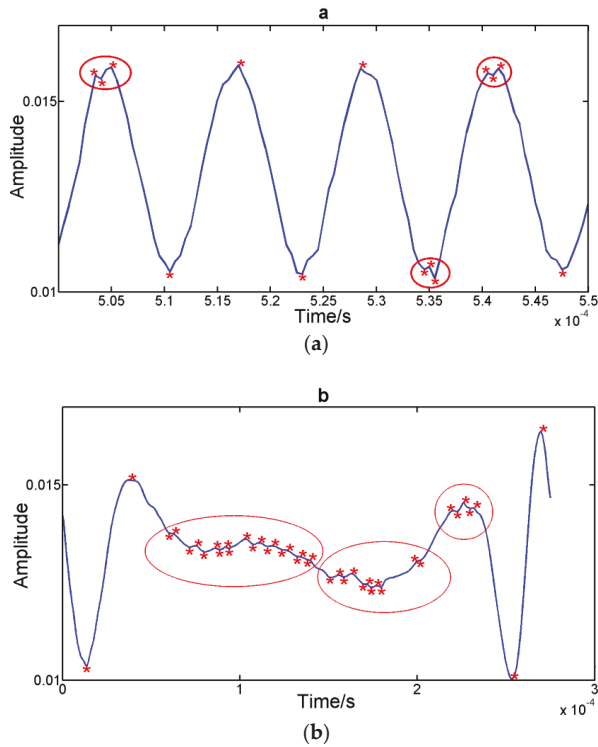
So, the measured distance can be described as:

$$L_m = (1 - 2\pi\alpha\beta v_g t) \frac{n_f L_a P}{n_{air} N} \tag{7}$$

The second item in the right side of Equation (7) is caused by dispersion mismatch. Consequently, according to distance offsets, the absolute distance can be corrected by Equation (7).

### 2.2. The Influence of the Erroneous Sampling Points in the Auxiliary Interferometer

In the paper, we take the extreme points of the auxiliary interference signal as the sampling points. Figure 2 shows the extreme points of the auxiliary interferometer output in the experiments. Obviously, because of the influence of noises and mode-hopping signals, some of the extreme points are incorrect (as shown in the red circle section). Thus, the parameters  $N$  and  $P$  in Equation (5) are different from their ideal values.



**Figure 2.** (a) Extreme points in the experiment when there are noises in the auxiliary interferometer. (b) Extreme points in the experiment when there is the mode-hopping signal in the auxiliary interferometer.

When there are noises and mode-hopping signals in the auxiliary interferometer, Equation (3) should be rewritten as:

$$2\pi\Delta f(k)\tau_a = (\pi + \Delta\Omega(k_1))k, k = 1, 2, \dots, N, k_1 = 1, 2, \dots, N_1 \quad (8)$$

where  $\Delta\Omega(k_1)$  is the phase deviation caused by all of the sampling points of the auxiliary interferometer output in the actual situation, and only part of the  $\Delta\Omega(k_1)$  is caused by the erroneous sampling points is not equal to zero, while  $N$  is the number of the ideal sampling points, and  $N_1$  is the total number of the sampling points in the actual situation. Thus, Equation (4) turns into:

$$I_m(k) = A_m \cos(2\pi f_0 \tau_m + \pi k \tau_m / \tau_a + \Delta\Omega(k_1) k \tau_m / \tau_a) \quad (9)$$

Equation (9) by FFT yields:

$$\left(1 + \sum_{k_2=1}^{N_2} \frac{\Delta\Omega(k_2)}{\pi N}\right) \frac{\tau_m}{2\tau_a} = \frac{P_1}{N_1} \quad (10)$$

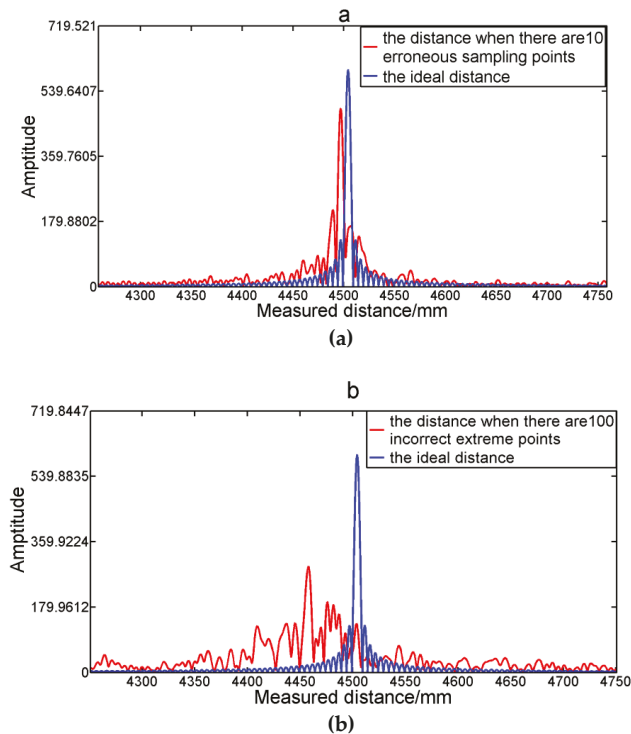
where  $N_2$  is the number of all of the erroneous sampling points in the auxiliary interferometer output, and  $P_1$  is the abscissa value of the peak point in the frequency spectrum of the measurement signal in Equation (9).

If  $N_2$  is much smaller than  $N$ , the second item on the left side of Equation (10) can be neglected. So, the measurement error caused by the erroneous sampling points of the auxiliary interferometer output can be shown as:

$$\Delta L_{m1} = L_{m1} - L_m = \frac{n_f}{n_{air}} L_a (P_1/N_1 - P/N) \tag{11}$$

where  $\Delta L_{m1}$  is the measurement error caused by the erroneous sampling points of the auxiliary interferometer,  $L_{m1}$  is the measured distance when there are erroneous sampling points, and  $L_m$  is the ideal measured distance when there are no erroneous sampling points. Figure 3a shows the comparison of the simulated result when there are 10 erroneous sampling points (the red line) and the ideal distance (the blue line). In our experiments,  $n_f$  is 1.4682,  $n_{air}$  is 1.0003, and  $L_a$  is 20.05832 m. When the ideal value of  $P$  is 9000,  $N$  is 60,000, and  $N_2$  is 10, the measured distance error is 758.9  $\mu\text{m}$ , which is agreement with the theoretical measurement error of 760  $\mu\text{m}$ , which is calculated by Equation (11). Obviously, the accuracy of the measurement distance is greatly affected. Besides, the main lobe energy of the spectrum decreases, and the side lobe's energy increase. Hence, it is essential to eliminate the influence of the erroneous sampling points in order to ensure the accuracy.

As is shown in Figure 3b, if  $N_2$  is 100, the second item on the left-hand side of Equation (11) can't be neglected. In this situation, we cannot obtain the accurate measured distance.



**Figure 3.** (a) The simulation result of 10 erroneous sampling points on accuracy when the ideal value of  $P$  is 9000 and  $N$  is 60,000. (b) The simulation result of 100 erroneous sampling points on accuracy when the ideal value of  $P$  is 9000 and  $N$  is 60,000.

2.3. The CRLB When There Are Interfering Signals and Noise in the Measurement Interferometer

Ignoring the influence of the erroneous sampling points in the auxiliary interferometer, when the system measures the non-cooperative target and there is noise in the measurement interferometer, Equation (4) turns into:

$$I_m(k) = A_m \cos\left(2\pi f_0 \tau_m + \frac{\pi k \tau_m}{\tau_a}\right) + \text{Noise}_m(k) + I_i(k) = A_m \cos(\zeta k + \varphi) + \text{Noise}_m(k) + I_i(k), \quad (12)$$

$$k = 1, 2, \dots, N$$

where  $\text{Noise}_m(k)$  denotes the Gaussian white noise (GWN) with an expectation of 0 and variance  $\sigma^2$ ,  $I_i(k)$  denotes the interfering signals, and  $\zeta$  is  $\frac{\pi \tau_m}{\tau_a}$ ,  $\varphi$  is  $2\pi f_0 \tau_m$ . The interfering signals  $I_i(k)$  are caused by the reflected lights from the circulator and the fiber end in Figure 1.

As shown in Figure 4, when our system measures a non-cooperative target, the measurement signal (the red line) becomes very weak, and the interfering signals (the blue line) are much stronger than the measurement signal. In addition, the frequencies of the interfering signals are not varying with the measured distances and are lower than the frequency of the measured signal, so we can use an appropriate low pass filter to delete them. Thus, the SNR of the measurement interferometer output can be shown as  $A_m^2/\sigma^2$ .

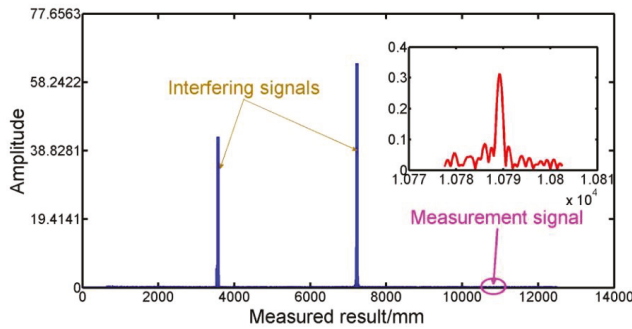


Figure 4. Measured result without cooperative target.

The Cramer–Rao lower bound (CRLB) is a method in the statistics and the variance of any unbiased estimator cannot be lower than the CRLB [23]. Therefore, we use the CRLB to evaluate our system. The likelihood function of  $I_m(k)$  can be expressed as:

$$L(I_m(k)|\zeta) = \frac{1}{(2\pi\sigma^2)^{N/2}} \exp\left\{-\frac{1}{2\sigma^2} \sum_{k=1}^N (I_m(k) - A_m \cos(\zeta k + \varphi))^2\right\} \quad (13)$$

The Fisher information is obtained as:

$$E\left(\frac{\partial^2 \ln L(I_m(k)|\zeta)}{\partial \zeta^2}\right) = \text{SNR} \sum_{k=1}^N (k \sin(\zeta k + \varphi))^2 \quad (14)$$

The CRLB on the variance of the measured distance is:

$$\text{Var}\{L_m\} \geq \frac{6n_f^2 L_a^2}{\pi^2 n_{air}^2 N^3 \text{SNR}} \quad (15)$$

Accordingly, when there is GWN in the measurement interferometer, the CRLB on the standard deviation of measured distance can be shown as:

$$\Delta L_{m2} \geq \frac{n_f L_a}{\pi n_{air} N^2} \sqrt{\frac{6(N - 1)}{SNR}} \tag{16}$$

When the measured target is not a reflector or a corner prism and the incident angle is not zero, the SNR should be low, and the error caused by the GWN in the measurement interferometer can't be ignored.

2.4. *The Adaptive Filtering Method, the Phase-Matching Mosaic Algorithm, and the Segmentation Mosaic Algorithm*

According to the analysis above, completely eliminating the erroneous sampling points of the auxiliary interferometer and the interfering signals in the measurement interferometer are essential in order to ensure the accuracy of our system. Thus, an adaptive filtering method and a phase-matching mosaic algorithm are proposed to ensure the accuracy, and a segmentation mosaic algorithm is used to improve the data-processing speed. The adaptive filtering method can be described as following.

- (1) Firstly, we make use of the Hanning window and the wavelet threshold filtering to depress the noises (such as the GWN) and the interfering signals in the auxiliary interferometer and measurement interferometer.
- (2) Then, we get the sampling points (all of the maximum and minimum extreme points of the auxiliary interference signal) and set appropriate thresholds to delete most of the erroneous sampling points caused by mode-hopping signals.
- (3) Lastly, according to the rule that the maximum and minimum values occur alternately, we remove the rest of the erroneous sampling points.

Step (1) is used to depress the noises and the interfering signals in the auxiliary interferometer and measurement interferometer. The Hanning window filtering method had been proved to be an effective method to depress the low and high frequency noise in an FSI-based system [10]. In our experiments, the measurement signal was very weak, and its frequency was varying with the measured distance, so we must use an adaptive filtering method to depress the GWN. The wavelet threshold filtering method is adopted in our paper. The mechanism of wavelet threshold filtering is based on the different properties of the wavelet coefficients of signals and noises on scales. To eliminate the noise, corresponding rules are adopted to deal with the nonlinear processing of the wavelet coefficients of the noise [24].

Step (2) and Step (3) are used to delete the erroneous sampling points caused by the mode-hopping of the ECL. Firstly, we get all of the sampling points in the auxiliary interferometer. The set of all of the sampling points can be described as:

$$Ep = Ep_{max} + Ep_{min} \tag{17}$$

where  $Ep$  denotes a set of all of the sampling points,  $Ep_{max}$  denotes a set of all of the maximum extreme points, and  $Ep_{min}$  denotes a set of all of the minimum extreme points. Then, we find that the amplitude and the time interval of the erroneous sampling points is obviously smaller than the correct sampling points. Firstly, we set the amplitude thresholds to delete part of the erroneous sampling points, and the rest of the sampling points can be written as:

$$Ep_1 = \{x | A(x_k) > 0.6 A_{ma} + 0.4 A_{mi} \vee A(x_k) < 0.6 A_{mi} + 0.4 A_{ma}, x \in Ep, k = 1, 2, \dots, S\} \tag{18}$$

where  $Ep_1$  denotes a set of the sampling points that contains part of the erroneous sampling points,  $A(x_k)$  is the amplitude of the extreme point,  $A_{ma}$  is the maximum amplitude of the auxiliary interference signal,  $A_{mi}$  is the minimum amplitude of the auxiliary interference signal,  $k$  is the index of the extreme



points, and  $S$  is the length of  $Ep$ . Then, we set the time interval thresholds to delete part of the erroneous sampling points, and the rest of the sampling points can be written as:

$$Ep_2 = \{x | (T(x_k) - T(x_{k-1})) > 0.6\Delta T_{me}, x \in Ep_1, k = 2, 3, \dots, K\} \tag{19}$$

where  $x$  denotes the sampling point in  $Ep_1$ ,  $T(x_k)$  is the time interval of the extreme point,  $\Delta T_{me}$  denotes the mean time interval of all of the points in  $Ep_1$ , and  $K$  is the length of  $Ep_1$ .

Lastly, we find that the rest of the erroneous sampling points do not follow the rule that the maximum and minimum values occur alternately. So, the set of the correct sampling points can be described as:

$$Ep_3 = \{x | (A(x_k) - A(x_{k-1})) (A(x_{k-1}) - A(x_{k-2})) < 0, x \in Ep_2, k = 3, 4, \dots, M\} \tag{20}$$

where  $Ep_3$  denotes a set of the correct extreme points,  $k$  is the index of the correct extreme points, and  $M$  is the length of the  $Ep_2$ . Thus, the measurement errors that are caused by the erroneous sampling points and the interfering signals are eliminated.

In addition, the phase-matching mosaic algorithm and the segmentation mosaic algorithm are used to eliminate the phase jump and improve the data-processing speed. After sampling the measurement signal with  $Ep_2$ , phase jump occurs at the starting and ending point of each mode-hopping signal in the measurement interferometer [25]. To solve the problem, firstly we find the starting point and ending point of each mode-hopping signal by the time interval. Then, we use the Hilbert translation to exact the phase of the 10 points before the starting point and the 10 points after the ending point and compare their phases one by one. Lastly, we stitch the closest two points. When the tuning bandwidth, tuning speed, and sampling frequency of the data acquisition card are set, the required number of the sampling points is fixed. The first step of the segmentation mosaic algorithm is dividing the signal of the fixed length to 10 segment signals of the same length, namely:

$$I_m(k) = I_m(k1) + I_m(k2) + \dots + I_m(k10) \tag{21}$$

Then, we use the adaptive filtering method and the phase-matching mosaic algorithm to progress them separately. Finally, the measured distance should be described as:

$$L_m = \frac{n_f \sum_{i=1}^{10} P_i}{n_{air} \sum_{i=1}^{10} N_i} L_a \tag{22}$$

Each segment signal is one-tenth of the fixed length, so the time that FFT spent is greatly reduced.

### 3. Simulation

When there are noises and the mode-hopping signals in the auxiliary interferometer, the output of the auxiliary interferometer can be described as:

$$\begin{aligned} I_a(t) &= S_a(t) \text{Noise}_a(t) + MHS_a(t) + Low(t) + High(t) \\ S_a(t) &= 0.5 \sin(2\pi(100000t - 100t^2 + 100t^3)) \\ \text{Noise}_a(t) &= \text{awng}(S_a(t), \text{SNR}) - S_a(t) \\ MHS_a(t) &= 0.2 \sin(10000\pi t) + 0.2 \sin(8000\pi t), \\ T &= 0.1, 0.10000004, 0.100001 \\ Low(t) &= 0.3 \sin(200\pi t) \\ High(t) &= 0.3 \sin(20000000\pi t) \end{aligned} \tag{23}$$

where  $I_a(t)$  simulates the actual output of the auxiliary interferometer,  $S_a(t)$  simulates the ideal output of the auxiliary interferometer,  $\text{Noise}_a(t)$  simulates the Gaussian white noise,  $\text{awng}$  is a function in Matlab that simulates the GWN,  $MHS_a(t)$  simulates the mode-hopping signals,  $Low(t)$

simulates the low-frequency noise,  $High(t)$  simulates the high-frequency noise,  $t$  is the sampling time, and  $-100t^2 + 10t^3$  simulates the variation caused by the tuning nonlinearity of the ECL. The sampling frequency of the simulation is 5 MHz, and the number of sampling points is  $1 \times 10^6$ , which is the same with the experiments in Section 4.

When FSI-based system measures *non-cooperative* targets, the SNR of the measurement interferometer is low. We set the SNR of the measurement interferometer to 5 dB, which is the same with the SNR in the experiments, so the output of the measurement interferometer can be simulated as:

$$\begin{aligned}
 I_m(t) &= S_m(t) \text{Noise}_m(t) + MHS_m(t) + Low(t) + High(t) + I_i(t) \\
 S_m(t) &= 0.2 \sin(2\pi(30000t - 30t^2 + 30t^3)) \\
 MHS_m(t) &= 0.1 \sin(10000\pi t) + 0.1 \sin(8000\pi t), \\
 t &= 0.1, 0.10000004, 0.100001 \\
 I_i(t) &= 5 \sin(2\pi(10000t - 10t^2 + 10t^3)) + 3 \sin(2\pi(10000t - 10t^2 + 10t^3)) \\
 \text{Noise}_m(t) &= \text{awng}(S_m(t), 5) - S_m(t)
 \end{aligned}
 \tag{24}$$

where  $I_m(t)$  simulates the actual output of the measurement interferometer,  $S_m(t)$  simulates the ideal output of the measurement interferometer,  $\text{Noise}_m(t)$  simulates the GWN,  $I_i(t)$  simulates the interfering signals, and  $-30t^2 + 30t^3$  is the variable caused by the ECL nonlinearity tuning. The ideal frequency is 30,000 Hz, and all of the signal processing was performed in MATLAB.

Figure 5 shows the comparison of the ideal frequency (the red line) with the simulated result of our method (the green line). When the SNR of the auxiliary interferometer ranges from 5 dB to 30 dB, the result of our method is the same with the ideal frequency. However, if the SNR is 0 dB, the result of our method is slightly smaller than the ideal frequency. In our experiments, the SNR of the auxiliary interferometer was about 25 dB, and our adaptive filtering method was effective in this situation.

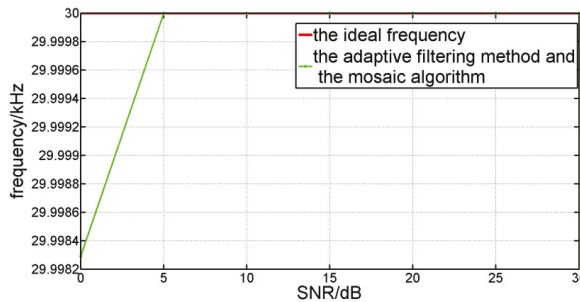


Figure 5. Comparison between the ideal frequency and the simulated result.

#### 4. Experiment and Analysis

Figure 6 shows the setup of our experiments. It includes two parts: the FSI-based absolute distance measurement system, and the interferometer (Renishaw XL-80). The ECL has been used for five years, and mode-hopping of the ECL (NewfocusTLB-6728) always happens. In experiments, the power of the ECL was 8 mW, and the scanning range of the ECL was 1515–1535 nm. The optical frequency was set to sweep as a triangle at tuning rate of 100 nm/s. The sampling frequency of the DAQ (Gage CSE161G4) was 5 MHz, and the number of sampling points was  $1 \times 10^6$ . The measured target was a 10-cent coin, and the incident angle of the measurement signal was 12 degrees, so the backward echo signal was very weak.

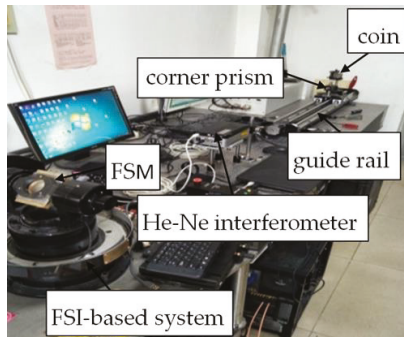


Figure 6. Experiment setup.

The He-Ne interferometer was used to verify the accuracy of our method. The cube-corner prism of the He-Ne interferometer and the target of the FSI system were both fixed on a slider. The slider moved along a precise guiderail, and the range of the guiderail was 1000 mm.

The measured result in Figure 4 consists of two parts. One was the distance from the ECL to the fast steering mirror (FSM) of our FSI-based system, and it can be obtained when the FSM was horizontal. Another was the measured distance from the FSM to the measured target, such as the results in Figures 7 and 8. In our experiments, the former was 8385.432 mm. The measured distance in Figures 7 and 8 was the measured result minus 8.385.432 mm.

Figure 7 shows the comparison of the experiment results between the method of Deng et al. [17] (part a) and the adaptive filtering method and the mosaic algorithm (part b) with the same experiment data. It proves the necessity and effectiveness of our method to deal with the noises and the mode-hopping signals.

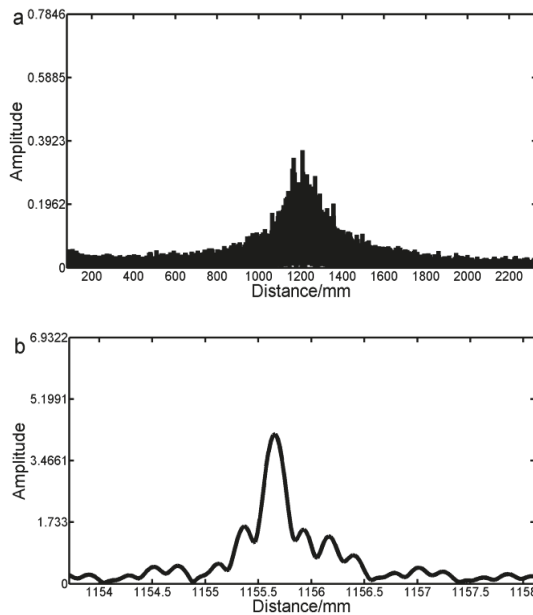


Figure 7. Part (a,b) respectively represent the experiment results of not using and using the adaptive filtering method and the mosaic algorithm with the same data.

In order to analyze the stability of our method, we measured the coin 20 times at one position and used our adaptive filtering method and mosaic algorithm to process the experiment data. As shown in Figure 8, the blue line denoted the results of 20 measurements, and the green line denoted the residuals. When there were noises, the mode-hopping signals and the interfering signals in our system, the statistical uncertainty of our method was  $5.7 \mu\text{m}$  over 20 measurements, which was the same order as the result of Lu et al. [16].

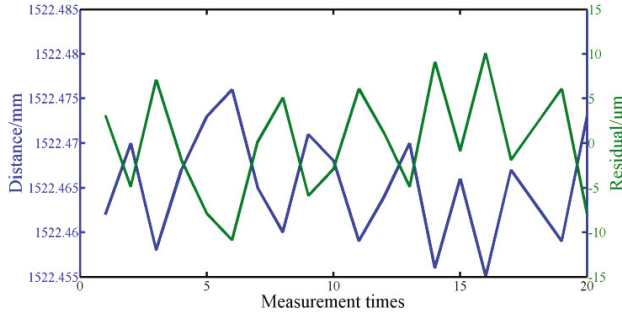


Figure 8. The measurement results of our FSI-based system.

To evaluate the accuracy of our method, the coin moved  $62.45 \text{ mm}$  every time, and totally moved seven times. The FSI-based system and the He-Ne interferometer simultaneously measured the displacement of the slider. So, our FSI-based system and the laser interferometer could use the same air refractive index. The measured results of the laser interferometer were used as the truth value of the moving distance. Figure 9 showed the difference between the measured results of our system and the laser interferometer. The max relative error of our method was  $9.6 \mu\text{m}$ . The relative error was attributed to the Abbe error, the air-path distance variation, and the optical dispersion error. In our experiment, the moving distance was less than  $500 \text{ mm}$ , so the Abbe error between the laser interferometer and our FSI-based system was less than  $0.5 \mu\text{m}$ . The air-path distance variation was mainly caused by the variation of the air refractive index. The temperature, pressure and humidity in the experiments were basically stable, so the error caused by the air-path distance variation was within  $1 \mu\text{m}$ . The optical dispersion error was originated from the optical frequency sweep of the ECL, and was compensated by Equation (7). According to Equation (7), the moving distance was  $0.06245 \text{ m}$ ,  $\beta$  was  $-23 \text{ ps}^2/\text{km}$ , and the tuning bandwidth was  $10 \text{ nm}$ , so the distance offset was  $2.3 \mu\text{m}$ .

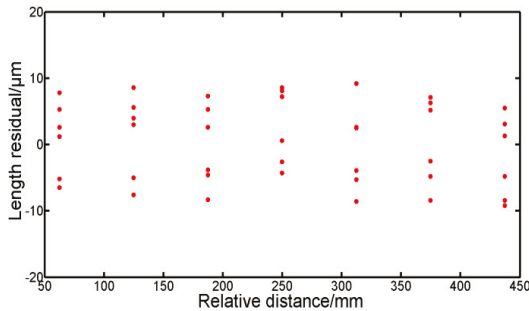


Figure 9. Distance residual between the measured results of our FSI-based system and the laser interferometer.

The CRLB was used to evaluate the performances of our system. In our experiments,  $n_f$  was  $1.4682$ ,  $n_{air}$  was  $1.0003$ ,  $L_a$  was  $20.05832 \text{ m}$ , the SNR was  $5 \text{ dB}$ , and the ideal value of  $N$  was about

60,000. According to Equation (16), the CRLB was about  $0.8 \mu\text{m}$ . Figure 10 showed the comparison of the measured root mean square error (MSEs) at eight positions with the CRLB. The red and pink data denoted the MSE obtained by the results of six independent measurements and the CRLB of ranging estimation, respectively. The max MSE based on our method at eight different placements was  $2.4 \mu\text{m}$ , which approached the CRLB of  $0.8 \mu\text{m}$ .

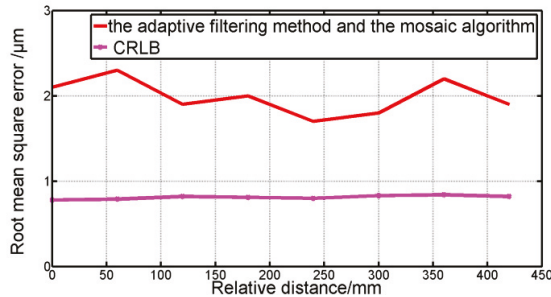


Figure 10. Comparison of the measured root mean square error with the CRLB.

## 5. Conclusions

In our FSI-based system, the ECL has been used for five years. Thus, the SNR of the ECL is low, and the mode-hopping signals often occur. In addition, the influence of the interfering signals cannot be neglected when the system measures non-cooperative targets. To ensure the precision of our system, it is essential to delete the erroneous sampling points in the auxiliary interferometer and the interfering signals and phase jump in the measurement interferometer. So, the adaptive filtering method, the phase-matching mosaic algorithm, and the segmentation mosaic algorithm are proposed to solve the problems and improve the data processing speed. The results of our simulation prove the efficiency and applicability of our methods. In the experiments, the target was a coin and the laser interferometer (Renishaw XL80) was used to verify the efficiency of our method. The stability of our system was  $5.7 \mu\text{m}$  over 20 measurements. The maximum deviation of the measured results between the FSI-based system and the He-Ne interferometer was  $9.6 \mu\text{m}$ , and the maximum mean square error of our method was  $2.4 \mu\text{m}$ , which approached the Cramer-Rao lower bound (CRLB)  $0.8 \mu\text{m}$ .

**Author Contributions:** Conceived the Method and Wrote the Paper, X.-T.X.; Designed the Experiments, X.-H.Q.; Improved the Experimental Results, F.-M.Z.

**Funding:** Natural National Science Foundation of China (NSFC) (No. 51675380, No. 51775379), National Key Research and Development Plan(2018TFF0212702).

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Barwood, G.P.; Gill, P.; Rowley, W.R.C. High-accuracy length metrology using multiple-stage swept-frequency interferometry with laser diodes. *Meas. Sci. Technol.* **1999**, *9*, 1036–1041. [[CrossRef](#)]
2. Stejskal, A.; Stone, J.A.; Howard, L. Absolute interferometry with a 670-nm external cavity diode laser. *Appl. Opt.* **1999**, *38*, 5981–5994.
3. Le, F.S.; Salvadé, Y.; Mitouassiou, R.; Favre, P. Radio frequency controlled synthetic wavelength sweep for absolute distance measurement by optical interferometry. *Appl. Opt.* **2008**, *47*, 3027–3031.
4. Gao, R.; Wang, L.; Teti, R.; Dornfeld, D.; Kumara, S.; Mori, M.; Helu, M. Cloud-enabled prognosis for manufacturing. *CIRP Ann. Manuf. Technol.* **2015**, *64*, 749–772. [[CrossRef](#)]
5. Abou-Zeid, A.; Pollinger, F.; Meiners-Hagen, K.; Wedde, M. Diode-laser-based high-precision absolute distance interferometer of 20 m range. *Appl. Opt.* **2009**, *48*, 6188–6194.


6. Kinder, T.; Salewski, K.D. Absolute distance interferometer with grating-stabilized tunable diode laser at 633 nm. *J. Opt. Pure Appl. Opt.* **2002**, *4*, S364–S368. [[CrossRef](#)]
7. Dale, J.; Hughes, B.; Lancaster, A.J.; Lewis, A.J.; Reichold, A.J.H.; Warden, M.S. Multi-channel absolute distance measurement system with sub ppm-accuracy and 20 m range using frequency scanning interferometry and gas absorption cells. *Opt. Express* **2014**, *22*, 24869–24893. [[CrossRef](#)] [[PubMed](#)]
8. Prellinger, G.; Meinershagen, K.; Pollinger, F. Spectroscopically in situ traceable heterodyne frequency-scanning interferometry for distances up to 50 m. *Meas. Sci. Technol.* **2015**, *26*, 084003. [[CrossRef](#)]
9. Wang, L.T.; Iiyama, K.; Tsukada, F.; Yoshida, N.; Hayashi, K.I. Loss measurement in optical waveguide devices by coherent frequency-modulated continuous-wave reflectometry. *Opt. Lett.* **1993**, *18*, 1095–1097. [[CrossRef](#)] [[PubMed](#)]
10. Zheng, K.; Liu, B.; Huang, C.; Brezinski, M.E. Experimental confirmation of potential swept source optical coherence tomography performance limitations. *Appl. Opt.* **2008**, *47*, 6151–6158. [[CrossRef](#)] [[PubMed](#)]
11. Baumann, E.; Giorgetta, F.R.; Deschênes, J.D.; Swann, W.C.; Coddington, I.; Newbury, N.R. Comb-calibrated laser ranging for three-dimensional surface profiling with micrometer-level precision at a distance. *Opt. Express* **2014**, *22*, 24914–24928. [[CrossRef](#)] [[PubMed](#)]
12. Ahn, T.J.; Kim, D.Y. Analysis of nonlinear frequency sweep in high-speed tunable laser sources using a self-homodyne measurement and Hilbert transformation. *Appl. Opt.* **2007**, *46*, 2394–2400. [[CrossRef](#)] [[PubMed](#)]
13. Coddington, I.; Swann, W.C.; Nenadovic, L. Rapid and precise absolute distance measurements at long range. *Nat. Photonics* **2009**, *3*, 351–356. [[CrossRef](#)]
14. Barber, Z.W.; Babbitt, W.R.; Kaylor, B.; Reibel, R.R.; Roos, P.A. Accuracy of active chirp linearization for broadband frequency modulated continuous wave lidar. *Appl. Opt.* **2010**, *49*, 213–219. [[CrossRef](#)] [[PubMed](#)]
15. Yuksel, K.; Wuilpart, M.; Mégret, P. Analysis and suppression of nonlinear frequency modulation in an optical frequency-domain reflectometer. *Opt. Express* **2009**, *17*, 5845–5851. [[CrossRef](#)] [[PubMed](#)]
16. Lu, C.; Liu, G.D.; Liu, B.G. Absolute distance measurement system with micron-grade measurement uncertainty and 24 m range using frequency scanning interferometry with compensation of environmental vibration. *Opt. Express* **2016**, *24*, 30215–30224. [[CrossRef](#)] [[PubMed](#)]
17. Deng, Z.; Liu, Z.; Li, B.; Liu, Z. Precision improvement in frequency scanning interferometry based on suppressing nonlinear optical frequency sweeping. *Opt. Rev.* **2015**, *22*, 724–730. [[CrossRef](#)]
18. Ahn, T.J.; Lee, J.Y.; Kim, D.Y. Suppression of nonlinear frequency sweep in an optical frequency-domain reflectometer by use of Hilbert transformation. *Appl. Opt.* **2005**, *44*, 7630–7634. [[CrossRef](#)] [[PubMed](#)]
19. Satyan, N.; Vasilyev, A. Phase-locking and coherent power combining of broadband linearly chirped optical waves. *Opt. Express* **2012**, *20*, 25213–25227. [[CrossRef](#)] [[PubMed](#)]
20. Lippok, N.; Coen, S.; Nielsen, P. Dispersion compensation in Fourier domain optical coherence tomography using the fractional Fourier transform. *Opt. Express* **2012**, *20*, 23398–23413. [[CrossRef](#)] [[PubMed](#)]
21. Liu, G.; Xu, X.; Liu, B.F. Dispersion compensation method based on focus definition evaluation functions for high-resolution laser frequency scanning interference measurement. *Opt. Commun.* **2017**, *386*, 57–64.
22. Gifford, D.K.; Soller, B.J.; Wolfe, M.S. Optical vector network analyzer for single-scan measurements of loss, group delay, and polarization mode dispersion. *Appl. Opt.* **2005**, *44*, 7282–7286. [[CrossRef](#)] [[PubMed](#)]
23. Gart, J. An extension of the Cramér–Rao inequality. *Ann. Math. Stat.* **1958**, *29*, 367–380. [[CrossRef](#)]
24. Hussain, M. Mammogram Enhancement Using Lifting Dyadic Wavelet Transform and Normalized Tsallis Entropy. *J. Comput. Sci. Technol.* **2014**, *29*, 1048–1057. [[CrossRef](#)]
25. Guang, S.; Wen, W. Single laser complex method to improve the resolution of FMCW laser ranging. *J. Infrared Millim. Waves* **2016**, *35*, 363–367.



© 2018 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).

Article

# Influence of Illumination Polarization and Target Structure on Measurement Sensitivity of Through-Focus Scanning Optical Microscopy

Yufu Qu <sup>1</sup> , Renju Peng <sup>1</sup>, Jialin Hao <sup>1</sup>, Hui Pan <sup>1</sup>, Jiebin Niu <sup>2</sup> and Jie Jiang <sup>1,\*</sup>

<sup>1</sup> School of Instrumentation Science and Opto-Electronics Engineering, Beihang University, Beijing 100191, China; qyf@buaa.edu.cn (Y.Q.); rjpeng@buaa.edu.cn (R.P.); hjl\_ace@buaa.edu.cn (J.H.); panhui0728@163.com (H.P.)

<sup>2</sup> Institute of Microelectronics of CAS, Chinese Academy of Sciences, Beijing 100029, China; niujiebin@ime.ac.cn

\* Correspondence: jiangjie@buaa.edu.cn

Received: 15 August 2018; Accepted: 27 September 2018; Published: 3 October 2018



**Abstract:** Unlike the optical information taken from a single in-focus image of general optical microscopy, through-focus scanning optical microscopy (TSOM) involves scanning a target through the focus and capturing of a series of images. These images can be used to conduct three-dimensional inspection and metrology with nanometer-scale lateral and vertical sensitivity. The sensitivity of TSOM strongly depends on many mechanical and optical factors. In this study, how illumination polarization and target structure affect the sensitivity of TSOM is analyzed. Firstly, the complete imaging procedure of the polarized light is investigated. Secondly, through-focus scanning results of different targets with two illumination polarizations are simulated using the finite-difference time-domain method. Thirdly, a few experiments are performed to verify the influence of illumination polarization and target structures on the sensitivity of TSOM. Both the results of the simulation and experiments illustrate an apparent influence of polarization on the sensitivity of inspecting the targets with center asymmetric structures. For enhanced sensitivity, illumination polarization should be perpendicular to the target texture. This conclusion is meaningful to adjust illumination polarization purposefully for different structure characteristics and improve the sensitivity of metrology.

**Keywords:** through-focus optical microscopy; illumination polarization; target structure; sensitivity

## 1. Introduction

With gradually increased applications of three-dimensional (3D) components in the fields of semiconductor industry and nanotechnology, obtaining 3D information of these targets is critical. However, conducting 3D topography optical metrology of nanoscale targets for general optical microscopy is becoming increasingly challenging due to severe diffraction.

Through-focus scanning optical microscopy (TSOM) is a novel and fast optical metrology. TSOM is not limited by the need to acquire an in-focus image in the conventional measurement. This method utilizes a TSOM image by scanning the target through the focus, and it extracts dimensional information of the targets by analyzing scattering intensity of the TSOM image and matching the TSOM image with the simulation results in the database [1]. Low resolution due to the diffractive limit in conventional optical metrology is consequently avoided. The lateral and vertical sensitivity of TSOM is quantified by the optical intensity range (OIR) of the TSOM image, which is defined as the absolute difference between the maximum and minimum intensity of the normalized TSOM image [2].

The sensitivity of TSOM strongly depends on many mechanical and optical factors [3–5]. The critical dimension of a target is usually smaller than 1/10 of wavelength of the illumination light;

thus, the interaction between incident light wave and target surface is complicated, and the spatial distribution of scattered light wave depends on illumination polarization and target structures [6]. TSOM images with different OIRs correspond to different spatial distributions of scattered light [7]. Consequently, the illumination polarization and structure characteristics of the target strongly affect the sensitivity of TSOM [8,9]. To address the lacking systematic analysis of these effects, many works attempt to optimize illumination polarization according to different targets through pre-tests [10,11]. The drawbacks are long preparation time and low inspection efficiency. To understand the light scattering around the target surface and conduct the inspection in the most optimized condition, the effects on TSOM sensitivity brought by illumination polarization and target structure are analyzed and summarized in this paper.

The reminder of this article is organized as follows: the complete image-forming procedure, from incidence of illumination to the imaging plane of the charge-coupled device (CCD), is analyzed in the second section. The simulation of the system is presented in the third section. A few experiments are performed in the fourth section to verify the simulation results. The simulation and experimental results are interpreted in the fifth section. A conclusion based on the analyses and results is proposed in the sixth section.

## 2. Principle Analysis

The principle analysis of TSOM is separated into three parts—illuminating, scattering, and imaging. Firstly, Köhler illumination in the TSOM system is explored. Secondly, the scattered electromagnetic field is analyzed after the illumination wave is reflected from a complex target. Thirdly, the image-scanning procedure along the optic-axis direction that collects an in-focus image and a series of defocused images of the target is investigated.

### 2.1. Köhler Illumination

A Köhler illumination system in reflective mode is widely used in TSOM. Under this condition, the illumination beam from each source point in the conjugate back focal plane of the objective lens converges to a few plane waves  $E_{i,j}$  with different angular directions [12,13].

These plane waves are similar with the definition of the angular spectrum. Thus, the Köhler illumination of the TSOM can be defined as a summation of these plane waves.

$$E(\mathbf{r}) = \sum_{i,j} E_{i,j} e^{-\hat{i}k_{ij} \cdot \mathbf{r}} \quad (i, j = \text{integer}), \tag{1}$$

where  $E(\mathbf{r})$  is the illuminating light wave near the object surface,  $\hat{i}$  is the imaginary unit here,  $k_{ij}$  is the corresponding wave vector of  $E_{i,j}$ , and  $\mathbf{r}$  is the location vector.

### 2.2. Scattered Electromagnetic Field from the Target

The incident illumination light on the target surface is scattered and forms a superimposed field with the incident field and a scattered field. The superimposed field is defined as follows [14]:

$$E(\mathbf{r}) = E^{(inc)}(\mathbf{r}) + E^{(sca)}(\mathbf{r}), \tag{2}$$

$$H(\mathbf{r}) = H^{(inc)}(\mathbf{r}) + H^{(sca)}(\mathbf{r}), \tag{3}$$

where the superscript (*inc*) is the incident field, and the superscript (*sca*) denotes the scattered field. Combined with Maxwell equations, Equations (2) and (3) can be rewritten as follows:

$$E(\mathbf{r}) = E^{(inc)}(\mathbf{r}) + \nabla \cdot \nabla \cdot \Pi_e(\mathbf{r}) + \hat{i}k \nabla \cdot \Pi_m(\mathbf{r}) - 4\pi P(\mathbf{r}), \tag{4}$$

$$H(\mathbf{r}) = H^{(inc)}(\mathbf{r}) + \nabla \cdot \nabla \cdot \Pi_m(\mathbf{r}) - \hat{i}k \nabla \cdot \Pi_e(\mathbf{r}) - 4\pi M(\mathbf{r}), \tag{5}$$



where  $\mathbf{P}(\mathbf{r})$  and  $\mathbf{M}(\mathbf{r})$  are the incident polarized intensity of the electric and magnetic vectors in the target, respectively.

$$\mathbf{P}(\mathbf{r}) = \eta(\mathbf{r})\mathbf{E}(\mathbf{r}), \tag{6}$$

$$\mathbf{M}(\mathbf{r}) = \chi(\mathbf{r})\mathbf{H}(\mathbf{r}), \tag{7}$$

where  $\eta(\mathbf{r})$  and  $\chi(\mathbf{r})$  are dielectric and magnetic susceptibilities, respectively.  $\Pi_e(\mathbf{r})$  and  $\Pi_m(\mathbf{r})$  are electric and magnetic Hertz potentials, respectively, as expressed in the following equations:

$$\Pi_e(\mathbf{r}) = \int_V \mathbf{P}(\mathbf{r}') \frac{e^{i\hat{k}|\mathbf{r}-\mathbf{r}'|}}{|\mathbf{r}-\mathbf{r}'|} d^3r', \tag{8}$$

$$\Pi_m(\mathbf{r}) = \int_V \mathbf{M}(\mathbf{r}') \frac{e^{i\hat{k}|\mathbf{r}-\mathbf{r}'|}}{|\mathbf{r}-\mathbf{r}'|} d^3r', \tag{9}$$

where  $V$  is the target's volume in space.

Equations (6), (7), (8), and (9) are substituted into Equations (4) and (5). The total field can be described by integral–differential equations in which each component of the electric and magnetic fields is coupled with others. The effect of coupling is negligible because the dimension of the target is much smaller than the wavelength. Therefore, this scattering should not be described with a simple scalar equation.

According to the relationship between  $\eta(\mathbf{r})$  and relative permittivity  $\epsilon_r(\mathbf{r})$ ,

$$\epsilon_r(\mathbf{r}) = 1 + \eta(\mathbf{r}), \tag{10}$$

the relationship between  $\epsilon_r(\mathbf{r})$  and refractive index  $n(\mathbf{r})$ ,

$$\epsilon_r(\mathbf{r}) \cdot \epsilon_0 = n^2(\mathbf{r}), \tag{11}$$

and the equivalent transformation from the topographic structure to the refractive index around the surface, the effect of illumination polarization and target structure on the spatial distribution of the scattered light is provided as follows:

$$\mathbf{E}(\mathbf{r}) = \mathbf{E}^{(inc)}(\mathbf{r}) + \nabla \cdot \nabla \cdot \int_V \mathbf{P}'(\mathbf{r}') d^3r' + i\hat{k} \nabla \cdot \int_V \mathbf{M}'(\mathbf{r}') d^3r' - 4\pi \left( \frac{n^2(\mathbf{r})}{\epsilon_0} - 1 \right) \mathbf{E}(\mathbf{r}), \tag{12}$$

$$\mathbf{H}(\mathbf{r}) = \mathbf{H}^{(inc)}(\mathbf{r}) + \nabla \cdot \nabla \cdot \int_V \mathbf{M}'(\mathbf{r}') d^3r' - i\hat{k} \nabla \cdot \mathbf{P}'(\mathbf{r}') d^3r' - 4\pi \chi(\mathbf{r}) \mathbf{H}(\mathbf{r}), \tag{13}$$

where

$$\mathbf{P}'(\mathbf{r}') = \left( \frac{n^2(\mathbf{r}')}{\epsilon_0} - 1 \right) \mathbf{E}(\mathbf{r}') \frac{e^{i\hat{k}|\mathbf{r}-\mathbf{r}'|}}{|\mathbf{r}-\mathbf{r}'|},$$

and

$$\mathbf{M}'(\mathbf{r}') = \chi(\mathbf{r}') \mathbf{H}(\mathbf{r}') \frac{e^{i\hat{k}|\mathbf{r}-\mathbf{r}'|}}{|\mathbf{r}-\mathbf{r}'|}.$$

For a target with given topographic features, calculating the specific marginal solution of Equations (12) and (13) provides the near-field scattered electromagnetic field theoretically. However, this procedure is complicated, in that obtaining analytic solutions is impossible. In this study, the finite-difference time-domain (FDTD) method is used to calculate the numerical solutions.

### 2.3. Imaging Procedure

The imaging system of the TSOM is the same as a bright-field microscopy. The complete imaging system includes an illuminating path and an imaging path. These two paths can be described by their corresponding optical transfer function (OTF). The OTF of the illuminating path is [6]

$$H_{ill}(k_{i,j}) = \exp \left[ \hat{i}D \sqrt{\left(\frac{2\pi}{\lambda}\right)^2 - |k_{i,j}|^2} \right], \quad (14)$$

where  $D$  is the illumination spot defocus and a variable here. The OTF of the imaging path is

$$H_{ima}(k_{k,m}) = \begin{cases} 1 & |k_{k,m}| \leq \frac{2\pi}{\lambda} NA \\ 0 & else \end{cases}, \quad (15)$$

where  $NA$  is the numerical aperture of objective lens. The complete system–structure interaction is described by

$$E_{k,m}^{out} = H_{ima}(k_{k,m}) \cdot S_{i,j,k,m} \cdot H_{ill}(k_{i,j}) \cdot E_{i,j}^{in}, \quad (16)$$

where  $S_{i,j,k,m}$  is a four-dimensional scattering matrix, and the two pairs of indices  $i, j$  and  $k, m$  describe the in-plane components of the wavevectors for the illuminating plane waves  $E_{i,j}^{in}$  and scattering plane waves  $E_{k,m}^{out}$ , respectively. Finally, as each illuminating plane wave component corresponds to a specific source point in Kohler illumination, and the light source is incoherent, the subsequent scattering plane waves are assumed as incoherent as well. Thus, each pixel of the image is generated from the incoherent sum of output plane waves scattered from a given location of the structure, as depicted in Equation (17).

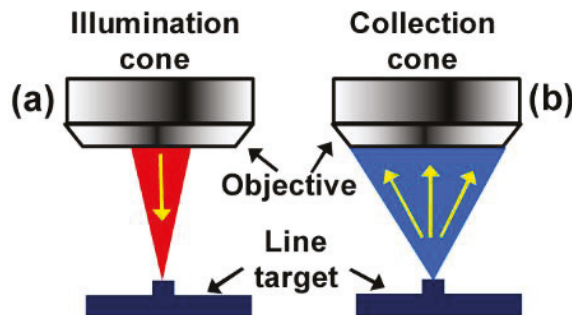
$$I(r) = \sum_{k,m} \left| E_{k,m}^{out} e^{ik_{k,m} \cdot r} \right|^2, \quad (17)$$

where  $r$  denotes the location vector toward the CCD plane.

## 3. Simulation

### 3.1. Target Structure

A commercial optical simulation program (i.e., FDTD) was used to calculate the light wave scattered by the target and the simulation of imaging process. Table 1 lists the structure parameters of six different targets and three imaging parameters, namely the illumination numerical aperture (INA), the collection numerical aperture (CNA), and the wavelength. As Figure 1 depicts, a given INA determines a specific illumination cone, and the collection cone is constrained by CNA.

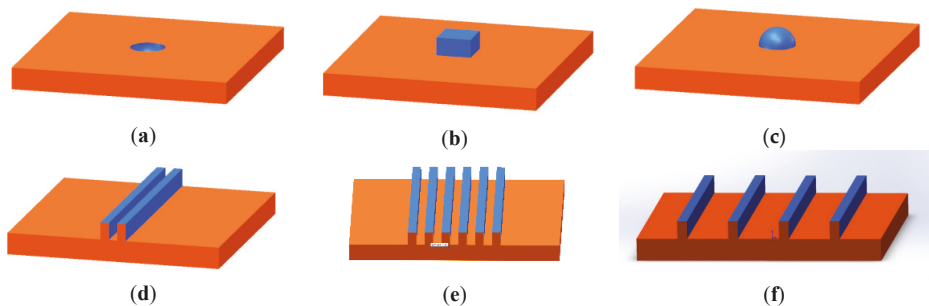


**Figure 1.** Schematic of (a) illumination numerical aperture (INA; red) and (b) collection numerical aperture (can; blue).

**Table 1.** Structure and imaging parameters used in the simulation.

Target	Structure Parameter		Imaging Parameter		
			Illumination Numerical Aperture	Collection Numerical Aperture	Wavelength (nm)
Nano-pit	Radius	100 nm	0.4	0.8	546
Nano-dot	Length	121 nm	0.4	0.8	546
	Width	121 nm			
	Height	71 nm			
Nano-bump	Radius	100 nm	0.4	0.8	546
Double silicon lines	Width	30 nm	0.4	0.8	546
	Height	50 nm			
	Pitch	20 nm			
	Sidewall angle	90°			
Dense line array	Width	30 nm	0.4	0.8	546
	Height	50 nm			
	Pitch	60 nm			
	Sidewall angle	90°			
Infinite line grating	Width	30 nm	0.4	0.8	546
	Height	50 nm			
	Pitch	600 nm			
	Sidewall angle	90°			

The materials of the four targets were all silicon, and the target structures are demonstrated in Figure 2, where different colors are used to highlight their structural features. Each model was illuminated by two kinds of light waves with the electric vector perpendicular or parallel to the incident plane. The structure of the first three targets was centrosymmetric, that is, the structures were consistent along X- and Y-axes, whereas the structure of the last target was center asymmetric. For convenience, the lines of the silicon line array were oriented perpendicular to the incident plane, so the porization of illumination light could be described as either perpendicular (0° polarization) or parallel (90° polarization) to the silicon lines.



**Figure 2.** Computer-aided design (CAD) models for (a) nano-pit, (b) nano-dot, (c) nano-bump, (d) double silicon lines, (e) dense line array, and (f) infinite line grating.

### 3.2. Simulation Procedure

The complete simulation was separated into three parts. In the first part, the interaction between the light wave and target was calculated with the FDTD method. In this part, the scattering matrix  $S_{i,j;k,m}$  depicted in Equation (16) is expected to be derived. The scattering matrix is dependent only on the simulated structure and wavelength of light, and it describes the complete interaction with two illumination polarizations. In the second part, an in-focus image and a set of defocused images were

simulated by changing the value of  $D$  in Equation (14). In the last part, a TSOM image was obtained using a particular method.

Attota et al. proposed the details of obtaining both the simulated and experimental TSOM images [10], and the procedure is restated below for convenience.

- Obtain an in-focus image and a series of defocused images of the target and a smooth background surface, respectively, as they are scanned along the axis direction.
- Normalize each through-focus image of the target and background as follows:

$$Normalized\ Image = \frac{Unnormalized\ Image}{Mean\ Value} - 1. \tag{18}$$

- For the in-focus or each defocused position, directly subtract the normalized background image from the normalized target image.
- Select the inspection region with a strip window on the subtracted normalized image at each focus height, and extract the intensity profile by averaging the image intensity along the width of the window.
- Stack the averaged intensity profile of each subtracted normalized through-focus image according to its corresponding focus height to construct a raw TSOM image, which contains three-dimensional information. X- (horizontal), Y- (vertical), and Z-axes represent the spatial position across the target, the focus position, and the optical intensity, respectively.
- Interpolate and smooth the raw TSOM image and add pseudo color. The result is the final TSOM image needed.

### 3.3. Simulation Results

The simulated TSOM images of the six targets in the illumination polarizations of  $90^\circ$  and  $0^\circ$  are illustrated in Figure 3. The relative difference of OIR ( $OIR_{rd}$ ) is calculated in Equation (19).

$$OIR_{rd} = \frac{|OIR_{0^\circ} - OIR_{90^\circ}|}{OIR_{90^\circ}}. \tag{19}$$

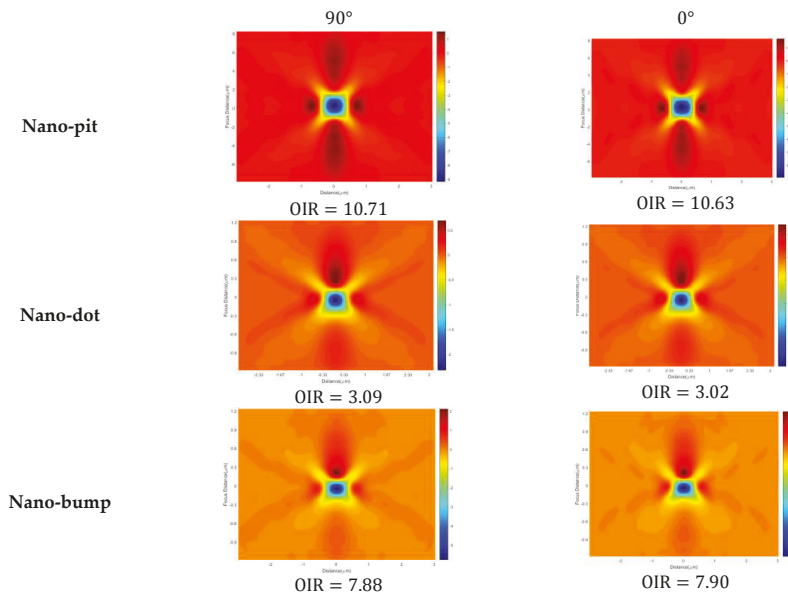
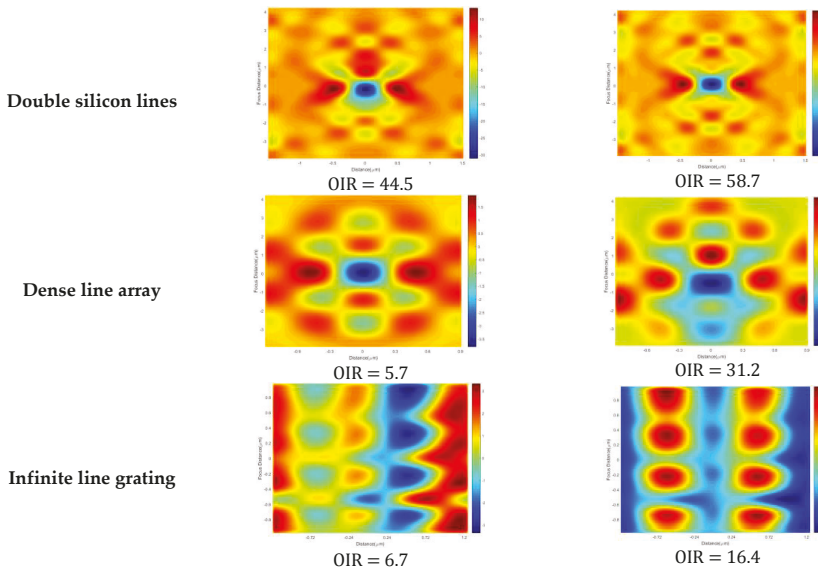


Figure 3. Cont.



**Figure 3.** Simulated through-focus scanning optical microscopy (TSOM) images of the six targets with different structural features with illumination polarizations of  $90^\circ$  (the first column) and  $0^\circ$  (the second column); OIR—optical intensity range; INA = 0.4.

Table 2 provides the results.

**Table 2.** Relative difference between optical intensity ranges (OIRs) with two illumination polarizations. TSOM—through-focus scanning optical microscopy.

Targets	Relative Difference (TSOM)
Nano-pit	0.74%
Nano-dot	2.3%
Nano-bump	0.25%
Double silicon lines	31.9%
Dense line array	447%
Infinite line grating	145%

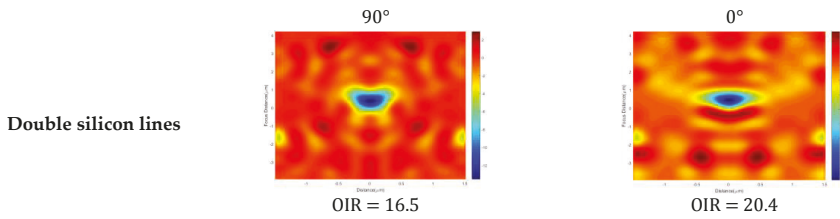
The first three targets were centrosymmetric. Under this condition, their TSOM images appear to be qualitatively similar. The corresponding OIRs in the two illumination polarizations were not strongly different, which illustrates a weak influence of polarization on the sensitivity of TSOM with regards to the targets of centrosymmetric features.

The others were center asymmetric, that is, the silicon lines were oriented parallel to the Y-axis, but periodic along the X-axis. Under this condition, an apparent influence of illumination polarization on the sensitivity of TSOM can be observed as the two TSOM images and corresponding OIRs were strongly different.

The effect brought about by illumination numerical aperture (INA) was also studied by simulating the TSOM of the double silicon line upon setting INA = 0.8 and keeping other parameters in line with those in the previous simulation. Figure 4 presents the simulation results.

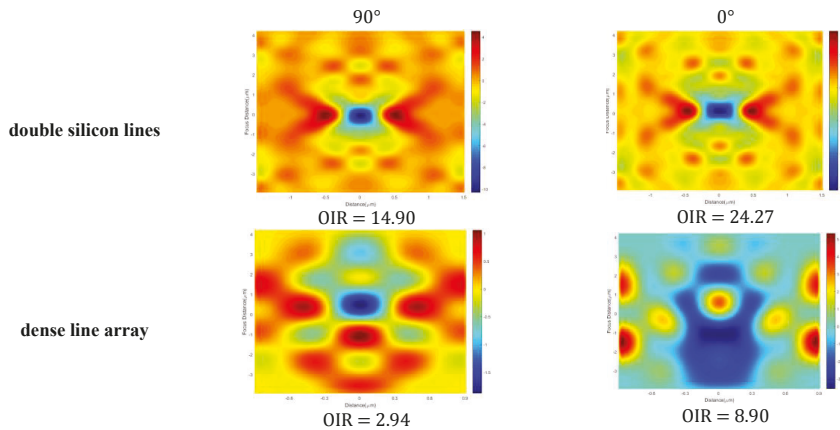
By comparing the results of Figures 3 and 4, it is demonstrated that the OIR of the TSOM image with the illumination numerical aperture of 0.8 is smaller than that with the illumination numerical aperture of 0.4. However, the two corresponding OIRs in both illumination polarizations are still different. It is concluded that the effect brought about by the illumination polarization and the

target structural feature exists as long as the metrology is performed in the same illuminating and imaging conditions.



**Figure 4.** Simulated TSOM images of double silicon lines with illumination polarizations of 90° (the first column) and 0° (the second column); INA = 0.8.

To analyze this phenomenon further, the line height of the double silicon line and dense line array were increased by 10 nm, but the other parameters are kept constant. This dimensional change produced a new TSOM image. The dimensional change was highlighted by a simple subtraction of the new and previous TSOM images and illustrated in a differential TSOM (DTSOM) image. The inspection of the vertical sensitivity of the dimensional change was directly quantified by the OIR of the DTSOM image. The DTSOM images in the two illumination polarizations are shown in Figure 5. The relative difference of OIR is provided in Table 3.



**Figure 5.** Simulated differential TSOM (DTSOM) images for silicon line array with illumination polarizations of 90° (the first column) and 0° (the second column). The line height was increased by 10 nm, but the line width and pitch were kept constant.

**Table 3.** Relative difference between OIRs with two illumination polarizations. DTSOM—differential TSOM.

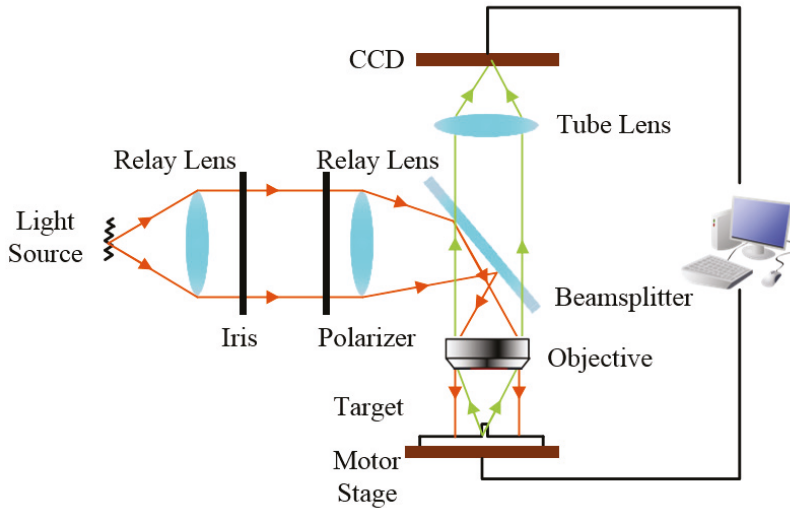
Target	Relative Difference (DTSOM)
Double silicon lines	62.9%
Dense line array	203%

## 4. Experiment

### 4.1. Experimental Set-Up

The optical scheme of the experiment is shown in Figure 6. A Zeiss Axio Imager M2 was used as the inspection device. The illumination numerical aperture (INA) and collection numerical aperture

(CNA) were both 0.8, and the magnification of the objective was  $100\times$ . The wavelength of illumination light was 555 nm. The optical path of illumination was designed as a Köhler illumination with a polarizer to produce the two kinds of linear polarizations mentioned above. A motor stage in closed-loop control was located beneath the microscope. The in-focus image and defocused images were collected by moving the motor stage vertically.



**Figure 6.** Schematic of the TSOM nanoscale metrology device, where the red ray is the illumination ray and the green ray is the imaging ray. The INA of the illuminating system is defined by the iris and the CNA of imaging system is defined by the objective. The motor stage and charge-coupled device (CCD) were controlled by a computer so that the CCD took one picture as soon as the motor stage moved a certain distance.

#### 4.2. Experimental Sample

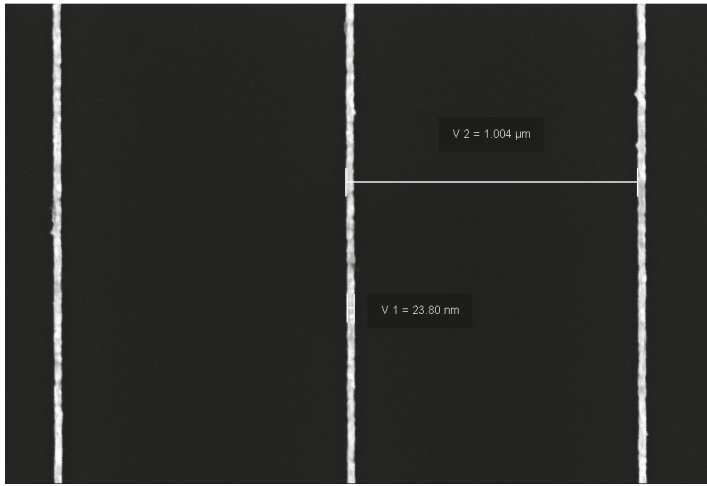
Two Au line arrays deposited on smooth silicon substrates were inspected. The line heights of the array were 50 nm and 70 nm, and the width of each Au line was approximately 20 nm. Each array was separated into 10 groups, and the nominal pitch of each group decreased from 1000 nm to 50 nm. The SEM of one Au line array is presented in Figure 7.

#### 4.3. Experimental Results

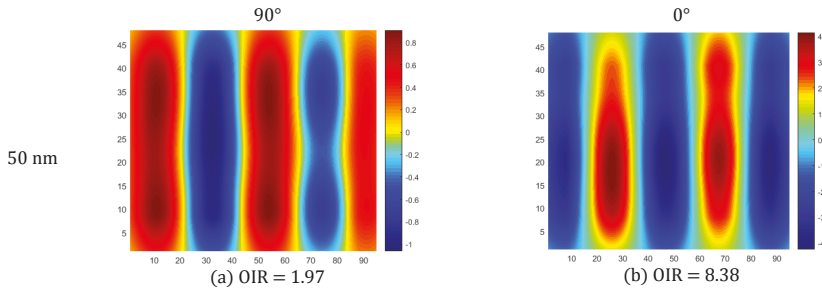
To verify our analysis and prove the effects of illumination polarization and structural features, the Au line array with the pitch of 1000 nm was inspected using TSOM with two illumination polarizations.

Figure 8 presents the simulation results of the Au line array with 50-nm height. The TSOM images were simulated with a 0.8 INA, 0.8 CNA, and 555-nm illumination wavelength, which was identical to the experimental conditions. Figure 9 shows the experimental results of the Au line arrays with 50-nm and 70-nm height.

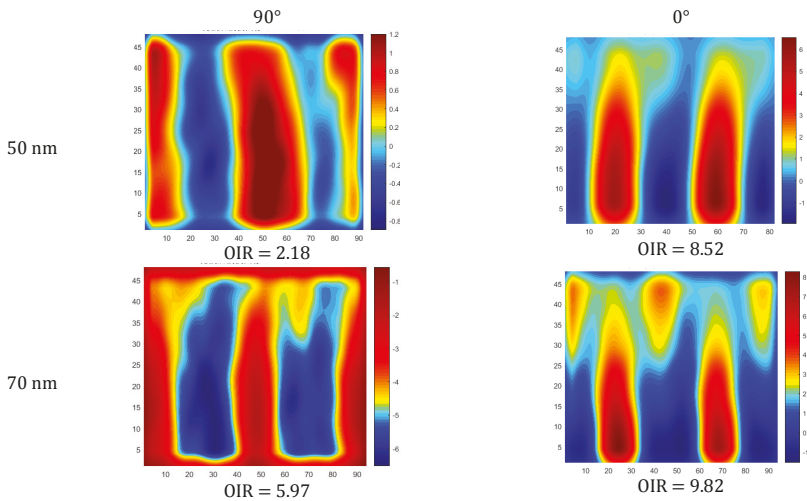
The microscope system noise is dependent on the stray light, blurs on the wafer, aberration of the microscope, and imprecision of the motion and location component of the stage [2,15]. The TSOM images of the system noise with two illumination polarizations were obtained by subtracting two background TSOM images corresponding to line heights of 70 nm and 50 nm. The TSOM images are demonstrated in Figure 10.



**Figure 7.** SEM of Au line array. In this figure, the width of the Au line is 23.6 nm, and the pitch of this group is approximately 1000 nm.

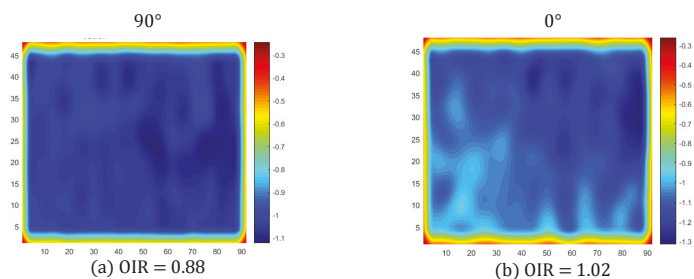


**Figure 8.** Simulated TSO images of Au line arrays with illumination polarizations of  $90^\circ$  (a) and  $0^\circ$  (b).



**Figure 9.** Experimental TSO images for the Au line arrays with illumination polarizations of  $90^\circ$  (the first column) and  $0^\circ$  (the second column). The line heights were 50 nm and 70 nm.





**Figure 10.** Experimental TSOM images of the system noise with 90° polarization (a) and 0° polarization (b).

Clearly, the intensity of signal and system noise are quantified by the OIR of the TSOM images of the target and noise, respectively. By dividing the OIR of the TSOM image of the target by that of the TSOM image of the noise, the signal-to-noise ratio (SNR) of TSOM is retained. The SNRs in the two illumination polarizations are presented in Table 4. According to the result, for the given target, the sensitivity and SNR apparently increase when the polarization is perpendicular to the Au lines.

**Table 4.** Signal-to-noise ratio (SNR) of each Au line array with different illumination polarizations.

Line Height	90° Polarization	0° Polarization
50 nm	SNR = 2.47	SNR = 8.45
70 nm	SNR = 6.47	SNR = 9.62

### 5. Analysis and Discussion

When light is scattered from the targets, its amplitude is changed. These changes depend on the sample material and shape, as well as the illumination polarization. As mentioned in Section 3.1, the structural features of nano-pit, nano-dot, and nano-bump were centrosymmetric. According to the analysis in Section 2, the function of their effective refractive index is also centrosymmetric, that is, in the polar coordinate

$$n(\rho, \theta) = n(\rho).$$

Therefore, the effective refractive indices of these three targets are constant with respect to the two illumination polarizations. For the silicon or Au lines, the function of the equivalent refractive index along the horizontal ( $X$ ) axis is inconsistent but periodic. From the simulation and experimental results, the TSOM image and its OIR of the centrosymmetric targets is similar regardless of the different illumination polarizations, but an apparent difference can be observed for center asymmetric targets, which supports our hypothesis.

From the data of the last three rows of Table 2 and the data of Table 3 (double silicon lines and dense line array), the relative differences between OIRs of TSOM and DTSOM increase as the number of lines increases or the pitch of the array decreases. This phenomenon can be explained through the effect of form birefringence, which leads to two effective refractive indices for two orthogonal polarizations. The form birefringence strongly depends on the sample structure even if the sample material is non-birefringent. As the number of lines increases or the pitch of the array decreases, the scattered light intensity of the perpendicular polarization varies more significantly across the line and space than that of the parallel polarization, which is eventually illustrated through the relative difference between OIRs with the two illumination polarizations.

### 6. Conclusions

In this study, the complete image-forming procedure of TSOM was modeled, and it included Köhler illumination in the TSOM, the scattered electromagnetic field reflected from a complex target,

and the image-scanning procedure. The FDTD method was performed to simulate the TSOM images of four targets with representative structural features, and the OIRs with two illumination polarizations were analyzed. The simulation and experimental results demonstrate that, for centrosymmetric targets, neither of the polarizations is effective for improving inspection sensitivity and SNR. On the contrary, for the asymmetric targets such as silicon and Au line arrays, the sensitivity and SNR apparently improved when the illumination polarization light was perpendicular to the lines. This result is valuable in TSOM applications for micro- and nano-inspection by saving time and improving measuring efficiency and precision.

**Author Contributions:** Conceptualization, Y.Q. and R.P. Methodology, J.J. Software, R.P. Validation, Y.Q., R.P., and J.H. Formal analysis, R.P. Investigation, R.P. Resources, H.P. and J.N. Data curation, R.P. Writing—original draft preparation, Y.Q. and R.P. Writing—review and editing, Y.Q. Visualization, R.P. Supervision, Y.Q. Project administration, Y.Q. Funding acquisition, Y.Q.

**Funding:** This research was funded by the Natural National Science Foundation of China (NSFC) No. 51675033.

**Conflicts of Interest:** The authors declare no conflicts of interest.

## References

- Ravikiran, A.; Ronald, G.D.; John, A.K.; James, E.P.; András, E.V.; Benjamin, B.; Erik, N.; Andrew, R. TSOM method for semiconductor metrology. *Metrol. Insp. Process Control Microlithogr.* **2011**. [[CrossRef](#)]
- Ravikiran, A. Noise analysis for through-focus scanning optical microscopy. *Opt. Lett.* **2016**, *41*, 745–748.
- Jun, H.L.; Jun, H.P.; Dohwan, J.; Eun, J.S.; Chris, P. Tip/tilt-compensated through-focus scanning optical microscopy. *Opt. Metrol. Insp. Ind. Appl. IV* **2016**. [[CrossRef](#)]
- An-Shun, L.; Yi-Sha, K.; Nigel, S. Through-focus technique for overlay metrology. *Metrol. Insp. Process Control Microlithogr. XXI* **2007**. [[CrossRef](#)]
- Timothy, F.C. Defect metrology challenges at the 11-nm node and beyond. *Metrol. Insp. Process Control Microlithogr. XXIV* **2010**. [[CrossRef](#)]
- Golani, O.; Dolev, I.; Pond, J.; Niegemann, J. Simulating semiconductor structures for next-generation optical inspection technologies. *Opt. Eng.* **2016**, *55*, 025102. [[CrossRef](#)]
- Ryoichi, H.; Nobutaka, K.; Hideaki, H.; Kenichi, T.; Masatoshi, H.; Hiroyuki, S. Study of EUV mask inspection technique using DUV light source for hp22nm and beyond. *Photomask Technol.* **2010**. [[CrossRef](#)]
- Bryan, M.B.; Richard, Q.; Yeung-Joon, S.; Hui, Z.; Richard, M.S. Optical illumination optimization for patterned defect inspection. *Metrol. Insp. Process Control Microlithogr. XXV* **2011**. [[CrossRef](#)]
- Stover, J.C. *Optical Scattering Measurement and Analysis*, 3rd ed.; SPIE Press: Bellingham, WA, USA, 2012; pp. 91–108.
- Ravikiran, A.; Hyeonggon, K. Parameter optimization for through-focus scanning optical microscopy. *Opt. Express* **2016**, *24*, 14915–14924.
- Ravikiran, A.; Peter, W.; John, A.K.; Bunday, B.; Victor, H.V. Feasibility study on 3-D shape analysis of high-aspect-ratio features using through-focus scanning optical microscopy. *Opt. Express* **2016**, *24*, 16574–16585.
- Jérôme, V.A.; Flavien, H.; Adam, C.; James, P. Uniform illumination and rigorous electromagnetic simulations applied to CMOS image sensors. *Opt. Express* **2007**, *15*, 5494–5503.
- Yeung, J.S.; Brian, M.B.; Lowell, H.; Richard, M.S.; Ravikiran, A.; Michael, T.S. Köhler illumination in high-resolution optical metrology. *Metrol. Insp. Process Control Microlithogr. XX* **2006**. [[CrossRef](#)]
- Marx, B.; Emil, W. *Principles of Optics*, 7th ed.; Publishing House of Electronics Industry: Beijing, China, 2009; pp. 608–613.
- Ryabko, M.; Koptyaev, S.; Shcherbakov, A.; Lantsov, A.; Oh, S.Y. Motion-free all optical inspection system for nanoscale topology control. *Opt. Express* **2014**, *22*, 14958–14963. [[CrossRef](#)] [[PubMed](#)]



© 2018 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).

Article

# Design, Measurement and Shape Reconstruction of Soft Surgical Actuator Based on Fiber Bragg Gratings

Yanlin He <sup>1,2</sup>, Lianqing Zhu <sup>1,2,\*</sup>, Guangkai Sun <sup>1,2</sup>, Mingxin Yu <sup>1,2</sup> and Mingli Dong <sup>1,2,\*</sup>

<sup>1</sup> Beijing Engineering Research Center of Optoelectronic Information and Instruments, Beijing Information Science and Technology University, Beijing 100192, China; heyianlin@bit.edu.cn (Y.H.); Guangkai.sun@buaa.edu.cn (G.S.); mingxinbit@gmail.com (M.Y.)

<sup>2</sup> Bionic and Intelligent Equipment Lab, Beijing Information Science and Technology University, Beijing 100192, China

\* Correspondence: lianqingzhu18@gmail.com (L.Z.); dongml@bistu.edu.cn (M.D.); Tel.: +86-010-8417-6477 (L.Z.)

Received: 19 August 2018; Accepted: 20 September 2018; Published: 30 September 2018



**Abstract:** Soft actuators are the components responsible for organs and tissues adsorptive fixation in some surgical operations, but the lack of shape sensing and monitoring of a soft actuator greatly limits their application potential. Consequently, this paper proposes a real-time 3D shape reconstruction method of soft surgical actuator which has an embedded optical fiber with two Fiber Bragg Grating (FBG) sensors. First, the design principle and the sensing of the soft actuator based on FBG sensors are analyzed, and the fabrication process of soft actuator which has an embedded optical fiber with two FBG sensors is described. Next, the calibration of the FBG sensors is conducted. Based on curvatures and curve fitting functions, the strategy of 3D shapes reconstruction of the soft actuator is presented. Finally, some bending experiments of the soft actuator are carried out, and the 3D shapes of the soft actuator at different bending states are reconstructed. This well reconstructed 3D shape of a soft actuator demonstrates the effectiveness of the shape reconstruction method that is proposed in this paper, as well as the potential and increased applications of these structures for real soft surgical actuators.

**Keywords:** micro fiber sensor; shape reconstruction; soft surgical robot; pneumatic actuator; modeling

## 1. Introduction

In recent years, soft actuators are widely used to implement surgical tasks due to its large-scale deformation and compliant adaptation to meet the complex elastic structures of the human organs and tissues [1–3]. The soft actuators can deform to absorb the energy generated by a variety of stimulation, including electrical charges, hydraulic actuating, pneumatic actuating and chemical reactions, etc. In particular, considering the advantages of easy fabrication, low material cost and lightweight, etc., the pneumatic actuating soft actuators are promising for surgical applications [4], and the use of soft manipulators in surgical operations could overcome the limitations of the current rigid surgical system. Consequently, compared with some traditional rigid and hard continuum surgical manipulators, the soft biological materials could considerably reduce the harm, which is caused by the robotic manipulators, and the soft surgical manipulators are inherently compliant and intrinsically safe in the interaction with humans.

While current researchers have emphasized the promising potential of soft manipulators in surgical applications, the lack of complicate shape measurement and monitoring for soft actuator are greatly limiting their practical application. In order to reliably and widely use these actuators in soft surgical manipulator, mastering the relationship between shape measurement and position

control of the actuator is indispensable. Considering the flexible bodies of the soft manipulator, some conventional rigid sensors, like encoders and strain gauges [5,6] have problems of complex cable networks, electromagnetic and noise signal interference, etc., and the rigid sensor will damage human tissue when the manipulator interacts with the human organs surround it [7–9]. The visual-based sensing systems are also inapplicable due to the narrow spaces and various obstacles in human body. At present, the embedded flexible sensing devices and methods that can be applied to the shape measurement and position control of soft surgical actuators are desired.

Different from above-mentioned sensors, the optical fiber sensors have advantages of small size, light weight, fast response, small transmission loss, not be affected with electromagnetic interference and easy to be buried into the interior of the material [10,11]. More importantly, the optical fiber sensor is harmless when in contact with human tissues and organs [12–15], and the optical fiber sensor can be embedded into the flexible body of the soft manipulator to realize its morphological sensing. Consequently, these distinctive characteristics of the optical fiber sensor make it more advantageous, and the prospects for soft surgical robotics and biomedical applications; and several types of research of optical fiber sensors have been investigated in recent years. In [12], the fiber sensors were embedded into a piece of garment, and the motion performance of the patient was evaluated by the angle of joint, which is monitored by the embedded fiber sensors; this sensing garment is comfortable for patients and has good response to the flexion motion of the elbow. As proposed in [13], the bending state of a manipulator was measured by an optical fiber sensing system and the mechanism of this sensing system based on a laser power modulation. Although this sensing system has good accuracy in measuring bending angles, bending radius and orientation of the flexible segment, its 3D shape reconstruction and visualization monitoring are not realized. In addition, this sensing system adopts three optical fibers to interpolate the bending radius and the elongation of the manipulator, and this undoubtedly increases the complexity of the theoretical models of the manipulator. In [12], some micro-optical sensors were integrated into soft robot arms to form their pose sensing system; this system couples the optical fibers with some passive cables mechanically in a basal unit that has sensing function, and a distance modulation array is integrated with the basal sensing unit. The pose of the soft robot arms can be calculated by the mechanical deformation to the voltage variation of the sensing signal. This sensing system can be used to measure the length and bending angle of soft robot arms, their 3D shape reconstruction and monitoring are not attained. Zhao et al. [15] designed a hardware and control method for a soft orthosis, the initial quantification of its force augmenting capabilities was performed, and its closed-loop control was implemented through the feedback from the optical-fiber sensors which was embedded into the actuator.

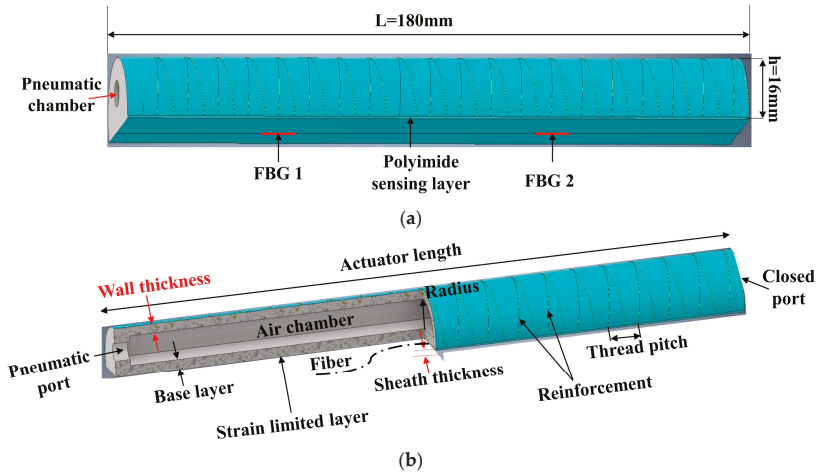
However, few studies have been conducted in real-time; 3D shape reconstruction and visualization monitoring of soft surgical actuator which are embedded a in single optical fiber with Fiber Bragg Grating (FBG) sensors. In this paper, to understand real-time shape monitoring of the soft actuator, a 3D shape reconstruction method based on two FBG sensors is proposed. First, an optical fiber with two FBG sensors is embedded into the soft actuator; then the calibration of FBG sensors is conducted, the spectrum and wavelength shift caused by the bending of the soft actuator is also measured; and finally, the 3D shape of the soft actuator is reconstructed based on the interpolation and curve fitting functions.

The remainder of this paper is organized as follows. In Section 2, we describe the mechanical design and modeling of the soft silicone actuator. In Section 3, we introduce the sensing theory and shape measurement method of FBG sensors. The experimental setup and testing of the soft actuator are elaborate in Section 4. Section 5 concludes the paper.

## 2. Design and Modeling of Soft Actuator with Embedded FBG Sensors

### 2.1. Structure Design and Analytical Modeling

Through adjusting some dimension parameters of the soft actuator, including its length, the diameter of the hemi-circle chamber, the pneumatic chamber's wall thickness, the fiber-winding angle and orientation, etc., The design details of the soft actuator are presents in Figure 1 [16].



**Figure 1.** Design of the proposed Fiber Bragg Grating (FBG) soft pneumatic actuator. (a) FBG-based polyimide sensing layer and its location. (b) Geometrical parameters and design variables of soft fiber-reinforced actuator.

The dimensions of the hemi-circular shape are achieved (see Figure 2a) using the cross-sectional area of  $\pi a^2/2$ , and the wall thickness of the actuator is  $t = a/4$ . With an input air pressure of  $P_{in}$ , the bending torques ( $M_a$ ) of internal air pressure relative to the closed port of the soft actuator are expressed as [17,18]:

$$M_a^{HC} = 0.34a^3P_{in} \quad (1)$$

Considering the characteristics of the hyper-elastic material of silicone and the dimension of the soft actuator; the relationship between the input pressure and the bending angle were analyzed. The variables in the model were actual actuator dimensions and material properties, that could be either measured or obtained from calibrations. From the model of incompressible Neo-Hookean (NH) material [4], the strain energy of the soft actuator can be written as:

$$W = \frac{\mu}{2}(I_1 - 3) = \frac{\mu}{2}(\lambda_1^2 + \lambda_2^2 + \lambda_3^2 - 3) \quad (2)$$

where  $\mu$  represents the initial shear modulus of the silicone,  $I_1$  represents the first constant of axial, circumferential and radial principal.  $\lambda_1$  represents the primary stretch in the axial direction of the actuator,  $\lambda_2 (=1)$  represents the circumferential strain, and it was ignored due to the fiber reinforcement constraint, and  $\lambda_1\lambda_2\lambda_3 = 1$  due to the incompressibility of the material, then it was obtained that:

$$\lambda_1 = \lambda, \quad \lambda_2 = 1, \quad \lambda_3 = \frac{1}{\lambda} \quad (3)$$

When supplying the pressure ( $P_1 > P_{atm}$ ) into the air chamber, the top wall of the actuator extended and the non-retractable layer restricted its bottom layer, the actuator bent with a radius  $R$  and angle  $\theta$  toward the bottom layer (see Figure 2). In addition, the fiber-reinforced structure was regarded as a

rigid constraint to the soft actuator, and the actuator was simplified as a homogeneous incompressible NH material [4] whose initial shear modulus is  $\bar{\mu}$ . In the model of the actuator, some dynamics pressurization effects were ignored, and the actuator had an identical bending curvature.

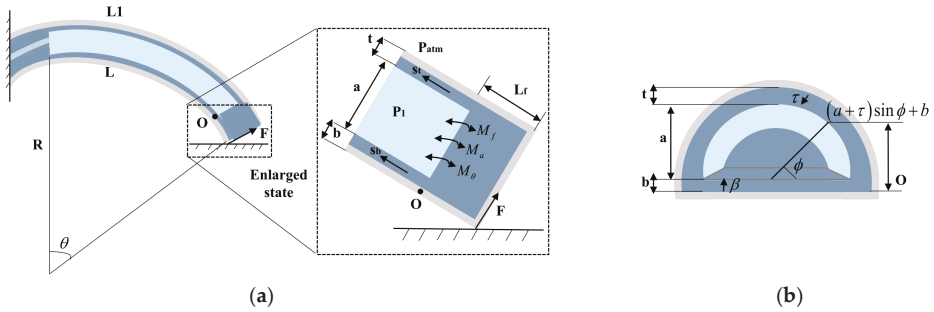


Figure 2. Different view of the soft actuator and its distal tip in bending state. (a) Side view. (b) Cross-sectional view.

The nominal main strain  $s_i$  can be defined by  $W, \lambda_i$ . Where  $p$ , the Lagrange multiplier is:

$$s_1 = \frac{\partial W}{\partial \lambda_1} - \frac{p}{\lambda_1} = \bar{\mu} \left( \lambda - \frac{1}{\lambda^3} \right) \tag{4}$$

$$s_2 = \frac{\partial W}{\partial \lambda_2} - \frac{p}{\lambda_2} = \bar{\mu} \left( 1 - \frac{1}{\lambda^2} \right) \tag{5}$$

$$s_3 = \frac{\partial W}{\partial \lambda_3} - \frac{p}{\lambda_3} = 0 \tag{6}$$

Within the range of stretches considered in soft actuator application ( $1 \leq \lambda < 1.5$ ), the axial stretch  $s_1$  this paper only considered was denoted as  $s$ .

The bending moment was induced by the internal strain of the soft actuator, the bending torque and the combined moment at each bending state can be expressed as:

$$M_a = 2(P_1 - P_{atm}) \int_0^{\frac{\pi}{2}} (a \sin v + b) a^2 \cos^2 v dv = \frac{4a^3 + 3\pi a^2 b}{6} (P_1 - P_{atm}) \tag{7}$$

$$M_\theta = \int_0^b 2s_\beta (a + t) L \beta d\beta + 2 \int_0^t \int_0^{\frac{\pi}{2}} s_{\tau,\phi} \left( (a + \tau)^2 \sin \phi + b(a + \tau) L d\phi \right) d\tau \tag{8}$$

where  $M_a$  is the strain of the soft actuator,  $M_\theta$  is the combined moment of the strain.

Considering the bending moment produced by the internal strain of the soft actuator, the bending torque  $M_a$  equal to the combined moment of the strain  $M_\theta$ , and the relationship between input pressure ( $P_{in}$ ) and the actuator's bending angle ( $\theta$ ) can be written as:

$$P_{in} = P_1 - P_{atm} = \frac{6M_\theta(\theta)}{4a^3 + 3\pi a^2 b} \tag{9}$$

where  $a$  is the radius of the actuator's air chamber,  $b$  is the bottom thickness of the actuator.

## 2.2. The Fabrication and Prototype of the Soft Actuator

In this section, we describe our procedure for molding the soft actuator. The design of the FBG-based optical fiber is embedded into the soft actuator through a polyimide tape, which acted as the strain limited layer; the bending behavior and deformed shape measurement of the soft actuator was realized by the FBG sensor. The soft actuator mainly includes five parts: A hemi-circle elastomeric air chamber which including the caps at the distal and proximal ends; an optical fiber with two FBG sensors is glued on the bottom surface of the silicone layer; some fiber windings are wrapped around the chamber of the actuator; and the whole actuator is encapsulated by a soft silicone sheath [11]. The radial expansion was limited and the linear extension was promoted by the reinforcement of the circumferential fibers, the linear elongation of the bottom surface of the soft actuator was restricted by the strain limited layer, so the part of the soft actuator extends while the strain limited layer restricts expansion along one surface, and thus generating a bending motion as the pressure was inputted into the soft actuator.

The fabrication of a soft actuator with embedded FBG sensors is performed through several steps. First, equal amount of Ecoflex-0050 A and B are mixed and poured into the mold #1 to fabricate the half round soft actuator with a thickness of 4 mm, and the silicone actuator is cured for 4 hours under indoor temperature. An Objet Connex 500 printer (Objet500 Connex3, Stratasys, Eden Prairie, MN, USA, 2016) is used to printing the two (#1, #2) molds of the soft actuator, as depicts in Figure 3. Second, the optical fiber with two FBG sensors is put on the bottom surface of the silicone layer, then a polyimide tape is glued on the bottom surface, and thus the optical fiber is fixed to the bottom of the silicone layer. Third, a single Kevlar fiber is used as fiber reinforcement and wrapped around the actuator body in a double helix pattern, and the actuator body is then encapsulated by a 1.0 mm thick silicone layer putting its entire assembly into the mold #2. Finally, the mold (#2) and the half steel rod are extracted from the soft actuator, then some additional silicone is poured on the top and bottom of the actuator in order to enclose the chamber and connect the pipe of pneumatic actuation.

The prototype of the soft actuator with different views are presents in Figure 4, as we can see from Figure 4, an optical fiber with two FBG sensors are embedded into the soft actuator through a polyimide tape, and the optical fiber with FBG sensors are well protected by the polyimide thin film layer. Thus, the FBG-based polyimide sensing layer is formed, and it is embedded into the soft actuator as the strain restricted layer, which served as a primary part to realize bending and shape measurement, and the FBG sensors are stretched with the bending motion of the soft actuator.

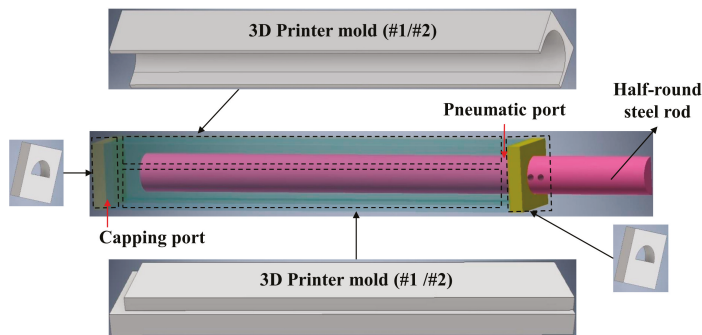


Figure 3. 3D printed mold of the soft actuator.

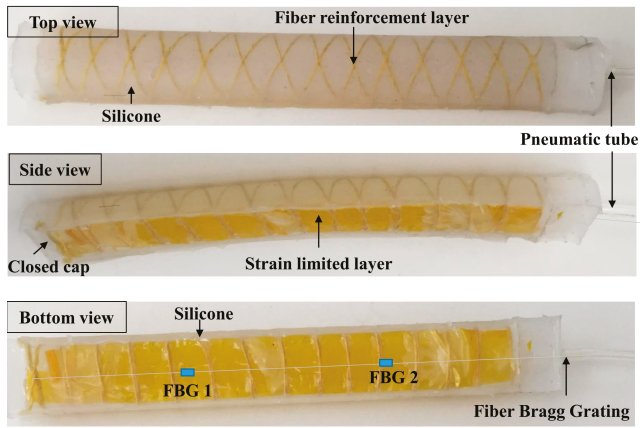


Figure 4. 3D printed mold of the soft actuator.

### 3. Principle of Sensing and Measurement of Embedded FBGs

According to the sensing principle of embedded FBG sensor, the relationship between the reflected wavelength and the grating period of FBG sensor can be expressed as [5,10]:

$$\lambda_B = 2n_{eff} \cdot \Lambda \tag{10}$$

where  $n_{eff}$  represents the effective index of the fiber’s refractive,  $\Lambda$  represents the grating period.

The grating period of FBG sensor will be changed due to the changes in the strain of the grating and the temperature, and then the reflected wavelength will be changed. The relation between the strain, temperature change, and the wavelength shift is governed by [6,7]:

$$\Delta\lambda_B/\lambda_B = (1 - P_\epsilon) \cdot \epsilon + (\alpha_T + \zeta) \cdot \Delta T \tag{11}$$

where  $\Delta\lambda_B$  is the change of wavelength,  $\epsilon$  is the strain of soft actuator,  $P_\epsilon$  is a constant value calculated from the photo-elastic coefficient of the fiber material,  $\alpha_T$  is the grating thermal coefficient of expansion,  $\zeta$  is the grating thermo-optical coefficient,  $\Delta T$  is the change of temperature.

Considering the soft surgical actuators are commonly used at room temperature with minor variation, so the influence of the temperature is ignored in this paper. Then Equation (11) can be written as Equation (12). The strain of soft actuator can be obtained from the wavelength shift of the FBGs and the photo elastic coefficient of the optical fiber [19–22];

$$\Delta\lambda_B/\lambda_B = (1 - P_\epsilon) \cdot \epsilon \tag{12}$$

In this study, two FBG sensors were integrated into the polyimide strain limited layer which was embedded into the soft actuator, the relative position and bending deformation of the FBG sensors in a microstructure of the soft actuator under an actuating force is schematically shown in Figure 5. The length and thickness of the microstructure is assumed as  $L$  and  $h$ , as Figure 5a shows.

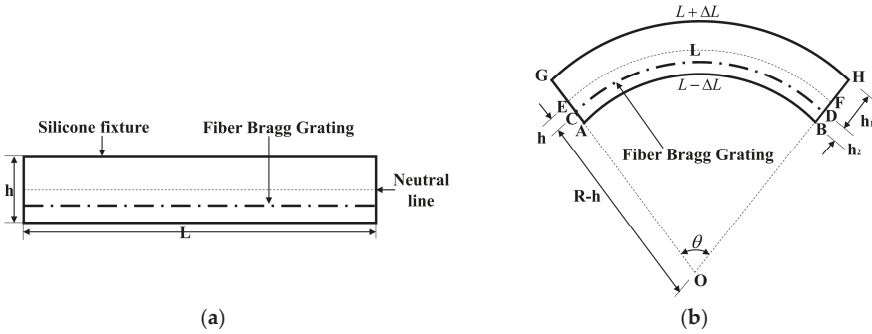
The pure bending model is an extensively used model representing the bending moment [10]. In this paper, an optical fiber with two FBG sensors embedded in an offset position from the neutral line, which is depicted in Figure 5b, is the dashed line E-F. The distance from the FBG sensor to the outer arc of the soft silicone sheet is  $h_1$ , and the distance from the FBG sensor to the inner arc of the soft silicone sensing sheet is  $h_2$ . When the optical fiber sensor is bending, as shown in Figure 5b,



the relationship between the bending radius, bending angle and the length of the neutral line can be expressed as:

$$L_{EF} = R \cdot \theta \tag{13}$$

where  $L_{EF}$  represents the constant value of the length of the polyimide tape in neutral line,  $R$  represents the radius of the neutral line,  $\theta$  represents the central angle of the neutral line.



**Figure 5.** Bending mechanism of the FBG-based optical fiber sensor. (a) Free-state without bending. (b) Bending situation.

As the pneumatic actuating forces are applied on the soft actuator, the inner side of the micro structure is pressed, shortening the length, and the outer side of the microstructure is pulled and extending the length. Under the bending condition, the length variation of both inner and outer side is assumed as the same, and the length of the neutral layer does not change. In this paper, the FBG sensors and the polyimide sheet are considered as undergoing the same bending, due to the optical fiber being very thin and glued together with the polyimide tape. The FBG sensor is compressed during bending, so the offset location ( $x$ ) of the FBG sensor leads to a slight change in the bending radius and is expressed as  $R - x$ , the length of FBG ( $L_{AB}$ ) under an applied curvature can be written as:

$$L - \Delta L = L_{AB} = (R - x) \cdot \theta = (R - (h_1 + h_2)/2) \cdot \theta \tag{14}$$

$$L + \Delta L = L_{GH} = (R + x) \cdot \theta = (R + (h_1 + h_2)/2) \cdot \theta \tag{15}$$

$$x = h_1 - (h_1 + h_2)/2 \tag{16}$$

where  $x$  represents the distance from the FBG to the neutral line. From Equations (14) and (15), the change of FBG length caused from the bending was obtained, and the grating period is changed proportionally to the bending angle  $\theta$ , the bending curvature of the microstructure can be written as [10]:

$$C = \frac{1}{R} = \frac{2 \Delta L}{h L} = \frac{2}{h} \varepsilon = \frac{\Delta \lambda_B}{\lambda_B (1 - R_\varepsilon) \cdot h} \tag{17}$$

where  $C$  is the bending curvature of the microstructure,  $R$  represents the curvature radius,  $\Delta L$  represents the length variation of both the inner and outer side of the microstructure,  $\varepsilon$  is the strain of the microstructure. For any FBG sensors,  $\lambda_B$ ,  $R_\varepsilon$ ,  $h$  are the constant, so the bending curvature and wavelength shift has a linear relationship, and this characteristic could realize the 3D shape reconstruction of the microstructure.

The shape reconstruction of the soft actuator includes curvature acquisition, curvature continuum, curve reconstruction and shape reconstruction, etc. Due to the curvature data being discrete and obtained from the FBG sensors, the curvature is serialized continuously through the linear interpolation and the fitted algorithm; then the geometrical structure of the soft actuator can be reconstructed. By using the idea of differential calculus, a line with fixed length can be divided into several uniform

sections, and the length of these sections expressed as  $L$ , namely  $O_1O_2 = O_2O_3 = L$ , and then become some arcs with the length of  $L$  when the microstructure is bending, and the center angle of the circular arc is determined by the length  $L$  and the curvature  $C$ . Assuming that the coordinate of  $O_2$  is  $(0, 0)$ , the coordinate of  $O_3$  can be calculated as  $(\sin\theta_2/C_2, (1 - \cos\theta_2)/C_2)$ , from the length  $L$  and curvature  $C$ , the  $\theta_2$  is equal to  $L \cdot C_2$ ; thus this algorithm can be applied for 3D shape reconstruction of the soft actuator [23–29].

The sensing surface formed by some sensing points can be reconstructed based on the motion coordinate system, as illustrated in Figure 6. Assuming that the original curved surface along the direction of  $(O_1, Z_1)$ , and could bend from the direction of curves  $O_1, O_2, O_3$ , the point  $O_2$  could be obtained from translation and rotation of the coordinate  $O_2^1$ . In the fixed length of arc, the coordinate system is set up in the tangent direction of the curve with points  $O_2$  and  $O_3$  respectively, and for the coordinate  $X_1-Z_1$  and  $X_2-Z_2$ , the direction and size of vector  $O_1O_2$  is invariable. The rotation angle  $\theta_1$  can be calculated from the curvature  $C_1$  (the coordinate of  $O_1$ ), and the rotation angle  $\theta_2$  can be calculated from the curvature  $C_2$  (the coordinate of  $O_2$ ).

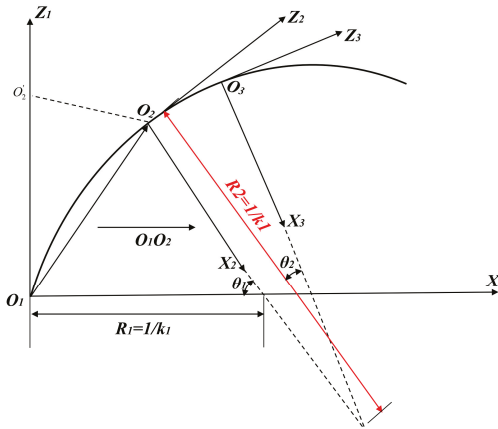


Figure 6. The algorithm of 3D shape reconstruction.

The relationship between each coordinate can be established by the rotation matrix, the ratio angle of the whole plane is fitted through the change of the curvature; thus the reconstruction and visualization of the soft actuator can be realized, and the rotation matrix can be expressed as:

$$\begin{bmatrix} X_2 \\ Z_2 \end{bmatrix} = \begin{bmatrix} X_1 \\ Z_1 \end{bmatrix} = \begin{bmatrix} \cos \theta_1 & -\sin \theta_1 \\ \sin \theta_1 & \cos \theta_1 \end{bmatrix} \tag{18}$$

Consequently, according to the curvature obtained from each sensing point, the 3D shape of the soft actuator is well reconstructed, and the real-time interaction surface with the soft actuator can be realized by using the multi-resolution mesh modeling tool. As introduced in [30–32], the 3D shape of the soft actuator can be obtained from the managing meshes with different refinement levels [30–32].

#### 4. Experiment Testing and Discussions

##### 4.1. Pressure and Bending Angle

In order to control the bending angle of the soft actuator, some experiments were carried out to evaluate its relationship between input air pressure and bending angle. The input pressure was controlled by a pneumatic valve in the experiment, and some experimental results were obtained

with the pressure of 0, 10, 20, 50, 80 kPa, and each experiment was repeated ten times. An AIRTAC VALVE (GPR30008L, AirTac, Taiwan, 2016) model was used to input pressure in the experiment; the received optical signals are translated into digital signals and extracted by the computer. A LabVIEW (LabVIEW 2014) program is developed to store the data and reconstruct the bending shape of the soft actuator [18–20]. In some robotic applications, the range of the soft surgical manipulation motion is more important than its actuation velocity, so the velocity of actuation is seldom considered. Consequently, for the sake of avoiding some dynamic oscillations, the bending motion of the soft actuator generated at a slow pressure rate of 0.2 Hz. Figure 7 shows some snapshots of the bending movement of the soft actuator, and an improved track learning detection algorithm was applied to reconstruct the motion trajectory of the soft actuator.

The average bending angle of the soft actuator was calculated from the same input pressure, as depicted in Figure 8. The theoretical and experimental results of the bending trajectories of the soft actuator are presents in Figure 9.

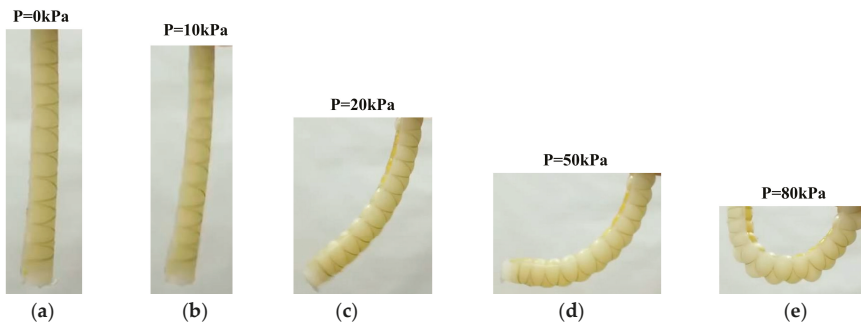


Figure 7. Experiment view of the actuator with different input pressure. (a) Bending state. (b) State 2. (c) State 3. (d) State 4. (e) State 5.

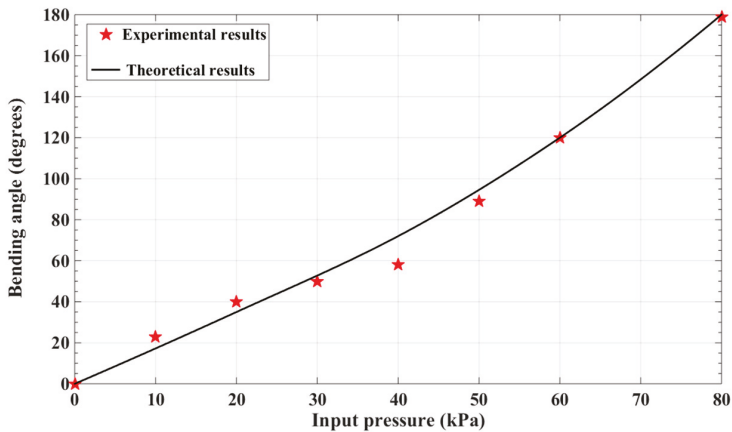


Figure 8. Input pressure against bending angle results of soft actuator.

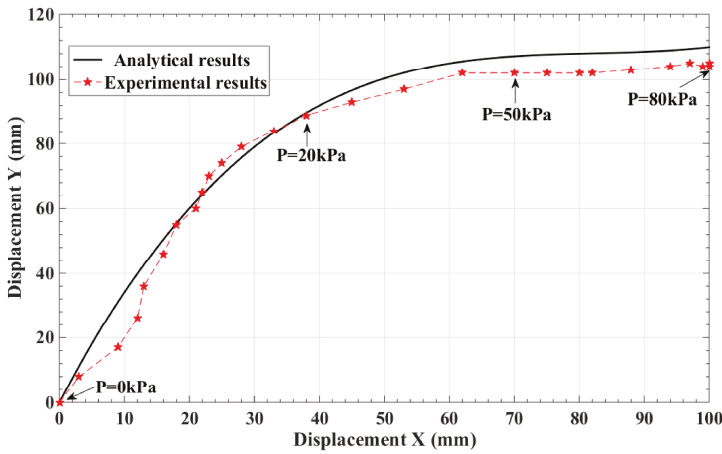


Figure 9. Motion trajectory of the actuator with different input pressure.

As depicted in Figure 8, due to the radial bulging effects of the base layer and extension of the soft silicone actuator, there exists some minor error between the analytical and experiment results, and a repeatable nonlinear pattern of results possibly caused by the gravity of the actuators, and the center of gravity of soft actuator is changing in different bending state. As depicted in Figure 9, although the actual motion trajectory of the soft actuator is not smooth and has some deviations, there was only a slight deviation between the theoretical and experimental results, so the design validity and effectiveness of the soft actuator was demonstrated.

#### 4.2. Shape Reconstruction of the Soft Actuator

In order to demonstrate the effectiveness of the shape reconstruction method proposed in this paper, some bending experiments of the soft actuator based on embedded FBGs were carried out. The experimental setup is illustrated in Figure 10. In the experiment, the original center wavelengths of two FBGs (FBG1, FBG2) are 1544 nm and 1546 nm respectively. The peak reflectivity of the FBGs are 70% and the response time of the FBGs is 70 ms. Two FBG sensors are located in the intermediate axis of the soft actuator, and their relative positions as illustrated in Figure 4, Section 2.

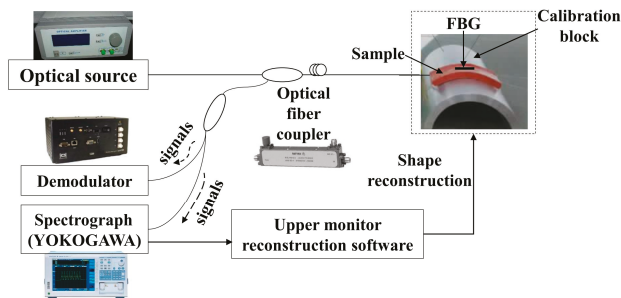


Figure 10. Experimental setup of the soft actuator with embedded FBGs.

As the bending curvature of the soft actuator varies with  $0\text{ m}^{-1}$ ,  $5\text{ m}^{-1}$ ,  $10\text{ m}^{-1}$ ,  $15\text{ m}^{-1}$ ,  $20\text{ m}^{-1}$  and  $25\text{ m}^{-1}$ , the wavelength and power of the FBG sensors are not the same as the original values. Since the grating characteristics of FBG sensors are varied from the strain loads of polyimide tape as it

bends, the power peaks of the FBGs can be identified accurately. As depicted in Figure 11, the power peaks of the two FBGs were around 1544.498 nm and 1546.611 nm at bending state 1, the curvature of the polyimide skin was increased when the soft actuator bent from state 1 to state 5, and the power peaks of the two FBGs was 1545.956 nm and 1547.541 nm at bending state 5. From Figure 11, it also indicates that the peak wavelength gradually shifts in the direction of a longer wavelength as the bottom surface curvature of the soft actuator increases from  $0\text{ m}^{-1}$  to  $25\text{ m}^{-1}$ , and the shift value of the two FBG sensors is about 1.458 nm and 0.93 nm, respectively. Consequently, the responses of the two FBGs are sensitive and instantaneous record the variation in power. The dynamic shape deformation of the soft actuator can be measured in real time, and the FBG sensors with different wavelengths can be distributed along a single optical fiber for multi-point shape sensing.

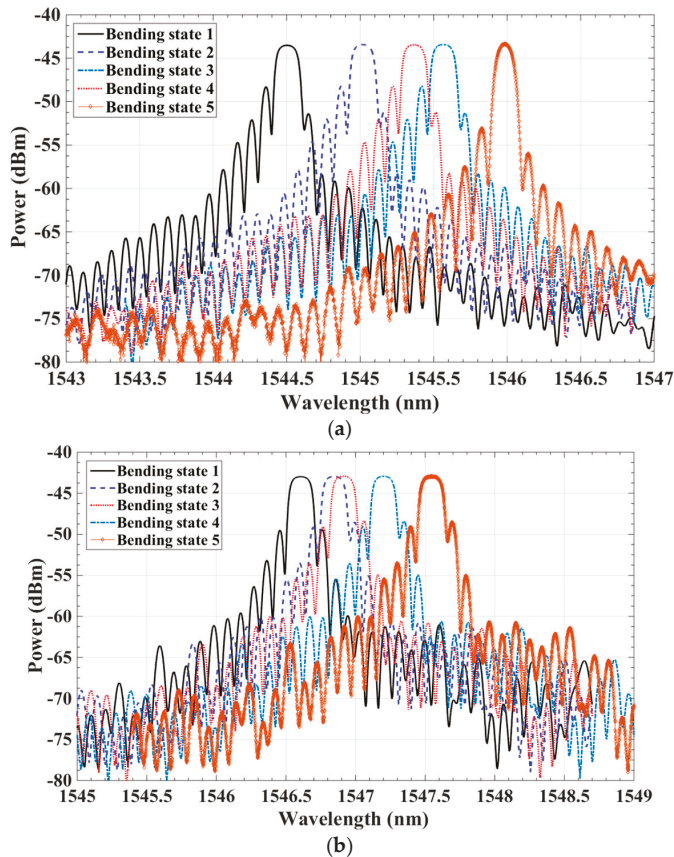


Figure 11. The wavelength shift of two FBG sensors at different bending states. (a) FBG 1. (b) FBG 2.

When the curvature of the FBGs increases from  $0\text{ m}^{-1}$  to  $25\text{ m}^{-1}$ , the relationship of the wavelength shift and the curvatures of the soft actuator are illustrated in Figure 12. All measurements have respective average values which were taken from six independent experiments. The bending curvature of the soft actuator at different bending states obtained from the wavelength shift of the FBG sensors, the interpolation algorithm, and the wavelength shift of the FBG sensors were different due to the bending produced strain loads at the different sensing positions. The results showed that the wavelength shift of the FBG sensors increased nonlinearly with the increasing bend in the curvature of the soft actuator. That is because the soft silicone material is extensible and its thickness is variable as

it sustains different driving loads and contact forces at different bending states. The thickness of the microstructure  $h$  at different sensing points varied and the strain loads induced by the deformation of the soft silicone actuator were different at these sensing points. According to Equation (17), when the bending curvature of the soft actuator increases from  $0 \text{ m}^{-1}$  to  $25 \text{ m}^{-1}$ , the wavelength shift and the bending curvature presents a nonlinear variation trend. The wavelength shift of the FBG sensors can be calculated by using the method of polynomial curve fitting, and the calibration equations can also be used to test the bending curvatures of the soft actuator.

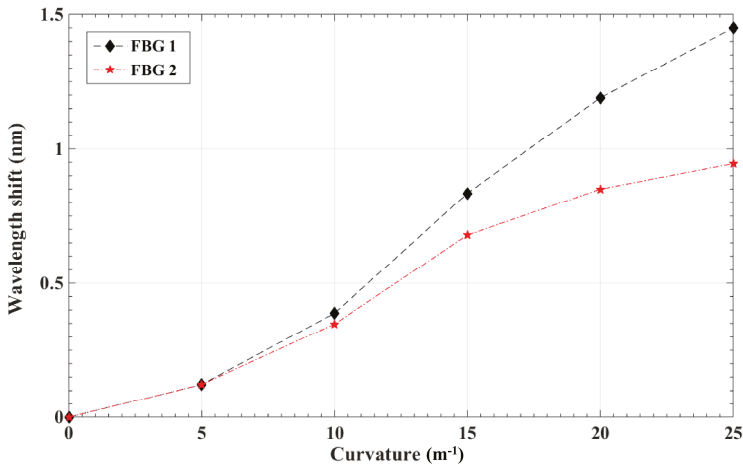


Figure 12. The wavelength shifts of two FBG sensors under different curvatures.

For the 3D shape reconstruction of the soft actuator at different bending states, the bending curvatures of the two FBG sensors were calculated from the wavelength shift. According to the data obtained from each sensing point, the 3D shape of the soft actuator is reconstructed through the interpolation and curve fitting functions. The 3D shapes reconstruction results of the soft actuator with different bending state numbered from 1 to 5 are presents in Figure 13. From the figure, the shape of the actuator is well reconstructed, and the spatial resolution determined by the density of FBG sensors in the sensing point is about 30 mm. Consequently, the feasibility and effectiveness of the 3D shape reconstruction method of the soft actuator based on embedded FBG sensors are demonstrated.

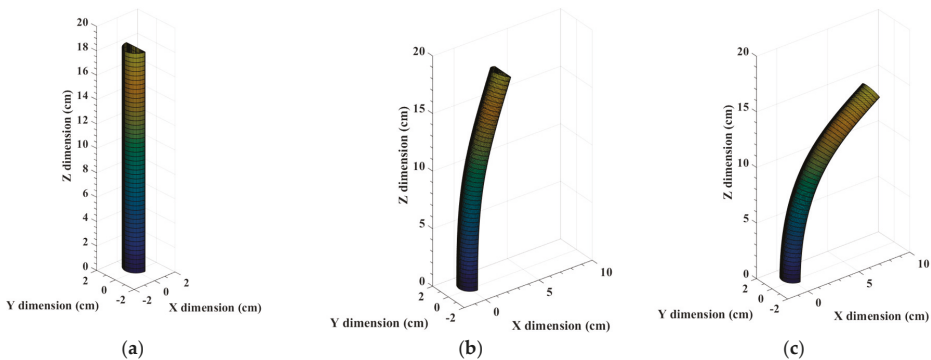
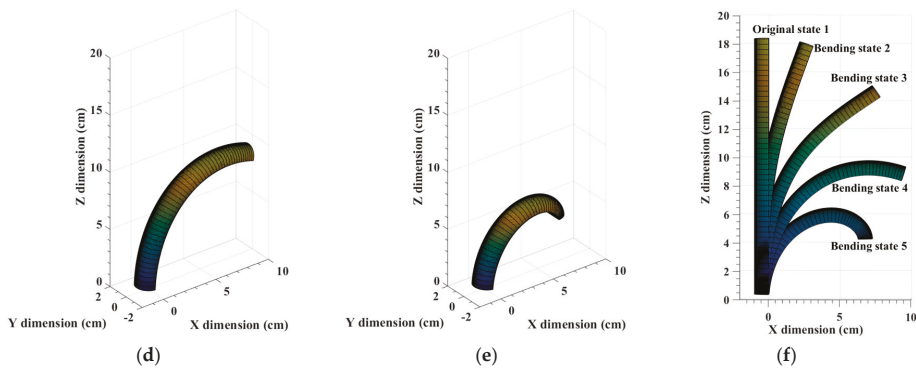


Figure 13. Cont.



**Figure 13.** Shape reconstruction of the soft actuator at different bending state. (a) Original state 1. (b) Bending state 2. (c) Bending state 3. (d) Bending state 4. (e) Bending state 5. (f) Bending state 1 to 5.

## 5. Conclusions

To implement real-time shape monitoring of soft surgical actuator, this paper proposes a 3D shape reconstruction method of a soft actuator that has an embedded optical fiber with two FBG sensors. First, the design and measurement principle of the soft actuator based on the FBG sensors are analyzed and described. An optical fiber with two FBG sensors are embedded into a soft actuator by a polyimide tape and encapsulated by the silicone layer. Second, the relation curve of the wavelength shifts of the FBG sensors and bending curvatures of the soft actuator are obtained, and which can be used to measure the bending curvatures of the soft actuator at different bending state. Then, the peaks of the reflected wavelengths and the variation of the reflection spectrums are acquired, the sensitive and instantaneous response of FBGs sensors are reflected. Finally, according to the data from the FBG sensing points, the 3D shape of the soft actuator is well reconstructed through the interpolation and curve fitting functions. This well reconstructed 3D shape of the soft actuator demonstrates the effectiveness of the shape reconstruction method that is proposed in this paper, as well as the potential and increased applications for real-time shape monitoring in the field of soft robotics and flexible biosensor monitoring. In the future, the dynamic behavior and closed-loop controlling of the soft actuator will be investigated, and the application of soft actuator in some surgical operations will be extended.

**Author Contributions:** L.Z. and G.S. conceived and designed the experiments; M.Y. performed the experiments; Y.H. analyzed the data and wrote the paper; M.D. contributed reagents/materials/analysis tools.

**Acknowledgments:** This work is supported by the Beijing Information Science and Technology University funded project (No. 1825008). This research project was also supported by the Changjiang Scholars and Innovative Research Team in University (No. IRT\_16R07), Importation and Development of High-Caliber Talents Project of Beijing Municipal Institutions (No. IDHT20170510), China Postdoctoral Science Foundation funded project (2018M631290).

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Ranzani, T.; Gerboni, G.; Cianchetti, M.; Menciassia, A. A bioinspired soft manipulator for minimally invasive surgery. *Bioinspir. Biomim.* **2015**, *10*, 035008. [[CrossRef](#)] [[PubMed](#)]
2. Zolfagharian, A.; Kouzani, A.Z.; Khoo, S.Y.; Moghadam, A.A.A.; Gibson, L.; Kaynak, A. Evolution of 3D printed soft actuators. *Sens. Actuators A Phys.* **2016**, *250*, 258–272. [[CrossRef](#)]
3. Luo, M.; Skorina, E.H.; Tao, W.; Chen, F.C.; Ozel, S.; Sun, Y.N.; Onal, C.D. Toward modular soft robotics: Proprioceptive curvature sensing and sliding mode control of soft bidirectional bending modules. *Soft Robot.* **2017**, *4*, 1–9. [[CrossRef](#)] [[PubMed](#)]

4. Polygerinos, P.; Wang, Z.; Overvelde, J.T.B.; Galloway, K.C.; Wood, R.J.; Bertoldi, K.; Walsh, C.J. Modeling of soft fiber-reinforced bending actuators. *IEEE Trans. Robot.* **2015**, *31*, 778–789. [[CrossRef](#)]
5. Sun, G.; Li, H.; Dong, M.; Lou, X.; Zhu, L. Optical fiber shape sensing of polyimide skin for a flexible morphing wing. *Appl. Opt.* **2018**, *56*, 9325. [[CrossRef](#)] [[PubMed](#)]
6. Nguyen, T.D.; Han, H.S.; Shin, H.Y.; Nguyen, C.T.; Phung, H.; Hoang, H.V.; Choi, H.R. Highly sensitive flexible proximity tactile array sensor by using carbon micro coils. *Sens. Actuators A Phys.* **2017**, *266*, 166–177. [[CrossRef](#)]
7. Ozel, S.; Keskin, N.A.; Khea, D.; Onal, C.D. A precise embedded curvature sensor module for soft-bodied robots. *Sens. Actuators A Phys.* **2015**, *236*, 349–356. [[CrossRef](#)]
8. Wang, H.S.; Zhang, R.X.; Chen, W.D.; Liang, X.W. Shape detection algorithm for soft manipulator based on Fiber Bragg Gratings. *IEEE/ASME Trans. Mechatron.* **2016**, *21*, 2977–2981. [[CrossRef](#)]
9. Payo, I.; Feliu, V.; Cortazar, O.D. Fibre Bragg grating (FBG) sensor system for highly flexible single-link robots. *Sens. Actuators A Phys.* **2009**, *150*, 24–39. [[CrossRef](#)]
10. Ge, J.; James, A.E.; Li, X.; Chen, Y.; Kwok, K.W.; Fok, M.P. Bidirectional soft silicone curvature sensor based on off-centered embedded fiber bragg grating. *IEEE Photonics Technol. Lett.* **2016**, *28*, 2237–2240. [[CrossRef](#)]
11. Kim, S.W.; Kang, W.R.; Jeong, M.S.; Lee, I.; Kwon, I.B. Deflection estimation of a wind turbine blade using FBG sensors embedded in the blade bonding line. *Smart Mater. Struct.* **2013**, *22*, 1–11. [[CrossRef](#)]
12. Silva, A.S.; Catarino, A.; Correia, M.V.; Frazão, O. Design and characterization of a wearable macro bending fibre optic sensor for human joint angle determination. *Opt. Eng.* **2013**, *52*, 992–999. [[CrossRef](#)]
13. Searle, T.C.; Althoefer, K.; Seneviratne, L.; Liu, H.B. An optical curvature sensor for flexible manipulator. In Proceedings of the 2013 IEEE International Conference on Robotics and Automation (ICRA), Karlsruhe, Germany, 6–10 May 2013; pp. 4415–4420. [[CrossRef](#)]
14. Sareh, S.; Noh, Y.; Li, M.; Ranzani, T.; Liu, H.B.; Althoefer, K. Macrobend optical sensing for pose measurement in soft robot arms. *Smart Mater. Struct.* **2015**, *24*, 125024. [[CrossRef](#)]
15. Zhao, H.; Jalving, J.; Huang, R.K.; Knepper, R.; Ruina, A.; Shepherd, R. A helping hand: Soft orthosis with integrated optical strain sensors and EMG control. *IEEE Robot. Autom. Mag.* **2016**, *23*, 55–64. [[CrossRef](#)]
16. Connolly, F.; Walsh, C.J.; Bertoldi, K. Automatic design of fiber-reinforced soft actuators for trajectory matching. *Proc. Natl. Acad. Sci. USA* **2017**, *114*, 51–56. [[CrossRef](#)] [[PubMed](#)]
17. Polygerinos, P.; Wang, Z.; Galloway, K.C. Soft robotic glove for combined assistance and at-home rehabilitation. *Robot. Autom. Syst.* **2015**, *73*, 135–143. [[CrossRef](#)]
18. Ranzani, T.; Cianchetti, M.; Gerboni, G.; De Falco, I.; Menciassi, A. A soft modular manipulator for minimally invasive surgery: Design and characterization of a single module. *IEEE Trans. Robot.* **2016**, *32*, 187–200. [[CrossRef](#)]
19. Laschi, C.; Mazzolai, B.; Cianchetti, M. Soft robotics: Technologies and systems pushing the boundaries of robot abilities. *Sci. Robot.* **2016**, *1*, 3690. [[CrossRef](#)]
20. Zhao, Z.; Soto, M.A.; Tang, M.; Thevenaz, L. Distributed shape sensing using Brillouin scattering in multi-core fibers. *Opt. Express* **2016**, *24*, 25211–25223. [[CrossRef](#)] [[PubMed](#)]
21. Guo, J.J.; Liu, X.Y.; Jiang, N.; Yetisen, A.K.; Yuk, H.; Yang, C.X.; Khademhosseini, A.; Zhao, X.H.; Yun, S.-H. Highly stretchable, strain sensing hydrogel optical fibers. *Adv. Mater.* **2016**, *28*, 10244–10249. [[CrossRef](#)] [[PubMed](#)]
22. Qiu, Y.; Shen, L.; Hu, W. Shape rebuilding and positioning method of search and rescue robot endoscope in ruin crack. *Chin. J. Sci. Instrum.* **2015**, *36*, 2782–2789.
23. Robertson, M.A.; Paik, J. New soft robots really suck: Vacuum-powered systems empower diverse capabilities. *Sci. Robot.* **2017**, *2*, 1–11. [[CrossRef](#)]
24. Cianchetti, M.; Ranzani, T.; Gerboni, G.; Nanayakkara, T.; Althoefer, K.; Dasgupta, P.; Menciassi, A. Soft robotics technologies to address shortcomings in today’s minimally invasive surgery: The stiff-flop approach. *Soft Robot.* **2014**, *1*, 122–131. [[CrossRef](#)]
25. Wang, Y.; Negahdaripour, S.; Aykin, M.D. Calibration and 3D reconstruction of underwater objects with non-single view projection model by structured light stereo imaging. *Appl. Opt.* **2016**, *55*, 6564–6575. [[CrossRef](#)] [[PubMed](#)]
26. Westbrook, P.S.; Kremp, T.; Feder, K.S.; Ko, W.; Monberg, E.M.; Wu, H.C.; Simoff, D.A.; Taunay, T.F.; Ortiz, R.M. Continuous multicore optical fiber grating arrays for distributed sensing applications. *J. Lightw. Technol.* **2017**, *35*, 1248–1252. [[CrossRef](#)]



27. Parent, F.; Loranger, S.; Mandal, K.K.; Iezzi, V.L.; Lapointe, I.; Boisvert, J.S.; Baiad, M.D.; Kadoury, S.; Kashyap, R. Enhancement of accuracy in shape sensing of surgical needles using optical frequency domain reflectometry in optical fibers. *Biomed. Opt. Express* **2017**, *8*, 2210–2221. [[CrossRef](#)] [[PubMed](#)]
28. Waltermann, C.; Doering, A.; Kohring, M.; Angelmahr, M.; Schade, W. Cladding waveguide gratings in standard single-mode fiber for 3D shape sensing. *Opt. Lett.* **2015**, *40*, 3109–3112. [[CrossRef](#)] [[PubMed](#)]
29. Feng, D.Y.; Zhou, W.J.; Qiao, X.G.; Albert, J. Compact optical fiber 3D shape sensor based on a pair of orthogonal tilted fiber bragg gratings. *Sci. Rep.* **2015**, 17415. [[CrossRef](#)] [[PubMed](#)]
30. Song, S.; Li, Z.; Yu, H.; Ren, H. Electromagnetic positioning for tip tracking and shape sensing of flexible robots. *IEEE Sens. J.* **2015**, *15*, 4565–4575. [[CrossRef](#)]
31. Quandt, B.M.; Fan, L.J.H.; Qian, J. Vibration shape testing of space sheet structure. *J. Shanghai Univ.* **2000**, *14*, 354–358.
32. Lee, K.K.C.; Mariampillai, A.; Haque, M.; Standish, B.A.; Yang, V.X.D.; Herman, P.R. Temperature-compensated fiber-optic 3D shape sensor based on femtosecond laser direct-written Bragg grating waveguides. *Opt. Express* **2013**, *21*, 24076–24086. [[CrossRef](#)] [[PubMed](#)]



© 2018 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).

Article

# Precision Manufacturing of Patterned Beryllium Bronze Leaf Springs via Chemical Etching

Rui-Jun Li <sup>1,\*</sup> , Peng-Yu Wang <sup>1</sup>, Dan-Dong Li <sup>1</sup>, Kuang-Chao Fan <sup>1,2</sup> , Fang-Fang Liu <sup>1</sup>,  
Li-Juan Chen <sup>1</sup> and Qiang-Xian Huang <sup>1</sup> 

<sup>1</sup> School of Instrument Science and Opto-electronics Engineering, Hefei University of Technology, Hefei 230009, China; 2016110023@mail.hfut.edu.cn (P.-Y.W.); lidandong@mail.hfut.edu.cn (D.-D.L.); fan@dlut.edu.cn (K.-C.F.); liuff@hfut.edu.cn (F.-F.L.); chenlj@hfut.edu.cn (L.-J.C.); huangqx@hfut.edu.cn (Q.-X.H.)

<sup>2</sup> School of Mechanical Engineering, Dalian University of Technology, Dalian 116024, China

\* Correspondence: rj-li@hfut.edu.cn; Tel.: +86-551-62901508

Received: 30 June 2018; Accepted: 22 August 2018; Published: 28 August 2018



**Featured Application:** Fabrication of high-precision patterned leaf springs or membranes for measuring equipment.

**Abstract:** Patterned leaf springs made of a beryllium bronze sheet are the key components of certain micro/nano contact probes. The accuracy of the probe is determined based on the precision of the formed pattern. However, a traditional manufacturing method using wire-electrode discharge machining (wire-EDM) is subject to poor tolerance at the sharp edges and corners. In addition, high energy consumption and costs are incurred for complex patterns. This paper presents a new chemical etching method for the manufacturing of a patterned leaf spring with high precision. Both the principle and process are introduced. Taguchi experiments were designed and conducted and the optimal process parameters were obtained based on the mean value and a variance analysis. Four V-shaped and some other complex patterned leaf springs were successfully fabricated. Comparison experiments concerning the characteristic parameters of the leaf spring were also conducted. The experimental results reveal that the patterned leaf springs manufactured through this method are much better than those achieved using wire-EDM. This manufacturing method can be used to fabricate different high-precision patterned leaf springs or membranes for coordinate measuring machines (CMM) probes and other measuring equipment.

**Keywords:** probe; leaf spring; chemical etching; beryllium bronze

## 1. Introduction

During the past two decades, a variety of micro-electromechanical systems (MEMS) have been fabricated through the rapid development of precision and ultra-precision machining technologies. Liu et al. (1999) described two microthermal shear-stress sensors, in which high aspect ratio cavities of  $(200 \times 250) \mu\text{m}$   $(2 \times 400) \mu\text{m}$  were fabricated using bulk silicon substrate etching and anodic bonding technology [1]. Alshehri et al. (2013) developed micro hot-film shear-stress sensors using surface micromachining techniques and fabricated a micro silicon-nitride diaphragm and polycrystalline silicon heat-sensing element [2]. Therefore, micro/nano measurement equipment is required to assess the fabricating quality of micro devices. The leaf spring is often used as a key component of high-precision measuring instruments, or micro-actuators, when an elastic and recoverable micromotion is required, such as the cantilever beam used in atomic force microscopy (AFM). Claverley et al. (2010) and Li et al. (2014, 2016 and 2016) developed several probing systems for micro/nano CMMs [3–5].

Leaf springs are used to transfer the 3D displacements of the ball tip into the motions of a mirror in a vertical translation and horizontal rotation. The performance of the leaf spring directly determines the repeatability, stiffness and isotropy of the probe. Because the beryllium bronze sheet used as an elastic membrane has better repeatability and stiffness, it is a more suitable material for the fabrication of patterned leaf springs. The flatness and edge uniformity of a leaf springs mainly affect the probe's repeatability. The higher the flatness and uniformity, the better the repeatability. The size precision of a leaf springs mainly affects the probe's stiffness and isotropy. According to the Equation (1) or Equation (2) in Reference [3], we can obtain that the relative change of the stiffness or anisotropy is less than 0.5% when the four V-shaped leaf spring's width have an error of 10 μm. Therefore, the target manufacturing accuracy of a leaf springs can be controlled in ±10 μm.

A common method for fabricating beryllium bronze leaf springs is wire-electrode discharge machining (wire-EDM), which produces a better edge quality than laser cutting. However, we found that wire-EDM still cannot satisfy the high-precision manufacturing requirements of a leaf spring owing to the following disadvantages: (1) poor edge uniformity and a large offset, (2) processing heat and anisotropic stress causing a deformation of the fabricated leaf spring, (3) difficulty in fabricating leaf springs with complex patterns and (4) a high cost. Other methods can be used to process beryllium bronze. For example, Liu et al. (2002) described a lithography technology for microstructure fabrication and fabricated micro-capillaries and a micro-valve [6]. Tajiri et al. (2001) described an electrotyping technology, which they used to develop a copper lining for RF (Radio frequency) cavity sputtering [7]. Schindler et al. (2001) summarized the fabrication of 3D micro-optical resist structures (e.g., micro-lens arrays and blazed Fresnel lens structures) using hard optical and optoelectronic materials through ion beam and plasma jet etching technology [8]. Finally, Bodeux et al. (2014) fabricated tetragonal tungsten bronze using RF magnetron sputtering [9]. Although high precision and a small size can be achieved through these methods, their practicality for the fabrication of leaf springs remains limited owing to high costs, low efficiency and difficult operations.

Aiming at the shortcomings of above methods, a chemical etching method that can be used to fabricate a beryllium-bronze based leaf spring was adopted for this study. This method is based on the chemical reaction of copper (Cu) with sodium persulfate (Na<sub>2</sub>S<sub>2</sub>O<sub>8</sub>, 2017, Taishan Chemical Factory Co., Ltd., Taishan, China). The advantages of this method include (1) high precision, (2) a lack of thermal and stress deformations, (3) easy operation and the ability to fabricate complex leaf spring patterns, (4) low energy consumption and (5) low cost. The results of a comparison of the different manufacturing methods available are summarized in Table 1 [10–14].

**Table 1.** Comparison of different manufacturing methods.

Item	Lithography	Electrotyping	Sputtering	Wire-Electro Discharge Machining (EDM)	Laser Cutting	Chemical Etching
precision	<10 nm	≤0.5 μm	10 nm	5–20 μm	50 μm	<10 μm
deformation	no	non-uniform thickness	no	thermal and stress	thermal and stress	no
complex patterns	easy	easy	easy	difficult	easy	easy
equipment cost (\$)	80,000+	1500+	15,000+	1500+	1500+	150–
power (kW)	1–100	3–10	3–20	2–5	2–10	<0.6
operation	very complex	very complex	very complex	complex	complex	very simple
space usage (m <sup>2</sup> )	>50	>50	>50	10–15	10–15	<5
production periodic	>5 h	>5 h	>5 h	<1 h	<1 h	≤2 h

## 2. Principle and Process Flow

### 2.1. Principle

The grade of the beryllium bronze sheet used in this study is C17200. The beryllium bronze's main element is Cu, other gradient includes: beryllium (1.9–2.15%), cobalt (0.35–0.65%), nickel (0.2–0.25%), silicon (<0.15%), iron (<0.15%) and aluminum (<0.15%). Cu can be corroded through a solution of Na<sub>2</sub>S<sub>2</sub>O<sub>8</sub>. The chemical reaction equation is as follows:



A beryllium bronze sheet is always used as the raw material of a leaf spring. A photoresist film (negative photosensitive material) can be used to protect the beryllium bronze sheet from corroding. The photoresist film used to cover a beryllium bronze sheet has to conform to the pattern of the designed leaf springs in advance. Therefore, the part of a beryllium bronze sheet with a photoresist film will be protected and maintained, whereas other parts will be etched off in a  $\text{Na}_2\text{S}_2\text{O}_8$  solution. Thus, patterned leaf springs can be obtained.

2.2. Process Flow

The process flow of this method consists of eight steps, namely, the pattern design, mask printing, pretreatment, pasting of photo-sensory membranes, exposure, development, etching and mold unloading, the details of which are as follows.

(1) Pattern design.

According to the actual demand of a particular type of measuring equipment, it is necessary to design the pattern of the leaf springs first and to then analyze and optimize the structural parameters of the pattern. Li et al. (2014) designed a four V-shaped suspending leaf spring pattern (shown in Figure 1) for their 3D contact-scanning probe [3,15].

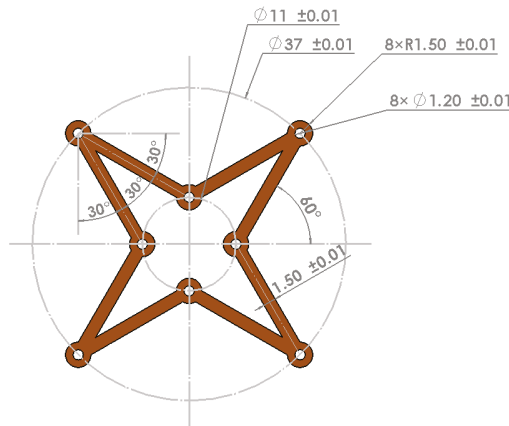


Figure 1. Four V-shaped leaf springs.

(2) Mask printing.

A high-precision mask is required to copy the designed pattern onto the beryllium bronze sheet. As a necessary intermediary, the precision of the mask directly determines the quality of the leaf springs. A professional film-printer of greater than  $600 \times 600$  dot per inch (DPI) was used to print the pattern onto the unsmooth surface of a piece of film. When printing, a part of the pattern should be transparent to allow ultraviolet light to pass through and the remaining part should be black. A printed mask is shown in Figure 2.

(3) Pretreatment.

Beryllium bronze is oxidized when exposed in air for a long period of time and forms an oxidation film on the surface. It is difficult for the oxidation film to be corroded by the solution of  $\text{Na}_2\text{S}_2\text{O}_8$ . Therefore, we have to clear the oxidation film using the following pretreatment procedures: select a piece of beryllium bronze with an appropriate size, polish it on both sides using fine waterproof abrasive paper (7000 #) with a grit size of far less than  $0.5 \mu\text{m}$  and wipe off the powders produced on the surface. We then need to place the polished beryllium bronze sheet into a solution (SanNy 3111,

2017, Dunhua, Inc., Shenzhen, China) for approximately 30 s to clean the oxidation film thoroughly. Finally, we need to rinse the beryllium bronze sheet using clean water and dry it. Images of an original beryllium bronze sheet and polished (using SanNy 3111) and cleaned sheets, are shown in Figure 3, respectively.

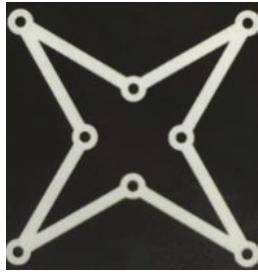


Figure 2. Printed mask.

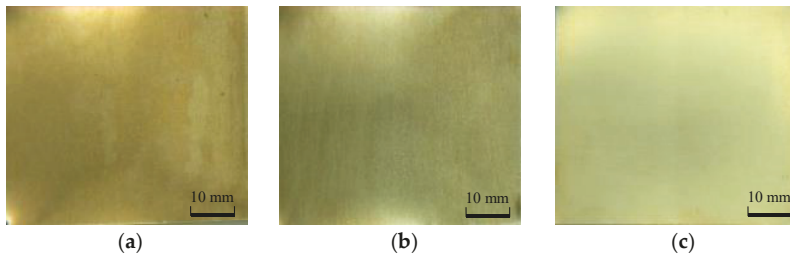


Figure 3. Images of beryllium bronze sheet: (a) original, (b) polished, (c) and cleaned.

#### (4) Pasting photo-sensory membrane.

The photo-sensory membrane (No.2138, 2017, Hemera, Inc., Shenzhen, China) we used was composed of a polyester, photoresist and polyethylene layers in sequence, as shown in Figure 4. The polyester layer is used to support and protect the photosensitive layer from being diffused when exposed to oxygen, which can destroy the light sensitivity of the photoresist. The polyethylene layer is used to cover the other side of the photosensitive layer to prevent it from being polluted by dust or other types of particles. The photoresist layer is the main body of the photosensitive membrane and is mostly made of a negative photosensitive material. The polyethylene layer needs to be torn off before attaching the photosensitive membrane onto the beryllium bronze sheet. The total thickness of the polyester layer and photoresist layer is 38  $\mu\text{m}$ . Both sides of the beryllium bronze sheet need to be pasted with a photosensitive membrane and the photoresist layer should be attached to the sheet evenly and tightly with no bubbles. To affix the photoresist to the beryllium bronze sheet firmly, the beryllium bronze sheet with the photosensitive membrane needs to be pressed by two flat glasses and heated by a hot air blower. Figure 5 shows a beryllium bronze sheet pasted using a photo-sensory membrane.

#### (5) Exposure.

This process is used to place the prepared mask on the smooth surface of a beryllium bronze sheet with a photosensitive membrane and to make sure that the side with the printed pattern of the mask film is close to the beryllium bronze sheet. Both the beryllium bronze sheet and the mask film are clamped using two flat glasses. The next step is to expose the beryllium bronze sheet to a 365 nm

ultraviolet light from a lamp of 48 W for approximately 2.5 min on both sides. The ultraviolet light can pass through the patterns of the mask film and solidify the photoresist. The color of the solidified photoresist is dark blue. The dark blue pattern shown in Figure 6 is what we want to achieve.

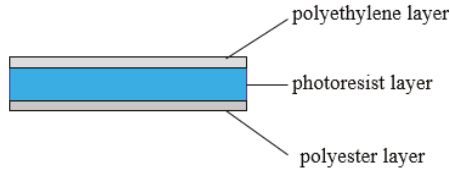


Figure 4. Sketch of photo-sensory membrane.

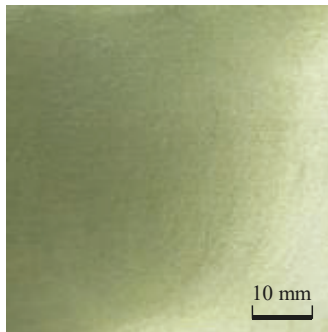


Figure 5. Beryllium bronze sheet with photo-sensory membrane.

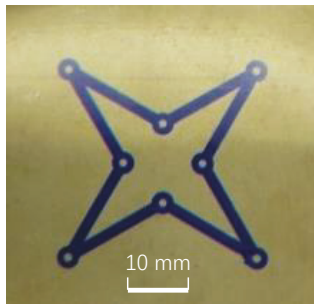


Figure 6. Exposed photo-sensory membrane.

#### (6) Development.

The developer, whose main component is sodium carbonate ( $\text{Na}_2\text{CO}_3$ , 2017, Taishan Chemical Factory Co., Ltd., Taishan, China), should be watered at a mass ratio of 1:100. The proper temperature of the sodium carbonate solution should be 30–40 °C. First, we must tear the polyester film of the exposed photo-sensory membrane off the beryllium bronze sheet and immerse it in the developer solution and then lightly scrub the patterned side of the beryllium bronze sheet with a brush to remove the unexposed photoresist. The final process is to rinse the beryllium bronze sheet using clean water and dry it again.

(7) Etching.

The developed beryllium bronze sheet is placed into a  $\text{Na}_2\text{S}_2\text{O}_8$  solution, the temperature of which is controlled and maintained at a stable value by using a thermostat water bath until the uncovered beryllium bronze is completely etched. Therefore, the leaf springs with a solidified photoresist will appear.

(8) Mold unloading.

The solidified photoresist on the leaf springs should be dissolved during this step. A special chemical for dissolving the solidified photoresist—the main component of which is sodium hydroxide (concentration NaOH, 2017, Taishan Chemical Factory Co., Ltd., Taishan, China)—should be watered at a mass ratio of 1:50 beforehand. The proper temperature of the sodium hydroxide solution is approximately 50 °C. The unfinished leaf spring is then soaked in the chemical solution, where it remains stable until the attached photoresists are cleared out, after which, the leaf spring is formed, as shown in Figure 7.

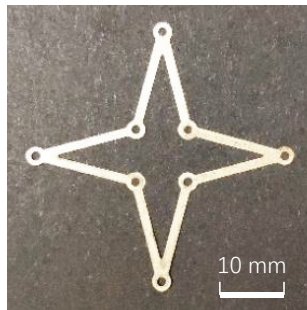


Figure 7. Finished leaf springs.

In term of the operational steps, this method (method A) is similar to the conventional photolithography (method B) [10]. However, there are essential differences between the two methods in practical, which include: (1) Method B adopts complicated photoresist coating and soft bake process to coat the photosensitive material on the wafer. Method A pastes the photo-sensory membrane on the beryllium bronze sheet directly by manual. (2) Method B uses scanning projection exposure method, which requires precision mechanical systems for assistance. Method A finish the exposure by irradiating with ultraviolet rays directly. (3) Method B usually uses plasma or sputtering process to do etching. Method A carries out the etching by sodium bicarbonate solution. (4) Method B has drying and wetting methods to unload the mold. Method A unloads the mold by sodium hydroxide solution.

### 3. Process Parameter Optimization

#### 3.1. Taguchi Experiments

The precision of a fabricated pattern on a leaf spring can be evaluated based on the roundness of its circle or arc, the straightness of the lines and the offset of the position. The roundness and straightness reflect formation errors of the pattern. An offset reflects a size error, such as the diameter, width, or position. Many factors are related to these three indicators. However, we found that the key processing parameters are the concentration and temperature of the etchant solution and the exposure time based on analysis and experience. The concentration of the etchant solution directly affects the etching speed and pH of the solution, thereby influencing the quality of the leaf spring. The temperature of the etchant solution will also affect the etching speed. A lower temperature will

slow down the reaction speed and a higher temperature will make the photoresist membrane fall off. An inadequate exposure time cannot solidify the photoresist completely. Instead, it will result in poor roundness or straightness and may even force the photo-sensory membrane to drop off. An overly long exposure time will bring about difficulties in the development and mold unloading and make the photoresist crisper. Therefore, proper experimental parameters are vital for us to fabricate high-precision leaf springs.

The Taguchi method was usually employed to obtain the optimal process parameters in different applications [16–20], which was also used in the present study. The roundness, straightness and offset of the leaf springs are regarded as three responsive variables. The three control factors are the concentration and temperature of the etchant solution and the exposure time. The levels of the control factors, presented in Table 2, were determined experimentally.

Table 2. Control factors and their levels.

Level	Factor A	Factor B	Factor C
	Concentration (Etchant: Water)	Temperature (°C)	Time (min)
1	1:3	50	2
2	1:4	60	2.5
3	1:5	70	3

The orthogonal table of L9 (3<sup>3</sup>) was used to design the test parameters. The parameters of each factor were filled into the orthogonal table according to the rules, as presented in Table 3.

Table 3. Orthogonal table of L9 (3<sup>3</sup>).

L9	Factor A	Factor B	Factor C
1	1	1	1
2	2	1	2
3	3	1	3
4	2	2	1
5	3	2	2
6	1	2	3
7	3	3	1
8	1	3	2
9	2	3	3

The signal-to-noise (S/N) ratio is applied in the Taguchi method as the quality characteristic of choice. In the pursuit of smaller characteristics of the roundness, straightness and offset of the leaf springs, a smaller-the-better (STB) type characteristic was used as the objective function (Equation (2)) in this experiment.

$$\eta = -10 \log\left(\frac{1}{n} \sum_{i=1}^n y_i^2\right) \tag{2}$$

Here,  $\eta$  is the signal to noise ratio (S/N: dB),  $n$  indicates the experimental period and  $y_i$  can be the roundness, straightness, or offset of the leaf springs.

According to the experimental requirements of the orthogonal table, nine group experiments manufacturing the leaf springs shown in Figure 1 were carried out and repeated five times for each group. One of the holes on the leaf springs was selected to evaluate roundness. One of the branches on the leaf springs was selected to evaluate straightness and offset. The selected hole and branch were measured each time using an image profile projector (2011, MUMA200, 3DFamily, Inc., Nanjing China, with a resolution of 1  $\mu$ m). The mean values of the roundness, straightness and offset are shown in Table 4. The values of  $\eta$  were also calculated and listed in Table 4.



**Table 4.** Mean values of the response variable and its S/N ratio.

No.	Average Roundness (μm)	$\eta_R$	Average Straightness (μm)	$\eta_S$	Average Offset (μm)	$\eta_O$
1	70	-37.80	21	-26.93	15	-23.77
2	73	-37.38	57	-35.40	15	-22.16
3	54	-33.67	21	-27.39	6	-16.19
4	32	-31.08	29	-27.58	7	-15.16
5	54	-34.99	50	-34.74	12	-22.32
6	59	-35.66	40	-32.76	10	-20.87
7	41	-33.16	37	-31.65	9	-19.48
8	34	-30.49	29	-30.39	6	-16.26
9	44	-32.76	34	-30.81	10	-19.57
mean value	51.2	-34.11	35.3	-30.58	10.0	-19.53

3.2. Mean Value Analysis

To find the optimal combination of the minimum roundness, straightness and offset, the mean values of the S/N ratio at the same level for each factor were calculated as shown in Tables 3 and 4. For example, the mean S/N ratio for roundness at level 2 of factor A is calculated using

$$\overline{\eta_{RA2}} = \frac{1}{3}(\eta_R(2) + \eta_R(4) + \eta_R(9)) \tag{3}$$

where R and A in the subscript denote the roundness and factor A, respectively. Factor A is set to level 2 only in experiments 2, 4 and 9. The mean S/N ratios for all levels of all factors can be obtained in a similar way, as shown in Table 5.

**Table 5.** Mean values at the same level of each factor and variable.

Response Variables	Level	Mean S/N Ratio		
		Factor A	Factor B	Factor C
roundness	1	-36.28	-34.01	-34.65
	2	-33.91	-34.29	-33.74
	3	-32.14	-34.03	-33.94
straightness	1	-29.91	-28.72	-30.03
	2	-31.69	-33.51	-31.26
	3	-30.95	-30.32	-31.26
offset	1	-20.71	-19.47	-20.30
	2	-19.45	-20.25	-18.96
	3	-18.72	-19.33	-19.33

According to the Taguchi principle, the maximum S/N ratio value should be considered when determining the main levels and main factors. From Table 5, we can see that the combination of factor levels (A<sub>3</sub>, B<sub>1</sub>, C<sub>2</sub>), (A<sub>1</sub>, B<sub>1</sub>, C<sub>1</sub>) and (A<sub>3</sub>, B<sub>3</sub>, C<sub>2</sub>) contribute to a minimization of the roundness, straightness and offset, respectively. Using the optimal factor-level combination, three groups experiments were conducted separately for the minimum roundness, straightness and offset, with each group repeated five times. Table 6 shows the results.

**Table 6.** Measured results of three variables.

No.	1	2	3	4	5	Combination of Factor Levels
roundness (μm)	52	46	33	38	49	A <sub>3</sub> , B <sub>1</sub> , C <sub>2</sub>
straightness (μm)	22	15	30	16	22	A <sub>1</sub> , B <sub>1</sub> , C <sub>1</sub>
offset (μm)	9	10	8	10	9	A <sub>3</sub> , B <sub>3</sub> , C <sub>2</sub>

Plugging the results shown in Table 6 into Equation (2), we can obtain the S/N ratio ( $\eta$ ) of the roundness, straightness and offset, shown in Table 7. The expectations of the S/N ratio ( $\eta_{RO}$ ,  $\eta_{SO}$  and  $\eta_{OO}$ ) of the roundness, straightness and offset can be calculated as

$$\eta_{RO} = \overline{\eta_R} + (\overline{\eta_{RA3}} + \overline{\eta_{RB1}} + \overline{\eta_{RC2}}) - 3\overline{\eta_R} \tag{4}$$

$$\eta_{SO} = \overline{\eta_S} + (\overline{\eta_{SA1}} + \overline{\eta_{SB1}} + \overline{\eta_{SC1}}) - 3\overline{\eta_S} \tag{5}$$

$$\eta_{OO} = \overline{\eta_O} + (\overline{\eta_{OA3}} + \overline{\eta_{OB3}} + \overline{\eta_{OC2}}) - 3\overline{\eta_O} \tag{6}$$

where  $\eta_R$ ,  $\eta_S$  and  $\eta_O$  are the mean values of the S/N ratio shown in Table 3. Here, the R, S and O subscripts denote the roundness, straightness and offset, respectively. In addition, the A, B and C subscripts indicate factors A, B and C, respectively and 1, 2 and 3 represent levels 1, 2 and 3. The relative error can also be calculated using Equation (7). Both the expectations and relative errors are presented in Table 7. From Table 7, we can see that all of the relative errors are less than 10%. Thus, the effectiveness of the factor-level combinations obtained is verified.

$$E_R = \frac{|\eta_O - \eta|}{|\eta_O|} \times 100\% \tag{7}$$

Table 7. S/N ratio and its relative error.

Item	$\eta_o$	$\eta$	Relative Error (%)
roundness	-31.67	-33.05	4.36
straightness	-26.95	-26.97	0.07
offset	-17.95	-19.59	9.13

### 3.3. Variance Analysis

Three factor-level combinations aiming at roundness, straightness and offset were obtained using a mean value analysis. We still need to calculate the contributions of each factor for the roundness, straightness and offset through a variance analysis, so as to achieve the final optimal process parameters. The variance (SS) for each variable and each control factor was calculated using Equation (8). The variances and contributions of the control factor for each variable were shown in Table 8.

$$SS_{VF} = 3 \sum_{i=1}^3 (\overline{\eta_{VF_i}} - \overline{\eta_V})^2 \tag{8}$$

where  $V$  indicates a variable, which can be roundness, straightness, or offset;  $F$  indicates the factor, which can be A, B, or C;  $i$  is the sequence number of the factor level; and  $\overline{\eta_V}$  and  $\overline{\eta_{VF_i}}$  are as shown in Tables 3 and 4, respectively.

Table 8. Contributions of the parameters.

Control Factors	Roundness		Straightness		Offset	
	$SS_R$	%	$SS_S$	%	$SS_O$	%
A	25.97	94.5	4.83	11.1	6.14	57.5
B	0.14	0.5	35.68	81.9	1.67	15.7
C	1.37	5.0	3.05	7.0	2.86	26.8
sum	27.48	100	43.56	100	10.67	100

From Table 8, we can see that the dominant control factor for roundness is factor A (concentration of etchant solution), the dominant control factor for straightness is factor B (temperature of the etchant solution) and the control factor's order of importance for an offset is A, C and B in sequence. Combining

the factor-level combinations above, we can see that the optimal process parameters are  $A_3$ ,  $B_1$  and  $C_2$ . In other words, the best mass ratio of etchant to water is 1:5, the best temperature of the etchant solution is 50 °C and the best exposure time is 2 min and 30 s.

## 4. Experimental Results

### 4.1. Leaf Springs with Complex Pattern

Using the optimal process parameters obtained through Taguchi experiments, different leaf springs with complex patterns were fabricated using the proposed chemical etching method, as shown in Figure 8. The thinnest line width is less than 0.5 mm. The leaf springs are also very flat and smooth without any warping or deformations. It is very difficult for wire-EDM to fabricate these complex leaf springs with thin lines and many disconnected regions. First, the thinnest width of the line wire-EDM that can be achieved is approximately 1 mm. A through-hole then needs to be drilled to allow a wire to pass through when cutting a newly disconnected region using wire-EDM. The more disconnected regions the leaf spring has, the longer the time, the higher the cost and the more energy consumption required. Finally, warping and deformation of the sheet are inevitable owing to the step-by-step process using.

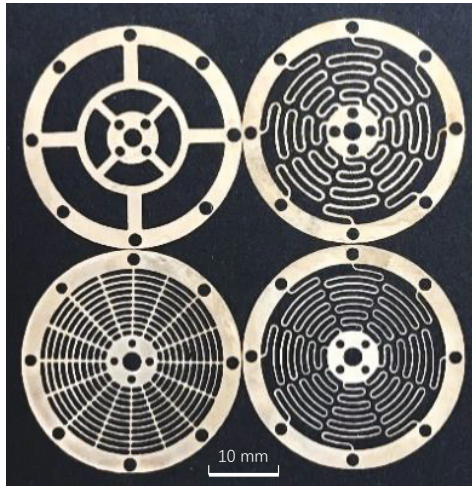
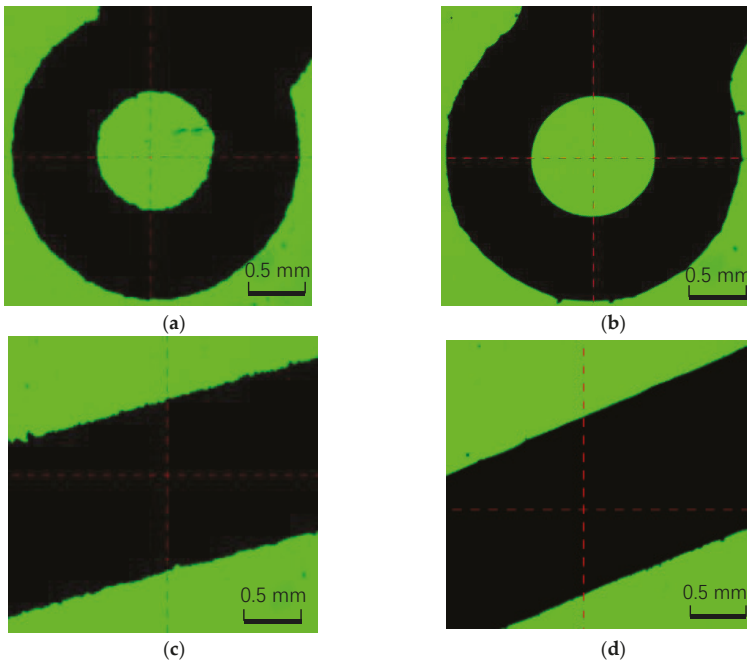


Figure 8. Leaf springs with a complex pattern.

### 4.2. Quality Measurements of Leaf Springs

Using the optimal process parameters obtained through the Taguchi method, the four V-shaped leaf spring Figure 1 shows was fabricated five times using an etching method. As a comparison, the same leaf spring was also fabricated five times using the wire-EDM method [21]. The images of a ring and line from the leaf springs fabricated using these two methods are shown in Figure 9. Their roundness, straightness and offset were measured using a MUMA 200 (2011, 3DFamily, Inc., Nanjing China), the results of which are provided in Tables 9–11, respectively.



**Figure 9.** A ring of a leaf spring using (a) wire-EDM, (b) etching and a line of a leaf spring using, (c) wire-EDM and (d) etching.

**Table 9.** Results of error in roundness.

No.	Etching ( $\mu\text{m}$ )	Wire-EDM ( $\mu\text{m}$ )	Difference ( $\mu\text{m}$ )
1	4	31	27
2	43	63	20
3	37	40	3
4	25	63	38
5	21	78	57
mean value	26.0	55.0	29.0

**Table 10.** Results of error in straightness.

No.	Etching ( $\mu\text{m}$ )	Wire-EDM ( $\mu\text{m}$ )	Difference ( $\mu\text{m}$ )
1	27	45	18
2	36	76	40
3	20	49	29
4	16	88	72
5	35	76	41
mean value	26.8	66.8	40.0

**Table 11.** Results of error in offset.

No.	Etching ( $\mu\text{m}$ )	Wire-EDM ( $\mu\text{m}$ )	Difference ( $\mu\text{m}$ )
1	7	17	10
2	6	15	9
3	4	19	15
4	4	23	19
5	7	20	13
mean value	5.6	18.8	13.2

We can see from Figure 9 that the edge of the leaf spring developed using an etching process is much smoother than that using wire-EDM. From the measured data in Tables 9–11, we can conclude that the roundness, straightness and offset of the leaf springs achieved through etching are much better than those achieved using wire-EDM.

The measurement uncertainty for the leaf springs is mainly composed of five parts. The first one is the measurement repeatability which reflecting the contributions from random errors. The second one is the indication error which reflecting the contributions from resolution, evaluation algorithm and so on. The third one is the attitude error which reflecting the contributions from the different attitudes of the leaf springs. The fourth one is the alignment error. The fifth one is the contributions from temperature and humidity which can be ignored when evaluating the form errors. One of the holes, and a rectangle leaf from the manufactured leaf spring, shown in Figure 1 were measured 10 times, using the proposed method, in a standard measurement environment ( $u_5 = 0$ ). Table 12 shows the measurement results and standard uncertainties ( $u_1$ ). From Table 12, we can also see that the manufacturing errors of both the diameter and the line width of the leaf springs shown in Figure 1 are less than  $9 \mu\text{m}$ . The indication errors of the image profile projector we used were investigated by measuring a calibration mask in different angles, the indication errors were shown in Table 13. Considering that the maximum indication errors include the attitude errors, the uncertainties ( $u_{23}$ ) contributed from indication errors and attitude errors have been obtained and shown in Table 13. The uncertainties ( $u_4$ ) contributed by alignment errors can be neglected because telecentric lenses were used in the image profile projector. Therefore, the measurement uncertainties shown in Table 14 have been achieved using the Equation of  $U = \sqrt{u_1^2 + u_{23}^2}$ .

Table 12. Results of repeated measurement.

No.	Diameter (mm)	Line Width (mm)	Roundness ( $\mu\text{m}$ )	Straightness ( $\mu\text{m}$ )	Offset ( $\mu\text{m}$ )
1	1.198	1.498	21	17	4
2	1.197	1.502	23	16	5
3	1.198	1.494	22	16	4
4	1.198	1.498	23	12	3
5	1.198	1.499	23	17	4
6	1.198	1.492	23	17	4
7	1.198	1.491	21	18	5
8	1.197	1.497	21	15	4
9	1.198	1.495	22	14	4
10	1.197	1.491	21	18	4
average	1.1977	1.496	22	16	4.1
standard deviation ( $u_1$ )	0.0005	0.0035	0.9	1.8	0.5

Table 13. Results of indication errors in different attitudes.

Angle	Diameter ( $\mu\text{m}$ )	Line Width ( $\mu\text{m}$ )	Roundness ( $\mu\text{m}$ )	Straightness ( $\mu\text{m}$ )	Offset ( $\mu\text{m}$ )
60°	1	2	1	1	1
120°	2	2	2	1	1
180°	2	1	2	2	1
240°	2	1	3	1	2
300°	1	3	1	2	1
360°	2	2	2	1	1
maximum value ( $\Delta$ )	2	3	3	2	2
$u_{23} (= \Delta / \sqrt{3})$	1.2	1.7	1.7	1.2	1.2

Table 14. Results of measurement uncertainties.

Item	Diameter	Line Width	Roundness	Straightness	Offset
$U (\mu\text{m}) (k = 2)$	2.6	7.8	3.8	4.3	2.6

#### 4.3. Leaf Spring Performance

Leaf spring is a key component of a micro/nano probe and determines the measurement precision of the probe to a great extent. The leaf spring shown in Figure 1 was manufactured by chemical etching method and applied to our probe [4]. Both the floating plate and the supporting ring connected to the leaf spring have a dimensional tolerance of  $0.1\ \mu\text{m}$ . The assembling accuracy of the probe on the leaf springs are controlled by the manufacturing accuracy since the symmetric design of the leaf spring and its connected attachments. An experimental setup shown in Figure 10 were built and the probe was tested. The probe was assembled in a frame of stand. A  $2\ \text{mm} \times 2\ \text{mm}$  square hole formed by four 0 grade gauge blocks with same horizontal sizes of  $30\ \text{mm} \times 9\ \text{mm}$  and thicknesses of  $2\ \text{mm}$ ,  $2\ \text{mm}$ ,  $1.6\ \text{mm}$  and  $1.7\ \text{mm}$  respectively, was used to contact the probe tip in lateral directions. As for the inner surfaces of the square hole, two parallel surfaces of them are the reference planes of the gauge block but the other two are not. Only the reference planes of the gauge block were used to contact the probe tip in the experiment. Therefore, rotation for the probe is required when changing a contact direction in horizontal. A round dial with an accuracy of  $0.5$  degree was equipped on the rotating shaft of the probe, so as to index the rotated angle of the probe. A 2D linear stage was used to change the initial position of the square hole manually. A high-precision 3D nano-positioning stage made by Physik Instrumente (2013, PI, model P-561.3 CD with a repeatability of  $2\ \text{nm}$  and a distance of travel of  $100\ \mu\text{m}$  in each direction, Germany) was used as a reference with an interval of  $2\ \mu\text{m}$ . 10 measurements were repeated at each point. The results (X2 and Y2) were shown in Figure 11. From Figure 11, we can see that the probe's measurement standard deviation is about  $16\ \text{nm}$ .

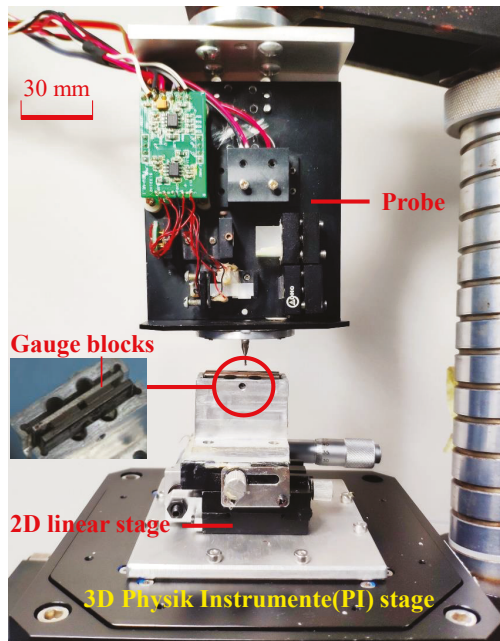


Figure 10. Photograph of the experimental setup.

In comparison, the same leaf spring was manufactured by wire-EDM, used in the same probe and similar experiments were conducted by the same experimental setup. The results (X1 and Y1) were also shown in Figure 11. In this case, the measurement standard deviation of the probe is about  $30\ \text{nm}$ . Compared with the leaf springs fabricated by the proposed method, the ones fabricated by EDM have

evident thermal deformation and poor edge uniformity, which cause an anisotropic and lead to worse measurement standard deviation of the probe. The experimental results demonstrate that the chemical etching method is a better choice in fabricating high-precision leaf springs for a micro/nano probe.

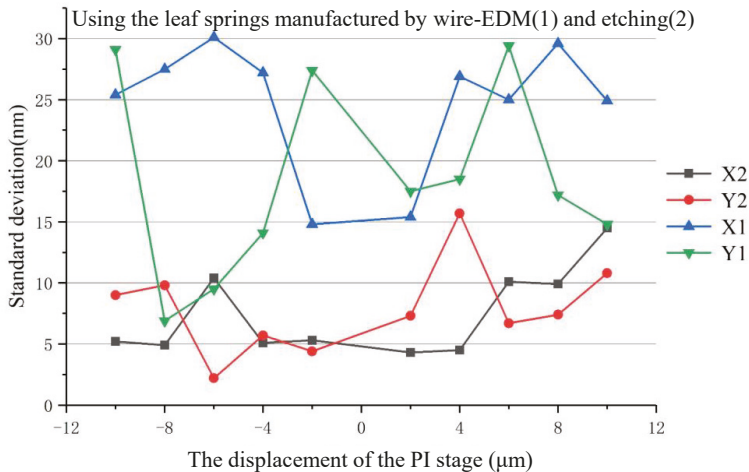


Figure 11. Results of the experiments.

## 5. Conclusions

A chemical etching method was proposed to fabricate the leaf springs for a CMM probe in this paper. The optimal process parameters were obtained using the Taguchi method, which includes a mass ratio of etchant to water of 1:5, temperature of the etchant solution of 50 °C and an exposure time of 2 min and 30 s. Compared with the wire-EDM method, this method has the advantages of high precision, no deformations, easy fabrication of complex patterns, low energy consumption, low cost and easy operation. This manufacturing method can be used to fabricate different high-precision springs for CMM probes and other types of measurement equipment.

**Author Contributions:** Conceived the Method and Wrote the Paper, R.-J.L.; Designed and Performed the Taguchi Experiments, P.-Y.W.; Improved the Experimental Results by Using a Thermostat Water Bath, D.-D.L.; Edited the Manuscript, K.-C.F.; Performed some Confirmatory Experiments, F.-F.L. and L.-J.C.; Designed the Confirmatory Experiments, Q.-X.H.

**Funding:** This study was funded by the National Natural Science Foundation of China (51675157, 51475131).

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Liu, C.; Huang, J.B.; Zhu, Z.; Jiang, F. A micromachined flow shear-stress sensor based on thermal transfer principles. *J. Microelectromech. Syst.* **1999**, *8*, 90–99. [\[CrossRef\]](#)
2. Alshehri, A.; Kraft, M.; Gardonio, P. Two-mass MEMS velocity sensor: Internal feedback loop design. *IEEE Sens. J.* **2013**, *13*, 1003–1011. [\[CrossRef\]](#)
3. Li, R.J.; Fan, K.C.; Miao, J.W.; Huang, Q.X.; Tao, S.; Gong, E.M. An analogue contact probe using a compact 3D optical sensor for micro/nano coordinate measuring machines. *Meas. Sci. Technol.* **2014**, *25*, 094008. [\[CrossRef\]](#)
4. Li, R.J.; Fan, K.C.; Huang, Q.X.; Zhou, H.; Gong, E.M.; Xiang, M. A long-stroke 3D contact scanning probe for micro/nano coordinate measuring machine. *Precis. Eng.* **2016**, *43*, 220–229. [\[CrossRef\]](#)
5. Li, R.J.; Xiang, M.; He, Y.X.; Fan, K.C.; Cheng, Z.Y.; Huang, Q.X.; Zhou, B. Development of a high-precision touch-trigger probe using a single sensor. *Appl. Sci.* **2016**, *6*, 86. [\[CrossRef\]](#)

6. Liu, W.; Cai, Q.; Guo, X.; Chen, Y.; Wang, L. Three dimensional micro-structure soft lithography technology. *Chin. J. Sens. Actuators* **2002**, *15*, 136–139.
7. Tajiri, K.; Nakamura, T.; Kabeya, Z.; Yamanaka, Y.; Naito, F. Development of an electroformed copper lining for accelerator components. *Electrochim. Acta.* **2001**, *47*, 143–148. [[CrossRef](#)]
8. Schindler, A.; Haensel, T.; Flamm, D. Ion beam and plasma jet etching for optical component fabrication. *J. Soc. Photo-Opt. Instrum. Eng.* **2001**, *4440*, 217–227. [[CrossRef](#)]
9. Bodeux, R.; Michau, D.; Josse, M.; Maglione, M. Dielectric properties of tetragonal tungsten bronze films deposited by RF magnetron sputtering. *Solid State Sci.* **2014**, *38*, 112–118. [[CrossRef](#)]
10. Chen, Y.F. Nanofabrication by electron beam lithography and its applications: A review. *Microelectron. Eng.* **2015**, *135*, 57–72. [[CrossRef](#)]
11. Li, Y.H.; Ding, G.F.; Mao, H.P.; Zhang, Y.H. The Development of microelectroforming process of LIGA/Quasi-LIGA technology. *Electron. Process Technol.* **2005**, *26*, 1–5.
12. Kelly, P.J.; Arnell, T.D. Magnetron sputtering: A review of recent developments and applications. *Vacuum* **2000**, *56*, 159–172. [[CrossRef](#)]
13. Ho, K.H.; Newman, S.T.; Rahimifard, S.; Allen, R.D. State of the art in wire electrical discharge machining (WEDM). *Int. J. Mach. Tool Manuf.* **2004**, *44*, 1247–1259. [[CrossRef](#)]
14. Jiang, H.H. Development and forecast of the laser processing technology application. *Optoelectron. Technol. Inf.* **2001**, *4*, 1–10.
15. Li, R.J.; Fan, K.C.; Zhou, H.; Wang, N.; Huang, Q.X. Elastic mechanism design of the CMM contact probe. In Proceedings of the Sixth International Symposium on Precision Mechanical Measurements, Guiyang, China, 10 August 2013; SPIE: Bellingham, WA, USA.
16. Lan, T.S.; Chuang, K.C.; Chen, Y.M. Optimization of machining parameters using fuzzy Taguchi method for reducing tool wear. *Appl. Sci.* **2018**, *8*, 1011. [[CrossRef](#)]
17. Krzysztof, M.; Katarzyna, M.; Sebastian, B.; Jakub, B.; Alicja, T.; Marta, M.; Martyna, L.; Szymon, P.; Beata, G. Optimization of a culture medium using the Taguchi approach for the production of microorganisms active in odorous compound removal. *Appl. Sci.* **2017**, *7*, 756. [[CrossRef](#)]
18. Vigneashwara, P.; Wahyu, C.; Tegoeh, T.; Gunasekaran, P. Predictive modelling and analysis of process parameters on material removal characteristics in abrasive belt grinding process. *Appl. Sci.* **2017**, *7*, 363.
19. Do, T.V.; Hsu, Q.C. Optimization of minimum quantity lubricant conditions and cutting parameters in hard milling of AISI H13 steel. *Appl. Sci.* **2016**, *6*, 83. [[CrossRef](#)]
20. Nguyen, H.T.; Hsu, Q.C. Surface roughness analysis in the hard milling of JIS SKD61 alloy steel. *Appl. Sci.* **2016**, *6*, 172. [[CrossRef](#)]
21. Fostex Manual CNC Wire Cut FZC Series Machine. Available online: <https://www.indiamart.com/proddetail/cnc-wire-cut-fzc-series-machine-15574858988.html> (accessed on 25 July 2018).



© 2018 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).



Article

# Calibration Method of Orthogonally Splitting Imaging Pose Sensor Based on General Imaging Model

Na Zhao <sup>1,2</sup>, Changku Sun <sup>1</sup> and Peng Wang <sup>1,\*</sup>

<sup>1</sup> State Key Laboratory of Precision Measuring Technology and Instruments, Tianjin University, Tianjin 300072, China; zhaonatj@tju.edu.cn (N.Z.); sunck@tju.edu.cn (C.S.)

<sup>2</sup> College of Computer and Information Engineering, Tianjin Agricultural University, Tianjin 300384, China

\* Correspondence: wang\_peng@tju.edu.cn; Tel.: +86-138-2086-1312

Received: 2 July 2018; Accepted: 6 August 2018; Published: 19 August 2018



**Abstract:** Orthogonally splitting imaging pose sensor is a new sensor with two orthogonal line array charge coupled devices (CCDs). Owing to its special structure, there are distortion correction and imaging model problems during the calibration procedure. This paper proposes a calibration method based on the general imaging model to solve these problems. The method introduces Plücker Coordinate to describe the mapping relation between the image coordinate system and the world coordinate system. This paper solves the mapping relation with radial basis function interpolation and adaptively selecting control points with Kmeans clustering method to improve the fitting accuracy. This paper determines the appropriate radial basis function and its shape parameter by experiments. And these parameters are used to calibrate the orthogonally splitting imaging pose sensor. According to the calibration result, the root mean square (RMS) of calibration dataset and the RMS of test dataset are 0.048 mm and 0.049 mm. A comparative experiment is conducted between the pinhole imaging model and the general imaging model. Experimental results show that the calibration method based on general imaging model applies to the orthogonally splitting imaging pose sensor. The calibration method requires only one image corresponding to the target in the world coordinates and distortion correction is not required to be taken into account. Compared with the calibration method based on the pinhole imaging model, the calibration procedure based on the general imaging model is easier and accuracy is greater.

**Keywords:** orthogonally splitting imaging pose sensor; general imaging model; radial basis function interpolation

## 1. Introduction

With the development of advanced industrial equipment intelligence and digitalization, the demand for accurate pose measurement, which can measure a wide range of space targets, is increasing. The visual pose measurement has the advantages of non-contact, simplicity, stability, and moderate accuracy. It is widely used in aerospace, machine building, robot navigation and other fields. At present, the visual sensors used in the visual pose measurement system include the area array CCD and the line array CCD. Compared with the area array CCD, the line array CCD has higher resolution and faster speed in vision measurement. But the line array CCD can only obtain a one-dimensional image, so multiple line array CCDs are combined to realize the spatial measurement of the target [1–3]. However, the target must be in the public view of all line array CCDs; the measurement range is thus limited in a small range. In order to satisfy the measurement requirements of a wide range, high precision and fast speed, the orthogonally splitting imaging pose sensor based on line

array CCD is designed. The new sensor is composed of a special optical imaging system and dual line array CCDs to simulate monocular measurement with an area array CCD. Compared with binocular or trinocular vision measurements, the new sensor can solve the problems of small measurement range, and has higher resolution to improve the accuracy of measurement.

The optical system of the pose sensor is composed of two parts: the imaging system and the beam splitting system. Responsible for imaging a target point as a two-dimensional image, the imaging system is composed of an aperture and spherical mirror group as an objective lens with a large field of view. The beam splitting system is responsible for dividing the two dimensional image into two one-dimensional images, which are received by the line array CCD in the corresponding direction. It is composed of a beam splitting prism and two cylindrical mirror groups.

Owing to the special optical imaging structure, for the optical system of the orthogonally splitting imaging pose sensor, there exists various types of distortion, such as the radial distortion caused by wide field of view, one-way error caused by cylindrical lens, and linear and non-linear error caused by assemble technology. These distortions will affect the geometry relationship between the target and its image, and the ideal model of pinhole imaging model cannot properly describe the actual optical imaging relationship. Normally calibration method based on pinhole imaging model ignores tangential distortion and takes radial distortion into consideration. The main calibration methods including Tsai two-step [4], Weng [5], and Zhang [6] requires at least two images corresponding to the target in the world coordinates. But problems do exist with these approaches. Pupil aberration and assembly error still exist. Tardif [7,8] proposed a rotational symmetry model to solve the problem of pupil aberration, but they could not solve the problem of assembly error.

With the above mentioned analysis, this paper proposes a calibration method based on a general imaging model [9–12] to calibrate the orthogonally splitting imaging pose sensor. General imaging model is a black box model. It describes the mapping relation between the incoming light and the pixel without considering distortion. Thus the model can apply to any optical imaging system. This paper adopts continuous vector-valued functions to express the mapping relation solved by radial basis function interpolation. And adaptively selected control points are selected by Kmeans clustering method to improve fitting accuracy. The calibration method requires only one image corresponding to the target in the world coordinates.

## 2. Materials and Methods

### 2.1. Mathematical Model

#### 2.1.1. Imaging Principle of the Orthogonally Splitting Imaging Pose Sensor

The orthogonally splitting imaging pose sensor consists of a spherical mirror, a beam splitting prism, cylindrical lens and line array CCDs as shown in Figure 1. The target point  $P_w(x_w, y_w, z_w)$  generate a two-dimensional image  $P_u(x_u, y_u)$  through the aperture diaphragm and spherical mirror. Through the beam splitting prism, two same 2d images are generated. Horizontal and vertical cylindrical lenses respectively compress the 2d images into 1d images, namely the X-direction image and Y-direction image. According to optical property of cylindrical lens, the beams parallel to the optical axis generate the image perpendicular to the generatrix direction. In order to realize imaging with a wide range, line array CCD's placement is parallel to the generatrix direction [13]. And horizontal and vertical line array CCDs receive the respective corresponding 1d image. Therefore, a 2d image is transformed to two orthogonal 1d images. The X-direction and Y direction of the line array CCDs respectively capture the image coordinate  $u$  and  $v$ . According to the two 1d image coordinates, 2d image coordinate of the target point  $I(u, v)$  can be restored as shown in Figure 2. Owing to the high resolution of the line array CCD, the 2d image has high resolution to improve the measurement precision of the system.

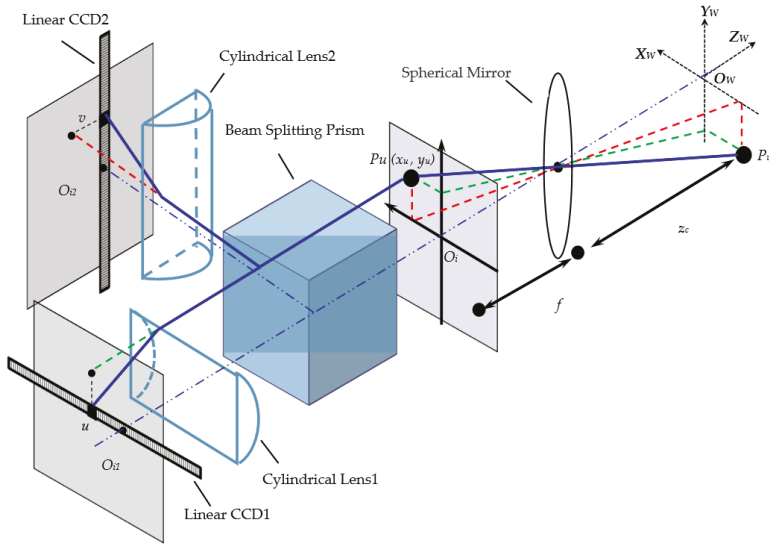


Figure 1. Structure of the orthogonally splitting imaging pose sensor.

### Image Coordinate Restore

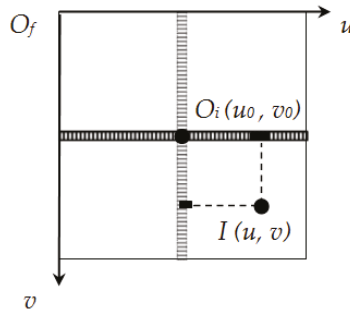


Figure 2. 2d image coordinate of the target point  $I(u, v)$  can be restored by two 1d images.

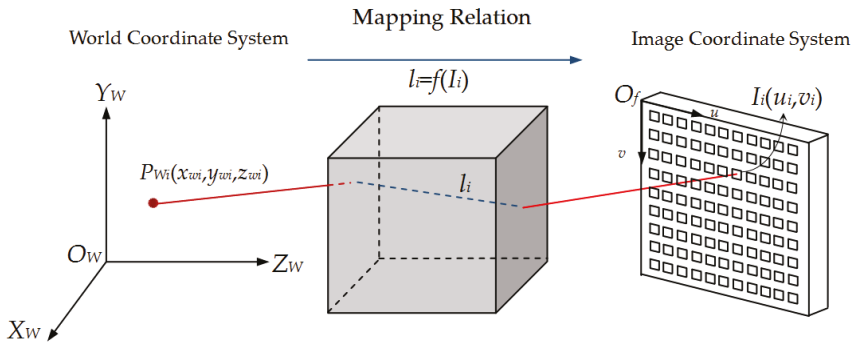
#### 2.1.2. Mapping Relation

In the general imaging model, the mapping relation between the target in a world coordinate system and a pixel in the image coordinate system can be described as a straight incoming line, as shown in Figure 3. Plücker coordinates can describe the incoming light with unique homogeneous coordinates in a five-dimensional projective space [14], as shown in Equation (1).

The coordinates of two points are  $(x_1, x_2, x_3)$  and  $(y_1, y_2, y_3)$  in a three-dimensional space. The line through the two points can be expressed by Plücker Coordinate as  $l_{ij} = x_i y_j - x_j y_i, x_0 = y_0 = 0$  Rewritten in vector form as

$$\vec{l} = (l_{01}, l_{02}, l_{03}, l_{23}, l_{31}, l_{12}) \tag{1}$$

Assuming the mapping relation is continuously changing, we can express the mapping relation by a vector valued function  $f(x)$ , which can be solved by radial basis function interpolation as shown in Figure 3.



**Figure 3.** General imaging model.  $P_{wi}$  is the world coordinate of the target point.  $I_i$  is the image coordinate corresponding to the target point.  $l_i$  is the mapping relation expressed by Plücker Coordinate.

2.1.3. Mathematical Model

Interpolation with radial basis function is used to solve the mapping relation. In order to guarantee positive definiteness and invertibility of interpolation matrix, the interpolation formula is improved as shown in Equation (2).

$$s(x) = a_0 + a_x^T x + \sum_{i=1}^M w_i \varphi(\|x - c_i\|) \tag{2}$$

In Equation (2)  $x$  represents an image coordinates  $I_i$ ,  $w_i$ ,  $a_0$ ,  $a_x$  are unknown parameters,  $\{c_i\} i = 1, \dots, M$  is a set of control points,  $\varphi$  is the radial basis function.

Rewriting Equation (2) in matrix form as

$$s(x) = \begin{pmatrix} \varphi(x) & p(x) \end{pmatrix} \begin{pmatrix} w \\ a \end{pmatrix} \tag{3}$$

In Equation (3),  $w$  and  $a$  are unknown parameters,  $\varphi(x)$  is the radial basis function expressed as  $\varphi_i(x) = \varphi(\|x - c_i\|)$ ,  $\{c_i\} i = 1, \dots, M$  is a set of control points,  $p(x)$  is homogeneous image coordinate expressed as  $p(x) = (1 \ u \ v)$ .

Supposing  $h_{wa} = \begin{pmatrix} w \\ a \end{pmatrix}$ , rewriting Equation (3) as

$$s(x) = \begin{pmatrix} \varphi(x) & p(x) \end{pmatrix} h_{wa} \tag{4}$$

Six independent interpolations based on radial basis function describe a vector-valued function  $f(x)$  as shown in Equation (5)

$$f(x) = (s_1(x), s_2(x), s_3(x), s_4(x), s_5(x), s_6(x))^T \tag{5}$$

Replacing  $s_i(x)$  in Equation (5) by Equation (4)

$$f(x) = \begin{pmatrix} \left( \begin{matrix} \varphi(x) & p(x) \end{matrix} \right) \underbrace{\begin{pmatrix} h_{wa}^{(1)} & \dots & h_{wa}^{(6)} \end{pmatrix}}_{H_{cam}} \end{pmatrix}^T \tag{6}$$

In Equation (6),  $H_{cam}$  is called camera matrix. Equation (6) can be simplified as Equation (7). And the mathematical model of the mapping relation can be expressed as Equation (7).

$$f(x) = \left( \begin{pmatrix} \varnothing(x) & p(x) \end{pmatrix} H_{cam} \right)^T \tag{7}$$

Based on the mathematical model, when the set of image points, the set of control points and camera matrix are known, the vector valued function can be solved for a given radial basis function. The set of image points and the set of control points can be gained by measuring, and camera matrix can be gained by calibrating.

## 2.2. Calibration

### 2.2.1. Mathematical Derivation

According to the mathematical model, the target point in the world coordinate system and the image in the image coordinate system should on the same line in Plücker coordinates.

According to the characteristics of the Plücker line [14], Equation (8) is satisfied if a point is on the Plücker line.

$$\underbrace{\begin{pmatrix} [w]_x & -I \\ 0^T & w^T \end{pmatrix}}_{Q(w)} f(x) = 0 \tag{8}$$

Substituting Equation (7) for Equation (8), there will be Equation (9).

$$Q(w) \left( \begin{pmatrix} \varnothing(x) & p(x) \end{pmatrix} H_{cam} \right)^T = 0 \tag{9}$$

According to the nature of the transposed matrix  $(AB)^T = B^T A^T$ , Equation (9) can be rewritten as Equation (10).

$$\underbrace{\begin{pmatrix} \varnothing(x)^T & p(x)^T \end{pmatrix}}_{r(x)} H_{cam} Q(w)^T = 0 \tag{10}$$

In Equation (10),  $\{w_i\} i = 1, \dots, N$  is the world coordinates set of target points,  $\{x_i\} i = 1, \dots, N$  is the image coordinates set of images.

According to the Kronecker product, Equation (10) is rewritten as Equation (11).

$$[Q(w) \otimes r(x)] \text{vec}(H_{cam}) = 0 \tag{11}$$

From the above equation, it can be concluded that if the image coordinates  $x$  and the world coordinates  $w$  are known, the camera matrix can be solved. If the image coordinates and world coordinates of  $N$  points are known, the above equation can be rewritten as Equation (12).

$$M \text{vec}(H_{cam}) = 0 \tag{12}$$

$$M = \begin{pmatrix} Q(w_1) \otimes r(x_1) \\ \vdots \\ Q(w_N) \otimes r(x_N) \end{pmatrix} \tag{13}$$

In order to exclude non-zero solutions, add constraints  $\| \text{vec}(H_{cam}) \| = 1$ . Then the problem of solving the camera matrix is transformed into solving the least squares solution [15] of the homogeneous equations. Find out  $\text{vec}(H_{cam})$  to gain the minimum value of  $\| M \text{vec}(H_{cam}) \|$ . According to the singular value decomposition method, the matrix  $M$  can be expressed as  $M = UD^T V$ , so  $\text{vec}(H_{cam})$  is the last column of matrix  $V$  based on derivation. Hence, if data set  $\{x_i \rightarrow w_i\} i = 1, \dots, N$  is known, the camera matrix  $H_{cam}$  can be solved out.

### 2.2.2. Experimental Apparatus

The experimental apparatus consists of a target, a peripheral component interconnect (PCI) controller (independent research and development) of the target, a motorized stage, a mechanical controller of the motorized stage, and an orthogonally splitting imaging pose sensor. Among them, the PCI controller of the target controls the LEDs, which are lightened up at different times, and the mechanical controller of the motorized stage controls the movement of the motorized stage. The experimental apparatus is shown in Figure 4. The main apparatus parameters are shown in the Table 1.

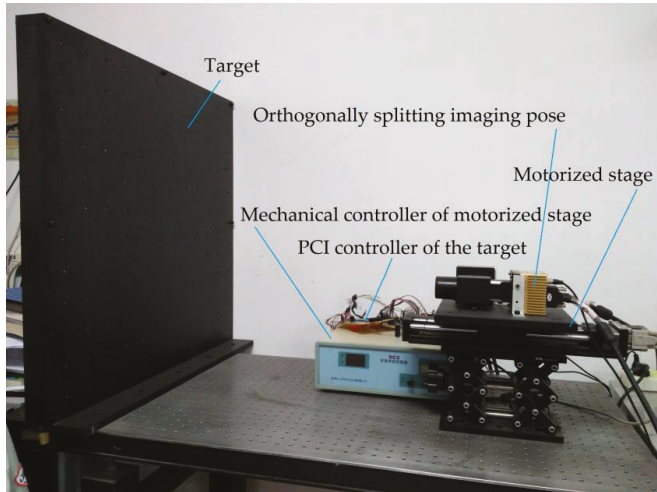


Figure 4. Experimental apparatus.

Table 1. Apparatus parameters.

Apparatus	Parameters
LED	The power is 1 W, the working current is 350 mA, and the working voltage is 3 ~3.8 V
Line array CCD	Resolution is 12,288 pixels, pixel size is 5 μm, mechanics is 76 × 76 × 56 mm <sup>3</sup>
Motorized stage	Ball screw drive mode, the guide rail adopts linear bearing, the resolution under the 8 subdivision is 2.5 μm, the maximum speed is 40 mm/s, and the repeated positioning accuracy is less than 5 μm

### 2.2.3. Data Acquisition Method

The data set is composed of N points' image coordinates  $x_i$  and N points' world coordinates  $w_i$ ,  $\{x_i \rightarrow w_i\} i = 1, \dots, N$ . The image coordinate of each point on the target can be measured by orthogonally splitting the imaging pose sensor and its numerical unit is pixel. The corresponding world coordinate can be acquired based on the distribution of points on the target. As shown in Figure 4, the horizontal distance between every two points is 60 mm, and the vertical distance between every two points is 60 mm. The point lying on the upper left corner of the target is set as the origin of the world coordinate system, then the x and y values of the world coordinate can be calculated. When setting the pose sensors in the initial position, the z value of the world coordinate is 0. On the motorized stage the pose sensor is driven to move away from the target for a distance of 20 mm, and the z value of the world coordinate is 20 mm. In this way the three-dimensional world coordinate can be obtained.

2.2.4. Calibration Procedure

Step one: Choose the appropriate radial basis function. Because radial basis function interpolation with shape parameters is better, this paper uses Multi-Quadric function (MQ function) with  $\phi(r) = (r^2 + \beta^2)^{1/2}$  and Gaussian function with  $\phi(r) = e^{-\frac{r^2}{\beta^2}}$  as the radial basis function to calibrate the camera, and  $\beta$  is the shape parameter.

Step two: The control points are adaptively selected based on the given data set using the Kmeans clustering method [16]. The radial basis function is an interpolation method of high-dimensional scattered data. The approximation and stability of the function are closely related to the distribution of control points. Therefore, the selection of control points is a very important part of the calibration method and directly affects the measurement accuracy. Thus the control points selected by Kmeans clustering method can represent the distribution characteristics of the data set.

Step three: Coordinate normalization. The radial basis functions chosen in this paper have shape parameters, and the choice of shape parameters depends on the distribution of the points in the coordinate system. In order to avoid the influence of the scale factor of the optical imaging system in terms of the coordinate values, the image coordinate and the world coordinate of each point must be normalized. The normalized data set is used for subsequent steps.

- (1) Normalization of the image coordinate. According to the affine transformation, the image coordinates  $x_i \ i = 1, \dots, N$  are normalized by  $u_i = \tau^{-1}(Ax_i + \alpha)$ .  $\tau$ ,  $A$ , and  $\alpha$  are normalized parameters which can be solved by Choleski decomposition method.
- (2) Normalization of the world coordinate. According to the affine transformation, the world coordinates  $w_i \ i = 1, \dots, N$  are normalized by  $q_i = \rho^{-1}(Bw_i + b)$ .  $\rho$ ,  $B$  and  $b$  are normalized parameters which can be solved by Choleski decomposition method.

Step four: Solve camera matrix  $H_{cam}$ . Calculate the calibration matrix  $M$  according to Equation (12), and then use the singular value decomposition method to solve the camera matrix  $H_{cam}$ .

Step five: Put the image coordinate  $x$  of any point and the solved camera matrix  $H_{cam}$  in Equation (7), the corresponding line coordinate can be fitted. Since the normalized coordinate values will change the coordinate values of the line space, the line coordinate obtained by using the interpolation should be converted to the original values according to Equation (14).

$$l^{(2)} = \begin{pmatrix} B & 0 \\ \rho^{-1}[b]_x B & \rho^{-1}det(B)B^{-T} \end{pmatrix} l^{(1)} \tag{14}$$

Step six: Repeat the fifth step until the line coordinate values of all points in the data set are fitted.

3. Results

3.1. Data Acquisition

Because of the experimental need, the calibration data set, control points set, and the test data set are acquired. The calibration data set is used to calibrate the pose sensor, and the control points set is selected from the calibration data set for calculating the radial basis function interpolation. The test data set is used to verify the accuracy of the calibration results.

3.1.1. Calibration Data Set Acquisition and Control Points Set Selection

According to the data acquisition method mentioned in Section 2.2.2, the calibration data set containing 358 points is acquired and the result is shown in Figure 5.

According to the Kmeans clustering method mentioned in Section 2.2.4, 179 control points are adaptively selected from the calibration data set. The selection results of control points set is shown in Figure 6.

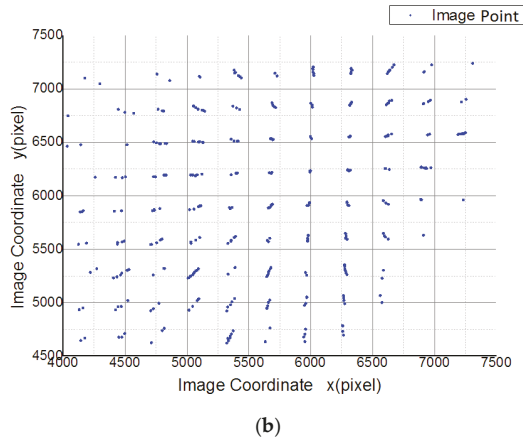
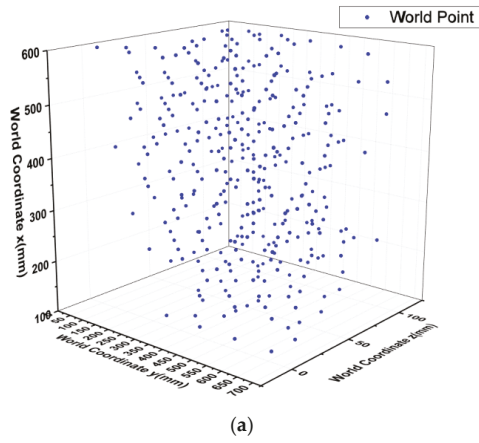


Figure 5. Calibration data set. (a) is the world coordinate of the calibration dataset. (b) is the image coordinate of the calibration dataset.

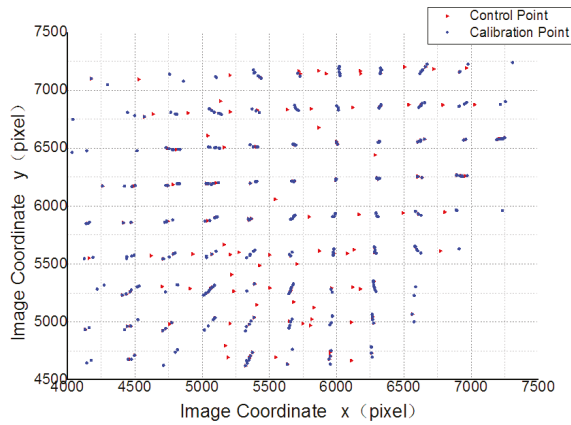
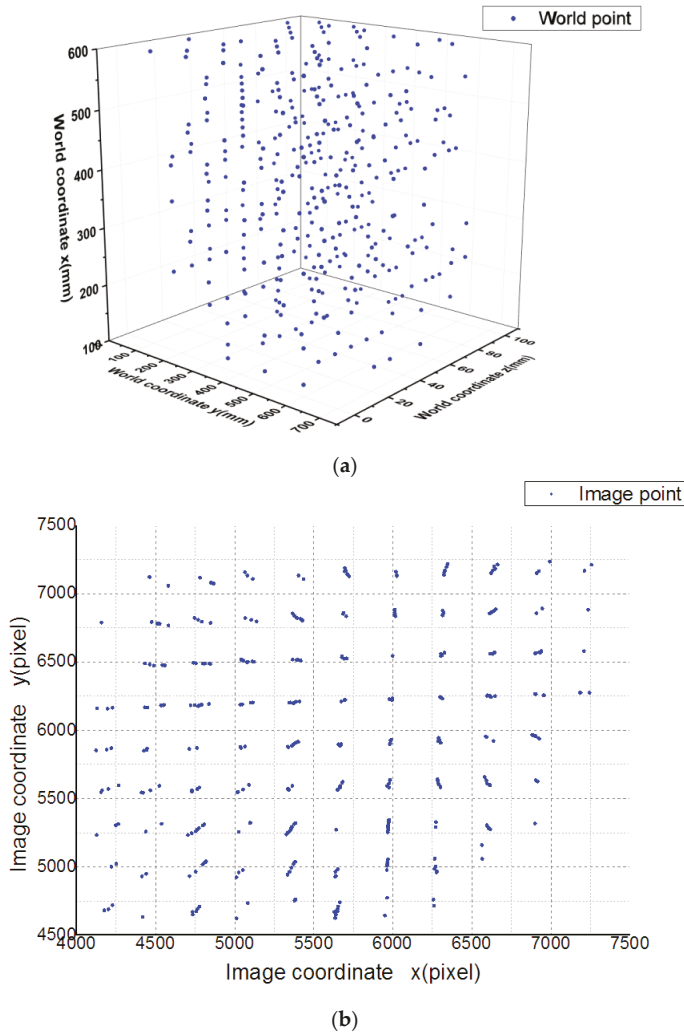


Figure 6. The selection results of control points set.



### 3.1.2. Test Data Set Acquisition

The acquisition method of test data set is the same as the acquisition method of the calibration set. The test data set containing 364 points is acquired and the result is shown in Figure 7. It is important that the test data set does not overlap with calibration data set.

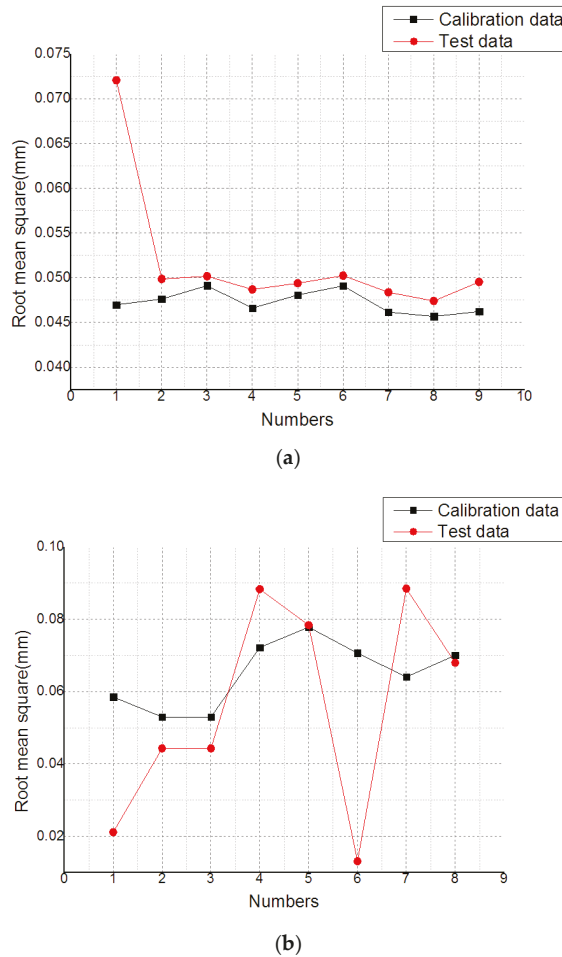


**Figure 7.** Test data set (a) is the world coordinate of the test dataset. (b) is the image coordinate of the test dataset.

### 3.2. Parameter Experiment

According to the calibration procedure proposed in Section 2.2.4, two different radial basis functions and different shape parameters are used to perform the interpolation. The calibration results are applied to the fitting calculation respectively for error evaluation of the calibration data and the test

data. The error evaluation is represented by the distance from the target point in the world coordinate system to the fitted Plücker line. The error evaluation is shown in Figure 8.



**Figure 8.** Analysis of different radial basis function interpolation and fitting accuracy with different shape parameters. (a) is the root mean square (RMS) calculated by using different shape parameters with different Multi-Quadric function (MQ function). (b) is the RMS calculated by using different shape parameters with different Gaussian function.

As can be seen from Figure 8, for the orthogonally splitting imaging pose sensor, when the MQ function is used as the radial basis function, the RMS of the calibration data set is similar to the RMS of the test data set. And when the Gaussian function is used as the radial basis function, the calibration result has a large difference between the calibration data set and the test data set. Therefore, the calibration accuracy of the MQ function is higher than the one of the Gaussian function.

From the experimental data, the shape parameters with the minimum difference between RMS of the test data set and RMS of the calibration data set are selected to calibrate the orthogonally splitting imaging pose sensor.

### 3.3. Calibration and Test

According to the selected shape parameter, the RMS of the calibration data set and test data set is shown in Table 2. It can be seen from the experimental results that the calibration method based on the general imaging model can meet the calibration requirements of the orthogonally splitting imaging pose sensor. And the MQ function is more suitable as an interpolation function.

**Table 2.** Error valuation of calibration dataset and test dataset.

	Calibration Dataset RMS (mm)	Test Dataset RMS (mm)
MQ function	0.048	0.049
Gaussian function	0.0778	0.0784

### 3.4. Comparison Experiment

The calibration method based on the general imaging model proposed in this paper and the calibration method based on the pinhole imaging model are applied to the orthogonally splitting imaging pose sensor, and the results of error valuation are compared.

The calibration method based on the general imaging model selects the MQ function and its shape parameter is selected according to the parameter experiment.

Since the pinhole imaging model is an ideal model and the pose sensor has a large distortion, the distortion correction is performed first, and then the Tsai two-step method is used for calibration. The first step uses a linear iterative optimization algorithm to calculate the initial values of the external parameters of the camera. The second step uses the optimization method to solve the internal parameters nonlinearly and further optimizes the external parameters. At last solve the world coordinate using parameters and calculate the RMS

The experimental results are shown in Table 3. It can be seen that the calibration accuracy based on the general imaging model is higher than the calibration accuracy based on the pinhole imaging model.

**Table 3.** Error valuation of general imaging model and pinhole imaging model.

	General Imaging Model	Pinhole Imaging Model
RMS (mm)	0.048	0.061

## 4. Discussion

The orthogonally splitting imaging pose sensor owns an autonomously designed optical imaging structure with assembly errors and various nonlinear and linear distortions. On the contrary, the pinhole imaging model cannot accurately describe the special optical imaging system. This paper proposes a calibration method based on general imaging models. The general imaging model introduces the Plücker coordinate to represent the mapping relationship between the world coordinate and the image coordinate and distortion does not need to be considered. It only cares about the mapping relationship between the incoming light and the pixel, and is suitable for the modeling of various optical systems. The method uses radial basis function interpolation to fit the mapping relationship between target points in the world coordinate system and image points in the image coordinate system. The calibration data requires that each target point in the world coordinate system only needs one corresponding image in the image coordinate system. Because the distribution of control points will affect the accuracy of radial basis function interpolation, this paper adopts Kmeans clustering method to adaptively select the control points set, which effectively improves the fitting accuracy. This paper uses the calibration method based on the general imaging model for the calibration of the orthogonally splitting imaging pose sensor. Through experiments, the appropriate radial basis function and its shape parameters are selected. The experimental results show that the MQ function

is more suitable for the orthogonally splitting imaging pose sensor as a radial basis function, RMS of calibration data set and RMS of test data set are 0.048 mm and 0.049 mm. Compared with the calibration method based on the pinhole imaging model, the calibration accuracy of the calibration method based on the general imaging model is higher. Therefore, the calibration method based on the general imaging model proposed in this paper is suitable for the calibration of the orthogonally splitting pose sensor. In the future, the calibration method based on the general imaging model can be applied to more types of optical imaging systems, and correspondingly interpolation methods should be explored to improve calibration accuracy. The application of the calibration method can be verified through more calibration experiments and test experiments.

**Author Contributions:** Conceptualization, C.S.; Writing-original draft, N.Z.; Writing-review & editing, P.W.

**Funding:** This research was funded by the National Natural Science Foundation of China grant number [No.51375339].

**Acknowledgments:** We thank Xintong Liu and Qian Yang for contribution in the preliminary work of the experiment.

**Conflicts of Interest:** The authors declare no conflicts of interest.

## References

1. Wang, P. Perspective-n-point pose measurement with two line array cameras. *Opt. Eng.* **2013**, *52*, 177–182.
2. Wu, J.; Wen, Q. The method of realizing the three-dimension positioning based on linear CCD sensor in general DSP chip. In Proceedings of the 2008 International Conference of the IEEE Engineering in Medicine and Biology Society, Vancouver, BC, Canada, 20–25 August 2008; IEEE: Piscataway, NJ, USA, 2008; pp. 2302–2305.
3. Ai, L.L.; Yuan, F.; Ding, Z.L. Measurement of spatial object's exterior attitude based on linear CCD. *Chin. Opt. Lett.* **2008**, *6*, 505–509.
4. Tsai, R.Y. A versatile camera calibration technique for high-accuracy 3D machine vision metrology using off-the-shelf TV cameras and lenses. *IEEE J. Robot. Autom.* **1987**, *3*, 323–344. [[CrossRef](#)]
5. Weng, J.; Cohen, P.; Herniou, M. Camera calibration with distortion models and accuracy evaluation. *IEEE Trans. Pattern Anal. Mach. Intell.* **2002**, *14*, 965–980. [[CrossRef](#)]
6. Zhang, Z. A flexible new technique for camera calibration. *IEEE Trans. Pattern Anal. Mach. Intell.* **2000**, *22*, 1330–1334. [[CrossRef](#)]
7. Tardif, J.P.; Sturm, P.; Trudeau, M.; Roy, S. Calibration of cameras with radially symmetric distortion. *IEEE Trans. Pattern Anal. Mach. Intell.* **2009**, *31*, 1552–1566. [[CrossRef](#)] [[PubMed](#)]
8. Li, W.; Li, Y.F. Single-camera panoramic stereo imaging system with a fisheye lens and a convex mirror. *Opt. Express* **2011**, *19*, 5855–5867. [[CrossRef](#)] [[PubMed](#)]
9. Grossberg, M.D.; Nayar, S.K. A general imaging model and a method for finding its parameters. In Proceedings of the Eighth IEEE International Conference on Computer Vision (ICCV 2001), Vancouver, BC, Canada, 7–14 July 2001; IEEE: Piscataway, NJ, USA, 2001; Volume 2, pp. 108–115.
10. Grossberg, M.D.; Nayar, S.K. The raxel imaging model and ray-based calibration. *Int. J. Comput. Vis.* **2005**, *61*, 119–137. [[CrossRef](#)]
11. Ramalingam, S.; Sturm, P. Minimal solutions for generic imaging models. In Proceedings of the 2008 IEEE Conference on Computer Vision and Pattern Recognition (CVPR 2008), Anchorage, AK, USA, 23–28 June 2008; IEEE: Piscataway, NJ, USA, 2008; pp. 1–8.
12. Miraldo, P.; Araujo, H.; Queiró, J. Point-based calibration using a parametric representation of the general imaging model. In Proceedings of the IEEE International Conference on Computer Vision, Computer Society, Barcelona, Spain, 6–13 November 2011; pp. 2304–2311.
13. Zhao, N.; Sun, C.; Wang, P.; Yang, Q.; Liu, X. Method of orthogonally splitting imaging pose measurement. In Proceedings of the International Conference on Optical Instruments and Technology 2017: Optoelectronic Measurement Technology and System, Beijing, China, 12 January 2018; p. 8.
14. Mann, C.J.H. *Geometric Fundamentals of Robotics Second Edition*; Springer: New York, NY, USA, 2005.

15. Quak, E.; Sivakumar, N.; Ward, J.D. *Least Squares Approximation by Radial Functions*; Society for Industrial and Applied Mathematics: Philadelphia, PA, USA, 1993.
16. Napoleon, D.; Pavalakodi, S. A new method for dimensionality reduction using k-means clustering algorithm for high dimensional data set. *Int. J. Comput. Appl.* **2011**, *13*, 41–46.



© 2018 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).

Article

# Non-Contact and Real-Time Measurement of Kolsky Bar with Temporal Speckle Interferometry

Shanwei Yang<sup>1</sup>, Zhan Gao<sup>1,\*</sup>, Haihui Ruan<sup>2</sup> , Chenjia Gao<sup>1</sup>, Xu Wang<sup>1</sup>, Xiang Sun<sup>1</sup> and Xin Wen<sup>1</sup>

<sup>1</sup> Key Laboratory of Luminescence and Optical Information of Ministry of Education, Beijing Jiao tong University, Beijing 100044, China; 15121635@bjtu.edu.cn (S.Y.); 16118443@bjtu.edu.cn (C.G.); 17118455@bjtu.edu.cn (X.W.); 16126101@bjtu.edu.cn (X.S.); 14126048@bjtu.edu.cn (X.W.)

<sup>2</sup> Department of Mechanical Engineering, The Hong Kong Polytechnic University, Hung Hom, Kowloon, Hong Kong, China; haihui.ruan@polyu.edu.hk

\* Correspondence: zhangaob@bjtu.edu.cn; Tel.: +86-10-5168-8333

Received: 17 April 2018; Accepted: 15 May 2018; Published: 17 May 2018



**Abstract:** In this paper, a new non-contact and real-time measurement system for Kolsky bars is presented. This system uses two sets of temporal speckle interferometry in-plane displacement measurement devices to replace two strain gauges of conventional Kolsky bars. The in-plane displacement measurement of the Kolsky bar is mainly intended to provide a new test method for the dynamic mechanical properties of small-size material samples with diameters below 2 mm. This method is non-contact, does not require any intermediate medium, and can make the Kolsky bar applicable to characterizing the dynamic mechanical properties of materials under higher strain rates and smaller size conditions. The measuring devices and principles are described. In addition, a preliminary experiment is carried out to demonstrate the performance of this new device.

**Keywords:** Kolsky bar; speckle; in-plane displacement measurement; wavelet transform; dynamic mechanical properties

## 1. Introduction

In the fields of aviation, aerospace, packaging, transportation, and other areas of military and civil engineering, the materials used will encounter stress events such as explosions, rapid collision, and other impact load. It is of great practical significance to study the dynamic mechanical properties of materials under high strain rate conditions for engineering design and application. The Kolsky bar (also known as Split Hopkinson Pressure Bar) apparatus is the most popular and convenient system for characterizing the high strain-rate behaviors of materials [1].

The typical Kolsky bar is using two strain gages that past on the incident bar and transmitter bar to measure the dynamic mechanical properties of a material under high strain rate conditions. This method requires good adhesion between the bar and the strain gauge. Apart from this, the measurement accuracy that the conventional Kolsky bar provides is limited. The support of real-time and high-accuracy dynamic mechanical properties of materials under high strain rate conditions is required with the rapid development of industry, especially for the machinery defense industries. For example, it is useful to employ bars with diameters of only a few millimeters for testing at higher strain-rates. However, traditional strain-gauge measurements of the longitudinal waves within the bars become impractical at these sizes, because the strain gauge and the bar cannot be reliably connected when the diameter of the bar is less than 3 mm [2].

Avinadav and Ashuach proposed the application of a fiber-based velocity interferometry measurement method to directly measure the actual velocities of the bars. The two fiber lasers

are irradiated onto the bars at a small angle, and then the Doppler-shifted light carrying the optical information is collected [3]. The dynamic stress–strain curves of the material are obtained from the measured velocity information by short-time Fourier transform and phase-based analysis. However, because the Kolsky bar is characterized by rapidly changing velocities, their system contains a  $3 \times 3$  single-mode fiber coupler to create three interferometric singles which are  $120^\circ$  out of phase relative to each other [4,5]. For a single velocity trace, it is necessary to manually choose the transition points between high and low velocity.

Casem and Grunschel proposed a method to permit wave separation of a Kolsky bar. The particle velocity at each strain gauge position is measured with PDV. This method can measure the strain of the scattered wave train present on each bar, which can extend the test time of the experiment so that more types of loads can be loaded into the specimen [6]. However, the data obtained by measuring the particle velocity with PDV will contain a significant bending wave component due to the errors introduced by the leading of wire motion. In order to account for bending, it is necessary to multiple the PDV measurements at each gauge position.

Casem and Zellner proposed TDI (transverse displacement interferometer) and NDI (normal displacement interferometer) for application to Kolsky bars. TDI is used to measure the displacement of the mid-point of the incident bar and provide measurements of the incident and reflected pulses. Similarly, NDI is used to measure the displacement of the free-end of the transmitter bar and provide a measurement of the transmitted pulse [7]. However, it is necessary to install a diffraction grating in the middle of the incident bar when measuring the incident pulse and the reflected pulse by TDI. The machining precision of the diffraction grating and the quality of the paste with the incident bar will have significant influence on the measurement accuracy of the whole experiment.

Fu and Tang used a fiber micro-displacement interferometer system for any reflector to directly measure the actual velocities of the bars. The main feature of this method was that the interfacial velocity profiles between the bars and specimen were determined instead of the reflected and transmitted strain pulse in the middle of bars [8]. However, before the experiment, the shutters must be designed with reflector slices, and the reflector slices must be fixed to the front and back ends of the specimen.

In this paper, we apply temporal speckle interferometry in-plane displacement measurement to a Kolsky bar. This method uses two sets of time-domain speckle interference in-plane displacement measurement devices to replace the strain gauges of a conventional Kolsky bar. Moreover, the wavelet transform, which is believed to be more suitable to process the non-stationary signal than the Fourier transform [9], is used to extract the dynamic mechanical properties of the measured specimen. This new method does not require any intermediate media, nor the pasting of devices, and has the advantages of being non-contact, real-time, and more reliable. Furthermore, the Kolsky bar can measure the dynamic mechanical properties of small specimens under higher strain rate conditions in this method. To the best of our knowledge, no such system has yet been presented.

## 2. Principles

### 2.1. Principle of the Kolsky Bar Based on the Velocity Method

As shown in Figure 1, a typical Kolsky bar includes a striker bar, incident bar, transmitter bar, specimen, and strain gauge 1 and strain gauge 2 which are pasted at  $X_A$  of the incident bar and  $X_B$  of the transmitter bar, respectively.

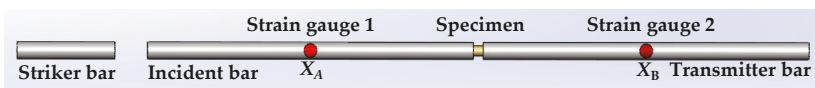


Figure 1. A typical Kolsky bar.

Before the measurement, the specimen is placed between the incident bar and transmitter bar. Using a hammer or other tools, the strike bar is driven so that it hits the incident bar at a certain speed, thereby generating an incident wave. The specimen is deformed at high speed under the loading of the incident wave, propagates the reflected wave into the incident bar, and propagates the transmitted wave into the transmitter bar. The incident wave and the reflected wave are recorded by strain gauge 1, and the transmitted wave is recorded by strain gauge 2. The dynamic mechanical properties of the specimen can be solved from the records of these waves by theoretical calculation.

If the length of the strike bar is  $L_0$ , the impact velocity is  $v$ , then its duration  $T_M$  is:

$$T_M = \frac{2L_0}{C_0} \tag{1}$$

The magnitude of the incident pulse is:

$$A_M = \frac{\rho_0 C_0 v}{2} \tag{2}$$

where  $\rho_0$  is the density of the bars.  $C_0$  is the longitudinal speed of wave in the bar, and can be expressed as:

$$C_0 = \left(\frac{E_0}{\rho_0}\right)^{1/2} \tag{3}$$

where  $E_0$  is the elastic modulus of the bars.

If the four parameters include stress  $\sigma(X_1, t)$  and  $\sigma(X_2, t)$ , particle velocity  $v(X_1, t)$  and  $v(X_2, t)$  can be measured, then the dynamic mechanical properties of the specimen can be calculated according to Formula (4)–(6).

$$\sigma(t) = \frac{A_0}{2A_S} [\sigma(X_1, t) + \sigma(X_2, t)] = \frac{A_0}{2A_S} [\sigma_I(X_1, t) + \sigma_R(X_1, t) + \sigma_T(X_2, t)] \tag{4}$$

$$\dot{\epsilon}(t) = \frac{v(X_2, t) - v(X_1, t)}{L_S} = \frac{v_T(X_2, t) - v_I(X_1, t) - v_R(X_1, t)}{L_S} \tag{5}$$

$$\epsilon(t) = \int_0^t \dot{\epsilon}(t) dt = \frac{1}{L_S} \int_0^t [v_T(X_2, t) - v_I(X_1, t) - v_R(X_1, t)] dt \tag{6}$$

where  $\sigma(t)$  is stress,  $\dot{\epsilon}(t)$  is strain rate,  $\epsilon(t)$  is strain.  $L_S$  is the length of the specimen.  $A_0$  and  $A_S$  are the cross-sectional area of the bar and specimen, respectively.  $\sigma_I(X_1, t)$  is incident wave,  $\sigma_R(X_1, t)$  is reflected wave, and  $\sigma_T(X_2, t)$  is transmitted wave.  $v_I(X_1, t)$ ,  $v_R(X_1, t)$  and  $v_T(X_2, t)$  are incident velocities, reflected velocities, and transmitted velocities, respectively.

According to the assumption, both the one-dimensional stress waves of the bar and the short specimen strain are evenly distributed over its length, and the three waves have the following relationship:

$$\epsilon_I + \epsilon_R = \epsilon_T \tag{7}$$

where  $\epsilon_I$ ,  $\epsilon_R$  and  $\epsilon_T$  is strain corresponding with incident pulse, reflected pulse, and transmitted pulse, respectively.

Moreover, there is a linearly proportional relationship between strain, stress, and particle velocity [10] in the elastic region of the Kolsky bar:

$$\sigma(X_1, t) = \sigma_I(X_1, t) + \sigma_R(X_1, t) = E_0[\epsilon_I(X_1, t) + \epsilon_R(X_1, t)] \tag{8}$$

$$\sigma(X_2, t) = \sigma_T(X_2, t) = E_0\epsilon_T(X_2, t) \tag{9}$$

$$v(X_1, t) = v_I(X_1, t) + v_R(X_1, t) = C_0[\epsilon_I(X_1, t) + \epsilon_R(X_1, t)] \tag{10}$$

$$v(X_2, t) = v_T(X_2, t) = C_0\epsilon_T(X_2, t) \tag{11}$$



In summary, the stress, strain, and strain rate of the specimen can be expressed [11] as follows:

$$\sigma(t) = \frac{E_0}{C_0} \frac{A_0}{A_S} v_T(t) \tag{12}$$

$$\varepsilon(t) = -\frac{2}{L_S} \int_0^t v_R(t) dt \tag{13}$$

$$\dot{\varepsilon}(t) = -\frac{2}{L_S} v_R(t) \tag{14}$$

where  $v_T$  and  $v_R$  are transmitted and reflected velocities as measured, respectively.

2.2. Principle of Temporal Speckle Interferometry In-Plane Measurement

The principle of the in-plane displacement measurement used in this article is shown in Figure 2. When the object moves in the z-direction, the optical paths of both the reflected light and transmitted light change consistently [12]. Therefore, the speckle remains unchanged.

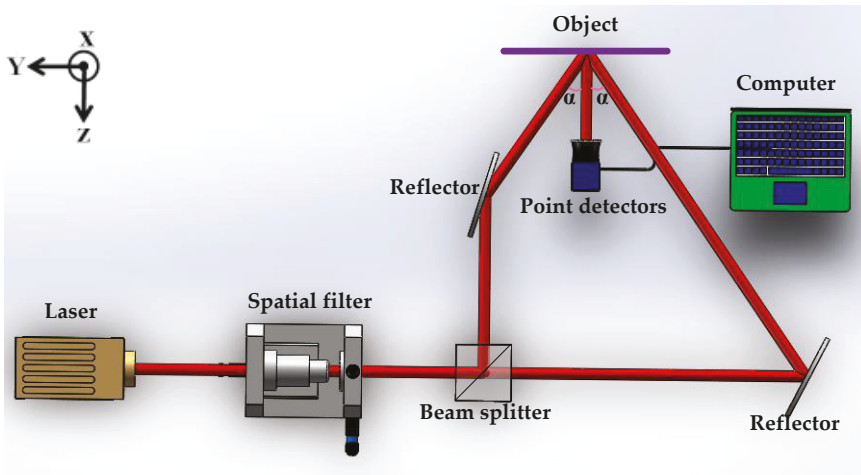


Figure 2. The optical path of in-plane displacement measurement.

When the object moves the  $y$ -direction, one of the optical path increases  $\Delta y(x, z, t) \sin \alpha$  and the other reduces  $\Delta y(x, z, t) \sin \alpha$ . Therefore, the optical path difference is as follows:

$$\delta = 2\Delta y(x, z, t) \sin \alpha \tag{15}$$

The phase change of the reflected light and the transmitted light is:

$$\Delta \phi_r = \frac{2\pi}{\lambda} \delta = \frac{4\pi}{\lambda} \Delta y(x, z, t) \sin \alpha \tag{16}$$

where  $\lambda$  is the wavelength.  $\alpha$  is the incident angle. Before the motion of the object, the intensity function of the speckle is:

$$I(x, y, t) = I_0(x, y) \{1 + V \cos[\Phi_0(x, y)]\} \tag{17}$$

where  $I_0$  is the average intensity of the interference field.  $V$  is the modulation visibility.  $\Phi_0(x, y)$  is the initial phase. Hence, after the motion of the object, the intensity function of the speckle is:

$$I(x, y, t) = I_0(x, y) \{1 + V \cos[\Phi_0(x, y)] \pm 4\pi \Delta y(x, z, t) \sin \alpha / \lambda\} \quad (18)$$

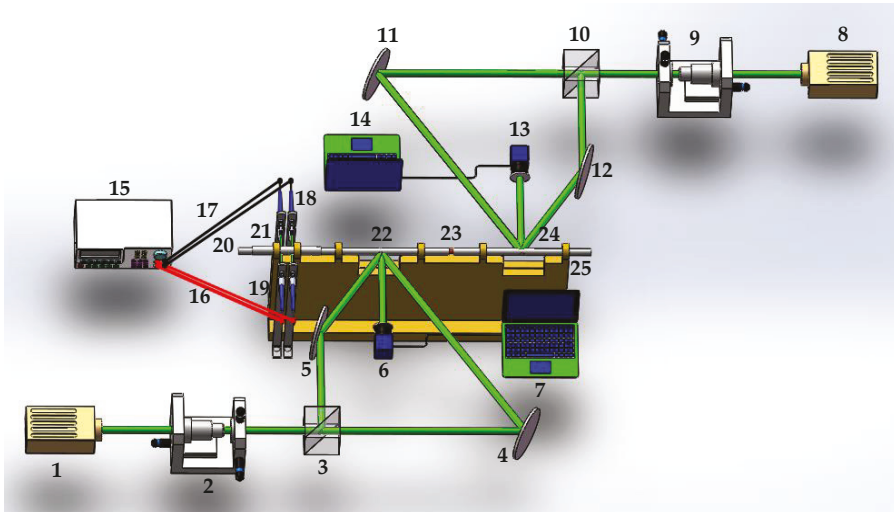
where  $\pm$  is the direction of the object movement.

### 2.3. Principle of a Kolsky Bar Based on Temporal Speckle Interferometry

As showing in Figure 3, a Kolsky bar based on in-plane displacement measurement combines two sets of in-plane displacement measuring systems and a Kolsky bar. The beam emitted from the laser is expanded and filtered by the spatial filter, and then the incident light is divided into two beams by the beam splitter. Both of them illuminate the surface of the incident bar or transmitter bar at the same angle. The back-scattered beams from the surface of the incident bar or transmitter bar form a speckle. The intensity of the speckle will change if the surface moves in the  $y$ -direction, which is captured by the detectors. The collected information is converted into a series of voltage values by the data acquisition device. By processing these voltage data with the wavelet transform and the phase unwrapping algorithm, the displacement of the incident bar or transmitter bar can be obtained according to the following formula,

$$\Delta x(t) = \frac{\lambda \Delta \Phi(t)}{4\pi \sin \alpha} \quad (19)$$

where  $\Delta \Phi(t)$  is a continuous phase based on time variation.



**Figure 3.** Principle of a Kolsky bar based on in-plane measurement; 1 and 8 are lasers; 2 and 9 are spatial filters; 3 and 10 are beam splitters; 4, 5, 11, and 12 are reflectors; 6 and 13 are detectors; 7 and 14 are computers; 15 is an oscilloscope; 16 and 17 are optical fibers; 18 and 19 are fiber detectors; 20 is the striker bar; 21 is the guide; 22 is the incident bar; 23 is the specimen; 24 is the transmitter bar; 25 is the support frame.

The velocity record is derived from the derivative of the displacement after Gaussian fitting. Therefore, the stress, strain, and strain rate of the specimen can be calculated according to the Formulae (12)–(14).

2.4. Signal Processing Procedure

Wavelet transform (WT) is a new transformation analysis method. WT inherits and develops the thought of short-time Fourier transformed localization, while overcoming shortcomings such as the window size not changing with frequency. It can automatically adapt to the requirements of time-frequency signal analysis and can focus on any detail of the signal to solve the difficult problem of the Fourier transform. It is the ideal tool for signal time-frequency analysis and processing.

The continuous wavelet coefficient [13] can be expressed as

$$W_f(m, n) = |m|^{-1/2} \int f(x)\psi^*\left(\frac{x-n}{m}\right)dx \tag{20}$$

where  $m$  is the scale parameter,  $n$  is the shift parameter,  $f(x)$  is the signal to be analyzed,  $\psi(x)$  is the mother wavelet, and  $\psi^*(x-n)/m$  is its conjugate function. The amplitude of  $W_f(m, n)$  is positively correlated with the similarity of the mother wavelet and the signal [14]. In this paper, Gaussian wavelet is chosen as the mother wavelet.

The amplitudes of the light field are given by the expressions,

$$A(m, n) = \sqrt{\left\{Im\left[W_f(m, n)\right]\right\}^2 + \left\{Re\left[W_f(m, n)\right]\right\}^2} \tag{21}$$

where  $Im\left[W_f(m, n)\right]$  is the imaginary part of  $W_f(m, n)$ ,  $Re\left[W_f(m, n)\right]$  is the real part of  $W_f(m, n)$ .

Hence, the phase is retrieved from the following equation:

$$\Phi(m, n) = \tan^{-1}\left\{\frac{Im\left[W_f(m, n)\right]}{Re\left[W_f(m, n)\right]}\right\} \tag{22}$$

Then, through phase unwrapping, the deformation information of the measured object is obtained [15]. According to Equation (19), the object displacements can be obtained. The whole signal processing flowchart is shown in Figure 4.

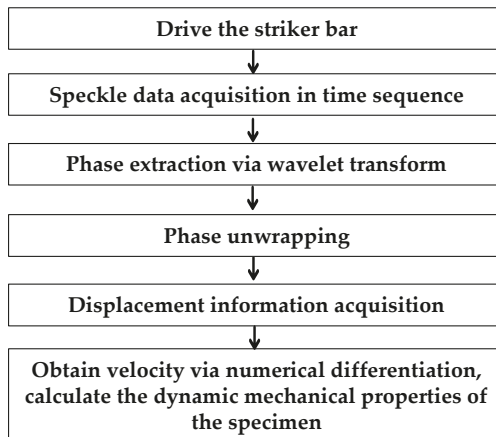


Figure 4. Signal processing flowchart.

### 3. Experiment and Results

The Kolsky bar based on in-plane displacement measurement is mainly intended to provide a new test method for the dynamic mechanical properties of small-size material samples. Therefore, the Kolsky bar is only 2 mm in diameter in the measurement system. A physical map of the guide, striker bar, transmitter bar, incident bar, and the sample is shown in Figure 5, and their sizes are shown in Table 1.

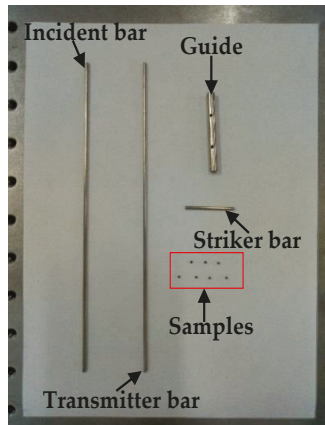


Figure 5. Physical map of each component in Kolsky bar.

Table 1. The size of each component in Kolsky bar.

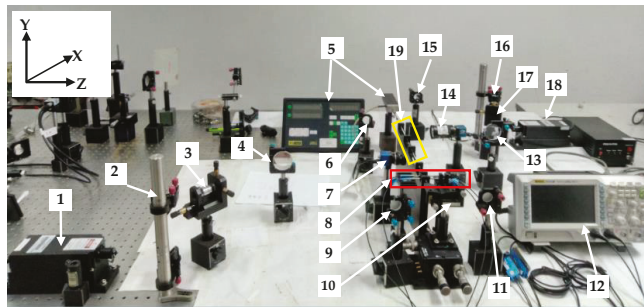
Part Name	Diameter (mm)	Length (mm)
Striker bar	2	30
Incident bar	2	200
Transmitter bar	2	200
Sample	1.5	0.75
Guide	3	50

#### 3.1. Verification Experiment of In-Plane Displacement Measurement Capability

According to Equations (12)–(14), the dynamic mechanical properties of the specimen depend on the velocity of the incident and transmitter bars, and the velocity is obtained by the displacement of the two bars after numerical differentiation, therefore, it is necessary to verify the displacement measurement capability of the constructed optical path. In other words, if the measurement accuracy of displacement can be guaranteed, then the dynamic mechanical properties of the sample can be considered accurate.

The verification experiment setup is shown in Figure 6. A single longitudinal mode green laser with a wavelength of 532 nm was used as the light source. The coherence length of the laser is greater than 50 m, and the beam diameter at the aperture is less than 1.5 mm. The detector (MER-030-120UC) is a color GigE Vision CCD camera (IMAVISION Company, Beijing, China) with a resolution of 656 × 492 pixels. The pixel size is 5.6 μm × 5.6 μm, and the numerical aperture NA of camera lens is 0.052 in the image plane, which gives a speckle size in the image plane of 6.1 μm according to Formula (23) [16],

$$\sigma = \frac{1.2\lambda}{2NA} \tag{23}$$



**Figure 6.** Verification experiment setup: 1 and 18 are green lasers; 2 and 16 are beam lifters; 3 and 17 are spatial filters; 4 and 13 are beam splitters; 5 is a grating ruler; 6, 9, 11, and 15 are reflectors; 7 and 14 are CCD; 8 is the trigger device; 10 is the translation stage; 12 is an oscilloscope; 19 is the SHPB device.

The translation stage (Thorlabs, NJ, USA) consists of the MAX302(/M) 3-Axis NanoMax Flexure Stage and three DRV3 Differential Micrometers, with a calculated fine resolution of 100 nm. The accuracy of the grating ruler is 0.1  $\mu\text{m}$ .

What needs to be emphasized here is that since the entire measurement is based on the speckle, it is necessary to determine whether the optical path that we have built can form speckle interference fringes. Since the diameter of the Kolsky bar we used here is very small, the speckle interference fringes of both the incident and transmitter bar cannot be observed directly even if it does exist, two iron plates of dimensions 40 mm  $\times$  30 mm  $\times$  2 mm were installed in the incident and transmitter bars to verify the optical path. In Figure 7, the left part is the original structure of the experimental setup, the incident bar and the transmitter bar are placed on the support base. The right part of the figure is the structure using two iron plates to replace the incident and the transmitter bars to verify the existence of speckle interference fringes. The iron plate is fixed to the support base by two screws. In the measurement process, applying a force along the  $x$ -direction to the iron plate, the surface of the iron plate will have a slight deformation in the  $x$ -direction. The CCD camera will record the entire process of the iron plate before and after the deformation. If clear speckle interference fringes can be seen from the processed image, it is proved that the optical path to be debugged meets the measurement conditions. It is noteworthy that the position of the spot on the iron plate is the same as the position on the incident and transmitter bars; this condition has been achieved when designing the support base structure of the Kolsky bar. This ensures that the replacement between the iron plate and the bar is only for more convenient observation of the existence of speckle interference fringes, it is not related to the accuracy of the system displacement measurement capability, nor the robust position of the measuring beams.

The observation result of the speckle interference fringes is shown in Figure 8. The grey-scale maps of fringes in the incident and the transmitter bars can be observed by real-time subtraction. From this we can prove that the measurement light path meets the speckle interference measurement conditions.

After the speckle verification is completed, we replaced the iron plate with the incident and transmitter bars continue to verify the in-plane displacement measurement capability of the measurement system.

To accurately test the in-plane displacement measurement capability of the measuring system, the verification experiment was performed without a sample and the striker bar was driven with a very low velocity of a few microns per second. The translation stages push the striker bar at a slow velocity so that it strikes the incident bar at a certain velocity. The intensities of the speckles on the incident bar and transmitter bar are changed in the time domain. These changes are collected by the CCD camera. The frame rate of the CCD camera is 120 frames/s. The whole measuring time is five seconds and an arbitrary point on the position of incident bar and transmitter bar is chosen as the test point, the

region of interest which is chosen arbitrarily is shown in Figure 9. At the same time, the grating ruler at the end of the transmitter bar will record the final displacement of the bar as a standard to verify the in-plane displacement detection capability of the test system.

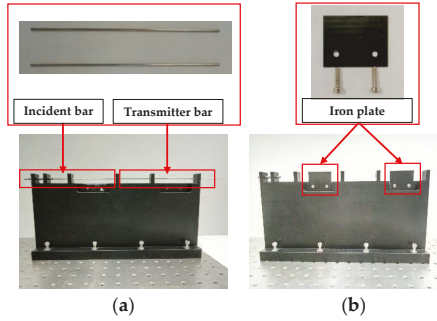


Figure 7. Iron plate instead of incident bar and transmitter bar: (a) Original structure; (b) Replacement structure.

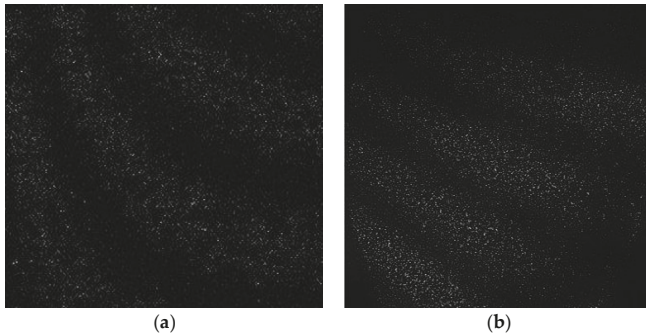


Figure 8. Speckle verification results from the position of incident bar and the transmitter bar: (a) Speckle on the incident bar; (b) Speckle on the transmitter bar.

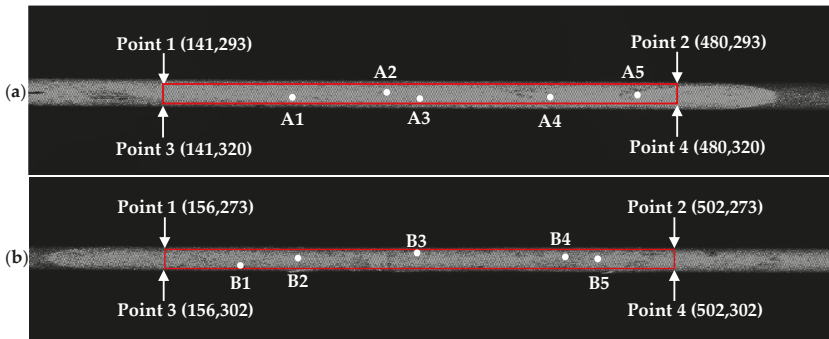
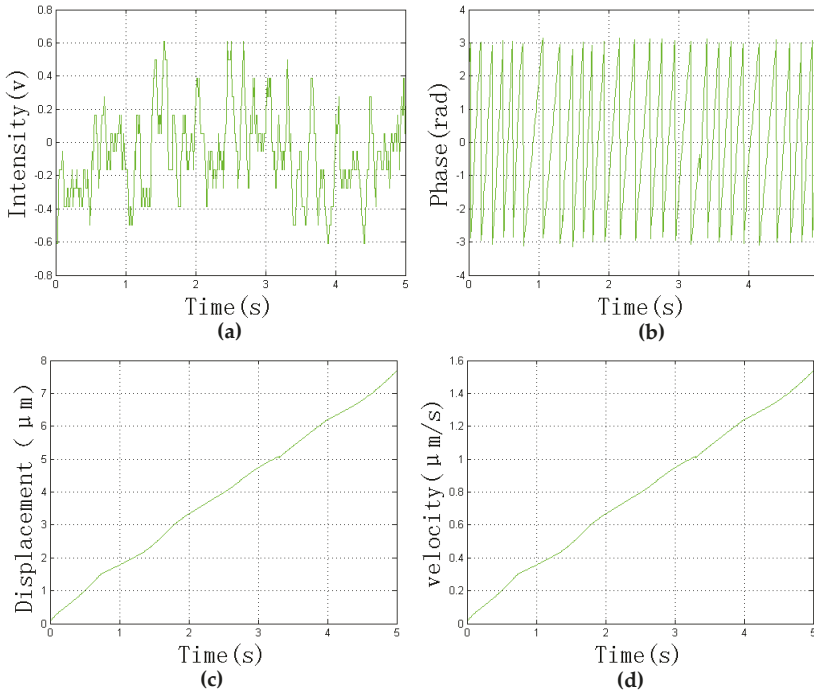


Figure 9. The region of interest: (a) incident bar; (b) transmitter bar.

In the verification experiment, the incident angle of the incident and transmitter bars is 70 degrees and 72.5 degrees, respectively. The actual displacement of the grating ruler is 7.7  $\mu\text{m}$ , and the actual average velocity of the bar is 1.54  $\mu\text{m}/\text{s}$ .

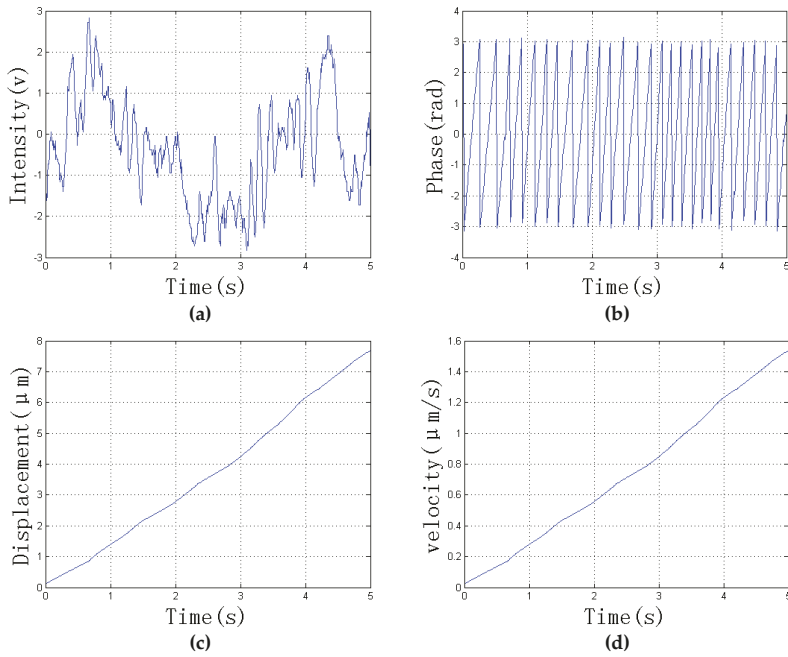
The displacement of the incident bar obtained by the CCD camera is shown in Figure 10. As can be seen from the figure, the displacement processed by our algorithm is 7.68  $\mu\text{m}$  and the average velocity is 1.536  $\mu\text{m}/\text{s}$ . Therefore, the measured error of displacement is 0.02  $\mu\text{m}$  and the measured error of velocity is 0.004  $\mu\text{m}/\text{s}$ . The relative error of displacement and velocity are both 0.26%.



**Figure 10.** The in-plane displacement of the incident bar: (a) the intensity distribution over time; (b) the discontinuous phase map obtained by wavelet transformation; (c) the measured displacement over time; (d) average speed.

The displacement of the transmitter bar obtained by the CCD camera is shown in Figure 11. As can be seen from the figure, the displacement processed by our algorithm is 7.68  $\mu\text{m}$  and the average velocity is 1.535  $\mu\text{m}/\text{s}$ . Therefore, the measured error of displacement is 0.02  $\mu\text{m}$  and the measured error of velocity is 0.005  $\mu\text{m}/\text{s}$ . The relative error of displacement and velocity are 0.26% and 0.32%, respectively.

In order to verify the repeatability of the measurement system, we have also designed the system for multi-detection points processing. The measured results of two bars are shown in Table 2, where A1 to A5 are the points on the incident bar, and B1 to B5 are the points on the transmitter bar, as shown in Figure 9. From Table 2 we can see the displacement error of the incident bar and transmitter bar are less than 0.04  $\mu\text{m}$  and 0.05  $\mu\text{m}$  and the velocity error is less than 0.008  $\mu\text{m}/\text{s}$  and 0.006  $\mu\text{m}/\text{s}$ , respectively.



**Figure 11.** The in-plane displacement of the transmitter bar: (a) the intensity distribution over time; (b) the discontinuous phase map obtained by wavelet transformation; (c) the measured displacement over time; (d) average speed.

**Table 2.** The comparison of two measurement results of multi-detection points.

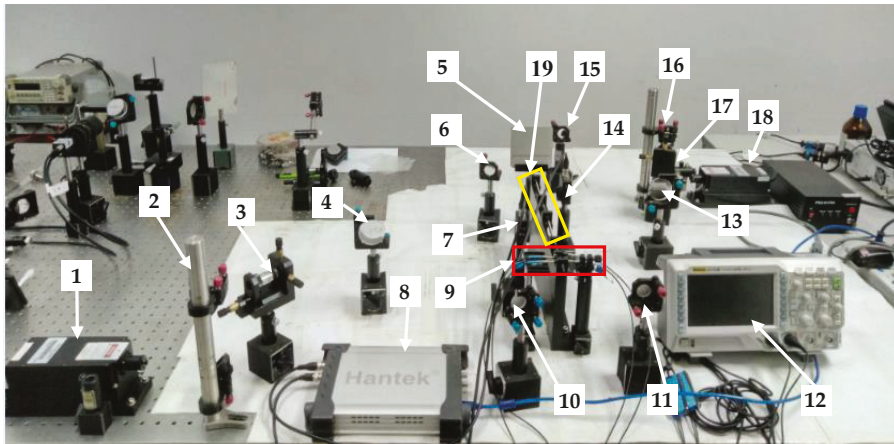
Actual Value	Coordinate	Displacement (μm)			Velocity(μm/s)		
		Value	Error	Relative Error	Value	Error	Relative Error
Displacement = 7.7 μm Velocity = 1.54 μm/s	A1(265,303)	7.69	0.01	0.13%	1.538	0.002	0.13%
	A2(333,296)	7.74	0.04	0.52%	1.548	0.008	0.52%
	A3(347,313)	7.71	0.01	0.13%	1.542	0.002	0.13%
	A4(419,311)	7.71	0.01	0.13%	1.542	0.002	0.13%
	A5(456,309)	7.66	0.04	0.52%	1.532	0.008	0.52%
	B1(211,302)	7.73	0.03	0.39%	1.546	0.006	0.39%
	B2(255,286)	7.65	0.05	0.65%	1.530	0.010	0.65%
	B3(380,273)	7.72	0.02	0.26%	1.544	0.004	0.26%
	B4(444,283)	7.71	0.01	0.13%	1.542	0.002	0.13%
	B5(465,294)	7.71	0.01	0.13%	1.542	0.002	0.13%

As can be seen above, the results of these verifications experiment prove that the measurement system has a very good in-plane displacement measurement capability, and the algorithm we developed can accurately calculate the information of displacement and velocity.

### 3.2. Experimental Determination of the Dynamic Mechanical Properties of Materials

The experimental setup for the determination of the dynamic mechanical properties of materials is shown in Figure 12.





**Figure 12.** Dynamic mechanical properties of materials experimental setup: 1 and 18 are green lasers; 2 and 16 are beam lifters; 3 and 17 are spatial filters; 4 and 13 are beam splitters; 5 is a baffle; 6, 10, 11 and 15 are reflectors; 7 and 14 are point detectors; 8 is a virtual oscilloscope; 9 is the trigger device; 12 is an oscilloscope; 19 is the SHPB device.

The specimen described here is made of 304 L stainless steel, with diameter and thickness of 1.5 mm and 0.75 mm, respectively. The sample is held between the incident bar and the transmitter bar with butter. The pole longitudinal wave velocity  $C_0$  in the bar is 4934 m/s. In order to improve the detection of changes in light intensity, here, we use a high-speed photodetector as the detector device. The high-speed photodetector is Thorlabs' Biased Photo-detectors (DET36A/M, Thorlabs, NJ, USA), a Si biased detector covering the 350 nm to 1100 nm wavelength range and rise times as fast as 1 ns, as shown in Figure 13. A virtual oscilloscope (DSO3062AL) produced by Hantek Company (Hantek, Qindao, China) was used to collect data.



**Figure 13.** Photodetector.

It needs to be emphasized that after the verification experiment was completed, we just removed the translation stage and the grating ruler, and the light path did not change. Therefore, the incident angle of the incident bar is still 70 degrees, and the incident angle of the transmitter bar is still 72.5 degrees.

The typical original signal for the experiment is shown in Figure 14, where the green curve is the light intensity change on the incident bar and the blue curve is the light intensity change on the transmitter bar. It can be seen from the figure that the light intensities on the two bars have a similar trend.

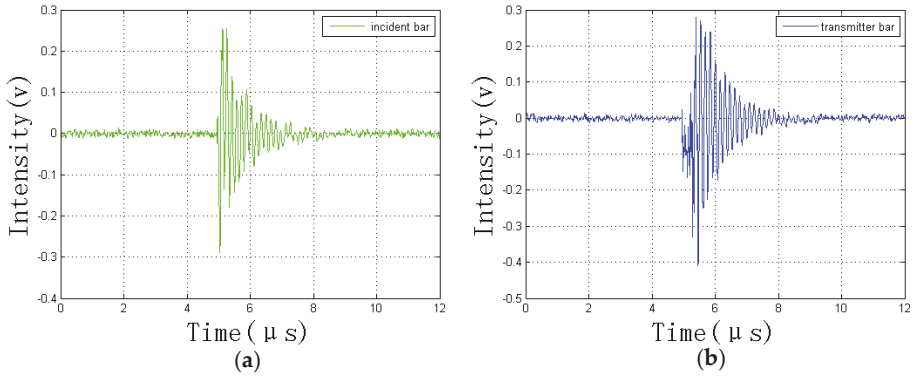


Figure 14. Original light intensity signal: (a) intensity of incident bar; (b) intensity of transmitter bar.

The dynamic mechanical properties of the specimen materials are shown in Figure 15.

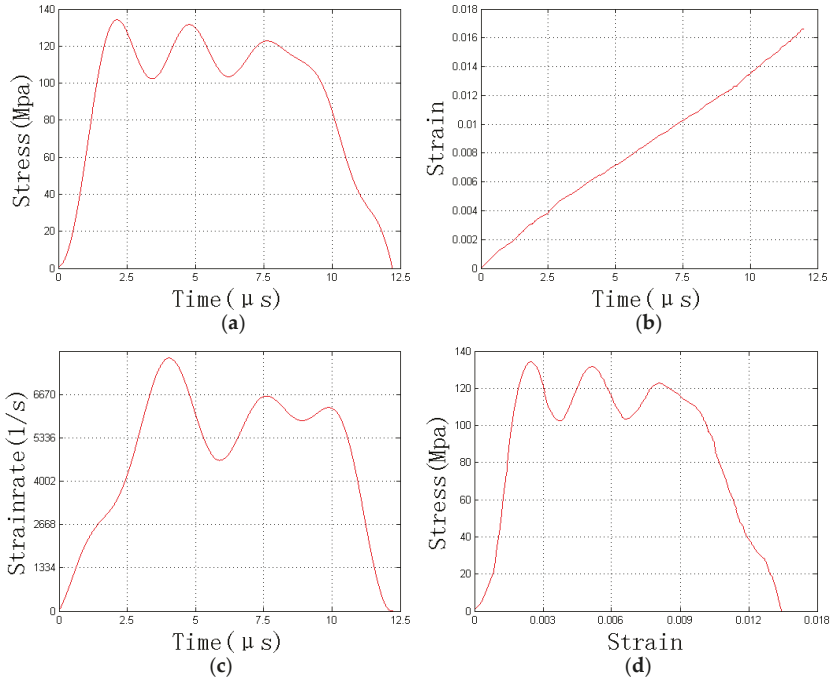


Figure 15. Dynamic mechanical properties of the specimen: (a) stress curve; (b) strain curve; (c) strain rate curve; (d) stress–strain curve.

Figure 15 shows that the maximum stress applied to the specimen during the experiment was 134.21 Mpa. As the stress increased from 0 to 134.21, the maximum strain of the specimen was 0.0166, indicating that the specimen is still in the elastic phase. The strain-rate increased from 0 to 7787(1/s). Figure 16 shows the three-dimensional surface map of stress, strain, and strain rate through the interpolation fitting algorithm.

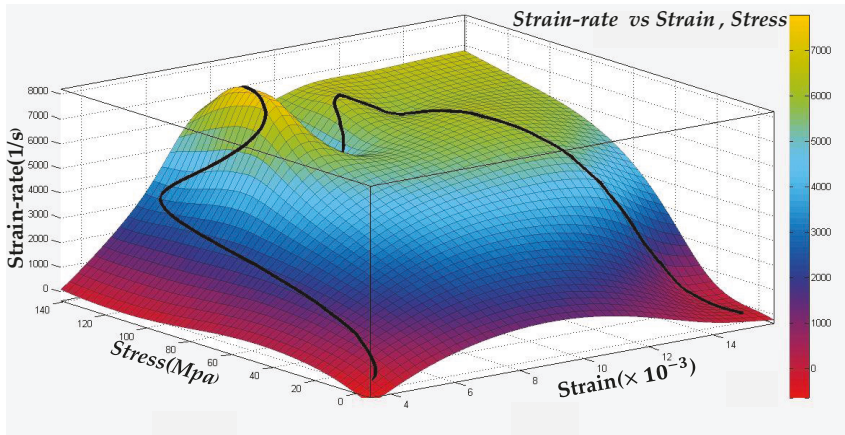


Figure 16. Three-dimensional surface map of stress, strain, and strain rate.

In the field of dynamic mechanics, a series of experiments to study the dynamic mechanical properties of materials are arranged according to the strain rate: medium strain rate experiments ( $10 \sim 10^2/s$ ), high strain rate experiments ( $10^2 \sim 10^4/s$ ), and ultra-high strain rate experiments ( $10^4 \sim 10^6/s$ ). As can be seen from the figure, the experiment we set belongs to the second case. Note that the measurement system of Kolsky bar with temporal speckle interferometry that we give here can also be used for ultra-high strain rate experiments by reducing the diameter of the bar and increasing the velocity of striking.

#### 4. Conclusions

With rapid development in the machinery industry and defense industry, the demand for dynamic mechanical properties of materials under higher strain rate and smaller size conditions is getting higher and higher. When the diameter of the bar is less than 3 mm, a traditional Kolsky bar is unable to provide reliable measurement because of the strain gauge and the bar cannot be reliably connected. A measurement system using a Kolsky bar with 2 mm diameter based on temporal speckle interferometry is presented in this paper. Compared with typical Kolsky bars, it can capture and process the dynamic mechanical properties of materials under higher strain rates and smaller size conditions. This method does not require any intermediate media, and has the advantages of being non-contact, real-time, and more reliable. It can be used for dynamic mechanical measurements of various solid materials, including metals, ceramics, and various composite materials, but it cannot be used for soft materials such as biological tissues. We believe that the application scope of the Kolsky bar based on temporal speckle interferometry will become more and more extensive through further research.

**Author Contributions:** All works with relation to this paper has been accomplished by all authors' efforts. Z.G., S.Y. and H.R. conceived and designed the experiments; S.Y., C.G., X.W., X.S., and X.W. performed the experiments; S.Y. analyzed the data; S.Y. and Z.G. wrote the paper.

**Funding:** This research was funded by National Natural Science Foundation of China (Grant No. 51675038) and the Internal Research Funds (4-ZZFJ and G-YBWT) of the Hong Kong Polytechnic University.

**Acknowledgments:** The authors acknowledge the financial supported from National Natural Science Foundation of China (Grant No. 51675038) and the Internal Research Funds (4-ZZFJ and G-YBWT) of the Hong Kong Polytechnic University.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Gama, B.A.; Lopatnikov, S.L.; Gillespie, J.W. Hopkinson bar experimental technique: A critical review. *Appl. Mech. Rev.* **2004**, *57*, 223–250. [[CrossRef](#)]
2. Casem, D.; Huskins, E.; Ligda, J.; Schuster, B. A Kolsky bar for high-rate micro-compression: Preliminary results. In *Dynamic Behavior of Materials*; Springer International Publishing: New York, NY, USA, 2016; Volume 1, pp. 87–92, ISBN 978-3-319-22451-0.
3. Avinadav, C.; Ashuach, Y.; Kreif, R. Interferometry-based Kolsky bar apparatus. *Rev. Sci. Instrum.* **2011**, *82*, 073908. [[CrossRef](#)] [[PubMed](#)]
4. Priest, R.G. Analysis of fiber interferometer utilizing  $3 \times 3$  fiber coupler. *IEEE J. Quantum Electron.* **1982**, *18*, 1601–1603. [[CrossRef](#)]
5. Choma, M.A.; Yang, C.; Lzatt, J.A. Instantaneous quadrature low-coherence interferometry with  $3 \times 3$  fiber-optic couplers. *Opt. Lett.* **2003**, *28*, 2162–2164. [[CrossRef](#)] [[PubMed](#)]
6. Casem, D.T.; Zellner, M.B. Kolsky bar wave separation using a photon doppler velocimeter. *Exp. Mech.* **2013**, *53*, 1467–1473. [[CrossRef](#)]
7. Casem, D.T.; Grunschel, S.E.; Schuster, B.E. Normal and transverse displacement interferometers applied to small diameter Kolsky bars. *Exp. Mech.* **2012**, *52*, 173–184. [[CrossRef](#)]
8. Fu, H.; Tang, X.R.; Li, J.L.; Tan, D.W. An experimental technique of split Hopkinson pressure bar using fiber micro-displacement interferometer system for any reflector. *Rev. Sci. Instrum.* **2014**, *85*, 045120. [[CrossRef](#)] [[PubMed](#)]
9. Qin, J.; Gao, Z.; Wang, X.; Yang, S. Three-dimensional continuous displacement measurement with temporal speckle pattern interferometry. *Sensors* **2016**, *16*, 2020. [[CrossRef](#)] [[PubMed](#)]
10. Wang, L.L. *Foundation of Stress Waves*; National Defend Industry Press: Beijing, China, 2005; Volume 2, pp. 52–54, ISBN 7-118-04015-0.
11. Chen, R.; Huang, S.; Xia, K.; Lu, F. A modified Kolsky bar system for testing ultrasoft materials under intermediate strain rates. *Rev. Sci. Instrum.* **2009**, *80*, 076108. [[CrossRef](#)] [[PubMed](#)]
12. Gao, Z.; Deng, Y.; Duan, Y.; Zhang, Z.; Wei, C.; Chen, S.; Cui, J.; Feng, Q. Continual in-plane displacement measurement with temporal wavelet transform speckle pattern interferometry. *Rev. Sci. Instrum.* **2012**, *83*, 015107. [[CrossRef](#)] [[PubMed](#)]
13. Daubechies, I. Ten lectures on wavelets (CBMS-NSF Regional Conference Series in Applied Mathematics). *Soc. Ind. Appl. Math.* **1992**, *377*. [[CrossRef](#)]
14. Wang, X.; Gao, Z.; Qin, J.; Zhang, X.; Yang, S. Temporal heterodyne shearing speckle pattern interferometry. *Opt. Lasers Eng.* **2017**, *93*, 76–82. [[CrossRef](#)]
15. Feng, Z.; Gao, Z.; Zhang, X.; Wang, S.; Yang, D.; Yuan, H.; Qin, J. A polarized digital shearing speckle pattern interferometry system based on temporal wavelet transformation. *Rev. Sci. Instrum.* **2015**, *86*, 61–87. [[CrossRef](#)] [[PubMed](#)]
16. Svanbro, A. In-plane dynamic speckle interferometry: Comparison between a combined speckle interferometry/speckle correlation and an update of the reference image. *Appl. Opt.* **2004**, *43*, 4172. [[CrossRef](#)] [[PubMed](#)]



© 2018 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).

Review

# Frequency-Shifted Optical Feedback Measurement Technologies Using a Solid-State Microchip Laser

Kaiyi Zhu <sup>1</sup>, Hongfang Chen <sup>2</sup>, Shulian Zhang <sup>1</sup>, Zhaoyao Shi <sup>2</sup>, Yun Wang <sup>3</sup> and Yidong Tan <sup>1,\*</sup>

<sup>1</sup> The State Key Laboratory of Precision Measurement Technology and Instruments, Department of Precision Instrument, Tsinghua University, Beijing 100084, China; zky14@mails.tsinghua.edu.cn (K.Z.); zsl-dpi@mail.tsinghua.edu.cn (S.Z.)

<sup>2</sup> College of Mechanical Engineering and Applied Electronics Technology, Beijing University of Technology, Beijing 100022, China; hfchen@bjut.edu.cn (H.C.); shizhaoyao@bjut.edu.cn (Z.S.)

<sup>3</sup> Beijing Institute of Space Electromechanical, Beijing 100076, China; 13366359922@163.com

\* Correspondence: Tanyd@tsinghua.edu.cn; Tel.: +86-10-62783440

Received: 29 November 2018; Accepted: 21 December 2018; Published: 29 December 2018



**Abstract:** Since its first application toward displacement measurements in the early-1960s, laser feedback interferometry has become a fast-developing precision measurement modality with many kinds of lasers. By employing the frequency-shifted optical feedback, microchip laser feedback interferometry has been widely researched due to its advantages of high sensitivity, simple structure, and easy alignment. More recently, the laser confocal feedback tomography has been proposed, which combines the high sensitivity of laser frequency-shifted feedback effect and the axial positioning ability of confocal microscopy. In this paper, the principles of a laser frequency-shifted optical feedback interferometer and laser confocal feedback tomography are briefly introduced. Then we describe their applications in various kinds of metrology regarding displacement measurement, vibration measurement, physical quantities measurement, imaging, profilometry, microstructure measurement, and so on. Finally, the existing challenges and promising future directions are discussed.

**Keywords:** laser feedback; precision measurement; frequency-shifted; solid-state laser

## 1. Introduction

Laser feedback, also known as laser self-mixing interference, was first applied as a displacement sensor by P.G.R. King in 1963 [1]. It is a physical phenomenon where part of the output laser is reflected or scattered by the external object returning back into the laser resonator to modulate the laser output power, phase, polarization states, and so on [2–6]. Unlike the traditional laser interferometry, laser feedback interference occurs inside the laser resonator, thus making the system concise and auto aligned. In the 1970s, the laser feedback effect of the laser diode (LD) became one of the important research fields due to the development of the optical communication technology. The theoretical system of the LD optical feedback was established by R. Lang, D. Lenstra, W.M. Wang, among others [7–10]. However, the sensitivity of the LD optical feedback is still not high enough to realize the non-cooperative measurement as the black or high transmission targets.

In 1979, Otsuka reported the external optical feedback of  $\text{LiNdP}_4\text{O}_{12}$  lasers produced using a rotating glass plate [11]. This kind of laser, known as the microchip laser, belongs to the class-B laser, of which the population decays slowly compared with the field [12]. In this case, the feedback signal can be amplified by a gain factor of the ratio  $\gamma_c/\gamma_1$ , where  $\gamma_c$  is the damping rate of the laser cavity and  $\gamma_1$  is the damping rate of the population inversion. According to the analysis of classic rate equation, the ratio  $\gamma_c/\gamma_1$  of a microchip laser can be high as  $10^6$ . The ratio is only  $10^3$

for an LD laser, and much lower in a gas laser. Therefore, the microchip laser has a high sensitivity compared with other kinds of lasers. It has been applied in various fields, such as displacement sensing [13,14], velocimetry [15], vibrometry [16,17], particle detecting [18], angle measurement [19], and laser parameters measurement [20,21].

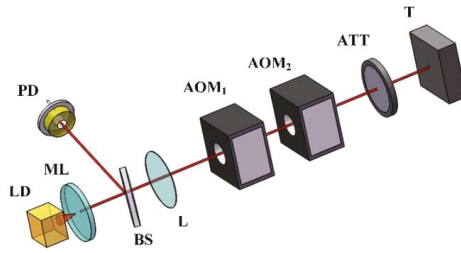
Imaging objects inside turbid media [22–25], such as the biological tissue and scattering liquid, is a challenging problem in many fields. Due to the non-contact, excellent contrast, high-resolution, nonionizing, and noninvasive, optical imaging is currently emerging as a promising method in medical imaging [26] (pp. 1–8) through the use of confocal microscopy [27–29], diffuse optical tomography [30,31], fluorescence spectroscopy [32,33], and optical coherence tomography [34–36]. For diffuse optical tomography, the penetration depth can be several centimeters into biological tissue, but recovering information from scattered photons is a challenge, and the spatial resolution is on the order of 20% of the imaging depth [26] (p. 249). Most optical imaging methods are based on ballistic photons, which will be rapidly decreased when the penetration depth increases in turbid media; thus, these methods are usually confined to imaging of a few millimeters in biological samples. More recently, the laser optical feedback tomography was demonstrated by Lacot [37], which realizes the surface imaging of a French coin immersed in 1 cm of milk. Combining the technology of confocal tomography and a laser feedback effect, Tan et al. proposed the laser confocal feedback tomography [38,39]. Due to the ultrahigh sensitivity of a solid-state microchip laser to external frequency-shifted feedback, it shows promise towards reaching a greater depth than other methods.

This paper provides an overall review of the research on frequency-shifted optical feedback measurements using a solid-state microchip laser. The remainder of the review is structured as follows. Section 2 reveals the basic principles and experimental setups of the laser feedback interferometer and laser confocal feedback tomography. Section 3 discusses the applications of the laser feedback interferometer in displacement sensing, vibration sensing, particle sensing, liquid evaporation rate measurement, refractive index measurement, and thermal expansion coefficient measurement. In Section 4, the microstructure imaging and measurement, profilometry, lens thickness measurement, and imaging, combined with other technologies, are presented based on the laser confocal feedback tomography. A summary and conclusion are given in Section 5.

## 2. Experimental Setup and Theoretical Analysis

### 2.1. Laser Frequency-Shifted Optical Feedback

The experimental setup of the laser frequency-shifted optical feedback is shown in Figure 1. A  $3\text{ mm} \times 3\text{ mm} \times 0.75\text{ mm}$  Nd:YVO<sub>4</sub> crystal plate is employed to form a laser resonator with the coating on both surfaces. The left surface is coated to be antireflective at the pump wavelength of 808 nm and highly reflective ( $R > 99.8\%$ ) at the lasing wavelength of 1064 nm; the output surface is coated to be have 5% transmittance at a wavelength of 1064 nm. The pump light, produced using a fiber-coupled single-mode laser diode, is focused onto the center of the Nd:YVO<sub>4</sub> crystal. The threshold of Nd:YVO<sub>4</sub> is about 27 mW, and the range of the pump light is 0–200 mW. The output laser is split by the beam splitter (BS). The reflective part is detected by a photon detector (PD), and the transmitted part is collimated by the lens (L). Two acousto-optic modulators (AOMs) with working frequencies of  $\Omega_1$  and  $\Omega_2$  are utilized at a differential configuration to modulate the frequency of the output laser. By adjusting the AOMs and target (T), the laser frequency is shifted by  $\Omega = 2|\Omega_1 - \Omega_2|$  after a round-trip as the measuring light. The attenuator (ATT) is inserted in the optical path to modify the feedback level.



**Figure 1.** Schematic diagram of the laser frequency-shifted optical feedback interferometer. LD: laser diode; ML: microchip Nd:YVO<sub>4</sub> laser; BS: beam splitter; PD: photon detector; L: lens; AOM<sub>1</sub>, AOM<sub>2</sub>: acousto-optic modulators; ATT: attenuator; T: target.

In the weak feedback level, the laser power fluctuation under the effect of frequency-shifted feedback can be simulated using the modified Lang–Kobayashi equation [40,41]:

$$\begin{aligned} \frac{dN(t)}{dt} &= \gamma(N_0 - N(t)) - BN(t)|E(t)|^2 \\ \frac{dE(t)}{dt} &= \frac{1}{2}(BN(t) - \gamma_C)E(t) + \gamma_C\kappa \cos(2\pi\Omega t - \omega\tau)E(t) \end{aligned} \tag{1}$$

where  $N(t)$  is the population inversion,  $E(t)$  is the amplitude of the laser electric field,  $N_0$  is the inversion particle number under a small signal,  $\gamma$  is the decay rate of the population inversion,  $B$  is the Einstein coefficient,  $\gamma_C$  is the laser cavity decay rate,  $\kappa$  is the effective laser feedback level,  $\omega$  is the optical running laser frequency, and  $\tau$  is the photon round-trip time between the laser and the target.

The stationary solution of the Equation (1) can be obtained by setting the electric field  $E$  and population inversion  $N$  to be constant. Without feedback ( $\kappa = 0$ ), the steady laser solutions are:

$$\begin{aligned} N_s &= \frac{\gamma_C}{B} \\ I_s &= |E_s|^2 = \frac{\gamma}{B}(\eta - 1) \end{aligned} \tag{2}$$

where  $I_s$  is the stationary intensity of the laser field, and  $\eta = BN/\gamma_C$  is the normalized pumping rate.

The power spectrum of the continuous pumped Nd: YVO<sub>4</sub> laser can be considered as slight fluctuations of the stable solutions. The electric field  $E(t)$  and the population inversion  $N(t)$  can be written as:

$$\begin{aligned} N(t) &= N_s + n(t) \\ E(t) &= E_s + e(t) \end{aligned} \tag{3}$$

Substituting Equations (2) and (3) into Equation (1) and neglecting the second-order terms, we obtain:

$$\begin{aligned} \frac{dn^*(t)}{dt} &= -\gamma\eta n^*(t) - 2\gamma(\eta - 1)e(t) \\ \frac{de^*(t)}{dt} &= \frac{1}{2}\gamma_C n^*(t) + \gamma_C\kappa \cos(2\pi\Omega t - \omega\tau) + \gamma_C\kappa \cos(2\pi\Omega t - \omega\tau)e^*(t) \end{aligned} \tag{4}$$

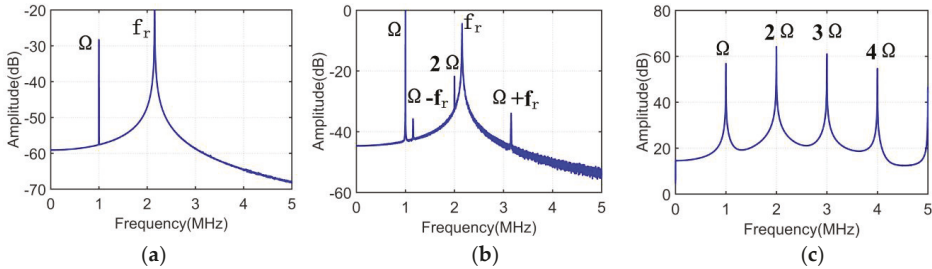
where  $n^*(t) = n(t)/N_s$  and  $e^*(t) = e(t)/E_s$  are the normalized variation of the population inversion and the electric field.

The variation of laser intensity versus time can be derived according to Equation (3) as:

$$\frac{\Delta I}{I} = \frac{2E(t)e(t)}{E(t)^2} \approx \frac{2e(t)}{E_s} = 2e^*(t) \tag{5}$$

Thus, the solution of Equation (4) denotes the power spectrum of a Nd: YVO<sub>4</sub> laser under frequency-shifted feedback. The results with different feedback levels using numerical solutions are shown in Figure 2. Three kinds of feedback levels are identified as weak, moderate, and strong feedback [42]. For the weak feedback level, there are two peaks at the relaxation oscillation frequency

$f_r$  and shifted frequency  $\Omega$ . When the feedback level is increasing, the signal to noise ratio (SNR) of the measuring light increases as well. However, the harmonic peaks ( $2\Omega \dots$ ) and parametric peaks ( $\Omega+f_r, \Omega-f_r \dots$ ) appear. Especially, once the  $\kappa$  increases to the strong regime, the oscillating peak at  $f_r$  disappears, and the signal at  $\Omega$  and its harmonics ( $2\Omega, 3\Omega, 4\Omega \dots$ ) oscillate. In summary, the laser power spectrum with two peaks at the relaxation oscillation frequency  $f_r$  and shifted frequency  $\Omega$  is the desirable case for actual measurement. The appropriate  $\kappa$  can be adjusted by the ATT in the measurement. For strong feedback, the power spectrum is raised and disturbed by the shifted frequency signal, which cannot be utilized as a sensor.



**Figure 2.** Numerical power spectrum of the Nd: YVO<sub>4</sub> laser with different feedback levels: (a) weak feedback, (b) moderate feedback, and (c) strong feedback.

The relative laser output power with the frequency of  $\Omega$  is given as:

$$\frac{\Delta I(\Omega)}{I_s} = \kappa G(\Omega) \cos(2\pi\Omega t - \phi + \phi_s) \tag{6}$$

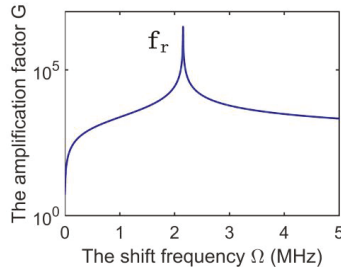
where  $\Delta I$  denotes the intensity modulation of the measuring light,  $\phi_s$  is a fixed phase, and  $\phi$  is the phase related to the external cavity length.  $G$  is the frequency-dependent amplification factor, which can be expressed as:

$$G(\Omega) = 2\gamma_c \frac{[\eta^2\gamma^2 + 4\pi^2\Omega^2]^{1/2}}{\left[4\eta^2\gamma^2\pi^2\Omega^2 + \left(4\pi^2f_r^2 - 4\pi^2\Omega^2\right)^2\right]^{1/2}} \tag{7}$$

The simulation based on the parameters of a Nd:YVO<sub>4</sub> laser is as follows.

The nearer the shift frequency  $\Omega$  is to the relaxation oscillation frequency  $f_r$ , the larger the amplification factor  $G$  is, and it can be as high as  $10^6$  according to Figure 3. Therefore, the laser frequency-shifted optical feedback has ultrahigh sensitivity and excellent performance in the detection of weakly scattering light. However, it should be noted that the shift frequency  $\Omega$  cannot coincide with the relaxation oscillation frequency  $f_r$  in the actual experiment. Otherwise, the optical power spectrum will be in chaos, similar to that under a strong feedback level. Both the stability of the power spectrum and the signal amplification should be considered in the laser feedback interferometry.

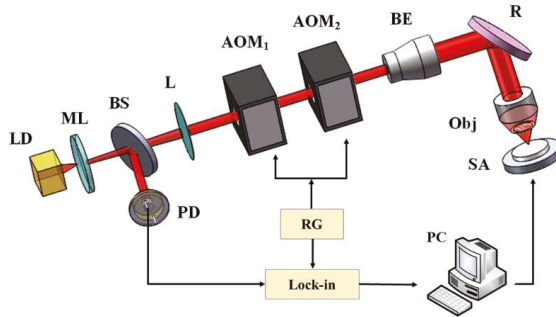




**Figure 3.** The relationship between the amplification factor  $G$  and the shift frequency  $\Omega$ .

2.2. Laser Confocal Feedback Tomography

The basic schematic of the laser confocal feedback tomography is shown in Figure 4. The optical path before the AOMs is the same as that of the laser feedback interferometer as shown in Figure 1. Then the modulated light with the shift frequency of  $\Omega/2 = |\Omega_1 - \Omega_2|$  is expanded by the beam expander (BE), reflected by the reflector (R), and finally, focused by the objective (Obj) onto the sample (SA). The reflected or scattered light by the SA returns back to the laser cavity along the same path, working as the measuring light with the shift frequency of  $\Omega$ .



**Figure 4.** Schematic diagram of the laser confocal feedback tomography. LD: laser diode; ML, microchip Nd:YVO<sub>4</sub> laser; BS, beam splitter; PD, photodiode; L, lens; AOM<sub>1</sub>, AOM<sub>2</sub>, acousto-optic modulators; BE, beam expander; R, Reflector; Obj, objective; SA, sample (phantom); RG, radio-frequency generator; Lock-in, lock-in amplifier; PC, computer.

By sending the signal of the PD to the lock-in amplifier (Lock-in) as the measurement channel, generating the electrical signal with the  $\Omega$  frequency from the radio-frequency generator (RG) as the reference channel, the intensity of the measuring light can be demodulated by the Lock-in. The optical power modulation of the measuring light can be induced as [38,39]:

$$\frac{\Delta I(\Omega)}{I_s} = I(u)\kappa G(\Omega) \cdot \cos(2\pi\Omega t - \phi + \phi_s) \tag{8}$$

$I(u)$  is the light intensity response function of the traditional confocal system, and is given as [27,43]:

$$I(u) = K_r \int_0^{v_d} \left| \int_0^1 \exp(iu\rho^2) J_0(v_2\rho) \rho d\rho \right|^2 v_2 dv_2 \tag{9}$$

where  $u$ ,  $\rho$ , and  $v_d$  are the normalized de-focus parameter, radial coordinate, and diameter of the pinhole, respectively, and  $K_r$  is the normalized constant.

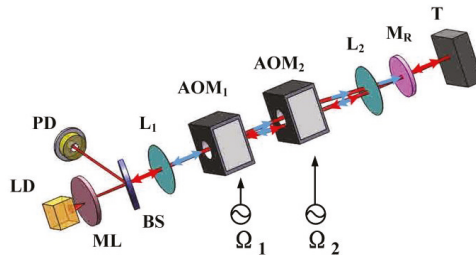
It is noted that there is no pinhole in Figure 4 because the laser in the setup is not only the light source but also the detector; therefore, the laser waist can be utilized as the pinhole filter. Thus, the structure of the laser confocal feedback tomography greatly simplified compared with the traditional confocal system.

### 3. Applications of the Laser Feedback Interferometer

#### 3.1. Displacement Sensing

Displacement measurement is the basic application of the laser feedback interferometer. According to Equation (6), one cycle in the laser intensity modulation corresponds to half the wavelength in the length of the external cavity, which is similar to the traditional laser interferometer. Due to the existence of the amplification factor  $G$ , the laser feedback interferometer has a high sensitivity that it does not need a retroreflector or corner prism set on the target. Therefore, it is advantageous to realize the axial displacement measurement where no additional optical devices are installed.

Wan [13] proposed a quasi-common-path configuration based on frequency shifting and multiplexing to compensate the air disturbance and the thermal effects of the components. The system structure is shown in Figure 5. It is innovative to put a reference mirror after the two AOMs to generate a feedback light. The original optical path without diffraction and the diffractive optical path with the shift frequency of  $\Omega/2 = |\Omega_1 - \Omega_2|$  constitute a round-trip as the reference light, which is shown in blue arrows.



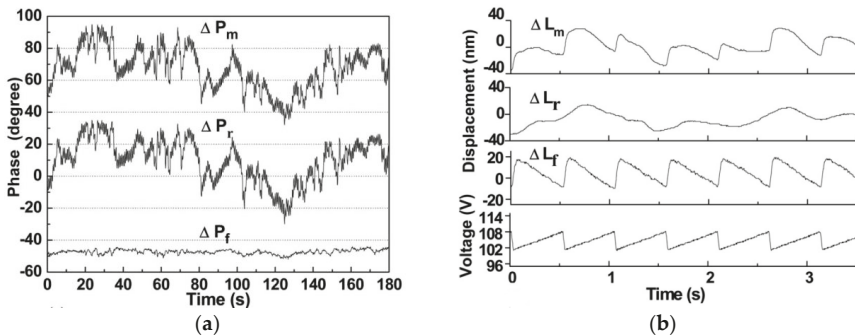
**Figure 5.** Configuration of the quasi-common-path laser feedback interferometer. LD: laser diode; ML: microchip Nd: YAG laser; BS: beam splitter; PD: photon detector; L<sub>1</sub>, L<sub>2</sub>: lens; AOM<sub>1</sub>, AOM<sub>2</sub>: acousto-optic modulators; M<sub>R</sub>: reference mirror; T: target.

The displacement information of the measuring light and reference light can be demodulated at the frequency  $\Omega/2$  and  $\Omega$ , respectively. The change of the displacement  $\Delta L$  is related to the phase variation  $\Delta P$  by:

$$\begin{aligned} \Delta L_m &= (c/2n\omega)\Delta P_m \\ \Delta L_r &= (c/2n\omega)\Delta P_r \\ \Delta L_f &= \Delta L_m - \Delta L_r = (c/2n\omega)\Delta P_f \end{aligned} \tag{10}$$

where  $\Delta P_m$  is the phase variation of the measuring light,  $\Delta P_r$  is the phase variation of the reference light,  $\Delta L_m$  is the corresponding measurement displacement,  $\Delta L_r$  is the corresponding reference displacement,  $\Delta P_f$  is the final phase variation, and  $\Delta L_f$  is the final displacement.

The piezoelectric transducer (PZT) is tested in the system as the T in Figure 5. The reference mirror (M<sub>R</sub>) is placed 10 mm before the PZT. In the experiment, the PZT is driven by a ramp wave signal. The measured  $\Delta P_m$  and  $\Delta P_r$ , and the corresponding  $\Delta L_m$ ,  $\Delta L_r$ ,  $\Delta P_f$ , and  $\Delta L_f$  are shown in Figure 6. It is noted that the PZT vibration waveform cannot be revealed from  $\Delta L_m$ ; however, it can be revealed from  $\Delta L_f$  accurately. The maximum nonlinear error of the  $\Delta L_f$  data is 1.8 nm, indicating that its short period resolution is better than 2 nm.

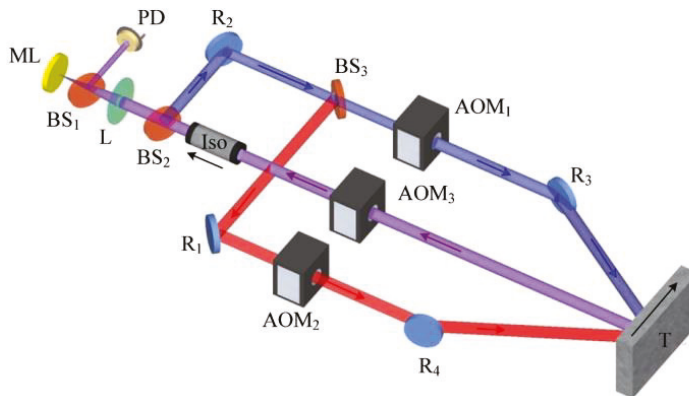


**Figure 6.** Results of the displacement measurement: (a) phase stability test results, and (b) measurement result of PZT vibration. The bottom curve shows the PZT driving signal. (Figure reproduced from Ref. [13]).

The quasi-common-path configuration has dramatically improved the performance of the laser feedback interferometer. However, two shortcomings remain to be improved. One is the measuring speed. The signal demodulation is based on the Lock-in in Ref. [13], which limits the measuring speed to be 100  $\mu\text{m/s}$ . Zhang and Ren [44] proposed a new signal processing method replacing the Lock-in to a phase meter, and the measuring speed is improved to 10 mm/s. Then, Zhang [45] improved the performance of the system by replacing the laser source with a Nd: YVO<sub>4</sub> crystal. The relaxation oscillation frequency increased from 300 kHz to 4.5 MHz, and the shifted frequencies were 2 MHz and 1 MHz, respectively. Finally, the measuring speed was improved to 120 mm/s.

The other shortcoming is the common path compensation. Due to the difference of the optical path between the measuring light and the reference light, the thermal effect produced by the AOMs is different; thus, quasi-common compensation cannot be completely eliminated. On the other hand, the position of the  $M_R$  limits the compensation effect when sensing the displacement over a long distance. Ren [46] proposed a ring optical path configuration of the laser feedback interferometer. The paths of the reference light and the measured light coincide completely; thus, the thermal compensation effect is improved. Zhang [47–49] demonstrated a common-path heterodyne self-mixing interferometry with polarization and frequency multiplexing. The two mutual independent orthogonal polarized lights are used as the measuring and reference lights, of which the optical paths completely coincide in space; thus, the effect of the AOMs thermal creep and air disturbance can be eliminated. The short-term resolution is better than 2.5 nm, and the long-term zero drift is less than 60 nm over 7 h. Xu [50] proposed a novel approach to realize full path compensation laser feedback interferometry for remote sensing. The displacement of a steel block at a distance of 10 m is measured, of which the stability is  $\pm 12$  nm over 100 s and  $\pm 50$  nm over 1000 s, and the short-term resolution is better than 3 nm.

More recently, the two-dimensional (2-D) displacement measurement based on the self-mixing interferometry is revealed [14]. The system is shown in Figure 7. Two measuring beams at the different shift frequencies are incident on the same spot on the target. By heterodyne demodulating the phases of the two beams and deriving the relationship between the phases' variations and the change of the in-plane and out-of-plane displacement, 2-D displacement measurement can be realized. Various movements in the track of Lissajous figures and random motion are measured in the experiments. The results show that the resolutions of the two dimensions are better than 5 nm and the standard deviation can be better than 0.1  $\mu\text{m}$ .



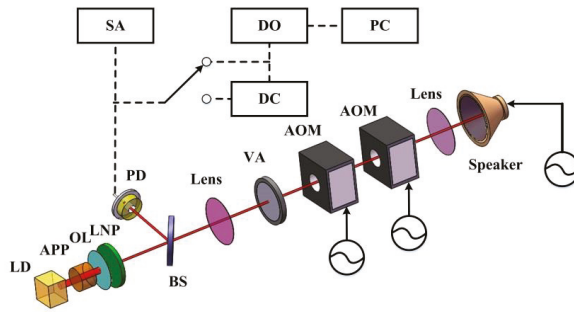
**Figure 7.** Schematic diagram of the 2-D displacement measurement system. ML is the microchip laser; BS<sub>1</sub>, BS<sub>2</sub>, and BS<sub>3</sub> are the beam splitters; L is the optical lens; AOM<sub>1</sub>, AOM<sub>2</sub>, and AOM<sub>3</sub> are the acousto-optic modulators; R<sub>1</sub>, R<sub>2</sub>, R<sub>3</sub>, and R<sub>4</sub> are reflectors; Iso is the optical isolator; T is the target; and PD is the photodetector. (Figure reproduced from Ref. [14]).

The simultaneous measurements of in-plane and out-of-plane displacements are a significant issue to be researched, such as the grating interferometry [51,52], speckle pattern interferometry [53,54], digital image correlation [55,56], and laser Doppler distance sensing [57]. The resolution of the grating interferometry is a few nanometers, and the accuracy can reach a submicron scale. However, the 2-D grating needs to be set on the target, which limits the application. A laser Doppler distance sensor is appropriate for the dynamic position measurements of the fast-moving object, and the resolution of which is only on the submicron order. Speckle pattern interferometry and digital image correlation are two noncontact and full field displacement measurement methods, which are widely used in industrial nondestructive detection. However, they are limited to the static and quasi-dynamic displacement measurement field; the measurement range and accuracy are related to the speckle size obtained and come to a compromise in the application. Compared with other methods, laser feedback interferometry has the advantages of compactness, non-contact, high resolution, and high accuracy. However, the characteristic of the single-spot measurement limits the application in the full field measurement. It is a promising method to be applied in the 2-D deformation of materials measurement and 2-D thermal expansion measurement, among others.

### 3.2. Vibration Sensing

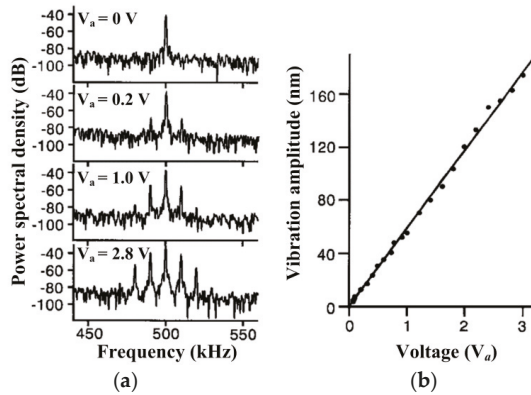
Besides the displacement sensing, the precise measurement of vibration has attracted wide attention as well. Vibration sensing can be utilized to analyze the dynamic characteristics of mechanical structures, fault diagnosis of mechanical systems, target identification, and sound visualization [58]. However, it is difficult to detect small vibration and displacement in many cases, especially the micro-vibration of non-cooperative targets at long distance.

Otsuka [16] proposed the real time nanometer vibration measurement using a self-mixing microchip solid-state laser. The experimental configuration is shown in Figure 8. The LiNdP<sub>4</sub>O<sub>12</sub> (LNP) crystal with a 1-mm-thick plane-parallel Fabry–Perot cavity that is utilized as the microchip laser. The pump light is transformed into a circular beam using the anamorphic prism pairs and is focused onto the LNP crystal via an objective microscope lens. The output light is frequency-shifted by two AOMs and impinged upon a speaker with the Al-coated surface that is placed 90 cm from the laser. A frequency demodulation circuit, a digital oscilloscope, and a radio-frequency spectrum analyzer are utilized in the signal processing.



**Figure 8.** Experimental configuration of self-mixing laser-Doppler vibrometry. LD, laser diode; APP, anamorphic prism pairs; OL, objective microscopelens; BS, glass-plate beam splitter; PD, photodiode receiver; SA, radio frequency spectrum analyzer; VA, variable attenuator; DO, digital oscilloscope; DC, frequency demodulation circuit; PC, personal computer. (Figure reproduced from Ref. [16]).

Figure 9a shows the power spectra for several voltages applied to the speaker. The measured vibration amplitudes ( $A_{v,m}$ ) at the modulation frequency of 8.42 kHz with the carrier frequency of 500 kHz are plotted in Figure 9b. The linear relation was  $A_{v,m}/V_a = 59 \text{ nm/V}$  for the speaker we used. The measurable minimum vibration amplitude was 1 nm according to the carrier-to-noise ratio in the absence of a voltage to the speaker. The velocity range in the present vibrometry was 1  $\mu\text{m/s}$  to 10 cm/s in the vibration frequency range 20 Hz to 20 kHz. The temporal evolutions of nanometer vibrations were measured by analyzing modulated output waveform using the Hilbert transformation. The almost unheard sound of music below a 20-dB pressure level is reproduced by the system.



**Figure 9.** (a) Power spectra (long-time average) for several applied voltages. (b) Maximum vibration amplitude versus voltage applied to the speaker. Carrier frequency, 500 kHz; modulation frequency, 8.42 kHz. (Figure reproduced from Ref. [16]).

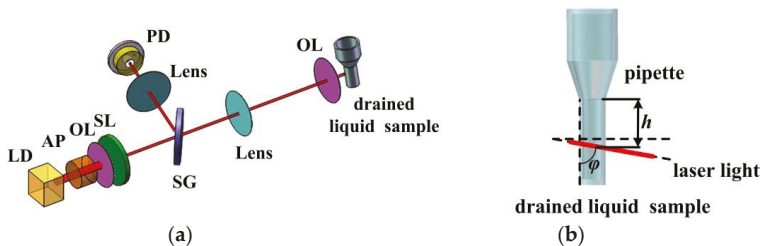
Furthermore, three-channel real time nanometer vibration [59] was successfully developed with three pairs of acoustic optical modulators and a three-channel frequency-modulated wave demodulation circuit, realizing simultaneous independent measurement of three different nanometer-vibrating targets. On the other hand, the vibration of targets placed 2.5 km away through single-mode optical fiber access was successfully measured [60], due to the effective long-haul self-mixing interference.

There are also other researchers applying self-mixing interference effects in the vibration sensing. Huang [61] proposed a vibration system extreme points model and self-mixing vibration sensor based

on the effect of LD; the amplitude-frequency response curve of the loudspeaker is drawn, and the value of the piezoelectric coefficient of PZT was obtained. Tao [62] presented a signal-processing synthesizing wavelet transform, and a Hilbert transform employed to the micro-vibration measurement based on the semiconductor laser self-mixing technology. The real-time micro vibration with a nanometer resolution and a much wider bandwidth than conventional modulation methods were proved. Dai [63] proposed the self-mixing interferometry in a fiber ring laser and its application for vibration measurement. The maximum error of the amplitude was about  $\lambda/10$ , and the maximum relative error of the frequency was about 10%.

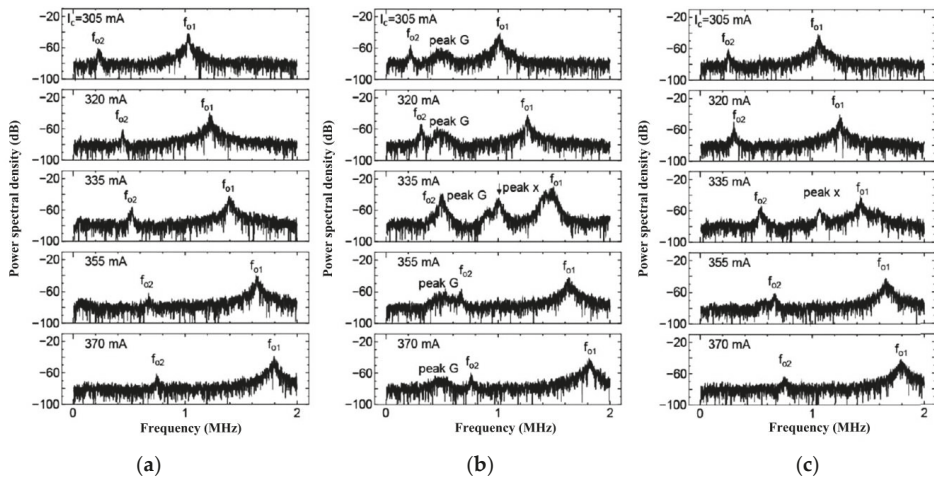
### 3.3. Particle Sensing

The self-mixing laser Doppler velocimeter is one of the promising technologies applied in the observation and detection of the particles flowing in liquid [64,65], which has been recognized as the simplest, most cost-effective, and most self-aligned metrology. The system is shown in Figure 10 using a drained suspension including polystyrene latex standard spheres (PLS). The output power is modulated by the frequency-shifted scattered light generated by the motion of moving targets. When the moving target moves at a uniform velocity, the Gaussian power spectrum can be observed, and the peak frequency corresponds to the Doppler shift frequency  $f_d$ .



**Figure 10.** (a) Experimental setup. LD, laser diode; AP, anamorphic prism pair; OL, objective lens; SL, Nd:GdVO<sub>4</sub> solid-state laser; SG, slide glass; PD, photodiode receiver; PC, personal computer. (b) Configuration used to detect drained suspension. (Figure reproduced from Ref. [64]).

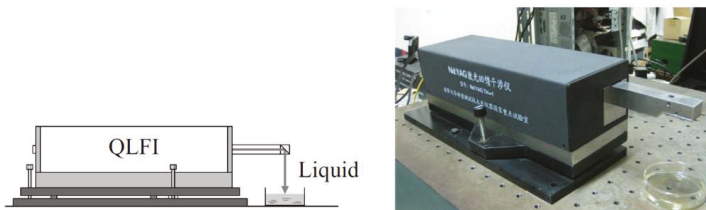
The angle of the incident laser light is set at  $\psi = 16^\circ$  to the surface of the drained suspension. Here it is noted that the Gaussian spectrum is marked by the noise if the intensity of the light scattered from the moving particles is very weak. In this case, the motion can be revealed in its higher harmonics when the frequency of the relaxation oscillation of the laser output,  $f_{o2}$ , is made to coincide with  $f_d$  by tuning the pump current of the laser ( $i_c$ ) at  $i_c = 335$  mA, as shown in Figure 11b,c. Thus, the ultrahigh sensitivity measurement of extremely weak scattered light can be achieved. The spectral peak G in Figure 11 reflects the motion of the particles passing through the incident laser light. The frequency of the Gaussian is  $f_d = 500$  kHz, which is related to the average velocity of the particles. The amplitude is associated with the intensity of the light scattered, which is proportional to the concentration of the particles.



**Figure 11.** (a) Power spectra of the laser output observed when the laser light is emitted into the atmosphere. Power spectra of the modulated wave observed in drained (b)  $7 \times 10^{-3}$  wt% and (c)  $7 \times 10^{-5}$  wt% PLS-water mixture with PLS 250 nm in diameter. (Figure reproduced from Ref. [64]).

### 3.4. Liquid Evaporation Rate Measurement

Liquid evaporation measurement is fundamental in many industrial applications and scientific research, such as quantitative analysis, and physical and chemical reaction process monitoring. The common methods used to measure the liquid level variation, including the capacitive sensors, the fiber liquid level sensor, the laser triangulation, etc., are contact-based or have limited accuracy. Tan [66] reported the application of real-time evaporation of the liquid measurement based on the laser feedback interferometer. The experimental system is shown in Figure 12. The structure of the instrument belongs to the quasi-common-path laser feedback interferometer [44]. A hollow arm with a right-angled prism fixed on the edge is installed at the interface of the interferometer, turning the laser beam 90 degrees and adjusting it incident perpendicularly onto the liquid surface. Four different transparent liquids, including distilled water, absolute alcohol, acetone, and ether, are measured in the system. The results show that real-time direct measurement of liquid evaporation and liquid level measurement with a nanometer order is realized.

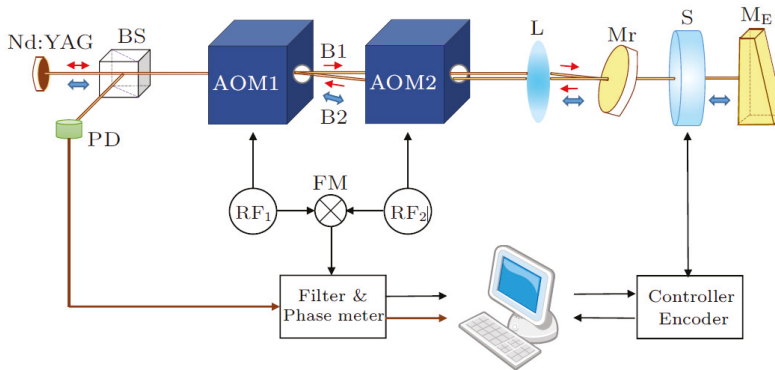


**Figure 12.** Quasi-common-path laser feedback interferometer (QLFI) monitoring the liquid level. (Figure reproduced from Ref. [66]).

### 3.5. Refractive Index Measurement

The refractive index  $n$  of a material is a physical quantity that describes how light propagates through the material. Thus, accurate measurement of material refractive index is significant for optical system design. Different methods have been developed. The minimum deviation method has the highest accuracy ( $10^{-6}$ ), but the sample needs to be prism-shaped, which is complicated and costly.

An Abbe refractometer is the most commonly used instrument, but the measurement range is limited from 1.3 to 1.7, and the accuracy is  $10^{-4}$ . In comparison, the interferometry methods have many advantages, such as easy sample processing, low cost, wide measuring range, high measurement accuracy, and they have been widely researched in the refractive index measurement field, including Michelson interferometry [67], Mach–Zehnder interferometry [68], Fabry–Perot interferometry [69] and plasmonic interferometry [70,71], etc. The common principle of these interferometry methods is to measure the optical path change caused by the sample rotation or displacement. However, these interferometers cannot avoid the environmental disturbance, and the measurement accuracies are hard to be further improved. Xu [72,73] proposed a method to measure the refractive index and thickness simultaneously based on the laser feedback interferometry. Compared with other methods, the influence of environmental disturbance can be eliminated by the quasi-common path structure of the laser feedback interferometry, and the absolute uncertainty in the refractive index measurement reaches  $\approx 10^{-5}$ . On the other hand, the refractive index measurement range is larger than the conventional interferometry due to the high sensitivity of the system. The system is shown in Figure 13.



**Figure 13.** Configuration of the refractive-index and thickness measurement using laser feedback interferometry. BS, beam splitter; PD, photo detector; AOM<sub>1</sub> and AOM<sub>2</sub>, acousto-optic modulators; L, lens; Mr, reference mirror; S, sample; ME, optical wedge (measurement mirror); RF<sub>1</sub> and RF<sub>2</sub>, radio frequency signal generators; FM, frequency mixer. (Figure reproduced from Ref. [73]).

By rotating a transparent parallel sample inside the external feedback cavity, the optical path difference  $\Delta L$  can be derived as:

$$\begin{aligned} \Delta L &= \frac{(\Delta P_m - \Delta P_r) \lambda}{2\pi} \\ &= d[\sqrt{n^2 - n_0^2 \sin^2 \theta} - n_0 \cos \theta - \sqrt{n^2 - n_0^2 \sin^2 \theta_0} + n_0 \cos \theta_0] \end{aligned} \tag{11}$$

where  $\lambda$  is the laser wavelength;  $d$  is the thickness of the sample;  $n$  is the refractive index to be measured;  $n_0$  is the refractive index of the air;  $\theta_0$  and  $\theta$  are the angles between the laser beam and the normal of sample surface before and after rotation, respectively; and  $P_m$  and  $P_r$  are the phase variations of the measurement light and the reference light, respectively, which are demodulated by the phase meter.

$\theta$  and  $\Delta L$  are measured at multiple angles in the experiment, then the overdetermined equation can be solved, and the refractive index of the sample can be obtained together with the thickness of the sample. Due to the high sensitivity of the laser feedback interferometry, the method can be used to measure the low transmittance materials, including calcium fluoride (CaF<sub>2</sub>), fused silica, and zinc selenide (ZnSe). The refractive indexes cover a large range from 1.42847 to 2.48272. The results demonstrate that the system has absolute uncertainties in the refractive index measurement of  $\approx 10^{-5}$ .

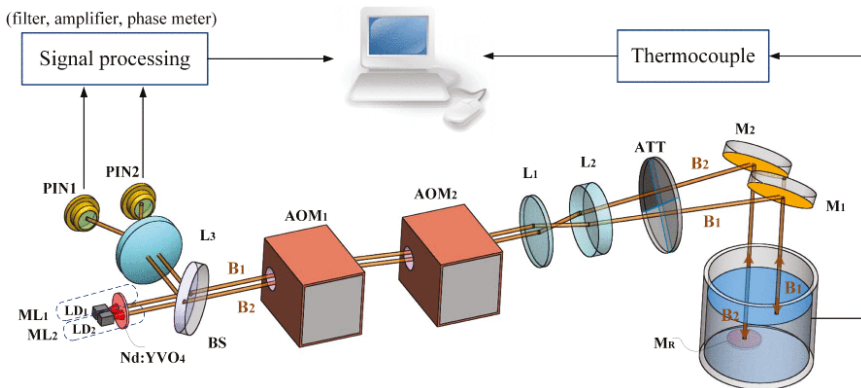


Xu [74] also reported a novel method to measure the liquid refractive index based on the double-beam laser frequency-shift feedback. The system is shown in Figure 14. Two parallel beams are adjusted to each monitor the displacement of the liquid surface and the  $M_R$  at the bottom of the tank. When the liquid level increases by  $\Delta h$ , the change of the two external cavity lengths are:

$$\begin{aligned} \Delta L_1 &= -n_0 \times \Delta h \\ \Delta L_2 &= (n - n_0) \times \Delta h \end{aligned} \tag{12}$$

where  $n_0$  is the air refractive index, and  $n$  is the liquid refractive index to be measured, which can be expressed as:

$$n = n_0 \times \left( 1 - \frac{\Delta L_2}{\Delta L_1} \right) \tag{13}$$

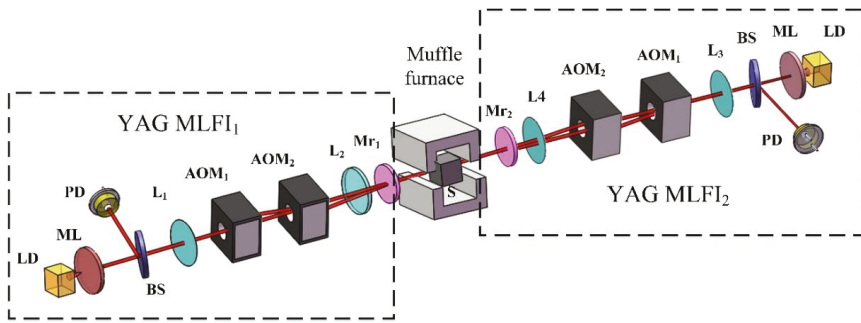


**Figure 14.** Experimental setup of the liquid refractive index measurement. LD<sub>1</sub>, LD<sub>2</sub>, laser diodes; ML<sub>1</sub>, ML<sub>2</sub>, microchip lasers; BS, beam splitter; PIN<sub>1</sub>, PIN<sub>2</sub>, detectors; AOM<sub>1</sub>, AOM<sub>2</sub>, acousto-optics modulators; L<sub>1</sub>, L<sub>2</sub>, L<sub>3</sub>, lenses; ATT, attenuator; M<sub>1</sub>, M<sub>2</sub>, M<sub>R</sub>, mirrors. (Figure reproduced from Ref. [74]).

Since the liquid refractive index depends entirely on the displacement measurement of the two beams, the refractive index measurement is also traceable. Thus, the results can be used as a reference for calibration of the liquid refractive index. Five different liquid samples, including the distilled water, ethanol, cyclohexane, silicone oil, engine oil, and NaCl solution with various concentrations, are measured in the experiment, which proves that the repeatability is better than 0.00005.

### 3.6. Thermal Expansion Coefficient Measurement

The thermal expansion coefficient is one of the most fundamental quantities of materials, which describes how the size of an object changes due to the temperature variation. The precise measurement of the thermal expansion coefficient is significant in basic scientific research and industrial application. Zheng [75,76] proposed a non-contact method to measure the thermal expansion coefficient of materials utilizing a pair of Nd: YAG microchip laser feedback interferometers (MLFIs). The schematic diagram of the system is shown in Figure 15.



**Figure 15.** Schematic diagram of the measurement of thermal expansion coefficients based on the YAG MLFIs. YAG MLFI<sub>1</sub> and YAG MLFI<sub>2</sub>, Nd:YAG microchip laser feedback interferometry systems; LD, laser diode; ML, microchip laser; BS, beam splitter; PD, photo detector; AOM<sub>1</sub> and AOM<sub>2</sub>, acousto-optic modulators; L<sub>1</sub>, L<sub>2</sub>, L<sub>3</sub>, and L<sub>4</sub>, lenses; Mr<sub>1</sub> and Mr<sub>2</sub>, reference mirrors; S, sample. (Figure reproduced from Ref. [76]).

The sample with a supporter is placed in the middle of the furnace chamber. Two symmetric laser feedback interferometers output two measurement beams, which are incident on each surface of the sample perpendicularly and coaxially, to compensate for the influence of the sample supporter distortion. Two M<sub>r</sub>s are set close to the muffle furnace, generating the two reference signals, revealing the air flow and thermal lens effect disturbances outside the furnace chamber. By subtracting it,  $S_{1m} - S_{1r}$  and  $S_{2m} - S_{2r}$  are the compensated displacement of the two laser feedback interferometers, respectively. Thus, the length change of the sample from  $T_0$  to  $T_1$  can be expressed approximately as:

$$\Delta S \approx S_{1m} + S_{2m} - (S_{1r} + S_{2r}) \tag{14}$$

where  $S_{1m}$  and  $S_{2m}$  are the displacements of the two measurement beams, and  $S_{1r}$  and  $S_{2r}$  are the displacements of the two reference beams. Then, the thermal expansion coefficient can be obtained as:

$$\alpha(T_1; T_0) = \frac{1}{L_0} \times \frac{\Delta S}{T_1 - T_0} \tag{15}$$

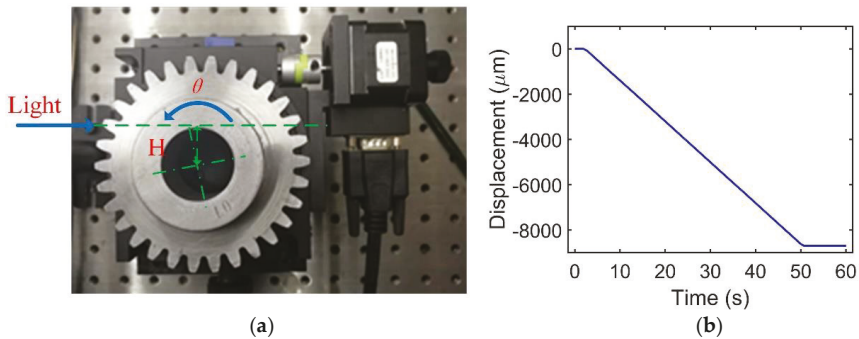
where  $L_0$  is the length of the sample at  $T_0$ .

The aluminum and the steel 45 samples are measured in the experiment from room temperature to 748 K, which proves that the measurement repeatability of thermal expansion coefficient is better than  $0.6 \times 10^{-6} \text{ (K}^{-1}\text{)}$  in the range 298 K–598 K and the high-sensitive non-contact measurement of the low reflectivity surface induced by the oxidization of the samples in the range of 598 K–748 K.

The existing methods for the thermal expansion coefficient measurement mainly include a mechanical dilatometer [77], speckle pattern interferometry [78], Fabry–Perot interferometry [79], etc. The mechanical dilatometer is an old and frequently utilized method, which has a wide temperature range but cannot satisfy the requirement for high-precision measurement. The performance of speckle pattern interferometry mostly depends on the algorithm model and speckle size; the measurement range and accuracy usually come to a compromise. Optical interference has a great performance in resolution and precision, but the measurable materials and working temperature range are limited. The thermal expansion coefficient measurement performed at the high temperature in our system solves the issues in special applications, such as aerospace materials, where the temperature of the environment can be as high as thousands of degrees, and the parameter of materials needs to be calibrated. It also shows the excellent advantages of the setup, because many measurement methods cannot be utilized at such high temperatures.

### 3.7. Gear Measurement

The possibility of the gear measurement using the laser feedback interferometer is discussed here. The gear is placed on a rotation stage with the measuring light incident perpendicularly at a certain eccentricity distance as shown in Figure 16. The unidirectional linear displacement curve is obtained in the experiment when the gear rotates in one direction. According to the Doppler theory, the integral of the velocity along the axis of light is the measuring displacement, which can be deduced as  $d = H \times \theta$ . Thus, the displacement measurement is related to the eccentricity distance and the rotation angles. The result is in agreement with the derived formula with the parameters of  $\theta = 25^\circ$  and  $H = 20$  mm. Therefore, the system can be utilized for the measurement of gear speed or the alignment error of the central gear shaft. In further research, the gear profile measurement is considered and hopefully to be realized in on-line monitoring.



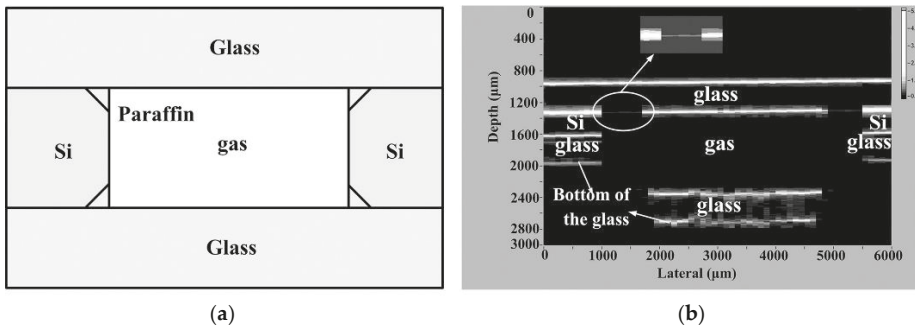
**Figure 16.** (a) Experimental system for the gear measurement;  $\theta$ , rotation angle;  $H$ , eccentricity distance. (b) Experimental result with the  $\theta = 25^\circ$  and  $H = 20$  mm.

## 4. Applications of the Laser Confocal Feedback Tomography

### 4.1. Microstructure Imaging and Measuring

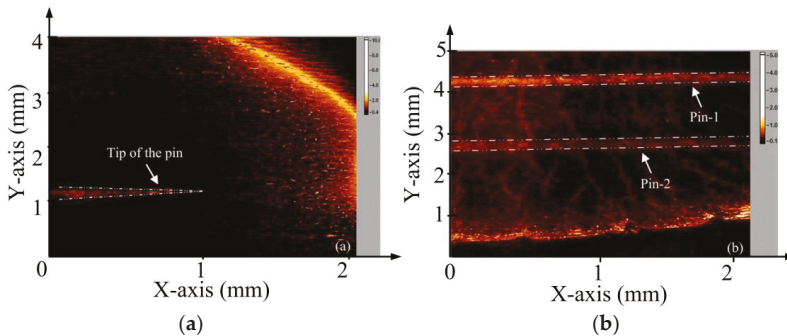
The schematic diagram of the laser confocal feedback tomography is shown in Figure 4 above. To realize the three-dimensional scanning in the inner structure of the sample, we fix the objective lens in a vertical translation stage to scan in the longitudinal direction, while the sample is set in the two-dimensional translation stage to get the horizontal movement.

The theoretical lateral resolution of the system is approximately given by  $\Delta x = 0.61\lambda / (\sqrt{2} \times NA)$ , for the objective with the  $NA = 0.42$  utilized in the experiment, corresponding to  $1.1 \mu\text{m}$ . The vertical resolution is evaluated via scanning the defocus curve and measuring the full width at half maximum (FWHM), which is about  $15\text{--}20 \mu\text{m}$ . The system can be applied in the imaging of the structure of the micro-electro-mechanical system [80,81]. For example, the sandwich type sample with the patterned layer etched in a silicon film of 1 mm and two pieces of 0.5 mm thick glass is measured. The 2-D cross-sectional imaging with the depth of 3 mm and lateral range of 6 mm is obtained as shown in Figure 17.



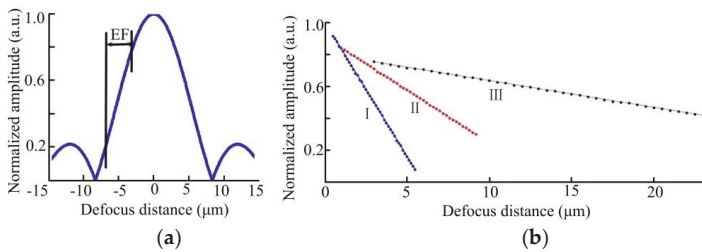
**Figure 17.** Measurement of the sandwich type sample: (a) the profile of the sample, and (b) the scanning image of the sandwich type sample. (Figure reproduced from Ref. [81]).

A biological sample, such as onion inside water, can also be measured using the laser confocal feedback tomography [39]. Two onion samples are adopted, one (named onion-1) containing a pin that is inserted 0.5 mm beneath its surface, and the other (named onion-2) containing two separated pins inserted 1 mm below its surface. The section of the pin’s tip immersed into the onion is selected to obtain its 2-D cross-sectional image in the X-Y plane. As illustrated in Figure 18, the cross-sectional images denoted by the double-dot-dash lines clearly indicate the profile of the pins’ tip. What is more, the position of the pins in the onions can also be confirmed.



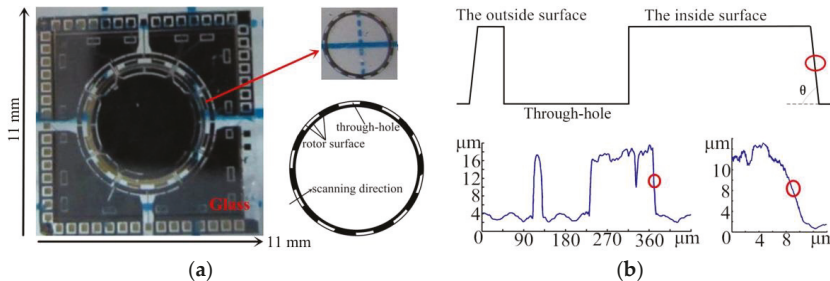
**Figure 18.** Cross-sectional imaging of the pins in the onions: (a) the pin’s tip image, and (b) two parallel pin’s image. (Figure reproduced from Ref. [39]).

Furthermore, Wang [82] improved the axial resolution to nanometers when measuring the microstructure inside the sample based on the laser confocal feedback tomography by using the linear region EF in the defocusing curve as shown in Figure 19. It should be noted that the linear region EF needs to be calibrated experimentally. First, the light is focused on the sample surface; then the objective is moved at a certain step in the longitudinal direction using a one-dimensional motorized scanning stage with the resolution of 0.5 nm, and simultaneously detects the amplitude A. The fitted line is obtained between the amplitude and the defocus distance. With the different objective lenses utilized in the system, different linear ranges and different axial resolutions are realized. The experimental measurement axial resolutions are 2, 5, and 11 nm with the objective NA values of 0.65, 0.55, and 0.3, respectively.



**Figure 19.** (a) The normalized defocusing curve, and (b) the calibrated linear range with different objective lenses. I, NA = 0.65, II, NA = 0.55, III, NA = 0.3. (Figure reproduced from Ref. [82]).

When measuring the samples, the light is focused on the interesting point first, and then the sample is scanned in the horizontal direction. During the scanning, the amplitude of the feedback light intensity varies with the structure variation. The amplitude variation can be converted to the distance according to the calibrated line. Thus, the structure of the sample is measured at last. A micro-gyroscope is measured as shown in Figure 20. The system can realize the noninvasive inner structure measurement through the protection glass. The tilt angle of the inner side of the rotor should be tested according to actual demands. According to the vertical and horizontal structure variations of the rotor edge,  $\tan(\theta)$  can be calculated, and then  $\theta$  can be obtained to estimate the verticality of the rotor edge. The results are in agreement with the value measured by the destructive method.



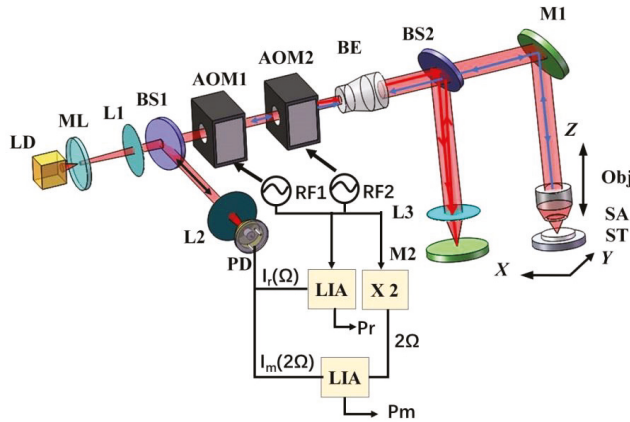
**Figure 20.** (a) Physical maps of the micro-gyroscope and the rotor. (b) Measurement results of the lateral scanning. (Figure reproduced from Ref. [82]).

It can be deduced that the laser optical feedback imaging is well-adapted for various fields and has been widely researched. Bertling [83] introduced the optical feedback interferometry for the two-dimensional visualization of acoustic fields. Several pressure distributions including progressive waves, standing waves, diffraction, and interference patterns are presented. Girardeau [84] applied the laser optical feedback imaging setup to the detection of ultrasound vibrations with nanometric amplitude. The transient-harmonics ultrasound vibrations propagating in water were detected at the air/water interface, which shows the potential for the detection of photoacoustic signals. Mowla [85] proposed a compact system using a semiconductor laser as both transmitter and receiver. Three phantoms containing macro-structural changes in optical properties were imaged, which revealed the discrimination ability between healthy tissue and a tumor. Hugon [86] proposed a galvanometric mirrors scanner moving the beam on the sample and optimized the scanner positioning to reduce the vignetting effects. A micro-structured silicon sample and a red blood cell were presented.

#### 4.2. Profilometry

In Section 4.1, only the amplitude information of the light is demodulated and utilized. When the system is applied to a surface measurement, the phase information can be demodulated, together

with the amplitude information, to reveal the micro-nano structure [38,87]. The system is shown in Figure 21. Both the intensity modulations  $I_m(2\Omega)$  and  $I_r(\Omega)$  can be obtained from the PD. By demodulating  $I_m(2\Omega)$ , not only the amplitude ( $A$ ), but also the phase ( $P_m$ ) can be obtained. At the same time, the phase information ( $P_r$ ) from  $I_r(\Omega)$  is used to compensate the  $P_m$  to achieve high accuracy in phase measurement, and thereby result in the high axial accuracy with nanometer resolution.



**Figure 21.** Schematic of the surface measurement system. ML: microchip laser; L1–L3: lenses; BS1 and BS2: beam splitters; PD: photodiode; RF1 and RF2: reference signal generators; LIA: lock-in amplifiers; AOM1 and AOM2: acousto-optic demodulators; BE: beam expander; M1 and M2: reflect mirrors; Obj: objective lens; SA: sample; ST: stage. (Figure reproduced from Ref. [87]).

When the profilometry is conducted, the objective lens is scanned in the longitudinal direction to obtain the defocus response curve, locating the sample surface. After that, scanning the sample in the horizontal direction and simultaneously detecting the feedback light intensity. When measuring the height difference  $\Delta h$  of two lateral positions on the sample surface, it obtains the integral number  $n$  of  $\lambda/2$  contained in  $\Delta h$  based on the amplitude ( $A$ ) variation of the laser intensity with the range of  $10 \mu\text{m}$ , and the fractional number  $s$  of  $\lambda/2$  using the phase measurement with a resolution of  $\approx 2 \text{ nm}$ . Finally, the height difference is computed as:

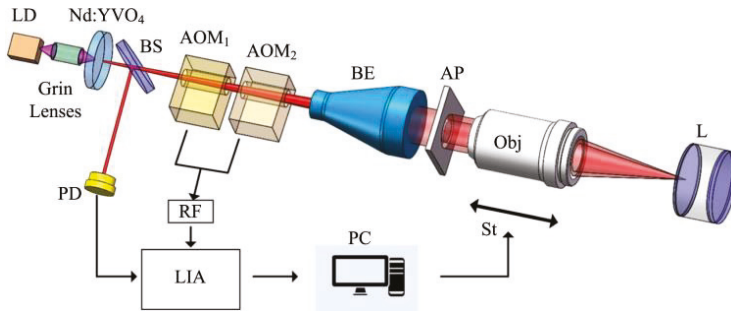
$$\Delta h = \zeta \times (n + s) \times \lambda/2 \tag{16}$$

where  $\zeta$  is a calibrated coefficient to take into consideration the influence of the high NA of the objective, whose value is carefully calibrated to be 0.93.

#### 4.3. Lens Thickness Measurement

Lens thickness measurements play an important role in the optical industry and optical systems. Various optical methods have been applied in the lens thickness noncontact measurement. A machine vision method [88] has advantages of continuous measurement, it is flexible in the production line, and the accuracy is less than 0.005 mm. Chromatic confocal sensors [89] are based on spectrally broadband light to realize a special optical probe. The measuring range and the accuracy are influenced by the designed probe and the dispersion of the measured object material. Low coherence interferometry [90] is utilized to measure the central thickness of soft and rigid contact lenses with the accuracy of a few micrometers. A laser differential confocal technique [91,92] realizes the high-precision measurement of the lens' axial space by detecting the absolute zero of an axial intensity curve. The measurement uncertainty is less than 0.05% for the thickness. Considering the high sensitivity of the laser feedback effect, it has great potential for the axial positioning, especially the multilayer lenses or the coated

lens with the high anti-reflecting film. Tan [93] proposed a new method to realize the lens thickness and air gap measurement, as shown in Figure 22. The annular pupil is inserted into the path before the objective lens to create the annular beam, in order to reduce the axial aberration and decrease the positioning error.

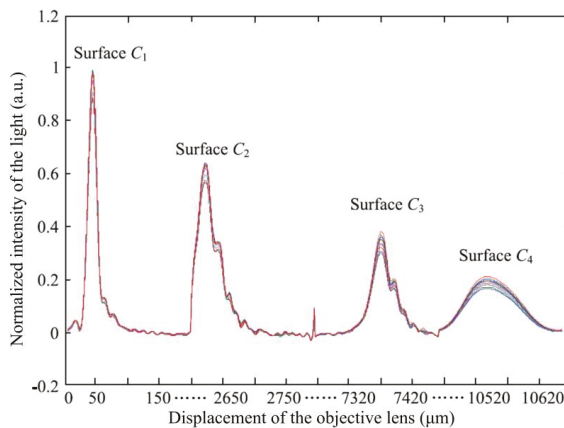


**Figure 22.** Schematic of the lens thickness measurement system. LD, laser diode; Grin Lenses, graded index lenses; BS, beam splitter; AOM<sub>1</sub> and AOM<sub>2</sub>, acousto-optic modulators; BE, beam expander; AP, annular pupil; Obj, objective lens; L, lenses; PD, photo detector; RF, reference signal generator; St, Stage; LIA, lock-in amplifier; PC, computer. (Figure reproduced from Ref. [93]).

When the objective lens moves along the optical axis, and the focused beam passes through the surfaces of the lenses, the defocus response curve can be obtained at each interface. The peak of the curve corresponds exactly to the beam focusing on the surface of the lens, as shown in Figure 23. Thus, the axial positioning can be realized via the synchronous acquisition of the light intensity and the displacement of the stage. Considering the refraction and reflection when the light passes through the lens, the relationship between the lens thickness and the axial positioning displacement can be deduced as:

$$d_i = \frac{\int_0^R T_i \cdot 2\pi\rho \cdot d\rho}{\pi R^2(1 - \epsilon^2)} \tag{17}$$

where  $\epsilon$  is the ratio of the inner and outer diameter of the AP,  $T_i$  is the ray tracing process function related to the lenses measured, and  $d_i$  is the displacement between each interface.



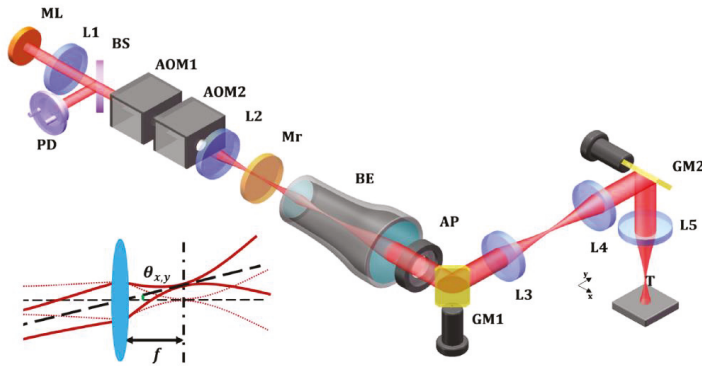
**Figure 23.** Experimental measuring results of two lenses. (Figure reproduced from Ref. [93]).

Different materials and kinds of lenses are measured in the experiment, including K9 plain glasses, fused silica plain glass, and K9 biconvex lens. The results prove that the uncertainty of the axial positioning is better than 0.0005 mm and the accuracy reaches the micron range.

4.4. Laser Confocal Feedback Imaging Combined with Other Technologies

4.4.1. Depth of Focus Extension in Laser Frequency-Shifted Feedback Imaging

For optical imaging, there is always a trade-off between the depth of focus (DOF) and the resolution [94,95]. Therefore, extending the DOF has drawn much attention, and various methods have been put forward. Numerically reconstructing a higher-resolved picture from a traditional beam is a main category of the methods, which is challenged by the decline of ballistic photons with the longer DOF. The laser feedback effect has the high sensitivity and could solve the problem. Thus, Lu [96] reported the depth of focus extension in laser frequency-shifted feedback imaging by filtering in the frequency domain. The system is shown in Figure 24. The measuring beam with the frequency shift after passing the BE is reflected by two galvanometric mirrors, then converged by objective lens L5; however, with the target located at the defocus plane.



**Figure 24.** Schematic diagram of the laser frequency-shifted feedback imaging system. ML, microchip laser; L1, collimator; BS, beam splitter; AOM1 and AOM2, acousto-optic modulators; L2, optical lens; Mr, reference mirror; AP, aperture; BE, beam expander; GM1 and GM2, galvanometric mirrors; L3 and L4, lenses forming the 4f system; L5, objective lens; T, target; PD, photodetector. The inset picture illustrates the physical meaning of the misalignment parameters  $\theta_x$  and  $\theta_y$ , where the dashed red line implies an ideal direction of the laser when  $\theta_x = \theta_y = 0$ . (Figure reproduced from Ref. [96]).

The original image is acquired via a 2-D scanning point by point, which can be described as the convolution of the point diffusion function (PSF) with the target plane. In theory, the Fourier transform of PSF (CTF) at a defocus length (DL) of  $L$  can be deduced as:

$$H(v, u) \propto \exp \left[ -\frac{\left(v - 2\frac{\theta_x}{\lambda}\right)^2 + \left(u - 2\frac{\theta_y}{\lambda}\right)^2}{\frac{2}{\pi^2 r_0^2}} \right] \times \exp \left\{ j\frac{\pi L \lambda}{2} \left[ -\left(v^2 + u^2\right) + 2\left(v\frac{2\theta_x}{\lambda} + u\frac{2\theta_y}{\lambda}\right) \right] \right\} \quad (18)$$

By filtering in the frequency domain, the new CTF can be obtained as:

$$H(v, u) \propto \exp \left[ -\frac{\left(v - 2\frac{\theta_x}{\lambda}\right)^2 + \left(u - 2\frac{\theta_y}{\lambda}\right)^2}{\frac{2}{\pi^2 r_0^2}} \right] \times \exp \left\{ j\pi L \lambda \left( v\frac{2\theta_x}{\lambda} + u\frac{2\theta_y}{\lambda} \right) \right\} \quad (19)$$

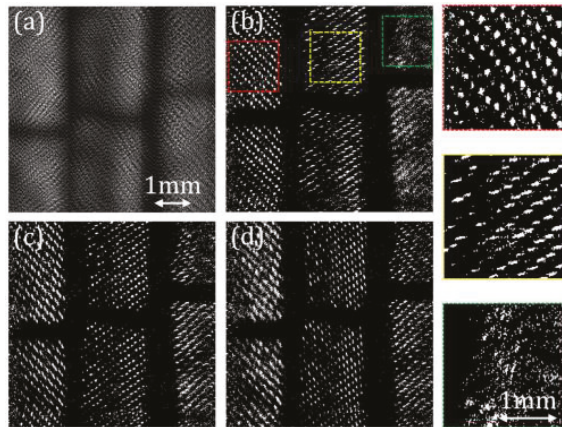


The inverse Fourier transform of the CTF corresponding to the new PSF can be calculated using:

$$h_a \propto \exp \left[ -\frac{(x + L\theta_x)^2 + (y + L\theta_y)^2}{\left(\frac{r_0}{\sqrt{2}}\right)^2} \right] \times \exp \left[ j2\pi \left( x \frac{2\theta_x}{\lambda} + y \frac{2\theta_y}{\lambda} \right) \right] \quad (20)$$

Therefore, a resolution of  $r_0/\sqrt{2}$  is obtained by filtering in the frequency domain, and it is possible to keep the resolution even in the far defocus plane.

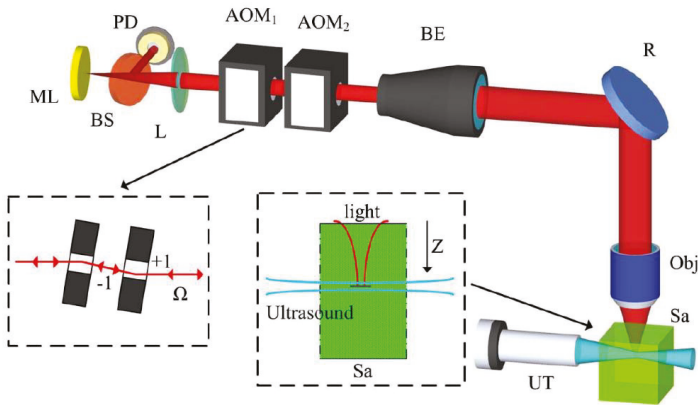
A three-dimensional target with three steps is tested in the experiment, and the processing of gradual refocusing is demonstrated as shown in Figure 25. The original blurred image is obtained using a 2-D scan. By filtering the image matrix in the frequency domain with a sequence of DL = 0.1:0.1:100 mm, three clear images of different steps can be obtained. The difference between the values of DL in filtering is exactly 5 mm, which equals the step height. In this way, the information along the optical axis can be obtained. The numerical experiments reveal that its depth of focus is capable of being extended to four times the length of the objective focal length.



**Figure 25.** Experiment results. (a) Image obtained by directly scanning. (b) Image after filtering with a DL of 48.5 mm. The three insets on the right are the magnified images that correspond to the rectangular regions labeled in the image. (c) Image after filtering with a DL of 53.5 mm. (d) Image after filtering with a DL of 58.5 mm. (Figure reproduced from Ref. [96]).

#### 4.4.2. Ultrasound Modulated Laser Confocal Feedback Imaging

When imaging objects inside turbid media, the number of ballistic photons decreases rapidly, compared to the scattered photons, as the penetration depth increases. As a result, the SNR is greatly reduced. The quality of the laser confocal feedback imaging is also affected by the scattered photons [39], and the generated speckles reduce the image contrast as well as the SNR. To solve the problem, Zhu [97] applied the ultrasound-modulated technology in the laser confocal feedback imaging. The system is shown in Figure 26. Except for the optical path focused into the sample, the ultrasound transducer is utilized, and the ultrasonic wave is focused into the same spot as well. As a result, the photons in the focal region are modulated with the shift frequency of the ultrasonic driving frequency [98,99]. Thus, the interesting photons are distinguished from the other scattered photons as the noise photons fall outside the focal region in the frequency domain. This is promising for reaching a larger penetration depth as well as a better SNR due to the reduction of noise photons.



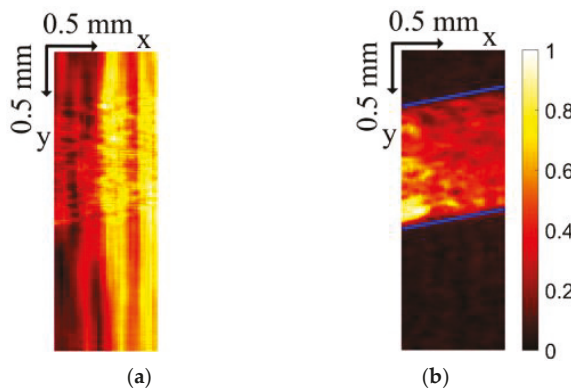
**Figure 26.** Schematic diagram of the ultrasound tagged laser confocal feedback imaging system. ML, microchip Nd: YVO<sub>4</sub> laser; BS, beam splitter; PD, photodiode; L, lens; AOM<sub>1</sub>, AOM<sub>2</sub>, acousto-optic modulators; BE, beam expander; R, reflector; Obj, objective; Sa, sample (phantom); UT, ultrasound transducer; Z, the axis of the incident light. (Figure reproduced from Ref. [97]).

When the measuring light returns to the laser cavity, the optical power modulation can be deduced as:

$$\frac{\Delta I(F_a + 2\Omega)}{I_s} = M_1 I(u) \kappa G(F_a + 2\Omega) \cdot \cos((F_a + 2\Omega)t - \phi + \phi_s) \quad (21)$$

where  $\Delta I$  denotes the intensity modulation of the measuring light,  $I_s$  is the output power of the solitary laser,  $M_1$  is the ultrasound one-sided modulation depth,  $I(u)$  is the light intensity of the confocal system,  $\kappa$  is the coefficient of the feedback light strength,  $G$  is the frequency dependent amplification factor,  $\phi_s$  is a fixed phase, and  $\phi$  is the phase related to the external cavity length.

The simulations and experiment are conducted in the paper. The contrasting experimental results with the traditional laser confocal feedback imaging are shown in Figure 27, proving that it is an effective method to realize a better SNR with the increasing of the scattering coefficient and focused depth. Compared with other optical methods, it is possible to reach both a larger imaging depth and a better SNR, especially in the turbid media.



**Figure 27.** Results of experimental imaging: (a) laser confocal feedback imaging, (b) ultrasound modulated laser confocal feedback imaging. (Figure reproduced from Ref. [97]).

## 5. Conclusions

In this paper, an overview of the laser feedback technology is presented. Two major configurations, the laser feedback interferometer and the laser confocal feedback tomography, are introduced theoretically, and a series of applications are described respectively.

The laser feedback interferometer has great potential because it does not need a retroreflector or corner prism set on the target. However, the stability and the performance indicators of the instrument are not as good as the He-Ne laser interferometer nowadays. Therefore, improving the performance of the laser feedback interferometer to be comparable with the traditional laser interferometer is a major research area, which includes speed improvement [44,45] and common path compensation [46–48]. On the other hand, research on the application is an important part to solve the scientific or industrial problem utilizing the ultra-high sensitivity, such as the measurement of refractive index, thermal expansion coefficient, liquid evaporation rate, etc. [66,72–76].

The laser confocal feedback tomography is proposed to reach a greater imaging depth compared with the confocal microscopy or optical coherence tomography [38,39]. It also shows great performance in the structure measurement of the micro-electro-mechanical system [80–82]. However, the system is also challenged to realize both high resolution and large penetration depth the same as the other optical methods. Therefore, combining other technologies with the advantages of the laser confocal feedback tomography is interesting research, such as the numerically reconstructing imaging to extend the DOF and ultrasound-modulated technology to improve the SNR [96,97].

In conclusion, the frequency-shifted optical feedback can be applied in a vast range of fields such as metrology, physical quantities measurement, imaging, microstructure measurement, and others. It is promising to be developed in further study, and in new fields combined with other technologies in future research.

**Author Contributions:** K.Z. initiated the manuscript. All authors participated in the revision of the manuscript. Y.T. provided overall supervision.

**Funding:** This research was funded by the National Science Fund for Excellent Young Scholars of China, grant number 51722506; the National Natural Science Foundation of China, grant number 61475082; and the National Natural Science Foundation of China, grant number 61775118.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

- King, P.G.R.; Steward, G.J. Metrology with an optical maser. *New Sci.* **1963**, *17*, 14.
- Taimre, T.; Nikolić, M.; Bertling, K.; Lim, Y.L.; Bosch, T.; Rakić, A.D. Laser feedback interferometry: A tutorial on the self-mixing effect for coherent sensing. *Adv. Opt. Photonics* **2015**, *7*, 570–631. [[CrossRef](#)]
- Li, J.; Niu, H.; Niu, Y. Laser feedback interferometry and applications: A review. *Opt. Eng.* **2017**, *56*, 050901. [[CrossRef](#)]
- Otsuka, K.; Kawai, R.; Asakawa, Y. Ultrahigh-sensitivity self-mixing laser Doppler velocimetry with laser-diode-pumped microchip LiNdP<sub>4</sub>O<sub>12</sub> lasers. *IEEE Photonics Technol. Lett.* **1999**, *11*, 706–708. [[CrossRef](#)]
- Timmermans, C.J.; Schellekens, P.H.J.; Schram, D.C. A phase quadrature feedback interferometer using a two-mode He-Ne laser. *J. Phys. E Sci. Instrum.* **1978**, *11*, 1023. [[CrossRef](#)]
- Besnard, P.; Jia, X.; Dalgliesh, R.; May, A.D.; Stephan, G. Polarization switching in a microchip Nd: YAG laser using polarized feedback. *J. Opt. Soc. Am. B* **1993**, *10*, 1605–1609. [[CrossRef](#)]
- Lang, R.; Kobayashi, K. External optical feedback effects on semiconductor injection laser properties. *IEEE J. Quantum Electron.* **1980**, *16*, 347–355. [[CrossRef](#)]
- Lenstra, D.; Van Vaalen, M.; Jaskorzynska, B. On the theory of a single-mode laser with weak optical feedback. *Phys. B+C* **1984**, *125*, 255–264. [[CrossRef](#)]
- Wang, W.M.; Boyle, W.J.O.; Grattan, K.T.V.; Palmer, A.W. Self-mixing interference in a diode laser: Experimental observations and theoretical analysis. *Appl. Opt.* **1993**, *32*, 1551–1558. [[CrossRef](#)]
- Bosch, T.; Servagent, N.; Lescure, M. A displacement sensor for spectrum analysis using the optical feedback in a single-mode laser diode. *Proc. IEEE Instrum. Meas. Technol. Conf.* **1997**, *2*, 870–873. [[CrossRef](#)]

11. Otsuka, K. Effects of external perturbations on LiNdP<sub>4</sub>O<sub>12</sub> lasers. *IEEE J. Quantum Electron.* **1979**, *15*, 655–663. [[CrossRef](#)]
12. Arecchi, F.T.; Lippi, G.L.; Puccioni, G.P.; Tredicce, J.R. Deterministic chaos in laser with injected signal. *Opt. Commun.* **1984**, *51*, 308–314. [[CrossRef](#)]
13. Wan, X.; Li, D.; Zhang, S. Quasi-common-path laser feedback interferometry based on frequency shifting and multiplexing. *Opt. Lett.* **2007**, *32*, 367–369. [[CrossRef](#)] [[PubMed](#)]
14. Zhu, K.; Guo, B.; Lu, Y.; Zhang, S.; Tan, Y. Single-spot two-dimensional displacement measurement based on self-mixing interferometry. *Optica* **2017**, *4*, 729–735. [[CrossRef](#)]
15. Otsuka, K. Self-mixing thin-slice solid-state laser Doppler velocimetry with much less than one feedback photon per Doppler cycle. *Opt. Lett.* **2015**, *40*, 4603–4606. [[CrossRef](#)]
16. Otsuka, K.; Abe, K.; Ko, J.Y.; Lim, T.S. Real-time nanometer-vibration measurement with a self-mixing microchip solid-state laser. *Opt. Lett.* **2002**, *27*, 1339–1341. [[CrossRef](#)] [[PubMed](#)]
17. Abe, K.; Otsuka, K.; Ko, J.Y. Self-mixing laser Doppler vibrometry with high optical sensitivity: Application to real-time sound reproduction. *New J. Phys.* **2003**, *5*, 8. [[CrossRef](#)]
18. Sudo, S.; Miyasaka, Y.; Nemoto, K.; Kamikariya, K.; Otsuka, K. Detection of small particles in fluid flow using a self-mixing laser. *Opt. Express* **2007**, *15*, 8135–8145. [[CrossRef](#)]
19. Zhang, S.; Tan, Y.; Zhang, S. Non-contact angle measurement based on parallel multiplex laser feedback interferometry. *Chin. Phys. B* **2014**, *23*, 114202. [[CrossRef](#)]
20. Szwaj, C.; Lacot, E.; Hugon, O. Large linewidth-enhancement factor in a microchip laser. *Phys. Rev. A* **2004**, *70*, 033809. [[CrossRef](#)]
21. Zhang, S.; Sun, L.; Tan, Y. Spectrum broadening in optical frequency-shifted feedback of microchip laser. *IEEE Photonics Technol. Lett.* **2016**, *28*, 1593–1596. [[CrossRef](#)]
22. Qin, C.; Feng, J.; Zhu, S.; Ma, X.; Zhong, J.; Wu, P.; Jin, Z.; Tian, J. Recent advances in bioluminescence tomography: Methodology and system as well as application. *Laser Photonics Rev.* **2014**, *8*, 94–114. [[CrossRef](#)]
23. Liba, O.; Lew, M.D.; SoRelle, E.D.; Dutta, R.; Sen, D.; Moshfeghi, D.M.; Chu, S.; de La Zerda, A. Speckle-modulating optical coherence tomography in living mice and humans. *Nat. Commun.* **2017**, *8*, 15845. [[CrossRef](#)] [[PubMed](#)]
24. Hillman, E.M.; Burgess, S.A. Sub-millimeter resolution 3D optical imaging of living tissue using laminar optical tomography. *Laser Photonics Rev.* **2009**, *3*, 159–179. [[CrossRef](#)] [[PubMed](#)]
25. Bychkov, A.; Simonova, V.; Zarubin, V.; Cherepetskaya, E.; Karabutov, A. The Progress in Photoacoustic and Laser Ultrasonic Tomographic Imaging for Biomedicine and Industry: A Review. *Appl. Sci.* **2018**, *8*, 1931. [[CrossRef](#)]
26. Wang, L.V.; Wu, H. *Biomedical Optics: Principles and Imaging*; John Wiley & Sons: Hoboken, NJ, USA, 2007; pp. 1–8, ISBN 978-0-471-74304-0.
27. Wilson, T. *Confocal Microscopy*; Academic Press: London, UK, 1990; pp. 1–64, ISBN 978-1-4899-1494-1.
28. Qiu, L.; Liu, D.; Zhao, W.; Cui, H.; Sheng, Z. Real-time laser differential confocal microscopy without sample reflectivity effects. *Opt. Express* **2014**, *22*, 21626–21640. [[CrossRef](#)] [[PubMed](#)]
29. Wang, Y.; Qiu, L.; Zhao, X.; Zhao, W. Divided-aperture differential confocal fast-imaging microscopy. *Meas. Sci. Technol.* **2017**, *28*, 035401. [[CrossRef](#)]
30. Durduran, T.; Choe, R.; Baker, W.B.; Yodh, A.G. Diffuse optics for tissue monitoring and tomography. *Rep. Prog. Phys.* **2010**, *73*, 076701. [[CrossRef](#)]
31. Eggebrecht, A.T.; Ferradal, S.L.; Robichaux-Viehoever, A.; Hassanpour, M.S.; Dehghani, H.; Snyder, A.Z.; Hershey, T.; Culver, J.P. Mapping distributed brain function and networks with diffuse optical tomography. *Nat. Photonics* **2014**, *8*, 448–454. [[CrossRef](#)]
32. Hong, G.; Antaris, A.L.; Dai, H. Near-infrared fluorophores for biomedical imaging. *Nat. Biomed. Eng.* **2017**, *1*, 0010. [[CrossRef](#)]
33. Sevcik-Muraca, E.M. Translation of near-infrared fluorescence imaging technologies: Emerging clinical applications. *Annu. Rev. Med.* **2012**, 217–231. [[CrossRef](#)] [[PubMed](#)]
34. Huang, D.; Swanson, E.A.; Lin, C.P.; Schuman, J.S.; Stinson, W.G.; Chang, W.; Hee, M.R.; Flotte, T.; Gregory, K.; Puliafito, C.A. Optical coherence tomography. *Science* **1991**, *254*, 1178–1181. [[CrossRef](#)] [[PubMed](#)]
35. Yi, L.; Sun, L.; Ding, W. Multifocal spectral-domain optical coherence tomography based on Bessel beam for extended imaging depth. *J. Biomed. Opt.* **2017**, *22*, 106016. [[CrossRef](#)]

36. De Boer, J.F.; Leitgeb, R.; Wojtkowski, M. Twenty-five years of optical coherence tomography: The paradigm shift in sensitivity and speed provided by Fourier domain OCT. *Biomed. Opt. Express* **2017**, *8*, 3248–3280. [[CrossRef](#)] [[PubMed](#)]
37. Lacot, E.; Day, R.; Stoeckel, F. Laser optical feedback tomography. *Opt. Lett.* **1999**, *24*, 744–746. [[CrossRef](#)] [[PubMed](#)]
38. Tan, Y.; Wang, W.; Xu, C.; Zhang, S. Laser confocal feedback tomography and nano-step height measurement. *Sci. Rep.* **2013**, *3*, 2971. [[CrossRef](#)] [[PubMed](#)]
39. Tan, Y.; Zhang, S.; Xu, C.; Zhao, S. Inspecting and locating foreign body in biological sample by laser confocal feedback technology. *Appl. Phys. Lett.* **2013**, *103*, 101909. [[CrossRef](#)]
40. Lacot, E.; Day, R.; Stoeckel, F. Coherent laser detection by frequency-shifted optical feedback. *Phys. Rev. A* **2001**, *64*, 043815. [[CrossRef](#)]
41. Lacot, E.; Hugon, O. Frequency-shifted optical feedback in a pumping laser diode dynamically amplified by a microchip laser. *Appl. Opt.* **2004**, *43*, 4915–4921. [[CrossRef](#)]
42. Tan, Y.; Xu, C.; Zhang, S.; Zhang, S. Power spectral characteristic of a microchip Nd:YAG laser subjected to frequency-shifted optical feedback. *Laser Phys. Lett.* **2013**, *10*, 025001. [[CrossRef](#)]
43. Gu, M. *Principles of Three Dimensional Imaging in Confocal Microscopes*; World Scientific Publishing Co. Inc.: Singapore, 1996; pp. 311–324, ISBN 981-02-2550-4.
44. Zhang, S.; Tan, Y.; Ren, Z.; Zhang, Y.; Zhang, S. A microchip laser feedback interferometer with nanometer resolution and increased measurement speed based on phase meter. *Appl. Phys. B* **2014**, *116*, 609–616. [[CrossRef](#)]
45. Zhang, S.; Tan, Y.; Zhang, S. High-speed Non-contact Displacement Sensor Based on Microchip Nd:YVO4 Laser Feedback Interferometry. In *Optical Sensors*; Optical Society of America: Washington, DC, USA, 2014; p. SeTh1D-4.
46. Ren, Z.; Li, D.; Wan, X.J.; Zhang, S. Quasi-common-path microchip laser feedback interferometry with a high stability and accuracy. *Laser Phys.* **2008**, *18*, 939–946. [[CrossRef](#)]
47. Zhang, S.; Zhang, S.; Tan, Y.; Sun, L. Self-mixing interferometry with mutual independent orthogonal polarized light. *Opt. Lett.* **2016**, *41*, 844–846. [[CrossRef](#)] [[PubMed](#)]
48. Zhang, S.; Zhang, S.; Tan, Y.; Sun, L. Common-path heterodyne self-mixing interferometry with polarization and frequency multiplexing. *Opt. Lett.* **2016**, *41*, 4827–4830. [[CrossRef](#)] [[PubMed](#)]
49. Zhang, S.; Zhang, S.; Sun, L.; Tan, Y. Fiber self-mixing interferometer with orthogonally polarized light compensation. *Opt. Express* **2016**, *24*, 26558–26564. [[CrossRef](#)] [[PubMed](#)]
50. Xu, L.; Tan, Y.; Zhang, S. Full path compensation laser feedback interferometry for remote sensing with recovered nanometer resolutions. *Rev. Sci. Instrum.* **2018**, *89*, 033108. [[CrossRef](#)]
51. Shi, L.; Kong, L.; Guo, D.; Xia, W.; Ni, X.; Hao, H.; Wang, M. Note: Simultaneous measurement of in-plane and out-of-plane displacement by using orthogonally polarized self-mixing grating interferometer. *Rev. Sci. Instrum.* **2018**, *89*, 096113. [[CrossRef](#)]
52. Hsieh, H.L.; Chen, J.C.; Lerondel, G.; Lee, J.Y. Two-dimensional displacement measurement by quasi-common-optical-path heterodyne grating interferometer. *Opt. Express* **2011**, *19*, 9770–9782. [[CrossRef](#)]
53. Yang, S.; Gao, Z.; Ruan, H.; Gao, C.; Wang, X.; Sun, X.; Wen, X. Non-Contact and Real-Time Measurement of Kolsky Bar with Temporal Speckle Interferometry. *Appl. Sci.* **2018**, *8*, 808. [[CrossRef](#)]
54. Barile, C.; Casavola, C.; Pappalettera, G.; Pappalettere, C. Remarks on residual stress measurement by hole-drilling and electronic speckle pattern interferometry. *Sci. World J.* **2014**, *2014*, 487149. [[CrossRef](#)]
55. Lu, H.; Cary, P.D. Deformation measurements by digital image correlation: Implementation of a second order displacement gradient. *Exp. Mech.* **2000**, *40*, 393–400. [[CrossRef](#)]
56. Hao, Y.; Bing, P. Three-dimensional displacement measurement based on the combination of digital holography and digital image correlation. *Opt. Lett.* **2014**, *39*, 5166–5169. [[CrossRef](#)]
57. Pfister, T.; Günther, P.; Nöthen, M.; Czarske, J. Heterodyne laser Doppler distance sensor with phase coding measuring stationary as well as laterally and axially moving objects. *Meas. Sci. Technol.* **2009**, *21*, 025302. [[CrossRef](#)]
58. Rothberg, S.J.; Allen, M.S.; Castellini, P.; Di Maio, D.; Dirckx, J.J.J.; Ewins, D.J.; Halkon, B.J.; Muyschondt, P.; Paone, N.; Ryan, T.; et al. An international review of laser Doppler vibrometry: Making light work of vibration measurement. *Opt. Laser Eng.* **2017**, *99*, 11–22. [[CrossRef](#)]



59. Ohtomo, T.; Sudo, S.; Otsuka, K. Three-channel three-dimensional self-mixing thin-slice solid-state laser-Doppler measurements. *Appl. Opt.* **2009**, *48*, 609–616. [[CrossRef](#)] [[PubMed](#)]
60. Otsuka, K. Long-haul self-mixing interference and remote sensing of a distant moving target with a thin-slice solid-state laser. *Opt. Lett.* **2014**, *39*, 1069–1072. [[CrossRef](#)] [[PubMed](#)]
61. Huang, Y.; Du, Z.; Deng, J.; Cai, X.; Yu, B.; Lu, L. A study of vibration system characteristics based on laser self-mixing interference effect. *J. Appl. Phys.* **2012**, *112*, 023106. [[CrossRef](#)]
62. Tao, Y.; Wang, M.; Xia, W. Semiconductor laser self-mixing micro-vibration measuring technology based on Hilbert transform. *Opt. Commun.* **2016**, *368*, 12–19. [[CrossRef](#)]
63. Dai, X.; Wang, M.; Zhao, Y.; Zhou, J. Self-mixing interference in fiber ring laser and its application for vibration measurement. *Opt. Express* **2009**, *17*, 16543–16548. [[CrossRef](#)]
64. Sudo, S.; Ohtomo, T.; Otsuka, K. Observation of motion of colloidal particles undergoing flowing Brownian motion using self-mixing laser velocimetry with a thin-slice solid-state laser. *Appl. Opt.* **2015**, *54*, 6832–6840. [[CrossRef](#)]
65. Ohtomo, T.; Sudo, S.; Otsuka, K. Detection and counting of a submicrometer particle in liquid flow by self-mixing microchip Yb: YAG laser velocimetry. *Appl. Opt.* **2016**, *55*, 7574–7582. [[CrossRef](#)] [[PubMed](#)]
66. Tan, Y.; Zhang, S.; Ren, Z.; Zhang, Y.; Zhang, S. Real-time liquid evaporation rate measurement based on a microchip laser feedback interferometer. *Chin. Phys. Lett.* **2013**, *30*, 124202. [[CrossRef](#)]
67. Ince, R.; Hüseyinoglu, E. Decoupling refractive index and geometric thickness from interferometric measurements of a quartz sample using a fourth-order polynomial. *Appl. Opt.* **2007**, *46*, 3498–3503. [[CrossRef](#)] [[PubMed](#)]
68. Harris, J.; Lu, P.; Larocque, H.; Xu, Y.; Chen, L.; Bao, X. Highly sensitive in-fiber interferometric refractometer with temperature and axial strain compensation. *Opt. Express* **2013**, *21*, 9996–10009. [[CrossRef](#)] [[PubMed](#)]
69. Lee, C.; Choi, H.; Jin, J.; Cha, M. Measurement of refractive index dispersion of a fused silica plate using Fabry–Perot interference. *Appl. Opt.* **2016**, *55*, 6285–6291. [[CrossRef](#)] [[PubMed](#)]
70. Li, X.; Tan, Q.; Bai, B.; Jin, G. Non-spectroscopic refractometric nanosensor based on a tilted slit-groove plasmonic interferometer. *Opt. Express* **2011**, *19*, 20691–20703. [[CrossRef](#)] [[PubMed](#)]
71. Cennamo, N.; Zeni, L.; Catalano, E.; Arcadio, F.; Minardo, A. Refractive Index Sensing through Surface Plasmon Resonance in Light-Diffusing Fibers. *Appl. Sci.* **2018**, *8*, 1172. [[CrossRef](#)]
72. Xu, L.; Zhang, S.; Tan, Y.; Sun, L. Simultaneous measurement of refractive-index and thickness for optical materials by laser feedback interferometry. *Rev. Sci. Instrum.* **2014**, *85*, 083111. [[CrossRef](#)]
73. Xu, L.; Tan, Y.; Zhang, S.; Sun, L. Measurement of refractive index ranging from 1.42847 to 2.48272 at 1064 nm using a quasi-common-path laser feedback system. *Chin. Phys. Lett.* **2015**, *32*, 090701. [[CrossRef](#)]
74. Xu, L.; Zhang, S.; Tan, Y.; Zhang, S.; Sun, L. Refractive index measurement of liquids by double-beam laser frequency-shift feedback. *IEEE Photonics Technol. Lett.* **2016**, *28*, 1049–1052. [[CrossRef](#)]
75. Zheng, F.; Ding, Y.; Tan, Y.; Lin, J.; Zhang, S. The approach of compensation of air refractive index in thermal expansion coefficients measurement based on laser feedback interferometry. *Chin. Phys. Lett.* **2015**, *32*, 070702. [[CrossRef](#)]
76. Zheng, F.; Tan, Y.; Lin, J.; Ding, Y.; Zhang, S. Study of non-contact measurement of the thermal expansion coefficients of materials based on laser feedback interferometry. *Rev. Sci. Instrum.* **2015**, *86*, 043109. [[CrossRef](#)] [[PubMed](#)]
77. Xu, H.; Liu, W.; Wang, T. Measurement of thermal expansion coefficient of human teeth. *Aust. Dent. J.* **1989**, *34*, 530–535. [[CrossRef](#)] [[PubMed](#)]
78. Casavola, C.; Lamberti, L.; Moramarco, V.; Pappaletta, G.; Pappaletta, C. Experimental analysis of thermo-mechanical behaviour of electronic components with speckle interferometry. *Strain* **2013**, *49*, 497–506. [[CrossRef](#)]
79. Li, C.; Liu, Q.; Peng, X.; Fan, S. Measurement of thermal expansion coefficient of graphene using optical fiber Fabry-Perot interference. *Meas. Sci. Technol.* **2016**, *27*, 075102. [[CrossRef](#)]
80. Xu, C.; Zhang, S.; Tan, Y.; Zhao, S. Inner structure detection by optical tomography technology based on feedback of microchip Nd:YAG lasers. *Opt. Express* **2013**, *21*, 11819–11826. [[CrossRef](#)]
81. Xu, C.; Tan, Y.; Zhang, S.; Zhao, S. The structure measurement of Micro-Electro-Mechanical System devices by the optical feedback tomography technology. *Appl. Phys. Lett.* **2013**, *102*, 221902. [[CrossRef](#)]
82. Wang, W.; Tan, Y.; Zhang, S.; Li, Y. Microstructure measurement based on frequency-shift feedback in a-cut Nd:YVO4 laser. *Chin. Opt. Lett.* **2015**, *13*, 121201. [[CrossRef](#)]

83. Bertling, K.; Perchoux, J.; Taimre, T.; Malkin, R.; Robert, D.; Rakić, A.D.; Bosch, T. Imaging of acoustic fields using optical feedback interferometry. *Opt. Express* **2014**, *22*, 30346–30356. [[CrossRef](#)]
84. Girardeau, V.; Jacquin, O.; Hugon, O.; Lacot, E. Ultrasound vibration measurements based on laser optical feedback imaging. *Appl. Opt.* **2018**, *57*, 7634–7643. [[CrossRef](#)]
85. Mowla, A.; Du, B.W.; Taimre, T.; Bertling, K.; Wilson, S.; Soyer, H.P.; Rakić, A.D. Confocal laser feedback tomography for skin cancer detection. *Biomed. Opt. Express* **2017**, *8*, 4037–4048. [[CrossRef](#)] [[PubMed](#)]
86. Hugon, O.; Joud, F.; Lacot, E.; Jacquin, O.; de Chatellus, H.G. Coherent microscopy by laser optical feedback imaging (LOFI) technique. *Ultramicroscopy* **2011**, *111*, 1557–1563. [[CrossRef](#)] [[PubMed](#)]
87. Wang, W.; Zhang, S.; Li, Y. Surface microstructure profilometry based on laser confocal feedback. *Rev. Sci. Instrum.* **2015**, *86*, 103108. [[CrossRef](#)] [[PubMed](#)]
88. Park, J.B.; Lee, J.G.; Lee, M.K.; Lee, E.S. A glass thickness measuring system using the machine vision method. *Int. J. Precis. Eng. Manuf.* **2011**, *12*, 769–774. [[CrossRef](#)]
89. Miks, A.; Novak, J.; Novak, P. Analysis of method for measuring thickness of plane-parallel plates and lenses using chromatic confocal sensor. *Appl. Opt.* **2010**, *49*, 3259–3264. [[CrossRef](#)] [[PubMed](#)]
90. Verrier, I.; Veillas, C.; Lépine, T. Low coherence interferometry for central thickness measurement of rigid and soft contact lenses. *Opt. Express* **2009**, *17*, 9157–9170. [[CrossRef](#)] [[PubMed](#)]
91. Wang, Y.; Qiu, L.; Yang, J.; Zhao, W. Measurement of the refractive index and thickness for lens by confocal technique. *Optik* **2013**, *124*, 2825–2828. [[CrossRef](#)]
92. Zhao, W.; Sun, R.; Qiu, L.; Shi, L.; Sha, D. Lenses axial space ray tracing measurement. *Opt. Express* **2010**, *18*, 3608–3617. [[CrossRef](#)]
93. Tan, Y.; Zhu, K.; Zhang, S. New method for lens thickness measurement by the frequency-shifted confocal feedback. *Opt. Commun.* **2016**, *380*, 91–94. [[CrossRef](#)]
94. Häusler, G. A method to increase the depth of focus by two step image processing. *Opt. Commun.* **1972**, *6*, 38–42. [[CrossRef](#)]
95. Fernández-Domínguez, A.I.; Liu, Z.; Pendry, J.B. Coherent fourfold super-resolution imaging with composite photonic-plasmonic structured illumination. *ACS Photonics* **2015**, *2*, 341–348. [[CrossRef](#)]
96. Lu, Y.; Zhu, K.; Li, J.; Zhang, S.; Tan, Y. Depth of focus extension by filtering in the frequency domain in laser frequency-shifted feedback imaging. *Appl. Opt.* **2018**, *57*, 5823–5830. [[CrossRef](#)] [[PubMed](#)]
97. Zhu, K.; Lu, Y.; Zhang, S.; Ruan, H.; Usuki, S.; Tan, Y. Ultrasound modulated laser confocal feedback imaging inside turbid media. *Opt. Lett.* **2018**, *43*, 1207–1210. [[CrossRef](#)] [[PubMed](#)]
98. Leutz, W.; Maret, G. Ultrasonic modulation of multiply scattered light. *Phys. B Condens. Matter* **1995**, *204*, 14–19. [[CrossRef](#)]
99. Wang, L.V. Mechanisms of ultrasonic modulation of multiply scattered coherent light: An analytic model. *Phys. Rev. Lett.* **2001**, *87*, 043903. [[CrossRef](#)] [[PubMed](#)]



© 2018 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).

# Comparison of Current Five-Point Cylindricity Error Separation Techniques

Wenwen Liu <sup>1,\*</sup>, Penghao Hu <sup>1</sup>  and Kuangchao Fan <sup>1,2</sup> 

<sup>1</sup> School of Instrument Science and Opto-electronic Engineering, Hefei University of Technology, Hefei 230009, China; hupenghao@hfut.edu.cn (P.H.); fan@ntu.edu.tw (K.F.)

<sup>2</sup> School of Mechanical Engineering, Dalian University of Technology, Dalian 116023, China

\* Correspondence: liuwenwen@hfut.edu.cn; Tel.: +86-138-6595-1024

Received: 8 September 2018; Accepted: 11 October 2018; Published: 16 October 2018



**Abstract:** Cylindricity is a kind of three-dimensional form distortion of a cylinder. An accurate in situ measurement of cylindricity is relatively complex because measuring and reconstructing cylindrical profile and evaluating out-of-cylindricity should be involved. Any method of in situ measuring cylindricity must solve a common issue, i.e., to eliminate spindle error motions and carriage error motions during measurement and reconstruction. Thus, error separation techniques have played an important role in in situ cylindricity measurement through multipoint detections. Although several valuable five-point methods for in situ measurement of cylindrical profile have been proposed up to present, namely the parallel scan, spiral scan, and V-block scan, there are obvious differences in many aspects, such as the arrangement of probes, error separation model, reconstruction method, adaptability to service environment, accuracy and reliability in practical application, etc. This paper presents the evaluation of their advantages and disadvantages in theory and the actual measurement based on the standard ISO 12180. Suggestions for best meeting the requirements of modern manufacturing and the most prospective one for industrial applications are also given.

**Keywords:** error separation technique; cylindricity; form measurement; in situ measurement

## 1. Introduction

Large rollers, whose sizes can sometimes reach a diameter of a few meters and a length of several meters, are employed in many heavy industries, e.g., papermaking, automobile, metallurgy, and shipbuilding. Such a roller should be characterized by high form accuracy, because its cylindricity errors would incorporate in the products, such as super calendar papers, roll-to-roll films, piston cylinders, and automobile steel sheets, etc. Using conventional roundness and cylindricity instruments may be impractical due to their large size and mass. Hence, there are increasing needs for accurate measurement of out-of-cylindricity of rollers in order to ensure the surface quality of the products. In situ measurements in production may be a good solution [1,2]. In general, a spindle supports the measured cylinder to rotate and a carriage slides along the longitudinal direction, which are indispensable in the in situ measuring system for the cylindricity. Raw probe signals representing radial deviations of the cylinder are always superposed by the radial error motions of the cylinder, which come from the spindle radial and tilt error motions, and the carriage straightness error motions [1–4]. Nevertheless, how to fully eliminate the radial error motions of the cylinder and straightness error motions of the carriage in the in situ measuring system is a great challenge [2,5]. Two basic issues are involved here: mathematical modeling of cylindrical profile and mathematical modeling of the measuring system.



1.1. Mathematical Modeling Cylindricity

The parameters and recommendations for measurement of cylindricity are described in the standard ISO 12180 [6]. The cylindrical profile can be considered as three superposed deviations: radial deviations, median line deviations, and cross-section deviations, as shown in Figure 1. Obviously, median line deviations denote out-of-straightness of the median line of a cylinder, and radial deviations and cross-section deviations denote the radius deviation and out-of-roundness of the cross-section, respectively.

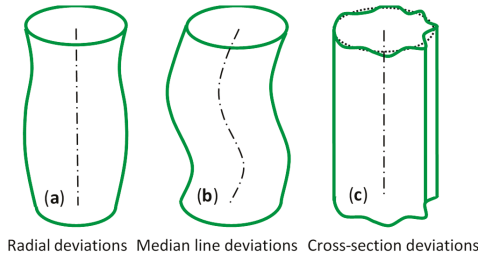


Figure 1. Components of the cylindricity deviations.

Cylindrical profile is made up of sequential cross-sectional profiles arranged in the axial direction, and each cross-sectional profile can be described as definite periodic signals with a base frequency equal to rotational frequency of the cylinder [5]. Therefore, the cylindrical profile  $r(z, \theta)$  can be described as a Fourier expansion:

$$r(z, \theta) = r_0(z) + \sum_{p=1}^{\infty} [a_p(z) \cos(p\theta) + b_p(z) \sin(p\theta)]. \tag{1}$$

where,  $z$  and  $\theta$  are respectively the axial and angular coordinates of the cylindrical profile, and  $r_0(z)$  is the radius of the cross-sectional profile.  $a_p(z) + jb_p(z)$ , which denotes the  $p$ th harmonic vector of the cross-sectional profile, can be expressed with the Legendre interpolation, i.e.,  $a_p(z) = \sum_{j=0}^{M_a} a_{pj}z^j$ , and  $b_p(z) = \sum_{j=0}^{M_b} b_{pj}z^j$ . Substituting  $a_p(z)$  and  $b_p(z)$  into Equation (1), the cylindrical profile  $r(z, \theta)$  can be rewritten as follows:

$$r(z, \theta) = r_0(z) + \left[ \left( \sum_{j=0}^{M_a} a_{1j}z^j \right) \cos(\theta) + \left( \sum_{j=0}^{M_b} b_{1j}z^j \right) \sin(\theta) \right] + \sum_{p=2}^{\infty} [a_p(z) \cos(p\theta) + b_p(z) \sin(p\theta)] \tag{2}$$

Term 2 in Equation (2) belongs to the first harmonic component of the cross-sectional profile. Let  $a_1(z) = \sum_{j=0}^{M_a} a_{1j}z^j$  and  $b_1(z) = \sum_{j=0}^{M_b} b_{1j}z^j$ , then,  $a_1(z) + jb_1(z)$  can be characterized as a vector from the rotating center in a cross-section of the cylinder to the least squares center (LSC) of the cross-sectional profile [5,7]. If the linear term  $(a_{10} + a_{11}z) + j(b_{10} + b_{11}z)$  about the Z-axis in  $a_1(z) + jb_1(z)$  is canceled, the cylinder as a whole will be tilted and moved, and the cylindrical profile will be not changed. Thus, the mathematical model of the cylindrical profile can be expressed by Equation (3).

$$r(z, \theta) = r_0(z) + r_1(z, \theta) + r_2(z, \theta) \tag{3}$$

where,  $r_1(z, \theta) = a_1(z) \cos(\theta) + b_1(z) \sin(\theta)$  is the first harmonic component of the cross-sectional profile. Obviously, a median line of the cylindrical profile, as shown Figure 1b, can be fitted through the LSC vectors  $a_1(z) + jb_1(z)$  in sequential cross-sections [5,6,8]. The 0th harmonic component,  $r_0(z)$ , represents average radius deviations of sequential cross-sectional profiles, as shown Figure 1a. The sum of the  $p$ th ( $p = 2-\infty$ ) harmonic components, denoted by  $r_2(z, \theta) = \sum_{p=2}^{\infty} [a_p(z) \cos(p\theta) + b_p(z) \sin(p\theta)]$ , represents the out-of-roundness of sequential cross-sectional profiles, as shown Figure 1c.

From Equation (3), a feasible method of measuring and reconstructing a cylindrical profile can be suggested. It is to accurately measure the radius deviations, roundness errors, and LSC vectors of sequential cross-sections. Moreover, the median line is fitted through the LSCs, and cylindrical profile can be reconstructed by assembling the measured radius deviation and out-of-roundness of each cross-section on the fitted median line.

1.2. Mathematical Modeling of the Measuring System

Cylindricity errors are three-dimensional form deviations, and accurate measurement of cylindricity is relatively complex, should include measuring and reconstructing cylindrical profile, and evaluating out-of-cylindricity. Any in situ measuring system must employ several probes mounted onto a carriage to sense the radial deviations of the cylinder while the cylinder is rotating and the carriage is sliding along the longitudinal direction. It is known that rotating of the cylinder or sliding of the carriage have inevitable five degree-of-freedom geometric error motions due to their manufacturing and assembly errors. Therefore, the method of measuring cylindrical profile must be able to eliminate error motions of the cylinder and carriage in data analysis. Error separation techniques (ESTs) have been used in measurement of the form errors, such as straightness, flatness, roundness, and so on. Error separation techniques ensure measurement accuracy by separating form errors and error motions with specific designed setup and special mathematical modeling. In general, the three-point roundness EST [3,4,9] can remove the radial error motions in the cross-section of the cylinder and accurately determine radius deviations and out-of-roundness of the cross-sectional profile of the cylinder [3,5,9]. However, the first harmonic of the cross-sectional profile are inseparable from that of the radial error motions in the cross-section of the cylinder [3–5,8,9]. The first harmonic of the cross-sectional profile denotes the vector from rotating center to the LSC of the cross-sectional profile. Median line of the cylinder is a spatial curve that is fitted through the LSCs of sequential cross-sectional profiles [6,8]. Therefore, three-point roundness EST is insufficient to accurately determine a cylindrical profile.

Modeling of a probe sensing the cylindrical profile is particularly foundational for cylindrical profile measurement. Reference [10] built such a measurement model, as shown in Figure 2. OXYZ is the global coordinate system, in which the cylinder coordinate system  $O_cX_cY_cZ_c$  and the carriage coordinate system  $O_sX_sY_sZ_s$  are located. Theoretically, there are ten terms of error motions, including the cylinder’s translational error motions  $e_{cx}$ ,  $e_{cy}$ ,  $e_{cz}$  and tilt error motions  $\tau_{cx}$ ,  $\tau_{cy}$ , as well as the carriage’s straightness error motions  $e_{sx}$ ,  $e_{sy}$  and three angular error motions of pitch  $\tau_{sx}$ , yaw  $\tau_{sy}$ , and roll  $\tau_{sz}$ . The vectors of the  $O_c$  and  $O_s$  in the OXYZ system are  $\mathbf{e}_c = [e_{cx}, e_{cy}, 0]$  and  $\mathbf{e}_s = [e_{sx}, e_{sy}, z_s]$ , respectively. Let the rotation angle of the cylinder be  $\theta$ . In the  $O_sX_sY_sZ_s$  system, the measuring point P locates at  $\mathbf{r}^s = [r_0 \cos(\varphi), r_0 \sin(\varphi), z_s]$ , and point P in the OXYZ system would be offset due to the carriage error motions. If  $e_{cz}$ ,  $\tau_{sz}$ , and quadratic terms of the errors are omitted, the actual position of point P in the OXYZ system is:

$$\zeta_x = r_0 \cos(\varphi) + e_{sx} + z_s \tau_{sy}; \zeta_y = r_0 \sin(\varphi) + e_{sy} + z_s \tau_{sx}.$$

In a similar way, in the  $O_cX_cY_cZ_c$  system, because of point P’s offsets caused by the error motions of the cylinder, the actual position of point P in the OXYZ system is:

$$\begin{aligned} \tilde{\zeta}_x &= r(z + z_s, \theta + \varphi) \cos(\varphi) + e_{cx} + (z + z_s) \tau_{cy}; \\ \tilde{\zeta}_y &= r(z + z_s, \theta + \varphi) \sin(\varphi) + e_{cy} + (z + z_s) \tau_{cx}. \end{aligned}$$

Obviously, the output of the probe indicates their difference in OXYZ system:

$$t(z + z_s, \theta) = (\tilde{\zeta}_x - \zeta_x) \cos(\varphi) + (\tilde{\zeta}_y - \zeta_y) \sin(\varphi) = [r(z + z_s, \theta + \varphi) - r_0] + [e_{cx} + z \tau_{cy} - e_{sx} + z_s (\tau_{cy} - \tau_{sy})] \cos(\varphi) + [e_{cy} + z \tau_{cx} - e_{sy} + z_s (\tau_{cx} - \tau_{sx})] \sin(\varphi). \tag{4}$$

where,  $[r(z + z_s, \theta + \varphi) - r_0]$  is the cylindrical profile. Equation (4) indicates that output of the probe is a mixture of cylindrical profile and all kinds of error motions in measuring system, in which  $[e_{cx} + z\tau_{cy} - e_{sx}]$ ,  $[z_s(\tau_{cy} - \tau_{sy})]$ ,  $[e_{cy} + z\tau_{cx} - e_{sy}]$ , and  $[z_s(\tau_{cx} - \tau_{sx})]$  are the mutual independent error motions. Therefore, the minimum number of the probes in order to fully eliminate the error motions is five [10]. From Figure 2 and Equation (4), the three-point roundness EST [3–5,9] arranges three probes in one cross-section of the cylinder, and the error motions  $[e_{cx} + z\tau_{cy} - e_{sx}]$  and  $[e_{cy} + z\tau_{cx} - e_{sy}]$  are sensed by the probe, and can be canceled. If all error motions are expected to be eliminated, a few probes must be arranged in the axial direction to sense the error motions  $[z_s(\tau_{cx} - \tau_{sx})]$  and  $[z_s(\tau_{cy} - \tau_{sy})]$ .

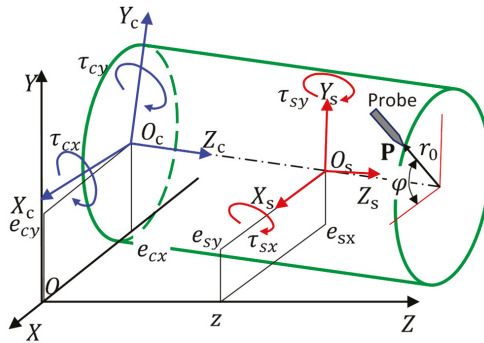


Figure 2. Probe senses cylindrical profile and error motions [10].

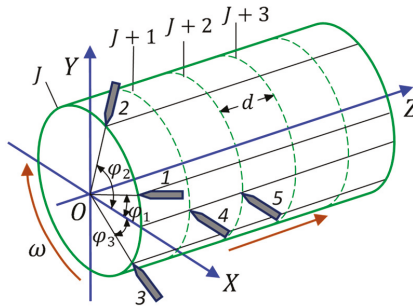
The literature on the subject of in situ measurement of cylindricity is limited. Nyberg [11] proposed the four-point EST for in situ measurement of cylindricity, which combines the concepts of two-point straightness and three-point roundness ESTs. The signals from the three probes were used to calculate radius deviations and out-of-roundness of each cross-sectional profile. Whereas the signals from the two probes placed at the same generatrix in the cylinder were employed to determine out-of-straightness of the generatrix. However, this method sidesteps the problem of accurately determining curved median line profile of the cylinder, although it was successfully verified on a dedicated measuring device by the experiments.

Several valuable five-point ESTs for in situ measurement of cylindrical profile have been reported [8,10,12,13]. They are clearly different in the arrangement of probes, error separation model, reconstruction method, adaptability to service environment, and accuracy of measuring and reconstructing. Related technical notes are presented in the following for comparing and evaluating pros and cons. Suggestions for best meets the requirements of modern manufacturing and the most prospective one for industrial applications will then be made.

## 2. Error Separation Models for In Situ Measurement of Cylindricity

### 2.1. Parallel Scan EST

Reference [8] presented a parallel scan EST for measuring cylindrical profile, as shown in Figure 3. Let the cylinder be divided by  $M$  cross-sections, and the axial distance between each two cross-sections be  $d$ . Five displacement probes are mounted onto the carriage, which moves along the  $Z$  direction. When the carriage is located at the  $J$ th ( $J = 1, 2, \dots, M$ ) position and the cylinder rotates one revolution, Probes 1–3 sense the mixed radial errors in the  $J$ th cross-section, meanwhile, Probes 4 and 5 sense the  $X$  directional mixed errors in the  $(J + 1)$ th and  $(J + 2)$ th cross-sections. When data collection is completed, the carriage moves a distance  $d$  along the  $Z$  direction and enters the  $(J + 1)$ th position. At this moment, Probes 1–3 locates in the  $(J + 1)$ th cross-section and Probes 4 and 5 locate in the  $(J + 2)$ th and  $(J + 3)$ th cross-sections, respectively. Data collection of new revolution will be carried on.



**Figure 3.** Schematic of parallel scan EST for measurement of cylindrical profile. The 1–5 are the displacement probes.

When the carriage locates at the  $J$ th ( $J = 1, 2, \dots, M$ ) position, the cylinder is rotating and Probes 1–5 collect data. It was verified that the  $p$ th ( $p = 0, 2, 3, \dots, N - 2$ ) harmonic vector  $\mathbf{R}(z_J, p)$  of the  $J$ th cross-sectional profile  $r(z_J, i)$  ( $i = 0, 1, \dots, N - 1$ ) of the cylinder could be detected with the three-point roundness EST [3–5,8,9] by the outputs of Probes 1–3, denoted by  $t_n(z_J, i)$  ( $n = 1, 2, 3$ ) ( $i = 0, 1, \dots, N - 1$ ). Therefore,  $r_0(z_J)$  and  $r_2(z_J, i)$  ( $i = 0, 1, \dots, N - 1$ ), indicating the radius deviation and out-of-roundness of the  $J$ th cross-sectional profile, can be accurately determined. In addition,  $\hat{e}_x(z_J, i)$  ( $i = 0, 1, \dots, N - 1$ ), indicating the separated result of the X directional error motions in the  $J$ th cross-section of the cylinder, can be obtained by the outputs of Probes 1 and 2, as shown in Ref. [8,14]. The first harmonic vector should be  $\hat{\mathbf{E}}_x(z_J, 1) = \text{DFT}[\hat{e}_x(z_J, i)]|_{p=1}$ , here, DFT is the discrete Fourier transform operator.

Simultaneously, Probes 4 and 5 sense the X directional mixed errors, including the  $(J + 1)$ th and  $(J + 2)$ th cross-sectional profiles, denoted by  $r(z_{J+1}, i)$  and  $r(z_{J+2}, i)$  ( $i = 0, 1, \dots, N - 1$ ), the X directional error motions in the  $J$ th cross-section  $e_{cx}(z_J, i)$  and the tilt error motion of the cylinder around the Y direction  $\tau_{cy}(J, i)$ , as well as the X directional straightness deviations of the carriage  $e_{sx}(z_J)$  and yawing  $\tau_{cy}(J)$ , as shown in Figures 2 and 3. From Equation (4), the outputs of Probes 4 and 5, denoted by  $t_n(J, i)$  ( $n = 4, 5$ ) ( $i = 0, 1, \dots, N - 1$ ), are given by

$$t_4(J, i) = r(z_{J+1}, i) - D_4 + \{[e_{cx}(z_J, i) + d \times \tau_{cy}(J, i)] - [e_{sx}(z_J) + d \times \tau_{sy}(J)]\} \quad (5a)$$

$$t_5(J, i) = r(z_{J+2}, i) - D_5 + \{[e_{cx}(z_J, i) + 2d \times \tau_{cy}(J, i)] - [e_{sx}(z_J) + 2d \times \tau_{sy}(J)]\}. \quad (5b)$$

Applying Discrete Fourier transform (DFT) to  $t_4(J, i)$  and  $t_5(J, i)$ , we may obtain the  $p$ th ( $p = 0, 1, \dots, N - 1$ ) harmonic vectors of the outputs of Probes 4 and 5, denoted by  $\mathbf{T}_4(J, p) = \text{DFT}[t_4(J, i)]$  and  $\mathbf{T}_5(J, p) = \text{DFT}[t_5(J, i)]$ , in which the first harmonic vectors are  $\mathbf{T}_4(J, 1)$  and  $\mathbf{T}_5(J, 1)$ , respectively.

$\Theta_{cx}(J, 1)$  ( $J = 1, 2, \dots, M$ ) is designated as the first harmonic vector of the X directional additional error motions induced by the tilt error motions of the cylinder around the Y-axis. Referencing the sequential three-point straightness EST [5,15], the iterative formulas used to extract  $\mathbf{R}(z_J, 1)$  ( $J = 1, 2, \dots, M$ ), representing the LSC vectors of each cross-sectional profile, were derived in in Appendix of Reference [14]. When the carriage locates at the first position, let  $\Theta_{cx}(1, 1)$  and  $\mathbf{R}(z_1, 1)$  be equal to “zero”. The first harmonic vector  $\mathbf{R}(z_J, 1)$  of the  $J$ th ( $J = 2, 3, \dots, M$ ) cross-sectional profile is extracted by:

$$\mathbf{R}(z_J, 1) = \mathbf{R}(z_{J-1}, 1) + \mathbf{T}_4(J - 1, 1) - \hat{\mathbf{E}}_x(z_{J-1}, 1) - \Theta_{cx}(J - 1, 1) \quad (6a)$$

$$\Theta_{cx}(J, 1) = \Theta_{cx}(J - 1, 1) - [\mathbf{T}_5(J - 1, 1) - \mathbf{T}_4(J - 1, 1)] + [\mathbf{T}_4(J, 1) - \hat{\mathbf{E}}_x(z_J, 1)]. \quad (6b)$$

The LSC vector of the  $J$ th ( $J = 1, 2, \dots, M$ ) cross-sectional profile is  $2R(z_J, 1)/N$  or  $2R(z_J, N - 1)/N$ , with which a curved median line of the cylinder can be accurately fitted, here,  $R(z_J, N - 1) = \text{conjugation}[R(z_J, 1)]$ . It is noteworthy that the curved median line fitted through the LSCs of each cross-sectional profile, as a whole, undergoes little movement and inclination because  $R(z_1, 1)$  and  $\Theta_{cx}(1, 1)$  are all set to “zero”, rather than their truth-values. The correctness of the curved median line profile will not be affected.

Now, by means of the three-point roundness EST and Equation (6),  $R(z_J, p)$  ( $p = 0, 2, 3, \dots, N - 2$ ) as well as  $R(z_J, 1)$  and  $R(z_J, N - 1)$ , indicating the  $p$ th ( $p = 0, 1, 2, \dots, N - 1$ ) harmonic vectors of the  $J$ th ( $J = 1, 2, \dots, M$ ) cross-sectional profiles, can be detected accurately. A reconstructed cylindrical profile, denoted by  $r(z_J, i)$  ( $J = 1, 2, \dots, M$ ) ( $i = 0, 1, \dots, N - 1$ ), can be achieved by;

$$r(z_J, i) = \text{IDFT}[R(z_J, p)] \quad (i, p = 0, 1, 2, \dots, N - 1). \tag{7}$$

Where, IDFT is the inverse discrete Fourier transform operator. Theoretical analysis and numerical validation about this EST have been performed in Reference [8]. The results verify that error motions of the cylinder and carriage are all removed, i.e., full harmonic error separation is realized. The LSCs of each cross-sectional profile of the cylinder are accurately extracted even if the error motions of the cylinder are not repeatable in each rotation. The spatial curved median line of the cylinder is accurately determined by fitting through the LSCs of the sequential cross-sectional profiles. The cylindrical profile conforming to the standard ISO 12180 [6], can then be reconstructed by Equation (7), as shown in Equation (3). The cylindrical form errors, not only cylindricity but also out-of-roundness, generatrix straightness, taper angle, and radius deviation, can be evaluated simultaneously by using the reconstructed cylindrical profile.

### 2.2. Spiral Scan EST

Reference [10] presented the spiral five-probe arrangement for measurement of cylindrical profile, as shown in Figure 4. Five displacement probes arranged in a spiral scanning-path are mounted onto the carriage, which moves along the Z direction. This system makes it possible to transform the cylindrical profile measurement to the one-dimensional straightness measurement of the spiral profile on the cylinder given that the rotation of the cylinder keeps pace with the movement of the carriage. Let the total number of scanning points per revolution be  $N$ , and the scanning angular interval should be  $\delta = 2\pi/N$ . When probes move along the spiral line, the variable  $z$  of the carriage moving along the Z direction is proportional to the variable  $\theta = i \times \delta$  of the cylinder rotating, i.e.,  $z = \theta \times T/(2\pi) = i \times T/N$ , where, constant  $T$  is the pitch of the spiral line. The angular positions of Probes 1–5 are  $\varphi_n$  ( $n = 1, 2, \dots, 5$ ) with respect to the X direction, and axial positions are  $z_{sn} = \tau \times \varphi_n/(2\pi)$  ( $n = 1, 2, \dots, 5$ ), respectively. When the cylinder rotates and the carriage moves synchronously, several error motions come up as shown in Figure 2, including the cylinder’s radial error motions  $e_{cx}(i)$ ,  $e_{cy}(i)$  and tilt error motions  $\tau_{cx}(i)$ ,  $\tau_{cy}(i)$ , as well as the carriage straightness error motions  $e_{sx}(i)$ ,  $e_{sy}(i)$  and pitching  $\tau_{sx}(i)$ , yawing  $\tau_{sy}(i)$  ( $i = 0, 1, \dots, N \times M - 1$ ). Here,  $M$  is number of total rotating circle of the spiral line on the cylinder. Let  $r(i)$  ( $i = 0, 1, \dots, N \times M - 1$ ) be the spiral line profile on the cylinder. According to Equation (4), the outputs of the five probes, denoted by  $t_n(i)$  ( $n = 1, 2, \dots, 5$ ) ( $i = 0, 1, \dots, N \times M - 1$ ), can be written as follow:

$$t_n(i) = r(i + k_n) - D_n + \{[e_{cx}(i) - e_{sx}(i) + z\tau_{cy}(i)] + z_{sn}[\tau_{cy}(i) - \tau_{sy}(i)]\} \cos(\varphi_n) + \{[e_{cy}(i) - e_{sy}(i) + z\tau_{cx}(i)] + z_{sn}[\tau_{cx}(i) - \tau_{sx}(i)]\} \sin(\varphi_n). \tag{8}$$

where,  $k_n = \varphi_n/\delta$  ( $n = 1, 2, \dots, 5$ );  $D_n$  ( $n = 1, 2, \dots, 5$ ) are the initial zero-setting values of the probes. Similarly to the three-point straightness EST based on DFT [5,16], the weighted output of Probes 1–5 should be  $t(i) = \sum_{n=1}^5 c_n t_n(i)$ . Let the weighted coefficients  $c_n$  ( $n = 1, 2, \dots, 5$ ) be  $\sum_{n=1}^5 c_n \cos(\varphi_n) = 0$  and  $\sum_{n=1}^5 c_n \sin(\varphi_n) = 0$ , and  $\sum_{n=1}^5 c_n z_{sn} \cos(\varphi_n) = 0$ ,  $\sum_{n=1}^5 c_n z_{sn} \sin(\varphi_n) = 0$ . The weighted output  $t(i)$  ( $i = 0, 1, \dots, N \times M - 1$ ) can be rewritten as:  $t(i) = \sum_{n=1}^5 c_n r(i + k_n) - D$ .

Where,  $D = \sum_{n=1}^5 c_n D_n$  is the weighted initial zero-setting value. Obviously, all error motions of the cylinder and carriage are removed from  $t(i)$ . Moreover, applying the DFT to  $t(i)$  and introducing the time-shift theorem, we can obtain the  $p$ th ( $p = 0, 1, \dots, N \times M - 1$ ) harmonic vectors of the weighted output  $t(i)$  ( $i = 0, 1, \dots, N \times M - 1$ ) as follows:

$$\mathbf{T}(p) = \text{DFT}[t(i)] = \mathbf{R}(p) \times \mathbf{\Omega}(p) - \mathbf{D}(p); \mathbf{\Omega}(p) = \sum_{n=1}^5 c_n \exp\left(\frac{jp\varphi_n}{M}\right). \tag{9}$$

where,  $\mathbf{D}(p) = \text{DFT}[D]$ . If the transfer function  $\mathbf{\Omega}(p) \neq 0$  ( $p = 0, 1, \dots, N \times M - 1$ ), the  $p$ th ( $p = 0, 1, \dots, N \times M - 1$ ) harmonic vectors of the spiral line profile on the cylinder are:

$$\mathbf{R}(p) = \frac{\mathbf{T}(p) + \mathbf{D}(p)}{\mathbf{\Omega}(p)}. \tag{10}$$

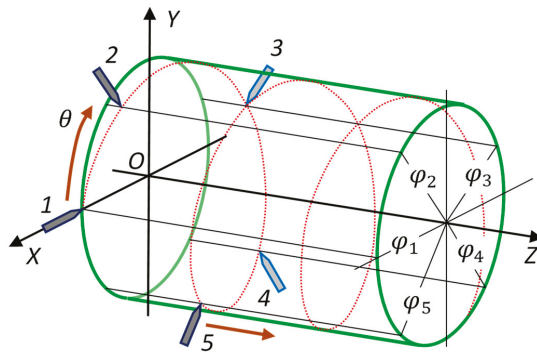


Figure 4. Schematic of spiral scan EST [10]. The 1–5 are the displacement probes.

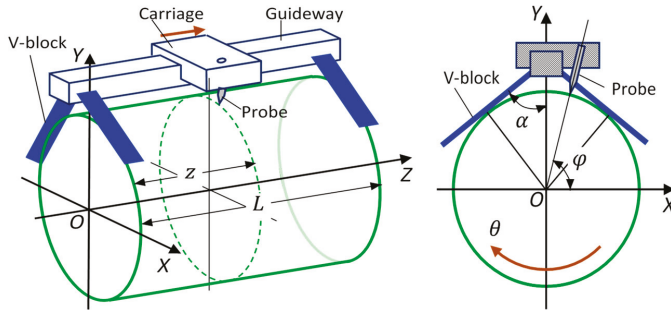
According to the properties of the DFT,  $\mathbf{D}(p)|_{p=0} \neq 0$  and  $\mathbf{D}(p)|_{p \neq 0} = 0$ . Similar to the parallel scan EST, the initial zero-setting value  $D_1$ – $D_5$  of Probes 1–5 is also allocated a value of zero or the nominal radius, which will not affect the spiral line profile on the cylinder. Applying the IDFT to  $\mathbf{R}(p)$  ( $p = 0, 1, \dots, N \times M - 1$ ), we may obtain  $\rho(i) = \text{IDFT}[\mathbf{R}(p)]$  ( $i, p = 0, 1, \dots, N \times M - 1$ ), indicating the spiral line profile on the cylinder. A reconstructed cylindrical profile should be:

$$r(z, \ddot{i}) = \rho(i) (i = 0, 1, \dots, N \times M - 1); z = i \times \frac{t}{N}; \ddot{i} = \text{remainder}(i/N) (\ddot{i} = 0, 1, \dots, N - 1). \tag{11}$$

It's worth noting that, different from the parallel scan EST, the spiral scan EST measures the one-dimensional straightness of the spiral profile on the cylinder, and reconstructs an integral cylindrical profile. It is difficult to distinguish the radius deviation and out-of-roundness of each cross-section, as well as the profile of the spatial curved median line of the cylinder.

### 2.3. V-Block Scan EST

Ref. [1,12,13,17] presented the V-block scan EST for measurement of cylindrical profile as shown in Figure 5. An element connecting the two V-blocks functions as guideway, along which a carriage moves. A displacement probe is mounted onto the carriage. When the cylinder driven by the spindle rotates, probe keeps collecting data. At the moment, two V-blocks should contact reliably with the surface of the cylinder, and the measuring system should avoid rotating around and moving in the Z-axis. Because five points are contact with the measured surface during measuring, this system may be regarded as an analogous five-point EST measuring system.



**Figure 5.** Schematic of V-block scan EST [1,12,13]. The left shows the measured cylinder in the OXYZ system; the right shows the cross-section of the cylinder in the plane of the V-block.

As shown in Figure 5, the OXYZ system was assumed in such a way that the Z-axis coincides with the axis of the nominal cylinder, and the plane determined by the X- and Y-axes coincide with the plane of the left V-block. The Y-axis is the angular bisector of the V-block. Let the total number of scanning points per revolution be  $N$ , and the scanning angular interval should be  $\delta = 2\pi/N$ . Let the angular position of the probe be  $\varphi = k\delta$  with respect to the X-axis, and the half angle of the V-block be  $\alpha = u\delta$ . Because the system rides on the cylinder during measuring, the radial and tilt error motions of the cylinder will have nothing to do with the output of the probe. However, as those two cross-sectional profiles that are contacted by two V-blocks have roundness errors, it would make the measuring system generate the radial and tilt error motions. At this moment, the probe situated in measured cross-section will collect the radial mixed errors, including the cross-sectional profile  $r(z, i)$  ( $i = 0, 1, \dots, N - 1$ ), the measuring system's radial error motions  $e_x(z, i)$ ,  $e_y(z, i)$  ( $i = 0, 1, \dots, N - 1$ ) in the cross-section where the probe is, as well as the carriage straightness error motions  $e_{sx}(z)$ ,  $e_{sy}(z)$  in the location where the probe is. According to Equation (4), the output of the probe  $t(z, i)$  ( $i = 0, 1, \dots, N - 1$ ) should be:

$$t(z, i) = r(z, i + k) + [e_x(z, i) - e_{sx}(z)] \cos(\varphi) + [e_y(z, i) - e_{sy}(z)] \sin(\varphi). \tag{12}$$

where, the initial zero-setting value of the probe is omitted due to without impact on accurate measurement. It is worth noting that  $e_x(z, i)$ ,  $e_y(z, i)$  ( $i = 0, 1, \dots, N - 1$ ), indicating the measuring system's radial error motions in measured cross-section, are caused by the cross-sectional profiles in contacted with the two V-blocks. There are obviously four points on the V-blocks in contact with the cross-sections defined by the plane  $z = 0, L$  of the cylinder. If the contacted points are regarded as the virtual probes, output of the virtual probes in the  $z = 0, L$  cross-sections should be "zero". It is obvious that the  $z = 0, L$  cross-sectional profiles  $r(z, i)|_{z=0,L}$  ( $i = 0, 1, \dots, N - 1$ ) control the radial error motions of the measuring system in the  $z = 0, L$  cross-sections, denoted by  $e_x(z, i)|_{z=0,L}$ ,  $e_y(z, i)|_{z=0,L}$  ( $i = 0, 1, \dots, N - 1$ ).

At this moment, if the carriage moves, and the probe locates in the  $z = 0, L$  cross-sections, respectively, the probe collects the radial mixed errors during cylinder rotating. The outputs of the probe should be  $t(z, i)|_{z=0,L}$  ( $i = 0, 1, \dots, N - 1$ ), as shown in Equation (12). Taking DFT to  $t(z, i)|_{z=0,L}$  ( $i = 0, 1, \dots, N - 1$ ), we will achieve the  $p$ th ( $p = 0, 1, \dots, N - 1$ ) harmonic vector of the  $z = 0, L$  cross-sectional profiles  $\mathbf{R}(z, p)|_{z=0,L}$ , as follow.

$$\mathbf{R}(z, p)|_{z=0,L} = \frac{\mathbf{T}(z, p) + \mathbf{M}(z, p)}{\mathbf{\Omega}(p)}|_{z=0,L}; \mathbf{\Omega}(p) = \exp(jp\varphi) - \mathbf{\Gamma}(p) \tag{13}$$

$$\mathbf{\Gamma}(p) = \left\{ \exp(jp\alpha) \left[ \frac{\cos(\varphi)}{2\cos(\alpha)} + \frac{\sin(\varphi)}{2\sin(\alpha)} \right] + \exp[jp(\pi - \alpha)] \left[ -\frac{\cos(\varphi)}{2\cos(\alpha)} + \frac{\sin(\varphi)}{2\sin(\alpha)} \right] \right\}$$

$$\mathbf{M}(z, p)|_{z=0,L} = \text{DFT}[e_{sx}(z) \cos(\varphi) + e_{sy}(z) \sin(\varphi)]|_{z=0,L}$$

where,  $\Omega(p)$  is the transfer function, and  $\mathbf{T}(z, p)|_{z=0,L} = \text{DFT}[t(z, i)]|_{z=0,L}$  are the  $p$ th ( $p = 0, 1, \dots, N - 1$ ) harmonic vectors of the output of the probe when it locates in the  $z = 0, L$  cross-sections.  $\mathbf{M}(z, p)|_{z=0,L}$  are the  $p$ th ( $p = 0, 1, \dots, N - 1$ ) harmonic vectors of the disturbances from the carriage straightness error motions when the probe locates in the  $z = 0, L$  cross-sections.

If stiffness of the measuring system is sufficient, the radial error motions of the measuring system in measured cross-section, denoted by  $e_x(z, i), e_y(z, i)$  ( $i = 0, 1, \dots, N - 1$ ), should be the linear combination of  $e_x(z, i)|_{z=0,L}, e_y(z, i)|_{z=0,L}$ . Here, their  $p$ th ( $p = 0, 1, \dots, N - 1$ ) harmonic vectors, denoted by  $\mathbf{E}_x(z, p) = \text{DFT}[e_x(z, i)]$  and  $\mathbf{E}_y(z, p) = \text{DFT}[e_y(z, i)]$ , can be achieved.

Now, taking DFT to Equation (12), we will obtain the  $p$ th ( $p = 0, 1, \dots, N - 1$ ) harmonic vector of the output of the probe when it locates in measured cross-section:

$$\mathbf{T}(z, p) = \text{DFT}[t(z, i)] = \exp(jp\varphi)\mathbf{R}(z, p) + \mathbf{E}_x(z, p) \cos(\varphi) + \mathbf{E}_y(z, p) \sin(\varphi) - \mathbf{M}(z, p). \quad (14)$$

Bringing  $\mathbf{E}_x(z, p), \mathbf{E}_y(z, p)$  and  $\mathbf{R}(z, p)|_{z=0,L}$  ( $p = 0, 1, \dots, N - 1$ ) into Equation (14), we can determine  $\mathbf{R}(z, p)$ , indicating the  $p$ th ( $p = 0, 1, \dots, N - 1$ ) harmonic vectors of the measured cross-sectional profile of the cylinder.

$$\mathbf{R}(z, p) = \left\{ \mathbf{T}(z, p) + \frac{(L - z)\mathbf{T}(0, p) + z\mathbf{T}(L, p)}{L\Omega(p)} \Gamma(p) \right\} \exp(-j p \varphi) + \Delta(z, p). \quad (15)$$

$$\Delta(z, p) = \left\{ \mathbf{M}(z, p) + \frac{(L - z)\mathbf{M}(0, p) + z\mathbf{M}(L, p)}{L\Omega(p)} \Gamma(p) \right\} \exp(-j p \varphi)$$

where,  $\mathbf{T}(z, p)$  is the  $p$ th ( $p = 0, 1, \dots, N - 1$ ) harmonic vector of the output of probe in measured cross-section.  $\mathbf{M}(z, p)$  is the  $p$ th ( $p = 0, 1, \dots, N - 1$ ) harmonic vector of the disturbances from the carriage straightness error motions in measured cross-section. A reconstructed cylindrical profile  $r(z, i)$  ( $i = 0, 1, \dots, N - 1$ ) can be achieved by taking IDFT to  $\mathbf{R}(z, p)$  ( $p = 0, 1, \dots, N - 1$ ), similarly to Equation (6). However, it is worth noting that Equations (13)–(15) are true if the transfer function  $\Omega(p) \neq 0$ , in addition, the disturbances from the carriage straightness error motions will impact the accuracy of the V-block scan EST.

### 3. Applicative and Comparative Analysis for In Situ Measurement of Cylindricity

Three error separation models for in situ measurement of cylindricity have been investigated as above. Theoretically, they all can be used to measure and reconstruct cylindrical profile. However, their validity and accuracy need to be verified by numerical verifications through experiments so as to further reveal characteristic features as well as main benefits and drawbacks.

#### 3.1. Experimental Verification and Analysis of Parallel Scan EST

An experimental system based on the parallel scan EST was setup as shown in Figure 6. The headstock and tailstock supported the cylinder. The carriage had three parallel sections separated by a distance  $d = 40$  mm. Five inductive displacement probes (TESA GTL 21 DC) were mounted onto the carriage, in which Probes 1–3 were mounted in the left section according to the three-point roundness EST, and Probes 4 and 5 were located in the middle and right sections. An angular encoder (RENISHAW RESM 52, 8192 lines, zero line, 5 subdivisions) was mounted on the rotor in the headstock to control the course of data acquisition so as to obtain uniform sampling on the circumference. The total number of sampling points per revolution  $N$  was 1024. In order to avoid the non-first harmonic suppression [3–5,8,14] and achieve the large error transfer function  $\Omega(p)$  ( $p = 0, 2, \dots, 15$ ), the angular positions of the Probes 1–3 were set respectively  $\alpha = 77.03^\circ, \beta = 86.84^\circ$  as shown in Figure 7. The rotational speed of the cylinder was 1 rpm.



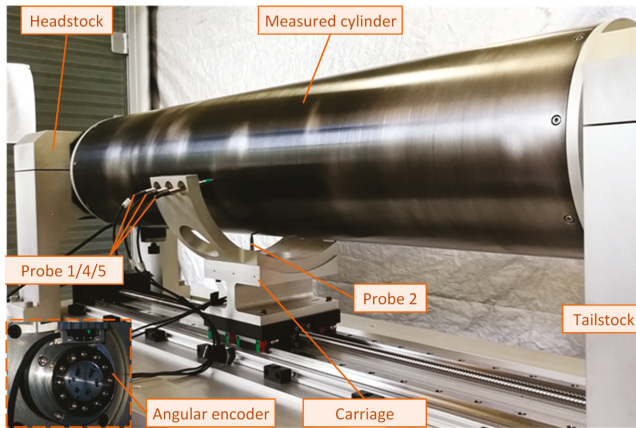


Figure 6. Photograph of the parallel scan EST experimental system.

It can be seen from Figure 3 and Equation (6) that the parallel scan EST is to simultaneously execute the three-point roundness and sequential three-point straightness ESTs [5,8] for the detection of the radius deviation, out-of-roundness, and LSC of each cross-section of the cylinder, respectively. A novelty of the EST is to accurately detect the spatial curved median line of the cylinder and to measure and reconstruct cylindrical profile consistent to the standard ISO 12180 [6]. The key to verifying the validity of the EST is to evaluate the accuracy of the measured curved median line profile of the cylinder. Currently, there is not a mature method to calibrate the curved median line profile, and it is difficult to investigate the error motions of the cylinder. Therefore, we will compare the repeatability of the curved median line profiles to prove the accuracy. In engineering practices, the high repeatability of measurement results will indicate high-precision of the measurements, and also high accuracy to some extent.

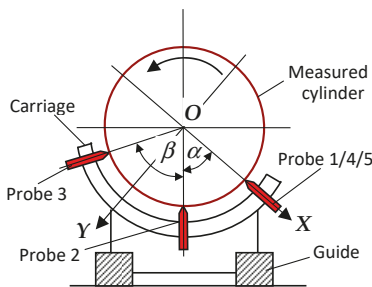


Figure 7. Angular positions of Probes 1–5.

Six measurement experiments were conducted on the experimental system for same cylinder shown in Figure 6, the number of the cross-section divided in the cylinder was  $M = 25$ . A set of software was developed to sequentially control the course of measurements and data collections of Probes 1–5. By calculating Equation (6) with iterative, the LSC vector  $2\mathbf{R}(z_j, N - 1)/2$  of the  $J$ th ( $J = 1, 2, \dots, M$ ) cross-sectional profile of the cylinder was extracted, with which the curved median line profile was fitted in each experiment, as shown in Figure 8. Although the profiles of the curved median line were quite consistent among in the six experiments, the variations of the extracted LSC vector of each cross-sectional profile were significant, being approximately within the range  $\pm 2.5 \pm 2.5 \mu\text{m}$ , which cannot be a convincing proof of the accuracy of the parallel scan EST for measuring cylindrical profile.

Theoretically, this EST can eliminate the radial and tilt error motions of the cylinder as well as the carriage straightness error motions. Actually, similarly to the sequential three-point straightness EST, the double integration calculations [15] as Equation (6) are executed for detecting the LSC vectors of each cross-sectional profile of the cylinder, which reduces the anti-interference capacity of the EST. Influences of various interferences on detecting the LSC vectors of each cross-section are amplified, which conduct the large variations in the measured median line profiles. In particular, when there are significant systematic interferences in the measuring system, such as inconsistencies in sensitivity and linearity, angular and positional errors of five probes, an obvious distortion of the measured median line profile will occur. Even more, the larger the total number of the cross-section in the cylinder is, the more serious the distortion is.

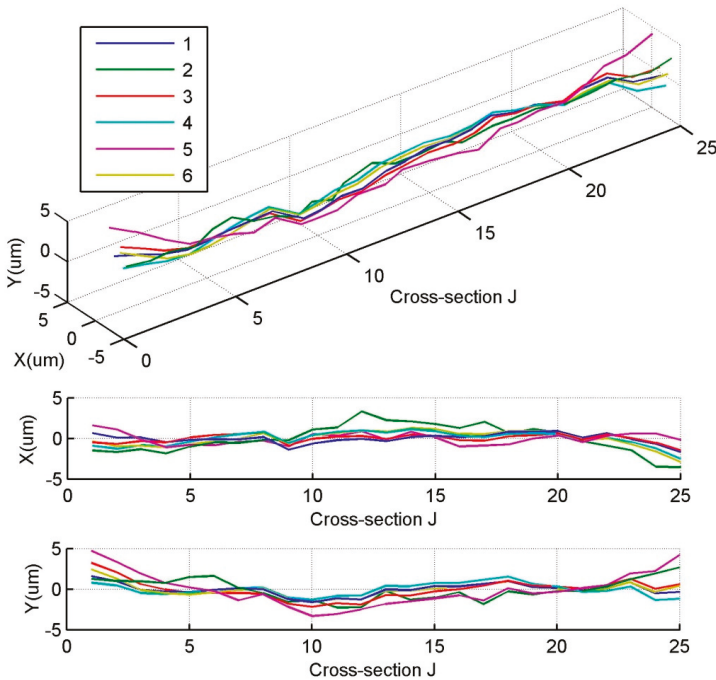


Figure 8. Median line profiles measured by the parallel five-point EST.

Comparatively, if the outputs of Probes 1–4 are used to measure and reconstruct the cylindrical profile, this EST will become an integrated EST of the three-point roundness and sequential two-point straightness. Here, the sequential two-point straightness EST [18] is executed for the detection of the LSC vector of each cross-sectional profile of the cylinder. Let us see Equation (6),  $\Theta_{cx}(J, 1)$ , which is the first harmonic vector of the additional X directional error motions induced by the tilt error motions of the cylinder around the Y-axis of each position, should be omitted, and the double integration calculations will be avoided. At this moment, the tilt error motions of the cylinder cannot be removed although the influences of various interferences on detecting the LSC vector of each cross-section are weakened. In engineering practices, the tilt error motions of the cylinder are much smaller, particularly in case if the cylinder is long and the rotation speed is low. Figure 9 shows the curved median line profiles measured by the outputs of Probes 1–4. Obviously, the measured profiles are in good consistency in six experiments, and the variation of the extracted LSC vector of each cross-sectional profile is within the range  $\pm 1 \pm 1\mu\text{m}$ , which shows a convincing proof of the accuracy of the parallel scan EST for measuring cylindrical profile.

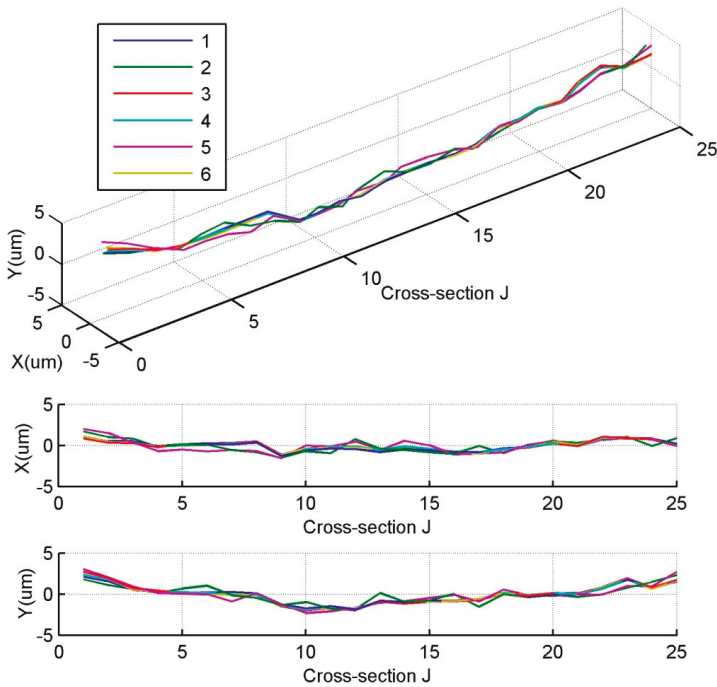


Figure 9. Median line profiles measured by the parallel four-point EST.

The accuracy of the extracted LSC vectors of sequential cross-sectional profiles by the four-point EST is higher than that extracted by the five-point EST if the tilt error motions of the cylinder are small and negligible. Figures 8 and 9 show that the profiles of the curved median line measured by two parallel ESTs differ slightly, and the conformity of the spatial median line measured by the parallel four-point EST is validated to be higher in the six experiments, which implies that the tilt error motions of the cylinder of the experimental system is small. Therefore, the parallel four-point EST is more suitable for in situ measurement of cylindricity of the large-scale roller, where the tilt error motions of the cylinder are indeed small in case that the cylinder is long and the rotation speed is low. On this basis, the cylindrical profile can be accurately reconstructed by Equation (7).

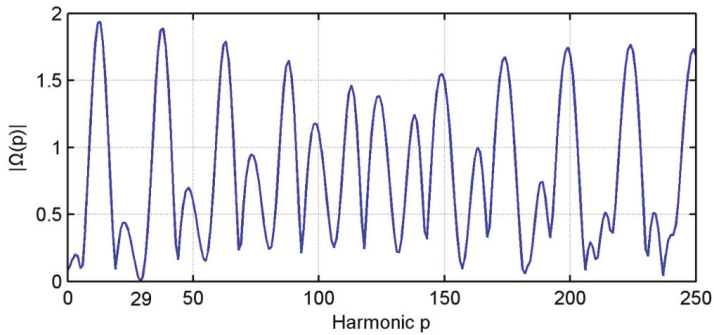
### 3.2. Applicative Analysis of Spiral Scan EST

Theoretically, the spiral scan EST for measuring the cylindrical profile must face two basic queries. One is harmonic suppression and the other is the multi-point straightness EST based on DFT. As stated above, Equation (9) would be true if the transfer function  $\Omega(p)$  ( $p = 0, 1, \dots, N \times M - 1$ ) was not “zero”. However, when  $p = M$ , the transfer function should be:

$$\Omega(M) = \sum_{n=1}^5 c_n \exp(j\varphi_n) = \sum_{n=1}^5 c_n \cos(\varphi_n) + j \sum_{n=1}^5 c_n \sin(\varphi_n) \equiv 0.$$

This means that the  $M$ th harmonic suppression happens in this EST, i.e., the  $M$ th harmonic component of the spiral line profile on the cylinder is unable to exactly be determined. In Reference [10], the partial transfer function  $\Omega(p)$  ( $p = 0, 1, \dots, 250$ ) was shown as Figure 10.

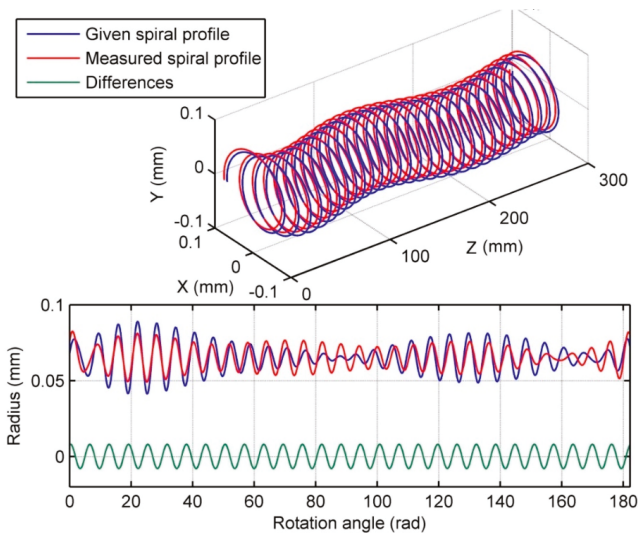
A numerical validation procedure has been developed in this study to help qualitative and quantitative evaluation of the cylindrical profile obtained by this EST. Whether the  $M$ th harmonic suppression impacts correct detection of spiral line profile on the cylinder should be verified first.



**Figure 10.** Transfer function  $\Omega(p)$  ( $p = 0, 1, \dots, 250$ ) of spiral scan EST. Here, angular position of five probes  $\varphi_n$  ( $n = 1, 2, \dots, 5$ ) is  $0^\circ, 395^\circ, 838.7^\circ, 1236.5^\circ, 1658.8^\circ$ , respectively, and pitch of spiral line is  $T = 10$ . Number of total rotating circles of spiral line is  $M = 29$  [10].

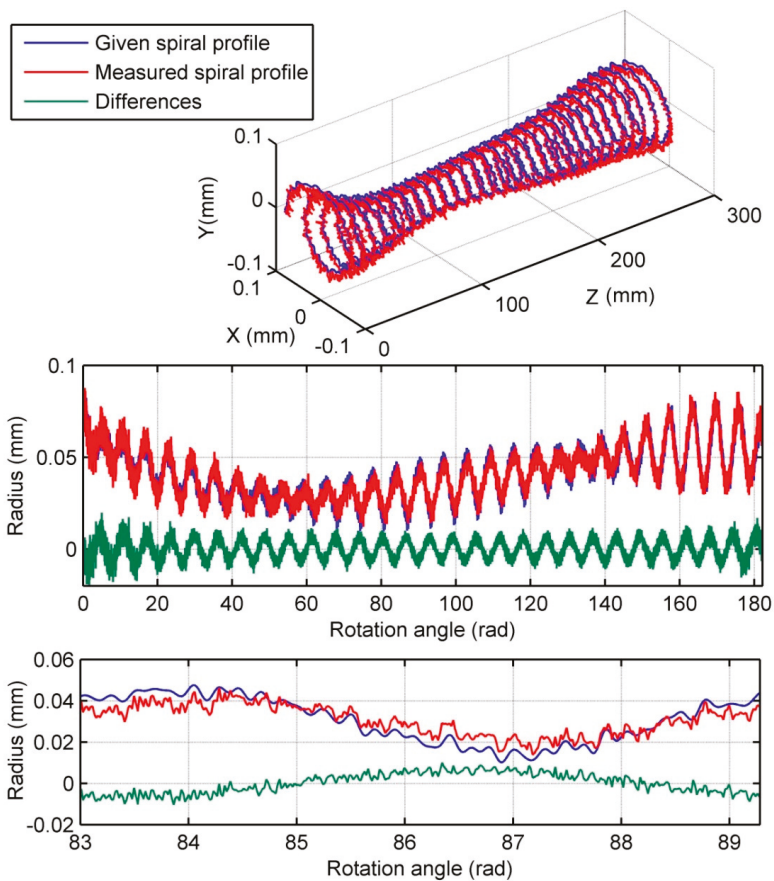
We created a set of given data to describe a virtual spiral line profile on the cylinder. A series of random data was used to simulate the radial error motions of two imaginary rolling bearings that supported the cylinder to rotate so that the different error motions in each revolution were simulated. A series of the other random data was used to simulate the carriage straightness error motions including yaw and pitch. The outputs of five probes were set according to Equation (8), and then, the procedures of reconstructing the spiral line on the cylinder were carried out according to Equations (9)–(11).

As shown in Figure 11, the given spiral line profile on the cylinder only with the median line deviations is similar to Figure 1b. The difference between the measured spiral line profile and given one presents periodicity, which is the 29th harmonic component of the spiral line profile on the cylinder. Such a difference causes the spiral line profile measured by this EST to slightly shift as a whole, yet the profile of the spiral line is not distorted. Therefore, the  $M$ th harmonic suppression of this EST will not impact the accuracy of measuring the spiral line profile.



**Figure 11.** Comparison of the measured spiral line profile with given one in the case of ideal measurement. Here, angular position of probes  $\varphi_n$  ( $n = 1, 2, \dots, 5$ ) is  $0^\circ, 395^\circ, 838^\circ, 1236^\circ, 1659^\circ$ , respectively, and pitch of spiral line is  $T = 10$  mm. Number of total rotating circles of spiral line is  $M = 29$ .

The measured spiral line profile as shown in Figure 11 is based on a kind of “ideal measurement”, i.e., based on an assumption that the straightness profile of the spiral line is closed, because the spiral scan EST is based on DFT in theory. In the case of “ideal measurement” during numerical validations, the sampling of five probes were carried out under the condition that the starting point and ending point of given spiral line profile were connected. However, it is impossible in actual measurements. If the angular distance between Probes 1 and 5 is  $\varphi = (\varphi_5 - \varphi_1)$  and the total rotation angle of the cylinder is  $2\pi M$ , Probes 2–4 cannot normally collect data in the angular range from  $(2\pi M - \varphi)$  to  $2\pi M$  during actual measurement. Now, how to estimate the output of Probes 2–4 in this angular range would affect the precision of the measured spiral line profile on the cylinder. Here, the outputs of Probes 2–4 in the angular range from  $(2\pi M - \varphi)$  to  $2\pi M$  are assigned with the data collected by Probe 1 in the angular range from 0 to  $\varphi$  so as to best agree with the requirement for closed profile. Figure 12 demonstrates the verified results of this EST in the case of “actual measurement”.



**Figure 12.** Comparison of the measured spiral line profile with given one in the case of actual measurement. Here, angular position of probes  $\varphi_n$  ( $n = 1, 2, \dots, 5$ ) is  $0^\circ, 395^\circ, 838^\circ, 1236^\circ, 1659^\circ$ , respectively, and pitch of spiral line is  $T = 10$  mm. Number of total rotating circles of spiral line is  $M = 29$ .

From Figure 12, it may be found that the given spiral line profile on the cylinder with the median line deviation, radius deviations and roundness errors, and the measured spiral line profile is

almost the same as the given one except for its slight translation. Their differences also present the periodicity of the 29th harmonic. Because the data substitutions of Probes 2–4 introduce some errors, the difference between the measured spiral line profile and given one has the high-frequency components. In addition, the measured spiral line profiles at both ends of the cylinder are deteriorated significantly. However, in engineering practice, we pay attention to low frequency parts of the measured profile, high frequency error in the measured profile can be eliminated by filtering. Therefore, the shortened effective measurement length of the spiral line profile is a weakness of this EST. It is worth noting that the above numerical verification results were obtained under the condition that the radial error motions of two imaginary rolling bearings and carriage error motions were only 20% of the given spiral line profile on the cylinder. If this ratio goes up, the confidence level of the spiral line profile measured by this EST will be greatly decreased.

In Reference [10], a prototype instrument was developed based on the spiral scan EST. Comparing the results of the instrument with that of a commercial roundness measuring machine indicated the feasibility of measuring out-of-roundness, but the instrument seemed to be not able to identify the curved median line profile of the cylinder. In the opinion of the authors of Reference [10], the  $M$ th harmonic suppression issue in this EST was worth in-depth examined. Whereas, they did not touch the issue of the multi-point straightness EST based on DFT.

### 3.3. Applicative Analysis of V-Block Scan EST

Frist, we should evaluate the validity and accuracy of the V-block scan EST based on the standard ISO 12180 [6]. From Equation (15), we seem to obtain the measured cross-sectional profiles  $\mathbf{R}(z, p)$  ( $p = 0, 1, 2, \dots, N - 1$ ). However, from Equation (13), it can be verified that the transfer function  $\Omega(p)|_{p=1, N-1} \equiv 0$ , as shown below.

$$\Omega(1) = \exp(j\varphi) - \left\{ \exp(j\alpha) \left[ \frac{\cos(\varphi)}{2\cos(\alpha)} + \frac{\sin(\varphi)}{2\sin(\alpha)} \right] - \exp[j(\pi - \alpha)] \left[ \frac{\cos(\varphi)}{2\cos(\alpha)} - \frac{\sin(\varphi)}{2\sin(\alpha)} \right] \right\} \\ = \exp(j\varphi) - \{ \cos(\varphi) + j\sin(\varphi) \} = 0$$

Therefore, Equations (13) and (15) are not true in this case, the first harmonic suppression occurs in the V-block scan EST. It can be seen that  $\mathbf{R}(0, p)|_{p=1, N-1}$  and  $\mathbf{R}(L, p)|_{p=1, N-1}$ , which indicate the first and  $(N - 1)$ th harmonic vectors of the  $z = 0, L$  cross-sectional profiles, cannot be accurately extracted. In other word, the LSCs vectors of the  $z = 0, L$  cross-sectional profiles are unknowable.

On the basis of routine,  $\mathbf{R}(0, p)|_{p=1, N-1}$  and  $\mathbf{R}(L, p)|_{p=1, N-1}$  will be set as “zero”. Then,  $\mathbf{E}_x(z, p)|_{p=1, N-1}$  and  $\mathbf{E}_y(z, p)|_{p=1, N-1}$  could be also regarded as “zero” because the radial and tilt error motions of the measuring system are controlled by the  $z = 0, L$  cross-sectional profiles. In view of the properties of DFT, there are  $\mathbf{M}(z, 0) \neq 0$  and  $\mathbf{M}(z, p)|_{p \neq 0} = 0$  in Equation (13). From Equation (15),  $\mathbf{R}(z, p)|_{p=1, N-1}$ , which indicate the first and  $(N - 1)$ th harmonic vectors of the measured cross-sectional profiles of the cylinder, are:

$$\mathbf{R}(z, 1) = \mathbf{T}(z, 1) \exp(-j\varphi); \mathbf{R}(z, N - 1) = \text{conjugation}[\mathbf{R}(z, 1)]. \tag{16}$$

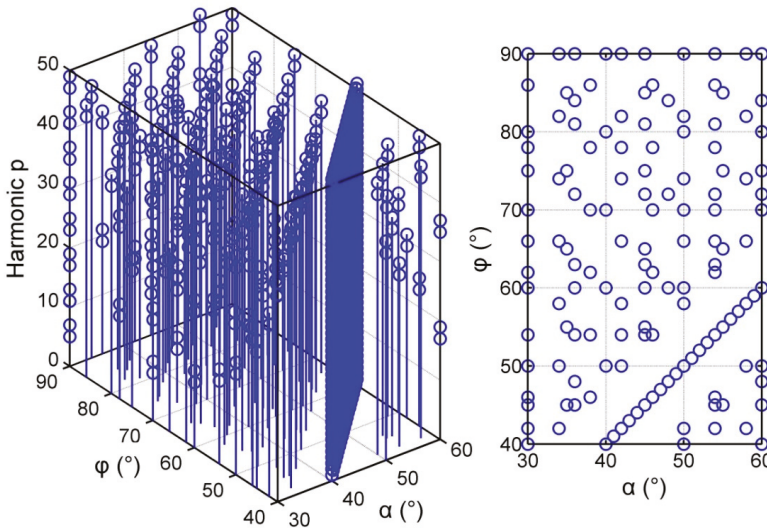
However,  $\mathbf{R}(0, p)|_{p=1, N-1}$  and  $\mathbf{R}(L, p)|_{p=1, N-1}$  are actually existent and are typical of the LSC vectors of the  $z = 0, L$  cross-sectional profiles in contact with two V-blocks. They induce  $\mathbf{E}_x(z, p)|_{p=1, N-1}$  and  $\mathbf{E}_y(z, p)|_{p=1, N-1}$ , indicating the first harmonic vectors of the radial error motions of the measuring system. Substituting  $\mathbf{E}_x(z, 1)$  and  $\mathbf{E}_y(z, 1)$  into Equation (14),  $\mathbf{T}(z, p)|_{p=1, N-1}$ , which indicate the first and  $(N - 1)$ th harmonic vectors of the output of the probe, can be rewritten as follow:

$$\mathbf{T}(z, 1) = \left[ \mathbf{R}(z, 1) - \frac{(L - z)\mathbf{R}(0, 1) + z\mathbf{R}(L, 1)}{L} \right] \exp(j\varphi); \mathbf{T}(z, N - 1) = \text{conjugation}[\mathbf{T}(z, 1)].$$

Equation above shows that  $\mathbf{T}(z, 1)$  comprises  $\mathbf{R}(z, 1)$  and linear combination of  $\mathbf{R}(0, 1)$  and  $\mathbf{R}(L, 1)$ . However, from Equation (16), the LSC vectors in the measured cross-sections  $2\mathbf{R}(z, 1)/N$  are just based

on the line connecting the actual LSCs of the  $z = 0, L$  cross-section, i.e.,  $2\mathbf{R}(0,1)/N$  and  $2\mathbf{R}(L,1)/N$ . Hence, the spatial curved median line fitted by  $2\mathbf{R}(z,1)/N$  or  $2\mathbf{R}(z,N-1)/N$  has a little movement and inclination, based on which the cylindrical profile can be reconstructed accurately even if the suppression of the first harmonic exists.

The non-first harmonic suppression [3–5,8,14], which denotes the transfer function  $\Omega(p)|_{p>1} = 0$ , will impact the measurement accuracy of the out-of-roundness of each cross-sectional profile of the cylinder. Whether suppression of the non-first harmonic exists in this EST will depend on the values of angular parameters of  $\varphi$  and  $\alpha$  as shown in Figure 5. Digital simulations were performed to verify how to configure parameters of  $\varphi$  and  $\alpha$  to avoid the non-first harmonic suppression. Figure 13 demonstrates the relationships between the angular parameters and the  $p$ th harmonic suppression. It may be found that if  $\alpha = 30^\circ$  and  $\varphi = 90^\circ$ , the  $p$ th ( $p = 5, 7, 11, 13, 17, \dots$ ) harmonic suppression occurs. Especially, when  $\alpha = \varphi$ , full harmonic suppression will occur. Let total number of scanning points per revolution be  $N$  and  $\delta = 2\pi/N$ ,  $\varphi = k\delta$  and  $\alpha = u\delta$ , the non-first harmonic suppression of this EST will be avoided provided that  $k \neq u$ , and the greatest common factor of  $(k - u)$  and  $(N/2 - 2u)$  as well as  $N$  is 1, e.g.,  $\alpha = 45^\circ$  and  $\varphi = 68^\circ$ . Therefore, this measuring system needs to carefully adjust the angular position of the probe so as to avoid the non-first harmonic suppression and get the transfer function  $\Omega(p)(p = 0, 2, \dots, 15)$  as large as possible.



**Figure 13.** Non-first harmonic suppressions and the values of angular parameters of  $\varphi$  and  $\alpha$ . If transfer function  $\Omega(p) = 0$  at given  $\varphi$  and  $\alpha$ , the  $p$ th harmonic suppression occurs. Here,  $\alpha = 30^\circ \sim 60^\circ$  and  $\varphi = 40^\circ \sim 90^\circ$ , harmonic order  $p = 2\text{--}50$ .

The 0th harmonic component  $\mathbf{R}(z, 0)$  denotes the radius deviations of the sequential cross-sectional profiles. From Equation (15), measurement accuracy of the radius deviations depends on:

$$\Delta(z, 0) = \left\{ \mathbf{M}(z, 0) + \frac{(L - z)\mathbf{M}(0, 0) + z\mathbf{M}(L, 0)}{L[\sin(\alpha) - \sin(\varphi)]} \sin(\varphi) \right\} \quad (17)$$

It is known that  $\mathbf{M}(0, 0)$  and  $\mathbf{M}(L, 0)$  come from  $e_{sx}(z)|_{z=0,L}$  and  $e_{sy}(z)|_{z=0,L}$ , i.e., the carriage straightness error motions in the  $z = 0, L$  cross-sections, in other word, the guideway’s slope in the  $XOZ$  and  $YOZ$  plane. Due to the existence of guideway’s errors at the  $z = 0, L$  cross-sections, the reconstructed cylindrical profile would produce a conical form.  $\mathbf{M}(z, 0) = \text{DFT}[e_{sx}(z) \cos(\varphi) + e_{sy}(z) \sin(\varphi)]|_{p=0}$ , indicating the 0th harmonic component of the disturbances

from the carriage straightness error motions in the measured cross-section, directly results in the measurement errors in the radius of the reconstructed cylindrical profile, as shown in Figure 1a.

In References [12,13,17], computer simulations of this EST for measurement of cylindrical profile were conducted to assess the non-first harmonic suppression. The results indicate that the 11th harmonic component exists when  $\alpha = 60^\circ$  and  $\varphi = 90^\circ$  were selected. From Figure 13, the  $p$ th ( $p = 11, 13, 23, 25, \dots$ ) harmonic suppression occurs in this case. It is a weakness that the straightness deviations of the guideway affect directly measurement accuracy of the cylindrical profile. It yields the necessity to compensate for the carriage straightness deviations in order to improve the measurement accuracy. In the authors' opinion, however, the compensation method should not change the fact that guideway's slopes in XOZ and YOZ planes would make the reconstructed cylindrical profile be a conical form. A measuring system allowing cylindricity measurements by this EST was constructed [12,13,17]. Some potential sources of systematic errors, such as, unequal angles of V-blocks, calibration error of the probe, deflection of the probe axis from its nominal orientation were analyzed. A series of experiments were conducted involving the comparison of measurement results of a group of cylindrical elements. The cylindricity deviation of each sample element was determined with two methods: V-block scan EST and high accurate radial method [19]. The conducted tests showed that the difference lies within the interval  $\pm 19\%$  (for a probability level  $p = 0.95$ ).

#### 4. Conclusions

Three five-point ESTs used for in situ cylindricity measurement were analyzed to determine their characteristic features as well as main benefits and drawbacks.

The following are conclusions drawn about the parallel scan EST:

1. Measurement of cylindrical profile was performed with five displacement probes. Outputs from five probes were used to measure radius deviations and roundness deviations of sequential cross-sectional profiles, as well as the LSC vectors of sequential cross-sectional profile, with which spatial curved median line profile of the cylinder was determined. Further, cylindrical profile could be accurately reconstructed by assembling each obtained cross-sectional profile onto the spatial curved median line.
2. The completeness of the measurement model in theory is promising. Accurate detection of the spatial curved median line of the cylinder highlights its novelty, and the error motions of the cylinder and carriage in measuring system are fully eliminated.
3. Detecting full harmonic components of sequential cross-sectional profiles is dependent on how the three probes in one cross-section are located relative to each other. Therefore, to ensure good anti-interference capability, angular arrangement of the three probes should be carefully selected so as to obtain large transfer function  $\Omega(p)$  ( $p = 0, 2, \dots, 15$ ).
4. The probes should be strictly calibrated and carefully chosen. Especially, inconsistencies of sensitivity and linearity among five probes should be overcome as much as possible, which is a main factor impacting on the precision of spatial curved median line of the cylinder.
5. Accuracy of the method has not been verified by calibration.
6. The method is suitable for on-machine measurement of cylindrical profile. Where, the carriage can be mounted onto the tool carriage. Measurement cost is relatively low.
7. The parallel four-point scan EST is more suitable for in situ cylindricity measurement of large-scale roller, where the tilt error motions of the cylinder are negligible because the roller is long and the rotation speed is low.

The conclusion about spiral scan EST can be summarized as follows:

1. Measurement of cylindrical profile is performed with five displacement probes, which are arranged in a spiral scanning-path. Cylindrical profile measurement is transformed to one-dimensional straightness measurement by the multi-point straightness EST based on DFT.



2. The multi-point straightness EST based on DFT is built in an assumption that the straightness profile is end-to-end. The actual spiral profile on the cylinder is impossible to achieve. Therefore, direct adoption of straightness EST based on DFT to measure and reconstruct spiral profile on the cylinder would weaken the theoretical completeness.
3. The way to fill a segment of data of Probes 2–4 at the tail end of measurement becomes a crucial factor to determine measurement accuracy. In light of the authors' experience, if straightness errors of actual spiral profile are significant in the outputs of the probes, it may be a good solution that a segment of data of Probes 2–4 at the tail end of measurement are assigned by the data collected by Probe 1 at the head end of measurement.
4. The  $M$ th harmonic suppression of the EST will not impact measurement accuracy of the spiral line profile. It only causes a slight shift of the measured spiral line profile.
5. Because a segment of data of Probes 2–4 at the tail end of measurement is uncertain, the accuracy of the measured spiral line profiles at both ends of the cylinder deteriorated significantly. Therefore, the shortened effective measurement length of the spiral line profile is a weakness.
6. Accuracy of the method has not been verified by calibration.
7. If this EST is employed in situ measurement of cylindricity, motion accuracy of the guideway and spindle should be high.

The conclusions about V-block scan method can be summarized as follows:

1. Measurement of cylindrical profile was performed with the measuring system made up of two V-blocks and a displacement probe. During measuring, two cross-sectional profiles that contact two V-blocks control the radial and tilt error motions of the measuring system, yielding repeatable error motions of the measuring system in each revolution. This makes the measured cylindrical profile immune to the error motions of the cylinder.
2. Error motions of the cylinder are removed, whereas, the carriage straightness error motions will directly impact the measurement accuracy of cylindrical profile, which would weaken the theoretical completeness.
3. Detecting full harmonic components of sequential cross-sectional profiles is dependent on selection angular values of the probe and V-block. Therefore, to ensure good anti-interference capability, angular value of the probe should be carefully selected so as to obtain large transfer function  $\Omega(p)(p = 0, 2, \dots, 15)$ .
4. The measurement uncertainty was assessed through statistical testing. The expanded uncertainty reached 19% compared to the results obtained with the highly accurate radial method. This EST was verified in practice using a model test stand, not on a large roller.
5. It is a portable system, and is suitable for in situ measurement.
6. During measurement, it is crucial to avoid the measuring system rotating around and moving in the  $Z$ -axis of the cylinder. How to decrease the impact from straightness deviations of the guideway is crucial to measurement accuracy.
7. Compensation for the straightness deviations of the guideway was proposed to improve the measurement accuracy. In the authors' opinion, however, it may be a good solution to diametrically arrange the second probe in the axial direction of the original probe.

Summing up, the parallel four-point scan technique is recommended for on-machine measurement cylindricity of large rollers, where spindle tilt error motions are negligible. Diametrical arrangement of the second probe in the axial direction of original probe will make the V-block scan EST be the most prospective technique for short roller. However, it is worth noting that, in actual multipoint EST, the main factor of impact on measurement accuracy is the inconsistency of characteristics between multiple displacement probes; therefore, probes should be carefully chosen.

**Author Contributions:** W.L. and K.F. conceived and investigated; P.H. designed the experimental system and performed the experiments; W.L. simulated, analyzed and wrote; K.F. modified.

**Funding:** This research was funded by [National Natural Science Foundation of China] grant number [51475132 and 51475133] and [Natural Science Foundation of Anhui Province of China] grant number [1508085ME81].

**Acknowledgments:** This study has been supported by the National Natural Science Foundation of China (Grant Nos.: 51475132, 51475133) and the Natural Science Foundation of Anhui Province of China (Grant No.: 1508085ME81). The authors would like to thank the other members of the research team for their contributions to this study.

**Conflicts of Interest:** The authors declare no conflicts of interest.

## Abbreviations

EST	Error separation technique
LSC	Least squares center
DET	Discrete Fourier transform
IDET	Inverse discrete Fourier transform

## References

1. Stępień, K. In situ measurement of cylindricity—Problems and solutions. *Precis. Eng.* **2014**, *38*, 697–701. [CrossRef]
2. Lee, J.C.; Yuki, S.Y.; Gao, W.; Oh, J.; Park, C.H. Precision evaluation of surface form error of a large-scale roll workpiece on a drum roll lathe. *Precis. Eng.* **2014**, *38*, 839–848. [CrossRef]
3. Ozono, S. On a new method of roundness measurement on the three-point method. *Proc. ICPE 1974, 1974*, 457–462.
4. Shi, S.Y.; Lin, J.; Wang, X.F.; Zhao, M. A hybrid three-probe method for measuring the roundness error and the spindle error. *Precis. Eng.* **2016**, *45*, 403–413. [CrossRef]
5. Li, S.Y.; Dai, Y.F.; Xie, X.H. *On-Side Measurement and Error Compensation Technology for Precision and Ultra-Precision Machining*; University of Defense Science and Technology Press: Changsha, China, 2007; pp. 95–279, ISBN 978-7-81099. (In Chinese)
6. ISO 12180-1 2011, Annex B: Geometrical Product Specifications (GPS)—Cylindricity—Part 1: Vocabulary and Parameters of Cylindrical Form—Consideration in the Assessment of Deviations from Cylindrical Form (International Organization for Standardization). Available online: <https://www.iso.org/obp/ui/#iso:std:iso:12180:-1:ed-1:v1:en> (accessed on 16 October 2018).
7. Zhang, L.; Zhao, Y.; Ba, L. Research and practice on the on-line measurement of cylindricity error in a grinding machine. *Key Eng. Mater.* **2008**, *359–360*, 176–180.
8. Liu, W.W.; Fan, K.C.; Hu, P.H.; Hu, Y. A parallel error separation method for the on-line measurement and reconstruction of cylindrical profiles. *Precis. Eng.* **2018**, *51*, 1–9. [CrossRef]
9. Muralikrishnan, B.; Venkatachalam, B.; Raja, J.; Malburg, M. A note on the three-point method for roundness measurement. *Precis. Eng.* **2005**, *29*, 257–260. [CrossRef]
10. Endo, K.; Gao, W.; Yiyono, S. A new multi-probe arrangement for surface profile measurement of cylinders. *JSME Int. J. Ser. C* **2003**, *46*, 1531–1537. [CrossRef]
11. Nyberg, T.R. *Dynamic Macro Topography of Large Slowly Rotating Cylinders (Acta Polytechnica Scandinavica, Mechanical Engineering Series)*; Finnish Academies Technology: Helsinki, Finland, 1993; Volume 108.
12. Adamczak, S.; Janecki, D.; Stępień, K. Cylindricity measurement by the V-block method—Theoretical and practical problems. *Measurement* **2011**, *44*, 164–173. [CrossRef]
13. Stępień, K.; Janecki, D.; Adamczak, S. Investigating the influence of selected factors on results of V-block cylindricity measurements. *Measurement* **2011**, *44*, 767–777. [CrossRef]
14. Liu, W.W.; Zeng, H.; Liu, S.L.; Wang, H.T.; Chen, W.Y. Four-point error separation technique for cylindricity. *Meas. Sci. Technol.* **2018**, *29*, 075007. [CrossRef]
15. Gao, W.; Yokoyama, J.; Kojima, H.; Kiyono, S. Precision measurement of cylinder straightness using a scanning multi-system. *Precis. Eng.* **2002**, *26*, 279–288. [CrossRef]
16. Li, C.J.; Li, S.Y.; Yu, J.M. High resolution error separation technique for in-situ straightness measurement of machine tools and workpiece. *Mechatronics* **1996**, *6*, 337–347. [CrossRef]
17. Adamczak, S.; Janecki, D.; Stępień, K. Qualitative and quantitative evaluation of the accuracy of the V-block method of cylindricity measurements. *Precis. Eng.* **2010**, *34*, 619–626. [CrossRef]

18. Zhang, L.; Zhang, Y. Accuracy analysis of time domain two-point and three-point EST linearity. *Chin. J. Sci. Instrum.* **1998**, *19*, 106–108.
19. Dawson, D.J.W. Cylindricity and its measurement. *Int. J. Mach. Tools Manuf.* **1992**, *32*, 247–253. [[CrossRef](#)]



© 2018 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).

MDPI  
St. Alban-Anlage 66  
4052 Basel  
Switzerland  
Tel. +41 61 683 77 34  
Fax +41 61 302 89 18  
[www.mdpi.com](http://www.mdpi.com)

*Applied Sciences* Editorial Office  
E-mail: [appls@mdpi.com](mailto:appls@mdpi.com)  
[www.mdpi.com/journal/appls](http://www.mdpi.com/journal/appls)





MDPI  
St. Alban-Anlage 66  
4052 Basel  
Switzerland

Tel: +41 61 683 77 34  
Fax: +41 61 302 89 18

[www.mdpi.com](http://www.mdpi.com)



ISBN 978-3-03921-713-7