PROCEEDINGS OF SPIE

SPIE—The International Society for Optical Engineering

Advanced Photonic Sensors and Applications

Robert A. Lieberman
Anand K. Asundi
Hiroshi Asanuma
Chairs/Editors

30 November–3 December 1999
Singapore

Sponsored by
SPIE—The International Society for Optical Engineering
Nanyang Technological University, Singapore

Cosponsored by
SPIE Singapore Chapter
Institute of Physics, Singapore
United States Air Force Asian Office of Aerospace Research and Development (AFOSR/AOARD)
United States Army Research Office-Far East (ARO-FE)

DISTRIBUTION STATEMENT A
Approved for Public Release
Distribution Unlimited

20000526 116

SPIE P PROCEEDINGS SERIES

Volume 3897

DTIC QUALITY INSPECTED 2

PROCEEDINGS OF SPIE
SPIE—The International Society for Optical Engineering

Advanced Photonic Sensors and Applications

Robert A. Lieberman
Anand K. Asundi
Hiroshi Asanuma
Chairs/Editors

30 November–3 December 1999
Singapore

Sponsored by
SPIE—The International Society for Optical Engineering
Nanyang Technological University, Singapore

Cosponsored by
SPIE Singapore Chapter
Institute of Physics, Singapore
United States Air Force Asian Office of Aerospace Research and Development (AFOSR/AOARD)
United States Army Research Office-Far East (ARO-FE)

Cooperating Organizations
National University of Singapore
SPIE Hong Kong Chapter
SPIE India Chapter
SPIE Japan Chapter
SPIE Korea Chapter
SPIE Taiwan Chapter

Volume 3897

Published by
SPIE—The International Society for Optical Engineering

SPIE is an international technical society dedicated to advancing engineering and scientific applications of optical , photonic, imaging, electronic, and optoelectronic technologies.

The papers appearing in this book comprise the proceedings of the meeting mentioned on the cover and title page. They reflect the authors' opinions and are published as presented and without change, in the interests of timely dissemination. Their inclusion in this publication does not necessarily constitute endorsement by the editors or by SPIE.

Please use the following format to cite material from this book:
Author(s), "Title of paper," in *Advanced Photonic Sensors and Applications*, Robert A. Lieberman, Anand K. Asundi, Hiroshi Asanuma, Editors, Proceedings of SPIE Vol. 3897, page numbers (1999).

ISSN 0277-786X
ISBN 0-8194-3499-X

Published by
SPIE—The International Society for Optical Engineering
P.O. Box 10, Bellingham, Washington 98227-0010 USA
Telephone 360/676-3290 (Pacific Time) • Fax 360/647-1445

Copyright ©1999, The Society of Photo-Optical Instrumentation Engineers.

Copying of material in this book for internal or personal use, or for the internal or personal use of specific clients, beyond the fair use provisions granted by the U.S. Copyright Law is authorized by SPIE subject to payment of copying fees. The Transactional Reporting Service base fee for this volume is $10.00 per article (or portion thereof), which should be paid directly to the Copyright Clearance Center (CCC), 222 Rosewood Drive, Danvers, MA 01923. Payment may also be made electronically through CCC Online at http://www.directory.net/copyright/. Other copying for republication, resale, advertising or promotion, or any form of systematic or multiple reproduction of any material in this book is prohibited except with permission in writing from the publisher. The CCC fee code is 0277-786X/99/$10.00.

Printed in the United States of America.

Contents

xi *Conference Committee*

PLENARY SESSION I

2 **Future technology and business opportunities in photonics: a view from the Optoelectronics Industry Development Association (OIDA)** [3897-201]
A. A. Bergh, Optoelectronics Industry Development Association (USA)

12 **Oxide-confined vertical-cavity surface-emitting lasers, quantum dots, and the Purcell effect: can scaling the mode size improve laser performance?** [3897-202]
D. G. Deppe, D. L. Huffaker, H. Huang, L. A. Graham, Univ. of Texas/Austin (USA)

PLENARY SESSION II

24 **Optical nonlinearities in semiconductors** [3897-203]
S. W. Koch, T. Meier, Philipps-Univ./Marburg (Germany)

36 **Advances in missile technology: impact of photonics** [3897-204]
A. S. Pillai, Defence Research and Development Organisation (India)

SESSION 1 PHOTONIC SENSORS MATERIALS

46 **Photochromic materials for holographic data storage (Invited Paper)** [3897-01]
R. A. Lessard, COPL/Univ. Laval (Canada); C. Lafond, COPL/Univ. Laval (Canada) and Univ. Blaise Pascal (France); F. Ghailane, COPL/Univ. Laval (Canada); M. Bolte, Univ. Blaise Pascal (France); A. Tork, Univ. Laval (Canada); I. Petkov, Univ. of Sofia (Bulgaria)

56 **Real-time diffraction efficiency in lithium niobate at 532 nm** [3897-02]
P. Jayanth, S. Y. Luen, Ngee Ann Polytechnic (Singapore)

63 **Metal-doped photosensitive materials: correlation between photoreactivity and holography** [3897-03]
M. Bolte, C. Pizzocaro, Univ. Blaise Pascal (France); R. A. Lessard, COPL/Univ. Laval (Canada)

SESSION 2 PHOTONIC SENSORS DESIGN AND FABRICATION

72 **Very selective volume holograms for spatial and spectral filtering** [3897-05]
N. O. Reinhand, Institute of Mechanical Engineering Problems (Russia); I. V. Semenova, A.F. Ioffe Physical-Technical Institute (Russia); A. Popov, LUMEX Ltd. (Russia)

79 **Fabrication and temperature dependence of As_2S_3 fiber long-period grating (LPG)** [3897-06]
M. Li, Y. Liao, R. Chen, Tsinghua Univ. (China); K. Yang, Beijing Glass Research Institute (China)

87 **Novel temperature compensation techniques for fiber Bragg gratings-based magnetostrictive sensors** [3897-07]
B. Yi, B. C. B. Chu, M. L. Mok, Z. Chen, K. S. Chiang, City Univ. of Hong Kong

94 **Bent long-period fiber gratings for sensor applications** [3897-08]
Z. Chen, K. S. Chiang, M. N. Ng, Y. M. Chan, City Univ. of Hong Kong; H. Ke, Tsinghua Univ. (China)

SESSION 3 BIOMEDICAL AND BIOCHEMICAL SENSORS I

106 **Biochip technology: a triad of micro-electro-mechanical (MEM), biochemical, and photonic technologies** [3897-09]
A. E. Chiou, National Dong Hwa Univ. (Taiwan)

115 **Precision blood-leak detector with high long-time stability** [3897-10]
C. Georgiadis, Univ.-Gesamthochschule Siegen (Germany); W. Kleuver, aicoss GmbH (Germany)

124 **Autofluorescence spectrum of human lung tissue by Monte Carlo modeling** [3897-11]
W. Zheng, S. M. Krishnan, Z. Huang, T.-C. Chia, Nanyang Technological Univ. (Singapore); G. Chiang, Singapore General Hospital; S. Lee, Nanyang Technological Univ. (Singapore)

131 **Blood cell counting and classification by nonflowing laser light scattering method** [3897-12]
Y. Yang, Nanyang Technological Univ. (Singapore); Z. Zhang, X. Yang, D. Jiang, Xi'an Jiaotong Univ. (China); J. H. Yeo, Nanyang Technological Univ. (Singapore)

139 **Fiber optic sensor for methane hazards** [3897-13]
V. Kumar, D. Chandra, Indian School of Mines

150 **Coating of polystyrene thin film on glass for protein immobilization in optical biosensor applications** [3897-14]
T. T. M. Tan, Y. Y. Gan, L. H. Gan, T. Yong, B. Zhou, Y. L. Lam, Y. Zhou, Nanyang Technological Univ. (Singapore)

158 **Laser-based instrumentation for medical diagnoses at visible wavelengths** [3897-15]
M. D. Waterworth, B. Tarte, A. J. Joblin, T. van Doorn, Queensland Univ. of Technology (Australia)

173 **NO_2 detection with a fiber optic evanescent wave sensor** [3897-16]
M. S. John, J. Thomas, K. P. Unnikrishnan, P. Radhakrishnan, V. P. N. Nampoori, C. P. G. Vallabhan, Cochin Univ. of Science and Technology (India)

SESSION 4 BIOMEDICAL AND BIOCHEMICAL SENSORS II

180 **Theoretical and experimental study on the optimum working wavelength for a fiber optic sensor based on the SPR used in monitoring the water quality** [3897-17]
D. Liu, Y. Huang, D. Huang, Huazhong Univ. of Science and Technology (China)

187 **Fiber optic Raman spectroscopy: an application for the in-situ measurement of water content in methanol** [3897-19]
A. K. Asundi, V. R. Veeredhi, Nanyang Technological Univ. (Singapore)

193 **Hyperspectral image sensor for weed-selective spraying** [3897-21]
F. Feyaerts, P. Pollet, L. J. Van Gool, P. Wambacq, Katholieke Univ. Leuven (Belgium)

204 **Portable system approach of monitoring plant nutrient deficiency using fiber optic spectrophotometry** [3897-22]
A. K. Asundi, J. W. Chen, D. M. He, Nanyang Technological Univ. (Singapore); O. W. Liew, Singapore Polytechnic

214 **Design of the integrated sensor system for two-dimension position based on CCD** [3897-23]
X. Liu, Y. Gu, Zhejiang Univ. (China)

SESSION 5 MICRO-OPTICAL METROLOGY

224 **Strain measurement in micrometrology (Invited Paper)** [3897-24]
D. Vogel, Fraunhofer Institut Zuverlassigkeit und Mikrointegration (Germany); R. Kuehnert, K&T Measuring Systems GmbH (Germany); B. Michel, Fraunhofer Institut Zuverlassigkeit und Mikrointegration (Germany)

239 **High-resolution stress and temperature measurements in semiconductor devices using micro-Raman spectroscopy (Invited Paper)** [3897-26]
I. De Wolf, J. Chen, M. Rasras, W. M. van Spengen, V. Simons, IMEC (Belgium)

253 **Full-field stress/birefringence analysis with a polarizing microscope** [3897-27]
A. K. Asundi, B. Zhao, Nanyang Technological Univ. (Singapore)

260 **Grid method for strain measurement in electronic packaging using optical, electronic, and atomic force microscope** [3897-28]
B. Zhao, A. K. Asundi, K. E. Oh, Nanyang Technological Univ. (Singapore)

SESSION 6 VIBRATION SENSORS

272 **Nondestructive diagnostics of layered structures: advanced signal analysis algorithms applied to vibrometric data** [3897-29]
E. P. Tomasini, F. Piazza, E. Esposito, M. Possanzini, Univ. degli Studi di Ancona (Italy)

283 **Torsional vibrations: a laser vibrometry approach** [3897-30]
G. M. Revel, E. P. Tomasini, Univ. degli Studi di Ancona (Italy)

SESSION 7 CURRENT/VOLTAGE SENSORS

290 **Problems of optical fiber voltage sensor in practical usage** [3897-33]
C. Wu, Y. Wang, X. Xin, North Jiaotong Univ. (China)

296 **Optical current transducer with the compensation for the polarization and the temperature disturbance** [3897-34]
D. Liu, J. He, D. Huang, Huazhong Univ. of Science and Technology (China)

304 **Compensation for the effect of incident-light-polarizing state in OCT** [3897-35]
X. Bi, H. Chen, G. Yan, Huazhong Univ. of Science and Technology (China)

SESSION 8 THREE-DIMENSIONAL INSPECTION

314 **New light sources and sensors for active optical 3D inspection (Invited Paper)** [3897-36]
W. Osten, W. P. Jüptner, Bremer Institut für Angewandte Strahltechnik GmbH (Germany)

328 **Automatic measurement equipment for quality control in the loop-slitting department** [3897-37]
S. Hußmann, Univ.-Gesamthochschule Siegen (Germany); W. Kleuver, B. Günther, J. Gröneweller, H. Rath, aicoss GmbH (Germany)

335 **Absolute thickness measurement using automatic fractional fringe order method** [3897-38]
M. Tsai, R. Tian, H. Huang, M. Itoh, T. Yatagai, Univ. of Tsukuba (Japan)

SESSION 9 CONFOCAL MICROSCOPY

342 **Separate measurement of geometrical thickness and refractive index by an interference confocal microscope (Invited Paper)** [3897-39]
T. Fukano, RIKEN—The Institute of Physical and Chemical Research (Japan) and Saitama Univ. (Japan); I. Yamaguchi, RIKEN—The Institute of Physical and Chemical Research (Japan)

355 **Beam shaping with component of cone axicons in confocal system** [3897-40]
W. Tian, Zhejiang Univ. of Technology (China); Z. Ding, Z. Bao, Z. Lu, Z. Liu, Zhejiang Univ. (China)

SESSION 10 MOEMS

364 **Analysis of Mach-Zehnder interferometric micro-opto-electro-mechanical (MOEM) pressure sensor** [3897-42]
T. Srinivas, P. K. Pattnaik, T. B. Narayana, A. Selvarajan, Indian Institute of Science/Bangalore

372 **Development of a pyroelectric thin film infrared sensor by micro-opto-electro-mechanical system (MOEMS) technology** [3897-43]
J.-J. Ho, Fortune Institute of Technology (Taiwan); Y. K. Fang, M. C. Hsieh, National Cheng Kung Univ. (Taiwan); C. Y. Chen, Fortune Institute of Technology (Taiwan); S. F. Ting, National Cheng Kung Univ. (Taiwan); S. M. Lin, Fortune Institute of Technology (Taiwan); K. H. Wu, Nan-Tai Univ. of Technology (Taiwan)

380 **Micro-electro-mechanical fiber optical switches** [3897-44]
A. Q. Liu, F. A. Chollet, National Univ. of Singapore; A. K. Asundi, J. Miao, Nanyang Technological Univ. (Singapore)

388 **Comparison of grating structures in optical fibers and integrated optics for MEMS applications** [3897-49]
S. R. Natarajan, V. M. Murukeshan, A. K. Asundi, Nanyang Technological Univ. (Singapore)

SESSION 11 MICROMETROLOGY

398 **Testing method for microprofile of optical supersmooth surface** [3897-50]
J. Li, Jiangxi Academy of Sciences (China); S. Xiao, Nanchang Univ. (China); X. Li, A. Ying, X. Zhang, A. Zhuo, Jiangxi Academy of Sciences (China)

404 **Thermal and mechanical properties of micromaterials using laser-optical strain sensors** [3897-52]
M. Anwander, A. Hadrboletz, B. Weiss, Univ. of Vienna (Austria); B. G. Zagar, Technical Univ. of Graz (Austria)

414 **Simultaneous strain and displacement measurements in micromechanics** [3897-53]
B. Zhao, A. K. Asundi, Nanyang Technological Univ. (Singapore)

424 **Usage of polarization for high-accuracy micrometrology sensors** [3897-54]
M. Totzeck, H. Jacobsen, H. J. Tiziani, Univ. Stuttgart (Germany)

436 **High-precision measurement scheme for half-wave voltage of Y-tap MIOC** [3897-55]
Y. Yang, W. Zhang, J. Ma, Beijing Univ. of Aeronautics and Astronautics (China)

SESSION 12 SPECKLE SENSORS

442 **Complex monitoring of monumental structures located in seismic and mining-influenced areas with the use of laser tilt sensors and acoustical holography of the ground** [3897-56]
A. Lipowczan, Z. Motyka, H. Passia, A. Szade, Central Mining Institute (Poland)

450 **Shearography as thermomechanical sensor for thermal stress determination of coatings at high temperatures** [3897-57]
K. J. Habib, Kuwait Institute for Scientific Research

453 **Mechanical behavior measurement of wood by ESPI method** [3897-58]
L. J. Jiang, A. K. Asundi, Nanyang Technological Univ. (Singapore); K. Winkelmann, Univ. of Canterbury (New Zealand)

457 **Pattern measurement of speckle noise in optical fiber and its application to sensors** [3897-59]
M. Taniguchi, Meijo Univ. (Japan); M. Oki, Ashikaga Institute of Technology (Japan); T. Takagi, Tohoku Bunka Gakuen Univ. (Japan)

SESSION 13 STRUCTURAL HEALTH MONITORING

468 **Multiplexed FBG sensors and their applications (Invited Paper)** [3897-60]
W. Jin, Hong Kong Polytechnic Univ.

480 **Approach to analysis and measurement of large-array Bragg grating sensors** [3897-61]
T. Srinivas, I. S. Das, A. Selvarajan, Indian Institute of Science/Bangalore

488 **Frequency domain technique for multiplexing of fiber Bragg grating sensors** [3897-62]
P. K. C. Chan, W. Jin, M. S. Demokan, Hong Kong Polytechnic Univ.

497 **Optical time-domain reflectometry for distributed sensing of the structural strain and deformation** [3897-63]
G. Xie, L. K. Seah, A. K. Asundi, Nanyang Technological Univ. (Singapore)

505 **Fiber optic vibration sensor-based smart civil structures** [3897-64]
J. Leng, Nanyang Technological Univ. (Singapore) and Harbin Institute of Technology (China); A. K. Asundi, Nanyang Technological Univ. (Singapore)

511 **Laser vibrometer application on bridge cable tension measurements** [3897-66]
S.-E. Chen, Univ. of Alabama/Birmingham (USA); S. H. Petro, Frederic R. Harris, Inc. (USA)

SESSION 14 NOVEL SENSOR DESIGNS I

522 **Low-cost optical liquid level sensors** [3897-67]
A. K. Ghosh, N. S. Bedi, P. K. Paul, Indian Institute of Technology/Kanpur

534 **Optical surface plasmon resonance sensor designs** [3897-68]
S. Panigrahi, N. B. Das, Regional Engineering College Rourkela (India); A. K. Hassan, A. K. Ray, Sheffield Hallam Univ. (UK)

543 **Separative structure ISFETs on a glass substrate** [3897-69]
L.-T. Yin, Chung Yuan Christian Univ. (Taiwan); J.-C. Chou, National Yunlin Univ. of Science and Technology (Taiwan); W.-Y. Chung, Chung Yuan Christian Univ. (Taiwan); T.-P. Sun, National Chi Nan Univ. (Taiwan); S.-K. Hsiung, Chung Yuan Christian Univ. (Taiwan)

552 **Optically differential phase detection using cross-talk mode of series Fabry-Perot sensors in PMDI** [3897-70]
Y.-L. Lo, C.-H. Wang, National Cheng Kung Univ. (Taiwan)

557 **Theory of an amplified closed-Sagnac-loop interferometric fiber optic gyroscope** [3897-71]
C.-Y. Liaw, Y. Zhou, Y.-L. Lam, Nanyang Technological Univ. (Singapore)

565 **Intensity-based fiber optic pressure sensor with an Au/NiCr/Si_3N_4/SiO_2/Si_3N_4 diaphragm** [3897-72]
J. Park, Keimyung Univ. (Korea); M.-G. Kim, Chungnam National Univ. (Korea); Y. Kim, K.-S. Kim, Korea Electrotechnology Research Inc.; I. Kim, Hansung Univ. (Korea)

570 **Surface roughness measurement with optical scatterometry** [3897-73]
J. Saarinen, Heptagon Oy (Finland) and Helsinki Univ. of Technology (Finland); I. J. Kallioniemi, A. Niinistö, Helsinki Univ. of Technology (Finland); A. T. Friberg, Royal Institute of Technology (Sweden)

578 **Bragg optical displacement sensor** [3897-74]
N. Khan, Univ. Putra Malaysia

SESSION 15 NOVEL SENSOR DESIGNS II

584 **Laser-diode interferometer: practice and application (Invited Paper)** [3897-100]
Y. Ishii, R. Onodera, Univ. of Industrial Technology (Japan)

597 **High-resolution velocimeter based on a fiber optic Sagnac interferometer** [3897-75]
F. Ruan, C.-Y. Liaw, Y. Zhou, Y.-L. Lam, S. H. Mei, Nanyang Technological Univ. (Singapore)

605 **pH sensitivity and hysteresis of A-WO_3 gate ISFET compared with different membranes** [3897-76]
J.-L. Chiang, National Sun Yat-Sen Univ. (Taiwan); J.-C. Chou, National Yunlin Univ. of Science and Technology (Taiwan); Y.-C. Chen, National Sun Yat-Sen Univ. (Taiwan)

614 **Development in high-stability of an optical fiber refractometer using path-matching differential interferometry** [3897-77]
Y.-L. Lo, H.-Y. Lai, W.-C. Wang, National Cheng Kung Univ. (Taiwan)

SESSION 16 POSTER SESSION

624 **Achromatic design strategies with diffractive optical elements** [3897-04]
L. Zhao, Y.-L. Lam, Y. Zhou, Z. Yun, Nanyang Technological Univ. (Singapore)

632 **Planar diffractive imaging element design** [3897-79]
Z. Yun, Y.-L. Lam, Y. Zhou, L. Zhao, Nanyang Technological Univ. (Singapore)

640 **Highly efficient diffraction gratings and resonant plasma layers at IR wavelengths** [3897-80]
J. H. Stiens, Vrije Univ. Brussel (Belgium); G. A. Nemova, Institute of Radio Engineering and Electronics (Russia); P. F. Muys, Laser Power Europe NV (Belgium); R. A. Vounckx, Vrije Univ. Brussel (Belgium)

650 **Development and application of micro-moiré interferometer** [3897-81]
K. E. Oh, G. B. Chai, A. K. Asundi, Nanyang Technological Univ. (Singapore); K. C. Chan, T. C. Chai, Institute of Microelectronics (Singapore)

658 **Photoexcitation-induced current sensing on semi-insulating GaAs using a tunneling microscope tip** [3897-82]
K. Kawashima, S. Takai, G. Kudou, H. Adachi, M. Takeuchi, K. Fujiwara, Kyushu Institute of Technology (Japan)

666 **Large-aperture continuous-phase diffractive optical element for beam transform** [3897-83]
Q. Tan, Y. Yan, G. Jin, M. Wu, Tsinghua Univ. (China)

674 **Micro-optics element for color separating** [3897-84]
J. Wang, Y. Yan, Q. Tan, G. Jin, Tsinghua Univ. (China)

686 **Dynamic and spectral responses of a micromachined Fabry-Perot interferometer-based spectrometer** [3897-86]
K.-K. Liu, Nanyang Technological Univ. (Singapore); S.-P. Lee, R.-J. Lin, Industrial Technology Research Institute (Taiwan); L. Hsu, National Chiao Tung Univ. (Taiwan)

692 **External-cavity semiconductor laser sensor** [3897-88]
O. L. Quan, Y. T. Seng, L. P. Lan, J. T. K. Wah, A. T. Sabaratnam, Ngee Ann Polytechnic (Singapore); G. M. Hegde, A. Selvarajan, Indian Institute of Science/Bangalore

700 **Nonintrusive surface inspection utilizing infrared light radiation** [3897-89]
H. A. Hamid, W. Wlodarski, F. Brennan, Royal Melbourne Institute of Technology (Australia)

708 **Detection of cracks in aluminum beams using fiber optic polarimetric sensor** [3897-90]
J. Ma, A. K. Asundi, Nanyang Technological Univ. (Singapore)

713 **Spectrally resolved white-light interferometry for profilometry with polarization phase shifter** [3897-91]
S. Suja Helen, M. P. Kothiyal, R. S. Sirohi, Indian Institute of Technology/Madras

719 **Method of measurement with high-speed external diameter outline** [3897-92]
B. Liu, Tianjin Institute of Laser Technology (China) and Tianjin OE TEC Co., Ltd. (China); B. Chen, K. Shi, Tianjin OE TEC Co., Ltd. (China)

728 **220-kV combined optical voltage and current transformer** [3897-93]
S. Luo, M. Ye, Y. Zhu, Y. Cui, Y. Xu, K. Li, C. Ou, Z. Chen, Huazhong Univ. of Science and Technology (China)

734 **Overcoming temperature offset and noise affection in optical current transducer** [3897-94]
G. Yan, Y. Huang, H. Chen, Z. Sun, Huazhong Univ. of Science and Technology (China)

742 **Effect of target position on signal amplitude in laser Doppler vibrometer** [3897-95]
V. Aranchuk, Scientific Ctr. for Machine Mechanics Problems (Belarus)

750 **Simulation and study on the temperature effect of the a-Ta_2O_5 ISFET** [3897-97]
J.-C. Chou, Y. S. Li, National Yunlin Univ. of Science and Technology (Taiwan)

758 **pH response of a-Si:H ISFET** [3897-98]
J.-C. Chou, J.-S. Lin, National Yunlin Univ. of Science and Technology (Taiwan)

767 **Precise measurement of the difference of the air-refractive indices between visible and near-infrared wavelengths using two-color interferometer** [3897-99]
I. Fujima, S. Iwasaki, National Research Lab. of Metrology (Japan); G. Xie, Nanyang Technological Univ. (Singapore); K. Seta, National Research Lab. of Metrology (Japan)

774 *Author Index*

Conference Committee

Conference Chairs

Robert A. Lieberman, Physical Optics Corporation (USA)
Anand K. Asundi, Nanyang Technological University (Singapore)
Hiroshi Asanuma, Chiba University (Japan)

Program Committee

Piero Bruscaglioni, University of Florence (Italy)
Michael A. Butler, Sandia National Laboratories (USA)
Arthur E. Chiou, National Dong Hwa University (Taiwan)
Brian Culshaw, University of Strathclyde (UK)
Kenneth T. Grattan, City University London (UK)
Roger A. Lessard, COPL/Université Laval (Canada)
Brian D. MacCraith, Dublin City University (Ireland)
Kazuo Maeno, Chiba University (Japan)
Wolfgang Osten, Bremer Institut für Angewandte Strahltechnik GmbH (Germany)
Ananth Selvarajan, Indian Institute of Science/Bangalore
Colin J. Sheppard, University of Sydney (Australia)
Rajpal S. Sirohi, Indian Institute of Technology/Madras
Swee Chuan Tjin, Nanyang Technological University (Singapore)
Dietmar Vogel, Fraunhofer-Institut für Zuverlassigkeit und Mikrointegration (Germany)
Wei Jin, Hong Kong Polytechnic University

Session Chairs

1 Photonic Sensors Materials
Ananth Selvarajan, Indian Institute of Science/Bangalore

2 Photonic Sensors Design and Fabrication
Roger A. Lessard, COPL/Université Laval (Canada)

3 Biomedical and Biochemical Sensors I
S. M. Krishnan, Nanyang Technological University (Singapore)

4 Biomedical and Biochemical Sensors II
Yan Zhou, Nanyang Technological University (Singapore)

5 Micro-optical Metrology
Wolfgang Osten, Bremer Institut für Angewandte Strahltechnik GmbH (Germany)

6 Vibration Sensors
Zhong Zhao Wei, Nanyang Technological University (Singapore)

7 Current/Voltage Sensors
Bing Zhao, Nanyang Technological University (Singapore)

8 Three-Dimensional Inspection
Werner P. Jüptner, Bremer Institut für Angewandte Strahltechnik GmbH (Germany)

9 Confocal Microscopy
Ingrid DeWolf, IMEC (Belgium)

10 MOEMS
Toyohiko Yatagai, University of Tsukuba (Japan)

11 Micrometrology
Dietmar Vogel, Fraunhofer-Institut für Zuverlassigkeit und Mikrointegration (Germany)

12 Speckle Sensors
Vadakke M. Murukeshan, Nanyang Technological University (Singapore)

13 Structural Health Monitoring
James Brownjohn, Nanyang Technological University (Singapore)

14 Novel Sensor Designs I
Puttappa Jayanth, Ngee Ann Polytechnic (Singapore)

15 Novel Sensor Designs II
Seriampalayam R. Natarajan, Nanyang Technological University (Singapore)

SPIE and Nanyang Technological University wish to thank the following for their contribution to the success of this conference:

- United States Air Force Asian Office of Aerospace Research and Development (AFOSR/AOARD)
- United States Army Research Office-Far East (ARO-FE)

Plenary Session I

Future Technology and Business Opportunities in Photonics: A View from the Optoelectronics Industry Development Association (OIDA) *

Dr. Arpad A. Bergh

Optoelectronics Industry Development Association, 2010 Massachusetts Avenue, NW, Suite 200, Washington, DC 20036, P:202-785-4426, F: 202-785-4428, email: aboida@osa.org

ABSTRACT

The "optoelectronics industry" is a collection of six or more distinct industries that all depend on OE technology. The major markets are in communication, imaging, storage and displays. This paper gives a brief overview of the anticipated paradigm shifts, the potential markets and the promising new technologies in various OE markets.

Keywords: optoelectronics industry, optical communications, imaging, optical storage

1. INTRODUCTION

The dictionary defines Optoelectronics as "a branch of electronics that deals with electronic devices for emitting, modulating, transmitting, and sensing light".
Today Optoelectronics has a broader meaning. It also incorporates Electro-optics, a branch of physics that deals with the effects of an electric field on light transversing it (e.g. gas and solid state lasers), and Photonics, a branch of physics that deals with the properties and applications of photons, especially as a medium for transmitting information. It also increasingly overlaps with many facets of Optics, a science that deals with light, its genesis and propagation.

This relatively new branch of science and engineering interchangeably deploys electrons and photons, to take advantage of their best attributes. It has already laid the foundation of the technology that enables the information age. Optoelectronics enables all functions of the information industry. It is essential in gathering the information (imaging), as well as in transporting and displaying it. It also has a role, along with other technologies, in storing and processing information. Due to their increasingly important role optoelectronics components already represent a substantial worldwide market exceeding by a factor of three the market for traditional optics, Fig 1.

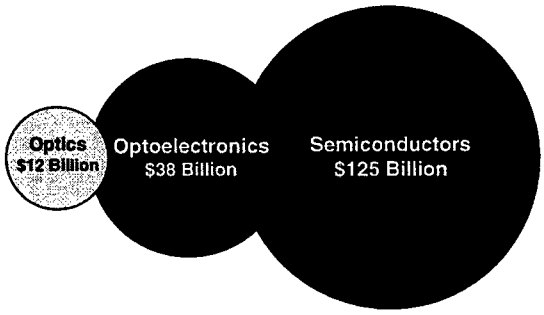


Figure 1. Worldwide Component Markets 1998

This is indeed remarkable when we consider that these elementary particles were only discovered in this Century.

Discovery of the electron:
It was in 1887, when J. J. Thomson working in the Cavendish laboratory at Cambridge measured the ratio of the charge e of an electron to its mass m by observing its deflection in both electric and magnetic fields. The discovery of electron is usually said to date from this historic experiment.

*Also published in *Proc. of SPIE* Vols. 3896, 3898, and 3899

Part of the SPIE Conference on Advanced Photonic Sensors and Applications
Singapore • November–December 1999
SPIE Vol. 3897 • 0277-786X/99/$10.00

Discovery of the Photon.

The discovery of the photon dates back to Planck's experiments in 1900 on the energy of photo-electronics leading to *h*, the Planck constant, 6.6254 x 10E-27 erg sec. This was followed in 1905 by the famous Planck-Einstein equation:

$E = h\nu$,

Where *h* is the Planck constant and E is the energy of a quantum of light (called photon by others at a later date).

Some of the major attributes of the two particles explain their application domains and are compared in the following table:

Attributes	Electrons	Photons
wavelength	3 cm – 30 m	500 nm
Frequency	10 MHz – 10 GHz	500 THz
Energy	40 neV – 40 μeV	2 eV
Propagation loss	High (in copper wire)	Low (in optical fiber)
Particle interaction	High	None

There are certain functions that can be uniquely provided by optoelectronics. In other areas OE competes with other technologies or assumes a supporting role as shown in the following table.

The Role of Optoelectronics in the Information Age Technologies

Dominant Enabling Technology	Competing with Other Technologies (Market Share)	Supporting Role Dominated by Electronics
Transmission	Sensors	Logic functions
Display	Storage	Processing
Imaging	Lighting	

2. OPTOELECTRONICS INDUSTRY

There is no single optoelectronics industry. It is, in fact, a combination of several industries supported by common technology and common infrastructure as shown in Fig. 2

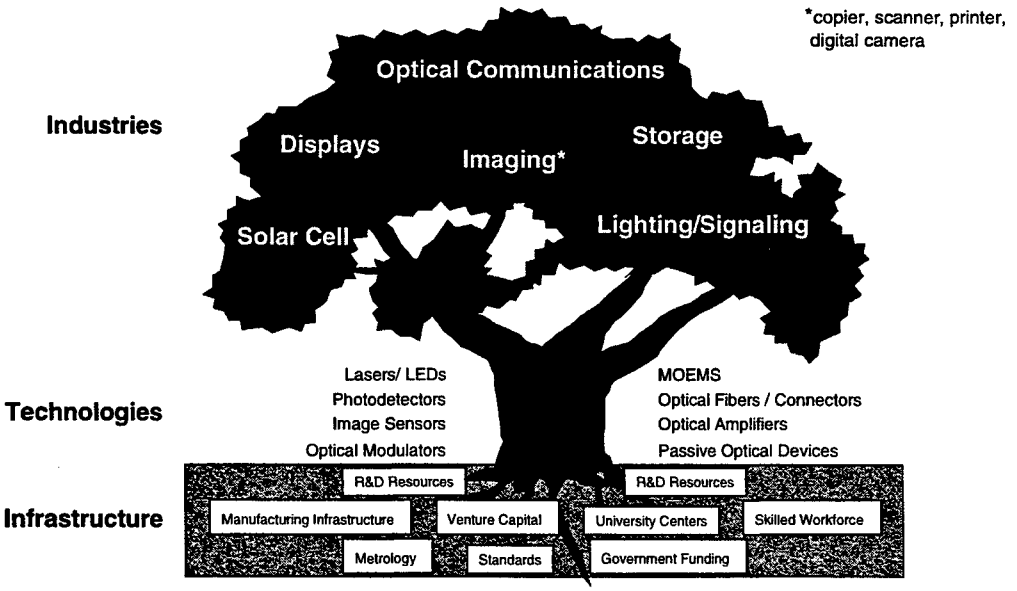


Figure 2. Optoelectronics Industries

Of the industry segments depicted in this figure communications, imaging, storage and displays have a large worldwide market, Fig 3 Emerging markets in optoelectronics are energy related and will have an increasingly important role as the world's known energy sources are depleted. Solar cells produce environmentally clean energy and solid state lighting has the promise of providing more efficient and environmentally benign replacement for the currently used incandescent and fluorescent light sources.

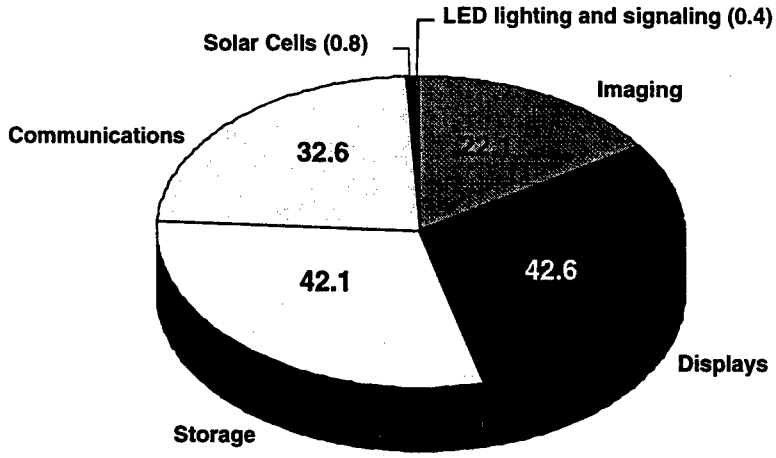


Figure 3. World OE Production ($B) 1998 (Total: $140.6 B, North America Estimate: $42.8B)

This paper gives a brief overview of the anticipated paradigm shifts, the potential markets and the promising new technologies in various OE markets.

3. OPTICAL COMMUNICATION

Driven by Internet and data services, the evolution of information networks continues at an unprecedented pace. The traditional telephone architecture is rapidly changing:

- Transmission is moving to all-digital before the end of the next decade
- Voice services are increasingly augmented by multi media (incl. imaging)
- Voice, relative to data, becomes insignificant part of transmission bandwidth
- Intelligence is moving from the central office to the perimeter of the network
- The stationary telephone is replaced by the mobile telephone
- Telecom and datacom networks converge.

All these changes are made possible by the evolution of the electronic and optoelectronic technologies. Multi-Gb/s SONET/SDH and WDM transmission, resulting in capacities in excess of one Terabit/sec on a single fiber meet the explosive demand for bandwidth. Simultaneously, the increasing dominance of Internet, Web and data services is causing a conversion from voice- to data-centric networks, with a need for high-speed and broadband transmission at the lowest-possible cost.

In addition to the ongoing research and development at universities and industry the government has also played a significant role in this evolution. Some of the fruits of DARPA sponsored technologies are being implemented commercially and are changing the communications landscape. On the service level there is the Internet, on the system level WDM is reaching all segments of the network and on the device level VCSELs promise to lower costs, facilitate the integration of electronic and photonic components and render optical networks more affordable. Simultaneously the 1996 Telecommunication Act is unleashing competition among service providers and the consumer is left with a bewildering array of options. The streets of major cities are under attack from competing carriers laying new cable, and mobile phones are spreading faster than before.

OIDA's roadmapping in optical communication started in 1992. The most recent studies took place in 1998 and 1999. The 1998 study concentrated on the anticipated demand on communication systems that can deliver the required services. The predicted growth rate of the various communication services is shown in Fig 4.

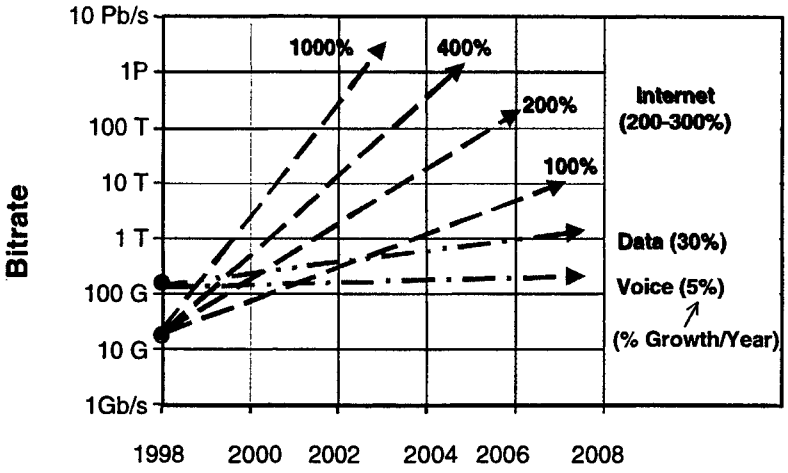


Figure 4. Growth Rate of Communication Services

This amazing growth rate is fueled by the evolution of the communication bandwidth which got a major boost in the mid-nineties from the commercial implementation of wavelength division multiplexing (WDM). To keep up with this growth and to deploy all the technology that has been demonstrated to date many practical and commercial issues must be addressed. The migration of the optical technology from the high performance backbone networks toward the high volume access network requires a drastic cost reduction of the OE components. The relentless cost reduction in the transmission network expressed in bit-km is shown in Fig 5.

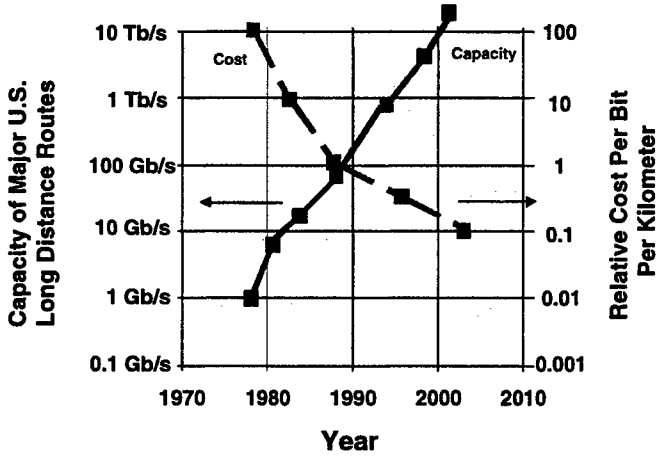


Figure 5. Transmission Cost Erosion, Ref. Yanming Liu, Corning

When the study was initiated two years ago it was obvious that communication would expand exponentially. It was also obvious that communication is segmented into several markets with sometimes overlapping and other times with divergent requirements.

The main areas of applications of the key OE components are:

- Long-Distance,

- Metro/Interoffice,
- Enterprise and Metro/Access,
- Local Area Network/Wide Area Network,
- Local Access Networks, and
- Backplane/Building Wiring.

In developing a roadmap estimates are made of the potential market volumes of the components in these various network segments, along with the key performance parameters and price targets as function of time. It is, however, exceedingly difficult to make predictions on the growth of the market volume of specific optoelectronic components. Nevertheless, using trend curves, latest technology developments and input from industry sources and analysts, a roadmap is constructed with the proviso that the predictions must be updated as data on actual equipment deployments and network applications become available.

The 1999 OIDA Roadmap Report delineated the OE components that require major industry efforts for their commercial realization. The following is a sample of key components required for optical networks with significant near term (1 to 4 years) market potential:

- Optical Crossconnects
- Optical Switches
- Optical Add-Drop Multiplexers
- Tunable Lasers
- Optical Amplifiers
- Optical Transceivers

An overriding requirement of these products is the simultaneous improvement in performance and ongoing cost reduction. The cost of components on the other hand depends on the volume of production. It is difficult to build high volumes when there is a proliferation of products for a given function. Optical switching for instance is one of the most desired functions but there are at least nine different approaches in which this can be accomplished. It will take several years and several generations of equipment before commonality is reached and standards are developed. In the meantime one can expect a steady increase in component performance with ongoing price erosion as shown in the next OIDA roadmap chart, Fig 6.

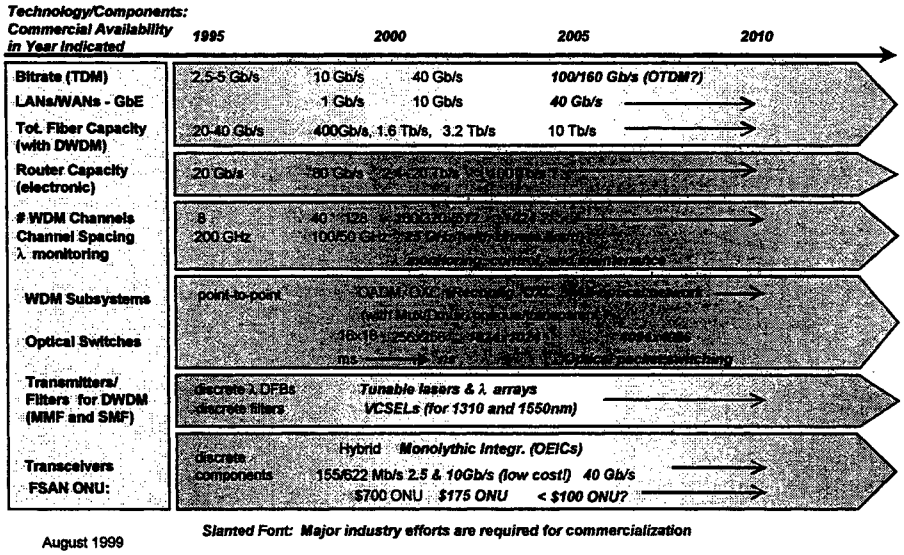


Figure 6. OIDA Communications Roadmap – 1999

All this activity will fuel the growth of the optoelectronic communications market, which in 1997 has already reached $ 30 billion worldwide. As indicated in Fig 7 this market is expected to double by 2001 with no saturation in site for many years to come.

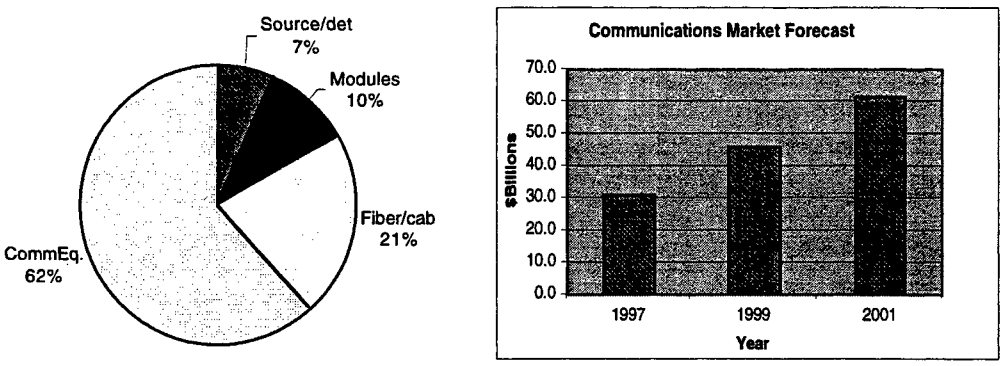


Figure 7. Optoelectronic Communications Market

4. IMAGING

Growth in the market value of the imaging industry has significantly lagged other technology sectors of the economy over the last five years as evidenced by the aggregate stock performance of imaging companies compared to the performance of the S&P 500 index. The Salomon Smith Barney Imaging Composite (SSIC), a weighted index of 50 imaging companies, widely used to track industry trends was up only 12%, versus more than 25% for the S&P 500 index.

The SSIC is composed of 5 sectors with Diversified Imaging/Photo and Document Processing companies representing 90% of the value of the composite. This Sector actually declined 4% last year, since many products in this sector are "mature" and experienced rapid price declines and decelerating volumes. Companies in this sector are looking to Advanced Imaging Applications as key growth drivers for future earnings.

Imaging companies are looking to Advanced Imaging Applications as key growth drivers for future earnings.

Advanced Imaging refers to the extension of conventional imaging products such as photographic film and cameras, video, copiers etc. as well as enabling entirely new capabilities through the application of digital technology. In the image capture area for example, electronic still cameras are replacing instant photography for scientific, insurance and real estate applications. Last year in Japan digital camcorders outsold both 8mm and VHS analog models. Worldwide digital camcorder sales totaled over $3B. In 1998, it is estimated that 1.1 million digital cameras were sold in the U.S., up from 740,000 the prior year. Those figures should expand rapidly as PC based imaging becomes more user friendly and photo-realistic digital printers are able to interface directly to digital cameras.

Also, digitization is expected to provide significant growth for traditional film-based imaging companies and photofinishers. Digitization refers to high-resolution scanning and digital printing of film images and photographs. While commercial labs and motion picture special effects houses have transitioned to a digital infrastructure over the last 10 years, minilabs and wholesale photofinishers are just now beginning to install equipment that allows in line digitization of film with minimal impact on workflow.

When this digital infrastructure is in place and broadband internet access is widely available, it will enable a wide range of internet based imaging products and services. Also, new partnerships and joint ventures between film companies, internet companies and chip manufacturers are expected to increase growth. Many advanced imaging applications will be enabled by the wide spread access of broadband communication to the internet, greater microprocessor performance and highly integrated optoelectronic components.

Once optical networks enable broadband access to the internet, a major beneficiary will be imaging with many opportunities for instant transfer of still and motion picture images.

While the potential market for military and industrial imaging applications at $3 to $5 billion is much smaller than digital photo-finishing and networked imaging services ($15 to $20 billion), the technology demands will be an important driver for advanced imaging capabilities. Current market forecasts are based on extrapolations and are probably too conservative.

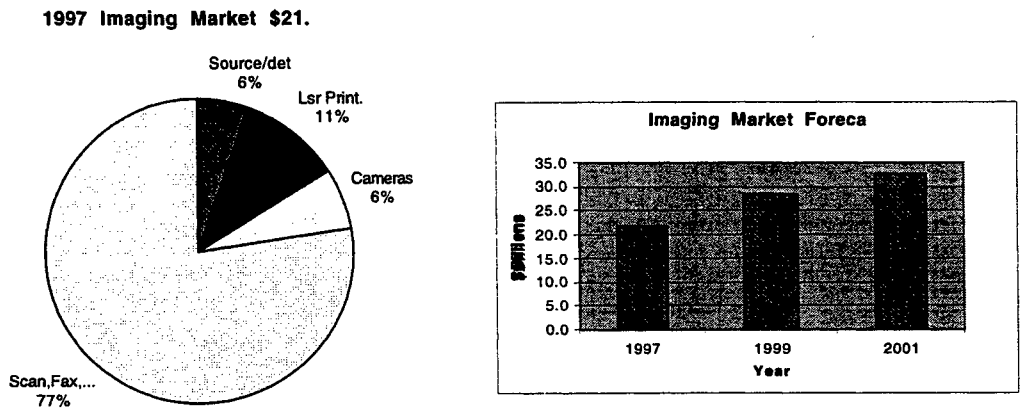


Figure 8. Optoelectronics Imaging Market

Typical areas of growth in imaging will benefit from

- Digital image format
- Digital signal processing with a great latitude of manipulating images
- Intelligent image capture/tracking
- Improved image sensor technology expanding the spectral and spatial resolution at an affordable cost
- Miniaturization, increasing portability and lowering power
- Fiber to the home, and
- Increased Small Office – Home Office use of imaging equipment.

Expanded Vision refers to the ability to outperform human vision.
Expanded Vision Characteristics are:

- Wider Field of View
- Broader Spectral Range (Uncooled detector)
- Increased Spatial Resolution
- Increased depth of field

These attributes are achieved by using advanced image sensors and by electronically processing the gathered information, such as in image fusion. In the future images acquired from multiple detectors will be combined using image fusion algorithms with integrated optoelectronics hardware systems to provide enhanced performance over a much broader spectral range than products are currently capable of delivering.

Night Vision Products
Night vision products have a long history dating back to the '60s. Refined military applications developed in the last two decades leading to many advanced products.
Currently military and law enforcement applications represent the largest segment of the night vision equipment market. In 1998 the market size for this segment exceeded $3B. However, in the 2003 –2005 time frame, automotive, surveillance and home security applications are expected to approach the size of the current military /law enforcement segment. Optoelectronic products are estimated to be roughly 1/3 the total market size for these applications.

Miniaturization in Imaging – "The Micro Camera"
Many of the applications for micro cameras are similar to those identified for digital cameras, which are currently enjoying double digit growth in the marketplace. The acceptance of digital cameras as solutions for advanced imaging applications has been paced by the ability of the technology to meet market needs better than existing technology i.e. traditional photography. Micro camera characteristics include:

- Significantly Smaller Size
- Higher Spatial Resolution (In Color!)
- Lower Power

Today, micro cameras are used largely by law enforcement, military and private security agencies for surveillance and home security applications. In total these represent a market size of $1B/yr. As digital cameras become a common peripheral for desktop and notebook PCs, demand for micro cameras for notebooks PDAs and other portable appliances will drive significant demand. In the 2003 –2005 time frame, these applications will represent the largest market segment followed by surveillance and inspection systems. In the table below, the optoelectronic products are estimated to be about 20% of the total market.

Intelligent Image Capture
The intelligent image capture system addresses the technology associated with the integration of a complex imaging system capable of automatically recognizing, isolating and tracking a target or object of interest in a scene in real time. Imaging applications identified for intelligent image capture also drive significant demand for digital image processing hardware as system complexity increases. The required image processing speeds for future systems might outpace Moore's Law.

Although many of the key enabling technologies for Intelligent Image Capture are algorithm based, the integrated systems designed for most applications drive significant demand for optoelectronic components.

Market Potential for Tracking Systems

Application	Estimated Total Market for Application in 2003	Estimated Market Size for OE Products
Weapons/Guidance Target Tracking	$5B	$1B
Surveillance / Security	$100M	$20M
Automotive	$1B	$200M
Consumer Web Camera	$500M	$100M

5. OPTICAL STORAGE

Data storage is an area where the optical approach competes with other technologies, such as magnetic hard disc, magnetic tape, semiconductor memory, etc.

Optical storage offers a reliable and removable storage medium with excellent archival lifetime at a low cost. Both optical recording and readout can be performed with a head positioned relatively far from the storage medium, unlike magnetic hard drive heads. This allows the medium to be reliable and removable, but the heavier head also leads to slower access time compared to hard disk drives. Consequently optical storage is limited to applications requiring reliability and removability, such as archival storage, software distribution, medical imaging, storing digital photographs, etc.

The market for disc storage exceeded $60 B/year in 1996 with a small fraction attributed to optical storage as shown in Fig 9

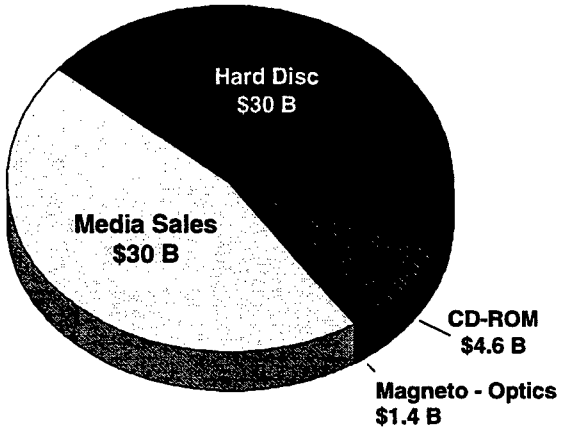


Figure 9. Worldwide Storage Market - 1996

With the rapid expansion of the Internet, server-based applications will emerge. Electronic commerce, medical imaging, libraries and corporate networks require modest access time (<10 ms) but very large storage capacities and reasonable transfer rates. These applications might give a further boost to the optical disc market. Optical disk market is also expected to grow for computer applications with the new DVD format gaining market share as shown in Fig 10

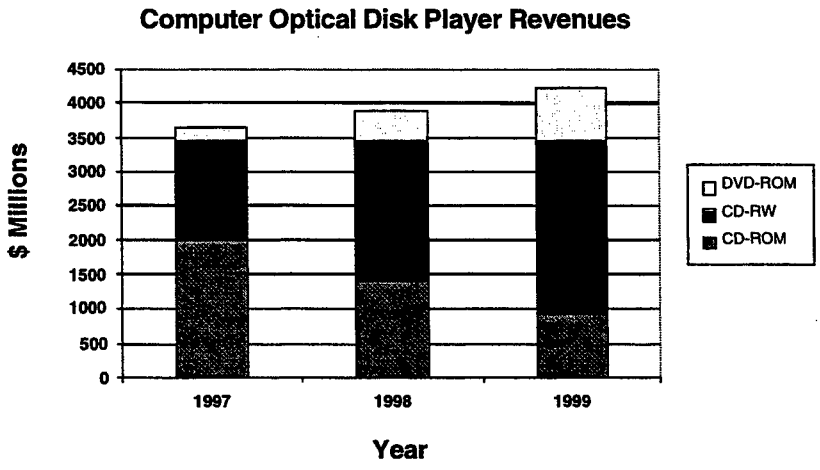


Figure 10. Optical Storage for Computers, Source: OITDA

It should be noted that storage technology is facing the same cost reduction pressure than all other market segments. Similar to the charts shown for communication, storage has a history of steady improvement in performance with a decline in unit price. For optical storage to maintain its position and increase its market share it must outperform competing technologies on both counts.

6. FUTURE TRENDS IN OPTOELECTRONICS

Future opportunities in optoelectronics will come in three areas:

- Improvements in technology – new breakthroughs
- Low cost, high volume manufacturing, and
- Integration with other technologies.

The first two items require no explanation. Integration requires the vertical integration of traditional academic disciplines. It should be a natural extension of optoelectronics which itself was created from the integration of optics and electronics. Future opportunities will rise from the integration with electronics, electronic processing (including software), micromechanics, etc as depicted in Fig 11

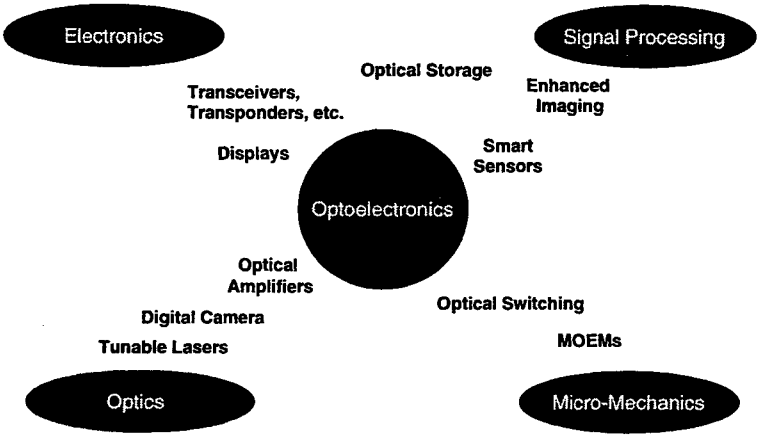


Figure 11. OE Centered Technology Integration

These are exciting times to be in optoelectronics. Our research and products will have major impact in many application areas from communication to transportation, medicine, biotechnology, electronic commerce, entertainment, etc. It will also support the leading technologies of the 21st Century, Fig 12.

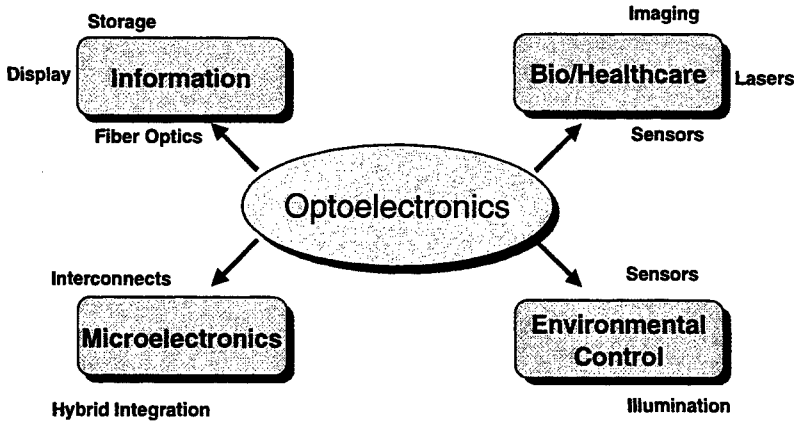


Figure 12. Leading Technologies for the 21st Century

Oxide-Confined Vertical-Cavity Surface-Emitting Lasers, Quantum Dots, and the Purcell Effect: Can Scaling the Mode Size Improve Laser Performance? *

D.G. Deppe, D.L. Huffaker, H. Huang, and L.A. Graham

Microelectronics Research Center
Department of Electrical and Computer Engineering
The University of Texas at Austin, Austin, Texas 78712-1084

ABSTRACT

The development of vertical-cavity surface-emitting lasers (VCSELs) has led to new types of low power, high efficiency light sources for data communication. The small size, low power, and surface-normal emission of VCSELs has enabled relatively dense two-dimensional arrays for highly parallel data communication and optical signal processing. In this paper we examine the issues of device scaling on VCSEL performance. We look specifically at what benefits may be derived from continued scaling of the active volume down to minimum sized dimensions, and what device schemes may be required to obtain the scaling. Laser rate equations are used to show that when the VCSEL mode volume is reduced to wavelength cubed dimensions, a significant improvement in modulation speed is predicted based on the radiative lifetime change due to the Purcell effect. However, several parasitic effects must be controlled in order to realize these benefits. Most important are control of the optical loss due to diffraction or scattering, and control of the electronic losses due to carrier diffusion and surface effects. Novel optical confinement schemes based on oxide-apertures, photonic bandgaps, and/or closely coupled two-dimensional arrays may be useful for controlling optical loss, while self-assembled quantum dots are attractive for controlling electronic diffusion to dimensions within the minimum optical mode volume.

Keywords: Semiconductor lasers, VCSELs, microcavities, cavity design, quantum dots, controlled spontaneous emission, high speed modulation.

1. INTRODUCTION

Vertical-cavity surface-emitting lasers (VCSELs) have emerged as one of the most important semiconductor lasers for low power optical interconnects.[1] Most commercial VCSELs are presently based on proton implantation to control the current injection into a small area of the optical cavity.[2] The optical mode volume is then defined by gain-guiding. This fabrication scheme has numerous proven advantages for manufacturing VCSELs including photolithographic patterning of the active area, planar processing, high device reliability, and high yield. On the other hand, in many ways devices based on gain-guiding have inferior performance to index-confined VCSELs, especially oxide-confined VCSELs.[3] For gain-guided VCSELs, diffraction loss from the optical mode limits scaling the active volume to very small sizes, and can lead to instabilities in the VCSEL threshold and operating characteristics due to thermally induced index-guiding. Because of this there is a wide-spread industrial research effort in developing commercial VCSELs based on oxide confinement, for which these thermal effects are reduced are eliminated by the built-in index guide of an oxide-aperture. As compared to proton implanted VCSELs, oxide-confined VCSELs have low threshold currents and lower threshold current densities,[3-9] higher wall-plug efficiency,[10-12] higher modulation speed,[13,14] and improved mode characteristics for both multi-mode and single-mode operation.[15,16] The lower power operation is crucial for two-dimensional arrays for reducing thermal cross-talk and increasing the VCSEL packing density. Oxide-confined VCSELs also have greatly improved scaling behavior. While gain-guided VCSELs that operate at 0.85 μm or 0.98 μm begin to show diffraction loss effects when the optical mode is reduced below ~10 μm diameter, oxide-confined VCSELs show good scaling behavior down to ~3 μm diameter. The improved scaling behavior is due to control of the diffraction loss by the oxide aperture.[17,18]

In this paper we examine the impact of scaling the VCSEL optical mode to even smaller sizes than that presently achieved, ultimately to volumes approaching a cubic wavelength. If such small optical modes can be achieved new VCSEL characteristics are predicted, especially for the modulation response. Simply scaling the active volume to small size can greatly reduce the lasing threshold current, since fewer electronic states must be inverted to achieve optical gain. This may be of interest for realizing extremely low threshold VCSELs for new ultralow power operations, such as for optical interconnects for dense focal plane arrays.[19] As we discuss below, though, a potentially more interesting result is the decreased radiative

*Also published in *Proc. of SPIE* Vols. 3896, 3898, and 3899

Part of the SPIE Conference on Advanced Photonic Sensors and Applications
Singapore • November–December 1999
SPIE Vol. 3897 • 0277-786X/99/$10.00

lifetime that results due to the Purcell effect, which can lead to high speed VCSEL operation even for bias levels at or below lasing threshold. High speed bias-free operation may then be possible, greatly improving VCSEL array performance and further reducing power consumption. Although the maximum power from such a small optical mode is reduced over today's commercial VCSELs, this power limitation can in principle be eliminated by using microarrays of Purcell enhanced VCSELs.

At present there are two serious device constraints that prevent the scaling of VCSELs to small optical modes. The first is optical loss. Depending on the mode confinement scheme, either diffraction or scattering loss usually accompanies a reduction in the optical mode size, and this is a major scaling limitation in today's VCSELs. In principle diffraction or scattering loss may be eliminated through lateral confinement using either very short cavities along with oxide apertures or photonic bandgaps for lateral confinement, or perhaps more simply by diffractively coupling the elements of a microarray. However, if optical loss is not controlled the gain required for lasing threshold in the small mode volume can exceed the maximum available from the active material. The second constraint is in the electronic confinement. Planar quantum well active regions can lead to lateral carrier diffusion lengths of several microns at room temperature, while electrons and holes must effectively be confined to the optical mode volume in order to realize the Purcell effect. Actually, the Purcell effect can reduce the diffusion length by shortening the recombination time, but it would be more effective to achieve tight lateral electronic confinement to sub-micron optical mode sizes. This can be obtained using quantum dot (QD) light emitters, for which lateral electronic confinement is "built-in" due to the self-organization of highly strained epitaxial films. The recent advances in QD lasers,[20-24] including QD VCSELs,[25-29] shows that this approach appears viable. Note that the diffusion problem is not of such a concern for infrared lasers, for which the optical wavelengths are a few microns within the semiconductor material.

Below we discuss the Purcell effect in oxide-confined VCSELs in detail. Experimental results on QD and planar quantum well microcavity light emitters and lasers aimed at characterizing the Purcell effect are described. Rate equations are used to predict the impact of the Purcell effect on VCSELs, both for lasing threshold and modulation response.

2. PURCELL EFFECT IN AN APERTURED-MICROCAVITY

It is well-known that an emitter's spontaneous emission rate is controlled by the volume and Q (loss rate) of the optical modes to which it's coupled.[30,31] Recently it was shown that the spontaneous lifetime can be controlled in apertured-microcavities[32-34] similar to those used for very low threshold VCSELs. Although Purcell's original arguments were brief and based on an ideal single mode cavity,[30] the VCSEL cavity typically contains many optical modes that exist within the emitters optical bandwidth. However, despite the existence of many optical modes, aperture confinement (or an etched pillar) can lead to one or two highly confined optical modes with sufficient Q to dominate the microcavity emission characteristics. This is illustrated in Fig. 1 for a VCSEL-type cavity based on distributed Bragg reflectors (DBRs). Each high index layer of the microcavity forms waveguide modes propagating in the plane of the layer, while apertured modes also exist that propagate back and forth vertically to the mirrors.

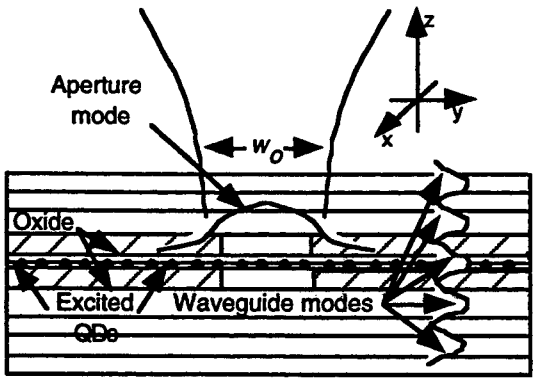


Fig. 1 Schematic illustration of an oxide-apertured microcavity based on QD light emitters and DBRs. Resonant optical modes exist due to the high index layers of mirrors and due to the double oxide apertures. However, for very small apertures the combined effects of Q and mode volume can lead to nearly 100% of the spontaneous emission being coupled into the aperture modes. (After Ref. [34])

When QD light emitters are used within the microcavity their coupling to the electromagnetic field depend on their positions within the cavity. Those QD emitters lying within the oxide-aperture interact with both the waveguide modes and the apertured-modes, while those QD emitters that lie outside the aperture interact with only the waveguide modes. We model the

QD light emitters as two-level systems. When only spontaneous emission is important for QD level decay, the total spontaneous from an emitter depends on the summation over all optical modes as given by

$$\frac{dN_2}{dt} = -\frac{2q^2\omega_d^2}{\hbar^2}\sum_m |\mathbf{d}\cdot\mathbf{A}_m(\mathbf{r}_d)|^2 \frac{(\gamma_d + \frac{\omega_m}{2Q_m})}{(\omega_d-\omega_m)^2 + (\gamma_d + \frac{\omega_m}{2Q_m})^2} N_2(t). \quad (1)$$

In Eq. (1), N_2 is the upper level population, m labels each cavity mode, q is the electronic charge, ω_d is the resonant frequency of the 2-level emitter, ω_m is the resonant frequency of mode m, V is the normalization volume of the cavity, $\mathbf{d}$ is the dipole vector strength, $\mathbf{A}_m(\mathbf{r}_d)$ is the normalized vector strength of the cavity field at the QD position, $\frac{\omega_m}{2Q_m}$ is the photon loss rate from mode m, and γ_d is the QD dephasing rate. The vector strength of the cavity field is normalized such that $\int_V d^3r\varepsilon(\mathbf{r})|\mathbf{A}_m(\mathbf{r})|^2 = \hbar/(2\omega_m)$, where $\varepsilon(\mathbf{r})$ is the material permittivity at $\mathbf{r}$.

The collection of the cavity emission can be restricted to the apertured-modes using the spectral dependence of the emission from the cavity. These apertured-modes have frequencies that increase with the transverse mode number. By making very small apertures, higher order transverse modes can be eliminated from the bandwidth of the QD emitters, so that only the two lowest order transverse modes collect significant emission. The spontaneous photon number in these lower-order modes satisfy the rate equation

$$\frac{dn_o}{dt} = -\frac{\omega_o}{Q}n_o + \sum_{n=1}^{N_{QD}} \frac{2q^2\omega_n^2}{\hbar^2}|\mathbf{d}_n\cdot\mathbf{A}_o(\mathbf{r}_n)|^2 \frac{(\gamma_n + \frac{\omega_o}{2Q})N_{2,n}(t)}{(\omega_n-\omega_o)^2 + (\gamma_n + \frac{\omega_o}{2Q})^2} \quad (2)$$

where the subscript n labels each emitter. We approximate the apertured-modes as Gaussian, and the QD emitters as having randomly oriented dipole moments. The coupling strength for these modes [Eqs. (1) or (2)] then becomes $|\mathbf{d}_n\cdot\mathbf{A}_o(\mathbf{r}_n)|^2 = \frac{|d|^2\hbar e^{-(x_n^2+y_n^2)/w_o^2}}{3\varepsilon\omega_o\pi w_o^2 L_z}$, with a mode volume given by $V = \pi w_o^2 L_z$ where w_o is the mode radius and L_z is an effective cavity length. It has previously been estimated that when an oxide-apertures is not present, 80 to 90% of the spontaneous emission is coupled to waveguide modes (see Fig. 1), and this emission is approximately independent of the aperture. In terms of the dipole moment, the spontaneous emission rate of the QD embedded in bulk material is $\frac{1}{\tau_{sp,o}} = \frac{q^2\omega_n^3 n^3|d|^2}{3\pi\varepsilon\hbar c^3}$, where n is the refractive index and c is the speed of light in vacuum. This means that for frequencies close to that of the lowest order apertured-mode, the position dependent summation over emitter positions given from Eq. (1) can be expressed by

$$\frac{1}{\tau_{sp}(x_n,y_n,\omega_n)} \approx \frac{1}{\tau_{sp,o}} + \frac{1}{\tau_{sp,o}}\frac{4c^3/n^3}{\omega_o\omega_n w_o^2 L_z}\frac{(\gamma_n + \frac{\omega_o}{2Q})e^{-(x_n^2+y_n^2)/w_o^2}}{(\omega_n-\omega_o)^2 + (\gamma_n + \frac{\omega_o}{2Q})^2}, \quad (3)$$

where the first term on the right in Eq. (3) reflects the fact that the summation over the waveguide modes leads to a spontaneous emission rate nearly independent of the cavity, and the second term is due to the Purcell enhancement by the apertured modes. For large apertures the fraction of spontaneous emission actually captured by the apertured modes is negligible, while for small apertures the enhancement into the apertured modes can dominate the spontaneous lifetime. On resonance the second term in brackets becomes $\frac{(\lambda_o^3/n^3)Q}{\pi^2(\pi w_o^2 L_z)}$ which is very nearly the mathematical form of the enhancement described by Purcell for a single mode cavity.[30] Aside from constant factors, the cavity field intensity decay measured at frequency ω_n after a short-pulse excitation is given by

$$\frac{\omega_o}{Q}n_o(\omega_n,t) \propto \int_0^\infty d\rho_n\rho_n \frac{(\gamma_n + \frac{\omega_o}{2Q})e^{-\rho_n^2/w_o^2}e^{-\frac{t-t_o}{\tau_{sp}(\rho_n,\omega_n)}}}{(\omega_n-\omega_o)^2 + (\gamma_n + \frac{\omega_o}{2Q})^2}. \quad (4)$$

The QD heterostructure has been used to characterize the Purcell effect for a micron sized oxide-apertured microcavity consisting of an 18 pair GaAs/AlAs DBR, an AlGaAs λ/2 cavity spacer, and a single upper λ/4 GaAs layer. The cavity spacer is $Al_{0.97}Ga_{0.03}As$ at the center of which is grown a single $In_{0.50}Ga_{0.33}Al_{0.17}As$ QD active region, with 100 Å GaAs layers and 135 Å grading layers immediately adjacent on either side. The apertured-microcavity is fabricated by pattering 5 μm squares in photoresist and reactive ion etching to a depth of 2900 Å to form mesas exposing both $Al_{0.97}Ga_{0.03}As$ layers of the cavity spacer. Lateral oxidation[35] is performed at 450 °C, after which a 5 pair $ZnSe/MgF_2$ DBR is deposited to complete the microcavity. The microcavity is measured at 10 K, for which the QD ground state emission wavelength is ~9700 Å with a spectral width of ~600 Å. Time resolved measurements are performed using a mode-locked Ti-Sapphire laser beam (pulse rate reduced to 5 MHz) focused with a microscope objective to a 5 μm diameter spot on the microcavity. Photoluminescence is collected through the same objective and time resolved using a grating spectrometer and photon counting module with rise and fall times of ~300 ps. The InGaAlAs/GaAs QDs have radiative lifetimes of ~2 ns and are conveniently used to obtain high sensitivity in the photodetection. Larger InGaAs/GaAs QDs that emit at longer wavelengths ranging from ~1.2 to ~1.3 μm at room temperature show shorter radiative lifetimes ranging from ~400 psec to ~800 psec. The longer wavelengths and shorter spontaneous lifetimes make these larger QDs of interest for fast VCSELs. The longer lifetimes and shorter wavelengths of the InGaAlAs/GaAs QDs make them convenient for characterizing the cavity effects.

Figure 2 shows photoluminescence decay for wavelengths around resonance for the 1 μm diameter apertured-microcavity. Curves (a) and (c) show off-resonance decays taken at 9800 and 9950 Å, while curve (b) shows the on-resonance decay for the lowest order mode at 9860 Å. There is a factor of ~2.3 increase in the emission rate at 9860 Å as compared to off-resonance wavelengths. The fact that wavelengths both shorter and longer than resonance show similar slower decays is clearly indicates that the increased rate at resonance is due to the microcavity. Considering the first 2 ns of decay, the on-resonance lifetime [curve (b) in Fig. 1] is 0.9 ns, compared with the off-resonance lifetimes of 2.2 ns at 9950 Å and 1.9 ns at 9800 Å. The off-resonance lifetimes are close to the 2.1 ns lifetime measured for the epitaxial sample before processing. Therefore, the spontaneous emission rate is increased for the spatially averaged emitter positions by a factor of ~2.3, with little inhibition off-resonance. Figure 3 shows spontaneous spectra and decay rates plotted versus wavelength for the 1 μm apertured-microcavity. The inset shows the emission over a greater wavelength range. The longer wavelength spectral peaks at 9860 Å

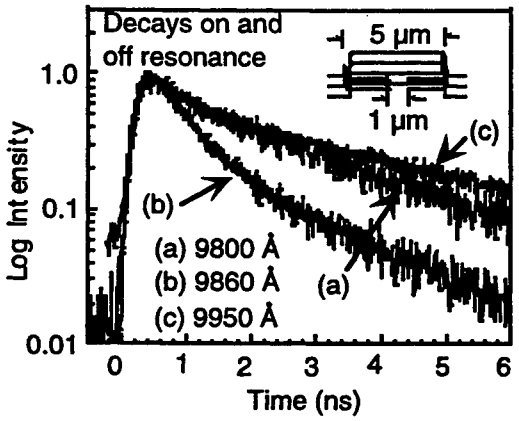


Fig. 2 Measured decay curves for a 1 μm apertured microcavity containing QD light emitters. Curve (a) for 9800 Å is for a wavelength shorter than resonance, curve (b) for 9860 Å is for a wavelength at resonance, and curve (c) for 9950 Å is for a wavelength longer than resonance.

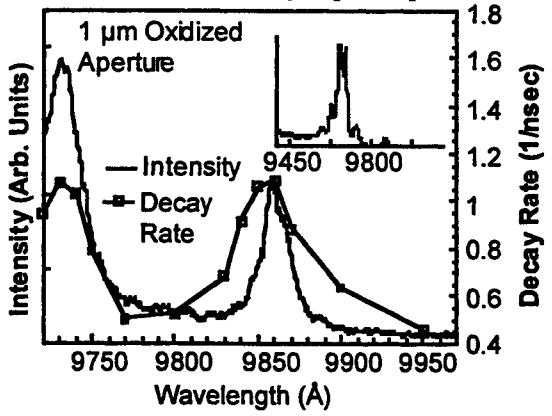


Fig. 3 Plot of measured decay rates versus wavelength and intensity versus wavelength for the 1 μm apertured microcavity.

and 9730 Å correspond to apertured-microcavity modes. The larger emission peaks starting at 9687 Å and visible to 9600 Å are due to emission from the oxide region (blue-shifted as compared to apertured modes) within the 5 μm mesa (see Fig. 1). Figure 3 shows that the spontaneous decay rate is enhanced at the lowest and next lowest order apertured-mode wavelengths of 9860 Å and 9730 Å. The spectral separation between the lowest and next lowest order modes depends on the mode area. The aperture-mode frequencies are approximately set by $\omega_m = (c/n)\sqrt{(\zeta_m / w_o)^2 + k_z^2}$, where the wavevector component k_z is

fixed by the mirrors. Bessel function modes satisfy the cylindrical symmetry of the cavity, so that we take ζ_o = 4.810 and ζ_1 = 7.664 for Bessel functions of the first kind. From Figs. 2 and 3, ω_0= 1.912x10^{15} rad/s and ω_1= 1.937x10^{15} rad/s, and the mode diameter is estimated to be $w_o = (c/n)\sqrt{(\zeta_1^2 - \zeta_0^2)/(\omega_1^2 - \omega_0^2)} \approx 1.8$ μm. The linewidth of the lowest order mode under continuous wave excitation is $\Delta\lambda$ = 15 Å which gives a $Q = \lambda_o / \Delta\lambda = \omega_o /(2\gamma_o) \approx 650$. Measurements to date suggest that electronic dephasing rates in the QDs can be <10^{11} s^{-1}, and we assume $\gamma_d < \frac{\omega_o}{Q}$ is satisfied.

From Eq. (4) we can also calculate the mode size required to achieve the spontaneous emission enhancement of 2.3 for the on-resonance wavelength. $\frac{\lambda_o^3 Q}{\pi^3 n^3 w_o^2 L_z} \approx 3.1$ provides a good match with the experimental data in Figs. 2 and 3. Given λ_o= 0.986 μm, Q= 650, an assumed value of n= 2.95 for the refractive index of the cavity spacer, and an effective cavity length of L_z= 0.75 μm, the estimated mode size is w_o= 0.7 μm. This Gaussian mode diameter of $2w_o$=1.4 μm is in rough agreement with the 1.8 μm diameter estimated from the spectral separation of the transverse modes assuming Bessel functions. From the calculations, the spontaneous emission rate enhancement for emitters placed at the center of the optical mode is then a factor of 3.1 compared to the factor of 2 for the emission rate averaged over emitter positions. This dependence on the emitter position can lead to spatial hole burning of the QD emitters at the mode center. It was noted in the introduction that diffraction loss can decrease the mode Q for very small optical modes. For the experimental data of Figs. 2 and 3 we expect that a larger enhancement of the emission rate can be obtained with an increased Q, but only slightly higher Q's have been achieved in similar cavities with more DBR pairs.

Although the scaling behavior of selectively oxidized VCSELs with planar quantum well active regions has been studied extensively and gives some information on the cavity loss rate for small apertures, the actual optical loss dependence on mode size can be difficult to extract from the experiment. Generally, the threshold current density increases as the oxide aperture is decreased and can prevent lasing for very small apertures. However, the increase in threshold current density is not only due to a reduction in the mode Q, electronic losses can also play a role. Even a relatively small diffusion coefficient of 10 cm^2/sec will lead to a diffusion length of ~1.7 μm. Ideally, if optical loss does not increase for the small aperture then neither should the threshold carrier density. On the other hand, the entire pumped area must be supplied through the small aperture, so that the current density through the aperture must increase to maintain the same carrier density. In addition, electronic losses often increase under and around the oxide aperture due to nonradiative recombination. These electronic effects due to diffusion make it difficult to clearly separate optical losses from electronic losses. However, a general trend is that the threshold current density is not so dependent on the aperture size as it is on the optical mode size. This is consistent with both the optical loss and electronic losses depending on the optical mode size, and for small apertures the larger optical mode both decreases the diffraction loss and increases the optical overlap with the injected carriers that suffer diffusion.

A more straightforward way to directly measure the optical loss dependence due to aperture and mode size is to the measure the spontaneous linewidth dependence on the optical mode size. Care must also be taken using this approach to avoid stimulated emission or absorption effects that narrow or broaden the "cold-cavity" linewidth. We have made such measurements for oxide-apertured VCSEL-type cavities that contain either QD or planar quantum well active regions. The measured results are shown in Fig. 4. The number of upper MgF/ZnSe DBR pairs are either five or seven. The QD cavities are undoped and measured in photoluminescence, while the cavities containing planar quantum wells are lightly doped and excited with current injection. An interesting trend is that the Q's for similar type cavities depend on the type of active region used. This suggests that absorption intrinsic to the planar quantum well active region may in fact limit the Q for high reflectivity. In Fig. 4 the measured Qs for the QD cavity with seven MgF/ZnSe DBR pairs exceeds 5000 for aperture sizes greater than 4 μm diameter, while it is only ~2000 for the planar quantum well active region. A similar reduction from ~3000 to ~1200 is found for five pairs. For each of the cavity types the Q drops significantly for aperture sizes smaller than 3 μm in diameter. The 3 μm diameter is also the size for which VCSEL threshold current densities begin to sharply rise with reducing aperture size. This shows that diffraction loss must indeed be considered for small apertures, but also shows that the VCSEL Q may contain subtleties that we do not yet understand.

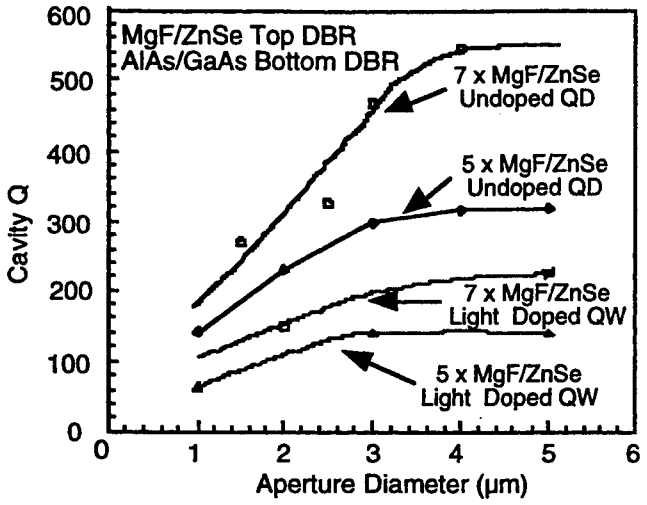


Fig. 4 Plot of Q versus aperture diameter for oxide-confined microcavities containing either QD or planar quantum well light emitters.

3. PURCELL EFFECT AND VCSEL CHARACTERISTICS

There has been a great deal of speculation about how laser thresholds will scale when the optical mode size is reduced due to microcavity confinement, with predictions that "zero-threshold" lasing will be obtained. However, the basis for some of these predictions are that the light versus current curve from the source can ideally be linear with 100% slope efficiency. This in itself is not the basis for a laser. A perfect "integrating sphere" would presumably give 100% spontaneous coupling efficiency from an emitter to a detector coupled through an emitting port, but this would not necessarily constitute a zero-threshold laser. The same is true for an ideal microcavity that can provide 100% spontaneous coupling efficiency to a detector, but with a Q that is insufficient to obtain lasing. Under extreme conditions the oxide-apertured microcavity that also uses strong QD electronic confinement such as shown in Fig. 1 can lead to nearly 100% spontaneous coupling efficiency. However, this microcavity emitter may not be a laser at all.

On the other hand, the increase in the spontaneous emission rate due to an oxide-apertured VCSEL-type cavity with a small enough volume to cause the Purcell effect can greatly improve the modulation rate of either a microcavity light emitting diode or a VCSEL. An interesting modification to the threshold also occurs, but the threshold current density is actually increased due to the increase in the spontaneous emission rate. Still the threshold current for very small optical modes and ideal electronic confinement such as provided by QDs can be very low, reaching sub-microamps. The very low threshold current is due to the small transparency current of the QD active region, while the optical gain can exceed that of a planar quantum well if inhomogeneous broadening is eliminated. The homogeneous linewidth from a QD emitter is measured to be ≤100 μeV at cryogenic temperatures, and is predicted to be ~2 to 3 meV or less at room temperature. While scaling the mode size has been studied for VCSELs, the impact on laser performance of changing the spontaneous emission rate by using a small mode volume has not yet been addressed. One reason is that experimentally, diffraction loss in part limits performance in single element VCSELs and is in itself an important scaling issue. Another reason though is that the VCSEL rate equations have not been put in a form that accurately accounts for the spontaneous lifetime dependence on the optical mode size.

A QD VCSEL microarray can be used to obtain milliwatt output powers typical of today's commercial VCSELs. Diffraction loss may also then be controlled through inter-element coupling. Below though we consider that the optical scattering losses are eliminated through strong lateral optical confinement, and consider the emission characteristics of a single element. From Fig. 1 we assume that the Gaussian optical mode has a free space wavelength λ_o = 1.3 μm, an effective mode area of πw_o^2, an effective length of L_z, and a photon loss rate of ω_o / Q. To maintain resonance L_z must depend on w_o with $L_z(w_o) = L_z(\infty) / \sqrt{1 - \lambda_o^2 / (\pi^2 n^2 w_o^2)}$, and we take $L_z(\infty)$= 0.59 μm and the cavity's average refractive index as n= 3.3. We hold ω_o / Q= 2x10^{11} sec^{-1} to isolate the effect of enhanced emission on laser performance.

The emitters are assumed to have a Gaussian density distribution with an effective area πw_{QD}^2 and a peak density n_{QD}= 1.5x10^{10} cm^{-2}. The spontaneous emission rate $1/\tau_{sp,o}$ is that which would occur without the cavity and for unity state

occupation by electrons and holes. $1/\tau_{sp,o} = 2.5\text{x}10^9$ sec^{-1} is consistent with the radiative lifetime recently measured for 1.3 μm QDs.[37] The average population inversion of QD states is given by $2\sqrt{x}-1$ with the upper state excitation given by x, where $0 \le x \le 1$. The $\sqrt{x}$ dependence assumes equal probabilities of electron and hole occupation in the QD ground state levels with random capture. We consider that ideally the QD emitters will be homogeneously broadened with a linewidth of $2\gamma_d$ due to dephasing. The QD VCSEL will likely require high Q, so that the adiabatic approximation $\gamma_d >> \omega_o/Q$ will be valid. For this case the QD spontaneous emission rate depends on the mode volume, dipole dephasing rate, and QD position within the mode, but not on the mode Q. Averaging over the emitter positions and retaining only leading terms gives the lasing mode's photon number rate equation as

$$\frac{dn}{dt} = -[\frac{\omega_o}{Q} - \frac{2\pi w_{QD}^2 w_o^2 n_{QD}}{w_{QD}^2 + w_o^2}\frac{\Omega^2}{\gamma_d}(2\sqrt{x}-1)]n + \frac{2\pi w_{QD}^2 w_o^2 n_{QD}}{w_{QD}^2 + w_o^2}\frac{\Omega^2}{\gamma_d}x \tag{5}$$

with the emitter population x driven by a current source $\frac{I}{q}$ given as

$$\frac{I}{q} = 2\pi w_{QD}^2 n_{QD}\frac{dx}{dt} + (\frac{2\pi w_{QD}^2 n_{QD}}{\tau_{sp,o}} + \frac{4\pi w_o^2 w_{QD}^2 n_{QD}}{w_{QD}^2 + w_o^2}\frac{\Omega^2}{\gamma_d})x + \frac{2\pi w_o^2 w_{QD}^2 n_{QD}}{w_{QD}^2 + w_o^2}\frac{\Omega^2}{\gamma_d}(2\sqrt{x}-1)n \tag{6}$$

and

$$\Omega = \sqrt{\frac{c\lambda_o^2}{2\pi n^3 \tau_{sp,o}\pi w_o^2 L_z}}. \tag{7}$$

We assume that the elements are designed so that $w_{QD}^2 = w_o^2$. The two spontaneous emission terms in Eq. (7) (in parenthesis and multiplying x) account separately for emission radiated into waveguide modes and into the aperture, just as above in Eq. (3). A modified Purcell effect (with $\gamma_d >> \omega_o/Q$) doubles the total spontaneous emission rate when $w_o = \frac{(\lambda_o/n)}{2\pi}\sqrt{\frac{(c/n)}{\gamma_d L_z}}$.

The threshold characteristics and the modulation response are calculated for a range of values of the mode diameter $2w_o$ and two different values of γ_d. We assume that the QD emitters are limited in linewidth only by homogeneous broadening, so that the spontaneous full-width at half-maximum energy spread is $2\hbar\gamma_d$ with $2\hbar\gamma_d$= 6.6 meV or $2\hbar\gamma_d$= 1.3 meV. Figure 5 shows the calculated thresholds for mode diameters ranging from 5 μm ≥ $2w_o$ ≥ 0.6 μm, and for the two different values of $2\hbar\gamma_d$. Threshold is taken as the pump rate needed to obtain a stimulated emission rate equal to the photon loss rate. Curves

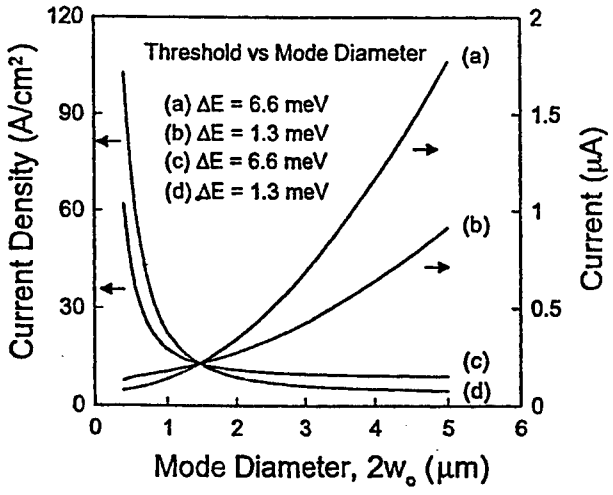


Fig. 5 Calculated threshold current and current density versus mode size for the QD VCSEL. The threshold currents versus mode diameter, $2w_o$, are shown in curve (a) for ΔE = 6.6 meV and (b) for ΔE = 1.3 meV, and the current density versus mode diameter is shown in curve (c) for ΔE = 6.6 meV, and (d) for ΔE = 1.3 meV. The other laser parameters held fixed are ω_o/Q = 2x10^{11} /sec, $\tau_{sp,o}$= 400 psec, n_{QD}=1.5x10^{10} /cm^2, λ_o = 1.3 μm, n = 3.3, L_z= 0.59 μm, and $w_{QD}^2/(w_o^2+w_{QD}^2)$ = 0.5.

(a) and (c) are the threshold current and current density for $2\hbar\gamma_d$ = 6.6 meV, and (b) and (d) are for $2\hbar\gamma_d$= 1.3 meV. For modes with $2w_o$ < 2 μm the threshold current density increases as the mode size decreases due to the increased spontaneous

emission rate. For our parameter values, the mode sizes for which $2w_o = \frac{(\lambda_o/n)}{\pi}\sqrt{\frac{(c/n)}{\gamma_d L_z}}$ are 0.7 μm for $2\hbar\gamma_d$= 6.6 meV, and 1.6 μm for $2\hbar\gamma_d$= 1.3 meV. Curves (c) and (d) show that for mode diameters $2w_o$ > 1.5 μm the threshold is less for $2\hbar\gamma_d$= 1.3 meV, but for $2w_o$< 1.5 μm the threshold is less for the larger linewidth of $2\hbar\gamma_d$= 6.6 meV. The transition occurs because the current density required for a fixed population inversion increases as the spontaneous emission rate increases. Since the spontaneous rate enhancement increases as the spontaneous linewidth decreases, the threshold for a given population inversion increases. For $2w_o$= 5 μm the calculated threshold current densities are 9.0 A/cm^2 for $2\hbar\gamma_d$= 1.3 meV, (c), and 4.7 A/cm^2 for $2\hbar\gamma_d$= 1.3 meV, (d). These are close to the experimental values for 1.3 μm QD edge-emitting lasers.[36]

The calculated small signal modulation responses from Eqs. (5) and (6) for different mode diameters at a bias level of twice threshold are shown in Fig. 6. The response curves in Fig. 3 are for $2w_o$ = 5 μm, 2 μm, or 0.6 μm. The spontaneous linewidth in (a) is $2\hbar\gamma_d$= 6.6 meV, and in (b) it is $2\hbar\gamma_d$= 1.3 meV. Both the mode volume and linewidth impact the VCSEL's modulation response through the cavity coupling Ω^2/γ_d. As $2w_o$ is reduced the relaxation oscillation peak is reduced in amplitude and the 3 dB bandwidth increases. Decreasing the spontaneous linewidth also decreases the relaxation oscillation amplitude and increases the 3 dB bandwidth.

Calculated bias-free pulse responses are shown in Fig. 7. The responses are for current pulses of unit step functions 500 psec long with amplitudes of either the steady-state transparency current value (x = 0.25) or twice the steady-state threshold value,

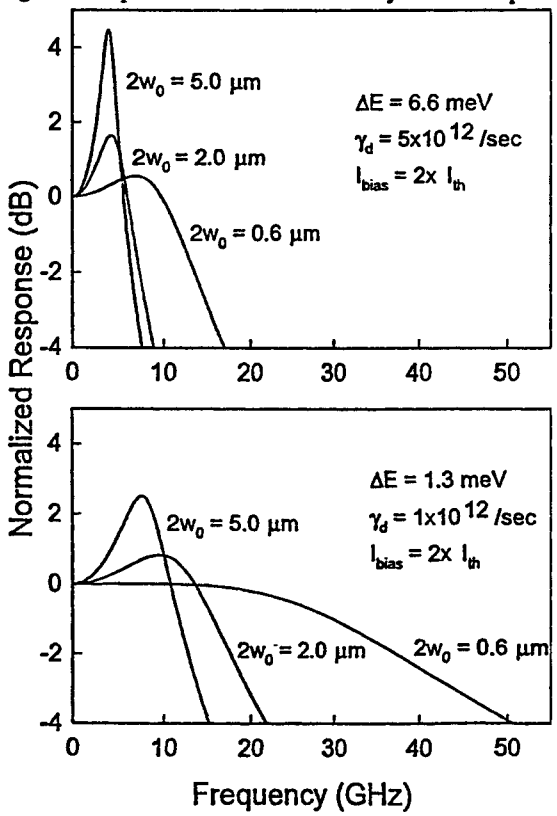


Fig. 6 Calculated small signal modulation responses for three different mode sizes of $2w_o$ = 5.0 μm, 2.0 μm, and 0.6 μm. (a) shows the responses for ΔE = 6.6 meV, and (b) shows the responses for ΔE = 1.3 meV. The other laser parameters are given in the Fig. 2 caption.

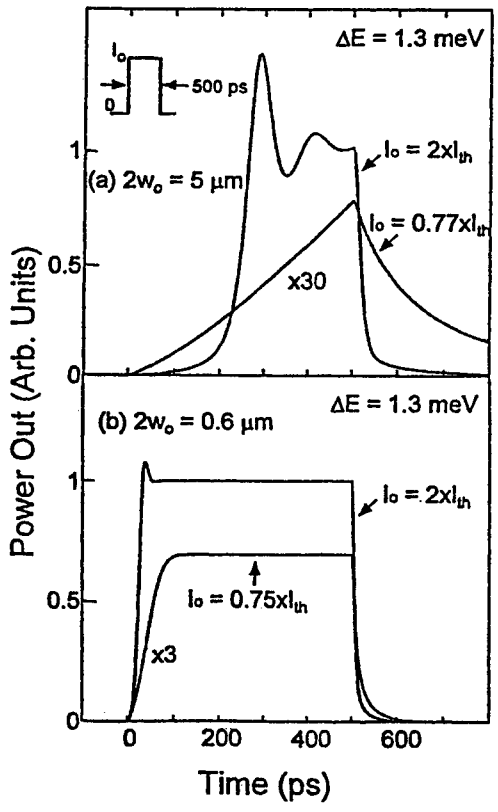


Fig. 7 Calculated bias-free responses for a 500 psec long current excitation pulse and ΔE = 1.3 meV, with the current amplitude equal to either transparency or twice the threshold current. (a) shows the responses for $2w_o$ = 5.0 μm, and (b) shows the responses for $2w_o$ = 0.6 μm. The other laser parameters are given in the Fig. 2 caption.

and two different mode sizes with $2w_o$ equal to either (a) 5 μm or (b) 0.6 μm. The spontaneous linewidths are held equal at $2\hbar\gamma_d$ = 1.3 meV to study only those effects due to mode size changes. From Fig. 7 (a) and (b) we see that for current pulses of ~0.7xI_{th} (transparency) the smaller mode size of $2w_o$ = 0.6 μm greatly improves the spontaneous pulse response, even allowing multi-gigabit data transmission. For current pulses of twice threshold the turn-on delay of ~250 psec for $2w_o$ = 5 μm in (a) is reduced to ~31 psec for $2w_o$ = 0.6 μm in (b). It's clear that the Purcell effect can have a major impact on the modulation response for microcavity QD VCSELs, especially around at bias levels around threshold.

4. SUMMARY

Scaling of the optical mode size has been examined for VCSELs. In order to take full advantage of the cavity effects strong lateral confinement must be achieved for both photons and electrons. The strong lateral confinement can be achieved using etched-pillars or oxide-apertures, but scattering and diffraction losses must be controlled to obtain high Q for micron-sized optical modes. Quantum dots are attractive for obtaining electronic confinement, but inhomogeneous broadening should be eliminated to take full advantage of cavity effects. For small sized QD VCSELs, the Purcell effect can greatly improve the modulation characteristics, perhaps even allowing high speed bias-free operation. Similar improvements in the modulation speed can be obtained using photonic bandgap cavities, but in many cases the microcavity QD VCSEL can be a direct replacement for today's commercial VCSELs.

ACKNOWLEDGMENTS

Various aspects of this work have been supported by the DARPA/Ultraphotonics program monitored by the AFOSR, and the DARPA supported Univ. of New Mexico OptoCenter.

REFERENCES

1. K. Iga, F. Koyama, and S. Kinoshita, "Surface emitting semiconductor lasers," *IEEE J. Quant. Electron.* **24**, pp. 1845-1855, 1988.
2. Y.H. Lee, B. Tell, K. Brown-Goebeler, J.L. Jewell, C.A. Burrus, J.M.Van Hove, "Characteristics of top-surface-emitting GaAs quantum-well lasers," *IEEE Phot. Tech. Lett.* **2**, pp. 686-688, 1990.
3. D.L. Huffaker, D.G. Deppe, K. Kumar, and T.J. Rogers, "Native-oxide defined ring contact for low threshold vertical-cavity lasers," *Appl. Phys. Lett.* **65**, pp. 97-99, 1994.
4. D.G. Deppe, D.L. Huffaker, T.-H. Oh, H. Deng, and Q. Deng, "Low-threshold vertical-cavity surface-emitting lasers based on oxide-confinement and high contrast distributed Bragg reflectors," *IEEE J. Quant. Electron.* **3**, pp. 893-904, 1997.
5. K.D. Choquette, R.P. Schneider, Jr., K.L. Lear, and K.M. Geib, "Low theshold voltage vertical-cavity lasers fabricated by selective oxidation," *Electron. Lett.* **30**, pp. 2043-2044, 1994.
6. Y. Hayashi, T. Mukaihara, N. Hatori, N. Ohnoki, A. Matsutani, F. Koyama, and K. Iga, "A record low threshold index-confined InGaAs/GaAlAs vertical-cavity surface-emitting laser with a native oxide confinement structure," *Electron. Lett.* **31**, pp. 560-561, 1995.
7. G.M. Yang, M.H. MacDougal, and P.D. Dapkus, "Ultralow threshold current vertical-cavity surface-emitting lasers obtained with selective oxidation," *Electron. Lett.* **31**, pp. 886-887, 1995.
8. B.J. Thibeault, E.R. Hegblom, P.D. Floyd, R. Naone, Y. Akulova, and L.A. Coldren, "Reduced optical scattering loss in vertical-cavity lasers using a thin (300Å) oxide aperture," *IEEE Phot. Tech. Lett.* **8**, pp. 593-595, 1996.
9. E.R. Hegblom, D.I. Babic, B.J. Thibeault, and L.A. Coldren, "Scattering losses from dielectric apertures in vertical-cavity lasers," *IEEE J. Sel. Top. Quant. Electron.* **3**, pp. 379-389, 1997.
10. K.L. Lear, K.D. Choquette, R.P. Schneider, Jr., S.P. Kilcoyne, and K.M. Geib, "Selectively oxidized vertical cavity surface emitting lasers with 50% power conversion efficiency," *Electron. Lett.*, vol. 31, 208 (1995).
11. R. Jager, M. Grabherr, C. Jung, R. Michalzik, G. Reiner, B. Wiegl, and K. Ebeling, "57 % wallplug efficiency oxide-confined 850 nm wavelength GaAs VCSELs," *Electron. Lett.* **33**, pp. 330-331, 1997.
12. B. Weigl, M. Grabherr, G. Reiner, and K.J. Ebeling, "High efficiency selectively oxidized MBE grown vertical-cavity surface-emitting lasers," *Electron. Lett.* **32**, pp. 557-558, 1996.
13. K.L. Lear, A. Mar, K.D. Choquette, S.P. Kilcoyne, R.P. Schneider, Jr., and K.M Geib, "High-frequency modulation of oxide-confined vertical-cavity surface-emitting laser," *Electron. Lett.* **32**, pp. 457-458, 1996.
14. B.J. Thibeault, K. Bertilsson, E.R. Hegblom, E. Strzelecka, P.D. Floyd, R. Naone, and L.A. Coldren, "High-speed characteristics of low-optical loss oxide-apertured vertical-cavity lasers," *IEEE Phot. Tech. Lett.* **9**, pp. 11-13, 1997.
15. B. Weigl, M. Grabherr, R. Michalzik, G. Reiner, and K.J. Ebeling, "High power single-mode selectively oxidized vertical-cavity surface-emitting lasers," *IEEE Phot. Tech. Lett.* **8**, pp. 971-973, 1996.

16. M. Grabherr, R. Hager, R. Michalzik, B. Weigl, G. Reiner, and K.J. Ebeling, "Efficiency single-mode oxide-confined GaAs VCSEL's emitting in the 850 nm wavelength regime," *IEEE Phot. Tech. Lett.* **9**, pp. 1304-1306, 1997.
17. G. R. Hadley, "Effective index model for vertical-cavity surface-emitting lasers", *Opt. Lett.* **20**, pp. 1483-1485, 1995.
18. D.G. Deppe, T.-H. Oh, and D.L. Huffaker, "Eigenmode confinement in the dielectrically apertured Fabry-Perot microcavity," *IEEE Phot. Tech. Lett.* **9**, 713-715, 1997.
19. W.B. Veldkamp, "Wireless focal planes 'On the road to amacronic sensors'," *IEEE J. Quant. Electron.* **29**, pp. 801 - 813, 1993.
20. N. Kirkstaedter, N. Ledentsov, M. Grundmann, D. Bimberg, V. Ustinov, S. Ruvimov, M. Maximov, P. Kop'ev, Zh. Alferov, "Low threshold, large To injection laser emission from (InGa)As quantum dots," *Electron. Lett.* **30**, pp. 1416-1417, 1994.
21. K. Kamath, P. Bhattacharya, T. Sosnowski, T. Norris, and J. Phillips, "Room temperature operation of $In_{0.4}Ga_{0.6}As/GaAs$ self-organized quantum dot lasers," *Electron. Lett.* **32**, pp. 1374-1375, 1996.
22. H. Shoji, Y. Nakata, K. Mukai, Y. Sugiyama, M. Suagawara, N. Yokoyama, and H. Ishikawa, "Room temperature CW operation at the ground state of self-formed quantum dot lasers with multi-stacked dot layer," *Electron. Lett.* **32**, pp. 2023-2024, 1996.
23. R. Mirin, A. Gossard, and J. Bowers, "Room temperature lasing from InGaAs quantum dots," *Electron. Lett.* **32**, pp. 1732-1733, 1996.
24. Q. Xie, A. Kalburge, P. Chen, and A. Madhukar, "Observation of lasing from vertically self-organized InAs three-dimensional island quantum boxes on GaAs (001)," *IEEE Phot. Tech. Lett.* **8**, pp. 965-967, 1996.
25. H. Saito, K. Nishi, I. Ogura, S. Sugov, and Y. Sugimoto, "Room-temperature lasing operation of a quantum-dot vertical-cavity surface-emitting laser," *Appl. Phys. Lett.* **69**, pp. 3140-3142, 1996.
26. D.L. Huffaker, O. Baklenov, L.A. Graham, B.G. Streetman, and D.G. Deppe, "Quantum dot vertical-cavity surface-emitting laser with a dielectric aperture," *Appl. Phys. Lett.* **70**, pp. 2356-2358, 1997.
27. J.A. Lott, N.N. Ledentsov, V.M. Ustinov, A.Yu. Egorov, A.E. Zhukov, P.S. Kop'ev, Zh.I. Alferov, and D. Bimberg, "Vertical cavity lasers based on vertically coupled quantum dots," *Electron. Lett.* **33**, pp. 1150-1151, 1997.
28. D.L. Huffaker, H. Deng, and D.G. Deppe, "1.15 μm Wavelength Oxide-Confined Quantum Dot Vertical-Cavity Surface-Emitting Laser," *IEEE Phot. Tech. Lett.* **10**, pp. 185-187, 1998.
29. Z. Zou, D.L. Huffaker, S. Csutak, and D.G. Deppe, "Ground State Lasing From a Quantum Dot Oxide-Confined Vertical-Cavity Surface-Emitting Laser," *Appl. Phys. Lett.* **75**, pp. 22-25, 1999.
30. E.M. Purcell, "Spontaneous emission probabilities at radio frequencies," *Phys. Rev.* **69**, p. 681, 1946.
31. K.H. Drexhage, "Interaction of light with monomolecular dye layers," in *Progress in Optics*, edited by E. Wolf (North-Holland, Amsterdam, 1974), Vol. XII, Chap. IV.
32. Q. Deng and D.G. Deppe, "Spontaneous lifetime change in a dielectrically-apertured Fabry-Perot microcavity," *Optics Express* **2**, pp. 157-162, 1998.
33. L.A. Graham, D.L. Huffaker, S.M. Csutak, Q. Deng, and D.G. Deppe, "Spontaneous lifetime control of quantum dot emitters in apertured microcavities," *J. Appl. Phys.* **85**, pp. 3383-3385, 1999.
34. L.A. Graham, D.L. Huffaker, and D.G. Deppe, "Spontaneous lifetime control in a native-oxide-apertured-microcavity," *Appl. Phys. Lett.* **74**, pp. 2408-2410, 1999.
35. J.M. Dallesasse, N. Holonyak, Jr., A.R. Sugg, T.A. Richard, and N. El-Zein, "Hydrolization oxidation of AlGaAs-AlAs-GaAs quantum well heterostructures," *Appl. Phys. Lett.* **57**, pp. 2844-2846, 1990.
36. L. Zhang, T.F. Boggess, D.G. Deppe, D.L. Huffaker, O.B. Shchekin, and C. Cao, "Dynamic response of 1.3 μm wavelength InGaAs/GaAs quantum dots," unpublished.
37. G. Park, D.L. Huffaker, Z. Zou, O.B. Shchekin, and D.G. Deppe, "Temperature Dependence of Lasing Characteristics for Long-Wavelength (1.3 μm) Quantum Dot Lasers," IEEE Phot. Tech. Lett. **11**, pp. 301-303, 1999.

Plenary Session II

Optical Nonlinearities in Semiconductors *

S. W. Koch and T. Meier

Department of Physics and Material Sciences Center,
Philipps University, Renthof 5, D-35032 Marburg, Germany

ABSTRACT

A microscopic many-body theory for the nonlinear optical response of semiconductors is reviewed. The importance of Coulomb interaction induced carrier correlations is demonstrated in excitonic pump-probe spectra. The influence of excitonic and biexcitonic contributions to coherent pump-induced absorption changes at the exciton frequency are discussed. Absorption changes induced by incoherent exciton and unbound electron-hole populations are studied.

Keywords: optical nonlinearities, many-body theory, Coulomb correlations, excitonic pump-probe

1. INTRODUCTION

The study of optical nonlinearities in semiconductors and semiconductor heterostructures is a field of active theoretical and experimental research. Microscopic theoretical descriptions of optical nonlinearities in semiconductors generally have to deal with the different relevant quasiparticles and their interactions. In this paper we mainly focus on the important influence of the many-body Coulomb interaction on near-bandgap optical nonlinearities in semiconductors.

One subject of particular interest that has been investigated intensively already in the 1980s and is still receiving considerable attention are nonlinear absorption and refractive index changes induced by the presence an electron-hole plasma. In the theoretical description of such nonlinearities many-body effects were included microscopically on a non-perturbative level.[1–5] This approach gives a good basic understanding of the weakening (bleaching) of excitonic absorption with increasing plasma density, the appearance of optical gain (negative absorption), as well as the corresponding dispersive nonlinearities caused by the presence of uncorrelated electron-hole pairs. These nonlinear population-induced absorption changes can be attributed to the effects of Pauli blocking (also called phase-space filling), as well as genuine many-body effects such as band-gap renormalization (electronic self energies) and screening of the Coulomb interaction potential.

Typically in optically excited semiconductors the quasi-particle scattering processes which govern the dephasing of optical excitations can be characterized by typical interaction times in the picosecond to sub-picosecond range. Hence, investigations of the coherent nonlinear optical response of semiconductors, which requires laser pulses shorter than the dephasing time, became generally possible only after the development of femtosecond laser sources in the late 1980s.

In the early ultrafast coherent pump and probe experiments particular attention was paid to the dynamical optical Stark effect of excitons in semiconductors.[6–8] To investigate this Stark effect, a strong pump pulse is tuned below the lowest excitonic transition and the resulting absorption changes are monitored by a weak probe pulse.[5] Analogous to atomic systems the experiments revealed that for large detuning the pump pulse leads to a blueshift of the excitonic resonance.[6–8] In a two-band semiconductor model and for relatively large detuning below the exciton, the experimentally observed blueshift could be well described on the basis of the semiconductor Bloch equations (SBE) in the time-dependent Hartree-Fock (TDHF) approximation[4,5,9,10]

Generally, the THDF-SBE are surprisingly successful in the analysis of coherent nonlinear optical effects in semiconductors and semiconductor nanostructures. With improved sample quality and experimental techniques, however, it became evident that many-body Coulomb correlations, i.e. many-body effects beyond the TDHF approximation, can lead to characteristic signatures in semiconductor nonlinearities that cannot be described with the TDHF-SBE.

Further author information: (Send correspondence to S. W. Koch)
S. W. Koch: E-mail: Stephan.W.Koch@physik.uni-marburg.de
T. Meier: E-mail: Torsten.Meier@physik.uni-marburg.de

*Also published in *Proc. of SPIE* Vols. 3896, 3898, and 3899

Part of the SPIE Conference on Advanced Photonic Sensors and Applications
Singapore • November–December 1999
SPIE Vol. 3897 • 0277-786X/99/$10.00

For example, the relevance of such correlations was clearly shown by analyzing the dependence of the measured signatures on the polarization directions of the incident pulses. Such experiments provided evidence for excitation-induced dephasing processes[11] and two-exciton (biexciton) contributions.[12–16]

To calculate the nonlinear optical response of semiconductors including two-exciton resonances it is necessary to include not only two-point functions, but also higher-order correlation functions in the theoretical description. For arbitrary excitation conditions this would mean that due to the many-body Coulomb interaction an infinite hierarchy of equations of motion has to be solved. One way to restrict this hierarchy to a finite number of correlation functions is to limit the theoretical analysis to a finite order in the optical field. Hence, only a finite number of interactions between the semiconductor and the light field takes place and thus only a finite number of electron-hole pairs is generated.[17,18]

In order to describe four-wave mixing and pump-probe experiments one needs to consider at least processes up to the third-order in the light field ($\chi^{(3)}$). In the coherent limit, where interaction with other quasiparticles, for example phonons, are neglected, it is sufficient to consider just two quantities (density matrices). These two are the single-exciton amplitude p and the two-exciton amplitude B, which describes the evolution of two interacting electron-hole pairs.[17,18] Nevertheless, solutions of the full coherent $\chi^{(3)}$-equations are computationally quite demanding since effectively a quantum mechanical four-body problem has to be solved. Thus, one often restricts the numerical analysis to one-dimensional model systems,[19–22] where however the parameters are chosen to represent realistic quantum-well configurations. Fully two-dimensional calculations are possible,[15] however, they require rather involved computational analysis and cannot conveniently be used for extended model studies in which dependencies of the results on large numbers of system parameters are investigated. In the coherent $\chi^{(3)}$-limit for excitation at and below the exciton resonance, the one- and two-dimensional results predict qualitatively similar absorption changes at the exciton which are in good agreement with experiments.[15]

Beyond the coherent limit additional occupation-type dynamical variables need to be considered. Up to third-order one then has to deal with four relevant variables, i.e. besides the coherences p and B also the pair occupation N and the exciton to two-exciton transition Z contribute.[19] Since Z is a six point function, solutions of the coupled equations for these four density matrices are numerically even more demanding, such that up to now only few such calculations have been reported. In Ref. 20 it has been shown that in the presence of low-density incoherent occupations N also the six point function Z is significant in order to describe exciton bleaching and to account for transitions from incoherent single- to two-exciton states.

In this paper we review some of our recent results made towards the understanding of nonlinear optical processes in semiconductors including many-body correlations, see Refs. 15 and 20–24. After a brief description of the theoretical approach in Section 2, we focus on the discussion of numerical results on excitonic absorption changes in different situations. Resonant and off-resonant excitation are discussed in Sections 3 and 4, respectively. In both cases we find distinct signatures induced by many-body correlations. In Section 5 absorption changes induced by incoherent populations are analyzed and a brief summary is presented in Section 6.

2. THEORY

To theoretically model correlation effects relevant in coherent two pulse experiments such as pump-probe and four-wave mixing spectroscopy, see Fig. 1(a), we compute the nonlinear optical response in the coherent $\chi^{(3)}$-limit.[14,17,18] This is done by numerically solving the relevant equations of motion for the single-exciton amplitude p and the correlated part of the two-exciton amplitude $\bar{B}$, see Ref. 21. Omitting the various sums and indices the basic structure of the equations of motion is given by

$$\frac{\partial}{\partial t}p = i\left[\omega_p p + \mu E(1-2p^*p) + Vp^*pp + Vp^*\bar{B}\right], \tag{1}$$

$$\frac{\partial}{\partial t}\bar{B} = i\left[\omega_B \bar{B} + Vpp\right]. \tag{2}$$

Here, ω_p and ω_B are energies of a single- and a two-exciton state, respectively, μ is the transition dipole, E the external laser field, and V the Coulomb interaction potential. In Refs. 15 and 21 we have numerically solved the microscopic versions of Eqs. (1) and (2), both for a fully two-dimensional model of the quantum-well structure[25] and for a simplified, quasi one-dimensional system. As shown in Ref. 15 the spectra obtained for both models are in good

qualitative agreement and explain correlation-induced signatures in the experimental differential absorption of high quality quantum wells.

Within our model we include the two spin-degenerate electron and heavy-hole bands using the circularly-polarized dipole matrix elements describing the optical interband transitions.[13,16] Besides the heavy holes also light-hole transitions can be incorporated.[23] The monopole-monopole Coulomb interactions between particles at sites i and j are given by V_{ij}. In our one-dimensional model,[15,21] the spatial variation is given by a regularized potential: $V_{ij} = U_0 d/(|i-j|d + a_0)$, where d is the distance between the sites and U_0 and a_0 are parameters characterizing the strength of the interaction and the spatial variation taken as $U_0 = 8meV$ and $a_0/d = 0.5$, respectively. In our numerical study we use a tight-binding model with electron and heavy-hole couplings of $J_e = 8meV$ and $J_{hh} = 0.8meV$ between neighboring sites. These system parameters result in an exciton binding energy of $8meV$ and a biexciton binding energy of $1.4meV$.[15]

The theory allows us to identify three types of optical nonlinearities.[15,21] The total differential absorption signal is the sum of a *Pauli-blocking term* ($\propto \mu E p^* p$ in Eq. (1)) and *Coulomb-induced many-body nonlinearities*, which can further be separated into a first-order term ($\propto V p^* p p$ in Eq. (1)) and higher-order correlations ($\propto V p^* B$ in Eq. (1)), see Refs. 15 and 21. Thus the total differential absorption $\delta\alpha$ can be written as the sum of three contributions[15,21]

$$\delta\alpha(\omega) = \delta\alpha_{pb}(\omega) + \delta\alpha_{CI,1st}(\omega) + \delta\alpha_{CI,corr}(\omega)\,, \tag{3}$$

where pb denotes the optical nonlinearity induced by Pauli-blocking. The terms denoted with CI are Coulomb-interaction-induced nonlinearities. $CI, 1st$ is the first-order (Hartree-Fock) term, and $CI, corr$ the higher-order correlation contribution. The lineshape is different for each of the three terms in Eq. (3) which therefore introduce distinct signatures in the differential absorption spectra.[15,21]

Besides the coherent response, we also analyze the importance of Coulomb correlations in situations where incoherent occupations are present. We focus on the fully incoherent situation and analyze absorption changes induced by incoherent occupations $\bar{N}$. Omitting the various sums and indices, which can be found in Ref. 20, the basic structure of the equations relevant in this case is given by

$$\frac{\partial}{\partial t} p = i\left[\omega_p p + \mu E(1-2\bar{N}) + Vp\bar{N} + V\bar{Z}\right], \tag{4}$$

$$\frac{\partial}{\partial t} \bar{Z} = i\left[(\omega_B - \omega_p)\bar{Z} + Vp\bar{N}\right]. \tag{5}$$

As before, we can also here distinguish between the different types of optical nonlinearities. Therefore the absorption changes induced by incoherent occupations $\bar{N}$ are given by the sum of a *Pauli-blocking term* ($\propto \mu E\bar{N}$ in Eq. (4)) and *Coulomb-induced many-body nonlinearities*, which can again further be separated into a first-order term ($\propto Vp\bar{N}$ in Eq. (4)) and higher-order correlations ($\propto V\bar{Z}$ in Eq. (4)), see Ref. 20. This allows us to write the total differential absorption $\delta\alpha$ as the sum over the three contributions as given by Eq. (3).

3. RESONANT EXCITATION

The differential absorption spectrum for resonant excitation at the exciton resonance with co-circularly polarized pump and probe pulses ($\sigma^+\sigma^+$) with a time delay of $\tau = 2ps$ is shown in Fig. 1(b). The differential absorption $\delta\alpha(\omega)$ is strictly negative in the vicinity of the exciton resonance corresponding to a pump-pulse-induced bleaching of the exciton resonance. Positive contributions to $\delta\alpha(\omega)$ appearing well above the exciton are related to excited state absorption induced by exciting unbound two-exciton states. Besides the total signal, Fig. 1(b) also displays the three optical nonlinearities according to Eq. (3) separately. It is shown that $\delta\alpha_{pb}$ is weak and corresponds purely to a bleaching of the exciton resonance. $\delta\alpha_{CI,1st}$ is very strong and is antisymmetric around the exciton resonance. The dispersive shape of the differential absorption corresponds to a blueshift of the exciton. $\delta\alpha_{CI,corr}$ is also mainly dispersive around the exciton resonance, but with opposite sign compared to $\delta\alpha_{CI,1st}$, i.e. this term yields a redshift. Besides contributions with resonances at the exciton energy, $\delta\alpha_{CI,corr}$ also includes terms having resonances at the transition energies to unbound two-exciton states. Therefore it is not completely antisymmetric around the exciton resonance. Including all three contributions to obtain the total signal via Eq. (3), strong cancellations occur between $\delta\alpha_{CI,1st}$ and $\delta\alpha_{CI,corr}$ and the resulting differential absorption shows a predominantly absorptive spectral shape around the exciton resonance. It can be inferred from Fig. 1(b) that the bleaching at the exciton resonance is dominated by Coulomb-induced nonlinearities $\delta\alpha_{CI,1st} + \delta\alpha_{CI,corr}$ and only weakly enhanced by $\delta\alpha_{pb}$.

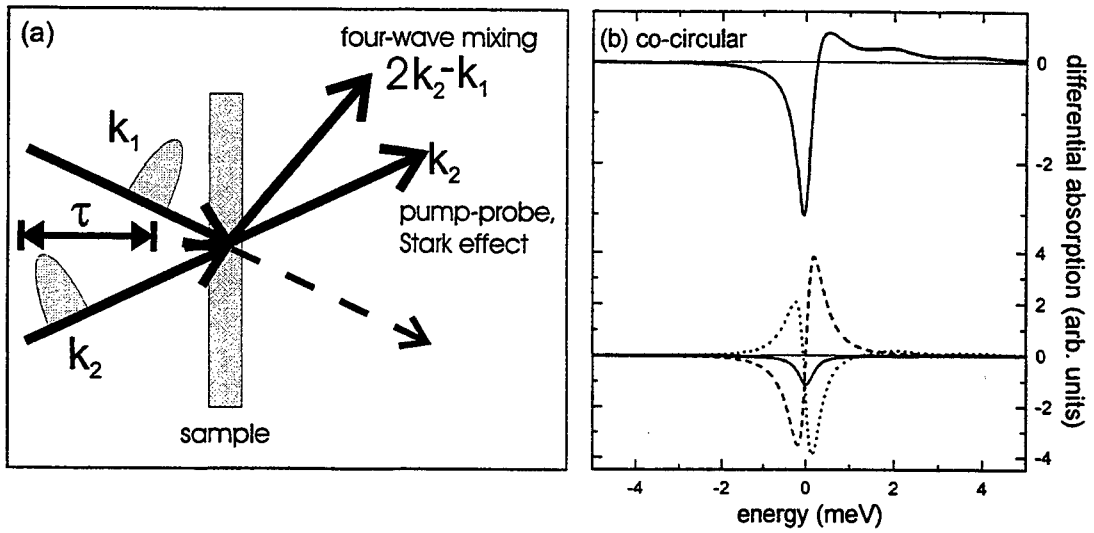


Figure 1. (a) Schematical illustration of nonlinear optical four-wave mixing and pump-probe experiments performed by excitation with two optical laser pulses having a time delay τ. (b) Total differential absorption spectrum $\delta\alpha(\omega)$ for resonant excitation at the exciton resonance and a time delay of $\tau = 2ps$ for co-circularly polarized pump (σ^+) and probe (σ^+) pulses. Also displayed in the lower part are $\delta\alpha_{pb}$ (solid), $\delta\alpha_{CI,1st}$ (dashed), and $\delta\alpha_{CI,corr}$ (dotted). The zero of the energy scale is chosen to coincide with the 1s-exciton resonance. The pump pulse has a duration of $1.18ps$ (full width at half maximum of pulse intensity). For further parameters see text.

Differential absorption spectra for time delays of $\tau = 2ps$, $0ps$, and $-2ps$ are displayed in Fig. 2 for both co-circularly ($\sigma^+\sigma^+$) and opposite-circularly ($\sigma^+\sigma^-$) polarized pump and probe pulses. As is shown in Ref. 21, for $\sigma^+\sigma^+$ excitation similar compensations among the Coulombic nonlinearities as displayed in Fig. 1(b) arise for all time delays. Within the coherent limit the signal amplitude is strongest for small delays and decreases with both increasing positive and negative delay, see Fig. 2. Apart from the change in strength, for $\sigma^+\sigma^+$ excitation the lineshapes are very similar for 0 and $2ps$ delay. The main visible difference between these two cases is that for $\tau = 0ps$ there is weak induced absorption appearing energetically below the exciton. This feature is reminiscent of the coherent spectral oscillations,[4] which dominate the spectra for negative delays, see dotted line in Fig. 2(a) for $\tau = -2ps$. For negative time delays the probe arrives before the pump pulse. With calculations performed in the limit of ultrashort pulses is can be shown that in this case the differential polarization $\delta P(\omega,\tau)$ which determines the absorption changes is multiplied by the phase factor $\exp(i(\omega_p - \omega)\tau)$, where ω_p is the exciton frequency. Since the differential absorption $\delta\alpha(\omega,\tau)$ is determined by the imaginary part of $\delta P(\omega,\tau)$ the phase factor introduces $\sin((\omega_p - \omega)\tau)/(\omega_p - \omega)$-like spectral oscillations of $\delta\alpha(\omega,\tau)$ around ω_p.[4,21]

If only heavy-hole transitions are important, $\sigma^+\sigma^-$ polarized excitation makes it possible to analyze pure correlations effects.[15,21,23] This is due to the fact, that for this geometry $\delta\alpha_{pb}$ vanishes, and also $\delta\alpha_{CI,1st}$ vanishes as long as the system is homogeneous, since neither of these contributions introduces any coupling between the subspaces of different spin states.[21,22] Differential absorption spectra for $\tau = 2ps$, $0ps$, and $-2ps$ and $\sigma^+\sigma^-$ excitation are displayed in Fig. 2(b). For zero and positive time delays also in this configuration the spectra show bleaching at and some excited state absorption above the exciton resonance. Since with opposite-circularly polarized pump and probe pulses also a bound biexciton is excited, we see in Fig. 2(b) additional strong induced absorption below the exciton resonance. The energetic difference between the maximum of this induced absorption and the bleaching is just the biexciton binding energy. For negative delays also in this polarization configuration which is entirely induced by Coulombic correlations the spectra are characterized by coherent spectral oscillations.[21] The presence of the oscillations in the correlation terms is in agreement with simplified analytical calculations presented in Ref. 21.

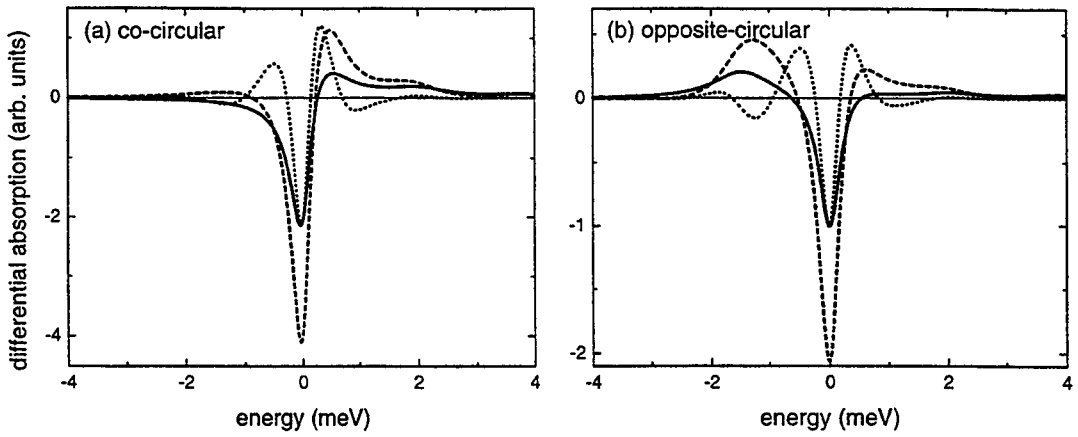


Figure 2. Total differential absorption spectra $\delta\alpha(\omega)$ for resonant excitation at the exciton resonance and a time delay of $\tau = 2ps$ (solid), $0ps$ (dashed), and $-2ps$ (dotted). (a) for co-circularly polarized ($\sigma^+\sigma^+$) and (b) for opposite-circularly polarized ($\sigma^+\sigma^-$) pump and probe pulses.

4. OFF-RESONANT EXCITATION

Here we investigate the influence of correlations on the excitonic optical Stark effect by tuning the pump pulse below the exciton resonance. Fig. 3 shows the resulting differential absorption spectra for co-circularly polarized excitation and various detunings of the central pump-pulse frequency below the exciton energy. The higher detunings ($5meV$ and $7.5meV$) exceed both the spectral width of the pump pulse and the homogeneous width of the exciton resonance. The solid line gives the full $\delta\alpha$, whereas the dashed line gives the result within TDHF where correlations are neglected, i.e. $\delta\alpha_{HF} = \delta\alpha_{pb} + \delta\alpha_{CI,1st}$. With increasing detuning the amplitude of the signal decreases and the bleaching at the exciton resonance develops into a dispersive shape corresponding to a blue-shift.[4,21] Fig. 3 directly shows that the importance of the carrier correlations diminishes rapidly with increasing detuning. We thus conclude that the time-dependent Hartree-Fock approximation provides a good description for this polarization configuration as long as off-resonant excitation is considered.

An interesting effect induced by the dynamics of Coulomb correlations is found by analyzing for a fixed detuning the polarization-dependence of the Stark effect. We choose the detuning of the pump pulse as $4.5\,meV$ below the exciton resonance and investigate the absorption changes for the various polarization configurations. Fig. 4(a)-(d) displays the theoretical results for $\sigma^+\sigma^+$, xx, xy, and $\sigma^+\sigma^-$ excitation. Whereas for the three cases (a), (b), and (c) the differential absorption corresponds to a blueshift, for $\sigma^+\sigma^-$ clearly a redshift appears. Since for $\sigma^+\sigma^-$ excitation both the Pauli blocking and the first-order Coulomb-induced nonlinearities, i.e. the Hartree-Fock contribution, vanish identically, the redshift of the differential absorption is purely induced by Coulomb correlations.[15]

The physical origin of the redshift can be analyzed in more detail by looking at the individual contributions to the signal, as displayed on the right panels in Fig. 4. For $\sigma^+\sigma^+$, xx, and xy polarization both the Pauli blocking and the first-order Coulomb terms always induce a blueshift, whereas the Coulomb correlations always correspond to a redshift, see Fig. 4(a)-(c). The fact that also for $\sigma^+\sigma^+$ excitation, where no bound biexcitons are excited, the correlation term alone corresponds to a redshift indicates that the correlation-induced reversal of the direction of the shift is not directly related to the existence of a *bound* biexciton. For $\sigma^+\sigma^+$, xx, and xy polarization, however, the influence of correlations is rather small and its redshift is always overcompensated by the Pauli blocking and the first-order Coulomb terms resulting in a net blueshift.[15]

To substantiate the claim that the existence of a bound biexciton is not necessary to obtain a correlation-induced redshift, we have performed additional calculations of the differential absorption spectra for $\sigma^+\sigma^-$ configuration. As discussed in Ref. 15, the bound biexciton contribution has been eliminated by artificially dropping the six terms containing the attractive and repulsive Coulomb terms between the two electrons and two holes appearing in the homogeneous part of the equation of motion for the two-exciton amplitude $\bar{B}$. In this case, also for $\sigma^+\sigma^-$ excitation no bound biexcitons exist. The approximation of neglecting the Coulomb terms that lead to the formation of bound

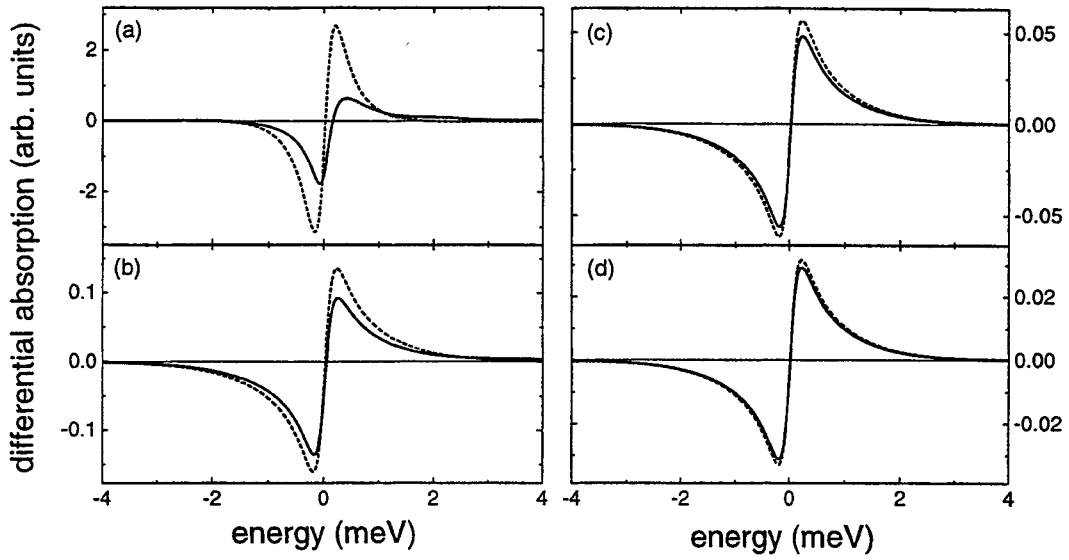


Figure 3. Differential absorption spectra for various detunings of the pump pulse below the exciton resonance with co-circularly polarized pump and probe pulses for time delay $\tau = 0ps$. The central frequency of the pump pulse is chosen (a) $1meV$, (b) $3meV$, (c) $5meV$, and (d) $7.5meV$ below the energy of the $1s$-exciton. The solid line gives the full $\delta\alpha$, and the dashed line the result of a Hartree-Fock calculation ($\delta\alpha_{pb} + \delta\alpha_{CI,1st}$) neglecting correlations ($\delta\alpha_{CI,corr}$).

two-exciton states is identical to the second Born-approximation (SBA), since only terms up to second-order in the Coulomb interaction are retained in the signal.[25,26] By solving a dynamic equation for $\bar{B}$, Coulomb memory effects are still included on this level.[15] The resulting spectrum displayed in Ref. 15 shows that in the SBA the signal amplitude is somewhat reduced, however, the redshift clearly persists. Further calculations performed for the SBA in Markov approximation did not reproduce the redshift. This allows us to conclude that it is clearly the *dynamics* of Coulomb correlations that is responsible for the presence of a redshift.[15] Our numerical results on the polarization-dependence of the optical Stark effect are in good agreement with experiments performed on high-quality InGaAs quantum wells with spectrally very narrow exciton linewidth, see Ref. 15.

In order to show that the redshift depends crucially on the energetic separation between heavy- and light-hole excitons and on the detuning we now extend our model to to include both types of valence bands using the typical selection rules and the three to one ratio of the oscillator strengths.[23] The in-plane dispersion of the valence band structure in quantum wells is modeled by considering heavy- and light-hole masses according to the Luttinger-Hamiltonian. For GaAs parameters we get $m_{hh} = m_0/(\gamma_1 + \gamma_2) = 0.112m_0$ and $m_{lh} = m_0/(\gamma_1 - \gamma_2) = 0.211m_0$. Further band-mixing effects are neglected here for simplicity. For the conduction band electrons we use $m_e = 0.0665m_0$ and take the same coupling of $J_e = 8meV$ as used above. The valence band masses then enter into the model by taking the nearest-neighbor coupling to be inversely proportional to the mass, which yields $J_{hh} = J_e m_e/m_{hh} = 4.75meV$ and $J_{lh} = J_e m_e/m_{lh} = 2.52meV$. The site energies for the light-holes are chosen relative to the heavy-holes to have an adjustable splitting between the heavy- and light-hole excitons. Within our model the heavy- and light-hole transitions are coupled by sharing a common electronic state and via the Coulomb coupling.

Fig. 5 shows calculated differential absorption spectra for co-circularly and opposite-circularly excitation for various splittings between the heavy-and light-holes excitons (a) and (b), as well as for various detunings (c) and (d). Considering a detuning of $4.5meV$ below the $1s$ heavy-hole exciton and a splitting of $15meV$ between the heavy- and light-hole excitons (which is the splitting present in the InGaAs quantum well sample investigated in Ref. 15) we reproduce the blueshift for $\sigma^+\sigma^+$ and the redshift for $\sigma^+\sigma^-$ excitation, see Fig. 5(a) and (b). For reduced heavy-light splitting the blueshift for $\sigma^+\sigma^+$ excitation remains unchanged. However, for $\sigma^+\sigma^-$ excitation the amplitude of the redshift strongly decreases with decreasing splitting. For a splitting of $1meV$ and less the redshift even changes

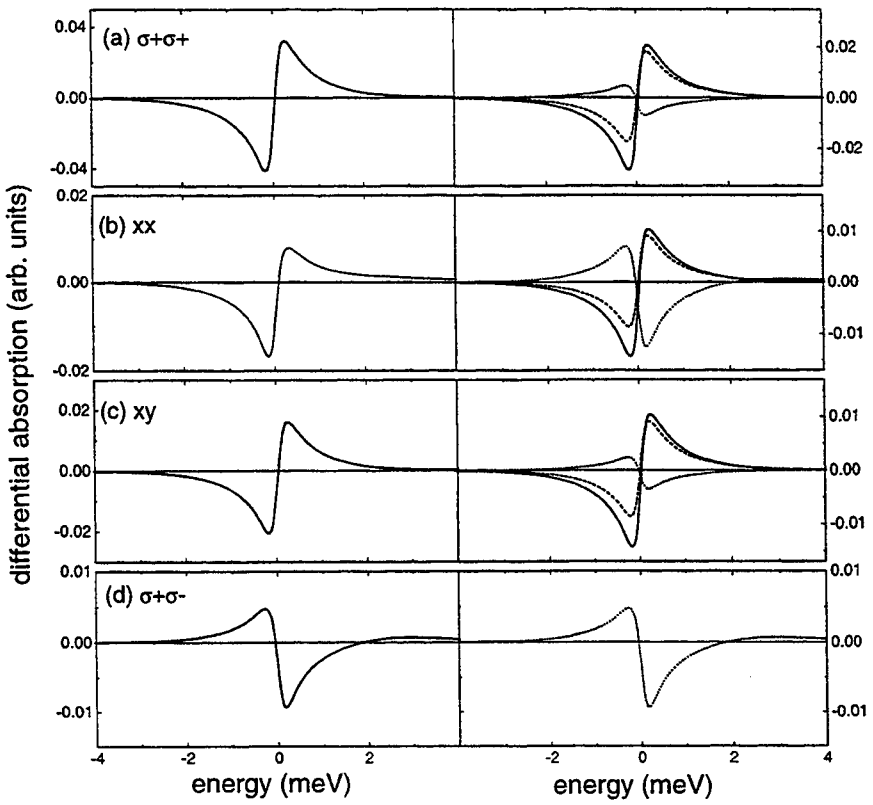


Figure 4. Differential absorption spectra for excitation $4.5\,meV$ below the exciton and time delay $\tau = 0\,ps$. (a) Co-circularly ($\sigma^+\sigma^+$), (b) linear parallel (xx), (c) linear perpendicular (xy), and (d) opposite-circularly ($\sigma^+\sigma^-$) polarized pump and probe pulses . The left panels show the full results, whereas the right panels show the three contributions to the signal: Pauli blocking (solid), first-order Coulomb (dashed), and Coulomb correlations (dotted).

into a blueshift. We thus conclude that the redshift should be well pronounced only in samples with a significant heavy-light splitting. For small splittings the coupling between heavy- and light-hole excitons leads to a blueshift which overcompensates the correlation-induced redshift of the heavy-hole exciton alone.

For a fixed heavy-light splitting taken as $15meV$ in Fig. 5(c) and (d) the differential absorption spectra depend strongly on the detuning. For co-circularly excitation the blueshift present for a detuning of $4.5meV$ survives also for larger detunings and is only reduced in amplitude with increased detuning. For opposite-circular excitation, however, besides a strong reduction in amplitude also the direction of the shift changes with increasing detuning. For very large detuning we again obtain a switch over from red- to blueshift. The reason for this behavior is that for very large detuning, heavy- and light-hole excitons are both off-resonant and we enter the regime of complete adiabatic following where the frequency dependent optical response decays only weakly with increasing resonance frequency.[27] Thus with increasing detuning the *relative* weight of the light-hole exciton increases, which then due to heavy-hole light-hole coupling induces the blueshift at the heavy-hole exciton.

The results presented in Fig. 5 clearly demonstrate that the Coulomb memory induced redshift for opposite-circularly polarized excitation is very sensitive to both the heavy-hole light-hole splitting and the pump detuning. Thus it should only be observable in samples with sufficiently large exciton splittings and only in a certain detuning range,[23] as was the case in the experiments presented in Ref. 15. The exact detuning and splitting ranges that allow to observe the redshift depend on the heavy- and light-hole masses and on the ratio of their oscillator strengths.

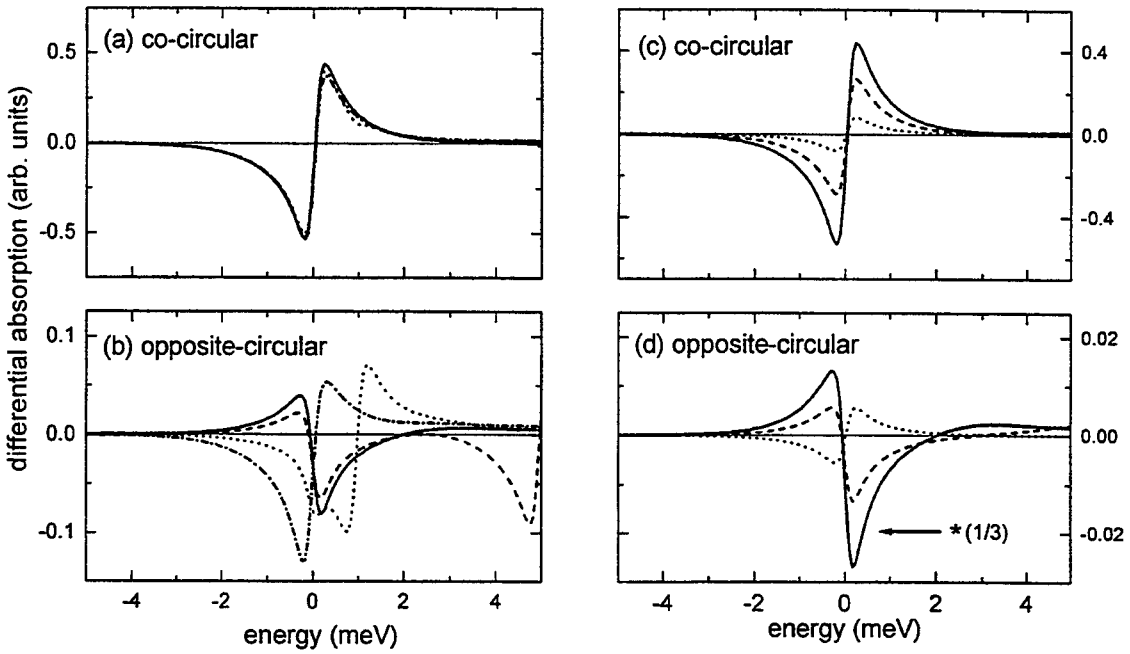


Figure 5. Differential absorption spectra including heavy- and light-holes for time delay $\tau = 0\,ps$. (a) Co-circularly and (b) opposite-circularly polarized pump and probe pulses, respectively, for a fixed detuning of $4.5\,meV$ below the $1s$ hh-exciton resonance and various splittings between the heavy- and light-hole excitons: $15meV$ (solid), $5meV$ (dashed), $1meV$ (dotted), and $0meV$ (dashed-dotted). (c) Co-circularly and (d) opposite-circularly polarized pump and probe pulses, respectively, for a fixed splitting between the heavy- and light-hole excitons of $15meV$ and various detunings below the heavy-hole exciton resonance. $4.5meV$ (solid), $6.75meV$ (dashed), and $18meV$ (dotted).

5. ABSORPTION CHANGES DUE TO INCOHERENT OCCUPATIONS

Numerical results on the absorption changes induced by incoherent exciton as well as electron-hole pair populations $\bar{N}$ have been presented in Ref. 20, where the full Eqs. (4) and (5) have been solved numerically for our model. The homogeneous part of the equation of motion of $\bar{N}$ contains the difference of energies of two single excitons. Thus $\bar{N}$ describes pair occupations and coherences.[19,28] As outlined in Ref. 20, the pair occupation can be expanded using the complete set of excitonic eigenstates Ψ_α, which are the eigenstates of the homogeneous part of the equation of p. For low-densities (non-degenerate limit) and in thermal equilibrium we assume that $\bar{N}$ is given by a summation over thermally populated exciton states, i.e. schematically

$$\bar{N} = \sum_\alpha \frac{\exp(-\epsilon_\alpha/k_B T)}{A} \Psi_\alpha^* \Psi_\alpha \quad , \tag{6}$$

with

$$n = \frac{1}{L} \sum_\alpha \frac{\exp(-\epsilon_\alpha/k_B T)}{A} \quad , \tag{7}$$

where ϵ_α is the energy of Ψ_α, L the length of the system, and A is a constant determining the total density n.[20] In Eq. (6) we have assumed that no coherences between different exciton states are present and that excitons of both optically coupled pairs of degenerate electron and heavy-hole subbands are equally populated.

Besides assuming that initially thermally populated excitons are present, we can also consider the situation where unbound electron-hole states are populated. To model this situation we simply replace the exciton wavefunctions Ψ_α in Eq. (6) by the complete set of single-particle eigenstates Φ_α with energies $\tilde{\epsilon}_\alpha$, which are the eigenstates of the homogeneous part of the equation of p if the electron-hole Coulomb attraction V is neglected. For generating the thermal electron-hole-pair occupation $\bar{N}$ we then use Eq. (6), where Ψ_α is replaced by Φ_α and ϵ_α by $\tilde{\epsilon}_\alpha$.[20]

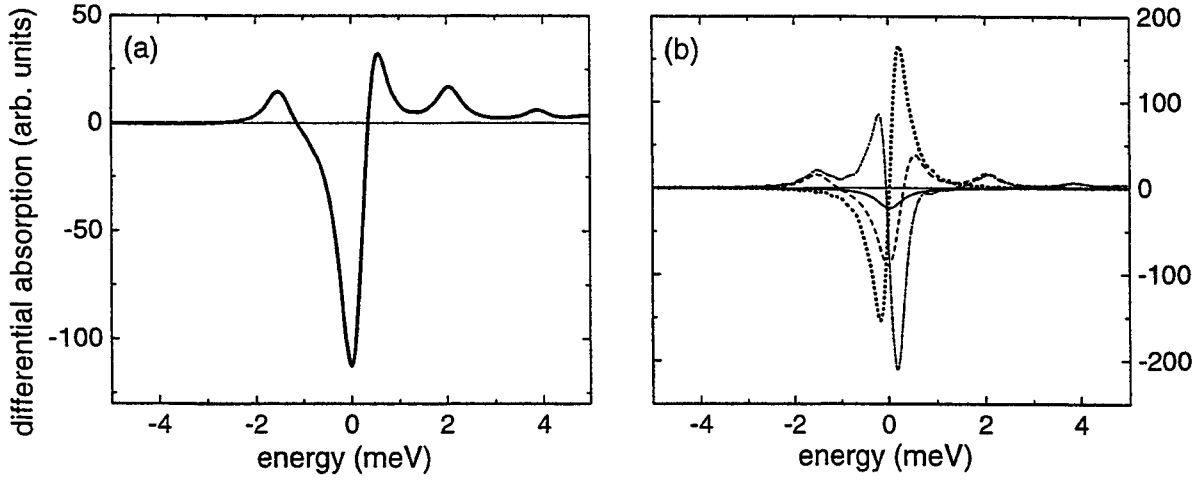


Figure 6. Differential absorption spectra induced by an incoherent occupation of the lowest exciton state. (a) total signal and (b) Pauli blocking (solid), first-order Coulomb terms (dotted), higher-order Coulomb-correlations (dashed-dotted), and sum of the two Coulomb terms (dashed). Dephasing times of $2ps$ for p and $\bar{Z}$ have been considered.

We start our numerical analysis with the low temperature limit. At $T = 0K$ only the energetically lowest exciton state, i.e. the optically active $1s$-exciton, is initially populated, and we calculate the differential absorption induced by an incoherent occupation $\bar{N}$ of this exciton viewed by a circularly polarized probe pulse. Fig. 6 show the numerically calculated differential absorption spectra at $T = 0K$. Besides the total signal displayed in Fig. 6(a) also the three contributions due to Pauli blocking and first- and higher-order Coulomb-interactions, as well as the sum of the two Coulomb terms are shown separately in Fig. 6(b). We see that Pauli blocking induces a pure and rather small bleaching of the exciton, whereas the first-order Coulomb-term has a dispersive shape at the exciton corresponding to the well-known blueshift.[10,29] As for the resonant coherent excitation, this blueshift is strongly compensated by the higher-order Coulomb-term ($\propto V\bar{Z}$), which besides excited-state absorption induced by exciton- to two-exciton transitions shows a redshift at the exciton resonance,[20] see Fig. 6(b). Adding the contributions we find that the total signal shows no net shift but bleaching of the exciton resonance, see Fig. 6(a). This bleaching of the exciton is mainly induced by the strong Coulomb-interaction-induced terms, and the Pauli blocking introduces only weak additional bleaching. Furthermore, we obtain excited state absorption above and below the exciton which is induced by unbound and bound biexcitonic two-exciton states, respectively. The results displayed in Fig. 6 clearly demonstrate that for a proper description of both exciton bleaching and excited-state absorption higher-order Coulomb correlations represented by the six-particle density matrix $\bar{Z}$ are important and should not be neglected.[20] The Pauli blocking- and Coulomb-interaction-induced signatures in the differential absorption including strong cancellation between the first- and higher-order Coulomb-terms and the dominance of the Coulomb-terms over the Pauli blocking in the presence of incoherent occupations[20] are very similar to the signatures obtained in the analysis of coherent pump-induced differential absorption changes, see Refs. 15 and 21.

Fig. 7 displays the differential absorption induced by (a) thermalized excitons and (b) thermalized electron-hole pairs for two temperatures of $T = 2K$ and $100K$. As shown in Ref. 20, also at elevated temperatures for both excitons and electron-hole pairs the different signatures of the three nonlinearities as discussed above remains. Fig. 7(b) shows that with raising the temperature of an occupation of unbound electron-hole pairs both the bleaching of the exciton as well as the amplitude of the excited state absorption decrease. The result obtained for thermalized exciton occupations, Fig. 7(a), looks similar. At very low temperatures, where only the lowest exciton is populated, one sees clear signatures of the individual two-exciton states. These positive contributions become less structured at elevated temperatures due to the fact that for increased temperatures also energetically higher excitons are populated. Their transitions to two-exciton states, which occur at different frequencies, are added to the signal induced by the lowest exciton. The tail of the excited state absorption towards low energies appearing for elevated temperatures in the

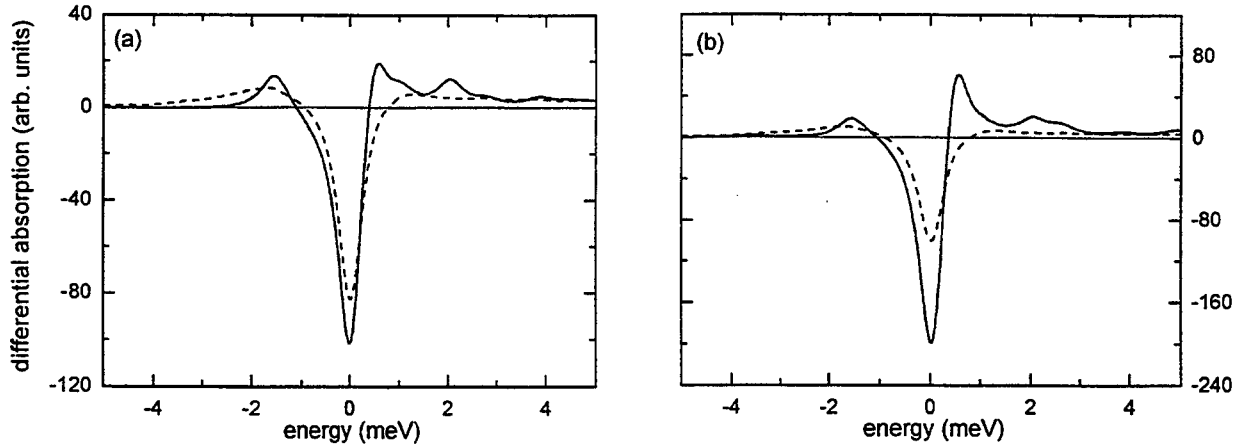


Figure 7. Differential absorption spectra for $T = 2K$ (solid) and $100K$ (dotted) induced by (a) thermally occupied excitons and (d) thermally occupied electron-hole states.

excited state absorption below the exciton directly shows that also energetically higher excitons are Coulombically-coupled to two-exciton states. Comparing the cases of unbound excitons and electron-hole pairs, Fig. 7(a) and (b), we note that for a fixed temperature the lineshapes of the differential absorption are quite similar. However, for thermalized electron-hole pairs the bleaching of the exciton resonance is stronger than for excitons. This stronger bleaching induced by electron-hole pairs compared to excitons means that there is stronger nonlinear coupling of low-energy electron-hole pairs to the 1s-exciton compared to the nonlinear coupling among the excitons themselves. For the Pauli-blocking these differences can be analyzed simply by using the exciton and electron-hole wavefunctions, as described in Ref. 29. If Coulomb-correlations are included in the description, on the other hand, due to the relevance of six particle correlations no such simple explanation is possible, since also two-exciton states (which can not be calculated analytically) are relevant. The temperature dependence of the bleaching induced by thermalized excitons and unbound electron and hole pairs obtained within our model has been investigated in detail in Ref. 20.

The dominance of the unbound electron-hole pair over the exciton contributions to the resonance bleaching is most likely the reason for the success of electron-hole plasma theories explaining experimentally observed exciton saturation at low temperatures.[1,3,5,29] This trend is expected to be even more pronounced in systems with dimensionalities greater than one because of the reduction of the Coulomb-correlation effects with increasing dimensionality.

6. SUMMARY

A microscopic theory capable of describing coherent excitation effects in semiconductors including many-body Coulomb-induced carrier correlations is reviewed. The theory is applied to analyze polarization-dependent ultrafast absorption changes at the 1s exciton resonance. For two-band semiconductors it is shown that the optical Stark effect may change from a blueshift for co-circularly polarized pump and probe pulses to a redshift for opposite-circularly excitation conditions. A theoretical analysis including heavy and light holes shows that the occurrence of the redshift for opposite-circularly polarized excitation depends sensitively on both the splitting between the heavy- and light-hole excitons and the pump detuning.

The theory is also applied to an incoherent situation. Whereas one could expect that correlations might not be important in this situation, strong correlation effects on the differential absorption induced by incoherent thermalized excitons and free electron-hole pairs are found. For low temperatures it is shown that populated free electron-hole pairs induce stronger bleaching at the exciton resonance than populated excitons of the same density.

In Ref. 21 the approach has been generalized to treat also effects of energetic disorder. Furthermore, the theory has been applied to excitons in microcavities,[21] yielding numerical results on reflection changes that are in good qualitative agreement with recent experiments.[30] The theoretical approach has further been extended to include effects of higher intensities (coherent $\chi^{(5)}$-limit) where it yields results on the excitonic differential absorption in good

agreement with experimental results.[31] In addition, the theoretical approach has been used to clarify the relevance of a new coherent technique called "coherent excitation spectroscopy" which is based on partly nondegenerate four-wave mixing.[24] Work is in progress dealing with the analysis of disorder-induced dephasing in four-wave mixing spectroscopy, which can only be understood properly if both disorder and correlations are treated microscopically.[22]

ACKNOWLEDGMENTS

We thank P. Thomas and our co-workers C. Sieh, F. Jahnke, and A. Knorr for many collaborations and valuable discussions. This work is supported by the Deutsche Forschungsgemeinschaft (DFG) through the Sonderforschungsbereich 383 and the Leibniz prize, and by the John von Neumann Institut für Computing (NIC), Forschungszentrum Jülich, Germany, through grants for extended CPU time on their supercomputer systems.

REFERENCES

1. H. Haug and S. Schmitt-Rink, "Electron theory of the optical properties of laser excited semiconductors," *Prog. in Quant. Electron.* **9**, pp. 3–100, 1984.
2. S. Schmitt-Rink, D. A. B. Miller, and D. S. Chemla, "Linear and nonlinear optical properties of semiconductor quantum wells," *Adv. in Phys.* **38**, pp. 89–188, 1989.
3. L. Bányai and S. W. Koch, "A simple theory for the effects of plasma screening on the optical spectra of highly excited semiconductors," *Z. Phys. B* **63**, pp. 283–291, 1986.
4. S. W. Koch, N. Peyghambarian, and M. Lindberg, "Transient and steady-state optical non-linearities in semiconductors," *J. Phys. C* **21**, pp. 5229–5249, 1988.
5. H. Haug and S. W. Koch, *Quantum Theory of the Optical and Electronic Properties of Semiconductors*, World Scientific, Singapore, 1994.
6. D. Fröhlich, A. Nöthe, and H. Reiman, "Observation of the resonant optical stark effect in a semiconductor," *Phys. Rev. Lett.* **55**, pp. 1335–1337, 1985.
7. A. Mysyrowicz, D. Hulin, A. Antonetti, A. Migus, W. T. Masselink, and H. Morkoc, ""Dressed excitons" in a multiple-quantum-well structure: Evidence for an optical stark effect with femtosecond response time," *Phys. Rev. Lett.* **56**, pp. 2748–2751, 1986.
8. A. V. Lehmen, D. S. Chemla, J. E. Zucker, and J. P. Heritage, "Optical stark effect on excitons in GaAs quantum wells," *Opt. Lett.* **11**, pp. 609–611, 1986.
9. S. Schmitt-Rink and D. S. Chemla, "Collective excitations and the dynamical stark effect in a coherently driven exciton system," *Phys. Rev. Lett.* **57**, pp. 2752–2755, 1986.
10. R. Binder, S. W. Koch, M. Lindberg, W. Schäfer, and F. Jahnke, "Transient many-body effects in the semiconductor optical stark effect: A numerical study," *Phys. Rev. B* **43**, pp. 6520–6529, 1991.
11. H. Wang, K. B. Ferrio, D. G. Steel, Y. Z. Hu, R. Binder, and S. W. Koch, "Transient nonlinear optical response from excitation induced dephasing in GaAs," *Phys. Rev. Lett.* **71**, pp. 1261–1264, 1993.
12. B. F. Feuerbacher, J. Kuhl, and K. Ploog, "Biexcitonic contribution to the degenerate-four-wave-mixing signal from a $GaAs/Al_xGa_{1-x}As$ quantum well," *Phys. Rev. B* **43**, pp. 2439–2441, 1991.
13. K. Bott, O. Heller, D. Bennhardt, S. T. Cundiff, P. Thomas, E. J. Mayer, G. O. Smith, R. Eccleston, J. Kuhl, and K. Ploog, "Influence of exciton-exciton interactions on the coherent optical response in GaAs quantum wells," *Phys. Rev. B* **48**, pp. 17418–17426, 1993.
14. V. M. Axt and A. Stahl, "The role of the biexciton in a dynamic density matrix theory of the semiconductor band edge," *Z. Phys. B* **93**, pp. 205–211, 1994.
15. C. Sieh, T. Meier, F. Jahnke, A. Knorr, S. W. Koch, P. Brick, M. Hübner, C. Ell, J. Prineas, G. Khitrova, and H. M. Gibbs, "Coulomb memory signatures in the excitonic optical stark effect," *Phys. Rev. Lett.* **82**, pp. 3112–3115, 1999.
16. W. Schäfer, D. S. Kim, J. Shah, T. C. Damen, J. E. Cunningham, K. W. Goosen, L. N. Pfeiffer, and K. Köhler, "Femtosecond coherent fields induced by many-particle correlations in transient four-wave mixing," *Phys. Rev. B* **53**, pp. 16429–16443, 1996.
17. M. Lindberg, Y. Z. Hu, R. Binder, and S. W. Koch, "$\chi^{(3)}$ formalism in optically excited semiconductors and its applications in four-wave-mixing spectroscopy," *Phys. Rev. B* **50**, pp. 18060–18072, 1994.
18. V. M. Axt and A. Stahl, "A dynamics-controlled truncation scheme for the hierarchy of density matrices in semiconductor optics," *Z. Phys. B* **93**, pp. 195–204, 1994.

19. V. M. Axt, K. Victor, and A. Stahl, "Influence of a phonon bath on the hierarchy of electronic densities in an optically excited semiconductor," *Phys. Rev. B* **53**, pp. 7244–7258, 1996.
20. T. Meier and S. W. Koch, "Excitons versus unbound electron-hole pairs and their influence on exciton bleaching: A model study," *Phys. Rev. B* **59**, pp. 13202–13208, 1999.
21. C. Sieh, T. Meier, F. Jahnke, A. Knorr, F. Jahnke, P. Thomas, and S. W. Koch, "Influence of carrier correlations on the excitonic optical response including disorder and microcavity effects," *Europ. Phys. J. B* , in press.
22. S. Weiser, T. Meier, J. Möbius, A. Euteneuer, E. J. Mayer, W. Stolz, M. Hofmann, W. W. Rühle, P. Thomas, and S. W. Koch, "Disorder-induced dephasing in semiconductors," submitted.
23. S. W. Koch, C. Sieh, T. Meier, F. Jahnke, A. Knorr, P. Brick, M. Hübner, C. Ell, J. Prineas, G. Khitrova, and H. M. Gibbs, "Theory of coherent effects in semiconductors," *J. Lumin.* , in press.
24. A. Euteneuer, E. Finger, M. Hofmann, W. Stolz, T. Meier, P. Thomas, S. W. Koch, W. W. Rühle, R. Hey, and K. Ploog, "Coherent excitation spectroscopy on inhomogeneous exciton ensembles," *Phys. Rev. Lett.* , in press.
25. F. Jahnke, M. Kira, S. W. Koch, G. Khitrova, E. K. Lindmark, T. R. Nelson, D. V. Wick, J. D. Berger, O. Lyngnes, H. M. Gibbs, and K. Tai, "Excitonic nonlinearities of semiconductor microcavities in the nonperturbative regime," *Phys. Rev. Lett.* **77**, pp. 5257–5260, 1996.
26. F. Jahnke, M. Kira, and S. W. Koch, "Linear and nonlinear optical properties of excitons in semiconductor quantum wells and microcavities," *Z. Phys. B* **104**, pp. 559–572, 1997.
27. R. Binder, S. W. Koch, M. Lindberg, N. Peyghambarian, and W. Schäfer, "Ultrafast adiabatic following in semiconductors," *Phys. Rev. Lett.* **65**, pp. 899–902, 1990.
28. G. Bartels, G. C. Cho, T. Dekorsy, H. Kurz, A. Stahl, and K. Köhler, "Coherent signature of differential transmission signals in semiconductors: Theory and experiment," *Phys. Rev. B* **55**, pp. 16404–16413, 1997.
29. S. Schmitt-Rink, D. S. Chemla, and D. A. B. Miller, "Theory of transient excitonic optical nonlinearities in semiconductor quantum-well structures," *Phys. Rev. B* **32**, pp. 6601–6609, 1985.
30. X. Fan, H. Wang, H. Q. Hou, and B. E. Hammons, "Biexcitonic effects in the nonperturbative regime of semiconductor microcavities," *Phys. Rev. B* **57**, pp. 9451–9454, 1998.
31. P. Brick et al. , unpublished.

Advances in Missile Technology - Impact of Photonics *

Dr. A. Sivathanu Pillai

Chief Controller R&D, Defence Research and Development Organisation
B Wing, Sena Bhavan, New Delhi - 110 011

ABSTRACT

This paper discusses the role and impact of advancements in Photonics Technology on the performance enhancement of guided missile weapon systems with specific reference to the development of Indian guided missiles programme. India is emerging as a technologically strong nation with core competence in Space, Missile and Nuclear technologies, advanced computing including supercomputers and software. Based on the realisation of the fact that high technology strength is the key to economic prosperity and military strength, India is progressing several high technology areas that help in attaining the global competitiveness. Photonics is identified as one of the important areas in this direction and hence high priority has been accorded for Research & Development in Photonics. This paper reviews the current trends and developments in missile technology and highlights some of the important developments in Photonics that have a force multiplying effect on the performance enhancement of guided missile systems.

Key words: High technology in India, Guided Missiles Technology, Photonics

1. INTRODUCTION

A study of the evolution of war weaponry reveals that, over the years, guided missiles have evolved as a weapon of choice for the modern warfare. During the early days of warfare, the wars were fought using human and animal strength and the role of weapons were mainly to augment this physical strength. At that time the theatre of war was confined only to the land and sea. This trend continued up to the18th century till the invention of gunpowder changed the nature of warfare. Rocketry made its beginning during the last quarter of the 18th century. A major milestone in the history of Rocketry occurred in 1792 AD, when the world's first war rocket was used by Tippu Sultan, Prince of Mysore, India, to defeat the British cavalry attack at Srirangapatanam, near Mysore in India. This has triggered the development of rocket as a war weapon. However, the real transformation happened almost 150 years later, towards the end of World War II, with the development of world's first guided missile - V2, by Germany. That was the beginning of revolutionary changes in the warfare that resulted in the development and use of a wide spectrum of guided missiles launched from multiple platforms and aided by variety of sensors. Gradually the theatre of war had extended to encompass the deep sea and outer space. High technology devices and intelligent and autonomous systems have emerged as the dominating features of the modern warfare. As a result of the transformation of the nature of warfare from being a human & animal warfare to weapon warfare to ultimately the high technology warfare, guided missile systems are at the core of this high technological warfare. There has been a continuous performance enhancement in every aspect of the guided missiles, as a result of the advancement of technologies in various fields. Photonics is one such area having significant impact on the performance enhancement of the guided missile weapon systems. This paper highlights some of the important developments in Photonics that have a force multiplying effect on the performance enhancement of guided missile systems.

*Also published in *Proc. of SPIE* Vols. 3896, 3898, and 3899

Part of the SPIE Conference on Advanced Photonic Sensors and Applications
Singapore • November–December 1999
SPIE Vol. 3897 • 0277-786X/99/$10.00

This paper is organised into three parts in the following sequence:

- A brief review of high technology growth in India, strength and core competence
- Integrated Guided Missile Development Programme, Missions, Technologies, and Global assessment
- Advances in Photonics and their impact on the performance enhancement of guided missile systems

2. TECHNOLOGY LEADERSHIP AND CORE COMPETENCE

A Nation is strong and gets its rightful place only if it has economic prosperity and comprehensive security. Realising that this can be achieved through technology, India has progressed several technological areas that helped the country to emerge as a technologically strong nation. During the last five decades, India has made an all round technological progress with many accomplishments. The green revolution and the operation flood made the country self sufficient in food production, milk, vegetables, fruits, cereals and other essential areas. The technology related to healthcare has resulted in increase of life expectance. The quest for tapping natural resources for generating power has given new impetus to the power sector. The nuclear tests in 1974 and in 1998 made India a nuclear weapon state, and India mastered harnessing of nuclear power to meet the growing demand for electricity. 500 MW fast breeder development and the target of 3000 MW power generation by 2000 are the immediate goals. By 2020 the nation is targeting 20000 MW nuclear power generation. India already has a 2 GFLOP super computer, which will reach 8 GFLOP by 2000 and teraflop speed by 2020. The nation is passing through a communication revolution towards reaching 20 million lines and soon to 100 million lines. With the established strength in computing systems, software and communication, large pool of talented software specialists, together, India is emerging as a strong nation in the area of information technology. India's current software export has already reached 1 billion US dollars.

The guided missiles Agni and Prithvi have demonstrated India's capability to develop high technology strategic systems without any help from outside. India made spectacular achievements in building its own launch vehicles, SLV-3, ASLV, and PSLV, and the GSLV which is getting ready will give India the capability to inject satellites into the geo-stationary orbit. The Indian satellites INSAT and IRS with 6 m resolution are among the best in the world. Recently PSLV launched German and Korean satellites providing cost effective launch services to the world. India is self-sufficient in space systems and their applications. In the area of aeronautics, the Indigenous light combat aircraft will take off very soon. India is dreaming of launching a reusable missile, deep penetration strike aircraft, and multipurpose aerospace vehicle - Hyperplane, harnessing the technological base established in aeronautics, space and missiles.

India is self-sufficient in production of all raw materials, special alloys, armour materials, electronic components and devices, computing systems and software for meeting the demands of the nation. Both public and private sector industries have been well knitted with R&D and academy to meet the demands. The Indian Defence R&D organisation, DRDO, achieved technology breakthrough in certain devices and computing systems, ASICs and MMICs, both Silicon and Gallium Arsenide based, PACE+ super computer, ANUPAMA microprocessor with state-of-the-art architecture to meet the requirements of strategic programmes. DRDO has developed core competence in many critical areas of armaments, combat vehicles, electronics and communications, aeronautics, missiles, life supporting systems and computing systems. This core competence is being used to develop state-of-the-art high technology systems for the armed forces.

India is blessed with abundance of manpower, natural resources and good value system. The R&D development cost varies from $1/10^{th}$ to $1/3^{rd}$ of the advanced nations depending on the type of programme. Also through various programmes, Indian specialists have learnt to work together with academic institutions and industries as partners, sharing each others strength to design, develop, and produce high technology systems in the country without any dependence from abroad.

3. INTEGRATED GUIDED MISSILE DEVELOPMENT

3.1 Mission & Technologies

The Integrated Guided Missile Development Programme (IGMDP) commenced in 1983 with an objective of design, development and leading to production of four types of missile systems namely, Prithvi, Trishul, Akash and Nag and technology demonstrator for long range system - Agni. Prithvi is a short range battlefield Surface to Surface missile launched from a mobile transporter. Prithvi uses liquid propellant rocket engines, light weight airframe, high accuracy strapdown inertial guidance system and electro-hydraulic control system. Trishul is a low level quick reaction surface to air missile system designed to defend the moving armoured columns from attacking aircraft and helicopters. The Naval version of Trishul is designed to defend the ships from the incoming sea-skimming missiles. Akash is a medium range surface to air missile with multi-target handling capability and ECCM features. Nag is a 3rd generation anti tank guided missile with a fire and forget and top attack capabilities. Nag has two state-of-the-art guidance systems, one using an Imaging Infra Red seeker and the other using a Millimetric wave active radar seeker. Agni is an Intermediate Range Ballistic Missile, which uses a specially developed carbon-carbon re-entry vehicle structure and re-entry guidance and control.

The technological goal of the IGMDP is to ensure that the systems are contemporary at the time of their induction into the Services. As a result, the technologies have been forecast 10 to 15 years in advance, concurrently developed, and channelled into the missile systems. The programme used several innovative management practices including the multi-institutional partnership, consortium, collaboration, technology empowerment, and developed all the critical technologies required for the programmes within the country using a network of R&D laboratories, academic institutions, public and private sector industries. The systems developed are of multi-user and multi-role in nature with contemporary performance.

3.2 Global Assessment

With a high supersonic manoeuvrable trajectory, high lethality, multiple field interchangeable warheads, and high accuracy, Prithvi is among the best in its class. With a quick reaction time of 6 sec, jamming proof Ka-band guidance, and anti-sea skimming missile capabilities, Trishul has contemporary performance in its class. The performance of Trishul is comparable to ADATS, CROTALE and BARAK. With multi-target engagement capability and modern homing guidance, Akash is state-of-the-art and comparable in performance to the PATRIOT. Nag is aiming to be the first of its kind in the world with fire & forget and top attack capabilities and a Tandem warhead capable of defeating composite and reactive armour.

3.3 Further advancements

Based on the above strength, India is working on several next generation technologies to enhance the performance of its guided missiles. The general directions of improvements include:

(a) Faster speeds towards hypersonic missiles
(b) Improved precision towards zero CEP
(c) Intelligent systems like the Terminally Guided Sub-munitions
(d) Capability of launching from multiple platforms
(e) Enhanced stealth
(f) Better ECM / ECCM capabilities

The above directions of improvement requires major technology advancements in the subsystems towards:

(a) Lower weight
(b) Less volume (Miniaturisation)
(c) Lower power consumption

(d) Better performance in terms of the processing efficiency, minimum errors, greater sensitivity and greater precision
(e) Minimum aerodynamic interference and better integration with missile
(f) Lower cost

It is interesting to note that many developments in Photonics could help in realising the above objectives. Some of them are discussed below.

4. ADVANCEMENTS IN PHOTONICS AND THEIR IMPACT ON PERFORMANCE ENHANCEMENT OF GUIDED MISSILES

4.1 Current applications of Photonics in Guided Missiles

Photonics has always been one of the important technologies used in various aspects of guided missiles. Some of the important areas where Photonics and related technologies play a crucial role in guided missiles include:

(a) Missile guidance applications

- Missile tracking sensors for command guidance
  (Ex : IR plume trackers used in 2nd generation anti tank guided missiles and IR imaging sensors used in surface to air missiles)
- Target tracking sensors for homing guidance
  (Ex : IR homing seekers for air to air missiles, IR imaging sensors for 3rd generation anti tank guided missiles)
- Inertial sensors for navigation systems used in surface to surface missiles and number of other applications
  (Ex: Fibre optic gyros, ring laser gyros)
- Imaging sensors for mid course guidance and terminal homing for long range missiles
  (Ex : Imaging sensors for terrain contour matching and terminal guidance)
- Illuminating source and homing sensors for semi-active homing applications, and range finders used in certain class of missiles

(b) Proximity sensors for initiating warhead detonation
(Ex: IR Proximity fuzes used in anti sea skimmer missiles, altitude sensors used in surface to surface missiles)

(c) Surveillance, target detection and target acquisition applications

- Thermal sights for target acquisition in 3rd generation anti-tank guided missiles
- Infra red search & tracking systems for naval applications
- Space based sensors for detection of missile launch point for Ballistic Missile Defence Systems
- Laser radars for precision tracking

(d) Equipment for command, control & communications

- Fibre optic communications
- Display systems

(e) Electro-optic counter measures & counter counter measure

(f) Design aids for missile subsystems
(Ex: Experimental stress analysis tools, Laser anemometers used in wind tunnel testing applications)

(g) Test & evaluation instrumentation for guided missiles
(Ex : Electro-optic tracking instruments used in missiles test ranges)

(h) Manufacturing of precision missile components and quality control
(Ex : Laser welding, equipment for inspection and measurement, holography for NDT)

The above list indicates that Photonics is important at every phase of the guided missiles lifecycle. Therefore any development in this technology has a wide ranging impact on the missile systems performance.

4.2 Advancements in Photonics and their impact on guided missiles

4.2.1 Some of the important developments in Photonics and related technologies with potential applications for guided missile systems include:

(a) Uncooled Imaging technology
(b) Focal plane array technology
(c) Conformal optics
(d) Adaptive optics
(e) Fiber optics technology
(f) Micro-Opto-Electro -Mechanical Systems (MOEMS)
(g) Optical Correlators
(h) Multi-Spectral fusion (fusion of 3-5μ and 8-12μ images)
(i) Multi-sensor Fusion (fusion of 3-5μ and 8-12μ images of IR and MM wave images)
(j) Real-time Image Processing

4.2.2 Uncooled Imaging

Uncooled thermal imaging systems are very important for missile systems as these system offer significant operational advantages and saving of weight, space as well as cost associated with the cooling system. Towards this the resistive microbolometer technology, ferro-electric bolometer technology, uncooled thermo electric linear arrays are becoming more important. The research effort in uncooled infrared technology is directed towards several other military applications including reconnaissance, surveillance and weapon sighting capabilities as well as for precision munitions and dispenser system applications and anti armour submunition programmes.

4.2.3 Focal Plane Array Technology

Focal Plane Arrays technology is crucial for missile seekers and other imaging sensors. The focus is on developing larger size arrays, higher resolution and higher sensitivity focal planes. The current research is also focused on reducing the pixel size and increasing pixel sensitivity using advanced materials and micro electromechanical device structures. The advancement in FPA technology has a direct impact on the guided missile systems. For example, target detection and lock on range of a 3rd generation anti tank guided missiles can be significantly increased by enhancing the performance of the FPA.

4.2.4 Conformal optics

Conformal optics minimises the aerodynamic interference of the missile IR domes and windows by shaping the optics to minimise the aerodynamic drag coefficient. This will increase the aerodynamic efficiency of the missile and thus increase its range. However, the design, manufacturing and testing of the conformal optics and development of techniques & methods for dynamic aberration correction is a great challenge that is being addresses by the current research.

4.2.5 Adaptive optics

Adaptive optics compensates for the turbulence induced phase distortions of optical waves propagating through the atmosphere. Adaptive optic systems typically consist of a wavefront phase sensor, focusing optics, a spatial light modulator (SLM) for correcting phase errors, imaging sensors, and the control and processing electronics. These systems improve the image quality by reducing the phase aberrations introduced when the wavefront travels through turbulent atmosphere or aberrations introduced by the optical system itself. Adaptive optic system is a growing area of interest for the guided missiles. Advanced technologies are now becoming available to make these systems lightweight, low power, and compact. The technologies that are making this possible include highly integrated low power electronics, and new processing architectures for error sensing and control, flexible high density packaging, and Micro-Opto-Electro-Mechanical Systems.

4.2.6 Fibre optic technology

The advancement in Fibre optic technology has multiple implications for guided missile systems. The most beneficial development is the enhanced communication abilities of the fibre optics, which are important for command, control and communication applications. Fibre optic guided missile and track via missile systems are some of the important developments in this direction.

Another major application of fibre optic technology is in the area of sensors. Fibre optic gyro used in the inertial navigation units of missiles is an example for this. The Fibre Bragg gratting sensors also called "Smart sensors" represent an exciting development in fibre-sensor technology. Fibre Bragg grating sensors can monitor the manufacture of a product or the condition of a structure in use by providing the real-time feedback by reflecting different wavelengths depending on the condition of the structure. Writing a periodic series of gratings (refractive-index modulations in the core of a single-mode fibre) using an ultraviolet beam generates a low-loss, highly reflective, wavelength-sensitive filter system. This non invasive process does not change the fibre's strength, electromagnetic-interference immunity or dielectric properties. Bragg grating sensors, when fabricated and annealed properly, have proven to be permanent, easily reproducible reflectors that can withstand high temperatures. The gratings respond to changes in temperature, compression and strain by changing the wavelength that they reflect, and their performance is not amplitude or intensity dependent. The sensors are small and can easily attached to surfaces to provide information on the structure's integrity. These sensors are very useful for monitoring the structural conditions during flight testing of guided missiles.

4.2.7 Micro- Opto-Electro-Mechanical Systems (MOEMS)

These systems represent another exciting development for missile and space applications. The saving in the weight, volume and power requirements are of great importance for the applications such as the terminally guided sub-munitions. Micro mirrors are the crucial enabling technology for these systems. Design techniques and fabrication processes for both individual micro mirrors as well as the arrays of micro mirrors are the focus of the current research efforts.

4.2.8 Optical correlators

Optical correlators relies on the Fourier transform property of a lens and a hologram of the Fourier transform of an object as viewed from a particular perspective, to establish a matching correlation coefficient. Optical correlators can be effectively used for automatic target recognition, missile guidance and tracking of the target scenes which do not have prominent features or high contrast with the background. The speed and innate parallel-processing capability of optics enable the correlator to simultaneously recognize and locate all objects in a scene almost instantaneously. These systems are very important for the missile guidance applications, especially the mid course and terminal guidance of long range surface to surface missiles.

4.2.9 Multi-Spectral fusion & Multi-Sensor fusion

The Multi-Spectral fusion methods greatly improve the detection and identification of low observable and camouflaged targets. Hence these methods are becoming more and more common for future applications. Missile seekers with two color detectors are already in operation. However for a battle field application, Multi-sensor fusion is emerging as a major technique to overcome limitations of IR and radar seekers, especially in a dense electronic / electro-optic counter measures environment. Dual mode missile seekers with both IR and radar imaging seekers are becoming important for future applications. Hence this is an important area for missile technology.

4.2.10. Real-time Image Processing

Due to the limitations imposed by the electronic processors for the real-time solutions to critical military application such as automatic target recognition, clutter rejection in infrared search and track (IRST) applications, vision-assisted piloting tasks in unmanned robotic vehicles (e.g., Remotely Piloted Vehicles, Unmanned Aerial Vehicles, or Unmanned Underwater Vehicles), and imaging Identification Friend of Foe, real-time image processing is an important technology. The intrinsic parallelism of optics is expected to circumvent these limitations in future. Towards this, the optronics processor holds a great promise for future real time applications for guided missiles.

4.2.11 The Indian scenario

India has already developed thermal Imagers based on 60 & 100 element linear arrays and 288x4 FPAs for MBT-FCS and Nag missile respectively. It is working on IRST system for Naval applications. A wide technology base exists within the country and also enough skilled manpower available in this field. Keeping in view the technology trends and the requirements of the country, DRDO is working in the development of several state-of-the-art systems including; Night Vision Devices based on Optical Amplification, Intelligent Search & Track system for Airborne platforms, High performance Thermal Imagers using staring FPAs/IR CCDs, Dual-bank/Bi-spectral 3-5 μm and 8-12 μm imagers, Thermal Imagers using uncooled detectors, Real time (on-line) image processing. Laser Instrumentation including Laser Designator/Range Finders using Laser diode pumped Nd:YAG laser, Eye-safe laser systems, Laser Range gated imaging, Laser proximity fuses, Blue-green lasers for under-water imaging, Coherent laser imaging, Sensors for smart ammunitions. Integrated Multi-spectral laser and Radar Warning systems and Autonomous EO Missile jammers with programmed auto-controls. Servo Control Systems will include LOS stabilization to an accuracy of 10μ rad. Photonics includes EO tracking/guidance system based on optical correlation techniques, Helmet mounted display using wave-guide HOE, EO sensor using IO chips, Real-time imaging system through turbulent media. Optical Designs include design of light weight optical systems using GRIN/Binary optics. Special attention is being given to the development of non-linear materials, lasers, IR transmitting materials, thermal sensors, electro-optical devices and transducers, within the country.

Some of the identified focal areas for R&D include :

(a) Photonic Devices

- Optical amplifiers for night vision communication.
- Associated memory devices using neural network for recognition of targets.
- Correlator, convolver, pulse compressor and tuneable filters for signal processing.
- Gallium Arsenide based integrated optical components such as phase-shifters, modulators.
- A to D converter for optical computing applications.
- Distortion free optical recognition.

(b) Photonic Materials

- Laser host materials such as Nd, YAG, Er YAG, CNGG, Alexandrite, Ti Sapphire, Nd Glass.
- Laser non-linear materials such as Lithium Nichate, KTP, BBO, LBO for parametric devices, BSO, BGO and KNB for OPC applications.
- Organic and polymer materials such as Urea, MNA, POM, etc. for SHG applications, MNA/PMMA for guided wave applications, organic dyes for Q-switching and mode locking applications.
- Semiconductor materials such as GaAs, CdTe, CdZnTe, MCT, etc.
- IR materials such as ZnS, ZnSe and Chalcogenide glasses for IR optical components.

(c) Photonic Technologies

- OPC techniques
- Diode laser pumping for solid state lasers
- Optical parametric oscillator
- Er doped passive fibre optical amplifiers
- Real-time optical processing
- Integrated optical waveguide technology
- Holographic optical elements

(d) Sensors

- Optical & IR sensors and thermal imagers for IR detection and thermal imaging.
- Fiber optic sensors , Fiber Optic gyros (FOGs) and Ring-Laser gyros (RLG)
- Multi-spectral and multi-mode homing seekers

5. CONCLUSION

Advances in Photonics technology holds great promise for the performance enhancement of guided missile systems. Realising this potential, the Indian Defence R&D identified Photonics as one of the high priority areas for the research and development. In this context, this paper discussed the various important developments in Photonics and highlighted some areas related to guided missiles. It is clear from the above discussions that Photonics is a great force multiplier for guided missile technology. This paper also briefly outlined the Indian scenario in the area of Photonics and the thrust areas for R&D. With an established core competence in space, missile and nuclear technologies, and information technology, India is poised for greater technological advancement through the synergistic growth of Photonics and guided missile technology.

ACKNOWLEDGEMENT

The author gratefully acknowledges the contributions of several specialists from various DRDO laboratories, projects and partner organisations for their help in evolving the technology vision 2010 for DRDO, identifying the thrust areas in many critical technological areas including Photonics. This paper draws upon those inputs and highlighted some of these areas relevant to the guided missiles.

REFERENCES

1. *Technology Vision 2020*, Technology Information, Forecasting and Assessment Council (TIFAC), New Delhi
2. *Technology Vision 2010 for DRDO*, internal report of Defence Research & Development Organisation, New Delhi
3. R.S. Balcerak, "Uncooled IR imaging: technology for the next generation", *Infrared Technology and Applications XXV*, Proceedings of SPIE, Vol. 3698, USA, 1999
4. R.R. Shannon, "Overview of conformal optics", *Window and Dome Technologies and Materials VI*, Proceedings of SPIE, Vol. 3705, USA, 1999
5. Jefffrey A. Sloan and Donald W. Small, "Design and fabrication of a miniaturized optical correlator", Optical Engineering, , Vol. 32, No.12, pp. 3307-3315, 1993
6. Keith E. Drundin, "Fiber Optic Sensors: Smart Technology for Safety", Photonics, pp. 106-107, January 1996.
7. Deborah Jackson, *A Structural Approach to the Photonic Processor*, RAND Note N-3399-RC

SESSION 1

Photonic Sensors Materials

Invited Paper

Photochromic materials for holographic data storage

Roger A. Lessard[a], Christophe Lafond[a,b], Fatima Ghailane[a], Michèle Bolte[b], Amir Tork[c], Ivan Petkov[d]

[a] Centre d'Optique Photonique et Laser (COPL)
Université Laval, Québec (Québec) G1K7P4 Canada
[b] Laboratoire de Photochimie Moléculaire et Macromoléculaire, UMR CNRS 6505
Université Blaise Pascal, F-63177 Aubière Cedex France
[c] CERSIM, Département de Chimie, Faculté des Sciences et Génie,
Université Laval, Québec, G1K7P4, Canada
[d] University of Sofia, Department of Organic Chemistry
1 James Bouchier Avenue, 1126 Sofia, Bulgaria

ABSTRACT

Photochromism and real-time holographic recording were characterized for two principal classes of photochromic compounds : spiropyrans and fulgides. For spiropyran molecules, important thickness and writing intensity effects were observed. Concerning fulgide, in PMMA matrix, the closed form presents a maximum of absorption centered at 525 nm upon irradiation at 365 nm. We have determined the photoreaction rate constants k_{UV} and k_{VIS} respectively for the coloring and bleaching process : $k_{UV}=1.2\times10^{-3}$ s^{-1} and $k_{VIS}=11.1\times10^{-3}$ s^{-1}. Photochemical fatigue resistance in different polymer matrices was investigated. We found a loss of 9, 11, 13 and 35% respectively in PS, CA, PMMA and PVK. Concerning holographic recording, we obtained diffraction efficiency $\eta=0.65\%$ in PMMA films 30 μm thick.

Keywords: Photochromism, photosensitive materials, fulgides, spiropyrans, polymethylmetacrylate, polyvinylcarbazole, real-time spectrometry, first-order kinetics, photochemical fatigue resistance, holographic recording, two-photon recording process.

1. INTRODUCTION

Photochromism is a reversible transformation of a chemical system between two states, presenting different absorption spectra. That transformation is induced by electromagnetic radiation, typically upon UV irradiation, specie A generated a deeply colored form B characterized by a higher absorbance in the visible range. For the reverse transition from the photogenerated B form to the inactivated state A, both thermal and photochemical effects (visible exposure upon green or red light) give a contribution with different quantum yields pertaining to the reaction $B \rightarrow A$[1].

In organic systems, well-known photochromic mechanisms include E/Z isomerisation (in azobenzene) and pericyclic reactions (in spiropyrans, spirooxazyns, fulgides,and diarylethenes).

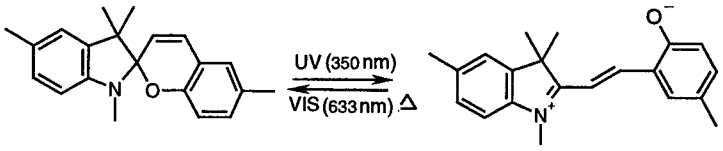
Organic photochromes are developed for both optical switching purposes and optical recording[2]. Therefore, among the required properties of such systems are full reversibility of the photoreactions, photochemical fatigue resistance, thermal stability, short response time and a method for non-destructive readout. While most of these specifications are met by several photochromic systems in particular by some fulgides[3], there is still many works concerning a non-destructive readout method. In this study, results concerning spiropyrans and fulgides will be presented.

In the past few years, many studies were appeared to find the optimum structures and properties, which will make organic compounds suitable for use as materials in the field of optical memories. Particular areas have been considered especially holography and two-photon recording. Indeed, this process represents an important application of photochromism in order to record 3D optical memory devices. In these conditions spiropyrans molecules are taken as example to illustrate this purpose.

Part of the SPIE Conference on Advanced Photonic Sensors and Applications
Singapore • November–December 1999
SPIE Vol. 3897 • 0277-786X/99/$10.00

2. SPIROPYRANS

They are usually colorless in one form **A** (closed form) because they are composed of two Π-electron moieties, which are orthogonal to each over. Absorption spectrum of form **A** lies in the UV/visible region ranging from 200 to 400 nm with an band located between 320 nm and 380 nm. Upon excitation in the UV range, bond cleavage allows for the rotation of the moieties. Simultaneously we observe the formation of a nearly planar structure corresponding to the photomerocyanine form **B** (figure 1). In this form, the molecule is fully formed by delocalized Π-electron system which is responsible for the red-shift observed in the absorption spectrum, from the UV to the 500 nm-600 nm region[4]. Spiropyrans present an important disadvantage. Indeed, the photomerocyanine form is not thermally stable and a back-conversion to the closed form A is taking place. This process can be observed by irradiation with visible light but the thermal process is more prevailing.



A (closed form) B (photomerocyanine)

Figure 1 : Photochemical reactions for spiropyran

2.1. Absorption Spectra

Figure 2 shows the absorption spectra of both forms in PVK matrix[5]. The closed form presents a maximum of absorption at 330 nm. Upon UV irradiation with Krypton ion laser at 350 nm, a strong visible band centered at 614 nm appears which characterized the formation of the photomerocyanine. The closed form is partially regenerated upon exposure with He-Ne laser at 633 nm.

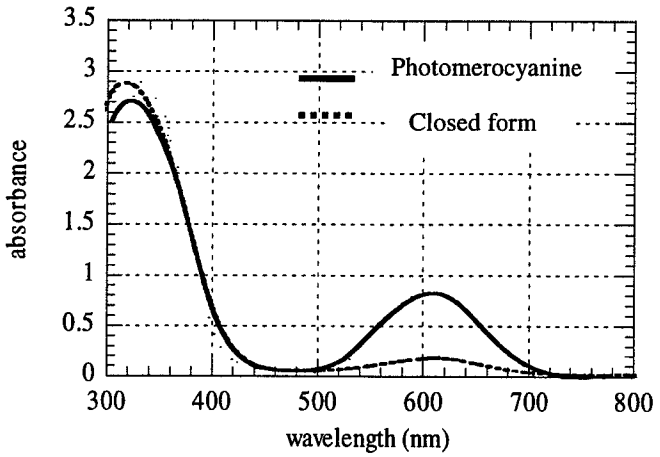


Figure 2 : Absorption spectra for both spiropyran forms :

2.2. Holographic Recording

The holograms were recorded with an UV line (λ_w=350 nm) of a Krypton ion laser. Argon ion laser at λ_r=488 nm performed the reading of the real-time diffraction efficiency : that wavelength is clearly located outside the spiropyran absorption band. The erasing beam came from a He-Ne laser at 633 nm.

2.2.1. Effect of the thickness on the diffraction efficiency

Figure 3 represents the real-time holographic growth profile of spiropyran dye-doped PVK film for two thicknesses : 12 μm and 50 μm.. Dye concentration is 5% by weight of polymer and the writing angle θ_W is 17.5°. The diffraction efficiency

depends strongly on the thickness. The highest diffraction efficiency (10.5%) is obtained for the thickness of 12 µm with an exposure energy of 300 mJ/cm^2. For the same energy, the diffraction efficiency is only 0.43% for a film 50 µm. thick.

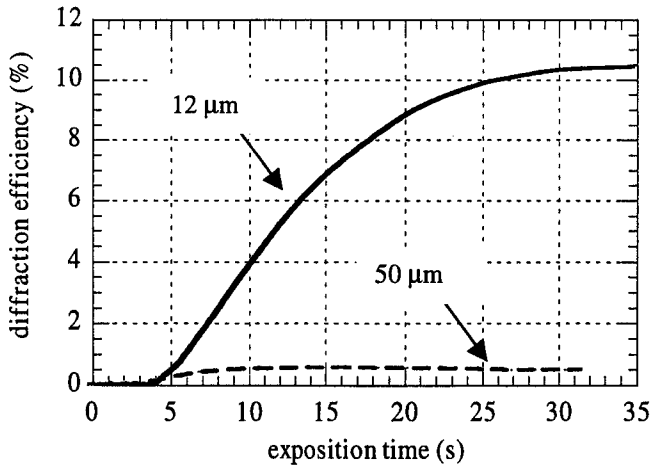


Figure 3 : thickness effect on the diffraction efficiency

2.2.2. Influence of writing intensity on the diffraction efficiency

Figure 4 represents the variations of the diffraction efficiency, in PVK matrix, for different writing intensities, (3 to 50 mW/cm^2). The thickness studied is 50 µm. That parameter increases with an increase in the intensity. The hologram formation is very slow until 10 mW/cm^2. It is much faster for intensities superior than 10 mW/cm^2. At an exposure energy lower than 5 mW/cm^2, the grating growth remains stable even after 15 s of exposure. At higher energies, the diffraction efficiency reached 1.7% of maximum value at an exposure energy of 50 mW/cm^2 but it decreases drastically after 7 s.

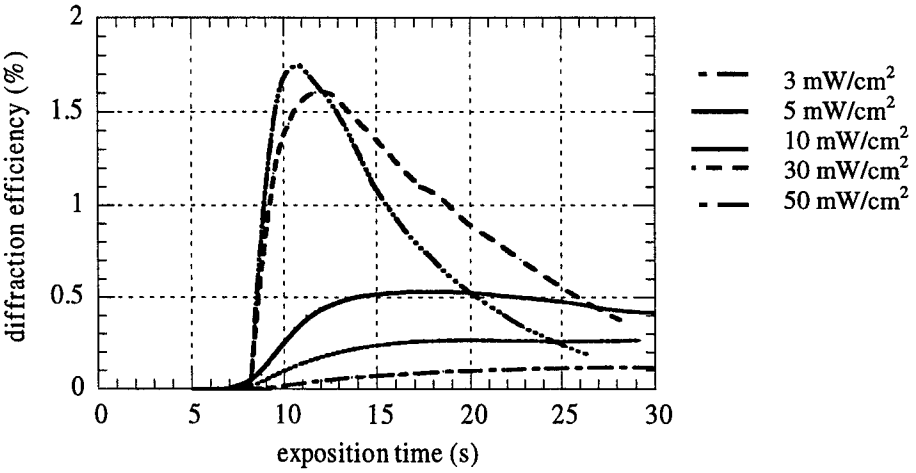


Figure 4 : effect of the writing intensity on the diffraction efficiency
film in PVK, thickness 50 µm, dye concentration 5%

2.2.3. Angular selectivity of spiropyran doped PMMA films

The angular selectivity $\Delta\theta$ is very important to achieve applications such as holographic multiplexing. As shown in figure 5, for thin film[6] in PMMA matrix (14 µm), $\Delta\theta = 6.3°$ and for thick film (250 µm), $\Delta\theta = 4.8°$.

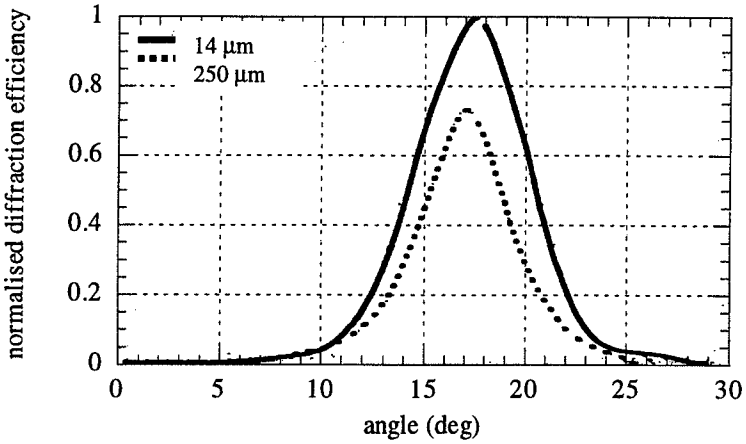


Figure 5 : angular selectivity in PMMA film

3. FULGIDES

A rather important class of photochromic molecules consists of fulgides. Fulgides show completely reversible photochromism. We studied the 2-(1-(2,5-Dimethyl-3-furyl)ethylidene)-3-(2-adamantylidene) succinic anhydride[7] **1** knew in commerce as Aberchrome 670 (figure 6). The 1E-isomer (open form) presents an absorption band only in UV (between 320 and 400 nm). Therefore, upon UV irradiation, a coloring process due to an electrocyclic ring-closure occurs, leading to the formation of the deeply colored 1C-isomer (closed form) which presents a strong absorption in the visible. Irradiation with visible light (bleaching process) induces back-conversion to the open form. However, upon UV irradiation, side reaction, such as E/Z isomerisation, occurs and implies formation of non photochromic Z-isomer but, in solid matrix, this process is less prevailing compared in solution [8].

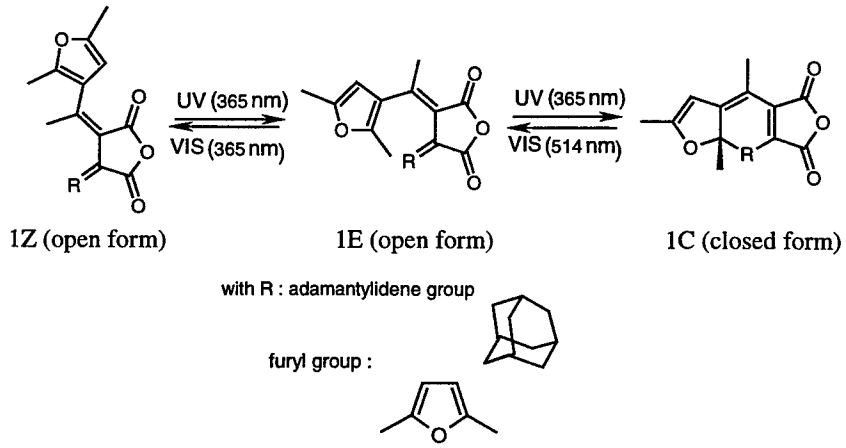


Figure 6 : Photochemical reactions of Aberchrome 670

Films Aberchrome 670 dye-doped polymer were studied in different matrix (PMMA, polymethylmethacrylate, PVK, polyvinylcarbazole, PS, polystyrene, and CA, cellulose acetate). Two approaches were considered.
In a first part, photochemical characterization was investigated (absorption spectral changes and kinetic study upon UV and visible exposure). Moreover, one of the indispensable properties required should be high resistance to photochemical degradation. In these conditions, we have compared the photochemical fatigue resistance of this fulgide in different polymer matrices.
In a second part, results in real-time holographic recording in PMMA and CA matrices were presented.

3.1. Theoritical Study

3.1.1. Kinetic

Equation (1) can be used to fit the experimental data. Each irradiation process has to verify separately the following equation [9].

$$-\ln\frac{A_t - A_e}{A_0 - A_e} = kt \qquad (1)$$

where, A_t represents the absorbance of the closed form in the visible range at the λmax and time t. A_e is the absorbance at the λ_{max} of the closed form measured at the maximum of formation (photostationnary state : PSS) for the coloring process and at the end of the reaction for the bleaching process. A_0 is the absorbance at time t=0 (for the bleaching process, this value was taken as the absorbance of the UV irradiation at the photostationnary state). k represents the photoreactions rate constant : k_{UV} and k_{VIS} for, the coloring and the bleaching process respectively.
If the photoreactions entirely follow first-order kinetics, the left-hand side of Eq.(1) must show a linear time irradiation dependence.

3.1.2. Photochemical fatigue resistance

In order to describe the fatigue of the polymer under cycled Write/Erase procedures, we have represented the changes in A_n/A_0 with UV/visible irradiation cycles numbers. A_0 and A_n are the closed form absorbance values obtained on the first and nth cycles.

3.2. Photochemical Results in PMMA Matrix

3.2.1. Experimental set-up for absorption spectra recording

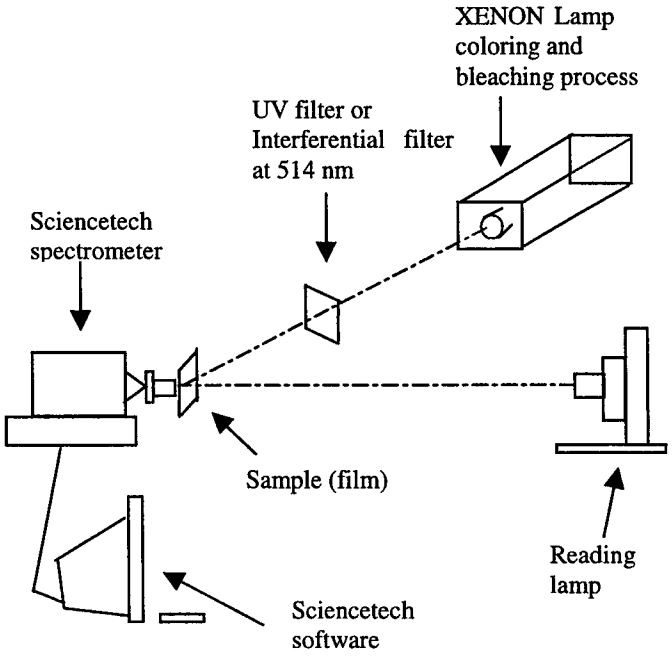


Figure 7 : Photochemical set-up

3.2.1. Absorption spectral changes upon UV/visible irradiation

Figure 8a represents the coloring process of Aberchrome 670 in PMMA film. UV exposure at 365 nm causes the 1E-isomer at 346 nm to decrease while a strong visible band centered at 525 nm, which characterizes the 1C-isomer formation, grows up simultaneously. The photostationnary state is reached after 60 minutes of exposure. The spectral changes, induced by irradiation at 514 nm are shown on figure 8b : the UV band is regenerated.

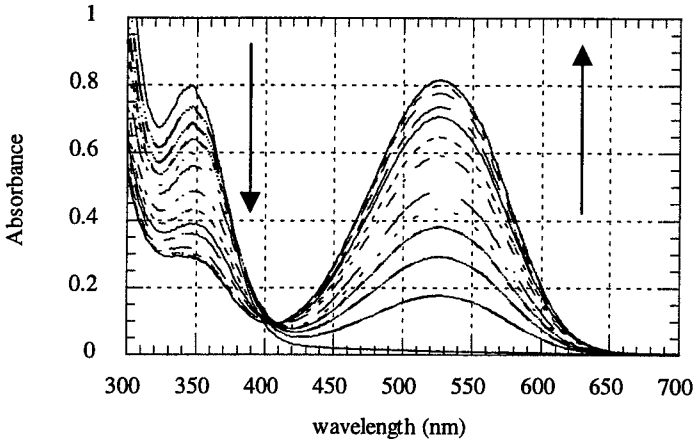


Figure 8a : Absoprtion spectral changes of Aberchrome 670 in PMMA film upon irradiation at 365 nm

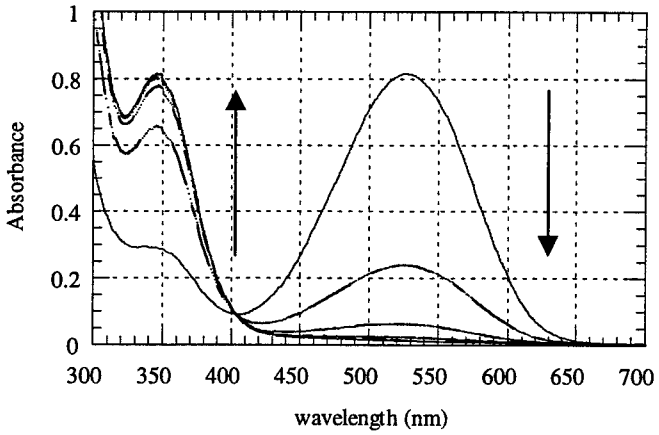


Figure 8b : Absoprtion spectral changes of Aberchrome 670 in PMMA film upon irradiation at 514 nm

3.2.2. Kinetic study

For the first UV/Visible cycle irradiation time dependence of the left-hand side of Eq(1) was represented in figure 9. We can observe that both irradiation processes obey to first-order kinetic: there is a linear dependence. Moreover the bleaching process is much larger than the coloring:

$$k_{UV}=1.2\times10^{-3}\ s^{-1} \text{ and } k_{VIS}=11.1\times10^{-3}\ s^{-1}.$$

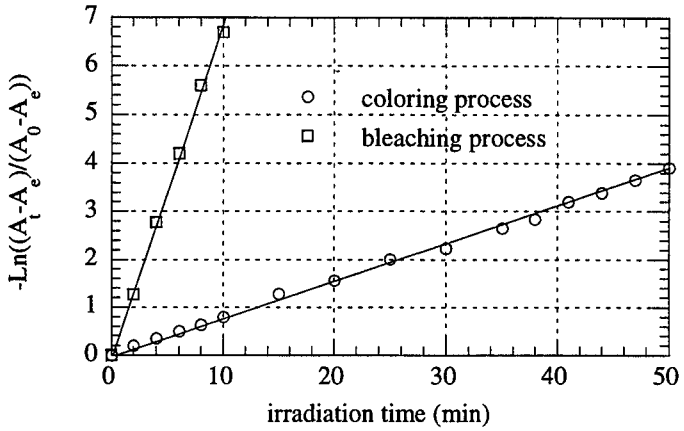


Figure 9 : irradiation time dependence of the left-hand side of Eq(1)

3.3. Photochemical Fatigue Resistance

The photochemical fatigue resistance of fulgide Aberchrome 670 was analyzed upon repeated UV and visible exposure cycles (figure 10). We have considered four polymer matrices : AC, PS, PMMA, and PVK.
This study shows an important matrix effect between PVK and the other polymers. Indeed we have obtained, in PVK, a loss of 35% of the C-isomer initial absorption after 10 cycles against 13, 11 and 9 respectively in PMMA, CA and PS. In these conditions, PVK is not a satisfactory candidate for the recording of repeated Write/Read/Erase cycles.
Moreover, Kaneko et al[10]. found a loss of 15% for Aberchrome 540 in PMMA matrix. This value is in agreement with our result (13%). Both Aberchromes only differ by substituent R, which represents isopropylidene and adamantylidene group respectively for Aberchrome 540 and 670. Consequently, this group does not induce photochemical fatigue resistance. This phenomenon is principally caused by the aryl group, which is, on both cases, a furyl group.

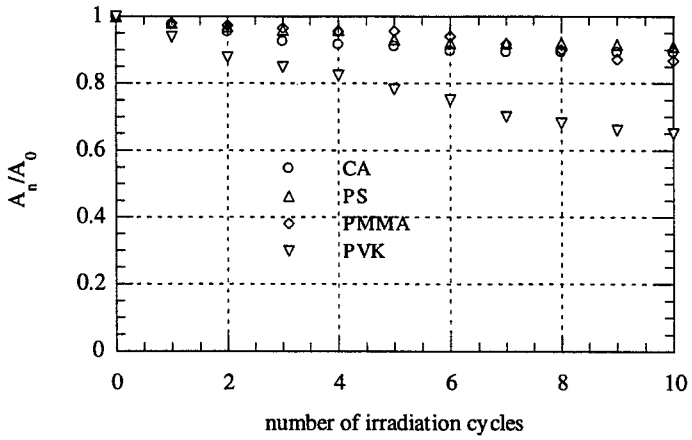


Figure 10 : Photochemical fatigue resistance of A670 in PMMA, CA, PS and PVK films after 10 UV/Vis irradiation cycles

3.4. Holographic Recording

Holographic recording using Aberchrome 670 doped PMMA films was investigated. To characterize a hologram, we have to define the diffraction efficiency[11] η:

$$\eta = \frac{I_1}{I_0}$$

I_1 represents the intensity of the diffracted first order beam and I_0 is the intensity of the incident beam.

3.4.1. Experimental set-up:

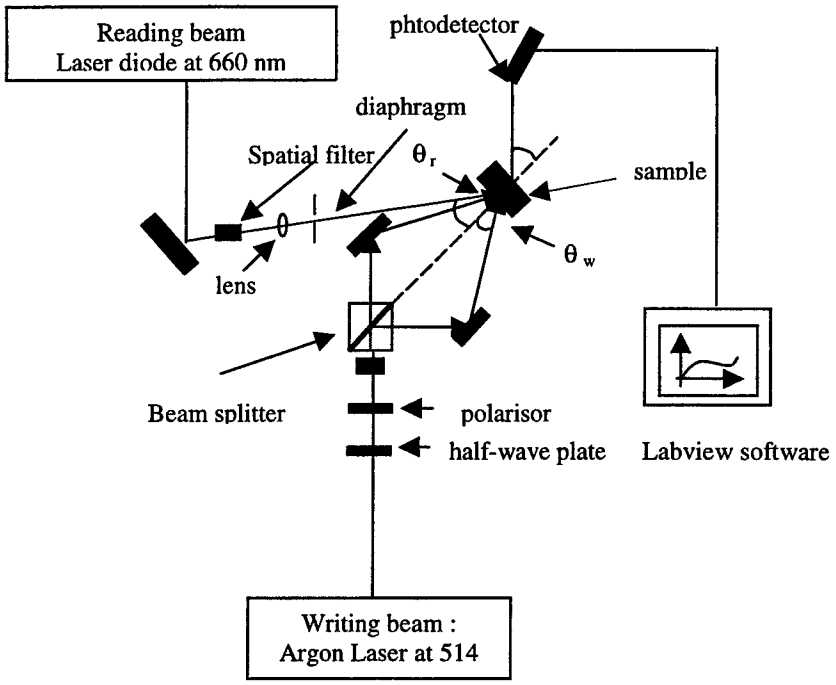


Figure 11 : Holographic recording-experimental set-up

UV irradiation at 365nm induced the formation of a strong absorption band centered at 525 nm. In these conditions, holograms can be written with Argon ion laser at 514 nm. Diode laser at 660 nm has been used for the reading beam, indeed, at this wavelength, the photochromic compound does not absorb.
Finally, the grating can be erased upon irradiation with UV light at 365 nm.
In this study thick holograms were only considered (Quality factor[12] Q=50 for films in PMMA matrix 20 microns thick). Moreover, with a writing angle θ_w=11°, holographic recording in PMMA, gives a grating with N=1100 lines/mm.

3.4.2. Results in PMMA matrix

The Argon and the diode lasers are polarized horizontally and vertically respectively.

First, a total of intensity of 3 mW/cm^2 illuminated the sample. In these conditions, we have not observed any diffraction beam for each dye concentration (1%, 2% and 5%) and for each thickness (20 and 30 μm).

The best results were found with an intensity of 10 mw/cm^2. The values are given in table 1 and 2.

TABLE 1. Diffraction efficiency η for A670 dye doped PMMA films thickness 20 μm

concentration	η (%)
1	/
2	0.4
5	0.5

TABLE 2. Diffraction efficiency η for A670 dye doped PMMA films thickness 30 μm

concentration	η (%)
1	/
2	0.6
5	0.65

The diffraction efficiencies measured are weak, one can observe an effect of dye concentration and thickness. For each thickness, no diffracted beam was obtained at a concentration of 1%.
If we compare the efficiencies in AC matrix, (concentration 5% and thickness 20 μm), we have obtained the same result, namely, a diffraction efficiency of 0.4%. We have even increased the thickness up to 120 μm and the diffraction efficiency has not exceed 0.6%.

4. Two-Photon Process

In optical data storage, an important area currently pursued is two-photon process because it provides a means for direct writing and access information in 3D space.
A molecule illuminated by UV or visible radiation is promoted to an excited state equal to the energy of the photon. An equivalent excited energy state can be populated by the simultaneous absorption of two photons.
To illustrate that process, Rentzepis and Dvornikov[4] wrote and read information in 3D memory device which contains photochromic molecules and especially spiropyran compounds.
To write information with spiropyran as recording medium, it requires excitation in the UV region of the spectrum, i.e., between 260 nm and 355 nm. Excitation to this state is provided by absorption of two-photon, either a 1064 nm photon, and a 532 nm photon which is equivalent to a 355 nm photon, or two 532 nm photon, corresponding to a 266 nm photon energy. Indeed 266 nm and 355 nm are the wavelengths of the absorption band of the spiropyran molecule.
When the 1064 nm beam is combined with a 532 nm beam, excitation of the molecule to the first excited state at 355 nm is possible. The written volume appears colored because the read form absorbs in the 550 nm region. The information can be stored not only a bit at a time but also all the information contained in a 2D page can be stored simultaneously.
The procedure to read the information written within the volume of the memory is similar to the write cycle except that the reading form absorbs at longer wavelengths than the written form so one or both laser beam must be at longer wavelengths than the ones used for writing. The "reading" is based on fluorescence. Indeed after the written form is excited by two-photon absorption, the molecule emits fluorescence. Erasing the information may be achieved either by increasing the temperature of the memory device or via irradiation by infrared light.

5. Discussion

By means of two approaches (photochemistry and holography), photochromic compounds fulgide ans spiropyran were characterized.

Spiropyrans molecules shows important thickness and writing intensity effect. For films 12μm and 50μm thick, the diffraction efficiency obtained is respectively 10.5% and 0.43 %. Angular selectivity $\Delta\theta$ was also measured for thin and thick PMMA films. The results obtained were 6.3° and 4.8° for respectively films 14μm and 250μm thick.

For fulgides, in PMMA matrix, the coloring and bleaching processes follow entirely first-order reaction with rate constants: $k_{UV}=1.2\times10^{-3}s^{-1}$ and $k_{VIS}=11.1\times10^{-3}s^{-1}$: bleaching process is much faster. In PVK matrix, the reaction order is identical and for the rate constants are of the same order.

Concerning the phenomenon of photochemical fatigue resistance it becomes evident that, PVK is not suitable for the recording of W/R/E cycles. As for the other matrices studied, photochemical losses are evaluated around 10% after 10 UV/Vis irradiation cycles. For CA, we have even observed stabilization after 7 cycles. However one will have to increase the number of exposition cycles to confirm this fact.

For the holographic study, diffraction efficiencies obtained in PMMA are weak (η=0.6% with a concentration of 5% and a thickness of 30μm). These facts show an effect of the concentration and thickness.

REFERENCES

1. H. Dürr, and H. Bouas-Laurent (eds), *Photochromism : Molecules and Systems*, p 5, Elsevier, Amsterdam, 1990.
2. A. Tomoda, A. Kaneko, H. Tsuboi, and R. Matsushima, "Photochromism of heterocyclic fulgides", Bull. Chem. Soc. Jpn., **65**, p. 1262-1267, 1992.
3. Y. Yokoyama, T. Yamane, and Y. Kurita, "Photochromism of a protonated 5-Dimethylaminoindolylfulgide : a model of a non-destructive readout for a photon mode optical memory", J. Chem. Soc., Chem. Commun., **24**, pp 1722-1724, 1991
4. A. S. Dvornikov, and P. M. Rentzepis, "Reaction kinetics of photochromic materials and their application to 3D optical memories", *Polymers in optics : physics, chemistry, and applications*, R. A. Lessard and W. F. Frank, **CR63**, pp. 239-261, SPIE, Bellingham, 1996.
5. R. A. Lessard, F. Ghailane, and G. Manivannan, "Photochromism and its applications in real-time holography", *Photoactive organic materials, science and application*, F. Kajar, V. M. Agranovich and C. Y.-C, Lee, **9**, pp. 325-341, Kluwer Academic, Dordrecht, 1996.
6. S. S. Xue. G. Manivannan, and R. A. Lessard, "Holographic and spectroscopic characterization of spiropyran doped polymethylmethacrylate films", Thin. Sol. Film, **253**, pp. 228-232, 1994.
7 C Lafond, R. A. Lessard, M. Bolte, and I. Petkov, "Characterization of dye-doped PMMA/PVK films as recording materials", *Proc. SPIE,* **3417**, pp. 216-227, 1998.
8 Y. Yokoyama, H. Hayata, H. Ito, and Y. Kurita, "Photochromism of a furylfulgide, 2-[1-(2,5-dimethyl-3-furyl)ethylidene]-3-isopropylidene succinic anhydride in solvents and polymer films", Bull. Chem. Soc. Jpn., **63**, pp. 1607-1610, 1990.
9 T. Tsuyioka, M. Kume, and. M Irie, "Photochromic reactions of a diarylethene derivative in polymer matrices", J. Photochem. Photobiol. A: Chem., **104**, pp. 203-206, 1997.
10 A. Kaneko, A. Tomoda, M. Ishizuka, H. Suzuky, and. R. Matsushima, "Photochemical fatigue resistances and thermal stabilities of heterocyclic fulgides in PMMA films", Bull. Chem. Soc. Jpn. **61**, pp. 3569-3573, 1988.
11 M. G. Moharam, T. K. Gaylord, and R. Magnusson, "Criteria for Bragg regime diffraction by phase grating", Opt. Commun., **32**, pp. 14-18, 1980.
12 H. Kogelnik, "Coupled wave theory for thick hologram grating", Bell Syst. Tech. J., **48**, pp 2909-2947, 1969.

Real-time diffraction efficiency in lithium niobate at 532 nm

P. Jayanth* and Sia Yii Luen
Electronics & Computer Engineering Department
Ngee Ann Polytechnic
535 Clementi Road, Singapore 599 489

Abstract

The formation of dynamic holograms in iron doped lithium niobate is studied with the 532 nm wavelength diode pumped solid state laser. It is observed that these holograms had higher diffraction efficiency in comparison with the holograms written with 488 nm. These holograms are then dynamically reconstructed with either the 532 nm or 633 nm. The choice of writing the holograms with the 532 nm and reconstructing with 633 nm leads to a whole lot of applications in real-time holographic interferometry and data storage is also discussed.

Keywords: Photorefractive crystals, Phase gratings, Holography, Photonics and Nonlinear optics

1. Introduction

Photorefractive effect (PR) is a nonlinear optical process that modifies the refraction properties of an electro-optic material when the material is subjected to nonuniform illumination. The PR owes its origin to the Pockels effect. The PR effect is explained as follows: a) Light causes migration of charges in the medium. b) the separation of charges produces a strong electrostatic field, of the order of 10^5 V/m. c) The electrostatic field causes a change in the refractive index of the crystal by the Pockels effect. The PR effect is a well-studied effect[1,2] and is important because it gives rise to large and saturated optical nonlinearity. It is known that the steady-state PR index change is independent of the total intensity of the incident beams, and instead depends on their relative intensity, but the speed of PR effect increases as the total intensity is increased.

The photorefractive effect is a well studied effect and is important because it gives rise to large and saturated optical nonlinearity. The formation of photorefractive index grating is shifted by $\pi/2$ with respect to in the interference of two optical waves. The photoexcited charge carriers diffuse or drift into the conduction band under the influence of the electric field from the brighter regions to the darker regions of the material. From Gauss's law this displacement of charges must in turn give rise to an electric field that opposes the movement of charges. This space-charge field is then responsible for modification of the refractive index of the material via the electro-optic interaction. The build up of the space-charge field will continue until dynamic equilibrium is reached. The rate of this process alone will be dependent on the overall intensity. In other words, the steady-state photorefractive index change is independent of the total intensity of the incident beams, and instead depends on their relative intensity. But the speed of photorefractive effect increases as the total intensity is increased.

One interesting property that has been observed in the formation of the refractive index-grating (phase grating) is that it is $\pi/2$ shifted with respect to the incident interference pattern. This nonlocal nature of the photorefractive effect is a consequence of Poisson's equation, $\nabla.\mathbf{E} = \rho/\varepsilon\varepsilon_o$. If, for example, the incident interference pattern is $I(x) \propto \cos(\mathbf{K}.\mathbf{x})$ then the charge density $\rho(x) \propto -\cos(\mathbf{K}.\mathbf{x})$ will create an electrostatic field $\mathbf{E(x)} \propto -\sin(\mathbf{K}.\mathbf{x})$. The electro static field $\mathbf{E(x)}$ is observed to by $\pi/2$ shifted relative to $I(\mathbf{x})$. The refractive index change, $\Delta n(\mathbf{x}) \propto -\mathbf{E(x)}$, is therefore $\pi/2$ shifted relative to the intensity pattern.

In two beam coupling (TBC) configuration, two beams from a laser source interfere in the volume of the PR crystal. The interference pattern causes nonlocalized $\pi/2$ shifted distribution of charges forming volume gratings in the crystal. These phase gratings are used for hologram storage in PR crystal. It has been recognized[3] that the volume nature of thick holograms permits the interference of an incident light beam with its own diffracted

* Correspondence: Email:puj@np.edu.sg; Telephone: (65)460 8581, Fax: (65) 467 1730

Part of the SPIE Conference on Advanced Photonic Sensors and Applications
Singapore • November–December 1999
SPIE Vol. 3897 • 0277-786X/99/$10.00

beam inside the recording medium. The resulting growth and decay characteristics of the phase gratings are explained by the dynamic theory developed by Ninomiya[4]. Holograms storage by TBC scheme had already been studied[1,5].

One of the reasons why photorefractive materials form a class different from other nonlinear materials is the formation of the $\pi/2$ shifted phase grating. The overall $\pi/2$shift between the illuminating intensity and refractive grating makes energy exchange in two wave mixing an intrinsic property of photorefractive materials. One of the waves can completely deplete the other wave and, practice, the gain or loss in energy of the waves is determined by the direction of the phase shift of the grating ($\pm\pi/2$). The dynamic read, write and erasure properties of the photorefractive crystals enable them to be used in real-time holographic applications. A hologram (phase gratings) is the record of the interference of an object wave (arbitrary wavefront) and a reference wave (a plane wave is chosen for convenience). The holograms can be either thermally or electrically fixed depending on the type of the photorefractive media.

This paper deals with the experimental investigations of the dynamic formation and erasure of the phase gratings formed in iron doped lithium niobate crystal using two lasers, viz. 532 nm and 633nm. The formation of the phase grating is carried out with the 532 nm laser while the erasure is carried out with both the 532 nm and 633 nm lasers. The diffraction efficiency at these two wavelengths is compared with the diffraction efficiency at 488 nm.

2. Experiment

Figure 1 shows the experimental setup. The laser beam from the diode pumped solid state (DPSS) YAG laser operating at 532 nm and with a overall power level of 500 mW is divided into two beams of equal intensity I_1 and I_2 at the beamsplitter. These two beams are made to interact in the iron doped lithium niobate crystal at an angle of 60° between them. An interference pattern is recorded in the $LiNbO_3$:Fe crystal. The formation and the decay of the phase grating is recorded by the detector interfaced to a digital oscilloscope. Initially both the writing beams I_1 and I_2 are allowed to interfere in the crystal. The light interference pattern causes the electrons to migrate from brighter region to darker regions. This leads to the formation of a phase grating in the volume of the crystal. The formation of the phase grating is recorded by intermittently blocking the writing beam, I_2. On blocking I_2, the first beam the second beam I_1 behaves as the reading beam. I_1 is diffracted from the phase grating formed in the crystal. The diffracted beam travels in the direction of I_2 to the detector. This beam is $I_1^{'}$ as shown in figure1. The beam $I_1^{'}$ is also known as the reconstructed beam and is a measure of the diffraction efficiency of the phase grating formed in the crystal. By monitoring $I_1^{'}$ the dynamic growth or formation of the phase grating and the dynamic erasure or decay of the phase grating can be studied. As mentioned earlier, to observe the dynamic growth of the phase grating the I_2 beam is intermittently blocked with the help of a mechanical chopper. While the dynamic erasure of the phase grating is monitored by continuously blocking the I_2 beam. Here the beam I_1 reads the phase grating as well as begins to dynamically erase it.

Figure 2 shows the dynamic formation of the phase grating using 532 nm and overall intensity of $I_1 + I_2 = 500$ mW. The behaviour of the growth of the phase grating is of an exponential type. It gradually increases to reach a maximum after which time it saturates. At this moment no further increase in the intensity of the reconstructed beam $I_1^{'}$ is observed. The growth of the of the phase grating is about 15 to 25 seconds to reach a value of 75% of the maximum value. The $LiNbO_3$:Fe crystal is exposed to one minute before determining the dynamic erasure time. During the dynamic erasure the I_2 beam is completely blocked and I_1 beam is allowed to illuminate the phase grating. The I_1 beam continuously erases the phase grating. Figure 2 shows the dynamic erasure with the intensity of I_1 being equal to 250 mW. It takes about 1 to 2 minutes to erase the phase grating to a value of 25% of the maximum. That is, this is the time taken for which the phase grating can be considered to have a very low diffraction efficiency of 1%. The y-axis is the normalized intensity in the power range of milliwatts.

The experiment of formation of the phase grating in $LiNbO_3$:Fe crystal is repeated. The crystal is exposed to the interference pattern at 532 nm for 60 seconds. The beam I_2 is blocked completely. The intensity of the beam I_1 is reduced to 4 mW with the help of the variable attenuator. The dynamic erasure is observed and is as shown in figure 3. Since the intensity of I_1 is decreased it takes longer time for the phase grating to decay and reach a value of 25% of the maximum saturation intensity value. This decay time is observed to be 5 to 10 minutes. The phase grating formation is once again repeated by exposing the crystal for 60 seconds with the laser wavelength of 532 nm and at an intensity of 500 mW. Now, both beams I_1 and I_2 are blocked. The phase grating is illuminated with a He-Ne laser (λ=633nm). The 633 nm laser beam is divided into two beams of equal intensity at the beamsplitter such that each of the beams copropagate with I_1 and I_2. The chopper is made use of to block

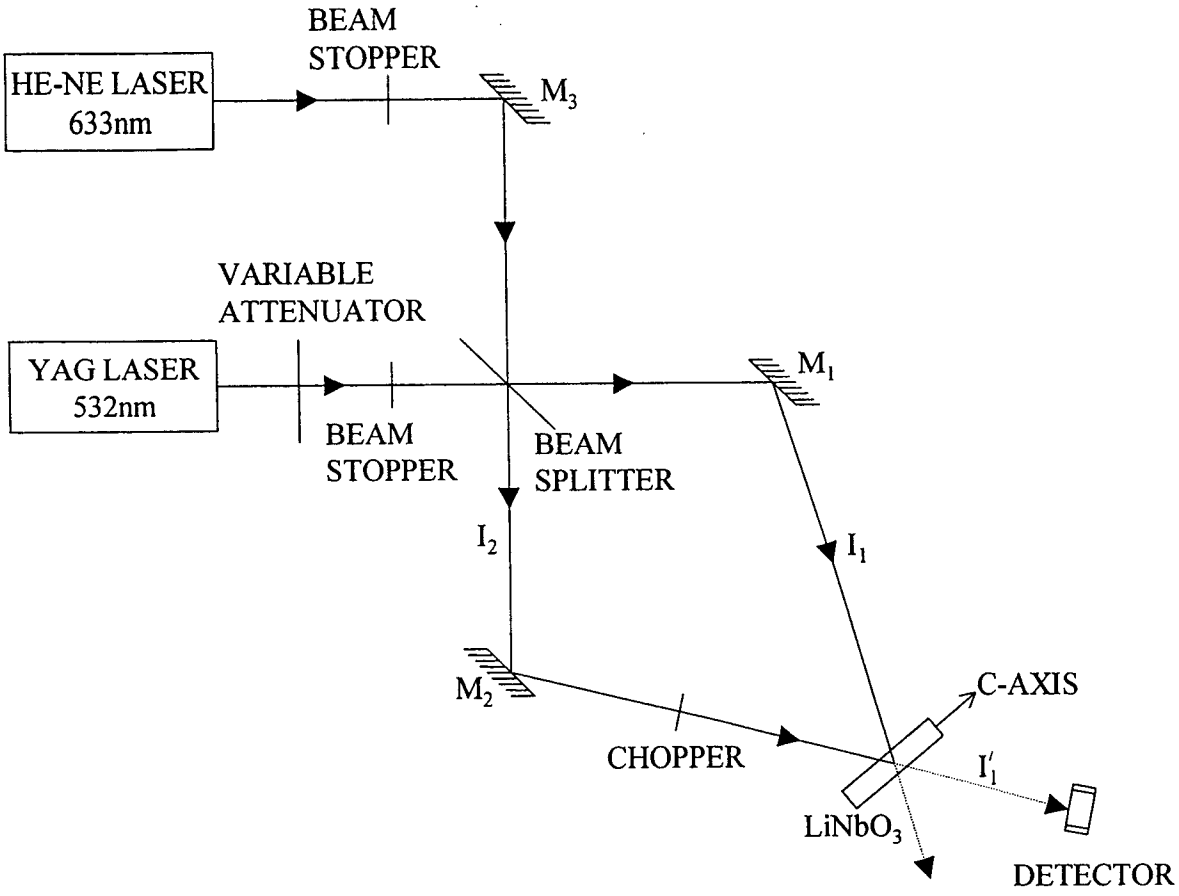


Figure 1. The experimental geometry to observe the dynamic growth and decay of the phase grating in iron doped Lithium Niobate crystal is shown. The 532 nm beam from the diode pumped solid state YAG laser is split into two beams at the beamsplitter. These two beams interfere in the crystal to form a phase grating in the crystal. The beam I_2 is intermittently blocked by the use of the chopper and the beam I_1 is diffracted from the phase grating to form I_1' beam, which is measured by the detector. The dynamic erasure of the phase grating is observed by completely blocking the I_2 beam and allowing the I_1 beam to diffract from the phase grating that is recorded in the crystal. The reconstruction of the hologram is also carried out at the 633 nm. The He-Ne laser is turned on while the YAG laser is switched off in and the I_2 beam is blocked in order to observe the dynamic erasure at 633 nm.

the He-Ne laser copropagating in the direction of I_2. Thus only the beam copropagating in the direction of I_1 is diffracted from the phase grating. The diffracted beam reconstructs the beam I_1' (λ=633nm). The newly reconstructed I_1' is spatially shifted with the previous I_1' by a small angle, which is governed by the Bragg's law ($2d\sin\theta=n\lambda$). The newly reconstructed I_1' beam is monitored. Figure 3 shows the dynamic erasure of the phase grating with the intensity of the He-Ne laser beam equal to 4 mW. The time for which the phase grating is erased to 25% of its full growth is observed to be about 8 to 9 hrs.

Thus it is observed that in $LiNbO_3$:Fe crystal the formation time is smaller than the dynamic decay times of the phase grating. Holograms can be recorded in $LiNbO_3$:Fe crystal in a few seconds, while its dynamic erasure during reconstruction is over several minutes. Of course, this depends on the intensity of the writing beams and reading beams. If the intensity of the reading beams or reconstruction beams is decreased then the holograms (phase grating) remain without being completely erased for a several tens of minutes. Our experimental observation also led us to believe that the writing beams and reading beams can be of different wavelengths. In the last experiment the writing beams was with λ=532 nm while the reconstruction beam was with λ=633 nm. It

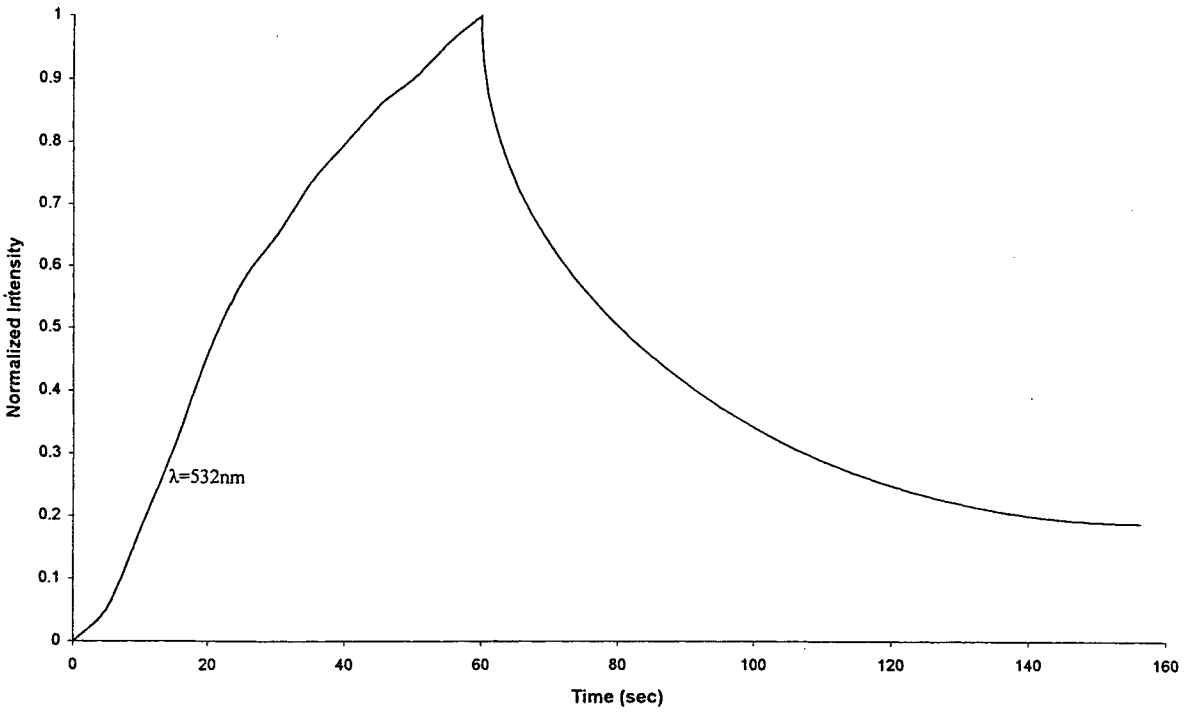


Figure 2: The growth and decay of the phase grating at 532 nm is shown. The overall intensity of the YAG laser during the writing of the phase grating was 500 mW. The dynamic erasure or decay is carried out by blocking the I_2 beam. The I_1 beam diffracted from the phase grating is the reconstructed I_1' beam, which is monitored.

was observed that it took more than one full day to completely erase the hologram when the reconstruction beam used was λ=633nm at an intensity of 4 mW in comparison to erasure time of a few hours when the wavelength λ=532nm was used at the same intensity. Thus it can be seen from figure 3 that for the same diffraction efficiency of the hologram written with 532 nm the erasure times with 633 nm is longer than with that of 532 nm. One of the applications of such a scheme can be used in real-time holographic interferometers. The hologram of a stationary object can be recorded with 532 nm. The reconstruction can be carried out with 633 nm. Now if the object is subjected to a small displacement then a far-field interference pattern is observed on the screen that is placed in the position of the detector. The interference pattern is a measure of the displacement of the object. Thus this scheme can be extended to several real-time holographic applications and data storage devices. Although permanent fixing of the holograms in $LiNbO_3$:Fe crystal by thermal means is also possible our scheme discusses the real-time applications and reusability of the crystal.

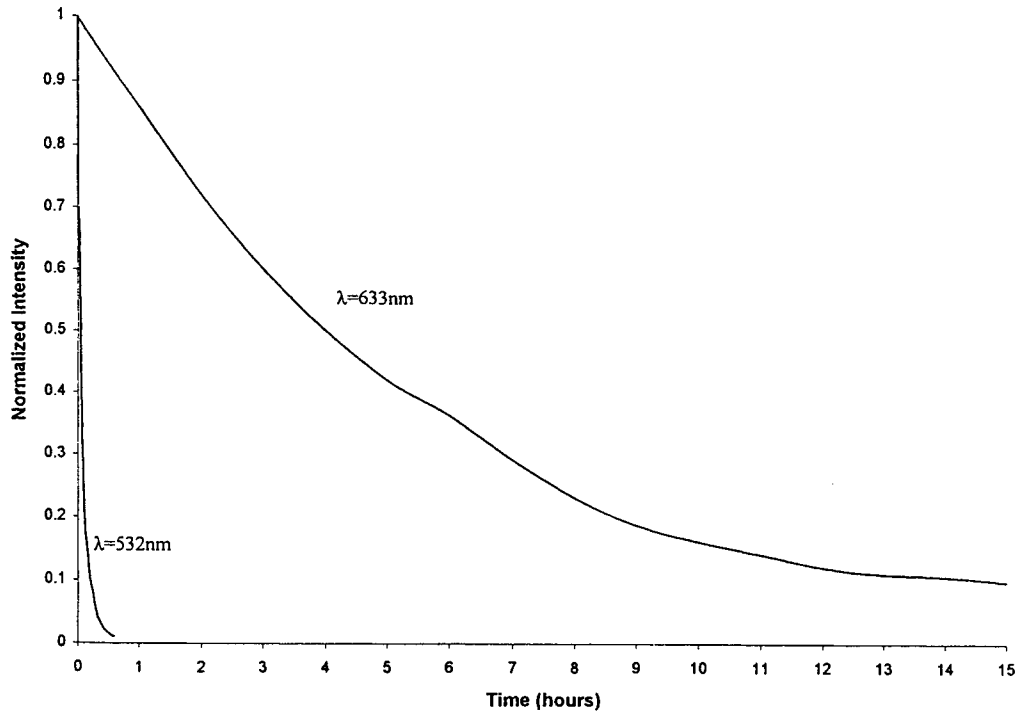


Figure 3: The overlay of the decay of the phase grating with the reconstruction beams of wavelengths 532 nm and 633 nm are shown. Each of these beams has an intensity of 4 mW. It is observed that although the reconstruction beams have the same intensity, the phase grating takes longer time to decay with λ = 633 nm beam in comparison with the λ = 532 nm beam.

Although the same intensity beams of 532 nm and 633 nm are used to reconstruct the hologram, it was observed that the 633 nm beam erased the phase grating at a much slower rate than that with the 532 nm. The explanation to this is because the quantum efficiency of the $LiNbO_3$:Fe crystal varies with different wavelengths. The relation of quantum efficiency with that of the wavelength is that quantum efficiency decreases linearly for longer wavelengths[6]. Hence the shorter wavelengths can erase the phase grating at a rate faster than longer wavelengths. The distortion of the reconstructed image on using a read beam of different wavelength is a subject of our further investigation.

The crystal is oriented such that the c-axis of the crystal is perpendicular to the to the angle bisector. The concentration of iron is 0.02% mol weight and the dimension of the crystal is 20 x 20 x 3 mm. The DPSS laser was built by Coherent Inc, and has a coherence length of 150 m for short-term time interval of 2 μs. A Melles-Griot He-Ne laser with coherence length of 13.5 cm was used. Because of the large coherence length of the 532 nm laser it was observed that the diffraction efficiency of phase grating was nearly 55% when compared to the diffraction efficiency[7] on using a 488 nm Argon laser which was nearly 20%. The results with 488nm (Argon laser) are not indicated in this paper.

3. Theory

In this section we will calculate the efficiency of the phase grating formed in the volume of the photorefractive crystal. The phase grating is formed by the interference of two beams whose intensities are I_1 and I_2. Let the two beam coupling field interacting in the medium be given as

$$\underline{E}(\underline{r}) = \underline{A}_1(r_1)e^{-i\underline{k}_1\cdot\underline{r}} + \underline{A}_2(r_2)e^{-i\underline{k}_2\cdot\underline{r}} + c.c \qquad (1)$$

The two waves with complex amplitudes $\underline{A}_i$ representing the envelope of the optical field, propagate in a medium and the resultant spatially periodic stationary index distribution is given as

$$n(\underline{r}) = n_0 + \frac{n_1}{2}e^{-(i\underline{K}\cdot\underline{r})+\phi} + \frac{n_1}{2}e^{(i\underline{K}\cdot\underline{r})+\phi} \qquad (2)$$

Where ϕ is the phase factor due to the medium and $\underline{K} = \underline{k}_1 + \underline{k}_2$ is the grating vector. The amplitudes are assumed to obey the slowly varying envelope approximation and the paraxial wave equation is given as

$$\nabla^2 \underline{E} + \omega^2 \mu\varepsilon(\underline{r})\underline{E} = 0 \qquad (3)$$

The dielectric constant ε(r), now has a contribution from the grating

$$\varepsilon(r) = \varepsilon_0 n^2(r) \cong \varepsilon_0\left[n_0^2 + (n_0 n_1 e^{-i(\underline{K}\cdot\underline{r}+\phi)} + c.c)\right] \qquad (4)$$

Substituting equations (1), (2) in (3), and neglecting the second order spatial derivative in comparison to the optical wavevector times the first derivative, ie., $d^2A/dr^2 << k\ dA/dr$. The second order differential equation reduces to the two first order equations given as

$$\cos\theta_1 \frac{dA_1}{dz} = -\frac{\alpha}{2}A_1 + i\frac{\pi n_1}{\lambda}e^{i\phi}A_2 e^{i(\underline{k}_1-\underline{k}_2+\underline{K})\cdot\underline{r}} \qquad (5.a)$$

$$\cos\theta_2 \frac{dA_2}{dz} = -\frac{\alpha}{2}A_2 + i\frac{\pi n_1}{\lambda}e^{-i\phi}A_1 e^{-i(\underline{k}_1-\underline{k}_2+\underline{K})\cdot\underline{r}} \qquad (5.b)$$

Where the loss terms $-(\alpha/2)A_{1,2}$ were added phenomenologically to account for absorption, $\lambda = 2\pi/(\omega\sqrt{\mu\varepsilon_0 n_0^2})$, and $\theta_1 + \theta_2$ is the angle between $\underline{k}_1$ and $\underline{k}_2$. Expressing the amplitudes in terms of magnitudes and phases by using the definition $A_j = \sqrt{I_j}e^{i\phi_j}$ and with Bragg condition being satisfied ie., $\underline{k}_1 - \underline{k}_2 + \underline{K} = 0$, the coupled wave equation (5) are now given as

$$\cos\theta_1 \frac{dI_1}{dz} = -\alpha I_1 + \frac{2\pi n_1}{\lambda}\sqrt{I_1 I_2}\sin(\psi) \qquad (6.a)$$

$$\cos\theta_2 \frac{dI_2}{dz} = -\alpha I_2 - \frac{2\pi n_1}{\lambda}\sqrt{I_1 I_2}\sin(\psi) \qquad (6.b)$$

The coupling on a point r depends on the local phase $\psi = (\phi_1 - \phi_2 + \phi)$. If the phase φ is zero, no power exchange takes place and if $\psi = \pm\pi/2$, the energy exchange is maximum. The diffraction efficiency of the fixed volume grating obtained by Kogelnik[8] for $\theta_1 = \theta_2$ is given as

$$\eta = \frac{I_2(l)}{I_1(0)} = \exp(\frac{-\alpha l}{\cos\theta_1}) \sin^2(\frac{\pi n_1 l}{\lambda \cos\theta_1}) \qquad (7)$$

4. Conclusion

The formation of dynamic phase gratings in iron doped lithium niobate was studied with the 532 nm wavelength diode pumped solid state laser. It was observed that these holograms had higher diffraction efficiency in comparison with the holograms written with 488 nm. These holograms are then dynamically reconstructed with either the 532 nm or 633 nm at equal intensity level of 2 mW. It was observed that the dynamic erasure with 633 nm is much slower than that with the 532 nm. Finally, the choice of writing the holograms with the 532 nm and reconstructing with 633 nm leading to a whole lot of applications in data storage which was also discussed.

Hence the shorter wavelengths can erase the phase grating at a rate faster than longer wavelengths. But the holograms reconstructed with a beam whose wavelength is different from the writing beams are observed to be either spatially separated or magnified. The distortion of the reconstructed image on using a read beam of different wavelength is a subject of our further investigation.

Acknowledgment

Thanks are due to Andrew T Sabaratnam, Centre for Optoelectronics, ECE Department and Lim Choo Min, Head, ECE Department, Ngee Ann Polytechnic, Singapore for constant support and encouragement.

References

1. Gunter P., and Huignard J.P., (Eds.), "Photorefractive materials and their applications Vol.- I and II", Springer Verlag, (1989).
2. Hall T.J. Jaura R., Connors L.M., and Foote P.D., Prog. Quant. Elect., **10**, 77, (1985).
3. Staebler D.L., and Amodei J.J., J. Appl. Phys. Lett., **43,** 1042, (1972).
4. Ninomiya Y., J.Opt.Soc. Am., **63**, 1124, (1973).
5. P.Jayanth, R.Krishna Mohan, C.K.Subramanian and P.S.Narayanan, "Optical Phase Conjugation by Two Wave Mixing in Iron doped Lithium Niobate", Opt. Commun., **116**, 175-178, (1995).
6. Gunter Phys. Rep. **93**, 199, (1982).
7. P.Jayanth and C.K. Subramanian, "Photorefractive effect in the Lithium Niobate crystals", J. of I.I.Sc., **76**,565-570, (1996).
8. Kogelnik H., Bell Syst. Tech., J., **48**, 2909, (1969).

Metal doped photosensitive materials :

Correlation between photoreactivity and holography

Michèle Bolte and Christine Pizzocaro

Laboratoire de Photochimie UMR 6505 CNRS-Université Blaise Pascal (Clermont-Ferrand)

F-63177 Aubière cedex (France)[1]

Roger A. Lessard

Centre Optique Photonique et Laser

Université Laval, Québec G1K 7PA (Canada)[2]

ABSTRACT

Dichromated polyacrylic acid (DCPAA) and dichromated polyvinylalcohol (DCPVA) have been investigated with a double approach : photochemical behavior and holographic results. The study focused on two particular systems : the effect of the addition of nitrilotriacetic acid (NTA) or dimethylformamide (DMF) to the photosensitive material. From the photochemical mechanism it was possible to assess that the complexing ability of both compounds plays a major role in the optimization of the holographic properties of the material. On the contrary, the presence of an electron donor like NTA in DCPAA is less favorable. A close correlation seems to be present between the photochemical behavior and the results obtained in holography.

Keywords : chromium doped system, photochemistry/holography, film processing, optimization

1. INTRODUCTION

The photoreactivity of transition complexes has been often used to elaborate or modify polymers in the domain of photography and more recently in holography. Actually the absorption of light by metallic complexes can induce polymerization and/or crosslinking of the material. In addition, transition complexes exist under different stable valence states and present absorption domain that makes possible the photoreaction upon excitation by visible light. Dichromate anion and iron (III) salt have been successfully used to sensitize the modifications of polymeric materials such as polyvinylalcohol, polyacrylic acid and gelatin. On exposure to appropriate light sources, the sensitizer undergoes a reduction of the metallic center whereas the oxidation causes the crosslinking of the polymeric matrix. Dichromated polymer materials such as dichromated poly(vinylalcohol) (DCPVA) and dichromated polyacrylic acid (DCPAA) have been employed to realize various applications such as transmission holograms (1-3), reflexion holograms (4), polymer waveguides (5) and computer generated holograms (6). Our interest in exploiting the photosensitive nature of DCPAA and DCPVA films led us to the necessity of investigating the primary photoprocess of chromium (VI) in the polymer film and correlating it to the properties of the resulting holograms. The effect of the presence of different additives able to improve the quality of the hologram performances is also of interest and was investigated. For instance it was already reported that the addition of dimethylformamide (DMF) strongly increased the diffraction efficiency of a hologram recorded in DCPAA (7).

---

[1] Telephone : ++33 (0)4 73 40 71 71 Fax ++33 (0)4 73 40 77 00 Email : boltem@mailhost.univ-bpclermont.fr

[2] Telephone : ++1 418 656 3436 Fax : ++1 418 656 2623 Email : ralessard@phy.ulaval.ca

Part of the SPIE Conference on Advanced Photonic Sensors and Applications
Singapore • November–December 1999
SPIE Vol. 3897 • 0277-786X/99/$10.00

2. EXPERIMENTAL

2.1. Materials and fabrication of dichromated polymer films

Ammonium Dichromate $(NH_4)_2Cr_20_7$ was of commercial origin (FLUKA) and used without further purification. Polyacrylic acid (ALDRICH) was in aqueous solution (35% in weight, $\overline{M}_W$=250.000), Polyvinylalcohol was ALDRICH 98% hydrolized, $\overline{M} \approx 170.000$. Dimethylformamide was purchased from ALDRICH (spectrometric grade). Nitrilotriacetic acid was an OMEGA product (99%). The preparation of films was essentially the same as reported earlier (8).

2.2. Apparatus and procedure

A high pressure mercury lamp with a Bausch and Lomb grating monochromator was used for irradiation. The light intensity was measured by ferrioxalate actinometry. UV visible spectrum were recorded on a CARY3 double beam spectrophotometer. ESR spectra were recorded on a BRUCKER ER200D spectrometer at 9.30 GHz with a modulation field of 100 Hz. A Xe Hanovia Lamp was used for irradiation in the ESR spectrometer cavity, the emission was filtered at λ>345 nm.
The holographic set up was described elsewhere (8).

3. BRIEF SURVEY ON CHROMIUM (VI) COMPLEX PHOTOCHEMISTRY

The absorption of a photon by dichromate (chromium (VI) species) $Cr_2O_7^{2-}$ leads to the formation of a short-lived electronically excited state $Cr_2O_7^{2-*}$. The excess of energy can be dissipated either physically or chemically. When the excitation takes place in a ligand to metal charge transfer transition, an electron located in an orbital of the ligand populates an empty orbital of chromium. In the resulting excited state, the metal is reduced and the organic ligand is oxidized (4) as a radical.
The quantum yield of the reaction that measures the efficiency of the photochemical reaction is defined by :

$$\Phi = \frac{\text{number of transformed molecules}}{\text{number of absorbed photons}} \text{ for the same period of time.}$$

The final chromium species is chromium (III) and what is generally calculated is the quantum yield of reduction of chromium (VI) into chromium (III). For some polymeric matrices like PVA or gelatin, the intermediate chromium (V) species is stable and the corresponding quantum yield $Cr_{VI \rightarrow V}$ can be calculated.

Remark : Chromium (VI), Cr^{6+}, represents a species whose electronic structure is in d°, it has lost 6 electrons and 3d orbitals are empty ; chromium (V), Cr^{5+}, and chromium (III), Cr^{3+}, have electronic structure with $3d^1$ and $3d^3$.

4. RESULTS AND DISCUSSION

4.1. Spectral features of dichromated films

Chromium (VI) in aqueous solution presents a UV visible spectrum depending on the pH. At pH<4, a maximum of optical density was observed at 350 nm with a shoulder at 430 nm whereas at pH>6 a maximum at 370 nm was observed and the shoulder was no longer present.
The addition of polyacrylic acid to the solution makes the pH decrease to pH<3 according to the concentration. On the contrary the addition of polyvinylalcohol did not affect the pH of the starting chromium solution (pH≈6).
Processing a film introduced some significant differences in the absorbing properties of chromium (VI) : namely for the molar extinction coefficient ε.

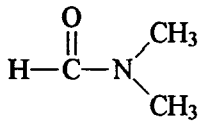
	DCPAA λ=350 nm	DCPVA λ=360 nm
$\varepsilon_{solution}$ (L.mol^{-1}.cm^{-1})	1600	2400
ε_{film} (L.mol^{-1}.cm^{-1})	1100	1800

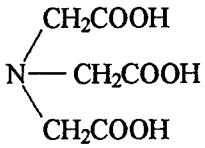
Processing a film causes a decrease in ε value, this phenomenon was already reported for dichromated gelatin (9). According to these results, it apears that processing a film strongly affects the photosensitive properties of the absorbing material.

4.2. Holographic results

Two sets of experiments were performed : holograms were recorded in photosensitive materials based on polymeric matrix of polyvinyl alcohol and polyacrylic acid. In both matrix two additives of different properties were used separately :

- dimethylformamide (DMF) and nitrilotriacetic acid (NTA)





Both compounds are complexing agents able to complex chromium (VI) but nitrilotriacetic acid is also known to be an electron donor through the lone pair of electron on the nitrogen atom.
When compounds like NTA or DMF were added to the starting solution of DCPVA or DCPAA, there were no detectable changes in the UV visible spectra neither in solution nor in the resulting film. But as a far as the "thermal" (in the dark) ageing is concerned, the two dichromated matrix behaved differently : the onset of an absorption at 575 nm due to chromium (III) was observed in DCPAA whereas the "thermal" reduction in DCPVA seemed to be of minor importance.
It is worth mentioning that the reduction of chromium also takes place in DCPAA alone but at a slower rate.

When DMF or NTA was added to DCPVA to record holograms, an increase roughly of the same order (≈20%) was observed for the value of the diffraction efficiency η (fig. 1) (10).

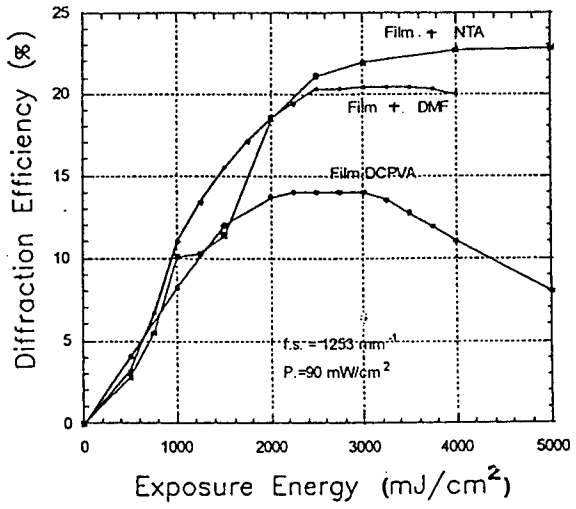


Fig. 1. Effect of DMF and NTA on the diffraction efficiency of holograms recorded in DCPVA.

In the case of DCPAA, an increase in the diffraction efficiency was observed again in the presence of DMF or NTA. However the effect is far less important for NTA (fig. 2) (11) :

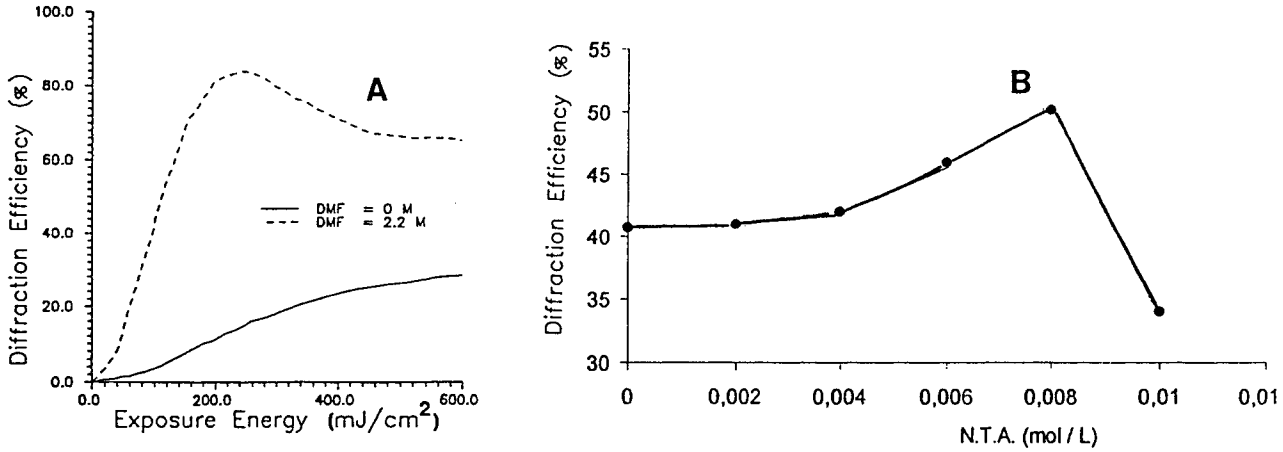


Fig. 2. Effect of DMF (A) and NTA (B) on the diffraction efficiency of holograms recorded in DCPAA.

All these results are in favor of a major role played by the complexing feature and more particularly that of dimethylformamide : the presence of DMF in a dichromated photosensitive material strongly improves the holographic performances of the material.
Accordingly, it was very important to try to understand the photochemical role of DMF and NTA in DCPAA or DCPVA films in order to confirm this hypothesis.

4.3. Photochemical reactions

The sequence of chromium (VI) reduction, the photochemical reaction occuring during the irradiation of dichromated photosensitive film, has been clearly established :

$$Cr^{6+} \xrightarrow[1\ \text{electron}]{h\nu} Cr^{5+} \xrightarrow[2\ \text{electrons}]{h\nu\ \text{and/or}\ \Delta} Cr^{3+}$$

It was also stated that the reduction into the final chromium (III) can occur thermally and/or photochemically in all investigated systems. It can be accelerated by using a developer.
The work performed on the dichromated systems DCPAA and DCPVA in the presence of DMF and NTA appears to back most of the results obtained in holography eventhough the study of some systems is still in progress.

4.3.1. DCPAA, DMF or NTA

The photochemical behavior of DCPAA/DMF films and the role of DMF were completely elucidated (12). DMF strongly increases both the rate of formation of chromium (V) and chromium (III). The previous photochemical investigation gave evidence for an acido-basic reaction between chromium (VI) and PAA enhanced by the presence of DMF. DMF involvement was due to its ability to complex chromium (VI) and to make the formation of chromium acrylate easier. After the first step yielding chromium (V), there was no longer any obvious role played by DMF: actually if the rate of chromium (V) formation was multiplied by approximately 30, the rate of chromium (III) was only 4 times higher.
The introduction of NTA into DCPAA film did not significantly affect the quantum yield of chromium (III) formation : a very slight decrease in the rate could even be observed. As far as chromium (V) was concerned, there was no accumulation in the polymeric matrix. It is reduced faster than it is produced as evidence by ESR spectroscopy upon irradiation (fig. 3).

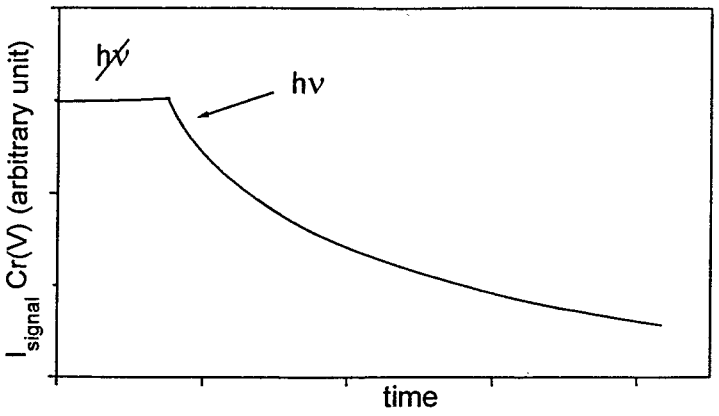
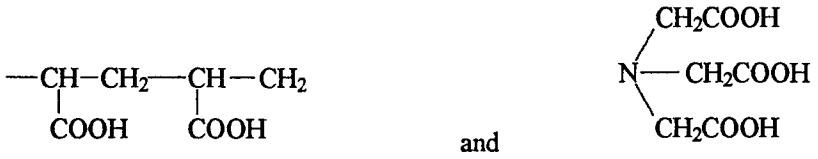


Fig. 3. Evolution of chromium (V) signal upon irradiation in DCPAA/NTA.

In this case, PAA and NTA are both able to complex chromium (VI) and to reduce it. They have similar chemical formulae and properties :



The competition between PAA and NTA to react with chromium (VI), complexation and then reduction, can account for the photochemical results. In the presence of DMF, only PAA presents reductive properties with respect to chromium (VI).
The competition between PAA and NTA was also confirmed by the strong effect of NTA concentration on the diffraction efficiency : it goes through a maximum and then drops down to a value lower than the starting η_0 value obtained in the absence of NTA.

NTA (mol. L^{-1})	η_{max} %
0	41
0.005	41
0.006	46
0.008	50
0.01	35

With DMF, even added at higher concentration (up to 2 mol. L^{-1}), no decrease was observed.

4.3.2. DCPVA, DMF or NTA

A DCPVA film presents a photochemical behavior quite different from that of DCPAA but similar to that of dichromated gelatin ; chromium (V), the primary chromium species resulting from the electron transfer from PVA to chromium (VI) is stabilized in the matrix.

The spectral evolution of a DCPVA film upon irradiation at 365 nm ($Cr_{(VI)}$ was the only absorbing species) is represented on figure 4 :

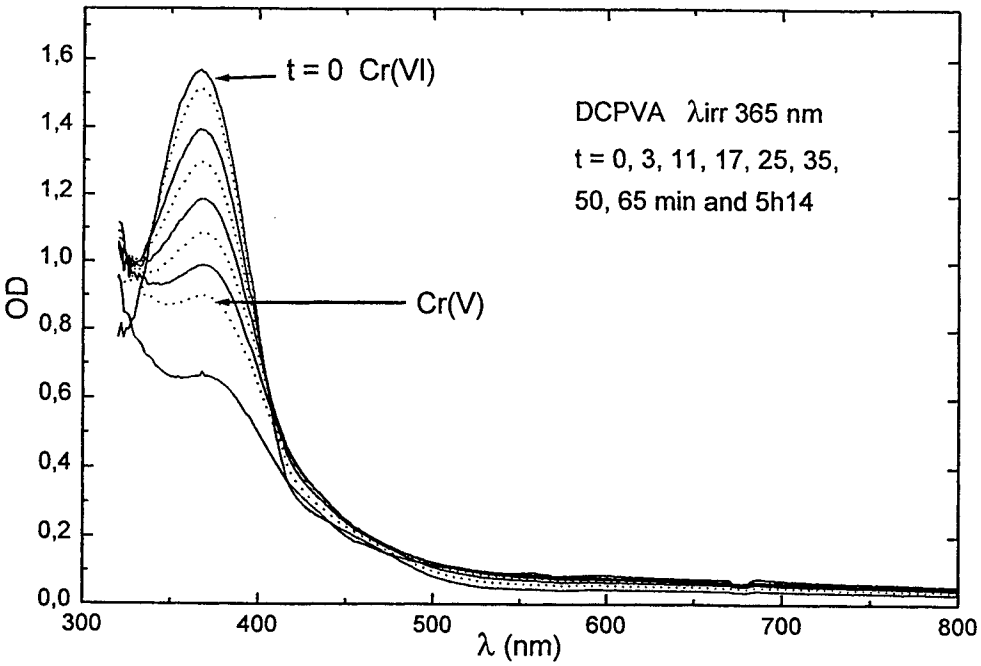


Fig. 4. Spectral evolution of a DCPVA film (λ_{exc}=365 nm).

A fast decrease with a subsequent very slow decrease was observed. The complete reduction of chromium (VI) into chromium (III) took days. An isosbestic point was present all along the fast period of the photochemical reaction, reflecting the one electron reduction of chromium (VI) into chromium (V). Accordingly, the UV visible spectrum obtained at the end of this process (65 min) can be assigned to pure chromium (V) in PVA matrix. The formation of chromium (V) with a maximum at 522 nm is represented on figure 5.

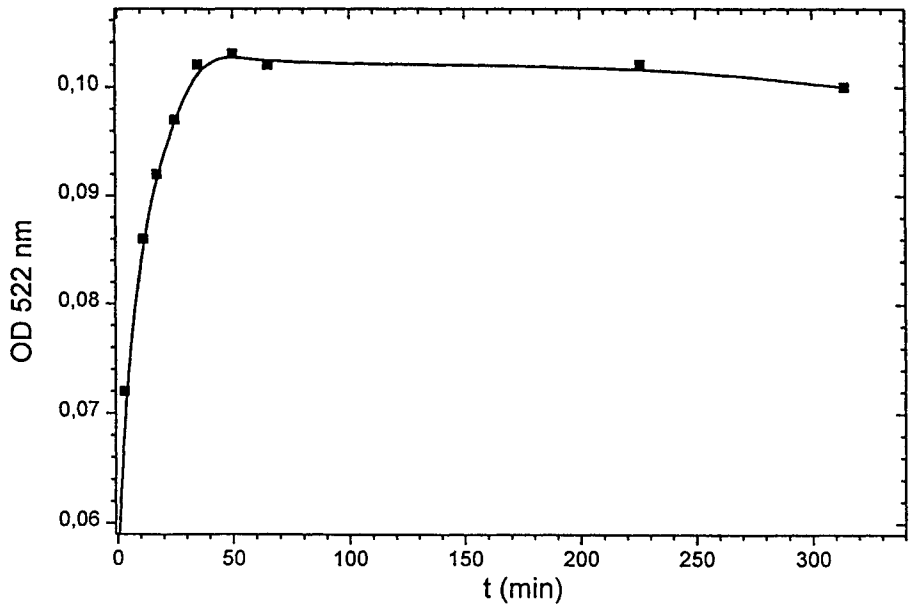
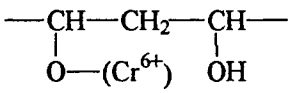


Fig. 5. Evolution of the optical density at 522 nm as a function of irradiation time.

The molar extinction coefficient of chromium (V) species was calculated and then the quantum yield of its formation :

	ε ($L.mol^{-1}cm^{-1}$)	$\Phi_{Cr(V)form}$
chromium (V) in PVA	at 365 nm : 1200 at 522 nm : 60	≈ 0.5

Similarly to what was obtained with gelatin, the Φ value of the first step of the reduction is very high and the transformation into final chromium (III) required days, photochemically and/or thermally (9).
The addition of NTA or DMF seems to only slightly affect the quantum yield of chromium (V) formation ; there is a slight increase but in every cases, the quantum yield is high $\approx$0.40 and comparable to that obtained with gelatin (9). The comparison between PVA and gelatin is also valid for holographic results : a previous work had shown that, with DCPVA, a post irradiation treatment made the diffraction efficiency strongly increase, from 20% up to 60% approximately (13).
The photochemical study of these systems is in progress : one of the major question being the nature of the reaction occuring between PVA and chromium (VI) : is it the formation of a chromate ester



as it was previously postulated (14) or some other reaction and then how additives like NTA or DMF can interfer in the process ? Some preliminary experiments with NTA have shown a strong effect on the diffraction efficiency : η value can even drop to 1 % when 0.02 M of NTA was added in DCPVA.

CONCLUSION

The complexing feature of an additive seems to play an important role in the properties of a dichromated material and contrary to what was postulated previously, the addition of a reductive compound in photosensitive material does not lead automatically to the improvement of the quality of the resulting hologram. It can even have a negative effect.
As the comparative work of the photochemical and holographic properties of dichromated systems proceeds, it clearly appears that a direct correlation exists between the two sets of results.

REFERENCES

1. R.A. Lessard, R. Changkakoti and G. Manivannan, "Metal ion doped polymer systems for real-time holographic recording", *Photopolymer Device Physics, Chemistry and Applications II, Proc. SPIE*, **1559**, PP. 438-448, 1991.
2. G. Manivannan, R. Changkakoti and R.A. Lessard, "Cr(VI) and Fe(III) doped polymer systems as real-time holographic recording materials", *Opt. Eng.*, **32**, pp. 671-676, 1993.
3. G. Manivannan, R. Changkakoti and R.A. Lessard, "Metal ion doped polymers as media for optical memories", *Polym. Advan. Techn.*, **4**, pp. 569-576, 1994.
4. R. Changkakoti, G. Manivannan, S. Ganguly and R.A. Lessard, "Fabrication and characterization of reflection holograms in dichromated poly(vinyl alcohol)", *Optical Memory and Neural Networks*, in press.
5. F. Trépanier, G. Manivannan, R. Changkakoti and R.A. Lessard, "Polymer waveguides : characterization through prism-film coupler method", *Can. J. Phys.*, **71**, pp. 423-428, 1993.
6. G. Manivannan, G. Lemelin, R. Changkakoti and R.A. Lessard, "Computer generated holograms on a metal ion doped polymer system : contact copying", *Appl. Opt.*, 33, pp. 3478-3481 (1994).

7. G. Manivannan and R.A. Lessard, "Trends in holographic recording materials", *Trends in Polymer Science*, **2**, pp. 282-290, 1994.
8. R.A. Lessard, G. Manivannan and R. Changkakoti, "Holographic multiplexing in polymer material: application in optical computing", *Optical Memory II. Proc. SPIE*, **2429**, pp. 1-11, 1994.
9. M. Bolte and C. Pizzocaro, "Photochemical behavior in dichromated gelatin", *Proc. SPIE*, **3347**, pp. 236-246, 1997.
10. J.L. Pereira and R.A Lessard, Rapport de stage de fin d'études, septembre 1998, COPL (Québec) - CUST (Clermont-Ferrand).
11. F. Michault and R.A. Lessard, Rapport de stage de fin d'études, septembre 1999, COPL (Québec) - CUST (Clermont-Ferrand).
12. C. Pizzocaro, R.A. Lessard and M. Bolte, "Charge transfer phenomena in holographic recording material : dichromated poly(acrylic acid), *Can. J. Chem.*, **76**, pp. 1746-1752, 1998).
13. R.A. Lessard, R. Changkakoti and G. Manivannan, "Polyvinyl alcohol based recording materials for holography and non linear optics", *Optical Memory and Neural Network, Proc. SPIE*, **1**, pp. 75-101, 1992.
14. H. Lee van Nice and R. Farlee, "A study of the photochemical reaction between polyvinylalcohol and aqueous chromate ion", *Polymer Engineering and Science*, **17**, pp. 359-365, 1977.

SESSION 2

Photonic Sensors Design and Fabrication

Very selective volume holograms for spatial and spectral filtering

Nadya Reinhand*[a], Irina Semenova[b], Alexander Popov[c]

[a]Institute of Mechanical Engineering Problems of the Russian Academy of Sciences,
V.O., Bolshoy pr., 61, St.Petersburg, 199178, Russia
[b]A.F.Ioffe Physical Technical Institute of the Russian Academy of Sciences,
26, Polytekhnicheskaya, St.Petersburg, 194021, Russia
[c]LUMEX Ltd., 19, Moskovsky pr., St.Petersburg, 198005, Russia

ABSTRACT

We discuss the application of narrowband holographic filters for aerosense technique including laser radars, communication devices, etc. It is wellknown that very thick volume holograms both transmission and reflective type being illuminated by light indicate very high angular (less than milliradian) and spectral (less than nanometer) selectivity. The variety of sensors require such selection. We report the experimental results of recording of selectors with nanometer bandwidth of spectral selectivity and milliradian bandwidth of angular selectivity recorded in photopolymer with diffusive amplification (PDA) of about millimeter thickness. We investigated the affect caused by introduction of holographic filter into receiving signal: the influence of filter selectivity property on the spatial frequency distribution of light passed through the atmosphere. We analysed the possibility to vary the shape of the output power spectral density (in order to reach superGaussian line shape) by means of hologram apodization.

Keywords: Spectral and angular selectivity, diffraction gratings, holographic filter, lidars, superGaussian spectral line.

1. INTRODUCTION

Variety of optical sensors require spectral and angular selection of the input or output laser beam. For instance, in airborne communication systems, in lidar technique, in telescopic observations the narrowband spectral filters are used in order to select the operating wavelength and suppress the unwanted part of spectrum. Commonly used optical filters include interference filters, interference polarisation filters, ruled gratings, fiber gratings, Fabry-Perot interferometers, and planar waveguide filters. All of them have their advantages and disadvantages, and the choice of the filter is determined by the particular task for which it will be used. Interference filters and ruled gratings possess rather low output efficiency for the selectivity bandwidth less than 1 nm, interference polarisation filter (in one wavelength output regime) is a large and heavy device, hardly amenable to miniaturisation. Fiber gratings can not be applied for wide aperture filtering. Fabry-Perot interferometers being an excellent narrowband filters are prone to wavelength drift. Planar waveguide filters suffer from high insertion losses and low extinction ratios.

The application of volume holographic gratings for spectral selection occurs very promising for this goal due to their high efficiency, spectral purity and rather low cost. The grating is recorded inside the photosensitive material by its illumination by interference pattern. Spectral selectivity is the inherent property of volume grating, the selectivity contour bandwidth vary in inverse proportion to the hologram thickness. In order to reach the values of spectral selectivity of about angstrom, the hologram thickness must be about millimeter. Usually the holographic photosensitive materials of millimeter thickness are associated with the variety of photorefractive crystals which were developed in recent years. Being the dynamic media, they can find applications in holographic memory systems, in real-time interferometry, etc. However the large range of applications require stationary media. The review of stationary holographic materials suitable for the recording of very

* Correspondence: Email: nadya@air.ipme.ru, Telephone: +7 812 3214765, Fax: +7 812 3214771

Part of the SPIE Conference on Advanced Photonic Sensors and Applications
Singapore • November–December 1999
SPIE Vol. 3897 • 0277-786X/99/$10.00

thick holograms was made in [1]. The list includes porous glass, silver glass, photopolymer with diffusive amplification (PDA), reoxan, and thick dichromated gelatin.

Angular and spectral selectivities of hologram are associated parameters [3]. Therefore any spectral selector carries out the selection of light spatial frequencies as well. In some applications this is a disturbing property that limits the incident and observation angles of the beam. However in some cases the angular selection can be useful in order to remove the unwanted scattered light, to clean the beam and to improve the signal quality in such a way. For instance, inhomogeneities in refractive index in atmosphere cause the refraction of beams passing through and the signal distortion. The improvement of the poor quality image obtained through the atmosphere can be carried out using digital or optical postprocessing [4]. The better approach is in application of an external control and feedback system introducing the dynamic correction of laser beam wavefront [5,6]. All of these systems are complicated and expensive ones. Our goal was to analyse the possibility of the spatial frequencies filtering of the laser beam passing through the atmosphere with turbulence by means of holographic spatial filter and to relate the affect of the holographic spectral filtering and the spatial frequencies selection in atmosphere.

Another goal was to analyse the possibility of formation of superGaussian beam by menas of holographic optical element. Such beams has advantages over regular Gaussian beams in several aerospace applications.

2. SELECTION CHARACTERISTICS

It was shown [2,3] that theoretically both reflective and transmission holograms can be applied for the recording of spectral and spatial frequencies selectors. In both cases we are working with the photosensitive medium of up to millimeter thickness (in order to reach up to angstrom selectivity). Basing on Kogelnik's couple wave theory [2] we calculated the dependence of bandwidth of the hologram selectivity contour $\Delta\lambda$ on its thickness T :

$$\Delta\lambda = \frac{2.7}{\pi} \cdot \frac{\sqrt{4n^2 - \lambda^2\gamma^2}}{T \cdot \gamma^2} \qquad (1)$$

for transmission holograms, and

$$\Delta\lambda = \frac{3.9\lambda}{\pi T\gamma} \qquad (2)$$

for reflection holograms, where n is the refractive index of the material, λ is the recording wavelength, and γ is the spatial frequency of the recording interferometric gratings. Figure 1 shows the obtained dependencies for the following parameters: n=1.52, γ_{trans}=1000 mm^{-1}, γ_{refl}=5000 mm^{-1}, λ=514 nm.

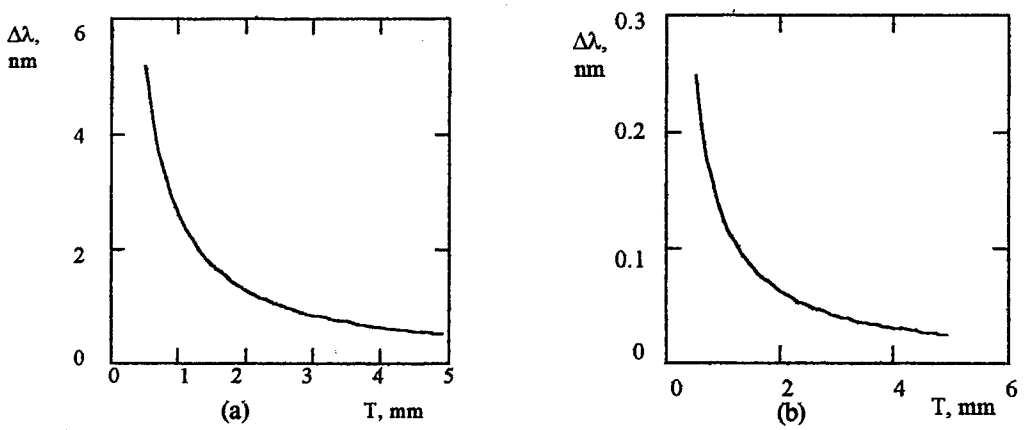


Figure 1. The obtained bandwidth of the spectral selectivity contour $\Delta\lambda$ for the hologram with the thickness of T for two types of holograms: (a) transmission and (b) reflection.

One can see from the figure that requirements on the thickness of holographic medium in order to record high efficient filters with the bandwidth of spectral selectivity of about 0.1 nm are rather hard to fulfil. For both types of holograms the required thickness comes to the value of millimeters. However, in general reflective holograms are more suited for the spectral selection. We reported [3] on the experimentally obtained reflective filters recorded in photopolymer with diffusive amplification (PDA). The reached values of spectral selectivity are about parts of nanometer (Figure 2). For the recording of holographic spectral selectors we used the optical scheme with opposite beams. The angle between the incident beam directions was about 140 degrees. The reflective holograms with the spatial frequency of the interferometric structure inside the photomaterial of about 4000 mm^{-1} were recorded using the photomaterials with 0.5-2 mm thickness. The obtained diffraction efficiency was from 80% to 93%. The experimentally obtained bandwidth of spectral selectivity was 0.2- 10 nm.

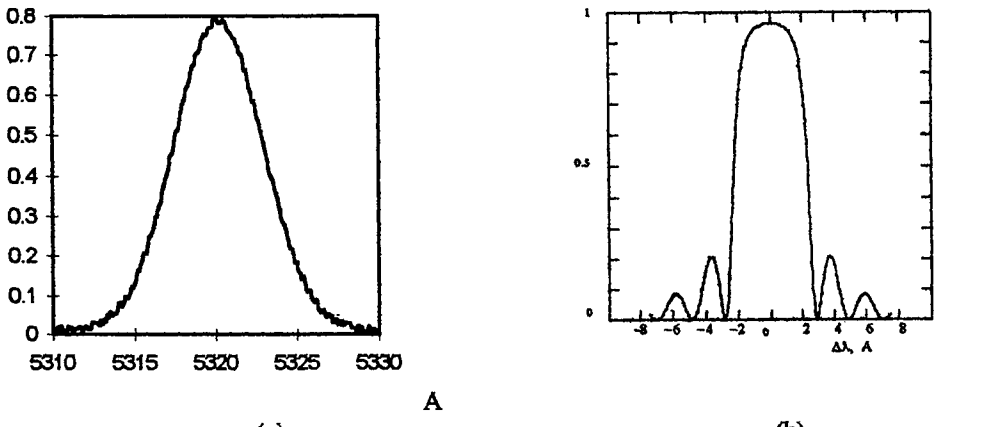


Figure 2. (a) Experimentally obtained contour of spectral selectivity for PDA sample of 0.5 mm thickness recorded in reflective configuration, λ=532 nm, (b) Theoretically calculated spectral selectivity contour for T=0.5 mm, λ=532 nm.

The corresponding angular selectivity of thick holograms can be calculated basing on coupled wave theory [2]. The highest value of angular selectivity can be reached in transmission configuration for the holograms thickness of several millimeters. We reported [4] on the experimentally obtained contours of angular selectivity of about milliradians in PDA samples (Figure 3). We recorded holographic spatial filters using the symmetric optical scheme with two equal intensity beams with flat wavefronts illuminating the photosensitive material from one side. The transmission holograms with the spatial frequency of the interferometric structure inside the photomaterial 900 - 2200 mm^{-1} were recorded using the photomaterials with 1-2 mm thickness. The obtained diffraction efficiency was from 80% to 97%. The experimentally obtained bandwidth of spatial selectivity was 0.5 - 2 milliradian.

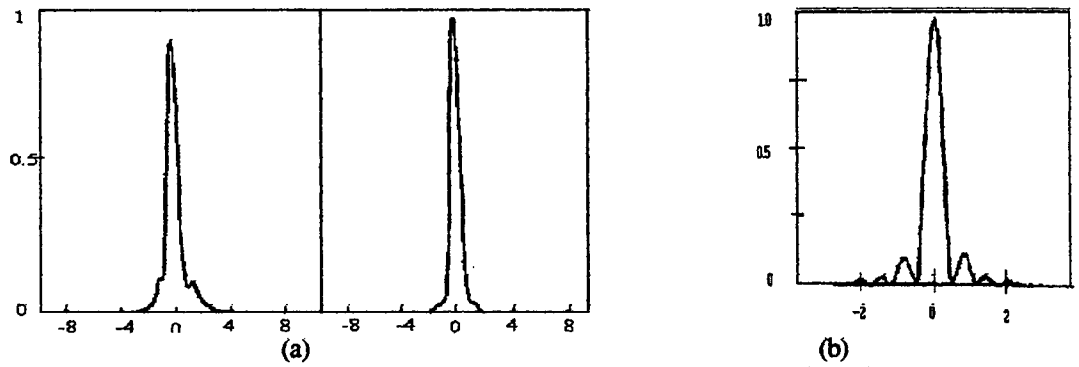


Figure 3. (a) The experimentally obtained angular selectivity contour for the PDA sample of 2 mm thickness recorded in transmission configuration, λ= 514 nm, (b) Calculated selectivity contour for the same parameters.

Therefore, to conclude this section we can say that the important characteristics of the typical holographic filters to be used in sensor systems for the selection of radiation are the following

Photomaterial thickness: T = 0.5 - 5 mm
Spectral selectivity: $\Delta\lambda$ = 0.1 - 10 nm
Angular selectivity: φ = 1-100 mrad

The spectral selection of particular wavelength range from the solar spectrum carried out by holographic filter is obvious. That concerns the angular selection, there were no reports so far on the possibility of spatial selection of radiation passed through atmosphere with turbulence using holograms or any other stationary (non-dynamic) gratings. For airborne sensor systems it is very important to know the atmosphere influence on the analysing optical signal. We decided to estimate the selective properties of holographic filters on the spatial frequencies of radiation passed through atmosphere with turbulence.

3. THE INFLUENCE OF THE FILTER ON THE CHARACTERISTICS OF RADIATION PASSED THROUGH THE ATMOSPHERE.

The spatial frequencies distribution in laser radiation passed through atmosphere is caused by the diffraction on turbulent eddies, the packets of air with a characteristic refractive index. The variation of the atmospheric refractive index is caused mainly by the spatial distribution of temperature [5] which happens due to differential heating of different portions of the Earth's surface by sun. For the isotropic turbulence the power spectral density S(k) of homogeneous turbulence, a function of wavenumber k, may be regarded as a relative number of eddies with characteristic size $L=2\pi/k$ [6]. Kolmogorov's theory of turbulence describes the power spectral density S(k) to be consisted of three separate regions. The first range of very small k (very large scale eddies) depends on large-scale geographic and meteorological conditions, it is not predicted by the theory. The second range, for k greater than k_0, the behavior of S(k) is determined by the physical laws governing turbulent flows and describing the break-up of large turbulent eddies into smaller ones. The scale size $L_0=2\pi/k_0$ is referred as the outer scale of the turbulence. The third range starts when k reaches another critical value k_m. At this scale region, very small turbulent eddies loose their energy due to viscous forces that results in a rapid drop of S(k) for $k>k_m$. The scale size $L_m=2\pi/k_m$ is called the inner scale of the turbulence. The resulting shape of the function S(k), so called von Karman spectrum, that includes the mentioned above models and an approximation for the function behavior at small k region, is described as follows [6]:

$$S(k) \cong \frac{0.033 \cdot C^2}{(k^2 + k_0^2)^{11/6}} \exp\left(\frac{-k^2}{k_m^2}\right) \qquad (3)$$

where C^2 is called the structure constant of refractive index fluctuations, it serves as a measure of the strength of the fluctuations. The variations of C^2 in horizontal direction depend on local terrain and wind conditions, so there is no analytical form for its description. The change of C^2 in vertical direction still depends strongly on atmospheric conditions at the time of the experiment, but its analytical approximation was proposed [7]. Typically the value of C^2 varies from $10^{-13}$to 10^{-18} $m^{-2/3}$ for the altitudes from 0 to 12 km decreasing with the increasing height. This is a consequence of decreasing temperature fluctuations at higher altitudes. Typical values of C^2 near ground vary from 10^{-13} $m^{-2/3}$ for strong turbulence to 10^{-17} $m^{-2/3}$ for weak turbulence, with 10^{-15} $m^{-2/3}$ often quoted as a typical average value.

The half of maximal diffraction angle θ_{max} for the laser light diffracted on the turbulent eddy with the size L is:

$$\Theta_{max} \cong \frac{\lambda}{2L} \qquad (4)$$

Holographic grating being introduces into the propagation way of laser beam cuts the spatial spectrum leaving the frequencies that correspond to the diffraction on eddies smaller than the some characteristic size. This characteristic size is determined by formula (4). Half of bandwidth of spatial selectivity contour for the best holographic filters is about milliradian. This means that the filter limits the diffraction angles in such a way as if only eddies larger than L_{min}~0.5 mm exist in the atmosphere. The corresponding value of wavenumber k_{cut} is about 10^4. One can see in Figure 4 the plot of the function S(k) and the wavenumber region limited by k_{cut} that can be cut holographic filter (hatched area). One can from the figure that holographic filter can only cut the part of spectrum that corresponds to eddy size L larger than L_0, the inner scale of turbulence. In non-disturbed atmosphere the contribution into spectrum distribution of such small size eddies is very small, and it can be neglected. However, the spectrum in disturbed atmosphere is very different, for example in the turbulent flow behind the aeroplane, and in this case the removal of the contribution made by small eddies can be essential.

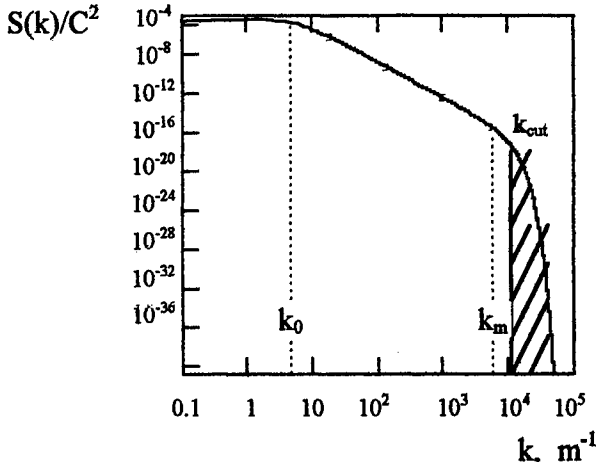


Figure 4. Power spectral density S(k) of refractive index fluctuations. Hatched area indicates the part of spectrum that can be cut by holographic filter.

4. THE CHANGE OF POWER SPECTRAL DENSITY OF LIGHT BY HOLOGRAPHIC GRATING

Holographic filter being three-dimensional grating influences on the power spectral density of the input light. If the grating strength (which depends on the distribution of photosensitive centers inside the material) is uniform throughout material depth, the diffraction efficiency as a function of wavelength (or angle) exhibits main maximum and several side maxima. The efforts of researchers were directed to suppress these side maxima, because they cause the widening of hologram selectivity and can be the origin of noises. It was shown [8] that non-uniform distribution of the grating strength in the material can lead to suppression of side maxima to some extent. The numerical calculations and experiments [9] showed the possibility to obtain bell-shaped contour of selectivity without side maxima.

We decided to analyze the possibility to obtain the superGaussian power spectral density distribution $F(x)= \exp(-x^4)$ of the output light by means of holographic grating. In lidar technique such beam may provide an advantage over Gaussian beam, because superGaussian beam has higher spatial modes in addition to the elementary lowest order Gaussian spatial mode. Those higher order modes help to reduce atmospheric speckle by providing "speckle diversity".

It should be pointed out that the selectivity property of hologram is associated with its thickness, which is limited to the value of about 10 cm. Therefore the formation of superGaussian beam by means of hologram is applicable for the lasers with very short pulses, the spectrum of which is rather wide.

We carried out our calculations under the assumption of gratings weakness to observe the general features of the effect. The diffraction efficiency is then a squared Fourier transform of the depth distribution of the grating strength [8]. The grating strength corresponds to the distribution of photosensitive centers inside the material. The results of the calculation are shown in Figure 5. One can see that the grating strength varies nearly as Gaussian function. However the phase of the grating has rather complicated dependence on the hologram depth.

There are at least two ways for practical realization of the photosensitive centers redistribution inside the material. The first one is photochemical bleaching of near -to-surface centers by means of incoherent pre-exposure of hologram by light strongly absorbed by photocenters. The intensity of light propagating into absorbing medium falls exponentially with penetration depth. The incoherent light converts the part of photosensitive centers into not-absorbing ones, and these centers will not contribute in the formation of the holographic grating during the recording. Another approach consists in redistribution of photosensitive centers caused by their diffusion into additional layers of material. For instance, the thin layer with uniform concentration of photosensitive molecules can be placed between two layers of basic material without photosensitive centers, and the redistribution of the molecules takes place due to the diffusion. The experimental results of

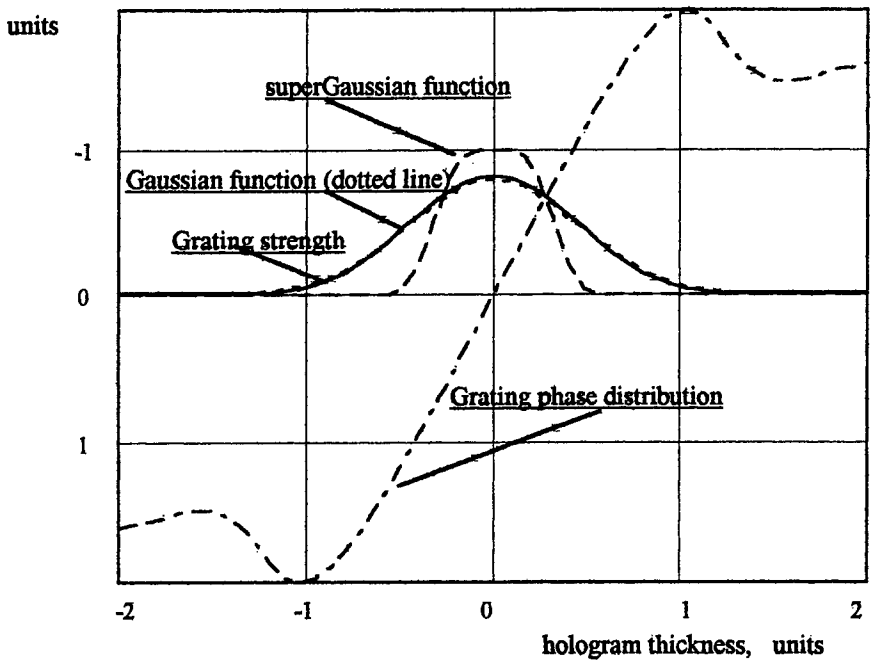


Figure 5. Distribution of the grating strength and phase distribution for holographic element serving to create the superGauss shape line of output beam.

recording of holographic filters using both of these technologies were reported in [9]. The Gaussian shape line without side maxima was obtained. It seems to be interesting to try experimentally to record the holographic filter which produces superGaussian line shape. According to our calculations (Figure 5) the distribution of photosensitive centers must be Gaussian, which can be reached by both photochemical and diffusion procedures. The change of the grating phase is rather hard to fulfill experimentally. Such distribution means that the grating must have different period in different layers of the photomaterial. Probably the interferometric maxima must have curved shape and diverge like a fan from one surface of hologram to another. The formation of such interferometric grating may be carried out by two waves with complex wavefronts.

5. CONCLUSIONS

We discussed the application of holographic filters for aerosense problems. We presented the results of experimental recording of transmission holographic optical filter with the bandwidth of spatial selectivity of about 1 milliradian and reflective holographic optical filter with the bandwidth spectral selectivity of 0.5 nm. Photopolymer with diffusive amplification (PDA) was used as a photosensitive medium. Such holographic filters can be applied in lidar technique, in communication systems, etc. We analysed the affect of holographic filter on the distribution of spatial frequencies of the atmosphere with turbulence. It was shown that holographic filter allows to cut the part of spectrum caused by the light diffraction on eddies less than millimeter in diameter. We also analysed the possibility to create the superGaussian beam by means of hologram. It was shown that the grating strength must change nearly as Gaussian function and besides that the phase distribution of the grating must exist.

ACKNOWLEDGMENTS

We are thankful to Dr. Yu.Korzinin for his help in carrying out this work. The support from the European Office of Aerospace Research and Development (EOARD) under the contracts #F61775-98-WE044 and # F61775- 98 WE130 is gratefully acknowledged.

REFERENCES

1. N.Reinhand, Yu.Korzinin, I.Semenova. Very thick holograms: manufacturing and applications, *Journal of Imaging Science & Technology*, 41, 3, 241, 1997.
2. R.J.Collier, C.B.Burckhardt, L.H.Lin, *Optical Holography*, Academic Press, New York and London, 1971, 689 pp.
3. N.Reinhand, I.Semenova, A.Popov, Highly selective holographic elements for spectroscopic applications, *Proceedings SPIE*, v.3637,1999, in press.
4. Yu.Korzinin, N.Reinhand, I.Semenova, J.E.Ludman, J.R.Riccobono, S.M.Shahriar, H.J.Caulfield, Two-dimensional holographic nonspatial filter, *Proceedings SPIE*, v.2688, pp.109-122, 1996.
5. S.F.Clifford, The classical theory of wave propagation in a turbulent medium, in *Laser beam propagation in the atmosphere*, J.W.Strohbehn editor, Springer-Verlag, Heidelberg, 1978.
6. J.W.Goodman, *Statistical Optics*, John Wiley & Sons, NY, 1985.
7. D.L.Fried, *J. Opt.Soc.Am*, 56, p.1380, 1966.
8. A.P.Yakimovich, Apodization of the selective response of 3D-holograms, *Optics and Spectroscopy*, v.53, 637-639, 1982.
9. A.P.Popov, A.V.Veniaminov, Yu.N.Sedunov, Photochemical and diffusional apodization of high-efficiency thick phase holograms, *Proceedings SPIE*, v.2215, 64-71, 1994.

Fabrication and temperature dependence of As_2S_3 fiber long period grating (LPG)*

Li Min Liao Yanbiao Chen Rongsheng Yang Kewu*
(Department of Electronics Engineering, Tsinghua University, Beijing 100084
*Beijing Glass Rearch Institute, Beijing 100062, P.R.China)

Abstract:

We report here a novel way to fabricate fiber LPG (long period grating) using moiré pattern amplitude mask (MPAM). Theoretical analysis of diffraction effect on MPAM has revealed that geometric optics rules are followed when area concerned is so close to diffractive grating that even Fresnel approximation can not be used. Experimental results of grating written in As_2S_3 fiber support the theoretic conclusion. And the temperature dependence is reported also.

Keywords: fiber grating, moiré Pattern, chalcogenide fiber

1. Introduction

Chalcogenide fiber is one of the most promising nonlinear optical media for all-optical switching because of its large nonlinearity and long interaction length. An all-optical switching has been previously demonstrated by M.Asobe et al with a required power of 3w and an As_2S_3 fiber with length of 1.3m [1]. And it is also proved that As_2S_3 glass possesses large group velocity dispersion (GVD) at communication wavelength [2]. Meanwhile, the FBG (Fiber Bragg Grating) technology in silica based fiber has been successfully demonstrated by using UV based on a photoinduced change in refractive index [3,4] and has attracted much attention because of its important applications, such as for optical filters, fiber lasers, etc. However the switching power required in LPG which couples light between co-propagating guiding and cladding modes is proved to be significantly less than that required in a fiber Bragg grating and well below the damage threshold of the fiber [7].

It is well known that chalcogenide glass exhibits a change in refractive index when exposed to visible light [5]. As_2S_3 is known to be a typical chalcogenide glass that exhibits remarkable photoinduced phenomena. An earlier report of grating fabrication in bulk chalcogenide glasses came about in 1994[6]. And afterwards, fabrications of FBG in chalcogenide fibers were reported separately by K.Tanaka in 1995[5] and M.Asobe in 1996[1].

We report here a novel method of the LPG fabrication in As_2S_3 fibers with MPAM, which is developed on the basis of photosensitivity measurement of As_2S_3 fibers and the comparison between As_2S_3 fiber LPG and silica fibers LPG. Such devices are promising to satisfy the request for the low switching power and ultrafast switching speed at the same time.

Fabrication of single-mode chalcogenide optical fiber is essential for such applications as fiber devices. Our experiments are based on the single mode As_2S_3 fibers, with wavelength at 1.55 μm and loss of 3dB/m, fabricated by Beijing Glass Research Institute.

2. Design of LPG in chalcogenide fiber

Before fabricating the LPG in As_2S_3 fiber, we first determine the photoinduced change in refractive index. We develop a heterodyne interferometer which consists of a linear frequency modulated LD as light source and a fiber Mach-Zehnder interferometer to measure the index change exactly [8]. Our experimental results proved a change of 4.5×10^{-4} in refractive index of As_2S_3 fiber after 35-minute-exposure to He-Ne laser of 30mw.

In order to design the period of LPG accurately, we solve the dispersion equations to get the propagation constants of the cladding modes. The propagation constants and periods of first six cladding modes are presented in table 1.

* This work is supported by the National Natural Science Foundation of China.

Table 1 the propagating constants and periods of first six cladding modes

Serial number of cladding modes	Propagating constant $\times 10^6$	Period (μm)
LP_{11}	9 77737	365
LP_{12}	9 77710	359
LP_{13}	9 77705	358
LP_{14}	9 77652	348
LP_{15}	9 77647	347
LP_{16}	9 77568	332

Because the refractive index is essential to design the period of grating and it is difficult to know exactly such large index as As_2S_3 of 2.4, we analyze the relation between period and refractive indexes and that between period and cladding modes. Main results are shown in figure 1a and 1b. It is clear that the error of 0.1 in fiber index causes only periodic error less than 5%, which results in small bias at the operating wavelength we care about.
Compared with Ge-doped oxide silica fibers, chalcogenide fiber gratings have shorter periods at the same operation wavelength, which means that the larger the refractive index, the shorter the period of grating.

3. Analysis of the diffraction intensity distribution of the moiré pattern

Moiré pattern is formed when two families of rays are overlapped. The pattern formed at the points of intersection is called moiré pattern. When two transmission gratings are overlapped with a small included angle in between, a set of fringes perpendicular to marks can be observed, which is just a kind of moiré pattern. The period between moiré pattern moves with the change of the included angle between two gratings. Therefore, we can produce amplitude masks making use of the character of moiré pattern, and gratings with different periods can be written in the fabric through the same mode.

3.1 Approximate calculation about diffraction

According to Kirchhoff diffraction theory, suppose a right-angle coordinate (x_0, y_0) is attached to grating A and another right angle coordinate (x, y) is attached to the observation screen and the axes of two coordinates are parallel to each other (see fig.2). The coordinate of inner point Q, through which transmitted beam pass the grating, is (x_0, y_0) and the coordinate of the point P in the observation screen is (x, y). And point P has amplitude [9,10]:

$$U(x,y)=\frac{i}{2\lambda}\int_{\Sigma}U(x_0,y_0)\frac{e^{-ikr}}{r}(1+\cos\theta)dx_0dy_0 \tag{1}$$

$$r^2=(x_1-x)^2+(y_1-y)^2+z^2$$

$$\cos\theta=\frac{z}{r}=\frac{z}{\sqrt{(x_1-x)^2+(y_1-y)^2+z^2}} \tag{2}$$

Where, $U(x_0,y_0)$ is the amplitude the plane $\sum$. Because θ is very small, $(x_1-x_0)^2 << z^2$ $(y_1-y_0)^2 << z^2$

Therefore

$$\frac{1}{2}(1+\cos\theta)\approx 1, r\approx z \tag{3}$$

Suppose the length and width of the grating are b and a, respectively, then $a << b$. As light passes through the slits of the grating, diffraction takes place mainly in the direction of width of slit (x_0). And it is just a simply one-dimension problem as the widths of slits are same and same effects are everywhere along the direction of length in the grating

The field distribution in the opening area of the slit can be simplified to a constant as parallel light impinges

perpendicularly onto the grating, so $U(x_0, y_0) = A = \text{CONSTANT}$ Formulary (1) can be rewritten as

$$U(x,y) = \frac{iA}{\lambda z}\int_{\Sigma} e^{-ikr}dx_0 \tag{4}$$

Considering $k = 2\pi/\lambda$, error introduced in is much bigger thanλ, the wavelength of incident light, if r in e^{-ikr} is substituted by z. And consequently an error of phase differences greater than 2π is induced. Therefore z can not substitute for r here

Because the distance between diffraction and observation screen(< 1mm) is quite short Fresnel approximation can not be used The binomial expansion is applied:

$$\sqrt{1+x} = 1 + x/2 - x^2/8 + \ldots \tag{5}$$

The induced phase error should be far less than 1 so does the phase error caused by omitting high order terms. According to our calculation following approximation can be used in the area of z~1mm:

$$r = z[1 + \frac{1}{2}(\frac{x-x_0}{z})^2 - \frac{1}{8}(\frac{x-x_0}{z})^4 + \frac{1}{16}(\frac{x-x_0}{z})^6] \tag{6}$$

(4) is expanded as:

$$\begin{aligned} U(x,y) &= \frac{iA}{\lambda z}\int_{\Sigma} e^{-ikz[1+\frac{1}{2}(\frac{x-x_0}{z})^2-\frac{1}{8}(\frac{x-x_0}{z})^4+\frac{1}{16}(\frac{x-x_0}{z})^6]}dx_0 \\ &= \frac{iA}{\lambda z}e^{-ikz}\cdot\int_{\Sigma}\cos\{kz[\frac{1}{2}(\frac{x-x_0}{z})^2 - \frac{1}{8}(\frac{x-x_0}{z})^4 + \frac{1}{16}(\frac{x-x_0}{z})^6]\}dx_0 \end{aligned} \tag{7}$$

Series development is carried out to the cosine function and the number of terms is restricted to confine the error of higher order terms within 0.2% limit, then equation (7) is changed to the integral of a multinomial. The distribution of diffractive power within 1mm area behind the grating A is calculated and the results are exhibited in figure 3.

It's quite clear in figure 3 that, when z is less than 1mm, the distribution of the light intensity approximates parabolic type and the intensity decreases drastically along the x-axis as the distance to center increases. If z continues to decrease (i.e. the observation screen moves to the direction of diffraction screen A), the speed by which intensity decreases along x-axis multiples as the distance to center increases. Figure (a) demonstrates that when z decreases to around $100\mu m$, intensity within the area of $100\mu m$ along x-axis decreases to less than 55 percent (here the width of slit is $a = 25\mu m$). So light propagates along straight line, and geometrical optics rule is followed in this area.

As the distance between diffraction and observation screen increases to large than 1mm Fresnel approximation can be used. Because distance is short, the relative intensity distribution in the area mentioned above is just the distribution of intensity close to diffraction peak shown in figure 3c

3.2 Gaussian light source[11]

Because the laser beam we use in the experiment is Gaussian shape, error with small intensity maximum and wider shape is introduced when we use planar wave as substitute in calculation. The error of shape can be expressed as

$$\Delta x = \frac{1}{2\pi}\cdot\frac{\alpha^2}{\omega^2} \tag{8}$$

where α is the width of the slit ω is the diameter of laser beam waist. Under the condition mentioned above the error is about 0.25 which can be neglected. So we can conclude that planar wave approximation does work.

3.2.1 Interference effect

According to interference theory of multi-rays, we have:

$$d\cdot\sin\theta = m\cdot\lambda \tag{8}$$

And the distance between two adjacent principal maximum fringes approximates the period of grating A because of the

small z about 1mm. Therefore, after the modulation of grating A, the distribution of light intensity almost has the same period as the grating A.

Given the same condition (i.e. z < 1mm), if light pass through grating B, which has a small inclined angle with grating A, after passing through A, then the light will experience the same modulation process again. The distribution of light intensity on the observation screen caused independently by grating B has the same period as itself if the distance in between is less than 1mm. Under such circumstance, the distribution of light intensity behind grating A and B is depended on moiré pattern effect.

3.2.2 Form of moiré pattern

Suppose the period of two gratings along x direction is Λ(space frequency $\xi=1/\Lambda$) and transmissivity is designated as ξ(shown in fig.4), according to Moiré pattern principle, the transmission intensity is expressed as:

$$\Phi = \Phi_0 \eta^2 = \Phi_0[1 + \cos 2\pi m_1 + \cos 2\pi m_2 + \frac{1}{2}\cos 2\pi(m_1 + m_2) + \frac{1}{2}\cos 2\pi(m_1 - m_2)] \quad (9)$$

Where, the fifth term refers to beat frequency, whose corresponding structure is called arithmetic fringe-i.e. Moiré pattern. Its series equation is : $q = m_1 - m_2$

Suppose two gratings are set so that their marks are symmetrical to y axis and the included angle between y axis are $\alpha$and $-\alpha$,respectively, the mark equations of two gratings are :

$$x\cos\alpha + y\sin\alpha = m_1 d \qquad x\cos\alpha - y\sin\alpha = m_2 d \quad (10)$$

Substitute them into ordinal equation and moiré pattern equation is expressed as:

$$2y\sin\alpha = dq \quad (11)$$

The fringes are a family of lines parallel to x-axis with the period of $d/2\sin\alpha$. Obviously, when the period of gratings is fixed, the spacing of moiré pattern moves with the change of included angle α between two gratings. That's the reason that gratings of different periods can be manufactured with a same moiré pattern mask.

So far we indicate that within the area close to amplitude mode of moiré pattern, the periodical distribution of transmission light intensity mainly results from the periodical modulation of moiré pattern. And in each period, the distribution of light intensity approximates Gaussian distribution.

4. Fabrication of the LPG in chalcogenide fiber with MPAM

Fig.5 is the experimental equipment. We use a He-Ne laser with intensity of 30mw as written source. After passing through collimating and focusing system, light beam is changed to a 30×0.5mm rectangular speckle parallel to the axes of the As_3S_2 fiber in which the grating is to be written. Then the speckle passes through the MPAM and impinges perpendicularly on to the As_3S_2 fiber. MPAM is very close to the fiber with a distance less than 1mm. Illuminating light is the ASE spectrum of Er-doped fiber amplifier EDFA with central wavelength of 1540nm 3dB bandwidth of 35nm and intensity of 14mw. After going through the As_3S_2 fiber, light beam from EDFA is received by spectrum analyzer (AQ-6315A, ANDO ELECTRIC CO., LTD). Figure 6a is ASE spectrum of EDFA 6b is the spectrum coming from the As_3S_2 fiber before the LPG is written; 6c is the spectrum after exposing for 3 hours.

Figure 6 illustrates the experimental results. A transmission dip emerges close to the designed wavelength of 1550nm (center wavelength of 1551.3nm), whose peak is above 5dB. A bias to design wavelength may be resulted from the periodic error in the MPAM. The results we got are in good agreed with the above theoretic analysis. The feasibility of writing fiber grating with MPAM is demonstrated. A new way to fabricate gratings in fiber is presented, which is promising for mass production.

5. Temperature dependence of As_2S_3 fiber

Considering the thermal characteristic of As2S3 glass, we try to find out the temperature dependence characteristic of As2S3 fiber from thermal experiments designed for As2S3 fiber LPG. It provided one way to proved the experiments from another attitude.

According to the periodic equation of LPG, we have the expression of wavelength shift with the temperature change:

$$\frac{\lambda}{T} = \frac{n_{eff}}{T}\Lambda + \Delta n_{eff}\frac{\Lambda}{T}$$

It is well known that the thermal-expanded coefficient of As2S3 glass is $24.6*10^{-6} m/\ C$.

We design the experiment to measure the wavelength change with the temperature increasement. Experimental results showed the wavelength shift of 2.15nm according to temperature difference of 1°C from 15°C to 50°C.

Calculation gives the coefficient of thermal-refractive index of As2S3 is 10-5/ °C, which is in good agreement with the As2Se3 glass whose physical and chemical characteristics are considered the most close to As2S3 glass and whose thermal refractive index is known. From another side, it gets to a similar result.

6. Conclusion

Since the beginning of 90's, researches on chalcogenide fibers have revived all over the world. The special photosensitive characteristic of chalcogenide fiber causes interests as well as their nonlinearity and dispersion characteristics. We forward a novel way to fabricate fiber LPG using moiré pattern amplitude mask (MPAM). Theoretical analysis of diffraction effect on MPAM has revealed that geometric optics rules are followed when area concerned is so close to diffractive grating that even Fresnel approximation can not be used. Experimental results of grating written in As_2S_3 fiber are in good agreement with the theoretic conclusion.

Our studies demonstrate the technical feasibility of fabricating LPG with MPAM. However, whether the shortest possible period is determined only by the periods of the gratings formed by the moiré pattern and applications of MPAM in fabrication of FBGs is still left for advanced researches.

References

1. M.Asobe, T.Ohara, Fabrication of Bragg grating in chalcogenide glass fiber using the transverse holographic method, Electron. Lett., 1996, vol.32, no.17, pp. 1611-3
2. M.Asobe, H.Kobayashi, laser-diode-driven ultrafast all-optical switching by using highly nonlinear chalcogenide glass fiber, Opt. Lett., 1993, vol. 18, pp.1056-58
3. K.O.Hill, Y.Fujii, Photosensitivity in optical fiber waveguides: application to reflection filter fabrication, Appl. Phys. Lett., 1978, vol.32, pp. 647-9
4. G.Meltz, W.W.Morey, Formation of Bragg gratings in optical fibers by transverse holographic method, Opt. Lett., 1989, vol.14, pp.823-5
5. K.Tanaka, Photoinduced dynamical changes in amorphous As_2S_3 films, Solid State Commun., 1991, vol.34, pp. 201-4
6. K.Shiramine, Photoinduced Bragg reflector in As_2S_3 glass, Appl. Phys. Lett., 1994, vol.64, pp. 1771-3
7. B.J.Eggleton, R.E.Slusher, All-optical switching in long-period fiber gratings, Opt. Lett., 1997, vol.22, pp.883-5
8. Li Min, Liao Yanbiao, Measurement of photoinduced change in refractive index, Academic Periodical Abstracts of China(Letters), vol.5, 1999, pp232
9. Liao Yanbiao, Physical Optics, Electronics Industry Press, Beijing, 1986
10. Yu Meiwen, Holographic Optics and Data Processing National Defense Industry Press 1984
11. Wei Guanghui Zhu Baoliang Laser beam Optics Beijing Industry University Press 1988

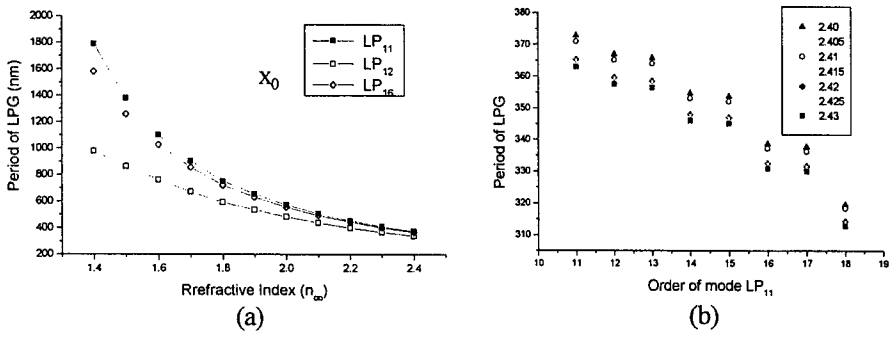


Fig.1 Relations between (a)period and different index, (b)period and cladding modes in grating

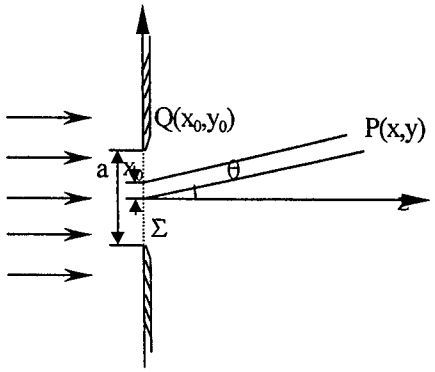


Fig.2 Sketch map of single slit diffraction

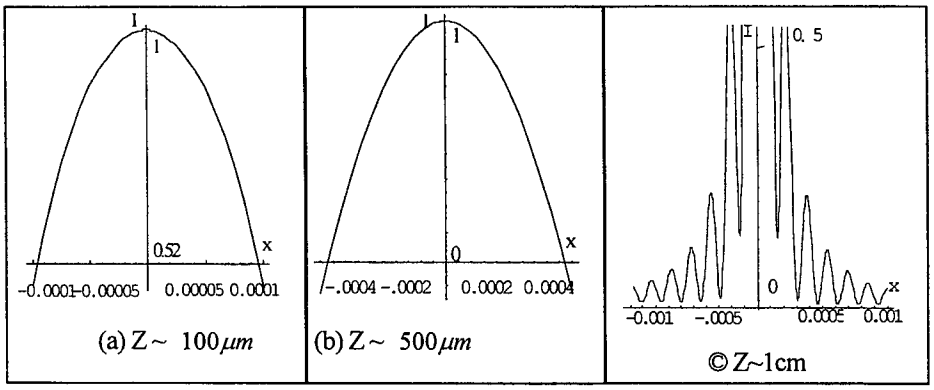


Fig.3 Relative intensity distribution of single slit diffraction

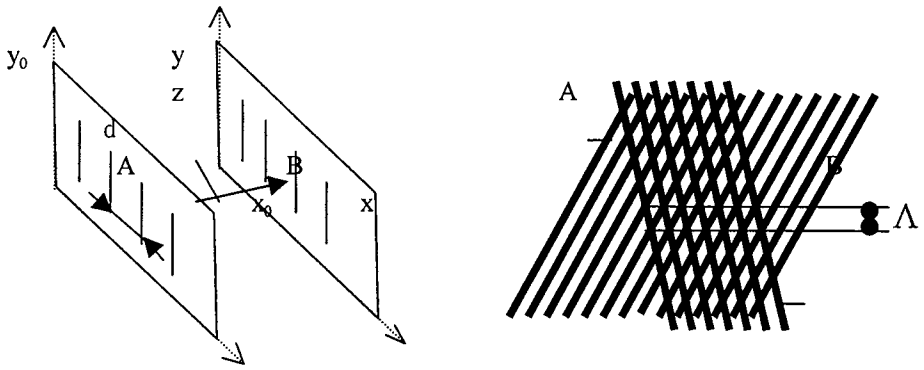


Fig.4: Form of moiré pattern

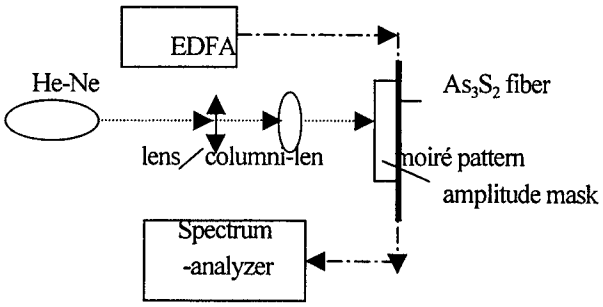


Fig.5 Experimental setup of grating writing

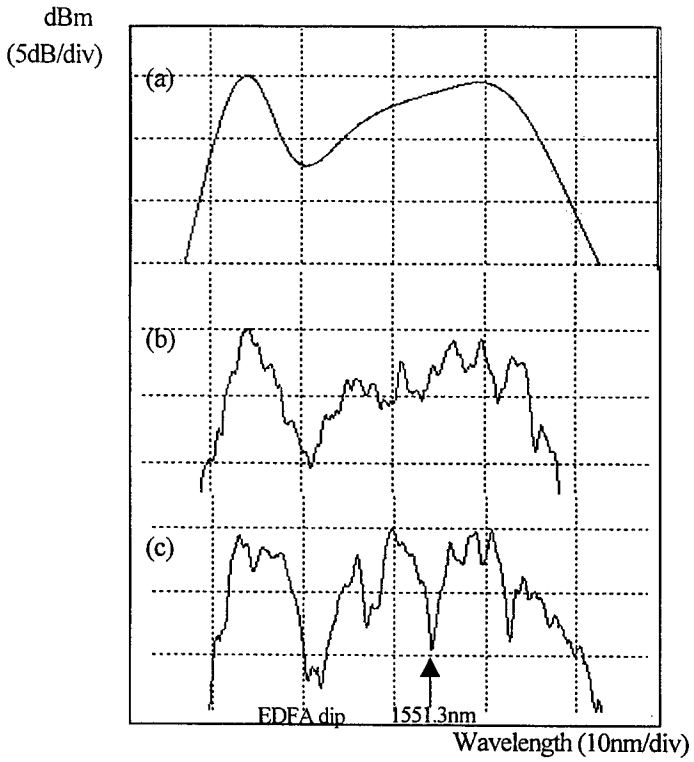


Fig.6 Spectrum before and after exposing

(a)ASE spectrum of EDFA (B)Before exposure (c)After exposure

Novel temperature compensation techniques for fibre Bragg gratings based magnetostrictive sensors

B. Yi, B. C. B. Chu, M. L. Mok, Z. Chen and K. S. Chiang

Optoelectronics Research Centre and Department of Electronic Engineering,
City University of Hong Kong, Hong Kong, China

ABSTRACT

Two simple techniques for the compensation of the temperature effects in fibre Bragg grating (FBG) based magnetostrictive sensors are demonstrated, where two FBG's are arranged in such a way that the magnetostrictive effects on the two gratings are added together, while the temperature effects are cancelled out. The two FBG's can be bonded in perpendicular to each other onto the surface of a magnetostrictive plate, or onto the surface of two different magnetostrictive bars that possess close thermal expansion coefficients but magnetostrictive coefficients of opposite signs. In both cases, the Bragg wavelength difference between the two gratings is insensitive to the temperature and measures the sum of the magnetostrictive effects experienced by the two FBG's. Using these techniques, we have been able to reduce the temperature sensitivity of the sensor by more than 10 times.

Keywords: Temperature compensation, fibre Bragg grating, magnetostrictive effect, magnetic field sensor, electric current sensor

1. INTRODUCTION

Fibre Bragg gratings as sensors are attracting considerable interest in recent years because of their small size, wavelength-encoded operation, and potential of large-scale multiplexing. Several fibre Bragg grating based magnetic field or electric current sensors have been proposed, which are based on magnetically induced circular birefringence[1] or magnetomotive-force generated tension along the gratings[2]. It is also possible to use a magnetostrictive material to physically elongate the grating under the influence of an applied magnetic field with a sensitivity of 2.0×10^{-2}nm/mT[3]. However, the large thermal responses of the magnetostrictive material and the grating itself make it difficult to develop this type of magnetic field sensors for practical application. Various temperature compensation schemes have been developed for fibre Bragg grating sensors[4-6], but most of these schemes require complicated arrangements and special devices.

In this paper, we report two simple methods to compensate for the temperature effects in grating based magnetostrictive sensors. In one approach, two FBG's are bonded in perpendicular to each other onto the surface of a magnetostrictive material. Under the longitudinal and transverse magnetostrictive effects, the Bragg wavelengths of the two gratings drift apart in the opposite direction. On the other hand, because the magnetostrictive material is thermally isotropic, a temperature change gives rise to identical Bragg wavelength shifts in the two gratings. The difference between the Bragg wavelengths of the two gratings is essentially temperature independent and, therefore, an accurate measure of the magnetostrictive effect. Alternatively, two FBG's can be bonded onto the surface of two different magnetostrictive bars that possess similar thermal expansion coefficients but magnetostrictive coefficients of opposite signs. In this way, the Bragg wavelength difference measures the total of the magnetostrictive effects of the two materials and is insensitive to the temperature.

2. PRINCIPLE

The experimental setups are illustrated in Fig. 1. In Fig. 1(a), two fibre Bragg gratings with almost identical Bragg wavelengths are bonded in perpendicular to each other onto the surface of a square-shaped (40mm×40mm×2mm) Terfenol-D ($Tb_{0.27}Dy_{0.73}Fe_2$) plate. Terfenol-D is a terbium-iron alloy that pocesses gigantic magnetostrictive effect, which is almost

Correspondence: Email: eebcbchu@cityu.edu.hk; WWW: http://ee.cityu.edu.hk/; Telephone: (852) 2788-8458; Fax: (852) 2788-7791

Part of the SPIE Conference on Advanced Photonic Sensors and Applications
Singapore • November–December 1999
SPIE Vol. 3897 • 0277-786X/99/$10.00

42 times larger than that of nickel. A magnetic field is generated by a DC current coil and applied to the Terfenol-D plate. In the presence of a magnetic field, the plate is subjected to longitudinal expansion in the direction of the magnetic field and transverse contraction in the direction perpendicular to the magnetic field. As a result, the Bragg wavelengths of the two FBG's will drift apart in opposite direction. At the same time, the temperature of the sensor could be changed as a result of heat dissipation from the current coil. Since the magnetostrictive material is thermally isotropic, the coefficients of thermal expansion are the same along both directions. Hence, a temperature change results in an identical wavelength shift on the FBG's. Measuring the difference of the Bragg wavelengths of the FBG's can eliminate the temperature effect.

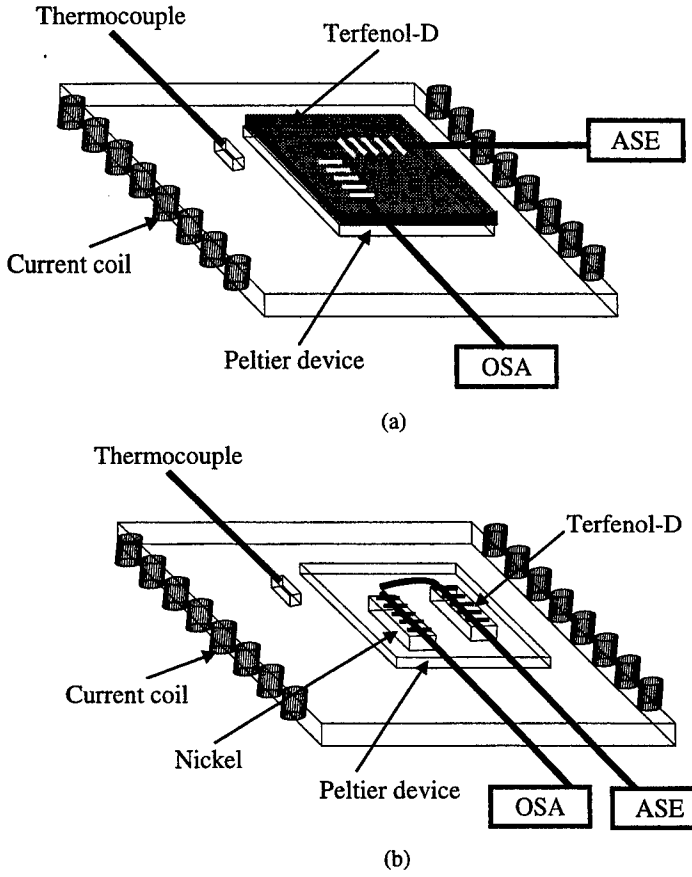


Fig. 1. Schematic diagrams of the experimental setups. (a) Two FBG's bonded in perpendicular to each other on a Terfenol-D plate. (b) Two FBG's bonded in parallel with each other on a Tefernol-D bar and a nickel bar, respectively.

The shifts $\Delta\lambda_{B1}$ and $\Delta\lambda_{B2}$ in the Bragg wavelengths λ_{B1} and λ_{B2} of the two gratings can be expressed as[7]

$$\frac{\Delta\lambda_{B1}}{\lambda_{B1}} = \rho_l H + (1 - p_e)\alpha_t \Delta T + \xi \Delta T \tag{1}$$

$$\frac{\Delta\lambda_{B2}}{\lambda_{B2}} = -\rho_t H + (1 - p_e)\alpha_t \Delta T + \xi \Delta T \tag{2}$$

where ρ_l (>0) and ρ_t (>0) are coefficients proportional to the magnetostrictive constants of Terfenol-D in the longitudinal and transverse directions, respectively, p_e is the photo-elastic constant of silica fibre, ξ *is* the thermo-optic coefficient of the fibre, α_t is the thermal expansion coefficient of Terfenol-D, H is the applied magnetic field, and ΔT is the temperature change. The difference in the Bragg wavelength shifts can be found by subtracting Eq. (2) from Eq. (1):

$$\frac{\Delta\lambda_{B1}}{\lambda_{B1}} - \frac{\Delta\lambda_{B2}}{\lambda_{B2}} = (\rho_l + \rho_t)H \tag{3}$$

Equation (3) shows that this method not only eliminates any undesirable temperature effect but also enhances the sensitivity to the magnetic field.

In the setup shown in Fig. 1(b), two FBG's are bonded onto the surface of a Terfenol-D bar with a size of 40mm×5mm×2mm and a nickel bar of the same size, respectively, which are placed in parallel with the direction of the magnetic field. As Terfenol-D and nickel possess similar thermal expansion coefficients, 12×10^{-6}/°C and 13.1×10^{-6}/°C, respectively[8, 9], but magnetostrictive coefficients of opposite signs, the difference in the Bragg wavelength shifts between the two FBG's is given by:

$$\frac{\Delta\lambda_{B1}}{\lambda_{B1}} - \frac{\Delta\lambda_{B2}}{\lambda_{B2}} = (\rho_l + \rho_n)H + (1 - p_e)(\alpha_t - \alpha_n)\Delta T \approx (\rho_l + \rho_n)H \tag{4}$$

where ρ_n (>0) is coefficient proportional to the magnetostrictive constant of nickel in the longitudinal direction and α_n ($\alpha_n \cong \alpha_t$) are the thermal expansion coefficients of nickel. The difference of the Bragg wavelengths of the gratings is thus a measure of the sum of the magnetostrictive effects of Terfenol-D and nickel.

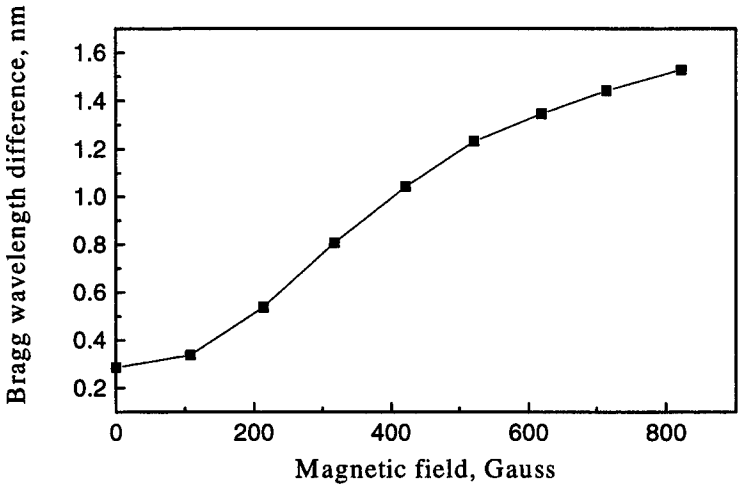


Fig. 2. The difference between the two Bragg wavelengths as a function of the applied magnetic field for the experimental setup shown in Fig. 1(a).

3. EXPERIMENTS AND RESULTS

The fibre Bragg gratings used in our experiments were fabricated in QPS photosensitive fibres by the phase-mask method in our laboratory. The amplified spontaneous emission (ASE) spectrum of an erbium-doped fibre amplifier (EDFA) was used as the light source and the output spectrum was measured with an optical spectrum analyzer (OSA). For the sensor shown in Fig. 1(a), the Bragg wavelengths of the two fibre Bragg gratings measured at 22°C were 1561.7nm and 1562.0nm, respectively. The Terfenol-D plate with the mounted gratings was placed inside a 12cm long solenoid that was formed by winding 258 turns of large current carrying copper wire around a rectangular-shaped plastic tube. A magnetic field of 820 Gauss could be produced with a DC current of 38 Ampere through the copper wire. Adjusting the magnitude of the applied

current to the solenoid varied the magnetic field. The measured difference between the Bragg wavelengths of the two gratings against the applied magnetic field is shown in Fig. 2. A difference of 1.2nm in the wavelength shifts is obtained with a magnetic field of 820 Gauss. The maximum sensitivity in the linear region of the response curve is 2.44×10^{-3} nm/Gauss, which is larger than the sensitivity 2.0×10^{-3} nm/Gauss measured for a Terfenol-D sensor using a single grating[3]. The improvement in sensitivity is due to the additional contribution from the grating mounted in the transverse direction of the magnetic field. The nonlinearity shown in Fig. 2 is believed to be the result of magnetostriction saturation at large magnetic field.

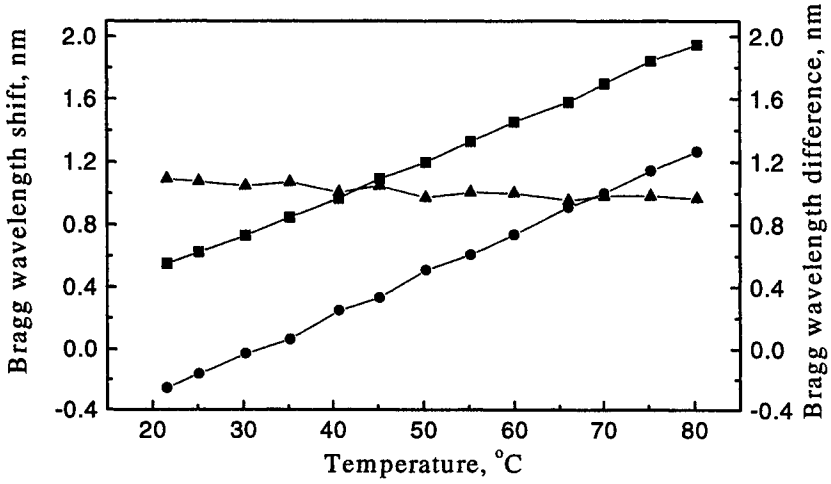


Fig. 3. The Bragg wavelength shifts of the two gratings and their difference against the temperature for the experimental setup shown in Fig. 1(a) with the magnetic field fixed at 317 Gauss.

—■— wavelength shift of the grating in the longitudinal direction
—●— wavelength shift of the grating in the transverse direction
—▲— wavelength difference between the two gratings

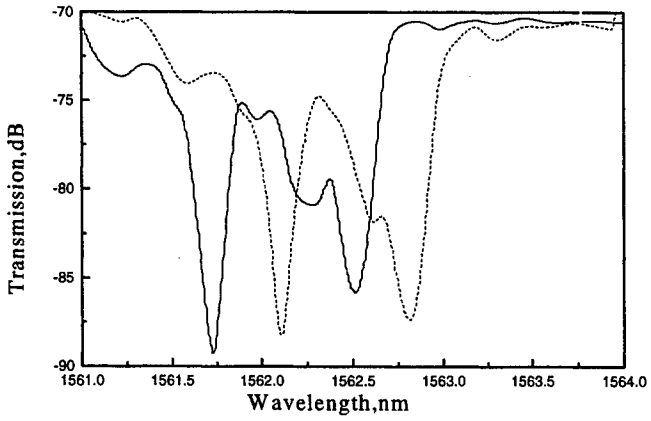


Fig. 4. Transmission spectra of the grating pair at different temperatures for the experimental setup shown in Fig. 1(a).

—at 22 °C
----at 45 °C

To demonstrate that the sensor was insensitive to the temperature, a Peltier effect heat pump was used to vary the temperature of the sensor from 22°C to 80°C. The magnetic field was fixed at 317 Gauss. The Bragg wavelength shift of each grating was recorded for every 5°C. Figure 3 shows the measured wavelength shifts of the gratings and the wavelength difference between the two gratings against the temperature. The temperature sensitivity for each bonded grating is approximately 2.6×10^{-2}nm/°C. On the other hand, the difference between the two Bragg wavelengths has a temperature sensitivity of -2.0×10^{-3}nm/°C, which is lower than that of each grating by an order of magnitude. It can be observed from Fig.3 that the Bragg wavelength difference decreases slowly as the temperature increases. This is due to the fact that the magnetostrictive coefficient of Terfenol-D decreases with increasing temperature[10]. Figure 4 shows the transmission spectrum of the gratings at 22 °C and 45 °C. It can be seen from Fig.4 that a temperature change results in an identical wavelength shift on the grating pair, as expected.

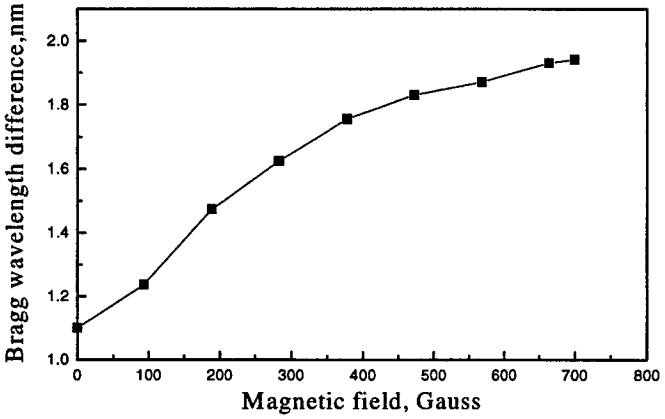


Fig. 5. The difference between the two Bragg wavelengths as a function of the applied magnetic field for the setup shown in Fig.1 (b).

For the sensor shown in Fig. 1(b), another pair of FBG's was used. The Bragg wavelengths measured at 22°C were 1563.5nm and 1564.8nm, respectively. These two FBG's were fusion-spliced together in series and bonded on the surfaces of a Terfenol-D bar and a nickel bar, respectively. The Terfenol-D and nickel bars were placed in parallel with each other inside the solenoid. As the magnetostrictive coefficients of Terfenol-D and nickel are of opposite signs, the Bragg wavelengths of the two gratings drift apart in opposite direction. Figure 5 shows the Bragg wavelength difference against the applied magnetic field. A difference of 0.8nm in the wavelength shifts was measured with a magnetic field of 700 Gauss. The maximum sensitivity in the linear region of the response curve is 1.8×10^{-3}nm/Gauss, which is smaller than the response measured from the sensor shown in Fig. 1(a). Although it was shown in Eq. (4) that the magnetostrictive effect should be enhanced, the magnetostrictive coefficient of nickel is too small, compared with that of Terfenol-D, to make the improvement noticeable in the measurement. Moreover, the measurement was also limited by the resolution (0.1nm) of the optical spectrum analyzer.

Figure 6 shows the measured wavelength shifts of the gratings and the wavelength difference between the two gratings against the temperature at a fixed magnetic field of 286 Gauss for the sensor shown in Fig. 1(b). The Bragg wavelength shift of each grating was also recorded for every 5°C. The temperature sensitivities of the gratings on Tefernol-D and nickel are 1.8×10^{-2}nm/°C and 2.0×10^{-2}nm/°C, respectively. The difference between the two Bragg wavelengths has a temperature sensitivity of -1.6×10^{-3}nm/°C. Similar to the results obtained from the other configuration, the temperature compensation is effective at low temperature. At high temperature, the magnetostrictive coefficient of Terfenol-D decreases gradually and the performance of the compensation method is somewhat degraded.

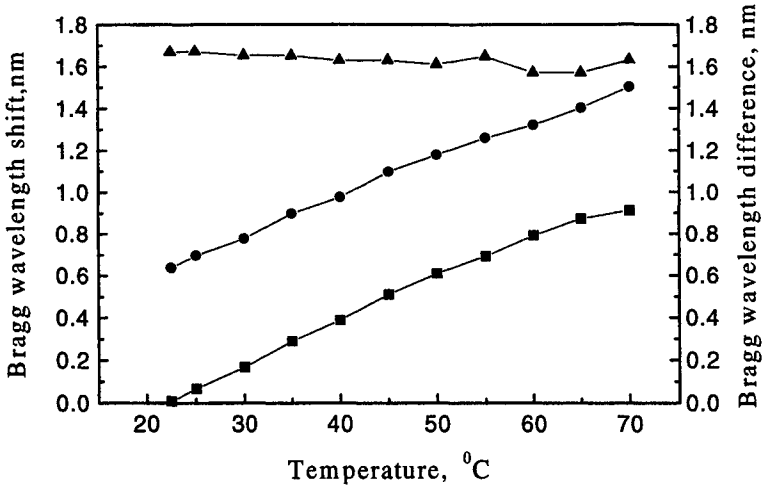


Fig. 6. The Bragg wavelength shifts of the two gratings and their difference against the temperature for the experimental setup shown in Fig. 1(b) with the magnetic field fixed at 286 Gauss.

—■— wavelength shift of the grating on the surface of nickel
—●— wavelength shift of the grating on the surface of Terfenol-D
—▲— wavelength difference between the two gratings

4. CONCLUSIONS

We have demonstrated two simple methods for temperature compensation of fibre Bragg grating based magnetostrictive sensors. With the configuration where the two fibre Bragg gratings are bonded in perpendicular to each other on a single plate of Terfenol-D, a large Terfenol-D plate is needed, which is expensive. The other configuration where two different magnetostrictive materials are used can lead to a more compact and cost-effective sensor. Both methods can eliminate the temperature effects on the sensor over a wide range of temperature and, at the same time, enhance the overall sensitivity of the sensor.

5. ACKNOWLEDGEMENT

This work is supported by the University Grants Committee (Hong Kong Government) under a research grant.

6. REFERENCES

1. A. D. Kersey, and M. J. Marrone, "Fibre Bragg grating high–magnetic-field probe", *Proceedings of 10th International Conference on Optical Fibre Sensors*, pp. 53 –56, Glasgow, 1994.
2. S. Jin, R. P. Espindola, H. Mavoori, T. A. Strasser, and J. J. DeMarco, "Magnetically programmable fibre Bragg gratings", *Electronics Letters*, **34**, (22), pp. 2158 - 2159, 1998.
3. J. L. Cruz, A. Diez, M. V. Andres, A. Segura, B. Ortega, and L. Dong, "Fibre Bragg gratings tuned and chirped using magnetic fields", *Electronics Letters*, **33**, (3), pp. 235 - 236, 1997.
4. G. W. Yoffe, P. A. Krug, F. Ouellette, and D. A. Thorncraft, "Passive temperature-compensating package for optical fiber gratings", *Applied Optics*, **34**, (30), pp. 6859 - 6861, 1995
5. F. M. Haran, J. K. Rew, and P. D. Foote, "A strain-isolated fibre Bragg grating sensor for temperature compensation of fibre Bragg grating strain", *Measurement Science and Technology*, **9**, pp. 1163 - 1166, 1998

6. V. Arya, D. W. Sherrer, A. Wang, and R. O. Claus, "Application of thin-film optical filters to the temperature compensation of optical fibre grating-based devices", *IEEE Transactions on Instrumentation and Measurement,* **46**, (5), pp. 1173 - 1177, 1997.
7. Y. L. Lo, and H. S. Chuang, "Measurement of thermal expansion coefficients using an in-fibre Bragg-grating sensor", *Measurement Science and Technology*, **9**, pp. 1543 - 1547, 1998.
8. A. E. Clark, and D. N. Crowder, "High temperature magnetostriction of $TbFe_2$ and $Tb_{0.27}Dy_{0.73}Fe_2$", *IEEE Transactions on Magnetics*, **21**, (5), pp. 1945 - 1947, 1985.
9. R. B. Ross, *Metallic materials specification handbook*, third edition, Chapman and Hall Ltd, 1980.
10. Data sheet of Terfenol-D from ETREMA Products Inc., U.S.A., 1998.

Bent long-period fiber gratings for sensor applications

Zhihao Chen[a], Kin Seng Chiang[a], Mei Nar Ng[a], Yuen Ming Chan[a], and Hong Ke[b]

[a]Optoelectronics Research Centre and Department of Electronic Engineering,
City University of Hong Kong, Hong Kong, China
[b]Department of Electronic Engineering, Tsinghua University, Beijing, China

ABSTRACT

The transmission characteristics of bent long-period fiber gratings (LPFG's) and their applications as sensors are studied experimentally. We demonstrate that, by bending a LPFG, two major rejection bands in the wavelength region of interest can be produced, whose relative strength can be controlled by the amount of bending. The wavelength separation between the two bands increases linearly with the lateral displacement of the bent LPFG and a linear tuning range over 40 nm is demonstrated. All these characteristics show no significant polarization dependence. We also find that, by mounting a bent LPFG on a proper material, the thermally induced bending effect can enhance or cancel the direct thermal effect. With this technique, we obtain temperature sensitivities of 0.01nm/°C and -0.35nm/°C, which are, respectively, about 1/5 and 7 times of that of a straight bare LPFG. A temperature-insensitive LPFG-based sensor for the measurement of the concentration of NaCl in water is demonstrated with a sensitivity of -0.35nm/molarity by keeping the temperature sensitivity at 0.01nm/°C from 15 to 70°C. By using the same technique, a strain sensitivity of -49nm/%ε is obtained, which is about 15 times of that of a straight bare LPFG.

Keywords: Fiber optics; Fiber optic sensors; Fiber gratings

1. INTRODUCTION

Long-period fiber gratings (LPFG's) have found important applications as optical filters, optical sensors, and gain equalizers in erbium-doped fiber amplifiers (EDFA's).[1-3] In general, the transmission spectrum of a uniform LPFG consists of only one major rejection band in the wavelength region of interest. For some applications, however, two or more rejection bands are required. For example, two rejection bands in the spectrum of a LPFG are needed for simultaneous measurement of strain and temperature.[4] To produce multiple bands in the transmission spectrum, one can cascade a number of different LPFG's.[5] It should be noted that multiple rejection bands arising from the coupling to various cladding modes of the fiber are usually present in the transmission spectrum of a LPFG, but these bands are spectrally widely separated.[1,4] Apart from cascading different LPFG's, it is possible to shape the transmission spectrum of a LPFG by introducing phase shifts along the grating.[6,7]

A LPFG-based sensor usually relies on the measurement of the shift in the resonance wavelength of the grating in response to a change in the physical parameter that is measured. However, the resonance wavelength of a LPFG is in general sensitive to a number of physical parameters (e.g., temperature, strain, refractive index of the surrounding medium, etc.).[1-10] It is of practical importance to develop means to separate various physical effects.

In this paper, we report our experimental results on bent LPFG's. We find that, by bending a uniform LPFG, new polarization-insensitive rejection bands in the transmission spectrum of the LPFG can be produced. We also describe the use of a bent LPFG to enhance or suppress the temperature effect, and demonstrate the principle of temperature-insensitive sensors based on bent LPFG's. The strain sensitivity of a bent LPFG is also discussed.

2. TRANSMISSION CHARACTERISTICS OF BENT LPFG'S

The LPFG used in our experiments was fabricated in a photosensitive fiber (QPS fiber) by exposing the fiber to 248-nm UV radiation from a KrF excimer laser through a chrome-plated amplitude mask. The UV energy was set at about 85mJ/cm^2 (at

Part of the SPIE Conference on Advanced Photonic Sensors and Applications
Singapore • November–December 1999
SPIE Vol. 3897 • 0277-786X/99/$10.00

a pulse repetition rate of 10 Hz) and the exposure time varied from a few minutes to several ten minutes. The transmission spectrum of the LPFG was monitored *in situ* during its fabrication.

In the first experiment, a uniform LPFG about 30-mm long was attached to the back of a flexible beam, which was supported freely at both ends. Bending was applied to the LPFG by pushing a micrometer translator towards the beam at its center. As shown in Fig.1, by measuring the lateral displacement h, we can determine the bend radius R of the LPFG by $R=(a^2+h^2)/2h$, where $2a$ = 80 mm is the distance between the two ends. Light from a pigtailed broadband ELED was launched into the LPFG that was subject to bending. The transmission spectrum of the LPFG was measured with an optical spectrum analyzer. In the early experiments,[10] it was found that, by bending a LPFG, the rejection band in the transmission spectrum of the LPFG shifted to longer wavelengths, and an increase in the bend radius caused a broadening of the bandwidth and a reduction in the strength of the band. Our experiments, however, reveal some new features that have not been reported previously.

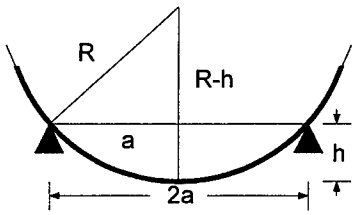


Fig. 1. Bending of a long-period fiber grating (LPFG).

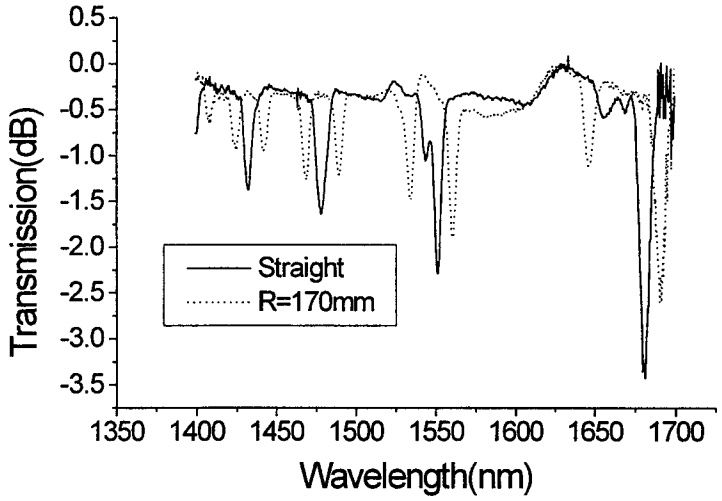


Fig. 2. Comparison of the measured transmission spectra of a straight LPFG and a bent LPFG.

The transmission spectra of our LPFG over the spectral range of the ELED are shown in Fig.2 for the cases of a straight grating and a bent grating (R=170 mm). For the straight grating, several orders of rejection bands can be identified in Fig.2 as distinct dips in the spectrum, which correspond to the resonance couplings between the fundamental mode and various cladding modes of the fiber. When the grating was bent, we observed that the dip in each order of the rejection band shifted to a longer wavelength with a reduction in strength and a new dip emerged at a shorter wavelength. It can be seen from Fig.2 that the bend-induced dip in the long-wavelength band (1640 – 1700 nm) is located far away from, and significantly weaker than, the original dip. This bend-induced dip could have been ignored in the previous studies where only the bend-induced wavelength shift of the major dip was studied.[10] In fact, by measuring the wavelength shift of the major dip in the

long-wavelength band as a function of the bend radius, we obtained results similar to those reported in Ref.10. On the other hand, bend-induced dips can be significant in the short-wavelength bands, as clearly shown in Fig.2.

In the second experiment, a LPFG about 18 mm long was bent and its transmission spectrum from 1500 nm to 1560 nm was measured at different bend radii. Figure 3 shows the transmission spectra at four different values of lateral displacement h: 0 mm, 1.0 mm, 2.5 mm, and 4.5 mm, which correspond, respectively, to bend radii of $+\infty$ (no bending), 800 mm, 320 mm, and 180 mm. The transmission spectrum of the straight LPFG (h=0 mm) consists of a major dip and a minor dip.

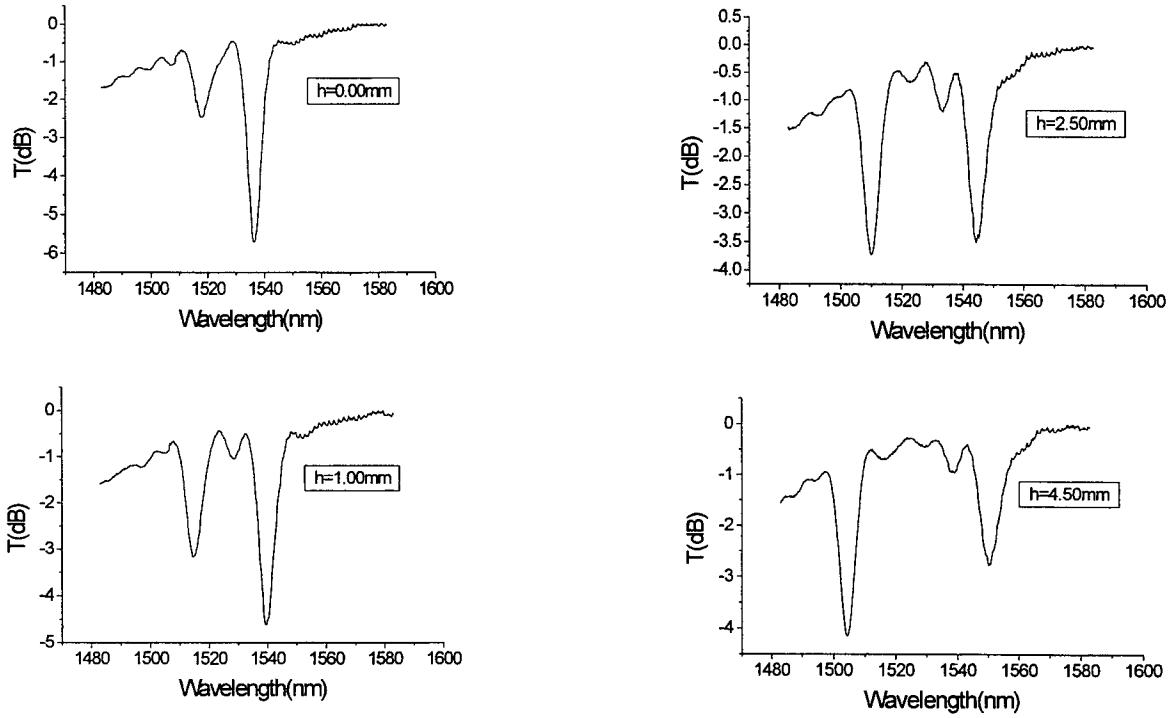


Fig. 3. Transmission spectra of a LPFG measured at different values of lateral displacement, showing the effects of bending on the resonance dips.

As shown in Fig.3, when the lateral displacement is increased, the major dip shifts to a longer wavelength with a slight increase in bandwidth while the minor dip grows and shifts to a shorter wavelength. In the case h=1.0 mm, the dip at the shorter wavelength is weaker than that at the longer wavelength. As the lateral displacement is increased to 2.5 mm, the two dips have practically equal strengths. As the lateral displacement is increased further, for instance, to 4.5 mm, the dip at the shorter wavelength becomes stronger than that at the longer wavelength.

The wavelengths of the two dips are found to vary linearly with the lateral displacement h, as shown in Fig.4. The wavelength separation of the two dips, i.e., the dip span, therefore, also varies linearly with h. A linear tuning range over 40nm can be achieved easily. Our results show that the two dips move with h in opposite directions at almost the same rate. The experiments were repeated with several different LPFG's and essentially the same results were obtained.

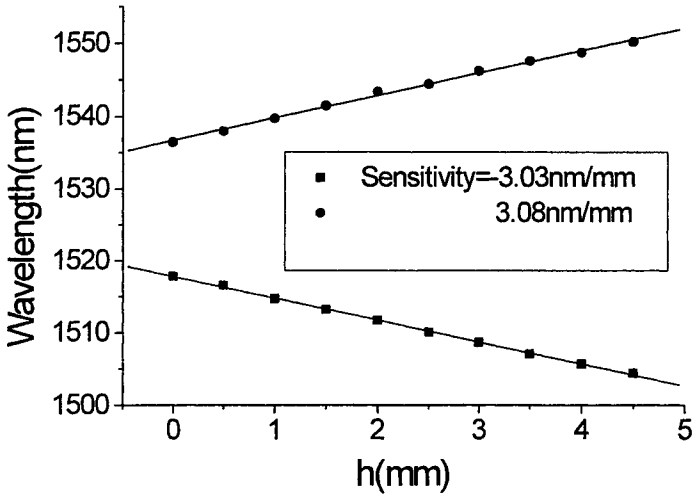


Fig. 4. The wavelengths of the two dips shown in Fig.3 vary linearly with the lateral displacement.

To study the polarization dependence of the bending effects, we used the amplified spontaneous emission (ASE) of an EDFA with an isolator (instead of an ELED) as the broadband source and inserted a bulk polarizer and a polarization controller between the source and the LPFG. Typical output spectra measured by rotating the polarizer at steps (covering a full rotation) for a bent LPFG are shown in Fig.5(a), while typical output spectra with the LPFG removed are shown in Fig.5(b) as the references. In both sets of results, the spectrum does not recover itself after a full rotation of the polarizer - there is an uncertainty of +/-0.5 dB across the spectrum. The results are found to be independent of the setting of the polarization controller. By comparing Figs.5(a) and (b), we can conclude that, within the measurement error, the transmission characteristics of the bent LPFG show no significant polarization dependence.

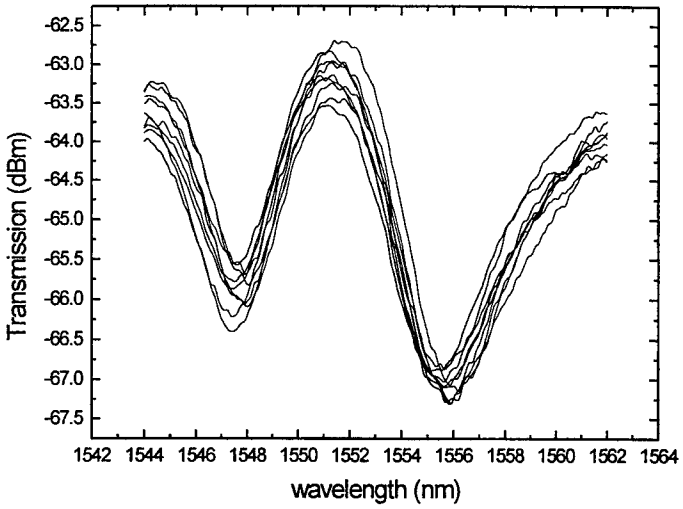
In the transmission spectrum of a LPFG, the major dips correspond to the resonance couplings between the fundamental mode and some particular cladding modes.[1] The couplings are effective only for cladding modes with field patterns of the right symmetry (i.e., cladding modes that can give large overlap integrals, and hence, large coupling coefficients). When the LPFG is bent, the symmetry in the refractive-index perturbation along the grating is broken and couplings to cladding modes with other symmetry properties can be significant or even become dominant. This may explain the generation of new dips in the transmission spectrum and the growth of such dips with bending. Bending can also change the bulk index of the fiber and produce chirping in the grating. These changes may account for the bend-induced wavelength shifts of the dips and the slight broadening of the bandwidths.

The temperature dependence of the transmission spectrum of a bent LPFG was also measured. We find that the wavelengths of the two dips in the transmission spectrum of a bent LPFG have the same temperature sensitivity (0.05nm/ °C), while the dip span is insensitive to the temperature. Typical measurement results are shown in Fig.6, where the dip span $\delta\lambda$ is plotted against the temperature at several bend radii. It is possible to build a temperature-insensitive sensor with a single bent LPFG by using the dip span as the parameter to be monitored.

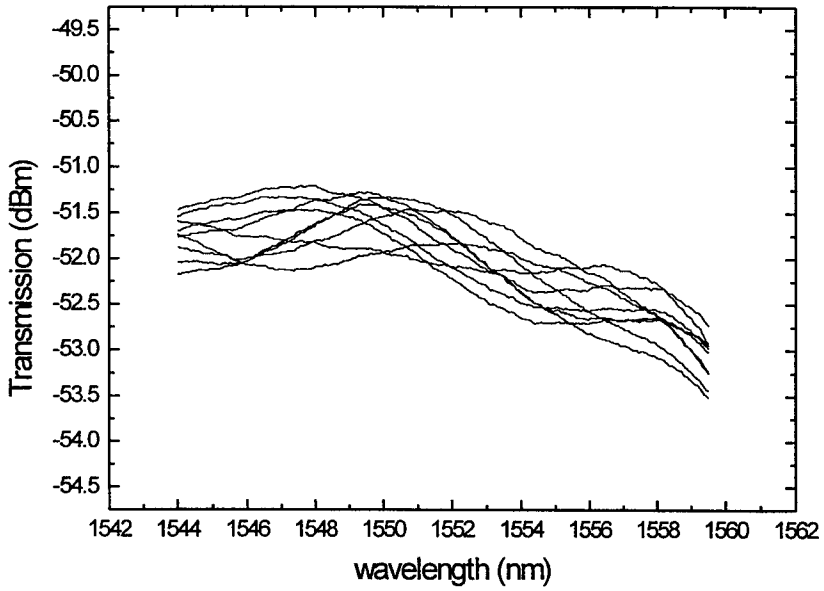
3. SENSORS BASED ON BENT LPFG'S

A LPFG-based sensor relies on the measurement of the shift in the resonance wavelength of a LPFG in response to a change in the physical parameter that is measured. For the sensor to be useful, it is necessary to separate various physical effects that can change the resonance wavelength of the LPFG. In this section, we discuss a method of enhancing or suppressing the temperature dependence of the resonance wavelength of a LPFG. A LPFG with an enhanced temperature sensitivity can be used as a temperature sensor, while a LPFG with a suppressed temperature sensitivity is useful for forming other types of sensors. It is possible to reduce the temperature sensitivity of a LPFG by using a proper host fiber.[11,12] It is, however, more

desirable to develop a simple means to reduce the temperature sensitivity of a LPFG in an ordinary fiber. Our approach is based on mounting a bent LPFG on a proper material in such a way that the thermally induced bending effect cancels the direct thermal effect. Our method can also be used to enhance the temperature and strain sensitivities of a LPFG.



(a)



(b)

Fig. 5. Typical output spectra measured at different polarizer angles for two cases: (a) with a bent LPFG; (b) without the LPFG. An uncertainty of +/-0.5 dB is seen in both cases.

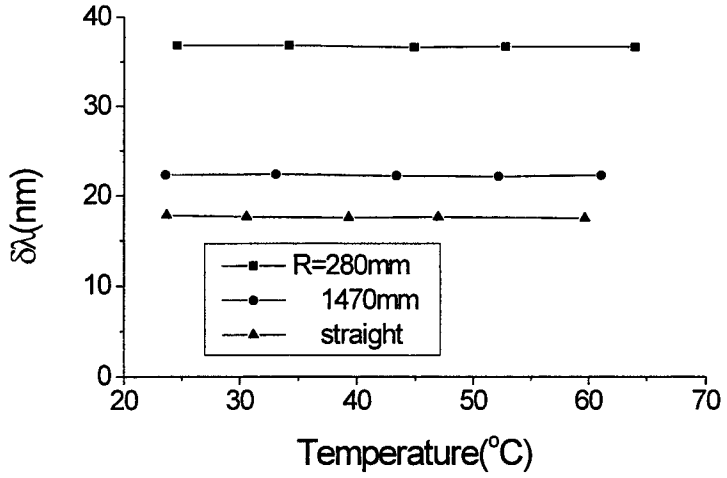


Fig. 6. The wavelength separation $\delta\lambda$ between the two dips in the transmission spectrum of a bent LPFG is insensitive to the temperature.

Figure 7 shows a bent LPFG mounted on a solid block. The temperature sensitivity of the LPFG can be expressed by

$$\frac{d\lambda^{(m)}}{dT} = \frac{d\Lambda}{dT}[n_{01} - n^{(m)}] + \Lambda\frac{d[n_{01} - n^{(m)}]}{dT} + \frac{\partial\lambda^{(m)}}{\partial R}\frac{\partial R}{\partial T} = A + \frac{\partial\lambda^{(m)}}{\partial R}\frac{\partial R}{\partial T} \quad (1)$$

where $\lambda^{(m)}$ is the resonance wavelength of the LPFG, Λ is the grating period, n_{01} and $n^{(m)}$ are the mode indices of the fundamental mode and the cladding mode, respectively, T is the temperature, R is the bend radius of the LPFG, $\partial R/\partial T$ is a function of R, T and α (the thermal-expansion coefficient of the mount), and $\partial\lambda^{(m)}/\partial R$ is a function of R.

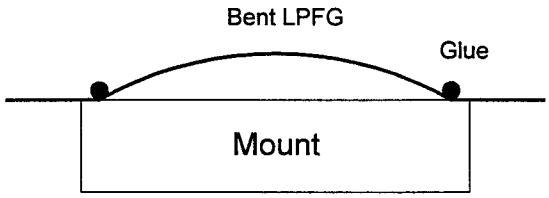


Fig. 7. A bent LPFG mounted on a solid block.

From our experimental results on a bent LPFG, the value of $\partial\lambda^{(m)}/\partial R$ is in the range from -0.01 to -1 nm/mm and decreases with increasing the bend radius of the LPFG. The sum of the first two terms in Eq.(1), denoted as A, is the inherent temperature dependence of the grating. The third term $(\partial\lambda^{(m)}/\partial R)(\partial R/\partial T)$ measures the effect due to the temperature dependence of the bend radius of the LPFG. The inherent temperature dependence A is about 0.05nm/°C. As can be seen from Eq.(1), we can control the temperature sensitivity of the LPFG by choosing a proper bend radius for the LPFG and a proper mount material, namely, a material with the proper thermal expansion coefficient. To obtain a temperature-insensitive LPFG, one should choose the bend radius R_o and the thermal-expansion coefficient α for the mount material such that

$$A + \frac{\partial \lambda^{(m)}}{\partial R} R_o \alpha = 0 \qquad (2)$$

For example, with A=0.04nm/°C, R_o=8.13m, and $\partial\lambda^{(m)}/\partial R$ =-0.3nm/mm, the thermal-expansion coefficient of the mount α should be 1.64×10^{-6}/°C, which is close to the thermal-expansion coefficient of stainless steel. On the other hand, if a LPFG with a large temperature sensitivity is needed, one should choose a mount with a thermal-expansion coefficient to make the left-hand side of Eq.(2) as large as possible in the temperature range of concern.

To demonstrate the principle experimentally, we mounted a 2cm long bent LPFG on a solid block as shown in Fig.7. Light from a pigtailed broadband ELED was launched into the LPFG. The temperature of the mounted LPFG was controlled with a Peltier heat pump. The transmission spectrum of the LPFG was measured with an optical spectrum analyzer and processed by a personal computer.

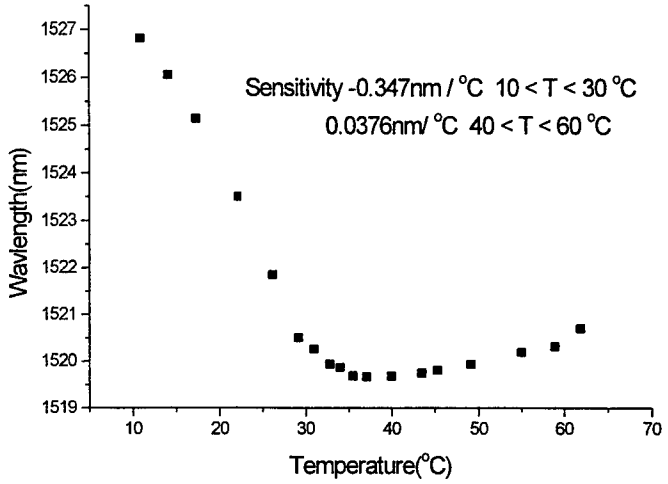
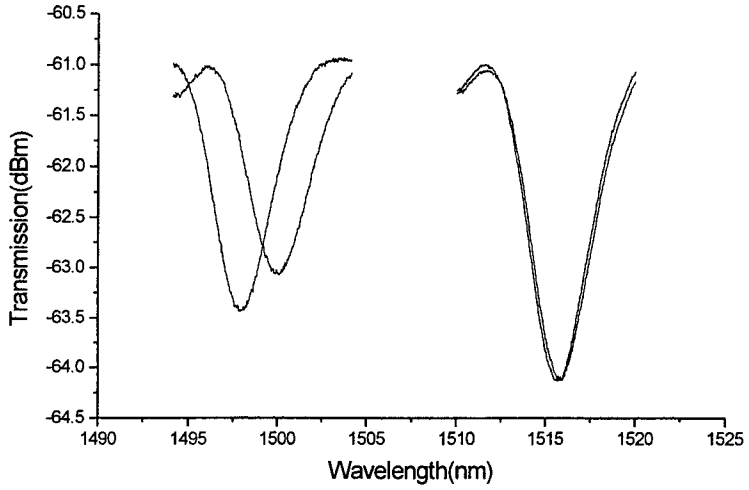


Fig. 8. Temperature dependence of the resonance wavelength of a slightly bent LPFG on an aluminum mount.

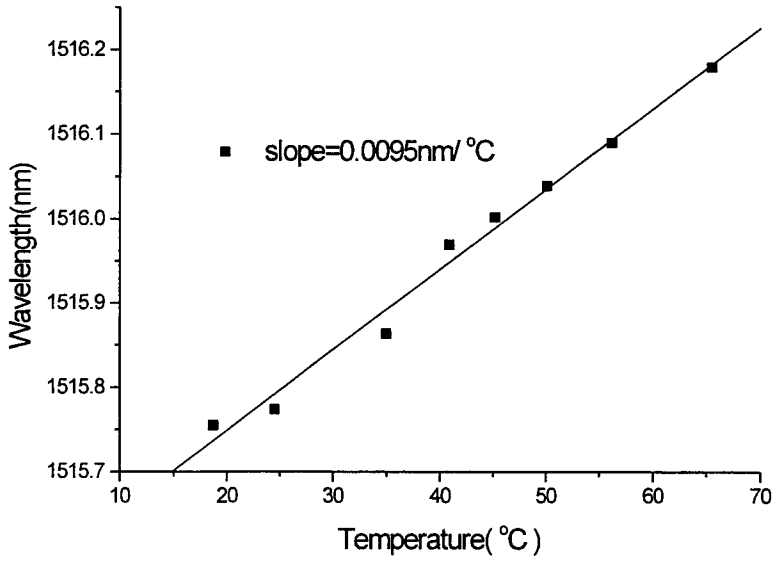
Figure 8 shows the measured temperature response of a slightly bent LPFG on an aluminum mount. When the temperature is less than about 30°C, the resonance wavelength shifts to shorter wavelengths as the temperature increases. The wavelength shift is approximately linearly proportional to the temperature change with a slope of -0.35nm/°C, which is about 7 times in magnitude of that of a straight bare LPFG. This shows that the thermally induced bending effect dominates over the direct thermal effect. In this case, we have $(\partial\lambda^{(m)}/\partial R)(\partial R/\partial T) \cong$ -0.4nm/°C. In the temperature range 35 - 40°C, the resonance wavelength becomes insensitive to the temperature. The wavelength shift due to the thermal change in the bend radius cancels largely the wavelength shift due to the direct thermal effect on the LPFG. When the temperature exceeds about 40°C, the resonance wavelength shifts to longer wavelengths. In this case, the bend radius becomes very large (i.e., the LPFG becomes nearly straight) and the temperature sensitivity is mainly contributed by the direct thermal effect on the LPFG. The temperature sensitivity in this range is about 0.038nm/°C, which is close to the temperature sensitivity of a straight LPFG, as expected.

We next mounted a bent LPFG on a stainless steel block and measured the temperature sensitivity of the resonance wavelength of the LPFG. As discussed in Section 2, there are two rejection bands in the transmission spectrum of a bent LPFG, which move apart from each other as the amount of bending increases. In the present experiment, both rejection bands were monitored. Figure 9(a) shows the two rejection bands of the bent LPFG on the stainless steel mount for a temperature change of 17°C. It can be seen from Fig.9(a) that the shift in the rejection band at the shorter wavelength is quite large while the shift in the rejection band at the longer wavelength is very small. The temperature sensitivity of the rejection band at the longer wavelength was measured more carefully and the results are shown in Fig.9(b). The measured temperature sensitivity is 0.01nm/°C from 15 to 70°C , which is only 1/5 of that of a straight bare LPFG.

We submerged the bent LPFG on the stainless steel mount into a NaCl solution to measure the concentration of NaCl in water. A variation of the NaCl concentration changed the refractive index of the solution and hence caused a shift in the resonance wavelength of the LPFG. To minimize the interference from the thermal effect, we measured the shift in the longer resonance wavelength of the bent LPFG, which was insensitive to the thermal effect. The experimental results are shown in Fig.10. A sensitivity of -0.35nm/molarity is obtained with the temperature sensitivity kept at 0.01nm/°C from 15 to 70°C.



(a)



(b)

Fig. 9. (a) Shift of the rejection bands of a bent LPFG on a stainless steel mount over a temperature change of 17°C. (b) Temperature response of the long-wavelength rejection band of the mounted LPFG.

We also measured the strain sensitivity of a bent LPFG on a solid mount. Figure 11 shows the experimental results for an unmounted straight LPFG and a mounted bent LPFG. A strain sensitivity of -49.2nm/%ε is obtained with the mounted LPFG, which is about 15 times of that of the unmounted straight LPFG.

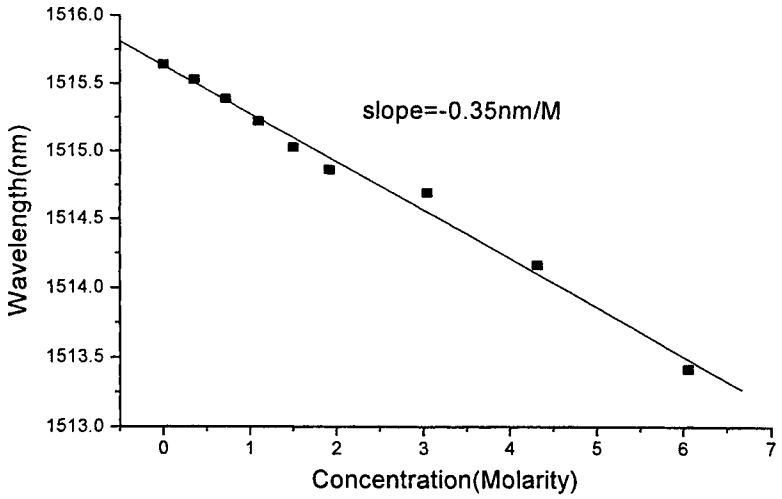


Fig. 10. Dependence of the wavelength shift on the NaCl concentration for the temperature-compensated LPFG sensor.

4. CONCLUSION

We have demonstrated experimentally that, by bending a LPFG, new rejection bands in the transmission spectrum of the LPFG can be produced. The wavelength separation between the new band and the original band and their relative strength depend on the bend radius. The strengths and the positions of these bands show no significant polarization dependence. We have proposed a simple method of controlling the temperature sensitivity of a LPFG based on mounting a bent LPFG on a solid block. We have shown that, by choosing the material for the mount properly, it is possible to enhance or suppress the temperature sensitivity of the LPFG over a wide temperature range. Using this technique, we have demonstrated a LPFG-based sensor for the measurement of the concentration of NaCl in water with a temperature sensitivity as low as 0.01nm/°C from 15 to 70°C. On the other hand, using the same technique, we have achieved temperature and strain sensitivities as large as 0.35nm/°C and -49.2nm/%ε, respectively. Our findings should lead to more practical applications with LPFG's in areas like EDFA gain flattening and optical sensors.

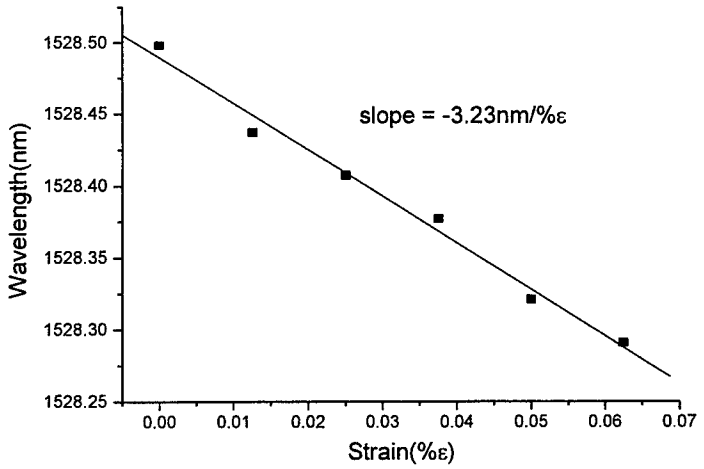
5. ACKNOWLEDGEMENTS

This research is supported by the University Grants Committee (Hong Kong Government) under a research grant. Z. Chen wishes to thank the support of Fujian Nature Science foundation.

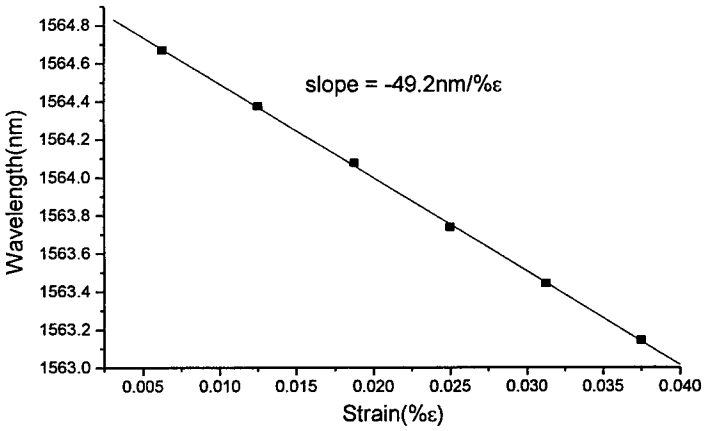
6. REFERENCES

1. A. M. Vengsarkar, P. J. Lemaire, J. B. Judkins, V. Bhatia, T. Erdogan, and J. E. Sipe, "Long-period fiber gratings as band-rejection filters", J. Lightwave Technol., vol. 14, pp58-65, 1996.

2. A. M. Vengsarkar, J. R. Pedrazzani, J. B. Judkins, P. J. Lemaire, N. S. Bergano, and C. R. Davidson, "Long-period fiber-grating-based gain equalizers", Opt. Lett., vol. 21, pp336-338, 1996.



(a)



(b)

Fig. 11. Dependence of the wavelength shift on the applied strain for (a) an unmounted straight LPFG and (b) a mounted bent LPFG.

3. V. Bhatia, T. D. Alberto, K. A. Murphy, R. O. Claus, and C. P. Nemarich, "Comparison of optical fiber long-period and Bragg grating sensors", SPIE vol. 2718, pp110-121, 1996.

4. V. Bhatia, D. Campbell, and R. O. Claus, "Simultaneous strain and temperature measurement with long-period gratings", Opt. Lett., vol. 22, no. 9, pp.648-650, 1996.

5. M. G. Xu, R. Maaskant, M. M. Ohn, and A. T. Alavie, "Independent tuning of cascaded long period fibre gratings for spectral shaping", Electron. Lett. vol.33, no.22, pp.1893-1895, 1997.

6. H. Ke, K. S. Chiang, and J. H. Peng, "Analysis of phase-shifted long-period fiber gratings", IEEE Photon. Tech. Lett., vol.10, no.11, pp.1596-1598, 1998.

7. F. Bakhti and P. Sansonetti, "Realization of low back-reflection, wide-band fiber bandpass filters using phase-shifted long-period gratings", in OFC'97, FB4, pp.349-350, 1997.

8. H. J. Patrick, A. D. Kersey, and F. Bucholtz, "Analysis of the response of long period fiber gratings to external index of refraction", J. Lightwave Technol., vol.16, No.9, pp.1606-1612, 1998.

9. B. H. Lee and J. Nishii, "Bending sensitivity of in-series long-period fiber gratings", Opt. Lett., Vol.23, No.20, pp1624-1626, 1998.

10. H. J. Patrick, C. C. Chang, and S. T. Vohra, "Long period fibre gratings for structural bend sensing", Electron. Lett. 34, 1773 (1998).

11. V. Bhatia, D K. Campbell, T D'Alberto, G. A. T. Eyck, D. Sherr, K. A. Murphy, and R. O. Claus, "Standard optical fiber long-period gratings with reduced temperature sensitivity for strain and refractive-index sensing", in OFC'97, FB1, pp.346-347, 1997.

12. K. Shima, K. Himeno, T. Sakai, S. Okude, A. Wada, R. Yamauchi, "A novel temperature-insensitive long-period fiber grating using a boraon-codoped-germanosilicate-core fiber", in OFC'97, FB2, pp.347-348, 1997.

SESSION 3

Biomedical and Biochemical Sensors I

Biochip technology: a triad of micro-electro-mechanical (MEM), biochemical, and photonic technologies

Arthur Chiou
Department of Electrical Engineering
National Dong Hwa University
Hualien, Taiwan

ABSTRACT

Biotechnology has recently been regarded as the technology of the 21st century. One important advancement of the biotechnology is the biochip technology which integrates the conventional biotechnology with semiconductor processing, micro-electro-mechanical (MEM), optoelectronic, and digital signal and image acquisition and processing technologies. In this paper, I give a brief introduction to the biochip technology in general, and describe the DNA micro-array technology as a specific example of this relatively new and burgeoning interdisciplinary research and development.

Keywords: Biochip, gene-chip, DNA-microarray, biosensor, bioinformatics, polymerase chain reaction (PCR), hybridization.

1. INTRODUCTION

We have all witnessed the impact of information science and technology in the last decades. As the information science and technology continue to advance at a tremendous rate through the next century, the biotechnology has also emerged as a new rising star, and has often been cited as the technology of the 21st century.[1,2,3] One of the key factors enabling the rapid advance and progress of the information science and technology is the semiconductor integrate circuit (IC) chip technology which allows low-cost, mass-production of millions of transistor circuits on a semiconductor chip of area on the order of a square centimeter. Likewise, the new biotechnology is likely to be escalated by the biochip technology[4-13] which integrates the conventional biotechnology with the semiconductor processing, the micro-electro-mechanical (MEM), optoelectronic, and digital signal and image acquisition and processing technologies.

The term "biochip" has appeared recently in many publications and has taken a variety of nuances of meaning. In the most generic sense, any device or component incorporating biological (or organic) materials, either extracted from biological species or synthesized in a laboratory, on a solid substrate can be regarded as a biochip. From the practical point of view and from the analogy taken from the IC chip, however, it often tacitly implies some miniaturation, usually in the form of micro-array format, and the possibility of low-cost mass production. Some of the examples that meet the qualification stated above include the electronic nose[14-18] (e-nose) or artificial nose chip, the electronic tongue,[19] the PCR (polymerase chain reaction) chip,[20-23] the DNA micro-array chip (or the gene chip),[4-13] and the (biochemical) lab-on-a-chip.[23-26] In this paper, I give a brief overview of the basic principle, the fabrication techniques, and potential applications of the DNA micro-array technology as a paradigm of the biochip technology. The main objective of this paper is to point out the intrinsic interdisciplinary nature of the biochip technology; a synergistic collaboration of scientists and engineers from different disciplines is essential to push this novel technology from a lab curiosity to practical devices and systems.[27,28]

2. THE DNA MICRO-ARRAY CHIP

2.1. BASIC CONCEPT

A DNA micro-array chip refers to a two-dimensional array of small reaction cells (each on the order of 100μm x 100μm) fabricated on a solid substrate. The solid substrate can be a silicon wafer, or a thin sheet of glass or plastic, or a nylon membrane. In each reaction cell, trillions of polymeric molecules of a specific sequence of single-stranded DNA fragment are immobilized as illustrated in Fig. 1. The DNA fragments can be either short (~ 20 to 25) sequences of oligonucleotides (oligos) or longer strands of complementary DNA (cDNA).[29,30] The specific sequence of oligos or bases (for example, ATTAGCT.....) in each cell is pre-selected or designed based on the intended application. The known sequences of single strand DNA fragments immobilized on the substrate are often called the probes. When unknown fragments of single strand

Correspondence: Email: aechiou@mail.ndhu.edu.tw; Tel: 886-38-662-500, x22201; Fax: 886-38-662-300

Part of the SPIE Conference on Advanced Photonic Sensors and Applications
Singapore • November–December 1999
SPIE Vol. 3897 • 0277-786X/99/$10.00

DNA samples (often called the target) react (or hybridize) with the probes on the chip, double strand DNA fragments are formed where the target and the probe are complementary according to the base pairing rule (A paired with T, and G paired with C).[29,30] To facilitate the diagnosis (or the analysis) of the hybridized chip, the target samples are often labeled with a tag, such as a fluorescent, a dye, or a radio-isotope molecule. When the targets contain more than one types of samples, each is labeled with its own distinguishable tag. Depending on the size of the array, the DNA micro-array chip described above provides a platform where the unknown target (or targets) can potentially be identified with very high speed and high throughput by matching with tens of thousands of different types of probes via hybridization in parallel.

Depending on the specific types of sequences of bases immobilized on the substrate, or often simply on personal preference, a wide variety of names such as, the DNA chip, the gene chip, the oligomer chip, or simply the biochip have been used in the literatures for this class of technologies. For simplicity, I use "DNA chip" as a generic term in this paper to represent any variation within this class of technologies.

Since the basic operating principle of a DNA micro-array chip is similar (but not identical) to that of other types of biosensors, such as the e-nose, it is useful to compare the two systems. In a typical biosensor, different types of bioreceptors are often immobolized on a solid substrate in the form of an array as illustrated schematically in Fig. 2. When a test target, usually in the form of a volatile gas or vapor, is passed towards the substrate its chemical constituents (or molecules) will adhere to different bioreceptors with varying degrees of affinity. A specific target is often characterized and identified by a unique distribution of the molecular mass over a set of bioreceptors. The molecular mass adhered to each reaction cell of bioreceptors is monitored by a proportional change in a certain physical properties such as the frequency of vibration (in a quartz micro-balance), or the phase retardation of the surface acoustic wave (in a surface acoustic wave SAW sensor). [14-17]

One major difference between the DNA chip and other types of biosensors is the specificity. In a typical biosensor such as the e-nose, the target molecules adhere to not just one specific type of receptors, but to many (and often all) of them with varying degrees of affinity. In the DNA chip, the formation of the double-strand DNA by the base-pairing rule is so specific that under favorable conditions even a single pair mismatch (known as single-nucleotide-polymorphism, or SNP) can be detected and identified.[31]

2.2. FABRICATION TECHNIQUES

DNA chips are often fabricated by one of the following three popular techniques: 1. Mechanical Spotting, 2. Ink-Jetting, and 3. Photolithography.[5-8,11-13] In the first two techniques, all the oligonucleotide sequences to be immobilized on the substrate are pre-synthesized (by conventional bio-chemical techniques), and stored in a set of individual containers. In the Mechanical Spotting technique, a robotic arm with an array of pins is used to transfer the pre-synthesized DNA sequences on to the solid substrate by first dipping into the containers and then touching the substrate. In the second approach, each pre-synthesized DNA sequence is injected to form a two-dimensional array on the substrate (by the ink-jet printing technology). In the third approach, the desired sequences of DNA fragment are synthesized in parallel, layer by layer, using photolithographic technology as is illustrated in Fig. 3.[32,33] In this approach, the substrate is first pre-coated with appropriate linkers and protectors. An ultraviolet (uv) light is used to illuminate the substrate through a photolithographic mask to remove the protectors at specific sites where one particular type of base (say A, for example) is immobilized (via proper biochemical reactions) as illustrated in Fig. 3(a). Likewise, a second mask is then used to immobilize another kind of base (T, for example) at other selected sites [Fig. 3(b)]. Four photolithographic masks are thus needed to fill the first layer with all the four types of bases. Successive rounds of deprotection (via uv illumination) and immobilization (via biochemical reactions) are carried out to synthesize the desired sequences of bases at each site, layer by layer, in parallel. For a sequence of "N" bases, the total number of photo masks required is 4N, in general. This approach is hence more suited for the synthesis of short sequences of bases (typically, no more than 30 bases). Some advantages and disadvantages of the three techniques are listed in Table 1. The specifications of selected examples of chips successfully fabricated by different chip manufacturers using these three techniques are listed in Table 2.

Another ingredient that has been added (by Nanogen) to the biochips is the fabrication of a two-dimensional array of electrodes (in the form of micro-strips) in concomitant with the array of reaction cells.[7,8,11] During the reaction (or hybridization) cycle, an appropriate positive voltage can be applied to selected sites (or lines of sites) to concentrate the negatively charge DNA targets to selected sites. After the hybridization, during the "wash" cycle, an appropriate negative voltage can be applied to the sites to drive away (or wash out) the mismatched DNA targets. This technique can potentially reduce the hybridization time from hours to minutes.

3. OTHER COMPONENTS OF BIOCHIP TECHNOLOGY

The biochip fabrication described in the previous section represents just one component of the biochip technology. From the application point of view, all the other components including the sample (or test target) preparation, signal detection and identification from the hybridized chips, and the visualization, interpretation, and management of the data are equally important. The technologies involved in each of these components are briefly outlined below in this section.

3.1. THE PCR (POLYMERASE CHAIN REACTION) CHIP

In many applications, the DNA samples to be tested (or identified) needs to be purified (from a mixture of other undesirable substances) and amplified to increase its concentration by orders of magnitude to a detectable level. The amplification of DNA samples is often accomplished by polymerase chain reaction (PCR)[20-22] in which a minute amount (the seed) of DNA molecules are repeatedly thermal cycled through a sequence of temperature stages as illustrated in Fig. 4. The concentration of the DNA molecules is doubled at the end each cycle, consisting of denaturation (the splitting of each double strand DNA molecule into two complementary single strands), primer annealing, and primer extension (the reconstruction of double strand DNA molecules form single strand components by supplying the proper reagents at the proper temperatures. The conventional PCR system can also be miniaturized and fabricated on a chip by micro-machining and microfluidic technologies.[23] A simplified schematic diagram of a continuous flow PCR chip is illustrated in Fig. 5.

3.2. SIGNAL READOUT FROM HYBRIDIZED DNA CHIPS.

The next step after the hybridization process is the signal detection or readout to determine the sites where the sequence of the unknown DNA targets complement (or match) that of the probes and stick to the site via the formation of double strand DNA molecules. The particular choice of signal detection technique goes hand in hand with the choice of the tag attached to the DNA target molecules (see Fig. 1 and the associated description in Section 2.1). For example, if one chooses to use fluorescent signal for detection, appropriate fluorescent molecules will be used as the tags; the signal readout system will then consist of an appropriate light source (and filter) for excitation, a photo-multiplier tube or a CCD camera (with an appropriate filter) for the detection of the fluorescent emission, and a personal computer for data acquisition and post processing. A schematic simplified diagram of such a signal detection system is shown in Fig. 6. Although one can, in principle, integrate either the CCD or the CMOS imager with the DNA micro-array on the same silicon wafer chip,[10,34] such an approach will unavoidably increase the complexity, and hence the cost, of the fabrication. Unless the fabricated chips are re-usable (with a reasonable large number of reusable cycles), such an approach may not be competitive from an economic point of view. Many commercial optical readout systems have been developed in recent years.[8,11,35] An alternative optical technique for signal detection is to use dye molecules as the tags and use colorimetry for signal detection and discrimination.[36] Techniques, other than optical, such as radio-isotopic, electronic, or time-of-flight mass spectrometric, have also been investigated.

3.3. BIOINFORMATICS

Depending on the size of the DNA micro-array, and the nature of the test targets, the raw signal (or data) obtained from the readout system described in the previous section may require a significant amount of post processing (including visualization, interpretation and management) to become useful information. Bioinformatics, the processing, visualization, interpretation, and the management of biological and biomedical data, has emerged as an important area of research.

3.4. POTENTIAL APPLICATIONS

Apparently, almost all applications of biotechnology or genetic engineering that rely upon information related DNA sequences (or gene sequences) will benefit from the biochip technology. Current research and development of biochip technology have focused on applications such as prognosis or diagnosis of genetic diseases, drug design and testing of drug efficacy, toxicology, and genetic agricultural products.[1,4,6,8]

4.0. SUMMARY AND CONCLUDING REMARKS

In this paper, I present a brief overview of the biochip technology to illustrate the interdisciplinary nature of this novel technology. The fabrication of biochip chips requires close collaboration of scientists and engineers form different

disciplines such as photolithography, micro-machine (including micro-electro-mechanical, or MEM, and microfluidic technologies), and biochemistry. The judicial choice of the specific DNA sequences on the chips is dictated by the target applications and may involve expertise from medical, pharmaceutical, or biological sciences. The technologies associated with the readout system may involve optics and opto-electronics, analytical chemistry, or semiconductor device physics. Finally, the processing, visualization, interpretation, and management of data requires the collaborative effort of computer scientists, mathematicians, and the end users for whom the specific chips are designed.

5.0. ACKNOWLEDGEMENT

This work is supported by the Electro-Optics Science and Technology Committee, National Science Council, Taiwan, under Contract Number NSC 88-2736-L-259-001. I also want to express my gratitude to Professor S. J. Y. Ting (Tunghai University) and Dr. Konan Peck (Institute of Biomedical Sciences, Academia Sinica) for leading me into this exciting world of Biotechnology.

6.0. REFERENCES

[1] "Biotechnology," in *Forbes ASAP*, 1999, pp. 44-109.
[2] F. B. Rudolph and L. V. McIntire, "Biotechnology: Science, Engineering, and Ethical Challenges for the 21st Century," Washington, D. C.: Joseph Henry Press, 1996, pp. 278.
[3] A. Kornberg, *The Goldern Helix: Inside Biotech Ventures*: University Science Books, 1996.
[4] "Biochips perform genetic analyses rapidly and economically," in *R&D Magazine*, 1999, pp. S-54.
[5] J. D. Hoheisel, "Oligomer-chip technology," *Tibtech*, vol. 15, pp. 465-469, 1997.
[6] D. Gerhold, T. Rushmore, and C. T. Caskey, "DNA chip: promising toys have become powerful tools," *TIBS*, vol. 24, pp. 168-173, 1999.
[7] G. Ramsey, "DNA chips: State-of-the art," *Nature Biotechnology*, vol. 16, pp. 40-44, 1998.
[8] A. Marshall and J. Hodgson, "DNA chips: An array of possibility," *Nature Biotechnology*, vol. 16, pp. 27-31, 1998.
[9] C. H. Mastrangelo, M. A. Burns, and D. T. Burke, "Microfabricated devices for genetic diagnostics," *Proceed. IEEE*, vol. 86, pp. 1769-1786, 1998.
[10] T. Vo-Dinh, J. P. Alarie, N. Isola, D. Landis, A. L. Wintenberg, and M. N. Ericson, "DNA Biochip using a phototransistor integrated circuit," *Analytical Chemistry*, vol. 71, pp. 358-363, 1999.
[11] http://www.gene-chips.com/
[12] http://www.cs.washington.edu/homes/jbuhler/research/array/
[13] http://www.the-scientist.library.upenn.edu/yr1999/may/profile1_990524.html
[14] H. T. Nagle, S. S. Schiffman, and R. Gutierrez-Osuna, "The how and why of electronic noses," *IEEE Spectrum*, vol. 35, pp. 22-31, 1998.
[15] W. Gopel and T. Weiss, "The design of smelling," *IEEE Spectrum*, vol. 35, pp. 32-34, 1998.
[16] H. Baltes, D. Lange, and A. Koll, "The electronic nose in Lilliput," *IEEE Spectrum*, vol. 35, pp. 35-38, 1998.
[17] T. A. Dickson, J. White, J. S. Kauer, and D. R. Walt, "Current trends in "artificial-nose" technology," *Tibtech*, vol. 16, pp. 250-258, 1998.
[18] A. D'Amico, C. D. Natale, A. Macagnano, F. Davide, A. Mantini, E. Tarizzo, R. Paolesse, and T. Boschi, "Technologies and tools for mimicking olfaction: status of the Rome "Tor Vergata" electronic nose," *Biosensors & Bioelectronics*, vol. 13, pp. 711-721, 1998.
[19] K. Toko, "Electronic tongue," *Biosensors & Bioelectronics*, vol. 13, pp. 701-709, 1998.
[20] K. B. Mullis, "The unusual origin of the polymerase chain reaction," *Sc. America*, vol. 270, pp. 56-65, 1990.
[21] K. B. Mullis, F. Ferre, and R. A. Gibbs, *The polymerase chain reaction*. Boston, MA: Birkhauser, 1994.
[22] E. A. Erlich, *PCR technology: principles and applications for DNA amplifications*. Oxford, UK: Oxford Univ. Press, 1992.
[23] A. T. Woolley, D. Hadley, P. Landre, A. J. deMello, R. A. Mathies, and M. A. Northrup, "Functional integration of PCR amplification and capillary electrophoresis in a microfabriacted DNA analysis device," *Anal. Chem.*, vol. 68, pp. 4081-4086, 1996.
[24] M. A. Burns, C. H. Mastrangelo, T. S. Sammarco, P. F. Man, J. R. Webster, B. N. Johnson, B. Foerster, D. K. Jones, Y. Fields, A. R. Kaiser, and D. T. Burke, "Microfabricated structures for integrated DNA analysis," *Proc. Nat. Acad. Sci. USA*, vol. 93, pp. 5556-5561, 1996.
[25] R. C. Anderson, Z. B. G. J. Bogden, T. D. Dawes, J. Winkler, and K. Roy, "Microfluidic biochemical analysis system," presented at Proc. 1997 IEEE Int. Conf. Solid State Sens. Actuat.. 1997.

[26] "Highly sensitive optical DNA assay system," in *New Technology Japan*, vol. 26, 1999, pp. 8-11.
[27] D. Malakoff, "Genomic, nanotech centers open $200 million push by Havard," *Science*, vol. 283, pp. 610-611, 1999.
[28] N. Metzger and R. N. Z. Malakoff, "Interdisciplinary research: from belief to reality," *Science*, vol. 283, pp. 642-643, 1999.
[29] C. R. Calladine and H. R. Drew, *Understanding DNA: The Molecules and How It Works*, 2nd ed. San Diego: Academic Press, 1997.
[30] G. R. Carter and S. M. Boyle, *All You Need to Know About DNA, Genes and Genetic Engineering: A Concise, Comprehensive Outline*. Springer, Illinois: Charles C Thomas Publisher, Ltd, 1998.
[31] P. N. Gilles, D. J. Wu, C. B. Foster, P. J. Dillon, and S. J. Chanock, "Single nucleotide polymorphic discrimination by an electronic dot blot assay on semiconductor microchips," *Nature Biotechnology*, vol. 17, pp. 365-370, 1999.
[32] G. McGall, J. Labadie, P. Brock, G. Wallraff, T. Nguyen, and W. Hinsberg, "Light-directed synthesis of high-density olionucleotide arrays using semiconductor photoresists," *Proc. Natl. Acad. Sci. USA*, vol. 93, pp. 13555-13560, 1996.
[33] S. P. A. Fodor, J. L. Read, M. C. Pirung, L. Stryer, A. T. Lu, and D. Solas, "Light-directed, spatially addressable parallel chemical synthesis," *Science*, vol. 251, pp. 767-773, 1991.
[34] P. Caillat, M. Belleville, F. Clerc, and C. Massit, "Active CMOS biochips: An electro-addressed DNA probe," presented at IEEE International Solid State Circuit Conference, 1998.
[35] S. J. B. Jr., R. Gangadharan, M. McMahon, J. Denman, R. Gonzales, L. G. Mandoza, and M. Eggers, "A proximal CCD imaging system for high-throughput detection of microarray-based assays," in *IEEE Engineering in Medicine and Biology*, 1999, pp. 120-122.
[36] J. J. W. Chen, R. Wu, P. C. Yang, J. Y. Huang, Y. P. Sher, M. H. Han, W. C. Kao, P. J. Lee, T. F. Chiu, Y. W. Chu, C. W. Wu, and K. Peck, "Profiling expression patterns & isolating differentially expressed genes by cDNA microarray system colorimetry detection," Genomics, 51(3), 313-324, 1998.

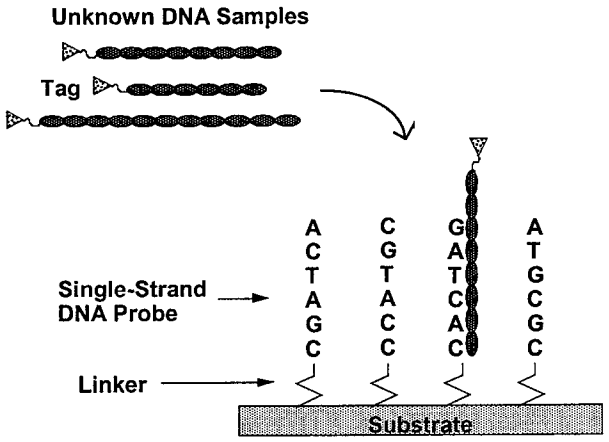


Fig.1 A schematic illustration of the basic principle of a DNA microarray chip.

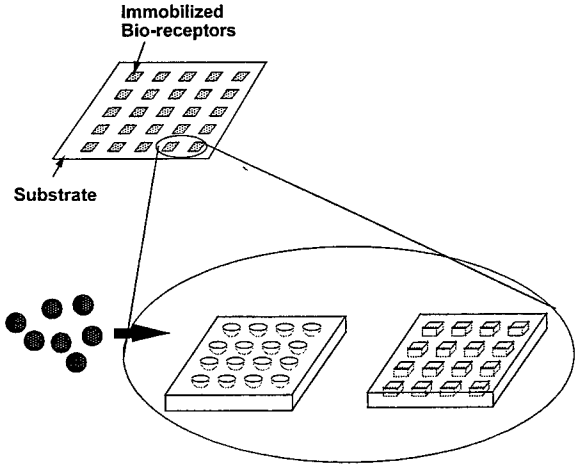


Fig. 2 A schematic illustration of a generic biosensor array.

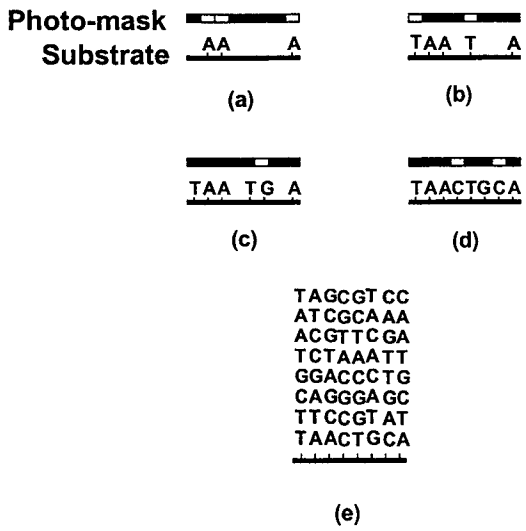


Fig. 3 In-situ synthesis of DNA fragments (or sequences of oligonucleotides) by photolithographic technique.

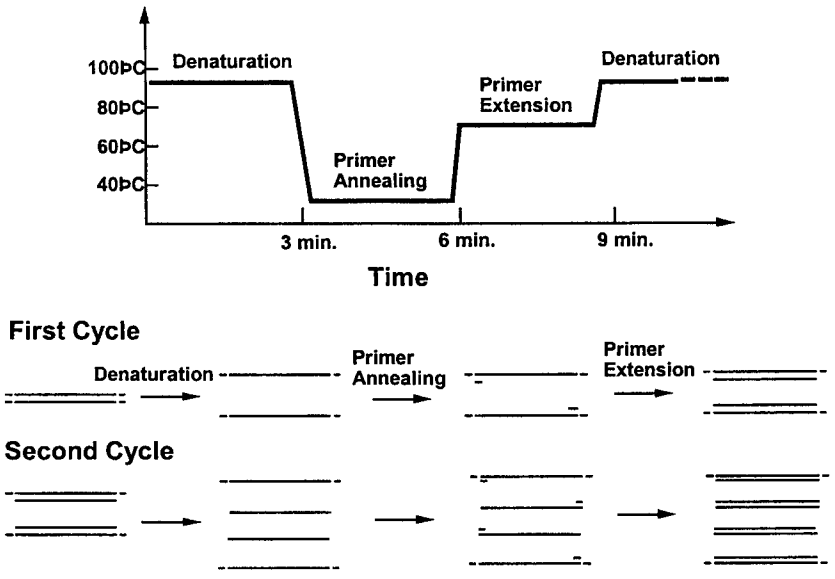


Fig. 4. Thermal cycles in polymerase chain reaction (PCR) for DNA amplification.

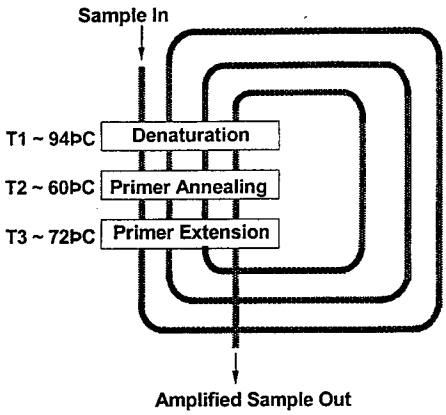


Fig. 5. A simplified schematic diagram of a continuous flow PCR chip

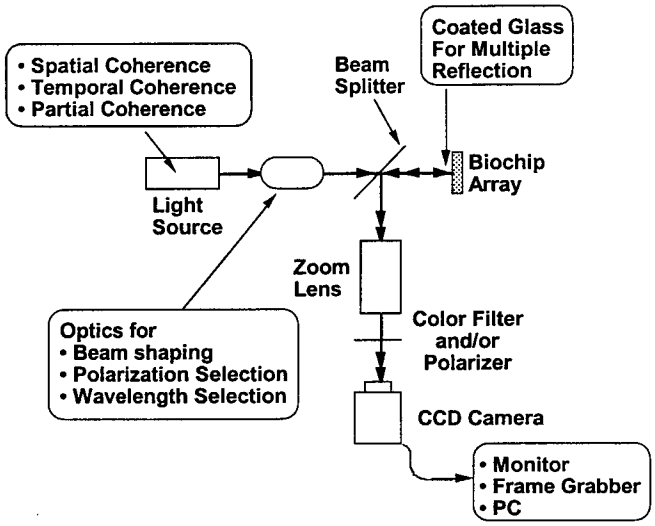


Fig. 6. A schematic illustration of a CCD-based imaging system for biochip signal detection.

Table 1. A comparison of some advantages and disadvantages of three different techniques for biochip fabrication.

Technique	Advantages	Disadvantages
Photolithography	• versatile • high density • large array size • mass production	• high up-front cost • time-consuming design • low yield
Mechanical Microspotting	• affordable • simple	• sample storage • low throughput
Ink-Jetting	• high throughput • high density • medium cost	• sample storage

Table 2. Specifications of selected examples of biochips fabricated by different chip manufacturers using three different fabrication techniques.

- **Photolithography**
  - **Affymetrix Chips**
    - **~ 1.6cm^2**
    - **~ 400,000 cells;**
    - **~ 10^7 oligonucleotides per cell**
- **Mechanical Microspotting**
  - **Synteni Chips**
    - **~ 3.6 cm^2**
    - **~ 10,000 cells;**
- **Ink-Jetting**
  - **Incyte Pharmaceuticals, or Protogene Chips**
    - **~ 10,000 cells per cm^2 (demonstrated)**

Precision blood-leak-detector with high long-time-stability

Christos Georgiadis[1], Wolfram Kleuver[2]

[1]Universität-Gesamthochschule Siegen, Zentrum für Sensorsysteme (ZESS)
Paul-Bonatz-Straße 9-11, 57068 Siegen, Germany
[2]aicoss GmbH, Birlenbacher Str. 14, 57078 Siegen, Germany

ABSTRACT

With this publication a precision blood-leak-detector is presented. The blood-leak-detector is used for recognition of fractures in the dialyzer of a kidney-machine. It has to detect safely a blood flow of $Q_{blood} > 0.5$ml/min to exclude any risk for the patient.

A lot of systems exist for blood-leak-detection. All of them use the same principle. They detect the light absorption in the dialyze fluid.

The actual used detectors are inferior to the new developed sensor in resolution and long-time stability. Regular tests of the existing systems and high failure rates are responsible for the high maintenance costs.

It was always necessary to adjust the single sensors to the belonging kidney-machine. That's not required for the new sensor. The new sensor reduces manufacturing-, installation- and maintenance-costs.

Furthermore this paper gives some detailed information about the accuracy, resolution and the long time stability of new sensor.

Keywords: kidney-machine; blood-leak-detector; dialyzer; optical design; precision inspection; high speed analog design

1 INTRODUCTION

The new precision blood-leak-detector has been developed for a kidney machine manufacturer. The sensor should be replace the previous detector. The interface is the same like the previous sensor and the integration in the existing system is possible without difficulties.

The control-unit of the kidney machine calculates the blood value. At the critical blood concentration an alarm will be triggered.

The use of the new sensor guarantees a better monitoring of fractures in the dialyzer of a kidney-machine.

A high flexible circuit architecture allows it to integrate the new sensor to other kidney machines.

Furthermore it is conceivable to process measurements in the catheter with the same principle to be always informed over the patients condition.

For further author information-
C.G. (correspondence): Email: geo@zess.uni-siegen.de
Telephone: ++49-271 / 740-2449; Fax: ++49-271 / 740-2336
W.K.: Email: kleuver@aicoss.de
Telephone: ++49-271 / 77309-50; Fax: ++49-271 / 77309-59

Part of the SPIE Conference on Advanced Photonic Sensors and Applications
Singapore • November–December 1999
SPIE Vol. 3897 • 0277-786X/99/$10.00

2 SYSTEM DESCRIPTION

To control the mechanical functionality of the dialyze membrane the manufactures use all the same principle. Differences between the realization of the optical path and the electrical circuit are deciding for the success of the functionality.

There exist two possibilities to realize the optical path the reflection method and the transmitted light method. These are the common methods.

The realization of the reflection method is very difficult. A lot of mechanical parameters must be adjusted. Furthermore is the component expenditure higher. The sequential higher costs makes this concept uninteresting. The basic structure of the reflection method is shown in Figure 1.

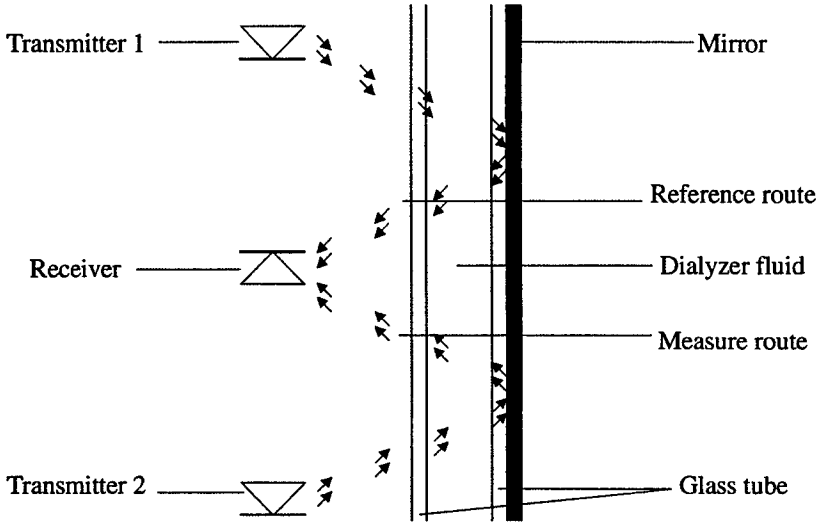


Figure 1: Basic structure of the reflection method

The realization of the electrical path is also many-sided. The used luminous source can be constant light or pulsed light. To minimize the secondary light source sensitivity the use of pulsed light is better.

For a high independent of temperature it is required to control the evolution of light. This method warranties also the high long time stability.

2.1 Measurement principle

The question is, how to minimize the influence of the dialyze fluid cloudiness and how to detect the blood concentration. The true absorption coefficient of blood with infrared light is lower as the one of the visible spectrum. In the visible spectrum is the difference between red light and green light high. Red light has almost the same true absorption coefficient as infrared light. The conclusion is that the operation with green light is requisite to detect the blood concentration.

In the following reflections it is important to consider the dispersion of the system and the requisite measuring range 0-10‰ of blood in the dialyze fluid.

In this case is the dispersion a result of the blood-dialyze mixture, because only the blood plasma dissolves. Further are the absorption and the dispersion dependent by the optical path length, the material constant and the concentration of solute.

The relation from green light, red light and infrared light in dependence from the concentration of blood in dialyze fluid is shown in Diagram 1.

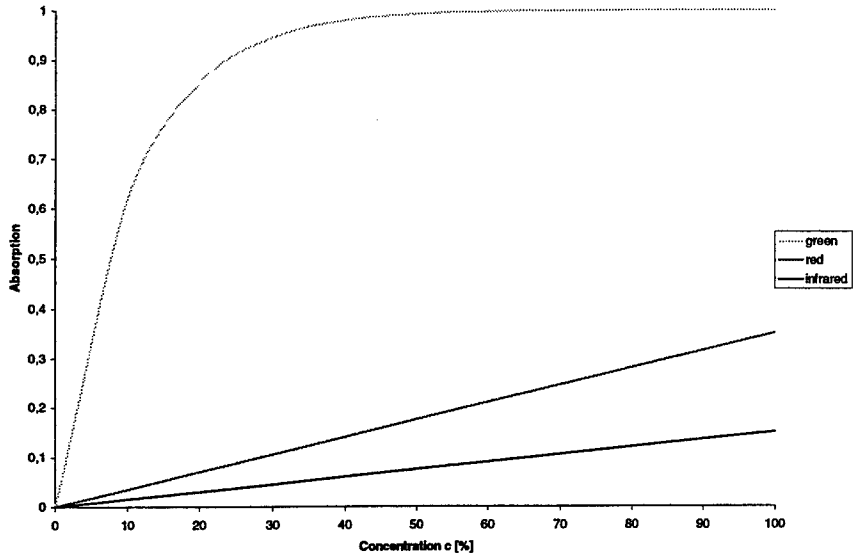


Diagram 1: Concentration of blood in dialyze fluid belonging to it absorption by different wavelengths

The general formula of absorption is: $A = 1 - e^{-\varepsilon c x}$

The extinction coefficient ε is conditional on the compound of the dialyze fluid, the blood density c on the membrane condition and the optical distance x on the developed sensor.

This formula is not correct for the existing application because the dispersion is neglected. The dispersion is also dependent on conditional, density and the optical distance.

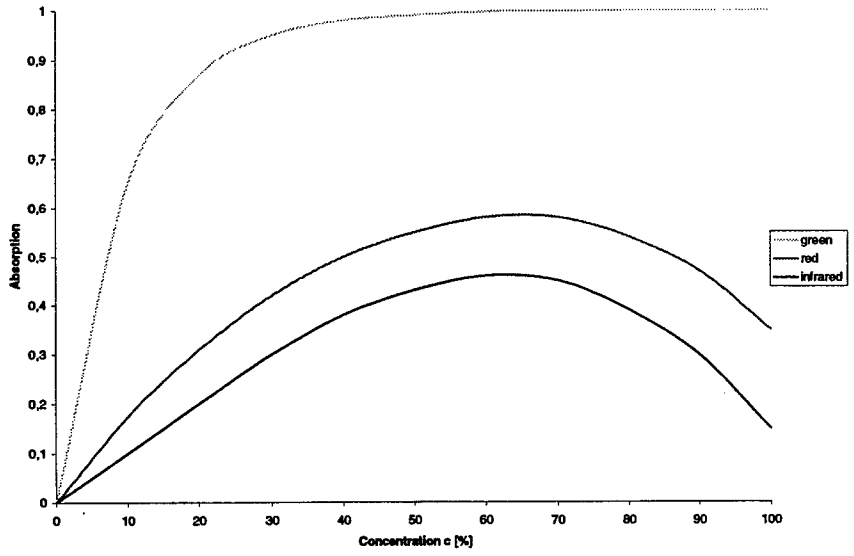


Diagram 2: Absorption and dispersion dependent on blood concentration and wavelength

The effect of scattering on the light transmitting capacity is shown in Diagram 2. The diagram illustrates that there is no more scattering left by a blood concentration of 100%. Only dispersion is effective.

For the existing application is only the measuring range of 0–10‰ interesting. At this range is the extinction of green light by dispersion smaller as by absorption. But the red and infrared light have a higher extinction by dispersion. The differences between the spectral regions are shown in Diagram 3.

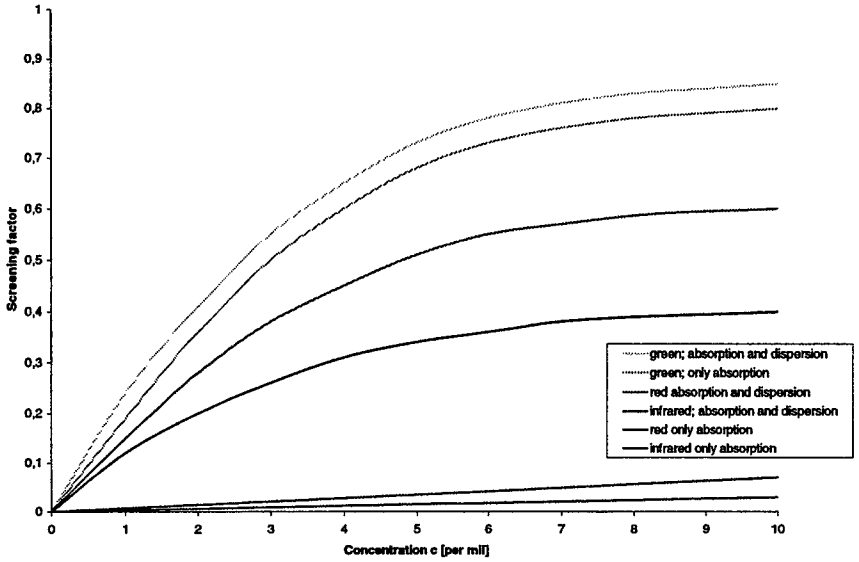


Diagram 3: Comparison of absorption and dispersion in dependence on possible measurement range

Important for further reflections are the errors of measurement. While testing the old common sensors four errors of measurement were determined:

- optical changing of light transmission ratio
- measurement medium changing without blood concentration
- beam intensity changing
- receiver sensitivity changing

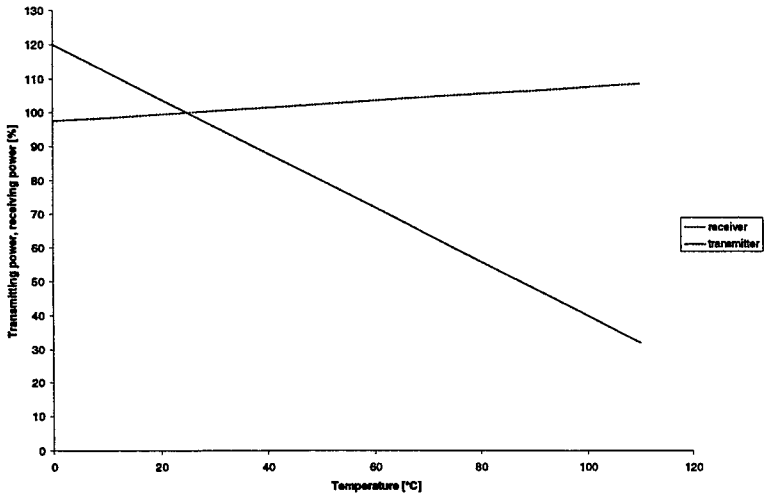


Diagram 4: Intensity of radiation I of a LED and luminous intensity S of a photodiode

The light transmission ratio can be confused by depositions on the glass panel. Other dirt accumulation can be exclude.

The variation of the measurement medium has three reasons:

- influence of air bubbles
- consistence of dialyze fluid
- clinical picture of the patient

The last two errors are conditional on temperature. When using a LED as transmitter and a photodiode as receiver occurrence the dependence on temperature as shown in **Diagram 4**.

2.2 Optical Path

To realize a new sensor without all this flaws many of tests had to performed. The result of the optical unit is shown in Figure 2.

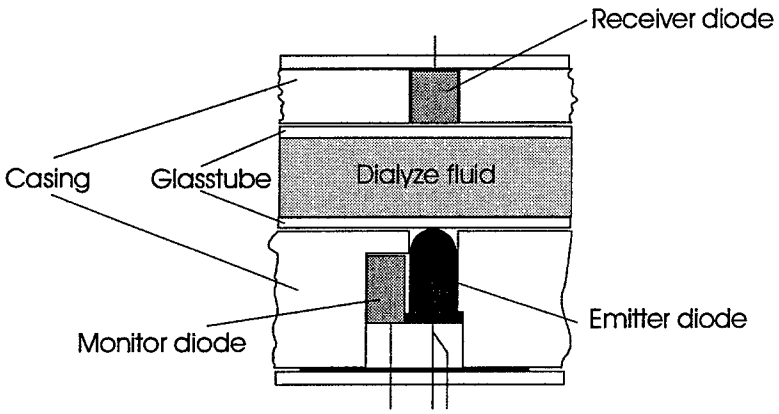


Figure 2: New optical unit

This application works with the transmitted light mode. It is easy to construct and requires no calibration. The mechanical costs are minimized.

From a drain plug it is possible to clean the glass tube. This is only required when depositions interferes the transmission factor. For this work is no service man necessary. The nurse can clean the glass tube with the enclosed scraper. The drain plug must be easily accessible.

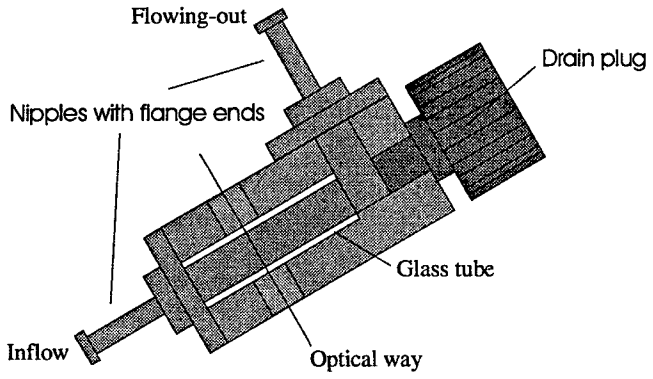


Figure 3: Mounting instructions

The connection to the kidney machine is very easy. At the closings are two nipples with flange ends. Above the flange ends the flexible tubes with the dialyze fluid can be pined up. The mounting in the kidney-machine must be happen with an angle of 30°. This guarantees that no air bubbles confuses the optical path length. Figure 3 shows this mechanical construction.

For the solution of the other errors is the electrical path responsible.

2.3 Electrical design

The functional block diagram of the electrical unit is shown in Figure 4.

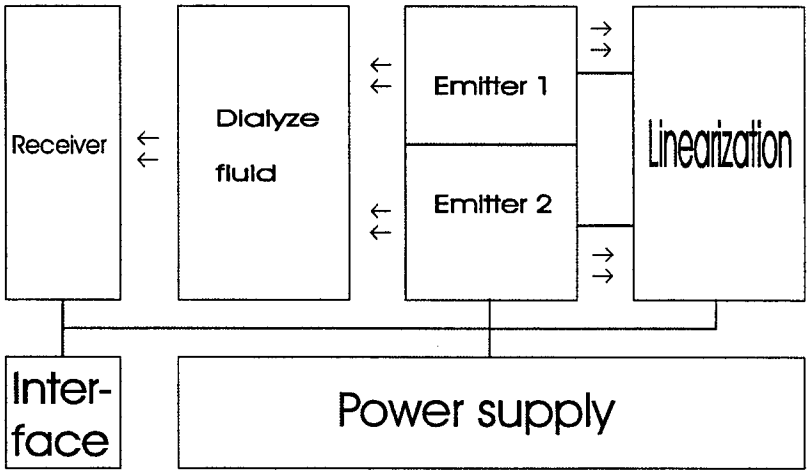


Figure 4: Functional block diagram

For best results it is important to use the same photo diodes in the receiver and transmitter unit. This reduces the dependence on temperature. For the best signal to noise ratio it is better to use a shielded photo diode case. Connecting the case to AGND guarantees the best results otherwise any kind of jamming, saturation and noise confuses will be amplified by the following componentry.

The next step is the conversion of the photo diode current to a voltage. The best possibility is to use a current to voltage converter designed with an operational amplifier. The connection diagram is shown in Figure 5.

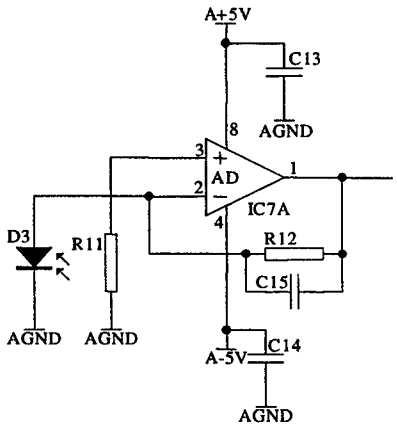


Figure 5: Preamplifier of the receiver

After the current-voltage-converter follows a rectifier. The characteristics of this circuit are the high speed and the high signal to noise ratio. To get the measuring signal it is necessary to complete the circuit with a low pass filter. Best results are reached by using high quality components with best dependence on temperature. The operational amplifier must have a low offset voltage. Figure 6 shows the used filter.

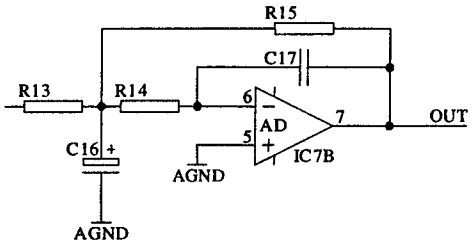


Figure 6: Low pass filter

For the data communication between the sensor and the kidney machine it is better to use digital signals. Therefore it is necessary to use an analog to digital converter with I^2C bus.

All this steps don't really reduces the temperature problems. The principal share takes the emitter unit as shown in **Diagram 4**. To compensate this problem it is necessary to regulate the optic performance.

This is possible with a second receiver. The used photo diode must be same as the other one. It is directly mounted by the LED. The result is a temperature compensated reference signal.

The cloudiness compensation of the dialyze fluid is now easy. With a two color LED it is possible to realize this. Here is the red light the reference signal and the green light the measure signal. The two LED's will be operated alternately. The head unit receives the measured value and adjusts the red light to a constant signal. The green signal is now equivalent to the blood concentration.

2.4 Design rules and results

The development and the operation of the new precision blood-leak-detector has shown that a lot of demands had to be arranged to the circuit. In the following the most important parameters are specified.

- Photodiodes:
  - identical temperature coefficient
  - stable mechanical fixing and absence of light to the printed circuit
  - filtered cathode voltage

- Preamplifier:
  - high speed and low noise Operational Amplifier
  - inspect with a square wave pulse generator that there is no overshoot; verify a frequency of 1MHz

- Rectifier:
  - high speed and low noise Operational Amplifier
  - temperature compensated components (low temperature coefficient)
  - high speed diodes

- Low pass filter:
  - low offset Operational Amplifier
  - temperature compensated components (low temperature coefficient)
  - low frequencies reduces output ripple

- AD-converter:
  - 12 bit resolution
  - I^2C bus to reduce cable amount
  - small package (SO8)

- LED:
  - Dual color emission
  - 20° radiation
  - opaque casing

2.5 Specification of the sensor system

Table 1 lists the specifications of the sensor system.

Model	Blood-leak-detector
Range of application	Recognition of fractures in the dialyzer of a kidney-machine
Measuring range	0 – 10 ‰
Accuracy	0,05 ‰
Resolution	0,003 ‰
Measuring rate	10Hz – 200Hz
Analog output	0 – 5V
Digital output	12 bit I^2C bus
Input level	TTL
Power supply	12 VDC ± 30 %
Current consumption	max. 30 mA
Fluid temperature	30°- 45°C
Ambient temperature	15°- 60°C
Storage temperature	-20°- 70°C

Table 1: Specifications of precision blood-leak-detector

3 CONCLUSION

- The developed sensor fulfills its requirements. It measures the blood concentration in a range from 0 – 10 ‰.
- The required speed is guaranteed. With a measuring rate of up to 200Hz and an accuracy of 0,05 ‰.
- High quality components and the regulated LED are responsible for the good thermal stability.
- The described sensor solves the given task and simplifies service- and maintenance-tasks.
- The sensor principle is also transmittable to other applications of the medicine technique.
- It is conceivable to process measurements in the catheter with the same principle to be always informed over the patients condition.

4 REALIZATION

The prototype of the new precision blood-leak-detector with high long-time-stability has already been integrated into the department of development and works fine. It is in operation service since october 1998. Two other devices were instaled in our department since november 1998. Figure 7 shows the complete kidney machine.

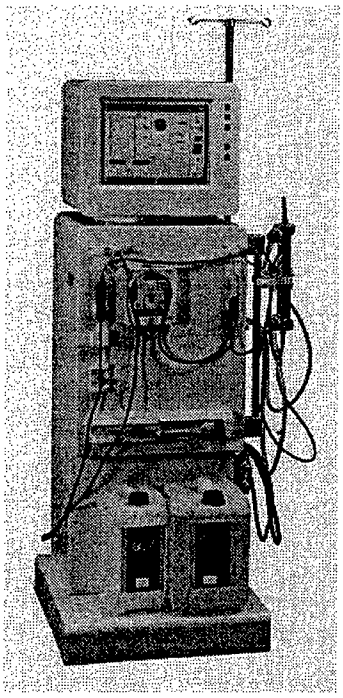


Figure 7: Kidney machine

5 ACKNOWLEDGMENTS

The authors are grateful to the center for sensor systems (Zentrum für Sensorsysteme, ZESS) at the university of Siegen for the financial and technical support.

6 REFERENCES

1. Gerthsen, Kneser, Vogel, "Physik" Springer-Verlag 15. Aufl. 1986

2. Medizintechnik

Autofluorescence Spectrum of Human Lung Tissue by Monte Carlo Modelling

Wei Zheng*[1], S.M. Krishnan[1], Zhiwei Huang[2], Teck-Chee Chia[2], Gilbert Chiang[3], Sing Lee[2]

[1]Biomedical Engineering Research Centre, School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore 639798
[2]School of Science, Nanyang Technological University, Singapore 259756
[3]Department of Pathology, Singapore General Hospital, Singapore 169608

ABSTRACT

Based on tissue microscopic fluorescence properties and optical parameters, a three-layer bronchus optical model was developed to calculate the excitation light (442 nm) distribution in the tissue and the fluorescence escape function from the tissue by Monte Carlo simulations. The contributions of each layer in the tissue to total autofluorescence signal observed at the tissue surface were also estimated. In addition, the fluorescence spectrum of the bronchial tissue was also modelled using the obtained tissue micro-spectroscopic properties. It is found that the theoretically modelled spectrum agrees well with the *in vitro* tissue fluorescence in the spectral range below 520 nm or above 600 nm.

Keywords: laser-induced autofluorescence, human lung tissue, Monte Carlo modelling

1. INTRODUCTION

Laser-induced autofluorescence (LIAF) spectroscopy is a promising technique for non-invasive detection of tumour and other diseased tissues. Recently several research groups have reported that the early diseases (e.g., displasia and carcinoma *in situ*) of lung tissues can be distinguished from normal tissues using the LIAF technique alone[1-3]. A good understanding of tissue autofluorescence spectral characteristics is essential for the accurate differentiation between normal and abnormal tissues using LIAF technique. Human tissue is an optically turbid medium, the fluorescence spectrum emitted from the tissue may be distorted by the absorption and scattering properties of tissues[4], and the spectral lineshape of the macroscopic autofluorescence may become different from the intrinsic spectra of tissue fluorophores. To enhance the understanding of the autofluorescence processes, theoretical modelling of light propagation in tissues is necessary.

*Correspondence: Email: ewzheng@ntu.edu.sg; Phone: +65-7906289; Fax: +65-7933318

Part of the SPIE Conference on Advanced Photonic Sensors and Applications
Singapore • November–December 1999
SPIE Vol. 3897 • 0277-786X/99/$10.00

In this study, the fluorophore micro-distribution and intrinsic fluorescence spectra of different layers in human bronchial tissue were measured under a novel microspectrophotometer (MSP) system. Based on the obtained experimental data and the available tissue optical parameters, a three-layer bronchus optical model was developed to model excitation laser light (442 nm) distribution in the tissue, and the autofluorescence escape process from the tissue as well as the contributions of different tissue layers to the total autofluorescence signal by Monte Carlo simulations. The theoretically modelled fluorescence spectrum of bronchial tissue was also compared with the *in vitro* autofluorescence spectra observed at the tissue surface.

2. METHODS

2.1 The Optical Model of Bronchial Tissue

Based on the fluorophore distribution and intrinsic fluorescent characteristics in tissues[5], human bronchial tissue can be simplified as a three-layer optical model, i.e., epithelium, mucosa/ submucosa, and cartilage. Table 1 outlines this model by providing data for tissue thickness, refractive index, and optical parameters at 442 nm for each bronchial tissue layer. The thickness and refractive index of each tissue layer were determined based on our microscopic measurements from bronchial tissue sections[5, 6]. To model the fluorescence escape process, optical parameters for the bronchus model at other wavelengths are also required. We compiled the data of each tissue layer at 25 different wavelengths from 460 to 700 nm in 10-nm intervals using the data from the literature[4]. For example, Table 1 also lists the optical parameters for fluorescence wavelength at 510 nm.

Table 1 Tissue optical parameters (μ_a, μ_s, g, n, d) at 442 and 510 nm of the three-layer bronchial optical model.

Optical Parameters	Epithelium	Mucosa/Submucosa	Cartilage
absorption coefficient μ_a (cm^{-1})	2.9[(a)] 1.9	10 3.8	6.7 1.8
scattering coefficient μ_s (cm^{-1})	343 295	253 231	290 265
anisotropy factor g	0.94 0.94	0.91 0.93	0.91 0.93
refractive index n	1.37	1.38	1.39
tissue thickness d (μm)	60	300/450	1100

(a) It is to be noted that for μ_a, μ_s and g, the first and the second sets of data correspond to the optical parameters at 442nm and 510 nm, respectively.

2.2 Total Observed Fluorescence $F(\lambda_{ex}, \lambda_{em}, z)$

Assuming that the tissue geometry is described by cylindrical coordinates: (r,θ,z), the observed flux of escaping fluorescence $F(\lambda_{ex}, \lambda_{em}, z)$, in units of W/cm^2, at the tissue surface can be calculated using the following formula[7]:

$$F(\lambda_{ex},\lambda_{em},r) = \int_0^D \int_0^{2\pi} \int_0^{\infty} \Phi(\lambda_{ex},r',z',\theta')\beta(\lambda_{ex},\lambda_{em},z')E\left(\lambda_{em},\sqrt{r^2+r'^2-2rr'\cos\theta'}\right) r'dr'd\theta'dz' \quad (1)$$

where Φ is the fluence rate distribution of the excitation light in the tissue, β is the intrinsic fluorescence coefficient, E is the fluorescence escape function which is the surface distribution as a function of radial position (r) of escaping photons from a point source of fluorescence at depth z and radial position r = 0 within a tissue of thickness d. (r,θ,z) corresponds to a position in the cylindrical coordinate system. λ_{ex} and λ_{em} represent the excitation and emission wavelengths, respectively. Note that the convolution in Eq(1) can be implemented numerically using discrete values for Φ and E (that were generated by the Monte Carlo simulation), and the experimentally determined β.

In this study, the *in vitro* autofluorescence spectra were measured using a wide illumination beam (10 mm) and a small collection spot (4 mm) at the centre of the illumination field[8]. Thus, the excitation light distribution could be simplified as a function of z only, i.e., Φ(z), and the fluorescence intensity will be the same in the tissue surface, independent of radial position, r. The Monte Carlo simulation program (MCML simulation code) [9] was used to calculate the excitation light distribution inside the bronchial tissue which was modelled as a three-layer structure. To calculate the fluorescence escape function, the source code was modified to simulate the light propagation process for an isotropic fluorescence point source buried at depth z inside the tissue. One million photons were launched in each simulation.

3. RESULTS AND DISCUSSION

3.1 Fluorophore Distribution inside Bronchial Tissue

Our micro spectroscopic fluorescence analysis[5] on bronchial tissue sections reveals that the natural fluorophore distribution inside the tissue is not uniform, but has a layer structure. The intrinsic fluorescence spectra of different bronchial tissue layers under the 442 nm laser excitation are shown in Fig.1. It is found that the mucosal/submucosal tissues have strong autofluorescence emissions, while the other tissue layers, particularly the cartilage tissue, fluoresce weakly. In addition, the intrinsic fluorescence spectra appear red-shifted, passing from the mucosa at 508 nm to the cartilage at 514 nm, suggesting that the fluorophores located in these layers may be of different types.

The relative fluorophore density function, ρ(z), in units of cm^{-1}, of each tissue layer can be determined based on the autofluorescence images captured by a digital CCD imaging camera. The relationship between the fluorophore density function ρ(z) and the tissue depth is shown in Fig.2. It is to be noted that the fluorophore density function of different layers was normalised to that of the cartilage layer. Based on the results, it is found that the mucosa fluoresces strongly, about eight times brighter than the cartilage tissue, the submucosa fluoresces about five times brighter than the cartilage, while the epithelium tissue is about two 2 times brighter than the cartilage. This also confirms that the mucosa is the most fluorescent tissue in bronchus.

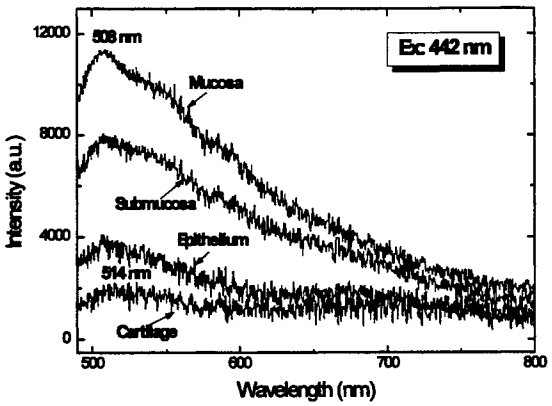


Fig. 1 Intrinsic fluorescence spectra of different tissue layers in bronchus under excitation laser light at 442 nm.

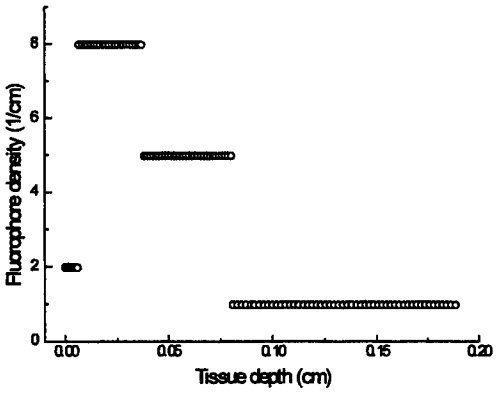


Fig. 2 Relative fluorophore density, ρ, as a function of tissue depth, z, inside bronchial tissue.

3.2 Intrinsic Fluorescence Coefficient $\beta(\lambda_{ex}, \lambda_{em}, z)$

The intrinsic fluorescence coefficient, $\beta(\lambda_{ex}, \lambda_{em}, z)$, denoted as the product of the absorption coefficient of the fluorophore, $\mu_{a.fluor}$ (cm^{-1}), and the quantum yield Y for fluorescence emission, can be obtained using[10]:

$$\beta(\lambda_{ex}, \lambda_{em}, z)=\rho(z) \times I_{intrin}(\lambda_{ex}, \lambda_{em}, z) \quad (2)$$

Within a tissue layer, β can be considered as a constant and therefore can be denoted as a function of z. The intrinsic spectra of different tissue layers measured with the microspectrophotometric (MSP) system were normalised to an equivalent overall integral, denoted as $I_{intrin}(\lambda_{ex}, \lambda_{em}, z)$, which is dimensionless. An example of the intrinsic fluorescence coefficient β (λ_{ex}=442 nm, λ_{em},=510 nm, z) inside the bronchial tissue, is shown in Fig. 3.

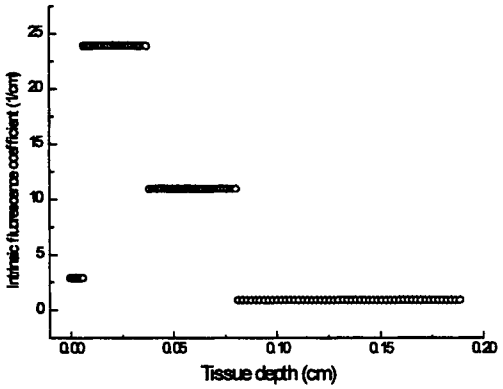


Fig.3 Intrinsic fluorescence coefficient, β(z), inside the model bronchial tissue for 510 nm fluorescence, under laser light excitation at 442 nm.

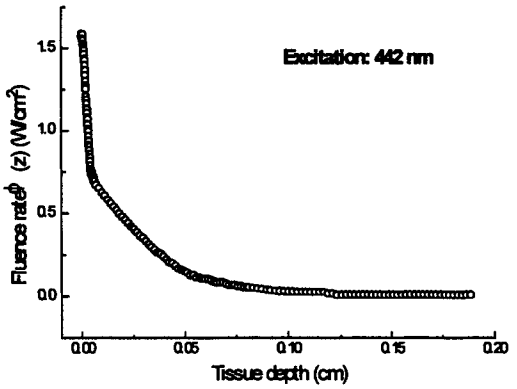


Fig.4 Laser light (442 nm) distribution Φ(z) in the bronchial tissue by Monte Carlo simulation. The normal incident beam is assumed as infinitely wide in diameter with power density of 1 W/cm^2.

3.3 Excitation Laser Light (442 nm) Distribution in Bronchial Tissue

Fig. 4 shows the Monte Carlo simulation results of the 442 nm excitation light distribution $\Phi(z)$ inside the tissue. It can be seen that that the epithelium and mucosa are the tissue layers mainly involved in the irradiation procedure and that, on considering an average mucosa thickness of 300 μm, only 15-30 % of the incident light fluence can reach the underlying structures, i.e., the submucosal layer. Very little amount of 442 nm light penetrates into the cartilage layer (tissue depth >800 μm), thus the cartilage layer has a negligible contribution to total measured autofluorescence signal under the 442 nm light excitation.

3.4 Fractional Contributions of Different Tissue Layers to Total Autofluorescence Signal

By the Monte Carlo method, simulations were performed in the one hundred fluorescence point sources at different depths inside the bronchus tissue (z ranged from 0 to 1000 μm), each emitting 1,000,000 photons isotropically. The point fluorescence sources were distributed along the z-axis. The fluorescence escape function E (λ=510 nm, z) versus the depth of the fluorescence source inside the bronchus tissue is shown Fig.5. It is found that the fluorescence escape efficiency decreases greatly while tissue depth increases.

The $\Phi(z)\cdot E(z)\cdot\beta(z)$ function versus tissue depth, z, is shown in Fig.6. Integration of the curve in Fig. 6 yields fractional contributions of different tissue layers in bronchus to the total fluorescence signal. It is found that 94% of the 510 nm fluorescence originates from the mucosal (85%) and submucosal (9%) tissues, about 6% comes from the epithelium tissue. The cartilage tissue contributes nearly zero to the observed fluorescence signal due to the fact that very little excitation light penetrates into this deep area and the fluorescence escape efficiency is also rather low from this tissue. Further experiments are currently underway to verify these theoretical predictions.

From the above modelling results, it can be seen that fluorophores located in the shallower layer are easier to detect than fluorophores in deeper tissue layers, mainly due to the excitation light distribution inside the tissue and the fluorescence escape efficiency. The back-scattering causes the excitation photons to pile up in the near surface region of the tissue, resulting in even higher light intensity in this region than the incident intensity (Fig. 4). The fluorescence escape efficiency near the tissue surface is also high compared to the deep tissue area. Thus the fluorescence light emitted near the surface is able to escape and reach the detector much more easily than fluorescence light emitted deep inside the tissue.

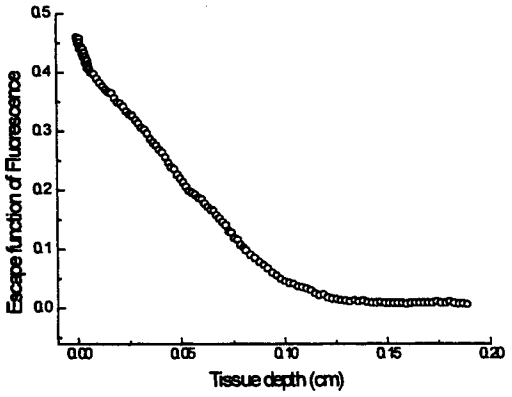


Fig.5 Fluorescence escape function E(z) versus the depth of the point source inside the tissue (λ_{em}=510 nm).

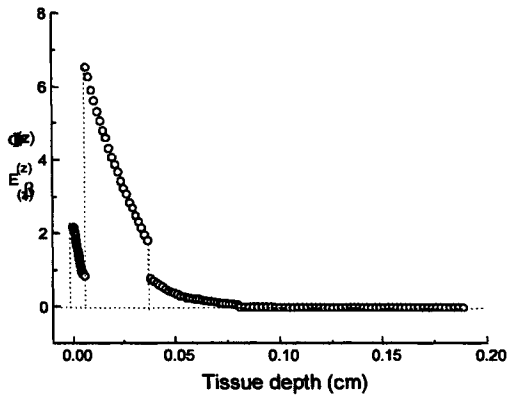


Fig.6 $\Phi(z)\cdot E(z)\cdot\beta(z)$ function versus tissue depth z.

3.5 Comparison of the Modelled Autofluorescence Spectrum with the Experimental Results

With a specific excitation/collection geometry[8], it is found that although the macroscopic *in vitro* autofluorescence intensity varies from different bronchial tissue specimens, the spectral maximum position of autofluorescence nearly remains constant at about 515 nm, and in most cases, the spectral lineshape of autofluorescence does not vary significantly. Fig.7 gives a typical laser-induced autofluorescence spectrum of bronchial tissue when excited by the 442 nm laser light. It can be seen that the macroscopic autofluorescence observed at the tissue surface is quite different from the intrinsic fluorescence spectra of bronchial tissues (Fig. 1).

For the purpose of comparison, the modelled tissue fluorescence spectrum by Monte Carlo simulation based on tissue optical parameters and microscopic fluorescence properties obtained in this study is also shown in Fig.7. Each curve is normalised to a maximum intensity of 1 at wavelength of 510 nm. It is found that the modelled spectrum shows a prominent green fluorescence peak at 510 nm, which is close to the experimental peak position of 515 nm. The 5-nm wavelength shift between the experimental and the modelled peak positions could have arisen from the errors in simulation, just using 10-nm interval. From Fig.7, it is observed that the theoretical curve lying in the regions of 460-520 nm and 600-700 nm fits well to the experimental data since light absorption by blood in these two regions is relatively small [11]. However, autofluorescence differences between the experimental and the calculated spectra exist over the strong haemoglobin absorption wavelength band of 530-590 nm[11, 12]. The experimentally observed spectrum shows double broad hollows in the 530-590 nm range. The divergence indicates that the data on microscopic properties of tissue for the 530-590 nm fluorescence in the theoretical modelling may not truly represent the actual situation in the bulk bronchial tissue. This suggests that the effect of different volumes of blood in tissue samples may have to be carefully considered in the theoretical modelling. Further refinements of the modelling to obtain better correlation with an experimental data are in progress.

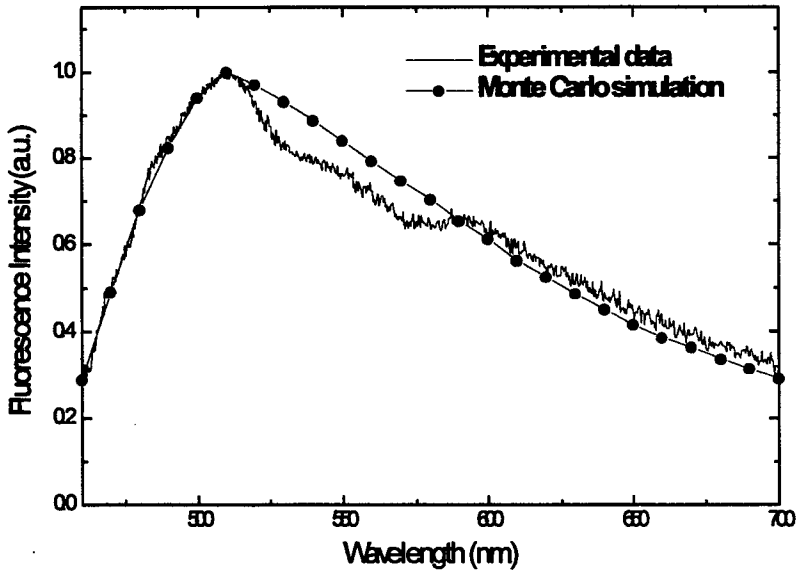


Fig. 7 Comparison of the modelled fluorescence spectrum by Monte Carlo simulation with the experimental *in vitro* autofluorescence spectrum.

4. CONCLUSIONS

In terms of the microscopic data on the histological structure of bronchial tissue, fluorophore micro-distribution, intrinsic fluorescence spectra and optical parameters, we can the theoretically predict the contributions of different histological layers to the observed fluorescence signal at the surface of tissue. The modelled fluorescence spectrum agrees well with the experimental results, indicating that the microscopic properties of tissue applied in the bronchus optical model can be well correlated with the macroscopic autofluorescence measurements. Therefore, the use of the theoretical optical bronchus model by Monte Carlo simulations, considering the microscopic properties of bronchial tissues and the intrinsic fluorescence distribution in tissue, will allow the optimisation of the experimental conditions that improve the modification of autofluorescence characteristics, thus approaching an *in vivo* autofluorescence measurements at bronchoscopy.

REFERENCES

1. G C Tang, A Pradhan, and R R Alfano, Lasers in Surg. & Med. 9: 290-295 (1989).
2. S Lam, J Hung, B Palcic. Lasers in Life Sciences 4:76-73 (1991).
3. S Lam, Timothy Kennedy, Michael Unger, et al., Chest 113 (3): 696-702 (1998).
4. J.Qu, C. MacAulay, S. Lam, and B. Palcic, Appl.Opt. 33 (31):7379-7405 (1994).
5. Wei Zheng, S.M. Krishnan, Zhiwei Huang, Teck-Chee Chia, Shusen Xie, Sing Lee, SPIE 3863 (1999), In Press.
6. Z.W.Huang, T.C.Chia, CH Diong, S.Lee, W.Zheng, S.S.Xie, SPIE, Vol.3195, 64-69 (1997).
7. M.Keijer, R.Richards-Kortum, S.L.Jacques, M.S.Feld, Appl. Opt., 28, 4286-4292 (1989).
8. Z.W Huang, PhD thesis, Nanyang Technological University, Singapore, December, 1998.
9. L.H.Wang, S.L.Jacques, Monte Carlo Modeling of Light Transport in Multi-layered Tissues in Standard C , University of Texas, M.D.Anderson Cancer Center, 1992.
10. H Zeng, C MacAulay, B Palcic, and D Mclean, SPIE 2135:94-103 (1994)
11. I.J.Bigio, J Boyer, T.M.Johnson, J.L.R.Mourant, R.Conn, et al., SPIE, Vol.2324, 46-54 (1994)
12. K.T.Schomacker, J.K.Frisoli, C.C.Compton, et al., Gastroenterology, 102, 1155-1160 (1992).

Study on Blood Cell Counting and Classification by Non-flowing Laser Light Scattering Mehtod

Ye Yang[a], Zhenxi Zhang[b], Xinhui Yang[b], Dazong Jiang[b] and Joon Hock Yeo[a]

[a]School of MPE, Nanyang Technological University, Nanyang Avenue, Singapore 639798
[b]Institute of Biomedical Engineering, Xi'an Jiaotong University, Xi'an, P.R.China 710049

ABSTRACT

A new non-flowing laser light scattering method for counting and classifying blood cells is presented. A linear charge-coupled device (CCD) with 1024 elements is used to detect the scattered light intensity distribution of the blood cells. A pinhole plate is combined with the CCD to complete the focusing of the measurement system. An isotropic sphere is used to simulate the blood cell. Mie theory is used to describe the scattering of blood cells. In order to inverse the size distribution of blood cells from their scattered light intensity distribution, Powell method combined with precision punishment method is used as a dependent model method (by assuming the blood cells' distribution fits to a special distribution in advance) for measuring red blood cells and blood plates. Non-negative constraint Least Square (NNLS) method combined with Powell method and precision punishment method is used as an independent model (by solving the blood cells' distribution without assuming it fits to a special distribution) for measuring white blood cells (red blood cells and blood platelets). The size distributions of white blood cells and red blood cells, and the mean diameter of red blood cells are measured by this method. White blood cells can be divided into three classes: lymphocytes, middle-sized cells and neutrocytes according to their sizes. And the number of blood cells in unit volume can also be measured by the linear dependence of blood cells (white and red blood cells) concentration on scattered light intensity

Keywords: Blood cells, counting and classification, non-flowing, laser light scattering, CCD, inversion algorithm

1. INTRODUCTION

Blood consists of plasma and blood cells. Blood cells comprise white blood cells, red blood cells and blood platelets. These three types of blood cells have different sizes, shapes, functions and numbers in blood. According to their shape difference, white blood cells can be further divided into two types: granule cells and nongranule cells. Granule cells can be further divided into neutrocytes, eosinocytes and basocytes. Nongranule cells can be divided into lymphocytes and monocytes. Blood cells counting and classification is a routine medical check item. It is very important both in medical diagnosis and therapy.

There are four methods for blood cells' counting and classification. One is the counting platelet method, which is a manual counting method. It has the advantages of low cost and ease of use, but has the disadvantages of high working intensity, ease of eyestrain, and high experience requirement for the operator. Another method is the hotoelectric nephelometric metood, which can only be used to measure the normal red blood cells.[1] Coulter Counter[1] and Flow Cytomether[2] are the two automatic blood cell counting and classification instruments being generally used. They use the Coulter method[1] and laser light scattering method[2] respectively. And Both of them based on the flowing principle, i.e., by measuring the single cell flowing through the measuring region. This two automatic counting methods have the advantage of accurate measurement, but have the common disadvantages of complicated structure, high price and ease of stop up because of the flowing principle. In order to overcome these disadvantages, a non-flowing laser light scattering method for counting and classifying blood cells is presented.

2. PRINCIPLE OF NON-FLOWING LASER LIGHT SCATTERING METHOD FOR COUNTING AND CLASSFLYING BLOOD CELLS

Non-flowing laser light scattering method for counting and classification of blood cells is based on the Mie scattering theory of spherical particles. Blood cell is simulated by an isotropic sphere.[3] The measurement set-up is shown in Fig.1.

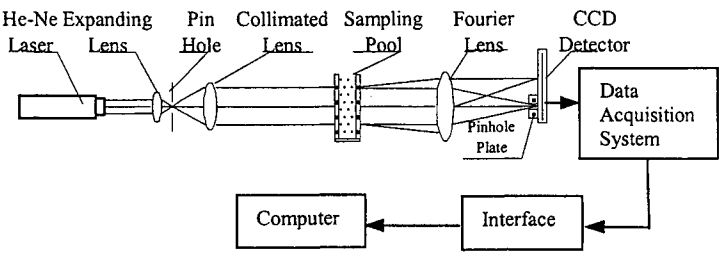


Fig.1 The principal of the non-flowing laser light scattering method for blood cell counting and classification

The narrow laser light emitted from the He-Ne laser is expanded by the expanding lens. When this expanded light going through the pinhole and collimated lens, it will become a uniform and wide parallel light. When this light incidents on the blood cells in the sampling pool, it will be scattered. The scattered light is received by the Fourier lens and forms a symmetrical pattern in the back focal plane of the Fourier lens. The linear charge-coupled device (CCD) detector, which is positioned in the back focal plane of the Fourier lens, converts the scattered light signals into electric signals. These analogue electric signals are collected and converted into digital signals by the data acquisition system. Then these digital signals are transferred into the computer through the interface circuit. After computer's analyzing and processing of these data, the size distribution of the white blood cells and red blood cells, the mean diameter of the red blood cells, and the blood cells number in the unit volume are given.

Let the intensity and wavelength in the medium of natural parallel light be respectively I_0 and λ ($\lambda = \lambda_0 / m_m$, λ_0 is the light wavelength in the vacuum, m_m is the refractivity of the medium). According to Mie theory,[4] in the condition of single scattering and elastic scattering, when this light incidents on the isotropic spherical particle of diameter D, at the scattered angle θ in the CCD plane which is positioned on the back focal plane of Fourier lens, whose focal length is f, the scattered light intensity $I(\theta)$ is

$$I(\theta) = \frac{\lambda^2}{8\pi^2 f^2} I_0 [i_1(m,\alpha,\theta) + i_2(m,\alpha,\theta)] \tag{1}$$

$$i_1 = S_1(m,\theta,\alpha) \bullet S_1^*(m,\theta,\alpha) \tag{2}$$

$$i_2 = S_2(m,\theta,\alpha) \bullet S_2^*(m,\theta,\alpha) \tag{3}$$

where i_1, i_2 are the intensity function; S_1, S_2 are the amplitude functions; S_1^*, S_2^* are the conjugate complex number of S_1, S_2; α is the particle size parameter ($\alpha = \pi D/\lambda$); $m = m_1 - im_2$ is the particle refractivity in the medium, when the particle has absorption, the image part m_2 of m is not zero.

For the cell group whose number is N, if each cell has the same diameter, then in the condition of single scattering and elastic scattering, the total scattered light intensity is N times of the single cell's. For multiple sized cell dispersion, the total scattered light intensity is equal to the sum of the light intensity of single cell,[5] i.e.

$$I(\theta) = \int_0^\infty (\frac{\lambda}{2\pi f})^2 I_0 \frac{i_1 + i_2}{2} W(D) dD \tag{4}$$

where $W(D)$ represents the number percentage of the cells in the range of some special size among the total cells.

In fact, the size of every cell group to be measured is between some range [D_{min}, D_{max}]. And the measuring of scattered light is also in some scattered angle range [θ_{min}, θ_{max}], it is determined by the radial measuring range of the CCD [r_{min}, r_{max}]. Where , in the case of small forward scattering angle range, $\theta_{min} = r_{min}/f$, $\theta_{max} = r_{max}/f$. Then integral Eq. (4) can be changed as a summation form

$$I(\theta_i) = I_0 \int_0^\infty \frac{\lambda^2}{8\pi^2 f^2}(i_1 + i_2)W(D)dD$$
$$= I_0 \sum_{j=1}^{M} \frac{\lambda^2}{8\pi^2 f^2}(i_1 + i_2)W_j \qquad (5)$$
$$i = 1,2,\cdots,L$$

where W_j is the number percentage of the cells whose diameter is D_j in the whole cells. M is the number of the cell size ranges. $L = (r_{\max} - r_{\min})/\Delta r$ is the number of the sampling units of the CCD, and Δr is the distance between each CCD sampling unit.

In Eq. (5), $I_0^{'}$ is a constant; let it be 1. Eq. (5) can be written as simple matrix form, i.e.

$$I = TW \qquad (6)$$

where $I=(I_1,I_2,\cdots,I_L)^T$ is the light intensity distribution vector, $W = (W_1, W_2, \cdots, W_M)^T$ is the size distribution vector, and

$$T = \begin{bmatrix} t_{1,1} & t_{1,2} & \cdots & t_{1,M} \\ t_{2,1} & t_{2,2} & \cdots & t_{2,M} \\ \vdots & & & \\ t_{L,1} & t_{L,2} & & t_{L,M} \end{bmatrix}$$

is the light intensity distribution coefficient matrix, the elements $T_{i,j}$ of T is

$$t_{i,j} = \frac{\lambda^2}{8\pi^2 f^2}(i_1(\theta_i, \alpha_j, m) + i_2(\theta_i, \alpha_j, m))$$
$$= \frac{\lambda^2}{8\pi^2 f^2}(i_1(r_i/f, \pi D_j/\lambda, m) + i_2(r_i/f, \alpha_j, m)) \qquad (7)$$

The physical meaning of Eq. (7) is, in the position of the CCD unit apart from the light axis r_i, the scattered light intensity caused by the cell of diameter D_j is $t_{i,j}$. Eq. (6) is the basic calculating equation of the non-flowing laser light scattering method.

So, when light distribution coefficient matrix T is known, CCD is used to detect the light intensity distribution vector I caused by the measured blood cells, then by solving the linear coupled equations $I = TW$ reasonably, the distribution of the blood cells group size distribution can be obtained. But as the approximation of the principle, the influence of the coherent noise and the system noise, the "ill-condition" of the equation is almost inevitable, so the usual methods used to solve the linear coupled equations can't be used to solve the distribution of blood cells. And the distribution function to be solved should be constrained, i.e., each element of the size distribution should be non-negative.

3. INVERSION ALGORITHMS FOR THE MEASUREMENT OF BLOOD CELLS SIZE DISTRIBUTION

3.1. Dependent model method

Dependent model method is the method which solve the blood cells' distribution by assuming it fits to a special distribution in advance. By solving the parameters of the distribution function, the distribution to be solved can be determined. According to reference [6] and [7], red blood cells size distribution and that of blood platelets are fit to normal distribution and logarithmic normal distribution respectively. So dependent model method can be used for red blood cells and blood platelets.

Take normal distribution for example, its distribution function is:

$$\varphi(D)=\frac{1}{\sqrt{2\pi}\sigma}\exp[-\frac{1}{2}(\frac{D-\overline{D}}{\sigma})^2] \tag{8}$$

where $\overline{D}$, σ are constants.

It is known from Eq. (8) that if $\overline{D}$, σ are determined, the distribution of the blood cells is determined. W_i is the number percentage of the cells with sizes between the diameter range [D_{iL}, D_{iH}] in the whole cells. The size distribution in the certain range[D_{iL}, D_{iH}] is:

$$W_i=\int_{D_{iL}}^{D_{iH}}\frac{1}{\sqrt{2\pi}\sigma}\exp[-\frac{1}{2}(\frac{D-\overline{D}}{\sigma})^2]dD \tag{9}$$
$$(i=1,\ 2,\cdots,L)$$

where L is the number of the sampling units of the CCD. D_{iL}, D_{iH} represent the upper and lower diameter limit of the i^{th} diameter range.

Substituting the value of W_i calculated from Eq.(9) to Eq. (6) $I=TW$, the calculated scattered light intensity in the sampling units of the CCD can be obtained. Adjusting the values of $\overline{D}$ and σ, a series of calculated scattered light intensity can be obtained. When the error between the calculated scattered light and the measured light intensity in the CCD sampling units reaches minimum, the values of $\overline{D}$ and σ can be determined. I.e., to determine the values of $\overline{D}$ and σ which make the objective function Eq.(10) minimum

$$f=\sum_{j=1}^{L}(I_{cj}-I_{mj})^2 \tag{10}$$

where I_{cj} represents the calculated light intensity value in the j^{th} detect unit,

I_{mj} represents the measured light intensity value detected by the j^{th} detect unit,

L is the number of the CCD sampling units,

suffix c and m represent the calculated and measured values of the light intensity respectively.

The adjustment of $\overline{D}$ and σ is an optimal search process. Powell method[8] and precision punishment method[8] are used in this investigation. Because Powell method[8] is accepted as the most effective method among the direct optimization methods, but it can only solve the non-constraint optimization problem. So the precision punishment method[8] is combined with Powell method to solve the constraint optimization problem. Table 1. shows some numerical simulation results. It is shown that this method can effectively solve the particle distribution that fit to special distribution, no matter what the initial values set.

Dependent model method has the advantages of high calculation speed and less influenced by the measurement error of the light intensity, but it cannot be used to the particle suspensions with random particle size distribution. However, independent model method can solve the distribution of blood cells without assuming it fits to special distribution.

Table.1. Numerical Simulation Results with different Initial Values

Set Value	$\overline{D}$	4.5				9.5			
	σ	6.0				3.0			
Initial Value	$\overline{D}^0$	1.8	3.6	15.0	25.0	1.8	3.6	15.0	25.0
	σ^0	3.0	8.1	12.0	20.0	3.0	8.1	12.0	20.0
Calculated Value	$\overline{D}^*$	4.50	4.50	4.50	4.50	9.50	9.50	9.50	9.50
	σ^*	6.00	6.00	6.00	6.00	3.00	3.00	3.00	3.00

3.2. Independent model method

Since independent model method can solve the distribution of blood cells without assuming it fits to a special distribution, it can be applied to the blood cells suspensions with random distribution.

In this paper, Non-negative constraint Least Square (NNLS) method[9] combined with Powell method and precision punishment method is used as an independent model. The objective function is changed as:

$$P(x,v)=\sum_{i=1}^{L}(I_{ci}-I_{mi})^2+v\sum_{j=1}^{M}\left|\min(-W_j,0)\right| \tag{11}$$

where

L is the number of the CCD sampling units,

I_{ci} represents the calculated light intensity value in the i^{th} detect unit,

I_{mi} represents the measured light intensity value detected by the i^{th} detect unit,

M is the number of the cell size ranges,

v is the punishment factor,

W_j is the j^{th} component of the size distribution vector.

suffix c and m represent the calculated and measured values of the light intensities respectively.

When this objective function reaches minimum, the values of W_j can be determined.

Numerical simulations and experimental results show that this method can calculate the particle suspensions not only with single-peak, but also with double-peak and multi-peak distribution accurately.[10] Table.2 shows the numerical simulation result of a double-peak distribution. Low calculation speed and high influenced by the measurement error of the light intensity are the two disadvantages of the independent model method.

White blood cells' size distribution is random, so the independent model method can be applied to them. And the independent model method can also be used to measure the distributions of red blood cells and blood platelets.

Table.2. Numerical simulation result of a particle suspension with double-peak distribution

Number	Size Group	Set Value	Calculating Value
1	0.6~0.8	0	0
2	0.8~1.1	0	0
3	1.1~1.5	0	0
4	1.5~2.1	0.10	0.099979
5	2.1~2.9	0.15	0.250003
6	2.9~3.9	0.15	0.149991
7	3.9~5.3	0	0
8	5.3~7.2	0	0
9	7.2~9.8	0	0
10	9.8~13.3	0	0
11	13.3~18.1	0	0
12	18.1~24.6	0.05	0.050001
13	24.6~33.4	0.30	0.299982
14	33.4~45.4	0.15	0.150044
15	45.4~61.7	0	0

4. EXPERIMENTAL METHODS AND MEAUREMENTS

Some experiments of standard particles (polystyrene latex provided by Beijing Chemical and Metallurgy Research Institute) and human blood suspensions are measured in Xi'an Jiaotong University using the experimental measuring set-up shown in Fig. 1.[10]

4.1. Experiment methods

A 5-mW helium-neon (He-Ne) laser of wavelength 632.8 *nm* is used. As CCD's photosensitive array has a high space resolution, constant geometrical size and identical interval between each element, it is very suitable for measuring the intensity distribution of the radial scattered light of the blood cells. A linear CCD with 1024 elements is used in this investigation to detect the scattering light intensity distribution of blood cells.

In order to overcome the difficulty of distinguishing scattered light of blood cells from the original incident light at the light axis and the disadvantage of CCD's ease of saturation, a pinhole plate combined with the CCD is used as a double-detectors' structure to complete the focusing of the experimental set-up. A silicon photoelectric cell is positioned behind the pinhole in the pinhole plate, when the output of the silicon photoelectric cell reaches the largest value, the focusing of the experimental set-up is completed.

Blood samples are randomly fetched from donors in Xi'an Jiaotong University hospital clinical laboratory. The blood sample is hepari anticoagulated venous blood. The number of the male subjects is the same as that of the female subjects. Donor's age ranges between 22 and 75. The blood diluents used are common diluents available at the university hospital having the following composition, i.e.,

Red blood cell diluent: 0.9% normal saline.
White blood cell diluent: 3% acetum.

Using different quantities of blood cell diluents, the blood is diluted to different concentration of red blood cells suspensions and white blood cells suspensions.

4.2. Standard particles suspensions measurements

The size distributions of the standard particle suspensions are measured.

Table.3. shows some experimental measurement results of the standard particle (polystyrene latex) suspensions by dependent model method and Fig.1 shows a standard particle (polystyrene latex) suspension's experimental result (the Coulter method's result is provided by the standard particle vendor) by the independent model method. It is shown that the experimental results fit well with the standard particles' nominal values.

When measuring standard particles, the results also fit well with the Coulter methods' (all the standard particles' Coulter method results are provided by the standard particle provider). The standard deviation of the reproducibility of the set-up is no more than 5% and the system uncertainty is less than 10% (compared with the nominal size of the standard particle).[10]

Table.3. Experimental Results of the Standard Particle (9.88μm) Suspension

Nominal Value	$\overline{D}$	9.88				
	σ	2.80				
Measurement Value	$\overline{D}^*$	9.37	9.76	9.41	9.32	9.68
	σ^*	2.56	2.67	2.87	2.84	2.55

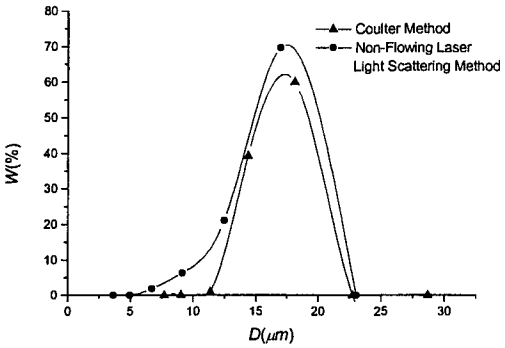


Fig. 2. Measurement results of a standard particle Suspension with particle's nominal diameter of 16.32μm.

4.3. Blood sample suspensions measurements

The size distributions of the red blood cells and white blood cells, and the mean diameter of the red blood cells' suspensions are measured; And white blood cells are classified as three classes.

The white blood cells are classified by measuring their size distribution. According to X. H. Lu,[11] the white blood cells between 4.67μm ~ 7.6μm sizes can be counted as the lymphocytes (actually, lymphocyte mostly between 5.6μm ~8.0μm). And the white blood cells with their sizes between 7.6μm ~10.1μm can be regarded as the middle-sized cells, which are mainly monocytes. But eosinophil, basophil, archecyte and heterocyte are also in this range (7.6μm ~10.7μm). So if the middle peak is dominant, it is important then, to take the smearing result as rule. And the white blood cells of sizes 10.1μm ~17.0μm are regarded as neutrocytes. So the white blood cells are classified as lymphocytes, neutrocytes and middle-sized cells according to their size distribution. By this method, the small and big cells accordance ratio are high, and the accordance ratio of middle sized cells is below 50%.

Here is one subject's measuring result. Gender: Female; Age: 74. The counting result of the counting plate method is: red cells number: 114 ten thousand/mm^4, lymphocytes: 27%, neutrocytes 72%, monocytes: 1.0%.

The red blood cells and white blood cells size distributions of this subject by the presenting method are shown in Fig.3 and Fig.4 (by independent model method). The red cells' mean diameter of this tester is 8.24μm. It is greater than the normal value and implied that this patient may has anaemia. The result of the white blood cells classification is: lymphocytes: 25.4%, neutrocytes: 61.7%, and middle-sized cells: 12.8%. The measuring result fits well with the counting plate method. Further works of comparing the presented method's result with other automatic methods' will do in the future.

The number of blood cells in unit volume can also be measured by the linear dependence of blood cells (white and red blood cells) concentration on scattered light intensity.[10]

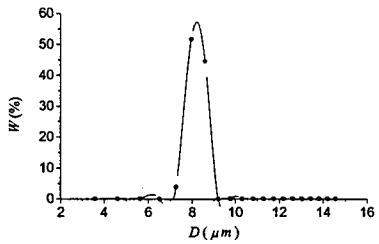


Fig.3. Red blood cells measuring result

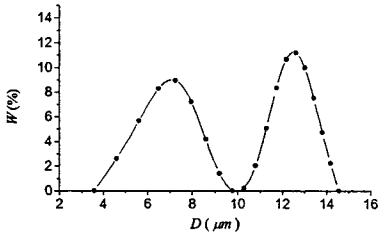


Fig.4. White blood cells measuring result

5. CONCLUSIONS

A new non-flowing laser light scattering method is presented for counting and classifying blood cells automatically. A linear charge-coupled device (CCD) is used to detect the scattered light intensity distribution of blood cells, and a pinhole plate is combined with the CCD as a double-detectors' structure to complete the focusing of the experimental set-up.

An isotropic sphere is used to simulate the blood cell. Mie theory is used to describe the scattering of blood cells. In order to inverse the size distribution of blood cells from their scattered light intensity distribution, Powell method combined with precision punishment method is used as a dependent model method (by assuming the blood cells' distribution fits to a special distribution in advance) for measuring red blood cells and blood plates. Non-negative constraint Least Square (NNLS) method combined with Powell method and precision punishment method is used as an independent model method (by solving the blood cells distribution without assuming it fits to a special distribution) for measuring white blood cells (red blood cells and blood platelets).

The size distributions of white blood cells and red blood cells, and the mean diameter of red blood cells can be measured by this method. White blood cells can be divided into three classes: lymphocytes, middle-sized cells and neutrocytes according to their sizes. The number of blood cells in unit volume can also be measured by the linear dependence of blood cells (white and red blood cells) concentration on scattered light intensity

REFERENCES

1. Li Qi Ying, Xu Mian Zhong, Kong Xiang Yun, *Practical clinical medical inspection*, Hu Bei People's Publishing House, Wu Han, 1980. (in chinese)
2. L. S. Cram, J. C. Martin, J. A. Steinkamp *et al*, "New flow cytometric capabilities at the national flow cytometry resource," *Proceedings of the IEEE* **80**(6), pp. 912-917, 1992.
3. Ye Yang, Zhenxi Zhang, Dazong Jiang, " The elastic light scattering of blood cells," *Sheng Wu Yi Xue Gong Cheng Xue Za Zhi* **14**(2), PP. 170-175, 1997. (In Chinese)
4. M. Kerker, *The Scattering of Light and Other Electromagnetic Radiation*, Academic press, New York ,1969.
5. H. C. Van de Hulst, *Light scattering by small particle*, Wiley Press, New York, 1981.
6. J.M.Steinke, A. P. Shephered, " Comparison of Mie theory and the light scattering of red blood cells," *Applied Optics* **27**(19), pp. 4027-4033, 1988.
7. Li Ling Zhi. "The scattering properties of blood platelet," *Photoelectron and Laser* **6**, pp.386-389, 1995.
8. R. Fletcher, *Practical methods of optimization*, John Wiley &Sons, New York ,1987.
9. L. Huang, *The linear algebra in system and control theory*, Science Publish House, Bei Jing,1984. (In Chinese)
10. Y. Yang, *The study on blood cell counting and classification by stationary suspending laser light scattering method*, Ph.D. Dissertation, Xi'an Jiaotong University, Xi'an, 1997. (In Chinese)
11. X.H.Lu, "Automatic blood cell analytical instrument," *Modern Medical Instrument and Its Application* **6**(4), pp.12-16, 1994. (In Chinese)

FIBER OPTIC SENSOR FOR METHANE HAZARDS[+]

V. Kumar[*] and D. Chandra

Department of Electronics and Instrumentation
Indian School of Mines, Dhanbad 826004, India

ABSTRACT

Different types of fiber optic methane sensor, especially for mining application, have been reviewed in this paper. Optical absorption and differential optical absorption techniques for the remote detection of methane gas using low-loss silica fiber have been discussed. Infrared fiber optic, sol-gel and correlation spectroscopy methods have been described in brief. Another noble technique based on attenuation of evanescent field has been enunciated using D-fiber. Merits and demerits of each technique and its suitability to mining industry have been highlighted. Optical fiber, being a dielectric, non-metallic and non-sparking, is an intrinsically safe media and is ideally suited to the hazardous environment present in mines.

Keywords: Fiber optic methane sensor, Mining hazards, Infrared sensor, Atmospheric study

1. INTRODUCTION

Methane-air-mixture is highly explosive between 5-14.5% and continuous monitoring of the methane gas in underground coal mines is very important for the safe and fearless operation of the mines. At present methane is routinely monitored between 0-5% using catalytic combustion type methane sensor, which is an electrically heated pellistor. In this, palladium is deposited on platinum wire which acts as catalyst for methane gas. Using hybrid thick-film technology, tin-oxide based methane sensor with improved sensitivity and reliability has been reported by several workers[1-7] but these are also based on heating effect of semiconductor oxides. Several other methods of detection of methane gas in underground mines have been described by different workers in the recent mine safety conference[8]. Catalytic combustion type methane devices are not intrinsically safe and poisoning of the catalyst may occur in certain environments such as grease, transformer oil, electrical insulator, Si, ester and phosphate. Cracking in pellistor occurs when methane percentage is more than 5% , which causes permanent damage of the bead. Lifetime of this sensor is also short, maximum upto 2 years. Frequent failure occurs in the system and it is very difficult to replace the sensor. It is understood that methane monitoring equipments worth crore of rupees are lying unused in different collieries in India mainly due to lack of sensors. Therefore, it has been thought of great interest to study fiber optic methane sensor (FOMS) for mining applications, which is intrinsically safe in nature. At present, the authors have a project from All India Council for Technical Education (AICTE), New Delhi to design and develop fiber optic methane sensor for mines.

Several workers[9-15] have tried to exploit the attractive features of fiber optics in mines. Recently, the authors[16-19] has explored the idea of various possible applications of fiber optic sensors in underground coal mine as well as in geophysics. In this paper, a brief description of various types of fiber optic methane sensor has been reviewed with special emphasis towards the suitability to underground coalmines. Different researchers[20-30] have used the absorbance in the optical power at wavelengths such as 1.33, 1.66, 2.25, 3.33, 3.39, and 7.7 μm for the detection of CH_4 gas using different methods. Recently Stewart et al.[31-33] and others[34,35] have reported fiber optic multi point sensor for the detection of methane gas. These methods are discussed in the following sections

[+]Some of the material presented here has been taken from our previous publication (Ref.16).

[*] Email - cme@ismine.ernet.in, Fax - 0091-326-203042/ 202380/ 206241, Phone - 0091-326-206247(Res)/ 202578 /202487/202577(Off)

2. OPTICAL ABSORPTION (OA) METHOD[36-45] OF DETECTION OF CH_4 GAS IN NIR REGION

The use of an optical fiber link to the optical absorption cell was first proposed by Inaba et al[36] using low-loss fibers. Fig 1 depicts schematically the block diagram of an optical system for remote detection of CH_4 gas based on OA method. The pulsed light from LED is guided along the optical fiber to the remote gas sensor cell. Collimated light is passed through the sensing volume, where it is selectively attenuated by any methane present. It is then refocused into the return fiber which guides it to detector through monochromator, Lock-in-amplifier and finally to recorder or microcomputer. Using an InGaAsP LED of 1.33 μm wavelength, multimode fibers of 50/125 μm core-clad diameter, transmission loss of nearly 1 dB/km, length 1 km + 1 km and Cell length 0.5 m, Inaba[37] detected upto 1000 ppm of CH_4 or 1/50 of lower explosive limit (LEL) were confirmed with resolution 0.3 nm. Recently, Nagai et al[38] have demonstrated long distance remote sensing of methane gas using a 1.65 μm LD and a single mode fiber.

In addition to CH_4, other polluting gases such as propane, ethylene, butane, ethane and ammonia have also been detected using this method. These have been described in different literatures by different workers[39-48].

3. DIFFERENTIAL OPTICAL ABSORPTION (DOA) METHOD OF DETECTION OF CH_4 GAS[37,44,46,49-51]

The block diagram of differential optical absorption method for the detection of CH4 gas is shown in Fig. 2. In this method, light absorption is based on two optical wavelength bands. The measurement is basically implemented by employing a laser source located at the control center, from where an optical fiber is connected to a multireflection white cell[52] with multi-pass geometry, placed at a remote point to be monitored. Zientkiewicz[13] has used a tight steel tube approximately 1 m long and 10 cm in diameter having gas inlet and outlet pipes as sample cell instead of white cell for detecting CH_4 in an underground mine. If extremely low-loss optical fiber is available in the frequency range concerned, a conventional nonlaser source such as LED and a multiple-wavelength discharge tube could also be utilised in place of LD which can be tuned to or coincide with two frequencies λ_X and λ_R in which at least one frequency is in absorption region of CH_4 gas i.e. 1.33 or 1.66 μm and other is at an adjacent reference wavelength say λ_R= 1.29 μm or 1.60 μm. Both the optical signals are passed through the sample cell and returned to a coupler.

The light from one part of the coupler is passed through a bandpass filter with the spectrum centered at 1.29 μm. The output of the filter serve as the reference signal and is used to monitor variation in power output of LED, or of the light launched into the filter. The light from the other port of the coupler passes through a bandpass filter with its spectrum centered at 1.33 μm which corresponds to the absorption peak of methane.

The output light level from the reference and the sensor ports was linearly converted into electrical signals with low noise photodiodes and a high impedance amplifier and subsequently was measured with a lock-in-amplifier. The presence of methane in the sensing volume results in a dip in the transmission. The magnitude of the dip is proportional to the concentration of methane.

Mohebati & King[50] have used diode lasers acting as both the light source and the wavelength selection device to eliminate the drift and noise of DOA fiber optic gas sensor. They have used two different diode lasers of widely different wavelengths 0.8 um and 1.33 μm and differential absorption has been measured using etalon to scan two modes of a diode laser. Recently, Jin et al[33,51] have shown the performance limitation of fiber optic methane sensor due to interfacing effect.

4. INFRARED FIBER OPTICS METHOD OF DETECTION OF CH_4 GAS[53-54]

The availability of zirconium fluoride (ZrF)[53] and chalcogenide IR transmitting fibers make possible the detection of CH_4 gas with absorption bands in NIR region. ZrF fibers can transmit 2-5 μm region with attenuation 10-40 dB/km. Matson and Griffin[54] have used ZrF IR fiber with core/clad diameter 250/300 μm, NA = 0.21, loss <20 dB/km, transmission range 0.3-4.5 μm, minimum bend radius 11 mm with protective coating of epoxy acrylate of 45 um thickness for detecting hydrocarbons. Based on IR transmitting fiber, Matson and Griffin[54] have

suggested three types of gas sensor cells: (i) a simple vapour absorption cell, (ii) a photoacoustic cell, and (iii) an expensive FTIR interface to a remote gas analysis cell. Only the first one is described here. Fig. 3 shows the block diagram of IR fiber optic vapour absorption sensor evaluation set up. A typical design is cylindrical tube with IR transmitting window at both ends. For the laboratory set up Matson and Griffin[54] have used a commercial FTIR spectrometer gas cell (10 cm long) as a probe. The ends of the cell were closed-off by CaF_2 windows (2 mm thick). Light from 3.39 μm He-Ne laser drived by the IR fiber was collimated into 1.0 inch diameter beam by a 2.5 inch focal length ZnSe lens. By adjusting the mirror, the light can be coupled into the second, or output, fiber in the tandem connector. At the end of the output fiber, a cooled InSb detector captures light and modulated at the source frequency. Signals from detector are fed into a lock-in-amplifier finally calibrated for the hexane and minimum detection limit was found to be 0.01%.

5. EVANESCENT WAVE SPECTROSCOPY METHOD OF DETECTION OF CH_4[53-55]

It has been observed that the amplitude of the field in cladding decay exponentially and is referred as an evanescent field. There are several types of IR absorption sensors based on evanescent field device. This device is commonly used for analysing liquid phase samples in commercial FTIR spectrophotometers. Recently[53-55], following three types of probes have been proposed using fluoride fiber for the detection of methane and propane gases using source between 3.2 - 3.6 μm region: (a) In one, a short section of cladding is removed from multimode fiber to access the evanescent field region. (b) In second, the gas diffuses through the pores of sol-gel or teflon cladding of fiber causing an attenuation of the evanescent field and thereby a reduction in transmittance of the fiber. (c) In the third, an evanescent wave methane detection has been suggested using D-fiber or cladding reduced locally by polishing. These are briefly described in the following sections:

5.1. Cladding removal with evanescent wave modulation[47,56]

In this method, a cylindrical or planar optical waveguide is used and its constructions is such that a part of the guided electromagnetic field extends beyond the physical boundary of the waveguide. An attenuation of the this field can be observed when an absorpting substance is present at the waveguide boundary as shown in Fig. 4. Any phenomenon that disturbs the intensity of the evanescent wave also produces modulation in the intensity of the fiber-guided light. A cylindrical waveguide and a 3.39 μm He-Ne laser has been used by Tai et al[56] for detection of CH_4 in nitrogen. In this experiment, a 40 cm length fused silica fiber served as waveguide. Several millimeter section of the fiber was stripped of its cladding to allow the evanescent field to penetrate into the ambient air. A representative diagram of this system is shown in Fig. 5. These authors[47,56] were able to detect CH_4 concentration up to 5% with this system.

5.2. Sol-Gel Cladding[47,57,58]

A layer of sol-gel makes an extremely versatile cladding, as shown in Fig. 6. The sol-gel process, is the same as that used in making glass and ceramics at low temperatures via hydrolysis and polymerization of organic precursors. A combination of metal alkoxides, solvent, catalyst and water are mixed to obtain homogeneity. The hydrolysis and condensation polymerization produce a viscous gel, which is an amorphous porous material containing liquid solvent in its pores. The gel is heated at less than 100^0C so that most of the liquid is expelled, leaving the porous oxide, and then densified at a higher temperature. The materials and their concentrations can be modified to vary both the material's refractive index and porosity. This process is especially suited to produce a thin layer deposited on the fiber core. An uncladed fiber covered by a layer of sol-gel can either be used directly or else a reagent with absorption or fluorescence varying with an external parameter can be trapped in the sol-gel.

Cladding with color variations[58] has also been used for detection of hydrocarbons. In this technique the color, and thus absorption spectrum of the cladding can be modulated by an external parameter.

5.3. Evanescent field detection of CH_4 gas using D-fiber[55,56,59-61]

D-fibers are produced by polishing a flat on the fiber preform to give a D-shaped cross-section, and then preform is subsequently pulled into fiber. D-fiber may be used as evanescent-wave CH_4 gas sensor by performing attenuation measurements on guided wave at a wave-length corresponding to 1.33 and 1.66 μm. Fig. 7 shows the block

diagram of a CH_4 gas sensor using D-fiber. In this setup a LD of 1.66 μm was coupled into a length of standard singlemode (SM) fiber spliced on to a 5 m length of D-fiber, core diameter of 3 μm and zero d-distance[59,60]. The D-fiber section has been inserted into the methane gas test chamber and output of the D-fiber is connected to a scanning Fabry-Perot (FP) resonator by a length of a multimode (MM) fiber. The detected signal from the FP could be monitored with an electronic spectrum analyser (ESA). The sensitivity of D-fiber was compared with that of an open path gas cell using the same detection system. Per equal length, the D-fiber was found to be less sensitive which can be improved by increasing index difference Δn (e.g. increasing Δn from 0.02 to 0.03, sensitivity increased twice). It has been shown by Muhammad et al.[59] that the sensitivity of D-fiber methane gas sensor can be realistically improved by overcoating the flat surface of fiber with a high index layer. Cross-section of overcoated D-fiber are shown in Fig. 8-a and 8-b.

6. CORRELATION SPECTROSCOPY FOR THE DETECTION OF GAS[62]

Recently, Dakin and Edwards[62] have reported gas sensor using real-time correlation spectroscopy, where a gas is used as a matched optical fiber to 'recognise' its own spectral absorption lines. The basic concept of correlation spectrometry involves the passage of light sequentially through two gas cells: a reference cell containing a known quantity of the gas to be detected, and a sampling cell where the presence of the gas is to be determined. An optical signal passing through both cells will suffer absorption due to the gas in each. If the absorption in the reference cell is periodically modulated, then the total absorption depends on whether the gas absorption lines in the sampling cell correlate with those in the reference cell gas. It is similar to the optical methanometer presently being used for the detection of methane.

7. DISCUSSION

The optical detection of methane gas by OA and DOA methods require an open cell with light propagating through it as the detection system. This requires a precise, expensive and delicate alignment of optics. Further it has disadvantage of transmitting the light to NIR region, so only the weaker overtone bands can be monitored. On the other hand the low-loss fiber (1 dB/km) have the advantage of transmitting the light to a longer distance which allow the sensor head to be sited remotely from the monitoring station, and offer the possibility of multiplexing a number of sensors. Thus, the OA and DOA methods of detection of methane gas may not be very much suitable for the underground mines.

In the case of fluoride fiber, their is no restriction of transmitting the light of NIR region only, it can transmit the wavelength between 2-5 μm or more which helps us in detecting main absorption band of methane gas at 2.27, 3.33 and 3.39 μm but the IR fibers are having more loss (20 dB/km), and hence cannot be used for remotely located gas cell at long distance in underground mines. These fibers are expensive and hygroscopic, and hence may not be much suitable for the underground mines where the environment is humid in addition to explosive and toxic. The IR fiber optic methane sensors are good for few meters measurements or in fiber optic instruments. Recently[54,63], a fiber optic FTIR instruments was made available that can support 100 m ZrF fiber and a gas cell. However, the prices of the fiber optic FTIR is more than $ 70,000. This appears very expensive for mining application. Correlation spectroscopy may not be suitable because it requires a gas filled cell for the reference which may leak after certain time.

An evanescent field device has the advantage that it can be both optically and mechanically coupled directly to an IR fiber optic link which makes rugged and low-light loss package. The absorption characteristics of D-fiber obviates the entire precision optics when used for methane detection. From above discussion it appears that the evanescent field detection of CH_4 gas using D-fiber may be best for the mining use. It is more rugged having less expensive optics and may be used for a long-distance network.

Parallel and sequential measurements[45] can be made employing a number of pairs of optical fibers connected to the individual sample cell kept at different location in underground mines for the monitoring of CO and CH_4 gases, and useful information regarding their concentration can be obtained. Using OTDR, a distributed gas sensor system can be designed specially for the underground coal mines. A laboratory model for fiber optic methane sensor is under development and will be appearing in forthcoming paper.

One of the major advantage of the optical gas detection system is that it is an intrinsically safe, flameproof and fail-safe. It will not give a misleading zero reading if CH_4 gas is in dangerous limit. A catalytic type detector may not detect CH_4 gas at all if catalyst has been contaminated.

ACKNOWLEDGEMENT:

The authors are grateful to Dr. D.K.Paul, Director, Indian School of Mines, Dhanbad, for continuous inspiration and All India Council of Technical Education (AICTE), New Delhi, for financial assistance for the project on 'Development of fiber optic methane sensor for mines'.

REFERENCES

1. H.Debeda, P.Mossok, C.Lucat, F.Manil and J.L.Aucouturier, "Methane sensing:from sensitive thick films to realiable selective device," *Meas Sci. Technol* **8**, 99, 1997.
2. H.Debeda, C.Lucat, F.Menil and P.Massok, "Selective thick film gas sensors," *Ann. Chim,* **20**, 481, 1995.
3. F.Menil, C.Lucat and H.Debeda, "Thick film rout to selective gas sensors", *Sensors and Actuators* **B45**, 415, 1995.
4. P.Dutronc, C.Lucat, F.Menil, M.Loesch, M.C.Horillo, I. Sagayo, J.Gutierrez, J.A.De Agapito, "A potentially selective methane gas sensor based on the differential conductivity responses of Pd and Pt-based tin oxide thick layer", *Sensors and Actaators* **B 15-16,** 384, 1993.
5. P.Dutronc, C.Lucat, F.Menil, M.Loesh and L.Combes, "A new approach to selectivity in methane gas sensing", *Sensors and Actuators* **B 15-16,** 24, 1993.
6. G.Sberveglieri, "Recent development in semiconducting thinfilm gas sensors", *Sensors and Actuators* **B23**, 103, 1995 and references therein.
7. R.Huck, U.Bottger, D.Kohl and G Heiland, "Spillover effects in the detection of H2 and CH4 by sputtered SnO2 films with Pd nad PdO deposits", *Sensors and Actuators* **17**, 335, 1989.
8. B.B.Dhar and B.C.Bhowmic (Editors), Proc. 27th Inter. Conf. on Safety in Mines Resch. Institutes, ICS-MRI '97, Vol.I, 20-23 Feb. (1997) New Delhi. Oxford & IBH Pub. Co., New Delhi, India, pp 321-373.
9. H.Ikeman, "Application of fiber optics in the mining industry", *The Canadian Mining and Metallurgical Bulletin*, **77**, 41, 1984.
10. Dipl.Ing.Reinhold Schutz, "Optical fiber conductors also in mining industry", *Gluckauf + translation*, **122**, 1986.
11. R.S.Romaniuk, J.Dorosz and F.Szczot, "Lightguide technology for adverse industrial environment" (review of selected university based work in Poland), J.Opt.Sensor, **2**, 163 (1987)
12. R.S.Romaniuk, F.Szczot and J.Dorosz, "Lightguide technology for adverse industrial environment", *Proc. SPIE, Int. Soc. Opt. Eng.(USA)* **721**, 19 (1986)
13. J.K.Zientkiewicz, Self-referenced fiber optic system for remote methane detection", *Proc. SPIE*, : *Fiber optics* **1120**, 284, 1989.
14. Ge.Linlin, "An optical fiber system for the multiparameter observation of underground water precursors", Prepared for the Int.Conf.on Opto-electronic Sci.& Engg. 1990.
15. B.J.Skutnik, M.H.Hodge, J.P.Clarkin, "Passivated optical fiber for harh environments", *Proc. SPIE, Fiber Optic Reliability: Benign and Adverse Environment* **842**, 1987.
16. V.Kumar and D.Chandra, "Fiber optic methane sensor for mines", *Optica Applicata,* **XXVII**, 279-290 1997.
17. V.Kumar and D.Chandra, "Possible applications of OTDR in the mining industry", *Mineral Resources Engg. (UK),* **7**, 5-15,1998.
18. V.Kumar, Satyaprakash and D.Chandra, "Application of fiber optic sensors in geophysics - A review", *Acta Geophysica Polonica*, **XLIII**, 343-79, 1995.
19. V.Kumar and G.M.Prasad, Fiber optic sensors for mines, Mining Technol. **71**, 187-88, 1981.
20. G.Hrzberg, *Molecular spectra and molecular structure II, Infrared and Raman spectra of polyatomic molecules,* Van Nostrand, New York , 1966.
21. R.A.McClatchey, W.S.Benedict, S.A.Clough, D.E.Burch, R.F. Calfee, K.Fox, L.S.Rothman and J.S.Garing, "AFCRL atmorpheric absorption line parameter compilation", Environmental Research paper 434 AFCRL-TR-73-0096 (Air Force Cambridge Research Laboratories, Bedford, Mass, 1973).
22. W.V.Norris and H.J.Unger, *Phys. Rev.* **43**, 467 1933.
23. K.Chan, H.Ito and H.Inaba, "Absorption measurement of $\gamma_2 + \gamma_3$ band of CH_4 at 1.33 μm using an InGaAsP light-emitting diode", *Applied Optics*, **23**, 3802, 1983..

24. K.Chan, H.Ito and H.Inaba, "Optical remote monitoring of CH_4 gas using low loss fiber link and InGaAsP light-emitting diode in 1.33 μm region", *Appl.Phys.Lett.*,**43**, 634, 1983.
25. S.Stueflotten, J.Christensen, S.Iversen, J.O.Hellvik, K. Almas, T.Wien and A.Graav, "An infrared fiber optic gas detection system", Proc. OFS'84, p. 87, 1984.
26. K.Chan, H.Ito and H.Inaba, "All-optical fiber based remote sensing system for near infrared absorption of low level CH_4 gas", *J. Lightwave Technol.* **LT5**, 234, 1984.
27. Y.Arakawa, H.Fukunaga and H.Inaba, "Fiber optic system for multipoint remote detection of inflammable gases", *Proc. OFS'86*, Tokyo, p.135, 1986.
28. J.P.Dakin, C.A.Wade, D.Pinchbeck and J.S.Wykes, "A novel optical fiber methane sensor", *J.Optical Sensors*, **2**, 261, 1987.
29. M.Saito, M.Takizama, K.Ikegawa and H.Takami, "Optical remote sensing system for hydrocarbon gases using infrared fibers", *J.Appl.Phys.* **63**, 269, 1988.
30. H.Tai, H.Tanaka and T.Yoshino, "Fibre-optic evanescent wave methane gas sensor using optical absorption for the 3.392 μm line of a He-Ne laser", *Opt.Letts.* **12**, 437, 1987.
31. G.Stewart, C.Tandy, D.Moodie, M.A.Morante and F.Dong, "Design of fiber optic multi-point sensor for gas detection", *Sensor & Actuators,* **B51**, 227-32 (1998).
32. G.,Stewart, A.Mencaglia and W. Philp, "Interferometric signals in fiber optic methane sensors with wave-length modulation of the DFB laser source", *J.Lightwave Technol.,* **16**, 43-53 (1998).
33. W.Jin, M.S.Demokan, Gstewart, W.Philp and B.Culshaw, "Performance limit of fiber-optic gas sensor from coherent back scattering", *IEE Proc., Optoelectronic (UK)*, **145**, 186-90 (1998).
34. A.G.Mignani, A.Mencaglia and M.Brenci, "Micro-optic probes for gas absorption measurement:design study and demonstrators", *Proc. SPIE- Int. Soc Opt. Eng.(USA),* **3278**, 229-33 (1998)
35. E.J.Lerner, "Fiber optic sensors monitoring environmental condition", *Laser Focus Worls (USA),***33**, 107-8 and 110-12 (1997).
36. H.Inaba, T.Kobayasi, M.Hiruma and M.Hamza, "Optical fiber network system for air-pollution monitoring over a wide area by optical absorption method", *Electron. Lett.* **15**, 749, 1979.
37. H.Inaba, *Optical and Laser Rremote Sensing*, Eds. D.K. Killinger and A.Mooradian, Springer Series in Optical Sciences, 39, Springer-Varlag, Berlin ,p 288, 1983.
38. H.Nagai, M.Aizawa and J. Ono, "Long distance remote sensing of methane gas using a 1.65 um DFB LD and a single mode fiber", *IEEE Laser and Electro-optics (LEOS)* 1995, Annual Meeting Conf. Proc. (Cat. No. 95CH35739), San Francisco, CA, USA, (New York, NY, USA: IEEE '95) **2**, p.316-17, 30-31 Oct. 1995
39. K.Chan, H.Ito and H.Inaba, "All optical remote monitoring of propane gas using a 5-km-long, low-loss optical fiber link and an InGaP light-emitting diode in 1.68 μm region", *Appl. Phys.Lett.* **45,** 220, 1984.
40. K.Chan, H.Ito and H.Inaba, "An optical fiber based gas sensor for remote absorption measurement of low-level CH4 gas in the near-infrared region", *IEEE, J. Lightwave Technol.*, **LT-2**, 234, 1984.
41. K.Chan, H.Ito and H.Inaba, "Optical fiber remote sensing system of low-level propane gas using 1.68 um InGaAs light-emitting diode", *Optics & Laser in Engg.*, **6**, 119, 1985.
42. K.Chan, H.Ito and H.Inaba, "10 km-long remote sensing of CH_4 gas by NIR infrared absorption", *Appl.Phys.* **B 138,** 11 (1985).
43. K.Chan, T.Furwya, H.Ito and H.Inaba, "Fully optical remote measurement of CH_4 gas in the near-infrared region using 5-km-long, low-loss optical fiber link", *Optical and Quantum Electro.* **17**, 153 1985.
44. K.Chan and H.Inaba, "Fiber optic remote gas sensing system by near infrared absorption", Ist Workshop on optical fiber sensors (WOFSI-1), July 12, 1985, Japan Soc. of Appl.Phys., p 1.
45. H.Inaba, "Remote sensing of environmental pollution and gas dispersal using low-loss optical fiber network system", *Laser/Optoelectronics in Engg., Proc. 7th Int. Congress Laser 85, Optoelektronic, Munich,* July 1985, Ed: W. Waidelich, Springer-Verlag, Berlin 1986.
46. M.C.Alarcon, H.Ito and H.Inaba, "All optical remote sensing of city gas through CH4 gas absorption employing low-loss optical fiber link and InGaAsP light-emitting diode in the near-infrared region", *Appl.Phys.* **B43**, 79, 1987.
47. A.G.Mignani, M.Brenci and A.Mencaglia, "Fiber optic sensors for environmental monitoring, Trends in optical fiber metrology and standard", Ed. D.D.D.Soares, NATO-ASI Series E: Applied Sciences, vol.285, Kluwer Academic Pub. p.691, 1995.
48. J.P.Dakin, "Optical fiber environmental sciences, Optical fiber sensors", *Springer Proc. in Phys. vol.44*, Eds. H.J. Arditty, J.P.Dakin and R.Th.Kersten, Springer-Verlag Berlin, Heidelberg 1989.
49. K.Chan, H.Ito and H.Inaba, "Remote sensing system for near-infrared differential optical absorption of CH_4 gas using low-loss optical link", *Appl. Opt.* **23**, 3415, 1984.

50. A.Mohebati and T.A.King, "Differential absorption fiber optic sensor", *Proc. SPIE, Fiber optic sensor III*, **1011**, 183, 1988.
51. W.Jin, G.Stwart and B.Coulshaw, "Performance limitation of fiber optic methane sensors due to interference effects", *J. Lightwave Technol.* **14**, 760, 1996.
52. J.U.White, " Long optical paths of large aperture", *J. Opt. Soc. Am.* **32**, 285, 1942; and "Very long optical paths in air ", *J.Opt.Soc.Am.* **66**, 411, 1976.
53. V.Ruddy, B.MacCraith and S.McCabe, "Remote flammable gas sensing using fluoride fiber evanescent probe; and Ruddy et .al Evanescent wave absorption spectroscopy using multimode fibers". (submitted for publication).
54. B.S.Matson and J.W.Grifin, "Infrared fiber optic sensors for the remote detection of hydrocarbons operating in the 3.3 to 3.6 micron region", (Private Communication).
55. B.Coulshaw, F.Muhammad, G.Stewart, S.Murray, D.Pinchbeck, J.Norris, S.Cassidy, M.Wilkinson, W.Williams, I.Crisp, R.Van Ewyk and A.McGhee, "Evanescent wave methane detection using optical fibers", *Electronics Lett.* **28**, 2232, 1992.
56. H.Tai, H.Tanaka and T.Yoshino, "Fiber optic evanescent wave methane gas sensors using optical absorption for the 3.392 micron line of a He-Ne laser", *Opt.Lett.* **12**, 437, 1987.
57. B.D.MacCraith, "Enhanced evanescent wave sensors based on sol-gel derived porous glass coating", *Sensor and Acuators* **B11**, 29, 1993.
58. M.Tabacco, Q.Zhou and B.Nelson, "Chemical, biochemical and environmental fiber sensor III", Ed. R.A.Lieberman SPIE Proc. Vol.1587, Bellingham (1992). Chemical sensor for environmental monitoring, p. 271.
59. F.A.Muhammed, G.Stewart and W.Jin, "Sensitivity enhancement of D-fiber methane gas sensor using high-index overlayer", *IEE Proc. J.* **140**, 115, 1993.
60. F.A.Muhammed and G.Stewart, "D-shaped optical fiber design for CH_4 gas sensing", *Electron Lett.* **28**, 1205, 1992.
61. F.A.Muhammed and G.Stewart, "Polarised finite difference analysis of D-fiber and application of chemical sensing", *Int. J. Optoelectron.* **7**, 705, 1992.
62. J.P.Dakin and H.O.Edwards, "Progress in fiber remote gas correlation spectroscopy", *Opt. Engg.* **31**, 1616, 1992.
63. S.C.Stinson, "*Advance media in applying IR sensor to process control*", CE & N , p.30, 1989.

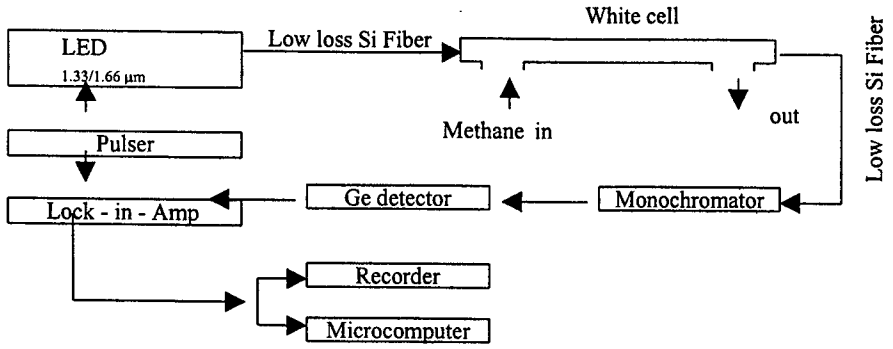


Fig.1: Block diagram of an optical system for remote detection of CH_4 gas based on optical absorption method.

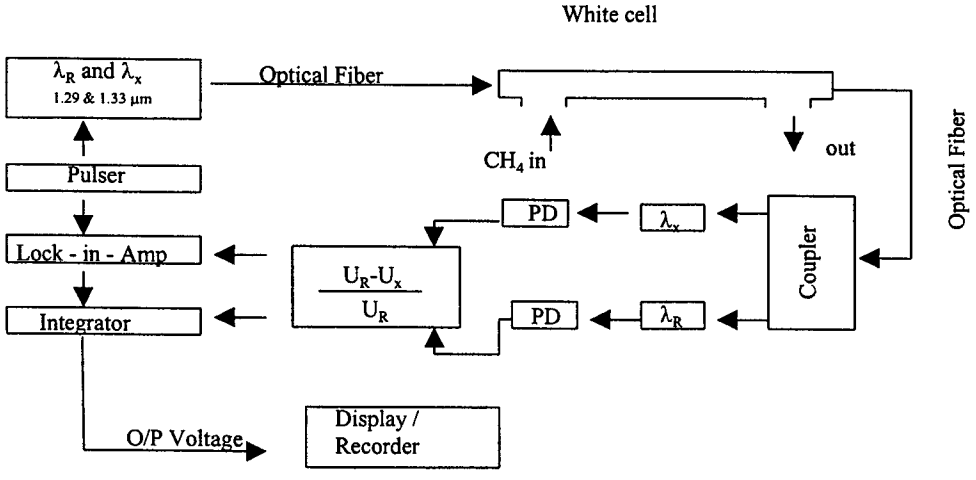


Fig.2: Block diagram of an optical system for remote detection of CH_4 gas based on differential optical absorption Method

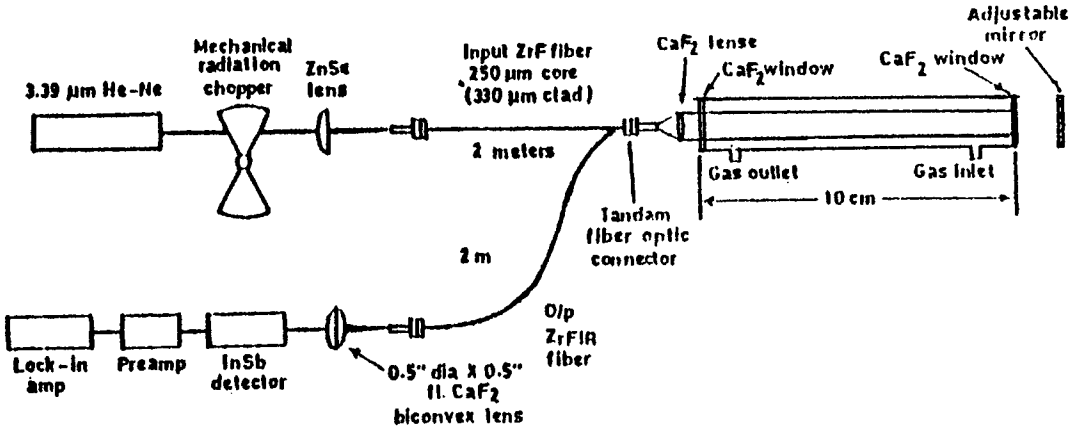


Fig.3: Block diagram of IR fiber optic absorption sensor for CH_4 detection.

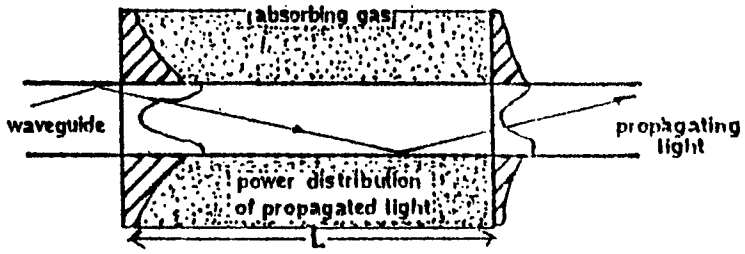


Fig.4: Pricniple of detection of absorbing substance using evanescent field. Power distribution of the propagated light through a waveguide is shown.

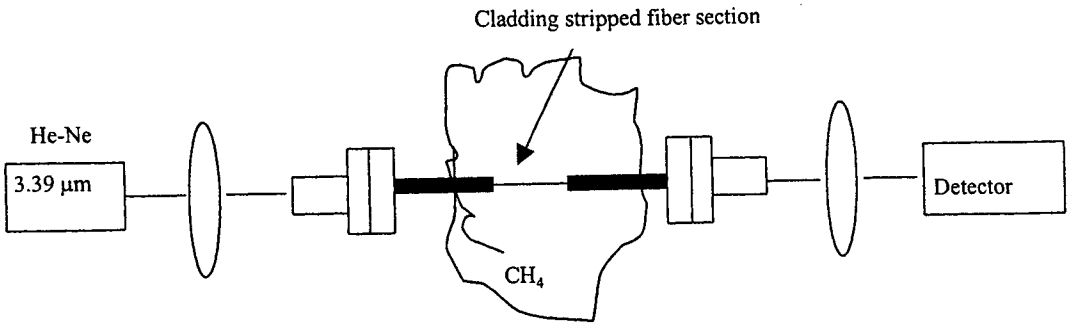


Fig.5: A device for the detection of methane based on the attenuation of the evenescent field.

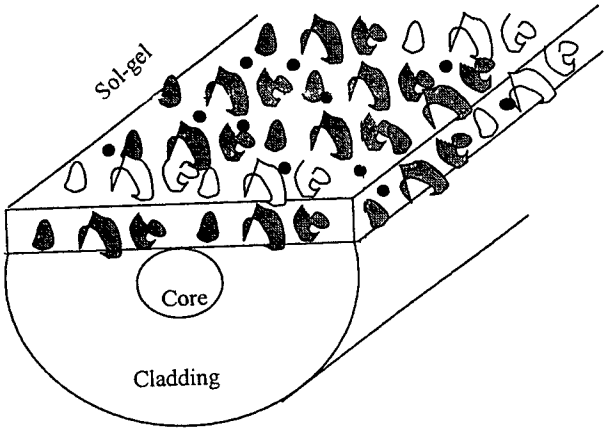


Fig.6: Sol-gel coating for fiber optic intrinsic sensing.

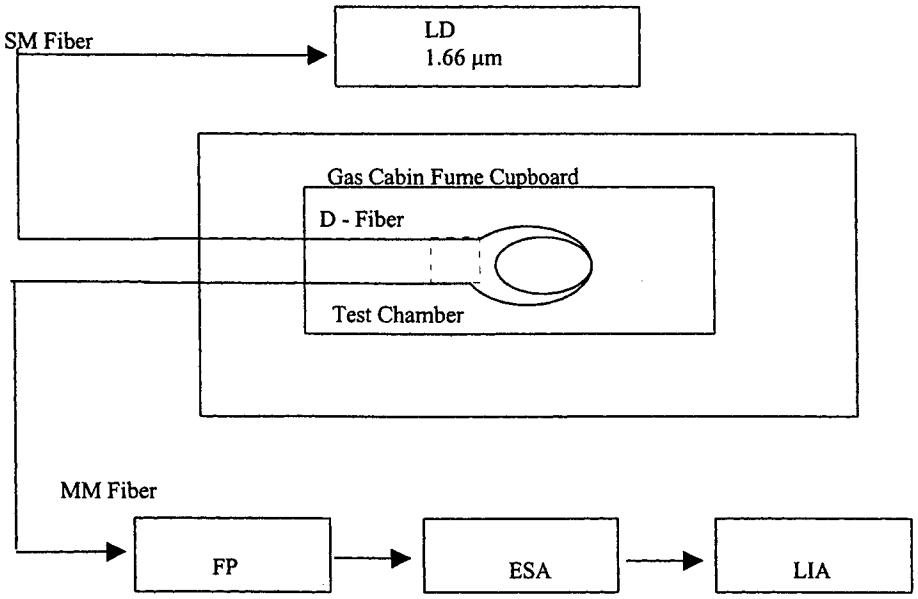


Fig.7: Block diagram for detction of methane gas usning D-fiber.

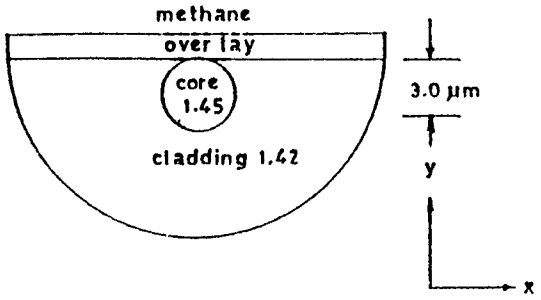


Fig.8a:Cross-section of overcoated D-fiber used for methane detection (not to the scale).

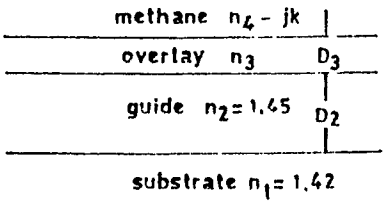


Fig.8b:Overcoated palner waveguide with thin high index overlay (not to the scale).

Coating of polystyrene thin film on glass for protein immobilization in optical biosensor applications

Timothy T.M. Tan[1], Yik Yuen Gan*[1], Leong Huat Gan[1], Thomas Yong[1], Bing Zhou[2], Yee Loy Lam[2] and Yan Zhou[2]

[1]School of Science; [2]School of Electrical and Electronic Engineering,
Nanyang Technological University, Nanyang Avenue, Singapore 639798, Singapore

ABSTRACT

Immobilizing protein on glass surfaces is typically more difficult and less efficient than on plastic surfaces. Proteins are readily adsorbed on plastic surfaces in a single step. To simplify protein immobilization efficiency on glass surfaces and enhance its efficiency, styrylsilane and polystyrene were coated on glass to serve as protein binding substrates.

The efficiency of protein immobilization on plain glass, styrylsilane coated glass, polystyrene thin film coated glass and polystyrene microtiter well were compared. It was found that styrylsilane and polystyrene thin film coated glasses have similar protein immobilization efficiency and kinetics as polystyrene microtiter wells commonly used in immunoassays. Both types of coated glass have a protein binding capacity of approximately 244 picograms/cm^2, polystyrene wells at about 222 picograms/cm^2 and plain glass, 89 picograms/cm^2.

Protein immobilization on glass was improved by coating the glass with styrylsilane or polystyrene thin films. This can be done easily and at low cost when compared to traditional methods, making this method very suitable for producing disposable optical fiber and waveguide biosensors. A model biosensor system utilizing recombinant Epstein-Barr viral proteins as antigen targets immobilized on a buried waveguide in the detection of nose and throat cancer is discussed.

Keywords: polystyrene thin film, protein immobilization, optical waveguide biosensor

1. INTRODUCTION

In recent years, we have seen a lot of research interest in the development of optical biosensors. A key feature of such a sensor is its one-step wash-free convenience and this is achieved through the evanescent interaction of light with the analyte[1]. Both optical fibres and optical waveguides have been used in the design of these sensors and it has been shown that using an optical waveguide, the concept of a one-step multi-analyte disposable optical biosensor can be realised[2-3].

For such an optical biosensor to be viable for commercial applications, the immobilization of proteins on the glass surface must be simple and of low cost. There exists an extensive array of protein immobilization methods[4-7], including adsorption, entrapment, protein crosslinking and covalent binding (or coupling). Physical adsorption and covalent binding are the most commonly adopted methods for optical biosensor applications because they are reasonably reproducible, have high protein loading and stability and do not affect immunological activity. Compared with physical adsorption, covalent binding is more predictable and has a higher capacity for protein immobilization. However, covalent binding is quite complex[8-9], and tends to result in greater non-specific binding.

While non-specific binding can be minimized by the application of blocking agents like bovine serum albumin (BSA), physical adsorption often provides comparable or sometimes better performance than covalent binding in terms of signal-to-non-specific binding noise ratio. In addition, it is simple to perform and is thus of low cost. This makes physical adsorption a

*Correspondence: Associate Professor Gan Yik Yuen, Division of Biology, School of Science, National Institute of Education, Nanyang Technological University, 469 Bukit Timah Road, Singapore 259756; Email: YYGAN@nie.edu.sg; Telephone: (65) 460 5331; Fax: (65) 469 8928.

Part of the SPIE Conference on Advanced Photonic Sensors and Applications
Singapore • November–December 1999
SPIE Vol. 3897 • 0277-786X/99/$10.00

desirable method for disposable device applications. Physical adsorption of proteins on glass surfaces is often done by treating the glass surface with a repel coat to make it hydrophobic and incubating protein solution on the surface overnight. Many types of immunoassays are typically performed in polystyrene microtiter wells. Proteins are easily adsorbed onto the polystyrene walls of the microtiter wells, mainly through hydrophobic interactions[5,10].

Noting that polystyrene has a much higher protein adsorption capacity than glass, we have managed to produce glass surfaces coated with a polystyrene thin film to investigate the utility of this method for protein immobilization on glass in optical biosensor applications. A comparison study was then done to determine the relative protein-binding profiles of glass (sulphuric acid cleaned), styrylsilane-coated glass and polystyrene-coated glass and polystyrene microtiter wells.

The amount of protein that is bound by physical adsorption on glass and polystyrene surfaces is too low for standard protein assay techniques. It was decided that the simplest method of determining the amount of protein bound to such surfaces was to use an enzyme that catalyses a color change. Horseradish peroxidase is often used in enzyme linked immunoassays and has several commonly used substrates. For example, the substrate TMB is colorless and develops a blue color proportional to the amount of enzyme present and the length of reaction. At acidic pH, the enzyme is inactivated and the blue color turns yellow and the absorbance of this yellow color can be read spectrophotometrically. The absorbance is proportional to the color change, and hence the amount of enzyme originally present.

2. MATERIALS AND METHODS

2.1 Materials

Soda lime glass slides (1mm x 26mm x 75mm) were cleaned by immersing them in concentrated sulphuric acid for 15 min, rinsed with deionized water and baked at 140°C for 15 min.

Dow Corning Z-6032 silane, molecular formula: $(CH_3O)_3$-$SiCH_2CH_2NHCH_2CH_2NHCH_2$ -$CH{=}CH_2HCl$) was prehydrolysed in methanol under acidic conditions by mixing 100 parts Z-6032 : 5 parts glacial acetic acid : 25 parts water. After 1 hr, the silane completely solubilized and was equivalent to 30% active silane. It was then diluted with water to 0.5% active ingredient. 1% dicumyl peroxide was then added to the prehydrolyzed silane solution. Glass slides cleaned as above were immersed for 2 min, then rinsed in deionized water, blown dry with filtered nitrogen and baked at 110°C for 10 min. The glass slides were thus uniformly coated with styrylsilane.

General purpose polystyrene pellets were dissolved in xylene to a final concentration of 1 mg/ml and stirred until the solution was clear and uniform. A thin film of polystyrene was coated on cleaned glass slides using spin-coating at 2000 rpm for 30 sec. Polystyrene spin-coated glass slides were incubated at 30°C overnight for the xylene to fully evaporate.

Styrylsilane and polystyrene-coated glass slides were prepared with both surfaces coated. 'Plain' glass slides were sulphuric acid cleaned only. Individual glass 'chips', approximately 1mm x 6mm x 13mm each, were cut from the glass slides by careful scoring with a diamond-tipped cutter. Glass chips significantly scratched, cracked or of incorrect dimensions were discarded. Care was taken to ensure chips were as closely matched in size as possible.

Polystyrene microtiter wells ("Maxisorp" type, Nunc A/S Denmark) were used as a protein coating positive control. Uncoated, clean microtiter wells were also used for color development in the HRP assay described below.

2.2 Adsorption of horseradish peroxidase conjugate

Sheep anti-human Immunoglobulin A horseradish peroxidase (HRP) conjugate (Silenus, Australia) was diluted 1:100 in carbonate buffer (15 mM sodium carbonate, 35 mM sodium bicarbonate, pH 9.6) to a final concentration of 75 µg/ml. The cut glass chips were immersed in the HRP conjugate solution at 4°C for 1, 4, 8, 16 and 24 hrs with occasional gentle agitation. For each time period, 3 chips were concurrently immersed in the HRP conjugate solution. After immersion, the glass chips were gently rinsed 3 times with carbonate buffer and stored at 4°C in fresh buffer without HRP conjugate until needed. As positive controls, polystyrene microtiter wells were similarly treated using 50 µl of HRP conjugate solution per well. Glass chips and microtiter wells treated with carbonate buffer only for 24 hrs at 4°C served as negative controls.

2.3 HRP Assay method

Glass chips were removed from carbonate buffer, blotted to remove excess liquid by briefly touching the edge to a clean lint-free tissue before placing them individually into clean, uncoated microtiter wells (400µl capacity). Positive control microtiter wells were aspirated and residual HRP solution removed by inverting microtiter plate and patting on a lint-free tissue. 200 µl of chromogen substrate buffer (100 µg/ml 3,3',5,5'-tetramethylbenzidine (TMB), 0.012% hydrogen peroxide, 20 mM citric acid, 60 mM sodium acetate, pH 6.0) was added to each well quickly and color was allowed to develop for exactly 30 min in the dark at room temperature before glass chips were removed and the reaction stopped with 100µl per well of 1M sulphuric acid. Absorbance was read spectrophotometrically at 450 nm with differential filter at 630 nm using a Bio-Rad Model 550 microplate reader. Clean wells incubated as above with substrate buffer served as blank controls. All samples were tested in triplicate.

2.4 Calibration curve of horseradish peroxidase concentration

HRP conjugate was serially diluted in carbonate buffer to final concentrations in the range 0 to 150 µg/ml. 10 µl of each dilution was placed at the bottom of clean microtiter wells and 200 µl of substrate buffer added, developed and absorbance measured as before. A calibration curve of amount of HRP conjugate protein (7.5 to 1500 picograms) against absorbance was drawn and used to estimate the amount of HRP conjugate bound to the glass chips and controls.

3. RESULTS

3.1 Polystyrene thin film coating

The thickness of the polystyrene film prepared as described was measured to be about 20Å. There was no detectable difference in the polystyrene thin film after immersion in deionized water for 24 hrs. The polystyrene thin film was highly reproducible and found to be even and durable.

3.2 HRP calibration curve

The amount of HRP conjugate in an unknown sample can be read directly off a semi-log plot (Figure 1) of HRP conjugate against the absorbance.

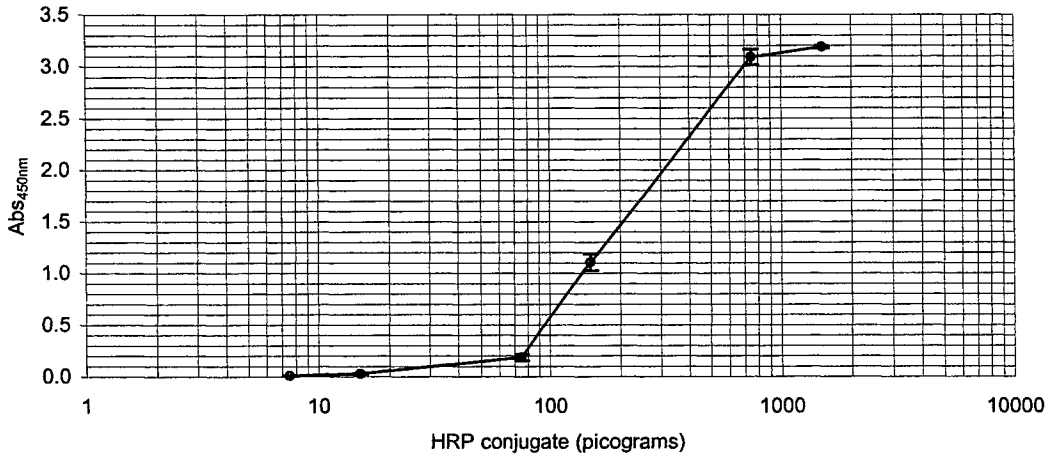


Figure 1: Semi-log plot of HRP calibration curve: spectrophotometric absorbance readings at 450nm (Abs_{450nm}) against amount of HRP conjugate in picograms. HRP assay was performed using the same conditions given in Section 2.3. The amount of HRP conjugate present in unknown sample can be estimated from the calibration curve given the absorbance value of the HRP assay.

3.3 HRP assay to determine amount of protein adsorbed over time

Uncoated, styrylsilane coated and polystyrene coated glass chips were tested for protein adsorption capacity using the HRP assay method described in Section 2.3. The kinetics of the adsorption of HRP conjugate to the various surfaces was determined by plotting the mean absorbance reading for each sample against adsorption time (Figure 2a-d). The absorbance readings for each type of surface immersed in carbonate buffer only, without any HRP conjugate present (negative control) effectively represented 'zero' adsorption time (0 hrs). The background (blank control) absorbance readings were negligible.

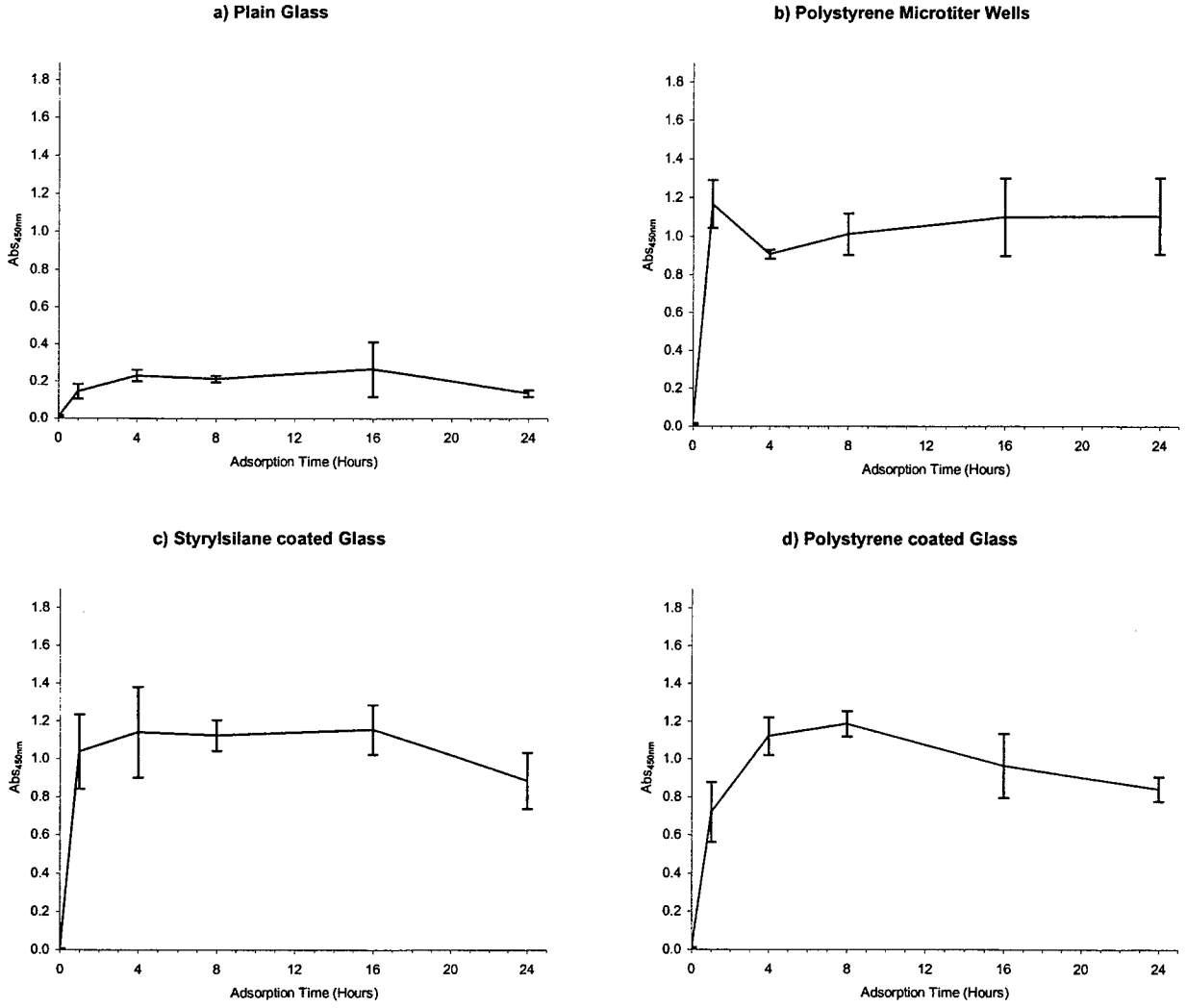


Figure 2: Spectrophotometric Abs_{450nm} of HRP assay using HRP conjugate immobilized over various adsorption time on a) plain glass, b) polystyrene microtiter wells, c) styrylsilane coated glass and d) polystyrene coated glass. Time zero data points reflect the negative control samples (immersed in carbonate buffer only). Mean Abs_{450nm} readings of 3 individual samples for each adsorption time were plotted, vertical bars indicate the standard deviation.

3.4 Calculation of protein immobilization capacity

The actual surface area of each glass chip (both flat surfaces) exposed to substrate buffer was approximately 0.90 cm^2. The calculation of the surface area of the glass chips ignored the edge thicknesses. The surface area of polystyrene wells adsorbed with HRP conjugate solution was 0.63 cm^2.

To determine the protein immobilization capacity of the various surfaces, the amount of HRP enzyme conjugate bound to each type of glass chip was estimated from the HRP calibration curve (Figure 1) given the absorbance. The amount of HRP conjugate (in picograms) was then divided by the surface area (in square centimeters) of adsorbed protein, to give the protein immobilization capacity in each case. Results were calculated using the data at 8 hrs adsorption time for all the samples.

The protein immobilization capacity of styrylsilane coated glass and polystyrene coated glass were nearly identical and was approximately 244 picograms/cm^2. In comparison, pure polystyrene microtiter wells had a protein immobilization capacity of 222 picograms/cm^2 and plain glass, 89 picograms/cm^2.

4. DISCUSSION

4.1 Comparison of protein immobilization capacity

As described in the introduction, the use of an enzyme as the model protein in this study of the amount of protein adsorbed on a given surface is ideal due to the intrinsic amplification of signal to allow convenient detection of very low levels of adsorbed protein. The enzyme used was actually a conjugate of the horseradish peroxidase (a globular protein) and an immunoglobulin protein molecule. The adsorption characteristics of different types of proteins vary and therefore the presence of both a globular protein and an immunoglobulin domain should represent an 'average' of the two. In any case, the results of the experiment show the relative protein adsorption capacity of the various surfaces.

Within the limits of experimental error, the protein adsorption capacity of styrylsilane coated and polystyrene coated glass are equivalent to that of pure polystyrene. The amount of protein adsorbed by plain glass is about 2½ times less. This is significant for the development of optical biosensor-based immunoassays because the relative amounts of bound protein is directly comparable to the amounts of bound protein (such as antigens) in existing assay systems utilizing polystyrene microtiter wells. The amount and relative concentrations of such bound proteins are critical parameters with narrow optimal ranges in immunoassays, where small variations result in wholly different results. A lowered capacity for adsorbed protein would likely cause the lower limit of detection to be too high to be useful, while excessive adsorbed protein affects immunoassays in myriad ways, such as steric hinderance of antigens and increased non-specific binding of antibodies.

4.2 Characteristics and kinetics of protein immobilization

It is noteworthy that the relative 'unpredictability' of physical adsorption can be seen in the standard deviation between tested samples. It appears that uncoated glass (acid cleaned) has relatively uniform adsorption characteristics, while the coated glass and pure polystyrene have approximately the same variability in adsorption. In practice, this variability in adsorption is acceptable and can be minimized through careful optimization and control of various parameters.

The protein adsorption kinetics are also very similar for the various surfaces with significant amounts of protein being immobilized within the first hour, even at 4°C, with equilibrium achieved at approximately 8 hours. The rate of protein absorption is of course proportional to the concentration of applied protein solution, and in this case, it was relatively high at 75 µg/ml. Typically, adsorption of globular proteins to polystyrene for use in immunoassays are applied at 100 ng/ml to 1 µg/ml and reach equilibrium at 16 to 24 hours. It is not currently known why the amount of adsorbed protein apparently falls slightly after 8 hours in the case of polystyrene coated glass, although it does not appear to be due to degradation of the polystyrene thin film.

4.3 Polystyrene thin film vs. styrylsilane coating

The chief attraction of polystyrene thin film coatings versus the more conventional styrylsilane coating of glass for the purpose of protein immobilization lies in its low cost. This factor contributes greatly to the viability of the disposable biosensor concept. The chemical similarity between the polystyrene thin film and pure polystyrene also makes development work easier in migrating existing polystyrene microtiter well based immunoassays. Therefore, polystyrene is very suitable as a coating for use in protein immobilization by physical adsorption onto glass surfaces.

4.4 Model biosensor application for the detection of nasopharyngeal carcinoma

We have previously developed a medical diagnostic kit based on an enzyme-linked immunoassay (ELISA) method for the early detection of nasopharyngeal carcinoma (NPC)[11]. NPC is also known as nose-and-throat cancer. The kit detects the serology (the presence of antibodies) to the Epstein-Barr Virus (EBV) in the blood of individuals with NPC. EBV is strongly implicated in the development of this disease. Early detection of specific EBV serology can save lives through the timely treatment of the disease.

Our interest in optical biosensors stems from the improvements in sensitivity and linearity of detection that fluorescence-linked immunoassays offer over enzyme-linked immunoassays, and the fact that such biosensor systems lend themselves readily to multichannel, multianalyte assay methods in a wash-free system. This is very important in the diagnostic applications like NPC serology since the immune response varies widely between individuals. Certain proteins from EBV elicit a strong antibody response and are key diagnostic antigens commonly used in ELISA and other methods[12]. However, NPC patients do not develop identical responses to all the antigens and this difference in immune response profile to the key antigens results in a certain percentage of false negatives when individual antigens are used in diagnostic tests. In our experience, it is possible to achieve far higher rate of diagnostic sensitivity (i.e. low false negatives) through the use of combined or mixed antigens, or through analysis of multichannel assays, where one blood sample is tested individually against two or more antigens. For example, a positive result in one or both antigens in a paired assay was taken to be a positive result.

Multichannel assays are inconvenient to perform using ELISA from a practical standpoint, and ELISA is only efficiently done using batches of blood samples. Testing of individual blood samples tends to be tedious and wasteful. The design of the biosensor under development (Figure 3) aims to simplify the testing for NPC through the simultaneous testing of multiple antigens against a single blood sample, in a one-step, wash-free system, that does not require laboratory-trained personnel to perform. A patterned biosensor containing an embedded waveguide utilizes fluorescence excited by the evanescent wave from total internal reflection of laser light. The fluorescence of positive wells can be differentially measured against negative controls to eliminate the need for a wash step.

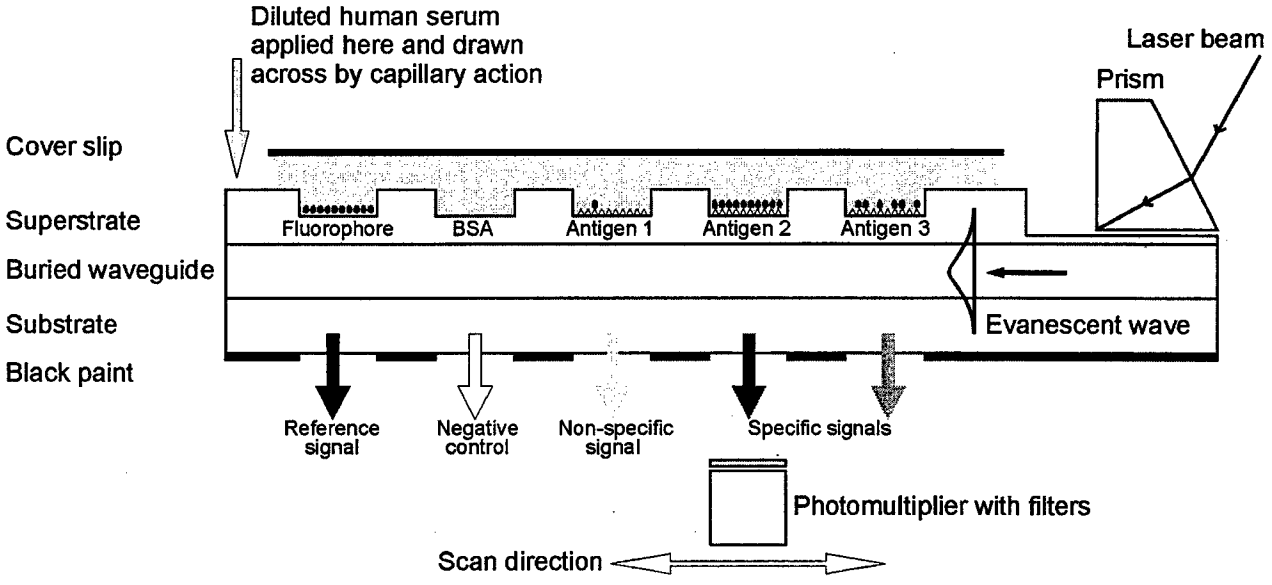


Figure 3: Schematic diagram showing the general features of an optical waveguide biosensor for fluorescence immunoassay applications. The sensor would necessarily incorporate positive (bound fluorophore) and negative (bovine serum albumin as blocking agent) controls and a single sample can easily be tested against multiple antigens. The presence of antibodies in human serum specific for the target antigens would result in a strong fluorescent signal that can be measured by the scanning photomultiplier. Differential measurements against the negative control signal allow a wash-free, single step application of sample.

We have been able to clone and produce many recombinant proteins from EBV through standard molecular biology techniques which have been successfully used as antigens in diagnostic immunoassays of NPC based on ELISA[11,13]. Several of these antigens are particularly useful and have very high diagnostic sensitivity and specificity in detecting NPC patients while minimizing false positives among normal individuals when used as single antigens. Analysis of the results from these existing assays strongly indicate that it would be possible to achieve nearly 100% detection rate should they be tested in a multichannel format. However, as with most immunoassays, the amount of antigens presented in the assay were carefully optimized and small variations in the amount of antigen dramatically lowers the sensitivity of the assay.

The application of a polystyrene thin film on the upper surface (superstrate) of the optical biosensor would greatly simplify the adsorption of antigens and control proteins to the wells, thus enabling the cheap and reproducible production of these units. The higher signal-to-noise ratio of physical adsorption methods over covalent-binding of proteins to the glass surfaces should greatly enhance the utility of this design and reduce the lower limit of detection to improve sensitivity.

5. CONCLUSION

We have demonstrated that proteins can be efficiently immobilized onto glass surfaces by first coating it with a thin polystyrene film. The performance of the polystyrene coated glass surfaces, in terms of the protein immobilization capacity and adsorption kinetics, has been found to be similar to that of pure polystyrene found in standard immunoassay microtiter wells, as well as a conventional method of protein immobilization onto glass using styrylsilane as surface treatment.

We therefore believe that this method of protein immobilization is very suitable for disposable optical biosensor applications because of its low cost, ease of manufacture and reproducibility. The similarity of polystyrene thin film on glass and pure polystyrene in terms of protein adsorption capacity and kinetics makes it easy to transfer existing immunoassay systems to glass-based biosensor applications. The results are applicable to optical fibers as well as planar waveguides because the physics (evanescent wave coupling in glass waveguides) and surface chemistry (coating and immobilization on glass) for both are the same.

ACKNOWLEDGEMENTS

This research was funded by the Academic Research Fund, Nanyang Technological University, Singapore, RG19/98 A/P Gan Yik Yuen.

REFERENCES

1. M. N. Kronick, and W. A. Little, "A new immunoassay based in fluorescence excitation by internal reflection spectroscopy", *J. Immunol. Methods*, **8**, pp. 235-242, 1975.
2. Y. Zhou, J. V. Magill, R. M. De La Rue, and P. J. R. Laybourn, "Evanescent fluorescence immunoassays performed with a disposable ion-exchanged patterned waveguide", *Sensors and Actuators B*, **11**, pp. 245-250, 1993.
3. P. J. R. Laybourn, Y. Zhou, R. M. De La Rue, W. Cushley, C. McSharry, and J. V. Magill, "An integrated optical immunosensor", *Uses of Immobilized Biological Compounds*, G. G. Guilbault and M. Mascini (eds.), Kulwer Academic Publishers, Netherlands, pp. 463-470, 1993.
4. M. P. Coughlan, M. P. Kierstan, P. M. Border, and A. P. E. Turner, "Analytical applications of immobilised proteins and cells", *J. Microbiol. Methods*, **8**, pp. 1-50, 1988.
5. I. Lundström, "Surface physics and biological phenomena", *Phys. Scrip.*, **T4**, pp. 5-13, 1983.
6. E. P. Plueddemann, *Silane coupling agents*, Plenum Press, New York, 1982.
7. M. Lynn, "Inorganic support intermediates: covalent coupling of enzymes on inorganic support", *Immobilised enzymes, antigens, antibodies and peptides*, edited by H. W. Weetall, Chapter 1, pp. 1-48, 1975.
8. S. K. Bhatia, L. C. Shriver-Lake, K. J. Prior, J. H. Georger, J. M. Calvert, R. Bredehorst, and F. S. Ligler, "Use of thio-terminal silanes and heterobifunctional crosslinkers for immobilization of antibodies on silica surfaces", *Anal. Biochem.*, **148**, pp. 408-413, 1989.
9. L. D. Krapivinskaya, G. B. Krapivinsky, and V. L. Ratner, "Immobilisation of immunoglobulin on a quartz surface by BrCN without loss of antigen binding capability", *Biosensors & Bioelectronics*, **7**, pp. 509-513, 1992.
10. V. Hlady, and J. Buijs, "Protein adsorption of solid surfaces", *Curr Opin Biotechnol*, **7**, pp. 72-77, 1996.

11. A. Fones-Tan, S. H. Chan, S. Y. Tsao, L. H. Gan, W. H. Tan, B. Li, P. W. Khong, and Y. Y. Gan, "Enzyme-linked immunosorbent assay (ELISA) for IgA and IgG antibodies to Epstein-Barr Virus ribonucleotide reductase in patients with nasopharyngeal carcinoma", *Int J Cancer*, **59**, pp. 1-4, 1994.
12. Y. Y. Gan, A. Fones-Tan, S. H. Chan and L. H. Gan, "Epstein-Barr viral antigens used in the diagnosis of nasopharyngeal carcinoma", *J. Biomed. Sci.*, **3**, pp. 159-169, 1996.
13. Y. Y. Gan, R. Hu, D. Chai, T. T. M. Tan, Y. H. Gan, S. H. Chan, S. Y. Tsao and L. H. Gan, "The distribution of Epstein Barr virus antigenic sites on the carboxyl terminal end of ribonucleotide reductase against nasopharyngeal carcinoma serum antibodies using immuno-absorption method", *J. Med. Virol.*, **58,** (In press), 1999.

LASER-BASED INSTRUMENTATION FOR MEDICAL DIAGNOSES AT VISIBLE WAVELENGTHS

BY

M.D. WATERWORTH *, B. TARTE, A.JOBLIN, T. VAN DOORN

QUEENSLAND UNIVERSITY OF TECHNOLOGY, AUSTRALIA.

- Currently at Avimo Electro-Optics Pte. Ltd. , 14, Fifth Lok Yang Road, Singapore 629763

ABSTRACT

The use of visible radiation in medical diagnostic procedures, initially attempted some 60 years ago, has only very recently become feasible, due to advances in source and detector technology, and the availability of elaborate computer-based imaging processing algorithms.

An instrument, operating at visible wavelengths, has been developed initially for the early detection of female breast disorders, and successfully tested on both phantom materials and on human subjects. Spatial resolutions of less than 100 microns are obtainable in the presence of severe scattering and intensity attenuation factors (of the order of 10^{-9}). Scan times of a few seconds only are required in the case of breast diagnoses, and the 3-D image is capable of reconstruction in near real time.

Application of this technology to other diagnostic procedures will be referred to briefly, and emphasis placed on the use of this non-ionising, non-invasive, completely safe, technique in medical fields.

KEYWORDS :

Visible wavelength laser instrumentation, safe diaphanography, breast cancer diagnoses, mammography,

INTRODUCTION :

Since the 1920's, the use of visible wavelength radiation has been postulated as a non-invasive soft tissue imaging modality. One application of such a technique is the examination of breast tissues for the early detection of carcinomas. As one of the major causes of early death in women, the detection of breast cancer , particularly in its pre-symptomatic stage, would be most beneficial.

Part of the SPIE Conference on Advanced Photonic Sensors and Applications
Singapore • November–December 1999
SPIE Vol. 3897 • 0277-786X/99/$10.00

In the late 1970's and early 1980's , infrared films and recording equipment permitted advances to be made, and as early ass 1980, commercial diaphanography units were available. Much of the current work uses light sources in the visible and infrared regions of the spectrum, from 500 to 1400 nm (Key et al 1991) for reasons of greater transmission than in adjacent spectral regions. Measurements by Key et al show that the scattering coefficients for specimens of normal and diseased breast tissue lie in the range 30 to 90 mm-1 at 500 nm., and decrease smoothly to between 10 and 50 mm^{-1} at 1100 nm.

Other workers (Ertafai et al, Ohlssen et al, Kaneko et al, Critchley et al) have demonstrated that very little light is transmitted in a straight path through breast tissue. Most of the photons are scattered extensively, thereby producing low resolution images in standard modes of operation. Figure 1 shows an image reconstructed tomographically from a simple object immersed in a scattering medium of 0.25% concentration of milk in water. Figure 2 is a schematic of visible wavelength propogation through a female breast, and Figure 3 shows the transmission through in vivo breast tissue, as measured by Key et al, and Waterworth et al.

Phantom materials, such as homogenised cows' milk, white paint, and Intralipid have been used to determine their propogation properties at a wavelength of 633 nm. This work is reported elsewhere (Waterworth et al), where the phantoms are compared with each other and with in vivo breast tissues. Table 1 summarises the results of these measurements, which assist in confirming the efficacy of laser diaphanography.

To illustrate the scattering processes occuring in tissues and phantoms, Figure 4 shows the scattering of a 633 nm. Laser beam through a 0.25% milk phantom. Figure 5 illustrates how the attenuation of this laser beam increases with milk concentration. Initially there is a Beer's Law relationship, but at higher concentrations the "effective" transmission is almost constant. A complete theoretical understanding of the physical processes involved here is the subject of on-going studies, and will result in a later publication.

EXPERIMENTAL CONCEPTS :

To improve the image quality upon reconstruction, it is necessary to reduce, or eliminate, the effects of scattering in tissues and phantoms. There remains then a residual attenuation factor of typically 10^{-8}, which can be managed by the use of suitable photon sources and detectors incorporating image intensification.

Time of flight measurements permit the selective detection of forward "scattered" photons, the angular acceptance angle being a function of detector time discrimination, and available photon levels within this time frame consistent with an acceptable signal to noise ratio. Coupled with these factors is the longitudonal spatial resolution requirements of an imaging system Δz, which is related to the time resolution t by

$$\Delta z = v \Delta t$$

where v is the velocity of wave propogation in the tissue/phantom medium. For the work reported here, a spatial resolution of 100 microns is required, resulting in a time resolution of approximately $4.4\ 10^{-13}$ seconds. Technology with this time resolution is currently readily available in the form of femtosecond pulsed lasers and corresponding detector resolutions, and is the ultimate goal of the commercialisation of the instrument under discussion here. However, due to financial constraints, slower systems with resultant lower resolutions are employed initially to verify the efficacy of the procedure.

EXPERIMENTAL ARRANGEMENT :

A schematic of the experimental apparatus used is shown in Figure 6. A Laser Photonics nitrogen dye laser, operating at 338 nm. Is used to pump the dye laser. Energy output from the dye laser is typically 250 microjoules in the wavelength range 500 to 960 nm., but is of course dye dependent. Pulse widths in the

output beam are typically 500 picoseconds, resulting in a minimum longitudonal spatial resolution of 11 cm. in a breast phantom.

The laser radiation passes through a controlled time delay system $\Delta t1$, employing fibre optics cables and retroreflectors, and thence into the breast phantom filled, in this case, with Intralipid solution. Approximately 1% of the incident beam is selected for further relative time delay $\Delta t2$ control to synchronise the detector "open time" with that of the laser pulse emission time and $\Delta t1$.

A Pockels Cell, controlled by a laser triggered spark gap, acts as an optical shutter for the system, the Pockels Cell having a rise time of 40 picoseconds. Suitable adjustment of $\Delta t1$, $\Delta t2$, and operation of the Pockels Cell, enables an effective sample time of near 40 picoseconds to be achieved. Due to the relatively large laser pulse width, the number of photons available in this short time interval is reduced well below that required to produce a high quality image, but nevertheless serves to illustrate the principles and to indicate problems which may occur with the faster technology when available. A second optical shutter will also be used later on the output side of the breast phantom.

Light transmitted through the phantom is collected by a CCD detector via an optical system, and the image of objects within the phantom displayed on the control computer screen. Provision is made for insertion of image intensifiers, and for the insertion of auto-correlation crystal technology which will enable time-averaging of signals to be performed, thus reducing the very stringent timing controls in the femtosecond system required in its absence.

To simulate tomographic imaging, test objects are rotated through prescribed angles within the phantom tank, and images recorded for each of these rotations. Sinograms are then produced, and the original object "reconstructed" using sophisticated image processing techniques developed at QUT.

Figure 7 shows the instrument with a 1 metre diameter phantomtank. Figure 8 consists of several image slices of a multi-bar object sited in the phantom tank, each slice corresponding to different parameters of focus, reconstruction parameters such as centre of rotation, etc., and illumination variations. It may be seen that, although spatial resolution is adequate, there are present in some cases artifacts which could be misleading vis a vis medical diagnoses. There are indications that the non-uniformity of incident illumination contributes significantly to these artifacts, a subject of further investigation.

With regard to the application of this technique to breast screening, an actual screening system is illustrated schematically in Figure 9. Essentially, it consists of a flying spot scanning on a rotatable axis, thus enabling the entire breast area to be scanned within a few seconds. Images of the breast will be produced in near real time, limited only by computer and software efficiency.

CONCLUSIONS :

Results obtained to date with this experimental arrangement confirm the efficacy of laser diaphanography for the early diagnosis of soft tissue abnormalities, and also confirm the requirements for a clinical instrument.

Future activities will involve the acquisition of a 'faster' laser referred to earlier in the paper, thus providing the possibility of 100 micron spatial resolution, along with a transmission differentiation of less than 1%. Achievement of these parameters will provide a significant advance in the early detection of soft tissue abnormalities, employing the non-ionising, completely safe, radiation from a laser at visible wavelengths. For the application of breast cancer diagnoses, it is further envisaged that there will be no need for compression of the breast tissue.

References

1. Cutler,M : Transillumination as an aid in the diagnosis of breast lesions, with special reference to its value in cases of bleeding nipples. Surg. Gynecol. Obstet. 48,729,1929.

2. Key,H., Davies E.R., Jackson,P.C, and Wells,P.N.T. : New approaches to transillumination imaging. Journal of Biomedical Engineering, 10, 113-118, 1988.

3. M.A. Duguay and A.T.Mattick, Ultrahigh Speed Photography of Picosecond Light Pulses and Echoes, Appl.Opt., Vol 10, No 9, 2163-2170, 1971.

4. Key, H., Davies, E.R., Jackson, P.C., and Wells, P.N.T. : Optical attenuation characteristics of breast tissues at visible and near infrared wavelength. Physics in Medicine and Biology, 36(5), 579-590,1991.

5. Key, H., Jackson, P.C., and Wells, P.N.T : Monte Carlo modelling of light propagation in breast tissue. Physics in Medicine and Biology, 36(5), 5910692, 1991.

6. Ertefai, S. and Profio, A.E. : Spectral transmittance and contrast in breast diaphanography. Medical Physics, 12(4), 393-400, 1985.

7. Ohlsson, B., Gundersen, J., Nilsson, D.M. : Diaphanography - A method for evaluating the female breast. : World Journal of Surgery, 4, pp 701-707, 1980

8. Flock, S.T, Wilson, B.C and Patterson, M.S. : Total attenuation coefficients and scattering phase function of tissues and phantom materials at 633nm. Medical Physics, 14(5), 835-841, 1990.

9. Peters, V.G., Wyman. D.R., Patterson, M.S., and Frank, G.L. : Optical properties of normal and diseased human breast tissue in the visible and near infrared. Physics in Medicine and Biology, 35(9), 1317-1334, 1990.

10. Arnfield, M.R., Matthew, R.P., Tulip, J., and McPhee, M.S., : Analysis of tissue optical coefficients using an approximate equation valid for comparable absorption and scattering. Physics in Medicine and Biology. 37(6), 1219-1230, 1992.

11. Lenz, P. : Spatial light distributions in tumours - Phantom measurements. : Medical Physics, 16(3), 326-332, 1989.

12. Profio, A.E. : Light transport in tissue. Applied optics, 28(12), 2216-2222, 1989.

13. Star, W.M., Marijnissen, J.P.A. and Van Gemert, M.J.C. : Light Dosimetry in optical phantoms and in tissues - 1 Multiple flux and transport theory. Physics in Medicine and Biology, 33(4), 437-454, 1988.

14. Madsen, S.J., Patterson, M.J., Wilson, B.C., Park, Y.D., Moulton, J.D., Jacques, S.L. and Hefetz, Y. : Time resolved diffuse reflectance and transmittance studies in tissue simulating phantoms - a comparison between theory and experiment. SPIE, Vol 1431, pp 42-51, 1991.

15. Cui W., Kumar C. and Chance B. : Experimental study of migration depth for the photons measured at sample surface. SPIE Vol 1431, pp 180-191, 1991.

16. Dilworth, D.S., Leith, E.N. and Lopez, J.L. : Three-dimensional confocal imaging of objects embedded within thick diffusing media. Applied Optics, Vol 30, No 14, pp 1796-1803, 1991.

17. Barbour, R.L. : Imaging in dense scattering media by diffusion tomography shows promise. Biomedical Optics, Nov 1992.

18. Watmough, D.J. : A light torch for the transillumination of female breast tissues. British Journal of Radiology. 55, 142-146, 1982.

19. Ohlsson B., Gundersen J. and Nilsson D. : Diaphanography - A method for evaluation of the female breast. World Journal of Surgery. 4, 701-707, 1980.

20. Jarry G., Debray S., Perez J., Lefebyre J.P., de Ficquelmont-Lpizos M. and Gaston A. : In vivo transillumination of the hand using near infrared laser pulses and differential spectroscopy. Journal of Biomedical Engineering, 11, 293-299, 1989.

21. Wang L., Liu Y., Ho P.P. and Alfano R.R. : Ballistic imaging of biomedical samples using picosecond optical kerr gates. SPIE Vol 1431, pp 97-101, 1991.

22. Berg R., Andersson-Engels S., Jarlman O. and Svanberg S. : Time-resolved transillumination for medical diagnostics. SPIE Vol 1431, pp 110-119, 1991.

23. Hebden J.C. : A time of flight breast imaging system - spatial resolution performance. SPIE Vol 1431, pp 225-231, 1991.

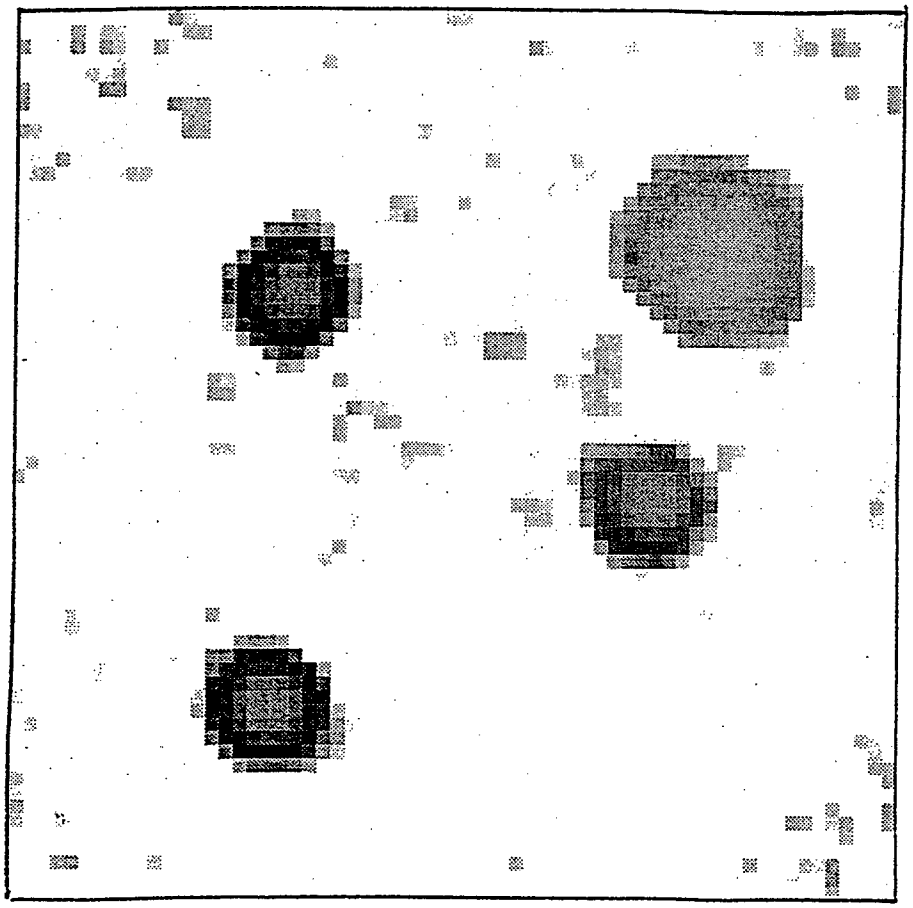


FIGURE 1 : FIRST IMAGE RECONSTRUCTION OF OBJECT IN 0.25% MILK CONCENTRATION PHANTOM, USING CW LASER RADIATION AT 633nm., SHOWING IMAGE BLURRING.

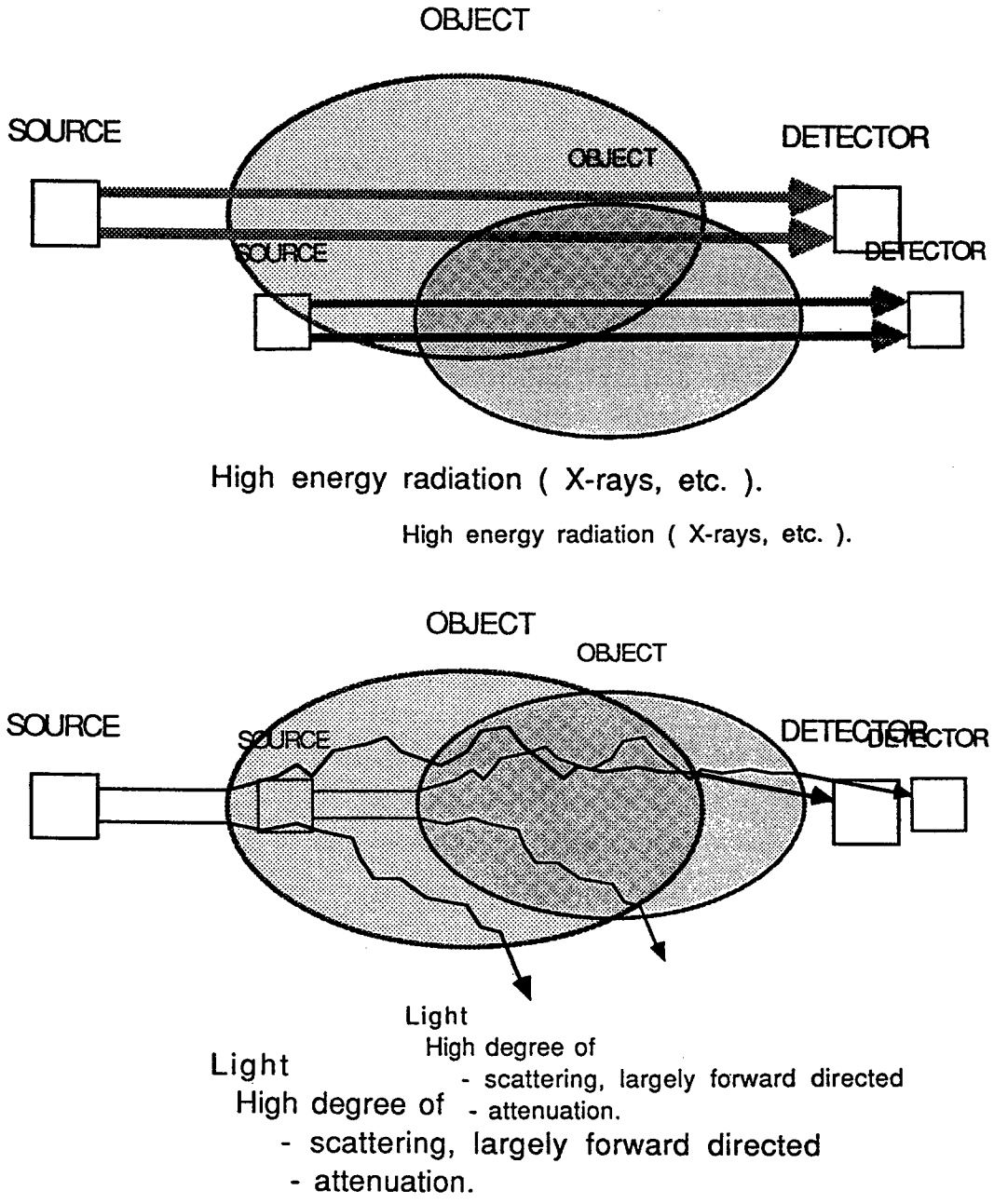


Figure 2 : Relative paths of X-radiation and visible radiation through soft tissue or highly scattering medium.

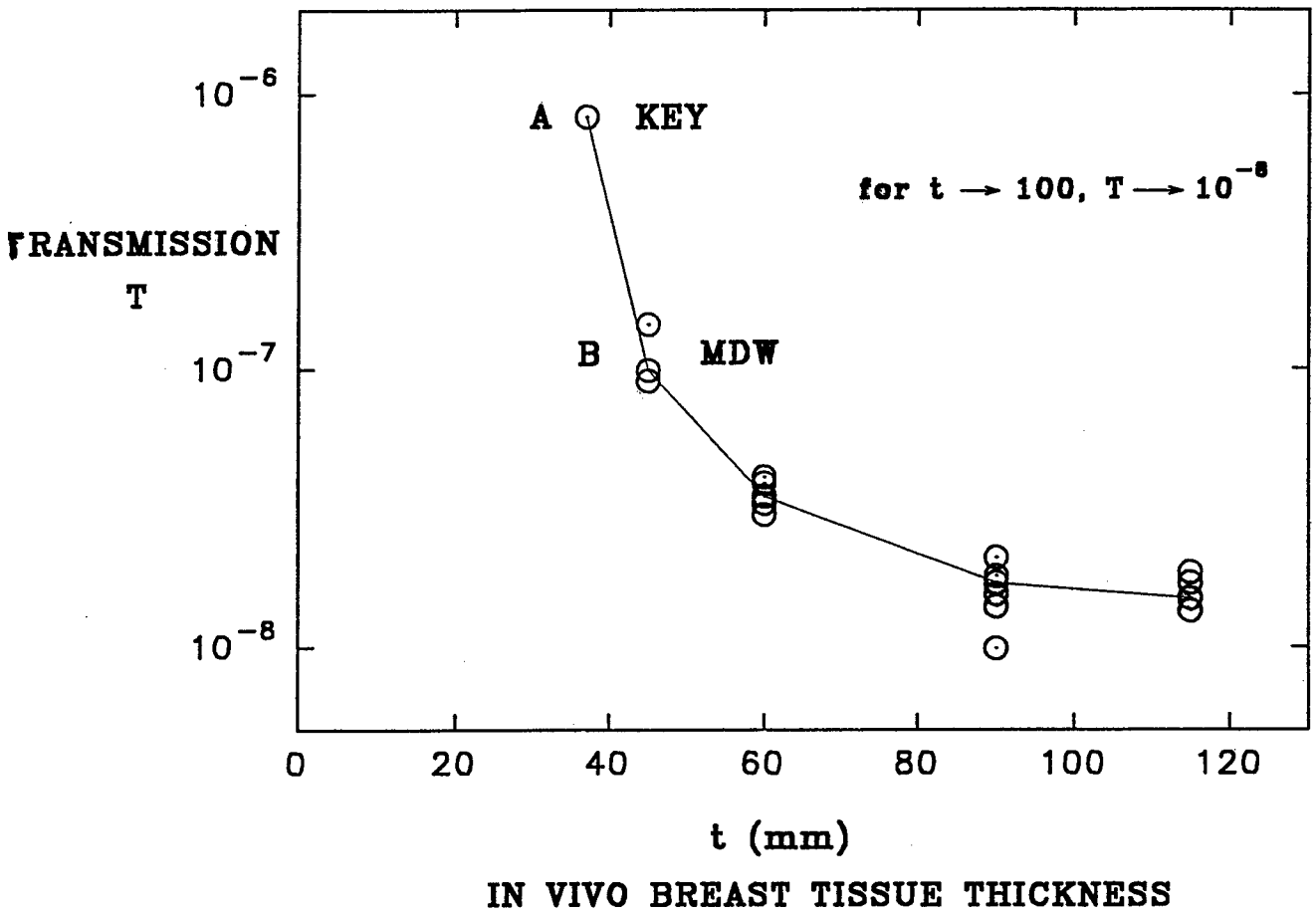


FIGURE 3 : MEASURED TRANSMISSION THROUGH IN VIVO BREAST TISSUES AT 633nm. BY KEY ET AL (1991) AND WATERWORTH ET AL (1993).

Table 1

Optical coefficients of phantoms and tissue

	μ_s mm^{-1}	μ_{eff} mm^{-1}	g	μ_a[a] mm^{-1}
pure milk	52±1	0.15±0.01	0.7[b]	4.8×10^{-4}
2% paint	44±1	0.27±0.02	0.7[c]	1.8×10^{-3}
8.5%I'pid	59±3	0.23±0.02	0.75[d]	1.2×10^{-3}
tissue[b]	10 - 100	1.20 (Mean)	0.8 - 0.95	0.01 - 1

(a) Calculated from values of μ_s, μ_{eff} and g using equation (2).

(b) Values taken from Wilson and Jacques 1990.

(c) By analogy with milk.

(d) Value taken from Flock et al 1992 for 10%-solids Intralipid.

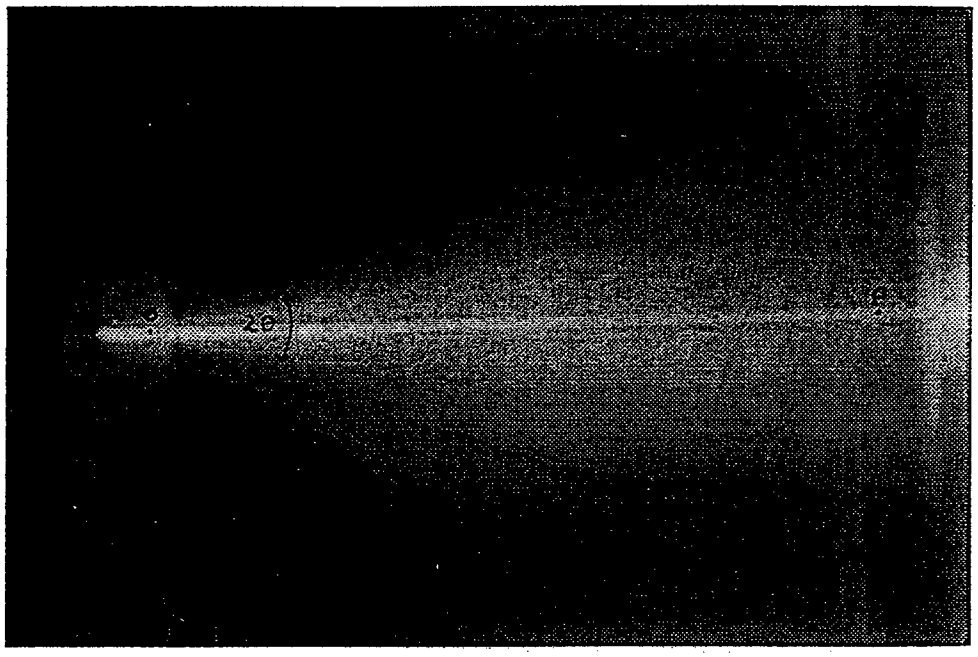


Figure 4 : Scattering and direct paths of 633 nm. Radiation through a 0.25% milk phantom.

Laser light intensity transfer function in 9.9mm thick tank as function of milk volume fraction.

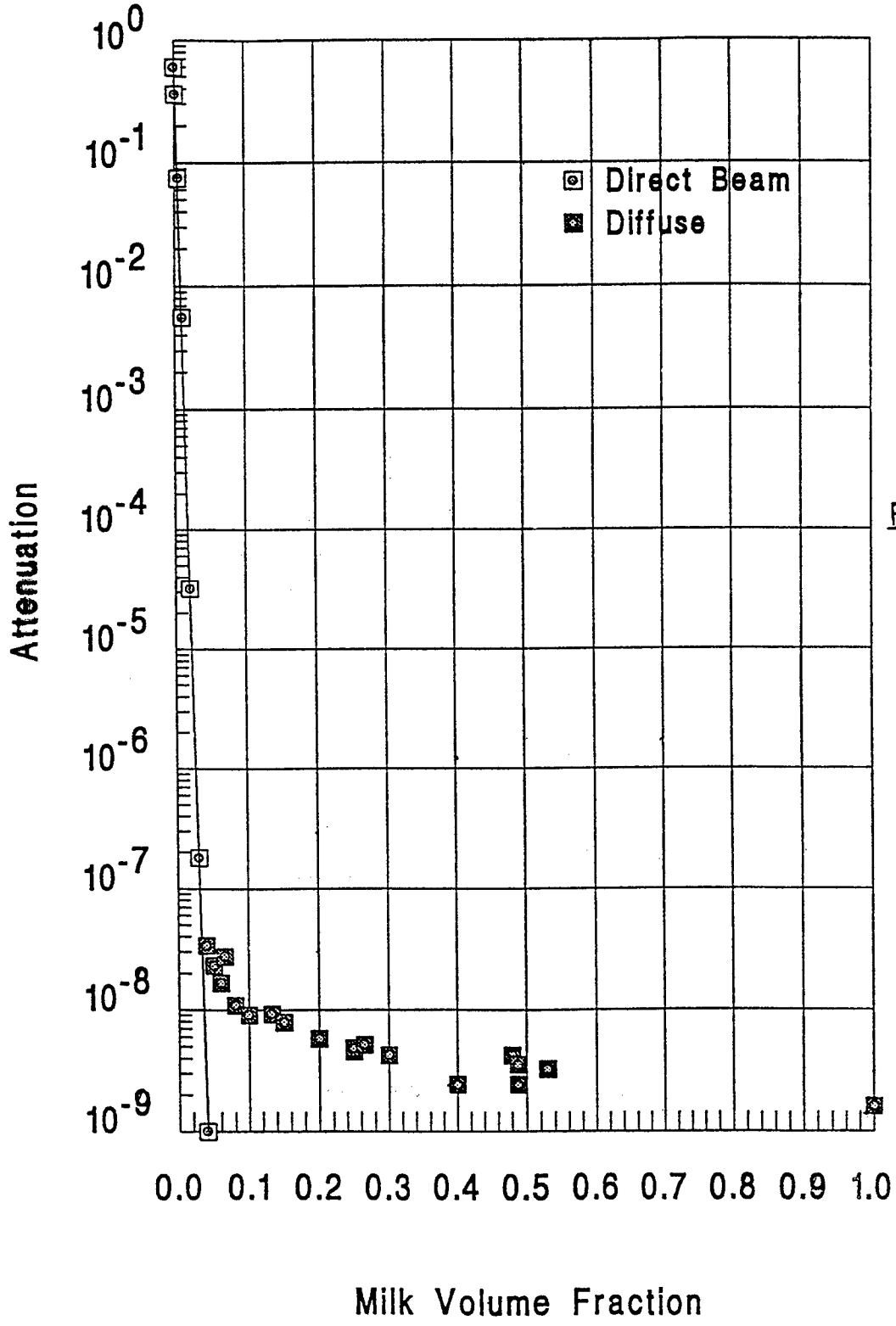
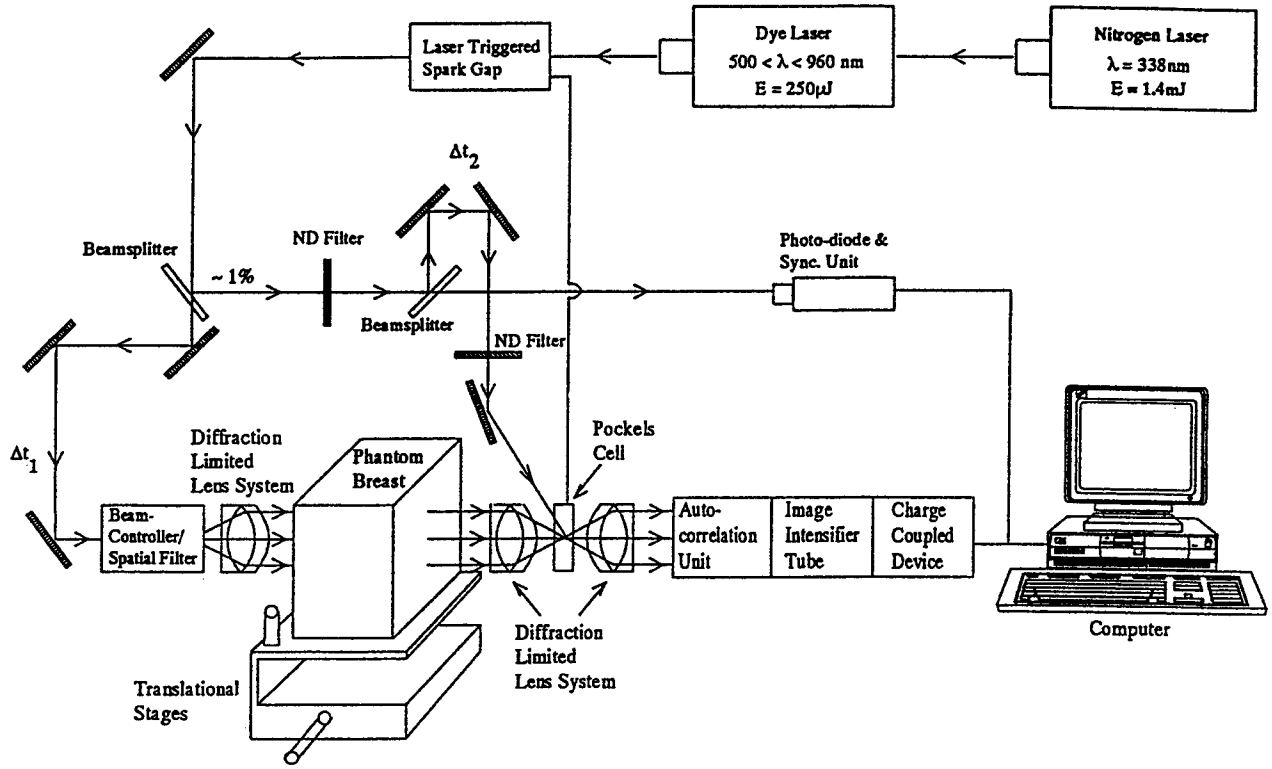


FIGURE 5



SCHEMATIC OF EXPERIMENTAL ARRANGEMENT FOR TIME OF FLIGHT IMAGING.

FIGURE 6

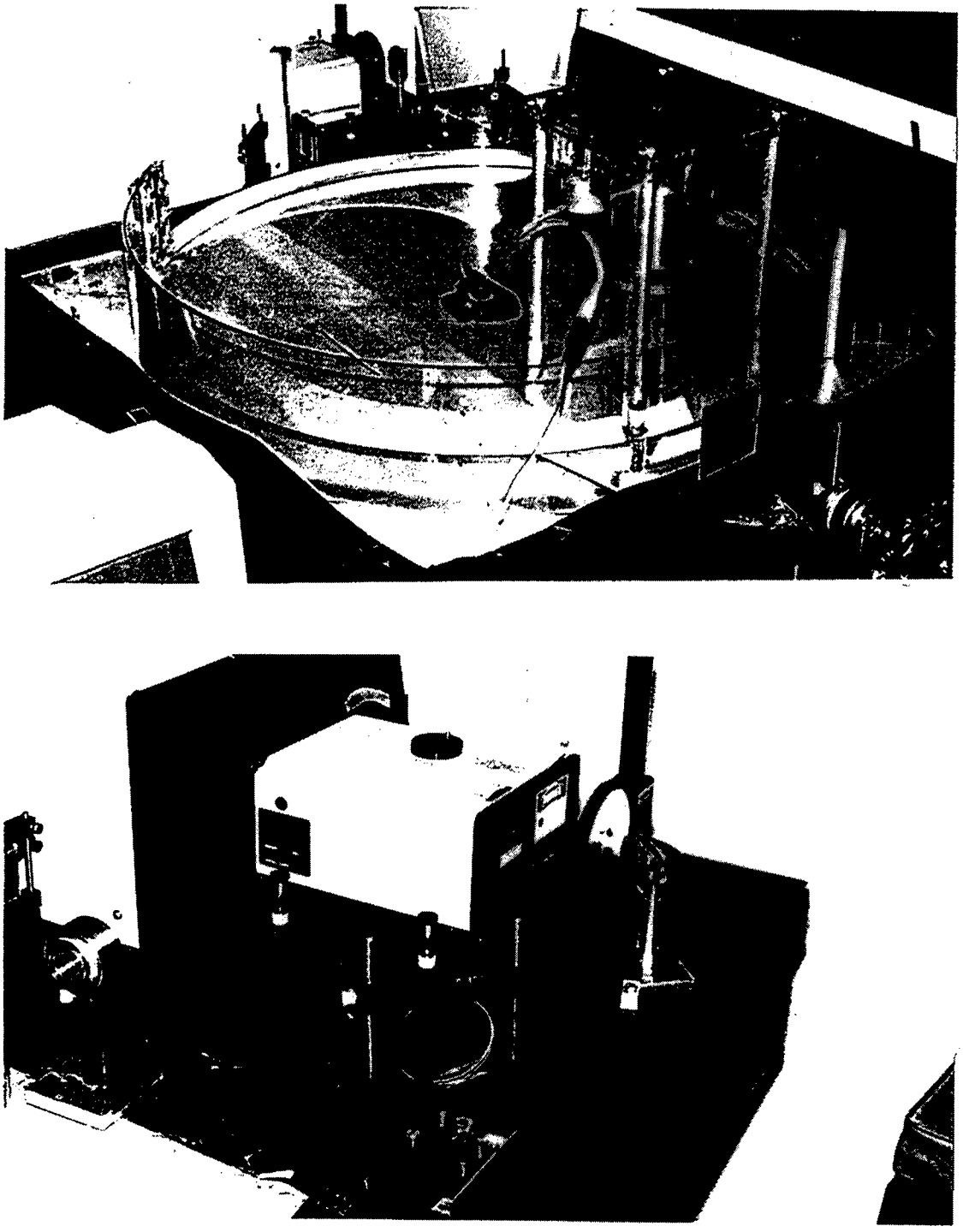


FIGURE 7 : PHOTOGRAPH OF EQUIPMENT USED WITH 1 METRE DIAMETER PHANTOM FOR TIME OF FLIGHT MEASUREMENTS.

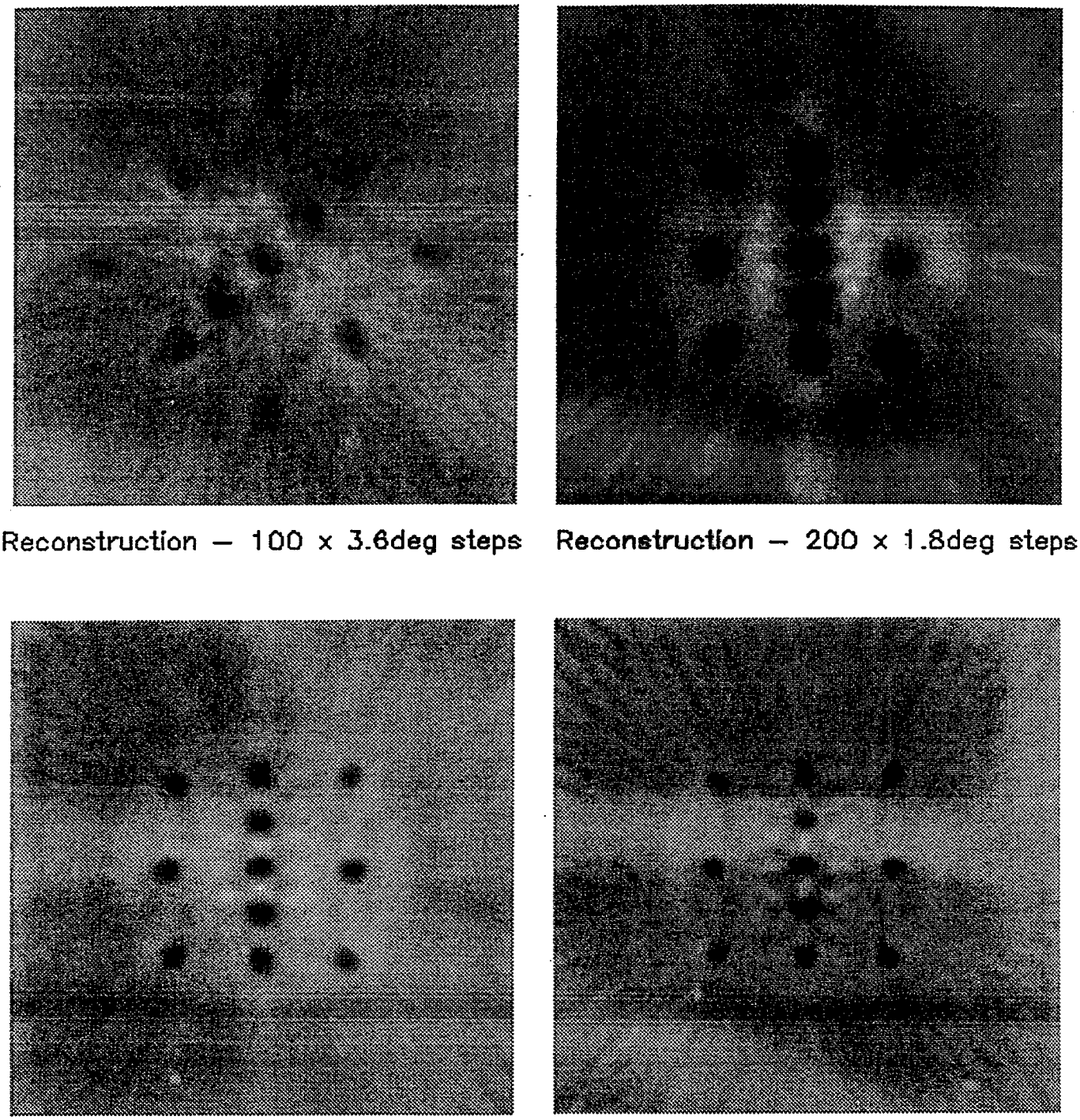


Figure 8 : Tomographic image reconstruction in various steps of sampling and focus.

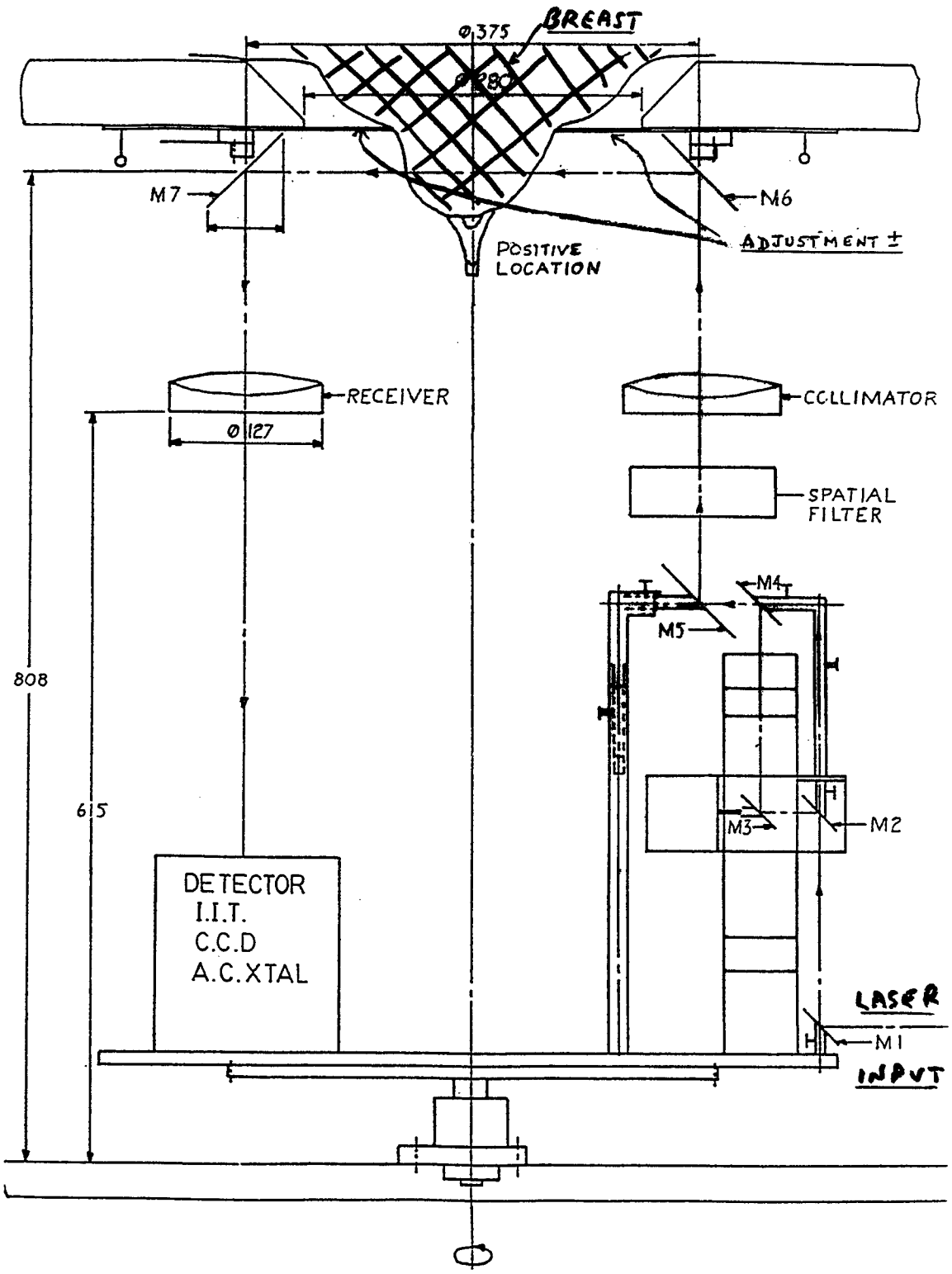


Figure 9 : Optical/mechanical layout of in vivo scanner of human female breasts.

NO_2 detection with a fibre optic evanescent wave sensor

M Shelly John, Jayan Thomas, K P Unnikrishnan, P Radhakrishnan*,
V P N Nampoori and C P G Vallabhan
International School of Photonics, Cochin University of Science and Technology,
Cochin 682 022, INDIA

ABSTRACT

A novel approach for the detection of nitrogen dioxide gas is described. This optical fibre based sensor (FOS) works on the principle of evanescent wave (EW) absorption phenomenon. EWs at the uncladded portion of a multimode fibre (200/380 μm) is utilized for the sensor development by replacing this region with a coating of Metallophthalocyanine (MPc), which is thermally deposited at a reduced pressure. MPcs are very sensitive to NO_2 gas and there is a change in the EW absorption in the NO_2 environment. Compared to other gas sensing devices, this is highly sensitive technique. The attraction of this FOS is its simple architecture and the easiness to implement.

Keywords: Fibre optic sensor, Evanescent wave, Nitrogen dioxide, Metalophthalocyanine

1. INTRODUCTION

The development of low loss optical fibres have revolutionized the field of telecommunication as well as the development of a variety of sensors. Immunity to electromagnetic interference, possibility of using them for remote measurements, multiplexing capability etc. made fibre optic sensors superior to other electronic sensors. Fibre optic sensors (FOS) have a wide range of applications and they are involved in a broad spectrum of fields like chemical sensing, smart structure applications, environmental pollution monitoring etc.[1,2,3] FOS is an effective tool in the determination and measurement of different physical parameters. A variety of FOSs are based on the evanescent wave absorption phenomenon and this is a thrust area where active research is going on. [4,5]

Phthalocyanines have a good sensitivity to the toxic gas NO_2 and these have added advantages like good chemical and thermal stability, that enhances its potential to be used in gas sensing systems. Researchers found metal substituted phthalocyanines as a very good material in the research and development activities to sense toxic as well as pollutant gases.[6] The molecular structure of the Metallophthalocyanine (MPc) is as shown in figure 1. Certain MPcs have shown good response to NO_2 in terms of the variation in the adsorption, electrical and optical behavior.[7,8,9] Exploiting the affinity of MPcs to NO_2, a variety of techniques have evolved to detect traces of NO_2 in the environment. Investigations were carried out in the field of NO_2 detection with different MPcs like lead phthalocyanine (PbPc) and copper phthalocyanine (CuPc) etc. by different groups. One of the techniques to detect NO_2 is by using Quartz Crystal Microbalance (QCM). This is achieved by forming Langmuir – Blogett (LB) film of MPc on a QCM oscillating at a particular frequency.[10] Sensing principle of QCM is due to the mass loading by adsorption of NO_2 on the MPc surface. Nitrogen dioxide, one of the electron acceptor gases forms donor – acceptor complex with the MPc. Mass increase due to NO_2 adsorption is manifested as a frequency change of QCM. The change in oscillation frequency is a measure of nitrogen dioxide gas concentration. Another technique to detect nitrogen dioxide is Surface Plasmon Resonance (SPR).[11,12] Very thin layers of Metallophthalocyanine are used in this technique for the detection of NO_2 which is formed on a thin metal layer. In this type, the changes in the optical properties of the active layer in response to external ambients are measured by monitoring the coupling of photons to surface plasmons at the interface between a metal and the active layer. Extensive investigations were carried out in the conductivity measurement of MPc in the presence of NO_2. MPc are p-type organic semiconductors and they show high sensitivity even to very low concentrations of electron – acceptor gases like NO_2. Hence NO_2 adsorption on the surface of film of MPcs is followed by charge transfer reaction which induces generation of charge carriers and enhancement of electrical conductivity.[9]

* Email : photonix@vsnl.com

The present paper discusses a new method to detect nitrogen dioxide in a gaseous sample. Nitrogen dioxide gas can be detected using an evanescent wave fibre optic sensor (EWFOS) by exploiting the evanescent waves (EW) at the uncladded portion of a multimode fibre having core diameter 200 μm and replacing that region with MPcs that are very sensitive to NO_2.

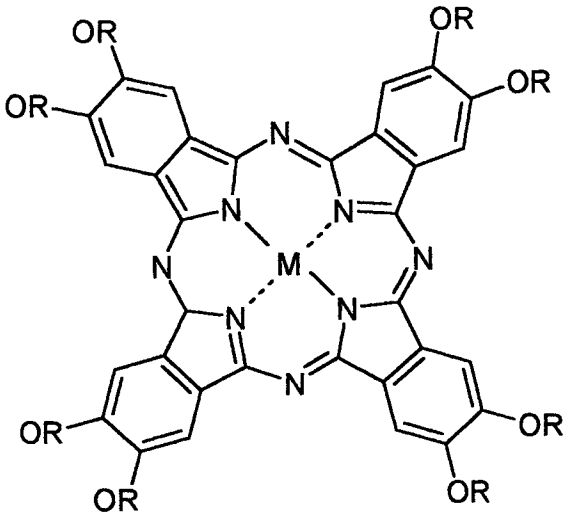


Figure 1 Molecular structure of Metallophthalocyanine M = Cu, Pb

2. EXPERIMENTAL SET-UP

A length of 0.04 m of the cladding of the multimode fibre is replaced with a coating of metallophthalocyanine (MPc) thin film which is produced by thermal evaporation at a reduced pressure of 10^{-5} mbar. This portion of the fibre acts as the sensor element of the FOS system. The structure of the sensor element is as shown in figure2. The MPc is uniformly coated to a thickness of 140 nm and this act as the sensing element.

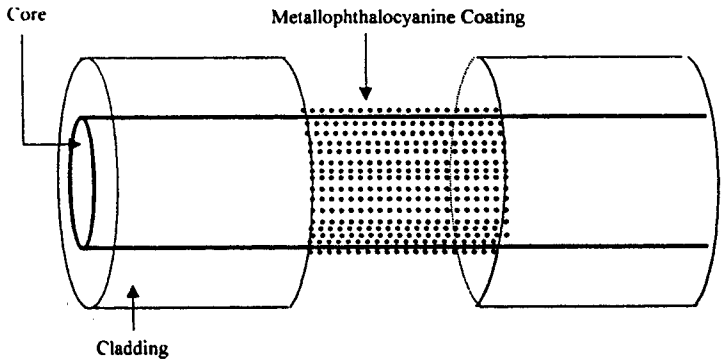


Figure 2 The sensor element

The schematic of the experimental set-up is as shown in figure 3. Light from a diode laser emitting at 670 nm is focused to one end of the fibre and the guided wave is detected using a fibre optic power meter (Meggar OTP

510) at the other end of the multimode fibre. The sensor element is placed in a gas chamber that can be evacuated to lower pressure. NO_2 gas flows through the chamber through the provisions as indicated in the figure. The investigations were conducted using the lead phthalocyanine (PbPc) in the sensor region as well as copper phthalocyanine (CuPc), at different NO_2 gas concentrations.

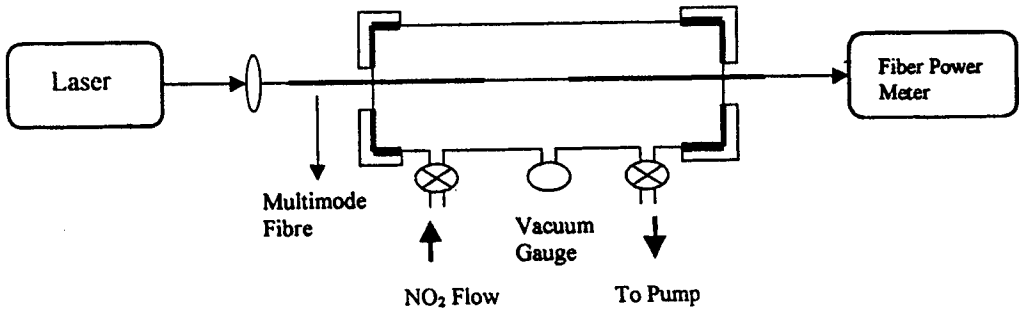


Figure 3 The schematic of the experimental set-up

3. RESULTS AND DISCUSSIONS

3.1 FOS with CuPc sensing element:

The rate of flow of NO_2 gas through the sensor chamber is kept constant. As soon as the NO_2 molecules reaches the sensor element then get adsorbed on the surface of the MPc. The absorption spectra of CuPc thin film given in figure 4 exhibits its peak at 697 nm.

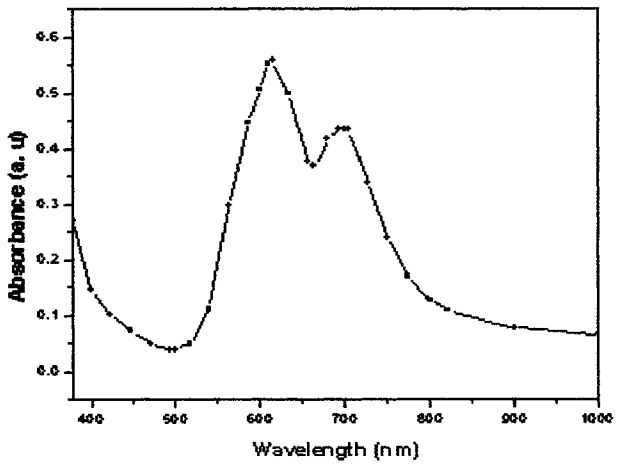


Figure 4 Absorption spectrum of copper phthalocyanine thin film

Adsorption of NO_2 molecules on CuPc leads to a shift in the absorption peak towards 670 nm, the propagating wavelength through the fibre and hence a variation in the evanescent absorbance. The power transmitted by an optical fibre, the cladding of which is replaced locally by an absorbing medium may be given as[13]

$$P(z)=P(0)\exp(-\gamma z) \quad \text{.....(1)}$$

where z is the length of the uncladded region of the fibre, P(0) is the power transmitted in the absence of an absorbing species and γ is the evanescent absorption coefficient.

The evanescent absorbance A of an uncladded fibre of length L surrounded by a medium is given by

$$A=\log\frac{P(0)}{P(z)}=\frac{\gamma L}{2.303}=\frac{r\alpha L}{2.303} \quad \text{.....(2)}$$

where α is the bulk absorption coefficient of the medium, when all bound modes are launched to the fibre. 'r' is the fraction of the power outside the core and is given by

$$r=\frac{4\sqrt{2}}{3V} \quad \text{.....(3)}$$

where V is the normalized frequency parameter of the fibre.

At the interface of the SiO_2 core and CuPc cladding in the sensor region , attenuated total internal reflections (ATR) take place. The variation of the output power with time for this sensor is given in figure 5. The two plots show the sensor response to two different nitrogen dioxide concentrations.This is atrributed to the blue shift in the absorption peak of the CuPc film on NO_2 adsorption. A similar blue shift has also been reported in Langmuir - Blodget CuPc film.[8]

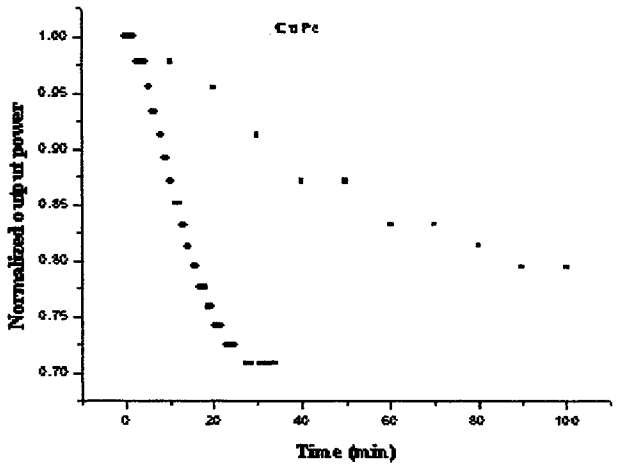


Figure 5 Variation of the output power with time for CuPc coated FOS at two different concentrations of nitogen dioxide

3.2 FOS with PbPc sensing element:

The variation of the output power of the light propagating in the FOS system with time with PbPc on the sensor element is as shown in figure 6. The plot shows the variation of the output light intensity with time at a particular concentration of nitrogen dioxide. It is observed that the variation is small at low concentrations of NO_2.

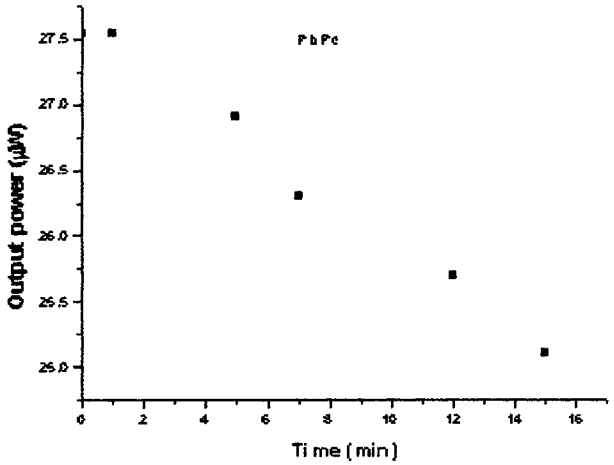


Figure 6 Variation of the output power with time for PbPc coated FOS

Using the QCM technique, S R Kim et al have studied the effect of central metal atom in metal phthalocyanine in sensing the NO_2 molecules[8]. They have established that CuPc is more sensitive than PbPc. Our observations using the EW technique confirms these results.

4. CONCLUSION

We have discussed a novel method to detect nitrogen dioxide quantitatively. The reusability of the sensor is under investigation. This evanescent wave fibre optic sensor seems to be promising for a wide range of applications. This is an effective tool for the remote sensing purpose.

ACKNOWLEDGEMENTS

MSJ acknowledges Dr. Sakthikumar, Binoy Paul, Pramod Gopinath and Narayanan Unni for their immense help given during this work. MSJ also acknowledges UGC (New Delhi) for the research fellowship. KPU and JT acknowledge CSIR (New Delhi) for their research fellowship.

REFERENCES

1. Mark A Arnold, "Fibre optic chemical sensors", Anal. Chem, **64**, pp.1015A - 1025A, 1992.
2. Murukeshan V M, Chan P Y, Asundi A, Ong Lin Sang, "Optimization of various parameters on the embedded polarimetric sensor for smart structure applications", *Optics and Optoelectronics, theory, devices and Applications*, O P Nijhawan, A K Gupta, A K Musla and Kehar Singh, **2**, pp. 1161-1166, Narosa Publishing house, New Delhi, 1998.
3. Kathleen M Leonard, "Development of a fibre optic chemical sensor for multicontaminant monitoring of environmental systems", Sensors and Actuators B, **24-25**, pp. 458-461, 1995.
4. M Shelly John, P Radhakrishnan, V P N Nampoori and C P G Vallabhan, " A fibre optic evanescent wave sensor for monitoring the rate of pulsed laser deposition of metal thin films", Meas. Sci. Technol., **10**, pp. N17-N20, 1999.
5. Deepa Jose, M Shelly John, P Radhakrishnan, V P N Nampoori and C P G Vallabhan, "An optical fibre based evanescent wave sensor for monitor the deposition rate of thin film", Thin Solid Films, **325**, pp. 264-267, 1998.

6. M E Azim-Araghi and A Krier, "The influence of ammonia, chlorine and nitrogen dioxide on chloro-aluminium phthalocyanine thin films", Applied Surface Science, **119**, pp. 260-266, 1997.
7. D Campbell, R A Collins, " A study of the interaction between nitrogen dioxide and lead phthalocyanine using electrical conduction and optical absorption", Thin Solid Films, **295**, pp. 277-282, 1997.
8. Seung Ryeol kim, Jong Duk Kim, Kyung H Choi, Young H Chang, "NO_2 sensing properties of octa (2 ethyhexyloxy) metallophthalocyanine L B films using quartz crystal micro balance", Sensors and Actuators B, **40**, pp. 39-45, 1997.
9. Shu Ji Qin, " The sensitivity to NO_2 of sandwich devices based on lead phthalocyanine and copper phthalocyanine ", Sensors and Actuators B, **3**, pp. 255-260, 1991.
10. S R Kim, S A Choi, J D Kim, K H Choi, S K Purk, Y H Chang, " The characteristcs of Metallophthalocyanine Mono-, multi layer and application to the gas sensor for NO_2", Synthetic Metals, **71**, pp. 2293 - 2294, 1995.
11. P S Vukusic, J R Sambles, " Cobalt phthalocyanine as a basis offor the optical sensing of nitrogen dioxide using surface plasmon resonance", Thin Solid Films, **221**, pp. 311-317, 1992.
12. D G Zhu, M C Petty and M Harris, " An optical sensor for nitrogen dioxide based on a copper phthalocyanine Langmuir - Blodgett film" Sensors and Actuators B, **2**, pp. 265-269, 1990.
13. P Radhakrishnan, V P N Nampoori and C P G Vallabhan,"Fibre optic sensor based on evanescent wave absorption", Opt. Engg., **32**, pp. 692-694, 1993.

SESSION 4

Biomedical and Biochemical Sensors II

Theoretical and experimental study on the optimum working wavelength for a fiber-optic sensor based on the SPR used in monitoring the water quality

Deming Liu*, Yajian Huang, Dexiu Huang

Dept. of Optoelectronic Engineering, Huazhong Univ. of Sci. & Tech., P. R. China

ABSTRACT

The optimum working wavelength of a fiber-optic sensor based on the surface plasmon resonance (SPR) used in monitoring the water quality is studied theoretically and experimentally in this paper. This sensing system was composed of a broad-spectrum source, a fiber sensor head and a micro-optic spectrograph. The sensing fiber adopted was a silica/polymer multi-mode optic fiber, with the 400 um core diameter and 0.36 number aperture. The calculation results showed that there is an optimum working wavelength range corresponding to the given sensor head configuration. Through adjusting the structure parameters the sensing wavelength could be selected to drop into the wavelength range of the commercial semiconductor optical sources, which is beneficial to improving the practicality and the reliability of the sensing system. This system was used to test the impurity of the fresh water and a sensitivity of 8.9×10^{-5},corresponding to 10 ppm, was reached.

Key words: fiber-optic sensor, water quality monitoring, surface plasmon resonance, optimum wavelength

1. INTRODUCTION

One of the major challenges of the world environment science is how to ensure people to be able to monitor the water quality effectively in the fields of industry, agriculture and daily life. In that way the water pollution can be controlled. There are several conventional ways using electrochemical technology to monitor the water quality. The electrochemical sensors are well established but suffer from electromagnetic interference (EMI). And the sensing systems are in general relatively large and expensive, and are inapplicable for remote-sensing applications. It is also impossible to realize real-time monitoring by these systems since all the conventional ways are based on sample analysis. Fiber-optic water-quality sensors offer however immunity to EMI and have been explored using a variety of detection schemes, including water-quality-induced changes in fluorescence, absorption and refractive index. In the first two cases, the linearity of the device response tends to be poor, while in the latter case, the index change is rather small, in the order of 1×10^{-3}. Although index changes on this scale can be measured interferometrically, an alternative, and perhaps more robust, approach is to use a technology based on a surface plasmon resonance (SPR) sensor, which can make the sensor more compact, cheap and applicable for remote sensing. With the help of a computer program, it could be possible to develop a sensing system for insitu monitoring the water quality.

* Correspondence: Email: dmliu@mail.hust.edu.cn; Telephone/Fax: 0086-27-87543355

Part of the SPIE Conference on Advanced Photonic Sensors and Applications
Singapore • November–December 1999
SPIE Vol. 3897 • 0277-786X/99/$10.00

2. THEORETICAL CALCULATION MODEL FOR SPR

Surface plasmon resonance is an optical phenomenon in which incident light excites a charge-density wave at the interface between a metal and dielectric. It has been intensively studied as a powerful technique to monitor the water and air pollution[1-4], in which at certain conditions the evanescent wave in the optical waveguide could excite (or couple with) the surface plasma wave (SPW) in the metal, propagating along the surface of the metal film, as shown in Fig.1. The propagation constant of the SPW can be described by the following equation,

$$K_{SPW} = K_0\sqrt{\frac{\varepsilon_m\varepsilon_s}{\varepsilon_m+\varepsilon_s}} = K_0\frac{n_m n_s}{\sqrt{n_m^{\,2}+n_s^{\,2}}}. \tag{1}$$

The conditions for plasmon resonance by the evanescent wave are satisfied when the propagation constant of the evanescent wave along the surface agrees with the propagation constant of the surface plasma wave. It can be showed by the following equation,

$$K_x = \frac{2\pi n_0}{\lambda}\sin\theta = K_{SPW} = K_0\sqrt{\frac{\varepsilon_m\varepsilon_s}{\varepsilon_m+\varepsilon_s}}. \tag{2}$$

The propagation constant of the evanescent wave varies with the incident angle and the incident light wavelength. As a result of the excitation of SPW, the reflected light intensity would be lost. By detecting the reflected light spectrum the refractive index of the sensed medium could therefore be determined. This mechanism could be used to detect the concentration of different species in water.

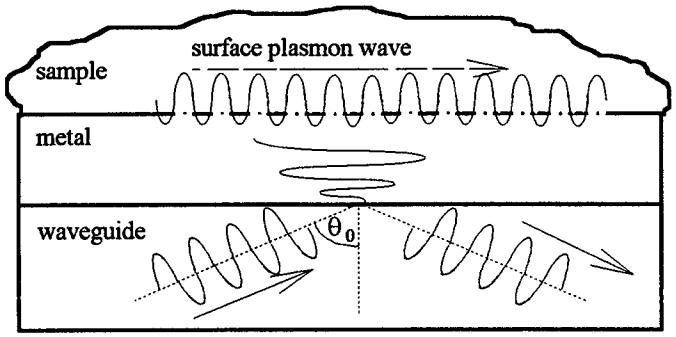


Fig.1 SPW in an optical waveguide

There are two methods for SPR sensing technology: the method of fixed wavelengths and modulated angle of incidence, the method of fixed angle of incidence and modulated wavelength. The normalized intensity of a light beam reflected on a metal surface of the thickness d with an incidence angle θ and a wavelength λ is described by the degree of reflection $R[\theta,d,n_f(\lambda), n_m(\lambda), n_s(\lambda)]$; whereas ,$n_f(\lambda)$ is the refractive index of the optical waveguide layer, $n_s(\lambda)$ is that of the substrate, $n_m(\lambda)$ is the complex refractive index of the metal.

In our experiment the fiber was selected as the sensing media. In this case the method of fixed angle of incidence and modulated wavelength is selected since the wavelength intensity distribution may be preserved in an optic fiber. And the angular intensity distribution of light will be indistinguishable due to mode mixing as a result of the inherent bending of the multimode fiber in practical sensing applications. Furthermore, since the sensor is fabricated on a multimode fiber, there is

not a fixed angle of incidence, but rather a range of incident angles that are allowed to propagate in the fiber. Figure 2 shows a typical fiber-optic SPR chemical sensor. Because of total internal reflection the evanescent field exists outside the fiber core. The fiber adopted is a silica/polymer multimode optic fiber, which has an aperture of 0.36. The core diameter is 400 μm, and the sensing length of the fiber is 10 mm. The buffer layer is MgF_2 film, which is 430nm thick.

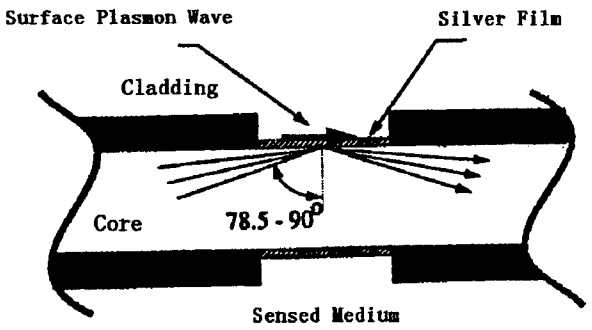


Fig.2 Illustration of the SPR fiber-optic sensing element

Due to the large diameter of the used fiber in contrast to the wavelength of the light, the propagation of the light within the sensing fiber can be described by common ray optical methods. The normalized intensity of a light beam reflected on a silver surface of the thickness d with an incidence angle θ and a wavelength λ is described by the degree of reflection $R(\theta, d, n_f(\lambda), n_m(\lambda), n_s(\lambda))$; whereas, $n_f(\lambda)$ is the refractive index of the fiber core, $n_s(\lambda)$ is that of the surrounding media, $n_m(\lambda)$ and is the complex refractive index of the metal. If the metal surface is coated with a thin adlayer, e.g. a polymer film, its thickness d_p and its refractive index $n_p(\lambda)$ have also to be considered on the degree of reflection which can be derived by a matrix method.

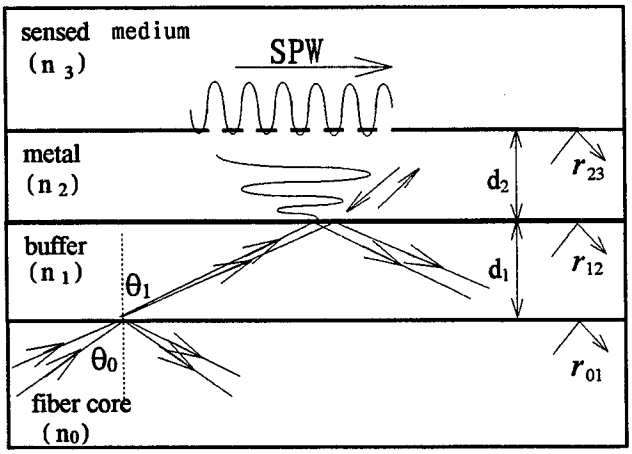


Fig.3 Theoretical calculation configuration

The theoretical calculation configuration is showed by figure 3, where n_0 is the refractive index of the fiber core, n_1 is the refractive index of the buffer layer, n_2 is the metal layer, n_3 is the refractive index of the sensed medium. In this configuration, the Snell theorem is expressed by,

$$n_0 \sin\theta_0 = n_1 \sin\theta_1 = n_2 \sin\theta_2 . \qquad (3)$$

By the equation above $\cos\theta_1$,$\cos\theta_2$,$\cos\theta_3$ can be derived,

$$\cos\theta_1 = \sqrt{1-(\frac{n_0 \sin\theta_0}{n_1})^2}, \tag{4}$$

$$\cos\theta_2 = \sqrt{1-(\frac{n_0 \sin\theta_0}{n_2})^2}, \tag{5}$$

$$\cos\theta_3 = \sqrt{1-(\frac{n_0 \sin\theta_0}{n_3})^2}. \tag{6}$$

The reflective coefficient from fiber core to buffer layer r_{01}, the buffer layer to metal layer r_{12}, and metal layer to sensed medium r_{23} can be deduced by the Fresnell theorem,

$$r_{01} = \frac{n_0 \cos\theta_1 - n_1 \cos\theta_0}{n_0 \cos\theta_1 + n_1 \cos\theta_0}, \tag{7}$$

$$r_{12} = \frac{n_1 \cos\theta_2 - n_2 \cos\theta_1}{n_1 \cos\theta_2 + n_2 \cos\theta_1}, \tag{8}$$

$$r_{23} = \frac{n_2 \cos\theta_3 - n_3 \cos\theta_2}{n_2 \cos\theta_3 + n_3 \cos\theta_2}. \tag{9}$$

The optical phrase thickness of the buffer layer and the metal layer are expressed by

$$\delta_1 = \frac{4\pi}{\lambda} n_1 d_1 \cos\theta_1, \tag{10}$$

$$\delta_2 = \frac{4\pi}{\lambda} n_1 d_1 \cos\theta_1. \tag{11}$$

The reflective coefficient from buffer layer to the sensed medium r_{13} can be deduced from equation (5), (6), (8),

$$r_{13} = \frac{r_{12} + r_{23} e^{j\delta_1}}{1 + r_{12} \times r_{23} e^{j\delta_1}}. \tag{12}$$

And the reflectance coefficient of the P mode incident light r_{03} can be deduced from equation (4), (7), (9),

$$r_{03} = \frac{r_{01} + r_{13} e^{j\delta_2}}{1 + r_{01} \times r_{13} e^{j\delta_2}}. \tag{13}$$

The reflectivity for a single reflection in the sensed fiber is expressed by,

$$R_\theta = r_{03} \times r_{03}^*. \tag{14}$$

The number of reflection in the fiber sensor area is a function of the mode propagation angle θ_0, as well as of the diameter of the fiber, d_f, and the length of sensing area L. This relationship is given by,

$$N_\theta = \frac{L}{d_f tg(\theta)}, \tag{15}$$

Based on this consideration the total degree of reflection R_{tot} is expressed by,

$$R_{tot} = (R_\theta)^{N_\theta}. \tag{16}$$

Within a multimode fiber the light beams propagate under different angles, The smallest angle is determined by the

numerical aperture (NA) and the largest is 90^0. Due to the large diameter of the used fiber in contrast to the wavelength of the light, the intensity distribution dependent on the incident angle can be approximated as equally distribution. The degree of reflection weighted by this distribution and added up are used for a theoretical determination,

$$R = \sum_{\theta=78.5}^{90}\{(R_\theta)^{N_\theta} \bullet \omega(\theta) \bullet \Delta\theta\} = \sum_{\theta=78.5}^{90}\{(R_\theta)^{N_\theta} \bullet \frac{1}{90-78.5} \bullet \Delta\theta\} \cdot \tag{17}$$

Thus it can be seen from the above equations, for a given sensor, R_{sen} depend only on the wavelength λ_{SPR} and the refractive index of the sensed medium n_s whereas other parameters are constant. If λ_{SPR} is decided under the condition that the reflectivity is minimum, the refractive index of tested medium can be determined also. It can be expressed by a mathematical method,

$$\frac{\partial R(n_s, \lambda)}{\partial \lambda} = 0 \cdot \tag{18}$$

3. RESULTS AND ANALYSES

The calculation is based on the above multiple-reflection theory. First, the refractive angles in different mediums could be obtained from the equations. Then the amplitude reflectance r*ij* at the boundary between the incident light from medium *i* to *j* could be calculated from the equations. The intensity reflectance R is shown as the function of the incident angle, the incident light wavelength, the thickness of metal layer, and the refractive index of different mediums.

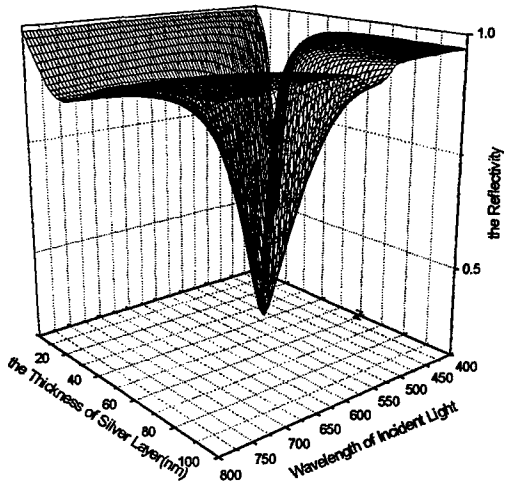


Fig.4 Curve of the reflectance vs the silver thickness and the wavelength

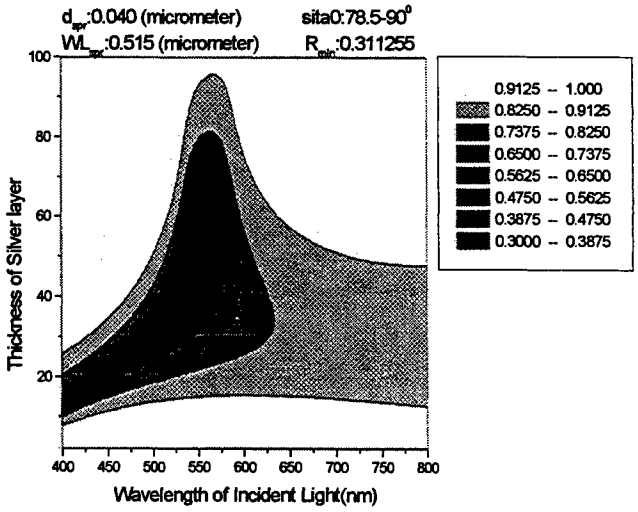


Fig.5 Gray-grade graph of the reflectance vs the silver thickness and the wavelength

The simplest SPR sensor is of a three-layer configuration, i.e. the substrate, the sensitive metal-layer (silver here) and the sample (layer). Figure 4 shows the three-dimension map of the intensity reflectivity versus the incident light wavelength and the thickness of the silver layer. The X-axis represents the incident light wavelength. The Y-axis represents the thickness of the silver layer. Figure 5 is the corresponding gray-grade map. The different gray-grade represents different rang of reflectivity value. This map shows that an absorbency peak is obtained at λ=514nm and d=38nm.It is apparent from the figures that the criterion for the thickness and the working wavelength is much rigorous in this case. And also, the optimum

working wavelength is not suitable for the transmission through the optical fiber. The calculation analyses show that a four-layer could change the excitation condition of SPR in a great extent.

In our experiment an appropriate buffer layer is sandwiched between the metal layer and the waveguide (core of the optic fiber here). It is possible to adjust the index of the buffer layer so that strong coupling between a waveguide mode and the SPW occurs. The calculation for this structure is as same as the sample structure without buffer layer.

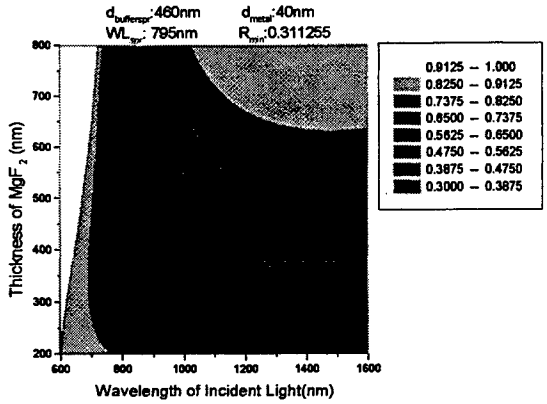


Fig.6 Gray-grade graph of the reflectance vs the sodium fluoride thickness and the wavelength

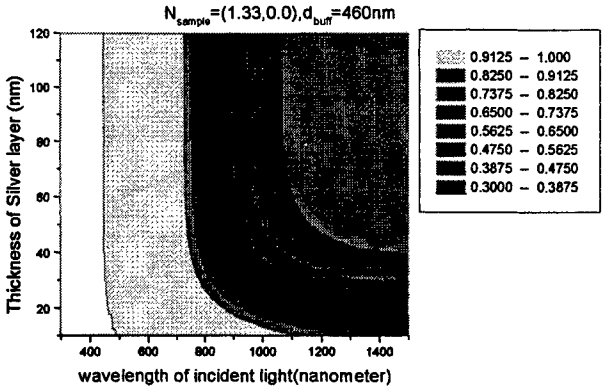


Fig.7 Gray-grade graph of the reflectance vs the silver thickness and the wavelength

Figure 6 shows the gray-grade graph of the reflectance vs the thickness of the sodium fluoride (MgF_2) and the working wavelength. Figure 6 shows the gray-grade graph of the reflectance vs the silver thickness and the working wavelength with the MgF_2 buffer layer. It is apparent from the figures that the criterion for the thickness and the working wavelength is much loose in this case. And also, the optimum-working wavelength has been moved to 850nm, one of the three widows for the transmission through the optical fiber. Through adjusting the index and the thickness of the buffer layer, the working wavelength could be also moved to 1310nm so as to be suitable for the long-haul transmission.

4. EXPERIMENTAL SETUP AND MEASUREMENTS

Fig.8 shows a block diagram of the experimental system based on a fiber-optic setup. The output light of a lamp is focused into an optic fiber. A mode scrambler is used to weed out all the higher modes of the optic fiber. The SPR sensing surface is enclosed by a flow cell constructed using a syringe with two syringe stoppers and inlet and outlet ports. The output of the fiber-optic sensor is connected to a fiber-optic spectrogram. The optic fiber chosen for simulation is a 2 oxidized silicon (SiO2)/polymer fiber which has a numerical aperture of 0.36. The silver layer is evaporated on the fiber core after eroding away the polymer cladding. The light resource is a wide band spectrum white light lamp.

Fig.9 is the measurement curve of the SPR using the above experimental setup. At the wavelength 1.38μm, a clear absorbent peek occurs. This working wavelength is a lit longer than what we predicted. It is however easy to be adjusted. The change range of the reference index in our experiment is 0.04, the corresponding wavelength range is about 45nm. Then the sensitivity for the detection of the reference index is 8.9×10^{-5} if using a spectrogram of 0.1nm resolution.

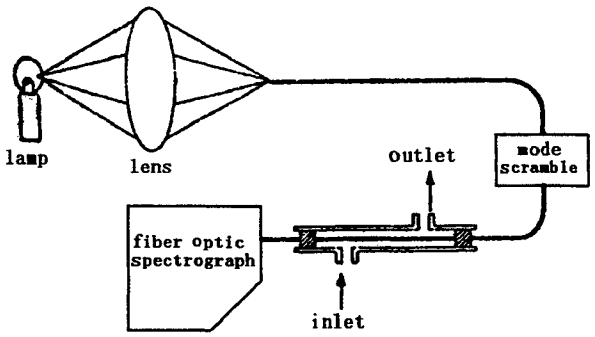


Fig.8 The experimental setup diagram

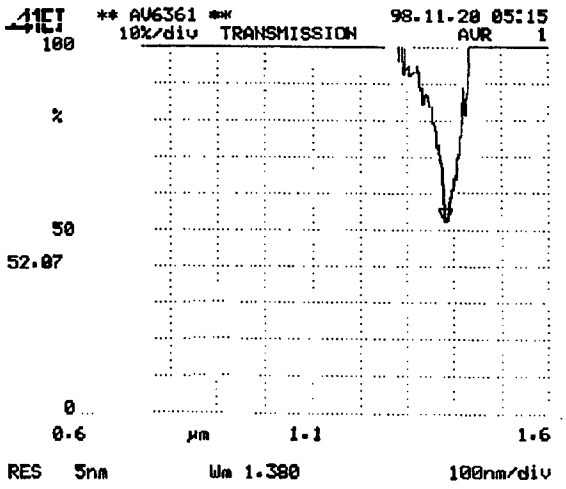


Fig.9 The measurement curve of the SPR

4. CONCLUSION

A fiber-optic sensing system based on the surface plasmon resonance (SPR) and designed for insitu monitoring the water quality was reported in this paper. A theoretical model for calculating the coupling efficiency between the fiber guide mode and the SPR mode was presented. The calculation results showed that there is an optimum working wavelength range corresponding to the given sensor head configuration. Through adjusting the structure parameters the sensing wavelength could be selected to drop into the wavelength range of the commercial transimission optical fiber, which is beneficial to improving the practicality and the reliability of the sensing system. This system was used to test the impurity of the fresh water and a sensitivity of 8.9×10^{-5}, corresponding to 10 ppm, was reached. The further work will result in an insitu environment-monitoring instruments based on SPR technique, which will find extensive application in the future.

ACKNOWLEDGEMENTS

Authors should be greatly appreciated for the support from the Natural Science Foundation of Hubei province, P.R.China.

REFERENCES

1. C.R.Lavers, J.S.Wilkinsor, a waveguide-coupled surface plasmon sensor for an aqueous environment, *Sensors and Actuators,* **B22,** pp.75-81,1994.
2. J.Wilkinson, Waveguide sensor detects trace levels of herbicidein water, *OLE,* pp.39, 1995.
3. Jose Melendez, Sinclair Yee, A commercial solution for surface plasmon sensing, *Sensors and Actuators,* **B35-36,** pp.212-216,1996.
4. Jose Melendez, Sinclair Yee, Development of a surface plasmon resonance sensor for commercial applications , *Sensors and Actuators,* **B38-39,** pp.375-379, 1997.

Fiber Optic Raman Spectroscopy an application for the In-situ measurement of water content in Methanol

Anand K Asundi., Vasudeva Rao Veeredhi.

School of Mechanical and Production Engineering
Nanyang Technological University
Republic of Singapore - 639798

ABSTRACT

This paper describes the results of the experimental investigation conducted to determine the percentage of water by volume in methanol. The suitability of a Fiber optic Raman spectrometer for in-situ measurement of percentage of water by volume in methanol is demonstrated. The measured intensities at the characteristic Raman shift corresponding to Methanol are found to be inversely proportional to the volume of water present in Methanol. The results obtained in the form of Raman spectra are presented in graphical form and average intensities in tabular form.

1. INTRODUCTION

During the last couple of decades it has been realized by many researchers that the conventional fossile fuels are fast depleting and hence there is a need for identifying alternate sources of fuels particularly a replacement for fuels used in transportation. Even though the idea of using methanol as a motor fuel is as old as the motor car itself it has not been considered seriously until late 80's (1, 2, 3, 4). Now it is well appreciated that the use of methanol as a motor fuel with some design modification in the IC engine fuel handling system and combustion chamber. Unlike gasoline, methanol can dissolve unlimited amounts of water (1). Alcohols in general are hygroscopic (attract and absorb water), and while this would enable methanol fires to be extinguished with water, the fuel requires careful dry handling. This undesirable property of methanol (in the present context) hinders the direct utilization of methanol as a transportation fuel. Therefore the measurement of purity of methanol in terms of dissolved water is an important criterion before it is used as a fuel. In other words the measurement of dryness of methanol is very important. The use of optical fuel sensors to determine the alcohol content in a mixture of methanol and gasoline is mentioned in the literature (6). But the use of fiber optic Raman spectroscopy to determine the percentage of water in methanol has not been reported till date. In the present investigation it is demonstrated that the suitability of fiber optic Raman spectroscopy for the in-situ measurement of water content in methanol.

2. METHOGOLOGY AND MEASUREMENT PRINCIPLE

The basic measurements involved in Fiber optic Raman spectroscopy are measurement of intensities of scattered radiation, from the material under investigation, as a function of wavelength or frequency. It is also possible to represent the measured intensity as a function of Raman shift. To accomplish this, the test sample is excited with a laser beam with approximately 785 nano-meters wavelength and the scattered radiation is detected with an array of detectors. The intensity of scattered radiation when represented as a function of Raman shift gives the signature of each material unique to that particular material. The details of the measurement system and the experimental procedure are given in the following section.

3. THE MEASUREMENT SYSTEM

The measurement system used in the present investigation basically consists of a diode laser source that can deliver 500mW power when operated at 785nm wavelength and a spectrometer with a linear array of silicon CCD detectors. The excitation

Part of the SPIE Conference on Advanced Photonic Sensors and Applications
Singapore • November–December 1999
SPIE Vol. 3897 • 0277-786X/99/$10.00

laser light is passed up to the test sample through a fiber optic cable. The reflected radiation from the test sample is captured and transmitted back to the spectrometer. The spectrometer shown in Fig.1 is interfaced to a personal computer through an ADC-500 card and an interface cable. Prior to using, the measurement system has been calibrated and tested for its accuracy and reliability by acquiring Raman spectra of Ethanol and compared with the literature. The Raman spectra obtained for Ethanol is shown in Fig.2 is in good agreement with the published literature and establishes the suitability of the fiber optic portable Raman spectrometer. The present investigation is focused on generating Raman spectra of Methanol with varying percentage of water.

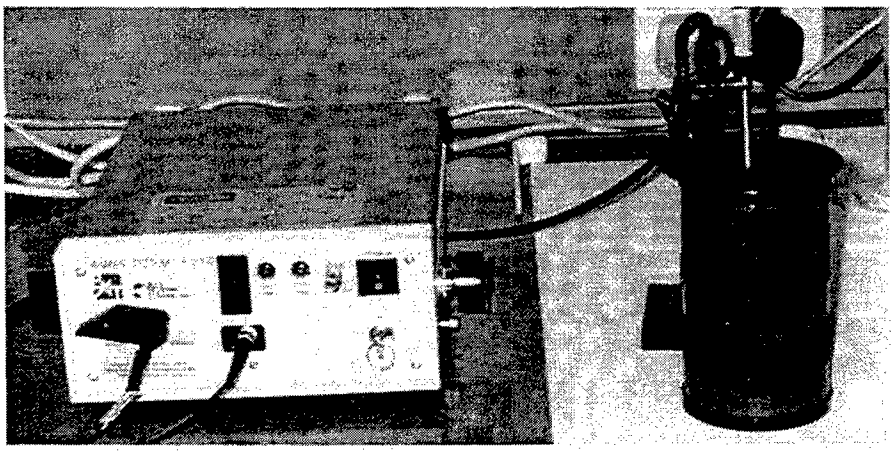


Fig.1 The Measurement System

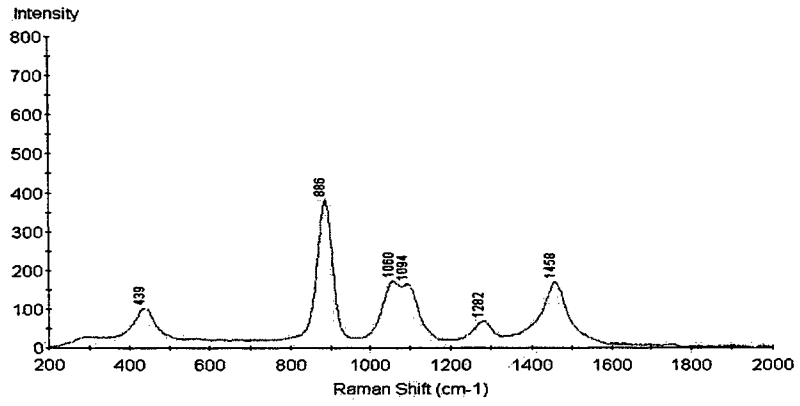


Fig.2 Raman Spectrum of Ethanol

4. EXPERIMENTAL INVESTIGATION

4.1 Test sample preparation:

The calibration test samples are prepared by mixing measured quantities of methanol and water. To prepare the samples, commercially available methanol (with purity better than 99.8%) is taken in to a glass beaker with 25ml pipette and tap water is mixed to it with known proportions. The test samples prepared with various concentrations are stored separately in

glass bottles. In the present investigation methanol with 0, 4.76, 8.77, 16.13, 27.77, 36.59, 43.48, 49.02, 53.57, 57.37, 60.60, 67.33, 70.27, 72.72, 74.81, percentage of water by volume are prepared for calibration purpose.

4.2 Test Procedure

The liquid test samples are taken in to a well sterilized 100 ml amber bottle with a rubber septum. The bottle is filled to half of its capacity with the sample under test and closed with the rubber septum. The fiber optic probe tip is inserted into the bottle through the hole in the septum until the probe tip slightly protrudes into the liquid test sample. To minimize the reflections of the laser light, from the side walls and the bottom, back into the probe's collection fibers, the probe is oriented such that it is approximately perpendicular to the bottom of the bottle. To minimize any possible noise, due to the background light in the room, the amber bottle is kept in a dark colored enclosure as shown in Fig.1.

Before initializing the spectrometer, an integration time of 500 msec and a laser wave length of 785 nm is selected in wave length mode. Once the above mentioned parameters are set and all the optical safety precautions are taken the laser is switched on. After 20 seconds the laser pulse, which is basically a Rayleigh line, will appear on the screen. To identify the wave length at which the laser exerts maximum power the cursor is moved to the position of maximum intensity. As the cursor moves, the wavelength and intensity corresponding to its position is displayed at the bottom of the screen. Once the cursor is centered on the Rayleigh line, laser wavelength displayed at the bottom of the screen is noted and entered in the wave length window along the top of the screen. Also the acquisition mode is switched from wavelength (nm) to Raman Shift (cm^{-1}).

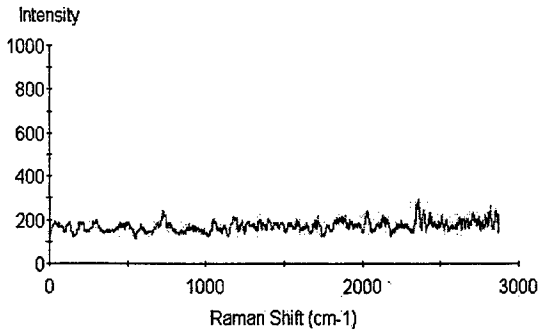


Fig. 3 Typical dark current.

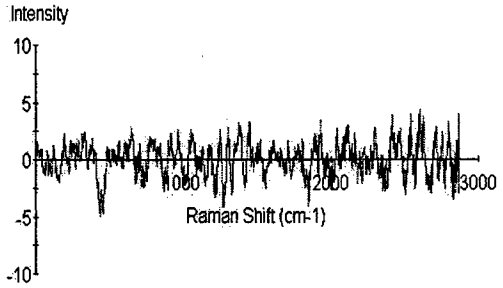


Fig. 4 Residual charge on CCD

Before recording Raman spectra, the dark current has been collected and stored. A typical spectra of a dark current is shown in Fig.3. This is carried out after turning the laser off and setting the integration time to 10,000 msec. The dark current appears on the screen after 20 seconds and it is stored and subtracted from the signals stored in the CCD at the end of each integration time (i.e. in the present case it is 10 seconds). The subtraction of the dark current makes the spectrum on the screen to near zero and is shown in Fig. 4. To get the actual Raman spectrum of the sample under test, the laser is switched on and the spectrum acquired. Since the laser has a built in 20 seconds delay, it is very unlikely that the start of the integration cycle synchronizes exactly with the moment of the system switching on. Therefore, the spectrum that appears on the screen within the first integration time corresponds to a partial change of the CCD array. The spectrum that appears on the screen at the end of the second integration cycle represents the Raman spectrum of the sample under analysis. The same procedure is repeated for all the calibration test samples and the Raman spectra acquire and stored on a hard disk for further processing.

5. RESULTS AND DISCUSSION

As mentioned earlier in section 4, validation test is conducted using ethanol by acquiring its Raman spectrum and is presented in Fig.2. Hamaguchi and Hirakawa [1988](7) acquired Raman spectrum of ethanol using Ar-ion laser with wavelength 488.0 nm at 100mW power output is in very good agreement with the present one. The comparison of the magnitude of intensity was not possible because, the data are not available from Hamaguchi and Hirakawa's work(7). Even if such data are available, it is not possible to compare the intensities since it is a function of integration time. However, the Raman shift, which is a key characteristic for a given material is in good agreement with the present results. This comparison establishes the reliability of the measurement system used in the present investigation. Few Raman spectra of methanol with varying percentage of water obtained during the course of the present experimental investigation are given in the Fig.5. The three spectra shown in Fig.5 correspond to 49, 57, 74% of water in methanol-water mixture. It is observed that the intensity of Raman signals decreases as the percentage of water in methanol increases. The intensity of the signals also fluctuates slightly between one integration cycle to the next integration cycle. To overcome this problem and reduce the error, ten readings of intensities corresponding to ten consecutive integration cycles are taken and averaged to represent the mean as a best value. The mean intensities thus obtained are plotted against the corresponding percentage of water in methanol in Fig. 6. It is observed that the variation of intensity with %of water in methanol can be represented with a linear curve arrived by the method of least squares.

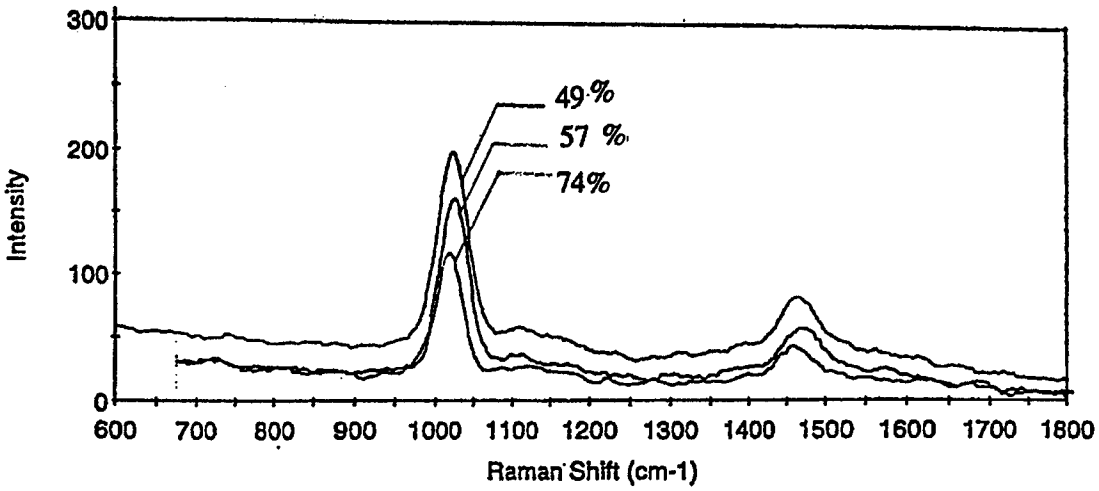


Fig. 5. Raman Spectra of Methanol and Water (% shown) mixture with various proportions

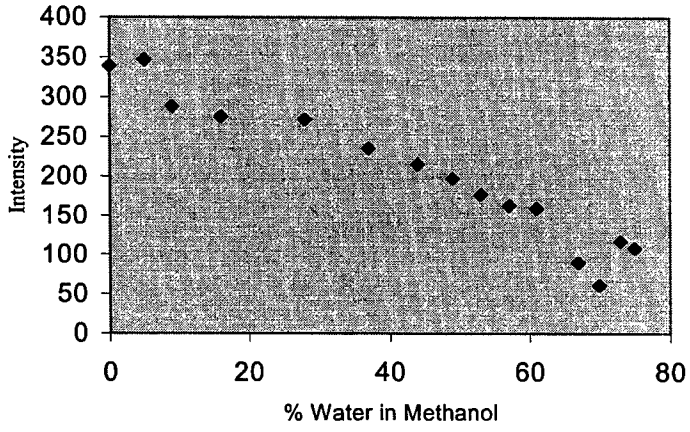


Fig. 6 Variation of Intensity as function of percentage of Water in Methanol

Table-1 shows the numerical date of intensity and the percentage of water present in methanol. In the light of the present investigation it is possible to in-situ evaluation of the purity of methanol using fiber optic Raman spectroscopy. It is observed that the variation of intensity with %of water in methanol can be represented with a linear curve arrived by the method of least squares. Table-1 shows the numerical date of intensity and the percentage of water present in methanol. In the light of the present investigation it is possible to in-situ evaluation of the purity of methanol using fiber optic Raman spectroscopy.

Table-1.

Sl.No:	Vol. Percentage of water in Methanol	Raman Shift (cm^{-1})	Maximum Intensity
1	0	1036	337.6623
2	4.76	1036	346.6844
3	8.77	1036	288.2521
4	16.13	1025	275.5289
5	27.77	1025	271.9645
6	36.59	1025	235.5377
7	43.48	1025	215.3867
8	49.02	1025	197.3555
9	53.57	1025	176.609
10	57.37	1025	162.7513
11	60.60	1025	159.1112
12	67.33	1020	89.89767
13	70.27	1020	61.17787
14	72.72	1020	117.0623
15	74.81	1018	108.2664

6. CONCLUSIONS

Raman Spectrum diagnostic tests conducted in the present investigation establishes the suitability of the fiber optic portable Raman Spectrometer for field and in-situ tests. Even though the spectrometer is portable, spectra obtained are on par with the spectra obtained with big and bulky table mount spectrometers. The measured intensities at the characteristic Raman shift corresponding to Methanol are found to be inversely proportional to the volume of water present in Methanol. The Raman-spectrum diagnostics can be conveniently applied to identification and detection of the water content in methanol.

7. SCOPE FOR FUTURE WORK

The possibilities for further application of this measurement system are as numerous as the number of alcohols that are amenable to Raman spectroscopic analysis. The measurement system used in the present experimental investigation is also

suitable to measure and quantify the methanol content in gasoline. The results can be directly used in the quality control where adulteration of gasoline with methanol is possible. Studies are underway in this direction and results will be available in the future communications. Experiments are underway to propose a suitable normalization method and would be proposed in future.

ACKNOWLEDGEMENT

The authors would like to acknowledge the support of grant MLC1/97 from the Nanyang Technological University, Republic of Singapore.

REFERENCES

1) Ingamells, J.C., and Lindquist, R.H., 1980., " Methanol as a Motor Fuel or a Gasoline Blending Component," SAE The Engineering Research for Advanced Mobility., Alcohols as Motor Fuels," Progress in Technology Series, Number-19. pp 123-134.

2) Mills. G. A., and Harney. B.M., 1974., "Methanol-The New Fuel from Coal," Chemtech, January, pp.26-31.

3) Reed. T. B., and Lerner. R. M., 1973., "Methanol a Versatile Fuel for Immediate Use," Science, Vol. 182, Dec. pp. 1299-1304.

4) Harris. W. D., and Davison. R. R., 1973, "Methanol from Cole," Oil and Gas Journal, Dec. pp. 70-72.

5) Norman D. Brinkman, E. Eugene Ecklund, and Roberta J. Nichols, 1990, " Fuel Methanol A Decade of Progress," SAE The Engineering Society For Advancing Mobility Land Sea and Space PT-36.

6) Poulton, M.L., 1994., "Alternative Fuels for Road Vehicles," Computational Fluid Mechannics Publications., Southampton, UK and Boston, USA.

7) Hamaguchi. H. and Hirakawa. A., 1988., "Raman Spectroscopic Methods" Gakukai Shupan center and Japan Spectrometric Society, Tokyo.

8) John R. Ferraro. and Kazuo Nakamoto., 1994., "Introductory Raman Spectroscopy" Academic Press. Inc.

Hyperspectral image sensor for weed selective spraying

Filip Feyaerts, Pascal Pollet, Prof. Luc Van Gool and Prof. Patrick Wambacq

Katholieke Universiteit Leuven, Kardinaal Mercierlaan 94, B-3001 Heverlee, Belgium

ABSTRACT

Recognizing, online, crops and weeds enables to reduce the use of chemicals in agriculture. First, a sensor and classifier is proposed to measure and classify, online, the plant reflectance. However, as plant reflectance varies with unknown field dependent plant stress factors, the classifier must be trained on each field separately in order to recognize crop and weeds accurately on that field. Collecting the samples manually requires user-knowledge and time and is therefore economically not feasible. The proposed tree-based cluster algorithm enables to automatically collect and label the necessary set of training samples for crops that are planted in rows, thus eliminating every user-interaction and user-knowledge. The classifier, trained with the automatically collected and labeled training samples, is able to recognize crop and weeds with an accuracy of almost 94 %. This results in acceptable weed hit rates (89 % or higher) and significant herbicide reductions (15-67 %). Spot-spraying on the weeds only becomes economically feasible.

Keywords: Hyperspectral imaging, precision agriculture, application, classification, clustering

1. INTRODUCTION

Growing environmental consciousness and economical benefit for the farmer are the driving forces behind *precision farming*. This trend encompasses efforts to decrease the use of herbicides. One of the most promising techniques towards that end is spot-spraying, i.e. to only spray where needed.

Initially, spot-spraying was focused on distinguishing vegetation from soil. Spraying would then be restricted to vegetation patches (weeds and crop) only. A patch-spraying system[1] could result in significant herbicide reductions.

Spot-spraying only on the weeds, however, will result in even higher herbicide reductions. To enable this type of spraying, crop and weeds must be discriminated. In fact, current practice in field spraying nowadays focuses on economical application of herbicides. Most of the herbicides will kill several weed classes, for instance the grasses. Mixing the herbicides, for reasons of cost effectiveness and time efficiency, makes it possible to kill all the weeds in one run. To enable herbicide reductions with this kind of spot-spraying, the samples need to be classified in only two classes: crop and weed. The weed samples need not necessarily be grouped into the individual weed classes.

To make the desired distinction between crop and weeds, shape, size and texture of several vegetation species were investigated.[2,3] This research was based on two-dimensional color images of the plant species. Plant species were found to have regular shape and texture, depending strongly on growth stage. Nevertheless, these techniques are too slow at the moment to be implemented in a real-time evaluation system, due to the mathematical complexity to characterize and recognize the plant species.
It was shown[4,5] that analysis of the spectral reflectance (color appearance) of plants can be used to make a distinction between vegetation classes. These first experiments were conducted using a spectrophotometer with integrating sphere, giving the average reflectance of the plant leafs. However, recent research[6] has shown the correlation between spatial and spectral resolution and possible crop recognition rates. Therefore, the light reflected by the plant should be analyzed with a spatial and spectral resolution as high as possible (read: as long as it remains economically feasible). The spectral information should also not be limited to the classical red, green and blue color bands. It was found that the near infra-red region is particularly interesting to discriminate between soil, crop and weeds. Classical multi-spectral measurement devices[1,7] suffer from low spatial resolution. The analyzed area of these devices is too large, mixing the spectral information of soil and/or crop and/or weeds, hiding the small spectral differences between the vegetation classes. Recent developments enable to discriminate online between crop and weeds based on the multi-spectral optical reflectance, characterizing the vegetation classes.[6] Selectively spraying on the weeds would therefore be possible.

Further author information: (Send correspondence to L.V.G.)
E-mail: Filip.Feyaerts,Pascal.Pollet,Luc.Vangool,Patrick.Wambacq@esat.kuleuven.ac.be

Part of the SPIE Conference on Advanced Photonic Sensors and Applications
Singapore • November–December 1999
SPIE Vol. 3897 • 0277-786X/99/$10.00

However, multi-spectral reflectances of vegetation plants show spatial and temporal variation due to different nutrient levels (N, H_2O, ...), competition with other vegetation plants, herbicides, insects, soil types and so on.[8] Training the classifier with spectral reflectances of all vegetation classes under all circumstances is not feasible. Gathering this world-representative set of samples is difficult – rather impossible – and due to the large variation of the plant reflectance over all circumstances, the classifier cannot discriminate anymore between crop and weeds. Selecting the samples from this world-representative data set that represent the situation on a certain field requires knowledge about the actual influencing factors in the field, which we have not. The only alternative is therefore to collect samples and train the classifier for each field individually. However, this procedure is extremely user-intensive. A representative set of vegetation samples for every vegetation class found on the field must be presented to the sensor. This requires user-knowledge and is time-consuming, returning this method practically and economically not feasible.

This paper presents a self-learning procedure to automatically label vegetation plants as crop or weed in case the vegetation samples are collected in fields with crops that are planted in rows. The presented procedure eliminates every user-interaction and requires no user-knowledge. These labeled samples are then used to train the classifier that will discriminate between the crop and the weeds with high accuracy.

2. SENSOR & ALGORITHMS

2.1. Reflectance sensor

The proposed down-looking weed sensor is based on the principle of diffraction.[9,10] It yields the reflectance spectrum for each point along a narrow linear stripe on the ground and consists of three parts: an objective lens, an imaging spectrograph and a camera. The principle of operation of the spectrograph is shown in Fig. 1.

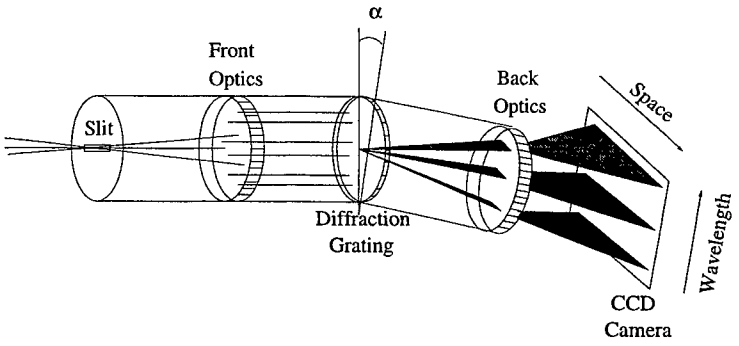


Figure 1. Working principle of imaging spectrograph

The basic principle is to project the diffracted spectral components of the light (obtained with a diffraction grating or a dispersing prism) of a small stripe extracted with the slit from a patch on the ground on a monochrome camera. In this way one of the axes of the camera acts as a spatial axis while the other camera axis is a spectral axis.

This type of spectrograph was chosen above other multi-spectral sensors because it has no moving parts, resulting in higher robustness with respect to vibrations of the spray boom. Next, all the spectral information about the analyzed surface is available at once, through diffraction of the reflected light. The speed limiting factor of spectral analysis is the frame rate of the camera. The parameters of this sensor can also be chosen to obtain a low cost device with reasonable performance. Finally, this sensor offers flexibility in selecting the wavelengths of interest which form the feature for classification. The spectral range can also be adjusted to cover interesting wavelengths to enable discrimination between crop and weeds.

We have constructed such a device with parameters given in Table 1. The spectral range was extracted from previous studies,[4,5] taking into account also the limited spectral sensitivity of a low-cost camera (typically 400-1000 nm). Within this range, wavelengths can easily be chosen using only the most significant spectral lines on the camera. Next, to keep power consumption low, the sensor was designed to operate under normal outdoors conditions with

Table 1. Design parameters of low cost (standard optical components) imaging spectrograph

Parameter	Value
F-number	3
Spectral resolution	35 nm
Available spectral range	400 - 1000 nm
Free spectral range	1 octave taken from the available range
Slit width	200 μm
Slit length	8 mm
Spatial resolution	depends on objective and height
Spectral bands	150

the sun as the only illumination source. With an objective lens with focal distance of $f = 3.5$ mm and from a height of 1 m, a stripe with a length of 3 m is analyzed. These are approximately the working conditions that we envisage.

2.2. Data processing

Crop and weeds can be discriminated based on the spectral reflectance of the plants.[4,5] A sensor and algorithms to analyze online the plant reflectance were previously proposed.[6,11] Combining row-detection and multi-spectral

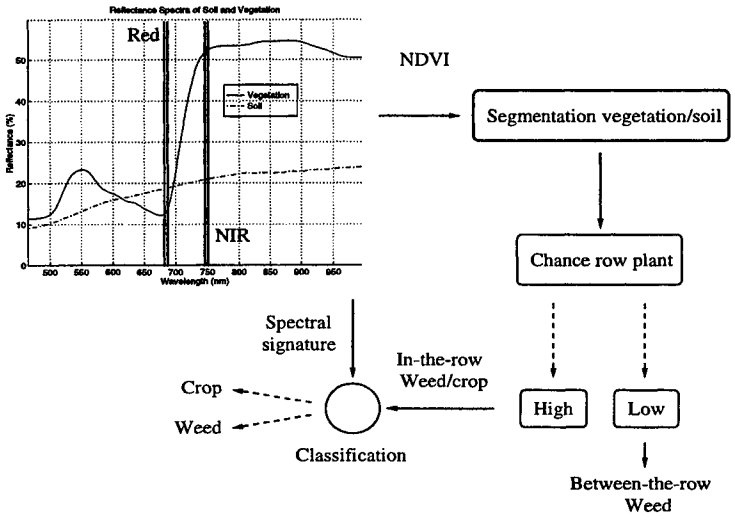


Figure 2. Data processing

analysis results in fast weed detection. Fig. 2 summarizes the processing steps.

First, the image is segmented into vegetation and background (soil) based on the Normalized Vegetation Index. This index, known from remote sensing, is based on the huge differences in reflectance in the red and near infra-red region for vegetation and soil, as can be seen in Fig. 2. The $NDVI$ can be written as in Eq. 1, with R_j the reflectance at wavelength j.

$$NDVI = \frac{R_{NIR} - R_{Red}}{R_{NIR} + R_{Red}} . \quad (1)$$

Based on the position of each vegetation sample in the scene (on the linear stripe), the chance of a sample being a row sample is calculated. To get reliable results, the positions of the vegetation samples of a few lines are combined.

Those lines are either subsequent analyses with the same sensor (subsequent lines in the moving direction) or analyses with different sensors along the spray boom. The previously published algorithm was adapted to generate chances in the range [0-1]. The present algorithm assumes that the position of the rows is Gaussianly distributed with known mean and variance. The distance between two rows is therefore Gaussianly distributed with the same mean and twice the variance. The algorithm works as follows:

1. Calculate the histogram for each base number *base* of occurrences for the remainders of each distance between two vegetation parts in the scene after division by some base number *base*. The following conditions speed up the algorithm and limit the memory usage without loss of performance.
   - The aforementioned distance is normalized: it is divided by the distance that equals the variation on the row-distance determined by the accuracy of seeding, growth stage of the crop, ...
   - Repeat the previous step for each base number *base*, calculated as the distance between every two vegetation parts in the scene.
   - The previous step is only carried out if the base number is in the interval $[0.33guess - 3guess]$ where *guess* is the estimated distance between the rows based on the real crop-dependent row-distance and a rough estimation of the height and angle-of-view of the sensor.
2. Determine the row-distance RD as the base number *base* for which the normalized sum of the smallest ($[0 \;\; 1]$, twice the variance) and largest ($[base - 2 \;\; base - 1]$, twice the variance) number of occurrences is highest.
3. The chance of sample A being a row sample is the average of the $p_{A,j}$ calculated as follows:.
   - Calculate the remainder of the distance between the sample A and another sample j after division by the estimated row-distance RD.
   - $p_{A,j}$ is the probability of getting this remainder, calculated using the Gaussian probability density function with mean 0 and variance σ^2(row distance).
   - σ^2(row distance) is calculated as two times the given variance on the given row-distance, calculated for the estimated row-distance RD.

Below a chosen threshold on this row chance, the sample is assigned a between-the-row (weed) label. Above the same threshold, the sample is assigned a crop or weed label depending on the multi-spectral classification result.

The optimal spot-spraying strategy would result in some herbicide reductions while still most of the weeds are sprayed. Crop should certainly be recognized accurately to enable herbicide reductions. On the other hand, the misclassification of one weed can be compensated recognizing other weeds – if any – in the analyzed (sprayed/not-sprayed) area.

3. CLASSIFIER INITIALIZATION

As mentioned in Section 1, a representative set of data samples must be collected to train the classifier for recognizing crop and weeds on a specific field. To collect the data samples manually, the user must recognize the different vegetation classes. This procedure takes a lot of time. End-users that we envisage using this technology, do not have time to spend collecting vegetation samples and do not always have the required knowledge to collect a representative set of data samples.

The proposed procedure eliminates these previous restrictions completely if the crop is planted in rows. This is the case for a lot of economically important crops such as maize or sugar beet. The procedure describes how an automatically collected, representative set of data samples can be labeled as crop or weed with high accuracy. Initializing the classifier based on this labeled set of samples will result in good crop recognition rates, as will be shown later on.

3.1. Feature selection

The training phase should be as short as possible. The end-users can and will not wait long, knowing that worst case savings on chemicals are rather small. As computation times depend on the feature space dimension, this dimension must be reduced without losing too much information.
The camera contained 786 rows of 576 pixels. After subsampling, the number of pixels was reduced by a factor four. As the imaging spectrograph was designed/optimized to work under normal outdoors conditions with the sun as the only illumination source (cf Section 2.1), only part of the camera was illuminated by the spectral decomposition. Only those camera lines were processed any further.

To eliminate redundancy in the data and to reduce the feature space dimension even more, a Principal Component Analysis (PCA) was carried out on the subsampled set of camera lines in the illuminated camera area. Only the most relevant principal components were used for further analysis, containing together just more than 95 % of the total original variance.

3.2. Cluster principles

The basic principle of the algorithm is that crop only appears in the rows while weed appears both in and between the rows. It can be detected if a plant is situated in a row or between the rows (cf Section 2.2). The row is assigned a certain width, depending on the positioning accuracy of the crop plants (the plants not planted perfectly in rows), the size of the plants, the distance between the rows (R), the distance between the crops in the row (S) and the local weed density (O). Above a certain weed density, the row-distance cannot be detected accurately anymore, so that the measured samples must be discarded.

Labeling all the samples in the rows as crop would result in all the weeds found in the rows being misclassified. The proposed cluster procedure performs much better on average.

The proposed cluster algorithm starts from the measured data set that consists actually of two subsets: the set of between-the-row samples and the set of in-the-row samples.

3.2.1. Tree-based cluster procedure

For a given data set, a sample is generally assigned the cluster label of the cluster to which the distance to the cluster center is smallest.

The k-means cluster algorithm is a stochastic process in which random cluster centers are chosen in a set of data samples. The initial cluster centers are shifted to stable positions, taking into account the (local) distribution of the data set samples in the feature space in the neighborhood of the actual cluster centers. Due to local density extremes and the definition of stability of the cluster centers (smallest distance between centers before and after shifting), different initializations can result in different resulting stable cluster centers.

An adaptation to the standard k-means cluster algorithm consists in choosing the initial cluster centers in both subsets of row-samples and between-the-row samples. Pure random initialization of the cluster centers in the total data set does not guarantee that initial centers are located in both subsets or even better in regions of crop and weeds in the feature space. The density of the crop samples depends on the number of crop and weed samples that were gathered and may be much lower, resulting in no initial cluster centers in the actual crop feature subset.
Based on the same idea, each subset of the original collected data set that is neither a weed cluster, nor a crop cluster, can be split in two subsets that may contain significantly more or less crop samples. The optimal binary split of a data set will result in a data subset containing mostly weed samples and a data subset containing mostly crop samples. This way, binary tree-like clustering is obtained of the original data set.

Classical criteria to determine the optimal number k of clusters include overall maximization of the between/within variance of the different clusters. As only the crop clusters should be discriminated from the weed clusters, this kind of criterium would be too strong.
The proposed criterium is to split the aforementioned clusters as long as the total number of crop labeled samples is lower than expected. Due to the measurement principle, the number of crop samples in the collected data set can be estimated quite accurately: the distance along the crop row(s) can be measured and combined with the distance between the crop plants in the row to estimate the number of measured crop plants.

3.2.2. Cluster labeling

After clustering the samples, each cluster must be given a label, crop or weed.

Each cluster j of the collected data set consists of NT_j data samples with NR_j row-samples and $NT_j - NR_j$ between-the-row (known weed) samples. The relative amount of row samples in cluster j (V_j) is given in Eq. 2. The corresponding percentage of known weed samples is easily calculated as $1 - V_j$.
$1 - V_j$ is a good first order estimate for the relative frequency of weed samples in the set of row-samples. A first order estimate for the relative frequency of the weeds in cluster j ($\phi_{j,\text{weed}}$) and the expected number of weed samples $E_{j,\text{weed}}$ can therefore be written as function of V_j. As we are dealing with a binomial distribution (crop/weed) with given probability, the variance on the number of weed samples in the cluster ($\sigma_{j,\text{weed}}$) is determined easily (Eq. 2).

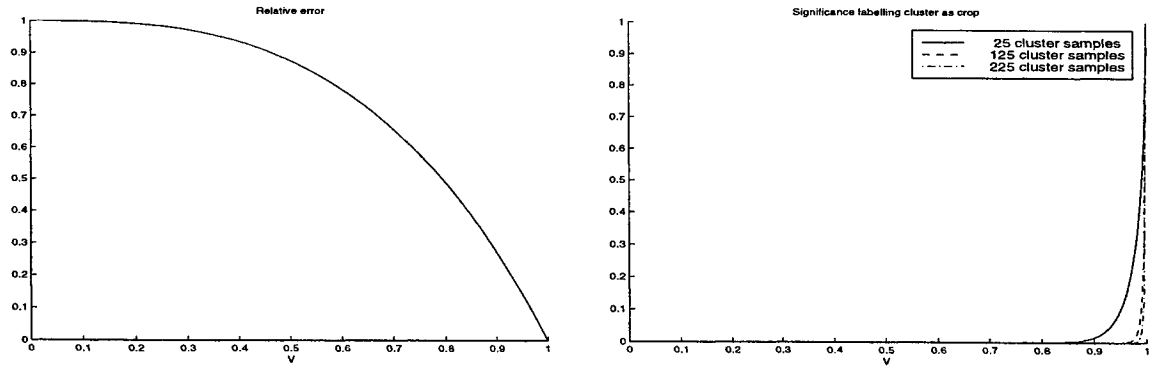
$$\begin{cases} V_j = \frac{NR_j}{NT_j} \ . \\ \phi_{j,\text{weed}} = \frac{NT_j - NR_j}{NT_j} + \frac{NR_j}{NT_j}\frac{NT_j - NR_j}{NT_j} = 1 - V_j^2 \ . \\ E_{j,\text{weed}} = \phi_{j,\text{weed}} NT_j = (1 - V_j^2) NT_j \ . \\ \sigma_{j,\text{weed}} = \sqrt{\frac{\phi_{j,\text{weed}}(1-\phi_{j,\text{weed}})}{NT_j}} = \sqrt{\frac{(1-V_j^2)V_j^2}{NT_j}} \ . \end{cases} \tag{2}$$

Assigning a crop label to each of the samples of the cluster will result in an estimated percentage of errors, given by the expected number of weed samples in the cluster ($E_{j,\text{weed}}$, Eq. 2). However, consider the following hypothesis:

> There are no weed samples in cluster j, given the percentage of row samples (or complementary given the percentage of known weed samples) in cluster j.

If the chance of measuring $\phi_{j,\text{weed}}$, assuming an average of zero weed samples and a variance that equals the variance of the binomial distribution, resulting from the measured percentage of known weed samples $\phi_{j,\text{weed}}$, is larger than $1 - S$, it is said that the hypothesis is true on a significance level of $1 - S$ percent. If this hypothesis is true on a significance level of more than $1 - S_{\text{reject}}$, it is assumed true (summarized in Eq. 3).

$$\frac{1}{\sqrt{2\pi}\sigma_{j,\text{weed}}} \int_{\phi_{j,\text{weed}}}^{\infty} e^{-\frac{\phi^2}{2\sigma^2_{j,\text{weed}}}} d\phi \begin{cases} > 1 - S_{\text{reject}} \longrightarrow \text{crop label} \ . \\ < 1 - S_{\text{reject}} \longrightarrow \text{weed label} \ . \end{cases} \tag{3}$$



a: Estimation of average percentage of errors

b: Significance level $1 - S$

Figure 3. Estimation of average percentage of errors and significance level for labeling the cluster as crop for varying percentage of in-the-row samples V_j

For a fixed rejection level S_{reject}, the minimal percentage of row samples (V_j) to assign a crop label to the cluster, depends on the number of samples in the cluster. Fig. 3 shows that, for a small rejection level S_{reject}, V_j should be rather high. As almost no clusters can be found with such large V_j, the minimal V_j for labeling the cluster as crop was set to 85 %. Resulting in lower significance (higher rejection level), some weed samples were labeled as crop.

However, labeling a clustered set of samples as crop might actually result in even more errors at low values of V_j. Due to the characteristic features of crop and weeds, most of the row-samples in the set of samples that are closely located to weed samples, will probably be weeds also. At higher values of V_j, the error will probably be lower as there are almost no known weed samples (between-the-row samples) around in the feature neighborhood, so that the row-samples are most likely crop.

The clusters that are split recursively are those of which no decision can be taken with reasonable significance but with a percentage of row samples as high as possible. Combination of Fig. 3 a and b learns that the number of errors for a given significance level is smaller for clusters with a larger number of samples. Clusters are therefore only split if at least 10 % of of the expected number of crop samples is missing.

3.3. Iteration of the clustering process

As the threshold on the crop significance-level is set high (low rejection level), most of the samples in the crop labeled clusters will actually be crop samples. Only a small number of weed samples will get wrong labels. Also as cluster are less often labeled as crop, weed labeled clusters may contain some false labeled crop samples.

A procedure was developed that takes this different sample labeling variance into account. Weed will mostly be labeled as weed whereas crop will be labeled as crop less often. Iteration with different initialization of the cluster centers of the aforementioned binary cluster procedure, for the same set of data samples, will result in small variation in labels for weed and higher variation for crop samples. A sample can be labeled as crop if the average of the crop labels (0 for weed, 1 for crop) exceeds an iteration dependent threshold. This threshold was found to perform well under different weed densities if set at (number of iterations)$^{-1}$.

Finally, it was ensured that all samples that were priorly known as weed, were given a weed label.

4. RESULTS

4.1. Measurement conditions

The spectral measurements were obtained under *real-field conditions*: outdoor measurements for actual plants on real fields. Soil and leafs showed natural variations in their orientations. The illumination (intensity, spectrum and direction of incidence) and the vehicle speed varied while gathering the spectral data samples.

The experiments in the rest of this paper were based on reflectance measurements of perfectly labeled plant samples on one field only. A statistically relevant number of samples was collected manually, covering the actual stress effects on that field. Nevertheless, special care was taken to gather a data set of only healthy looking crop and weed samples.

One of the most economically important crops, sugar beet, was selected for the experiments. The data set consisted of 905 samples of *Beta vulgaris L.* (common beet), together with *Poa annua L.* (annual bluegrass, 1830 samples), *Plantago lanceolata L.* (narrow-leaf plantain, 1412 samples), *Stellaria media L.* (common chickweed, 1019 samples), *Chenopodium album L.* (lambsquarters, 867 samples) and *Polygonum persicaria L.* (Redshank, 988 samples). The plants, fully grown, varied in age from 4 to 8 weeks.

Measurements were performed using a monochrome 1/2" (4.8 x 6.4 mm) CCD-camera (MX5 of Adimec) coupled to the imaging spectrograph (parameters of Table 1). The processing unit consisted of an 166 MHz Pentium PC with 32 MByte RAM. The sun was the only illumination source. Natural variations of the illumination (intensity and spectral composition, caused mainly by clouds and shadow) made it necessary to measure a reference object at all times. The samples were classified based on a feature (set of reflectances) of selected wavelengths. This set of wavelengths was optimized for the discrimination of the measured crop and weed samples.

Table 2. Standard sugar beet field parameters and spray settings

Distance between the rows R	45 cm
Distance between crops in the rows S	18 cm
Spray width	50 cm
Spray length	7.5 cm

4.2. Experiments

In the following experiments, it is investigated how the split and combine clustering algorithm performs. The performance is investigated for different weed densities and a fixed row width/tolerance W. The number of iterations and the threshold on the crop significance ($V > 0.85$) were constant over the experiments.

The results are compared to perfect labeling: number of correctly labeled samples, effect on sample classification, weed hit rate and herbicide reduction.[11] To estimate the weed hit rate and herbicide reduction, standard conditions on sugar beet fields (given in Table 2) were assumed. The spray resolution is determined by the spray width and the spray length. The spray width is the size of the spray pattern along the axis of the spray boom. The resolution in the driving direction (spray length) is determined by the size of the spray pattern in that direction, the on/off frequency of the spray nozzle in combination with the driving speed and the setup time to get a stable spray pattern. An almost planar spray pattern is extremely small in the driving direction and is stable within milliseconds. An achievable on/off frequency, limited by fluid dynamics, is 15 Hz. Driving at a speed of 4.05 km h^{-1} results in a spray length of 7.5 cm.

For the experiments, a data set was created synthetically containing crop and weed samples in selected amounts. From the chosen number of crop samples, an imaginary field length* and corresponding field area† can be associated. From the given number of weed samples and the imaginary field area, an average weed density can be calculated. Vice versa, the measurement system allows to determine the analyzed field length, local weed density and analyzed field area. Real life measurements would give the same information as for the synthetic data sets.

In the following experiments, classification results are given for a test set of samples with the classifiers being trained on the labeled train samples. Samples in the separated train and test sets were selected at random from the original collected data set.

4.3. Perfectly labeled samples

The data set, as previously mentioned, was originally perfectly labeled. Classifying the set of samples with known labels enables to compare the performance of the cluster algorithm. The following classification principles were compared.

4.3.1. Minimal distance classifiers.[12]

Parametric statistical methods based on this principle assume that sample reflectances are Gaussianly distributed round class averages. The sample is assigned to class X if the feature is closest to the expected feature of class X, in an Euclidean sense (**ED**), or in a Mahalanobis sense (**MD**).

4.3.2. Multi-layer neural network with nonlinear mapping between input and output space (MLNLM).[13]

The neural network consisted of three layers in which each layer was fully connected to the next one (5-8-2 neurons). The MLNLM has the ability to discriminate between the crop and weeds (it has only two neurons for output). Therefore, the classifier is not able to detect how many types of weeds are presented, neither can it discriminate between them. That could be useful if it is necessary to use different herbicides (or different concentrations of herbicide) for each weed individually. The training procedure was implemented with a back-propagation learning rule using momentum and an adaptive learning rate (Matlab™Neural Network Toolbox).

*The imaginary field length can be considered the length of one crop row that would result in the chosen number of crop plants given the distance between neighboring crop plants in the row (S).

†The imaginary field area is calculated as the imaginary field length multiplied with the distance between the rows (R).

Table 3. Success rates, weed hit rate and herbicide reduction for originally perfectly labeled crop and weed samples using different classifiers

	ED	MD	MLNLM
Classify			
Sugar beet	89.0	91.7	96.7
Weed	99.4	99.1	98.8
Average	94.2	95.4	97.8
Spraying effect			
Hit rate	99.7	99.6	99.4
Herbicide reduction	24.9	25.4	26.2

Table 3 gives the success rates for classification of sugar beet and weed as crop and weed and corresponding weed hit rates and herbicide reductions using each of the proposed classifiers. The MLNLM classifier has an average sample classification success rate of 97.8 %, an expected weed hit rate of 99.4 % and a herbicide reduction of 26.2 % for an average weed density of 37 weed plants m^{-2}. It is the best classifier as it gives the best hit rate and herbicide reduction results and allows also fast evaluation. Next, it allows black-box training on a set of labeled samples. Especially for this last reason, the estimated classification success rate and corresponding weed hit rate and herbicide reduction in the rest of the experiments are only calculated using this MLNLM classifier.

4.4. Proposed binary tree-based labeling

Every crop sample and part of the weeds, depending on the ratio of row tolerance W and distance between the rows R, would occur in the rows. Every crop sample and this portion of the weeds, selected at random, were given a row-label. The rest of the weeds were given a between-the-row label (known weed samples). The row detection algorithm proposed in Section 2.2, enables this kind of labeling.

The weed density was varied from very low (12 weed plants m^{-2}) to relatively high (86 weed plants m^{-2}). At this last weed density, for the given spray resolution, three weed plants are present, on average, in the spray rectangle. Possible herbicide reductions are rather small on fields with higher weed densities and are therefore not simulated.

Table 4. Labeling accuracy, classification success rates, weed hit rate and herbicide reduction for varying weed densities

Weed density (# m^{-2})	12	37	86
Proposed labeling			
Sugar beet	96.9	97.0	98.6
Weed	95.0	95.7	97.4
Weed in row as weed	70.0	74.6	84.6
$\overline{k}$	6	9	14
Classify			
Sugar beet	96.5	96.1	98.2
Weed	91.2	87.5	60.8
Average	93.9	91.8	79.5
Spraying effect			
Hit rate	92.7	93.5	89.1
Herbicide reduction	67.8	30.6	14.5

Weed density (# m^{-2})	12	37	86
Labeling (k fixed at 50)			
Sugar beet	90.1	90.3	97.5
Weed	93.1	99.2	94.5
Weed in row as weed	91.0	95.0	67.0
$\overline{k}$	50	50	50

Table 4 gives the percentages of correct labeled crop and weed samples, the classification success rate and corresponding weed hit rate and herbicide reduction for different weed densities that resulted from the proposed binary tree-based cluster procedure and a cluster procedure with constant number of clusters ($k = 50$).

Labeling accuracy for the proposed cluster algorithm is systematically better than clustering with a fixed number of cluster centers. Especially the crop samples are labeled more accurately, which is preferable to enable possible herbicide reductions. As explained before, a large number of clusters could result in too many clusters containing too little samples to give a significant crop or weed label. Closely related to lower crop label accuracy, significantly more row-weeds will be labeled as weed using this fixed number of clusters.

The labeling accuracy is almost insensitive to the weed density. A low increase in accuracy with increasing weed density could be explained as relatively more weed samples between the rows enable to cluster the weed samples more accurately due to a higher density of known (weed) samples. For the same reason, crop could also be labeled more accurately.

Corresponding classification success rates do not differ significantly from the case in which the sample labels are perfectly known. The success rate for the crop samples follows the same tendency as the crop labeling accuracy. However, they show a decrease in recognizing the weeds with increasing weed density. Weed hit rates are still acceptable for all weed densities. Herbicide reductions decrease, as expected, with increasing weed density. Due to an increase in the classification errors, the herbicide reduction remains significant, even for the higher weed density.

5. CONCLUSION

The proposed cluster algorithm can be used to label as set of data samples, collected with the proposed multi-spectral image sensor. Prior row or between-the-row labels that were assigned to each of the samples suffice to eliminate every user-interaction and user-knowledge in collecting and labeling the data set.

Based on this labeled data set, a classifier can be trained to recognize crop and weed. This classifier can be used to selective spray on weeds only. Resulting in acceptable weed hit rates (89 % or higher) and significant herbicide reductions (15-67 %) for the simulated weed densities, spot-spraying becomes economically feasible.

ACKNOWLEDGMENTS

Our special thanks go to the Flemish Community (VLIM-project), the Belgian Ministry of Agriculture (IWONL-project) and Teejet Technologies (Spraying Systems), our industrial partner, for the financial support to this project.

REFERENCES

1. W. L. Felton and K. R. McCloy, "Spot spraying," *Agricultural Engineering Research* **73(6)**, pp. 9–12, 1992.
2. R. Gerhards, A. Nabout, M. Sökefeld, W. Kühbauch, and H. A. N. Eldin, "Automatische erkennung von zehn unkrautern mit hilfe digitaler bildverarbeitung und fourier transformation," *Journal of Agronomy and Crop Science* **171**, pp. 321–328, 1993.
3. E. Franz, M. R. Gebhardt, and K. B. Unklesbay, "Algorithms for extracting leaf boundary information from digital images of plant foliage," *Transactions of the American Society of Agricultural Engineers* **32(2)**, pp. 625–633, 1995.
4. E. B. Knipling, "Physical and physiological basis for the reflectance of visible and near-infrared radiation from vegetation," *Remote Sensing of Environment* , p. 5, 1970.
5. F. Hahn and A. Y. Muir, "Weed-crop discrimination by optical reflectance," *Proc. of the IV International symposium on fruit, nut, and vegetable production engineering* **1**, pp. 95–103, 1993.
6. F. Feyaerts, P. Pollet, L. V. Gool, and P. Wambacq, "Sensor for weed detection based on spectral measurements," in *Remote sensing*, W. L. P.C. Robert, R.H. Rust, ed., *Proc. of the IV International Conference on Precision Agriculture* **B**, pp. 1537–1548, 1998.
7. R. J. Haggar, C. J. Stent, and S. Isaac, "A prototype hand-held patch sprayer for killing weeds, activated by spectral differences in crop/weed canopies," *Agricultural Engineering Research* **28**, pp. 349–358, 1983.
8. G. A. Carter, "Responses of leaf spectral reflectance to plant stress," *American Journal of Botany* **80(3)**, pp. 239–243, 1993.

9. D. E. Battey and J. S. Slater, "Compact holographic imaging spectrograph for process control applications," *Proc. SPIE* **2609**, pp. 60–64, 1993.
10. E. Herrala, T. Hyvarinen, M. Aikio, and J. Lammasniemi, "Imaging spectrometer for process industry applications," *Proc. SPIE* **2248**, pp. 33–40, 1994.
11. P. Pollet, F. Feyaerts, L. V. Gool, and P. Wambacq, "Weed detection based on structural information using an imaging spectrograph," in *Remote sensing*, W. L. P.C. Robert, R.H. Rust, ed., *Proc. of the IV International Conference on Precision Agriculture* **B**, pp. 1579–1592, 1998.
12. K. Fukunaga, *Introduction to statistical pattern recognition, 2nd edn*, Academic press, New York (N.Y.), 1990.
13. D. Rumelhart, G. E. Hinton, and R. J. Williams, "Learning internal representations by error propagation," *Parallel Data Processing* **1(8)**, pp. 318–362, 1986.

A Portable System Approach of Monitoring Plant Nutrient Deficiency using Fibre Optic Spectrophotometry

Anand K. Asundi[1], Jun Wei Chen[1], Duo Min He[1], Oi Wah Liew[2]

1. School of Mechanical and Production Engineering, Nanyang Technological University, Singapore 639798
Email : masundi@ntu.edu.sg or p7218361d@ntu.edu.sg or mdmhe@ntu.edu.sg;
Web: http://www.ntu.edu.sg/home/masundi
Fax : **65-791-1859**

2. Chemical Process and Biotechnology Dept,
Singapore Polytechnic, Singapore 139651
Email : OWLiew@sp.ac.sg

ABSTRACT

In this paper, a portable sensing system is developed using fibre optic spectrophotometry (FOSpectr) principle for measuring and detecting of stresses induced in plants due to nutrient deficiencies. Chlorophyll fluorescence in plants is used to monitor the effects of nutrient stress in plants. As this method aims at providing an early detection and warning of nutrient deficiencies, it gives an alternative to argument current semi-quantitative and destructive methods of nutrient analysis. Our early papers[1] had demonstrated significant differences in the color reflectance of plants' leaves when plants were subjected to various nutrient-deficient media. Developed using off-the shelf components, this digital sensing optical system could measure and detect the slight variation in the plants' reflectance and hence its chlorophyll levels. These relative levels of chlorophyll are determined by measuring the plants' color reflectance of light while using the wavelength of the healthy plants as a reference for comparison. This system comprises of a miniature spectrometer containing 1024 CCD detectors covering a visible light spectrum of wavelength ranging from approximately 400nm to 800nm and a reflective probe. A laptop with a PCMCIA A/D data acquisition card is used in conjunction with a customised program.

The application of this portable system allows immediate measurement as compared to conventional sampling methods using chemical test-kit or using multi-spectral imaging which is a time consuming and labor-intensive process. It could also be extended to incorporated into an automated, closed looped, controlled system providing online monitoring for high tech agriculture farms.

Keywords: Hydroponics, plant nutrient analysis, nutritional disorder, visual symptoms, Fibre Optic Spectrophotometry (FOSpectr), Fibre optics sensing

Part of the SPIE Conference on Advanced Photonic Sensors and Applications
Singapore • November–December 1999
SPIE Vol. 3897 • 0277-786X/99/$10.00

1. INTRODUCTION

Early detection of nutrient stress *before* the appearance of deficiency symptoms in plants is of paramount importance to ensure good crop growth. Specific nutrient disorder (deficiency or toxicity) induces distinct plant symptoms can be identified by trained personal when the plants show visible symptoms in coloration of their leaves after a period of days. However, this can be difficult and subjective due to crop variation and compounded visual anomalies when more than one nutrient is deficient[1]. Semi-quantitative assessment of specific ions is possible to ensure that the tissue elemental content is at sufficiency level. This generally involves sampling of plant parts like leaves, petioles, bark and fruits at specific growth intervals. The most convenient and facile field method available for assessing levels of soluble nutrients in plant sap or extract is rapid colorimetric tests using Merckoquant test strips[2]. This is a destructive method that requires the removal of plant parts. More recently, chlorophyll fluorescence has also been used to monitor the effects of nutrient stress in plants[3,4].

However, it is found that pre-visual indication stress could be detected with proper equipment and technique. By using FOSpectr methodology, it is possible to provide early detection in spectral reflectance variation of plants chlorophyll level and also quantifying the symptoms, thereby assisting growers in rectifying the problem before it is too late as the crop would already have exhibited loss in yield and vigor. Previous work has shown that it is possible to detect a nutrient deficient in both nitrogen and calcium as early as day 4 of the experiment, well before the development of visual symptoms on the 8th day. With this methodology, a highly sensitive, portable system for detection of plants chlorophyll levels is developed for immediate measurement and read-off. This system is developed for non-destructive field test for early detection of plants' nutrient deficiency prior to the appearance of visual symptoms. Once again, the effects of nitrogen and calcium deficiency in plants are chosen to illustrate the usage of this system.

2. SYSTEM CONFIGURATION

2.1 FOSpectr Set-Up

Developed using off-the shelf components, this digital sensing optical system can measure and detect the slight variation in the plants' reflectance. The relative levels of chlorophyll are determined by measuring the plants' color reflectance of light while the wavelength of the healthy plants is used as a reference for comparison with plants of same species. This system comprises of a miniature spectrometer consisting of 1024 CCD detectors covering a wavelength ranging from approximately 400nm to 800nm, giving an accuracy of 0.5nm. A 12V white light source from tungsten halogen lamp and a bifurcated fibre optic cable (see **Figure 1** for cutaway of reflective probe) for light transmission to the measurand and spectrometer is also used. Communication is established between the spectrometer and the laptop using a PCMCIA A/D data acquisition card via a customised program (**Figure 2** shows the block diagram while **Figure 3** shows the actual configuration).

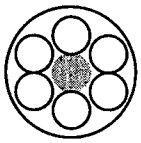


Figure 1: Cutaway of probe shows optical fibers, 6 illumination probes and 1 reading probe to the spectrometer

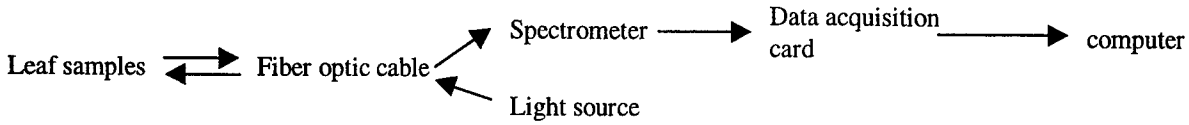


Figure 2: Block diagram of reflective optical fiber instrument

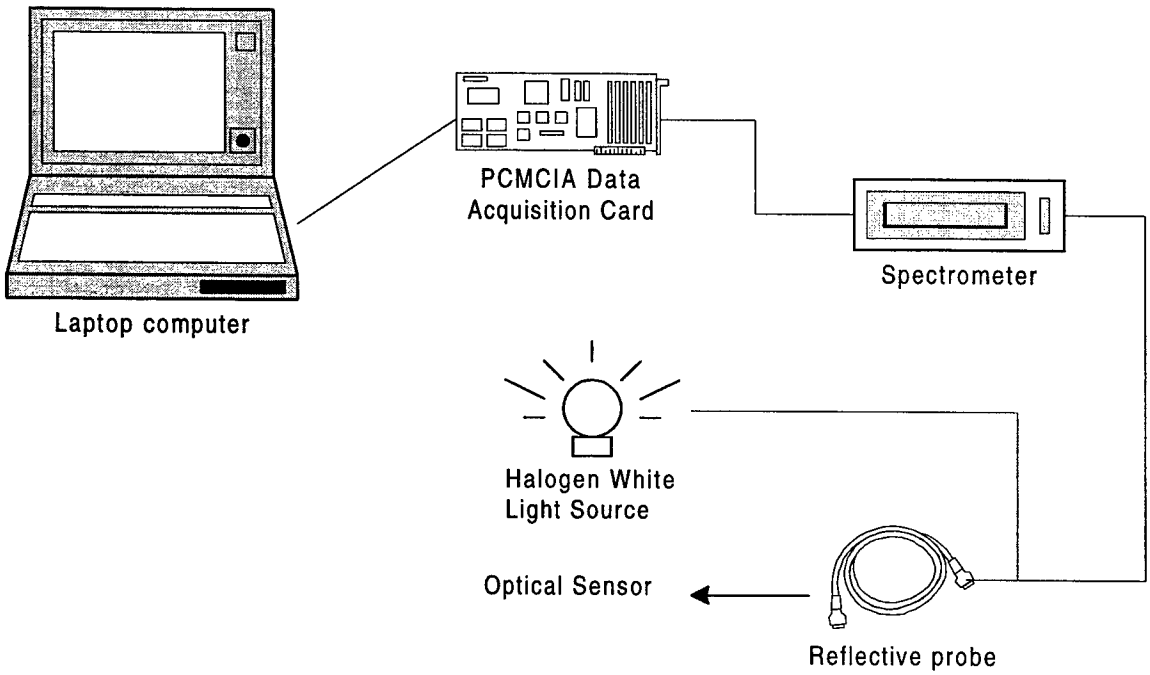


Figure 3: Schematic diagram for actual system configuration of portable system using FOSpectr analysis for plants' chlorophyll level

Figure 4: Plant samples subjecting to various nutrient solutions

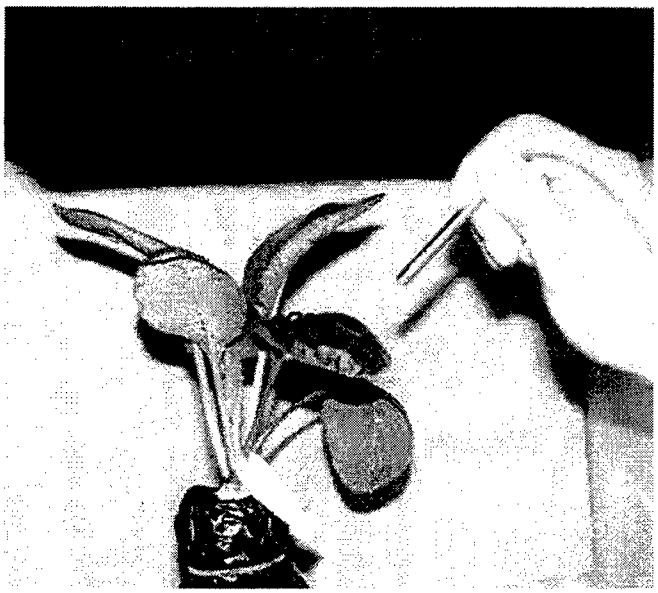


Figure 5: Chlorophyll level of leaves taken by portable system

2.2 Experimental Design

Seeds of *Brassica chinensis* var *parachinensis* (Bailey) were germinated in sponge cubes and the seedlings were maintained in COMPLETE hydroponics nutrient solution (based on a modification of Arnon and Hoagland's solution) for 26 days. At the commencement of the experiment, the plants were grown individually bottles of either "COMPLETE", "0% N" or "0% Ca" nutrient solution. Thirty-two plants were allocated to each of the three nutrient solutions (See **figure 4**). Of these, twenty plants were assigned to non-destructive methods to monitor treatment effects using the new portable system based on FOSpectr analysis. **Figure 5** shows the acquiring of leaf reflectance using the portable system. This was carried out daily on both mature and young leaves. The remaining 12 plants for each treatment were assigned to destructive analysis by extraction of plant sap to monitor the nitrate and calcium levels of mature and young leaves every four days. On the ninth day of the experiment, the three different nutrient media were replaced by a fresh batch of "COMPLETE" nutrient solution. Both FOSpectr analysis and destructive methods of data collection as described above were carried out for a further nine days.

3. RESULTS

3.1 FOSpectr analysis

FOSpectr analysis on the leaves' reflectance ($\approx 550\text{nm} \leq \lambda \leq 570\text{nm}$) showed first indication of nutrient deficiency on day 4 (Referred **Figure 6 & Figure 7**). In the case of nitrogen deficiency, the differences were detected in the old leaves and not in the young leaves. On the other hand, calcium deficiency was reflected in FOSpectr measurement of young leaves and not in the old leaves.

3.2 Tissue elemental levels

Referring to **Figure 8,** nitrogen-deficient plants showed obvious reduction in nitrate levels in old leaves after 5 days in 0% N medium. Significant drop in nitrate levels in young leaves appeared only after 9 days. In contrast, there was no obvious decline in calcium levels for both old and young leaves of plants growing for nine days in calcium-deficient medium.

3.3 Physiological responds

Tracking of physiological and morphological changes during the trial period were carried out by measurements of old and young leaves by FOSpectr, nitrate and calcium levels in plant sap by the appropriate Merckoquant tests (destructive testing), plant height, number of leaves and diameter of the largest leaf. FOSpectr analysis demonstrated significant differences within four days after the plants were placed in nutrient-deficient media. Recovery of the nutrient-stressed plants could also be detected after transferring the plants in complete nutrient solution. In contrast, plant response to nitrogen and calcium deficiency in terms of reduced growth and tissue elemental levels was slower and less pronounced. Figure 8 shows the inconsistency of nitrate/ calcium level taken by using Merckoquant chemical tests kit.

4. DISCUSSION

In the case of nitrogen deficiency, significant changes in the reflected wavelength from mature leaves and not from the young leaves were observed. This was in agreement with the expected location of physiological changes associated with deficiency of a mobile nutrient element and with previous results. With plants maintained in "0% Ca", calcium adequacy in the old leaves was indicated by FOSpectr but simultaneous deficiency was detected in the young leaves. This was again in agreement with the fact that calcium is an immobile nutrient element whereby indications of a deficiency were expected first in the young leaves. However, the difference in reflectance wavelength of young leaves in calcium-deficient plants compared with control was not as striking as for old leaves in nitrogen-deficient plants. This can be explained by the fact that chlorophyll breakdown is associated with nitrogen deficiency whereas loss of membrane integrity is more a consequence of calcium deficiency. It was also notable that the wavelength where these changes were observed was lower for the old leaves compared with the young leaves. Furthermore, the old leaves of control plants grown in COMPLETE nutrient solution showed a gradual increase in wavelength over the 18-day growth period, a trend that was not observed for the young leaves. This was attributed to the senescence process of the older leaves associated with their natural development towards moribundity.

The results illustrate key issues in determining nutrient deficiency and sampling procedures for data acquisition. Firstly, FOSpectr would be particularly useful for early detection of deficiencies where the nutrient elements are involved in the maintenance of chloroplast integrity, for example nitrogen, manganese, magnesium and iron. Secondly, the portable system makes it possible for fast and immediate analysis of the leaf reflectance data collected, hence giving ample of time for the farmer to made necessary rectification in his/ her nutrient solution. However, as the decision to perform FOSpectr analysis on old or young leaves is dependent on the deficient element in question and also the wavelength at which nutrient adequacy is indicated for old leaves also depend on the age of the plant. Hence, the results clearly demonstrated the need to establish baseline reference points for various growth periods when attempting to interpret the data collected. In addition, plant analysis would still be necessary for identification of the specific element that is deficient.

5. CONCLUSION

In conclusion, the application of this portable system allows immediate measurement as compared to conventional methods of nutrient analysis which are both a time consuming and labour-intensive process. FOSpectr analysis was found to be useful for detection of deficiencies for both mobile and immobile elements at an earlier stage compared with growth and sap analysis. This ecological environment monitoring is non-destructive and highly sensitivity, especially in allowing farmers to monitor the recovery of the plants growth under deficient levels of particular essential element, simulating early remedial action to asses the benefits of early detection of nutrient stress. There are good prospects for developing this on-line tracking method to be linked to appropriate computer software for quantitative analysis. Furthermore, it could be extended to incorporated into a new fully automated hydroponics system based on optical feedback control of environmental and biological parameters, providing online monitoring for high tech agriculture farms.

ACKNOWLEDGEMENTS

This work was supported by Singapore Polytechnic and School of Mechanical and Production Engineering, Nanyang Technological University and the Ministry of Education, Singapore, through grant MLC1/97.

REFERENCES

1. Oi Wah Liew, William SL Boey, Anand K Asundi, Jun Wei Chen, Duo-Min He, *Fiber Optic Spectra-Photometry (FOSP) monitoring of plant nutrient deficiency in hydroponic culture condition,* SPIE Vol. 3740, June 1999, p. 186-190
2. C. Bould, E.J. Hewitt, and P.Needham, *Diagnosis of Mineral Disorders in Plants (Volume 1)*, Her Majesty's Stationery Office, London, 1983.
3. A. Scaife and M. Turner, *Diagnosis of Mineral Disorders in Plants (Volume 2)*, Her Majesty's Stationery Office, London, 1983.
4. G. Ouzounidou, "The use of photoacoustic spectroscopy in assessing leaf photosynthesis under copper stress: correlation of energy storage to photosystem II fluorescence parameters and redox change of P_{700}", *Plant Science*, 113(2), pp. 229-237, 1995.
5. S. Ciompi, E. Gentili, L. Guidi, and G.F. Soldatini, "The effect of nitrogen deficiency on leaf gas exchange and chlorophyll fluorescence parameters in sunflower", *Plant Science*, 118(2), pp. 177-184, 1996.

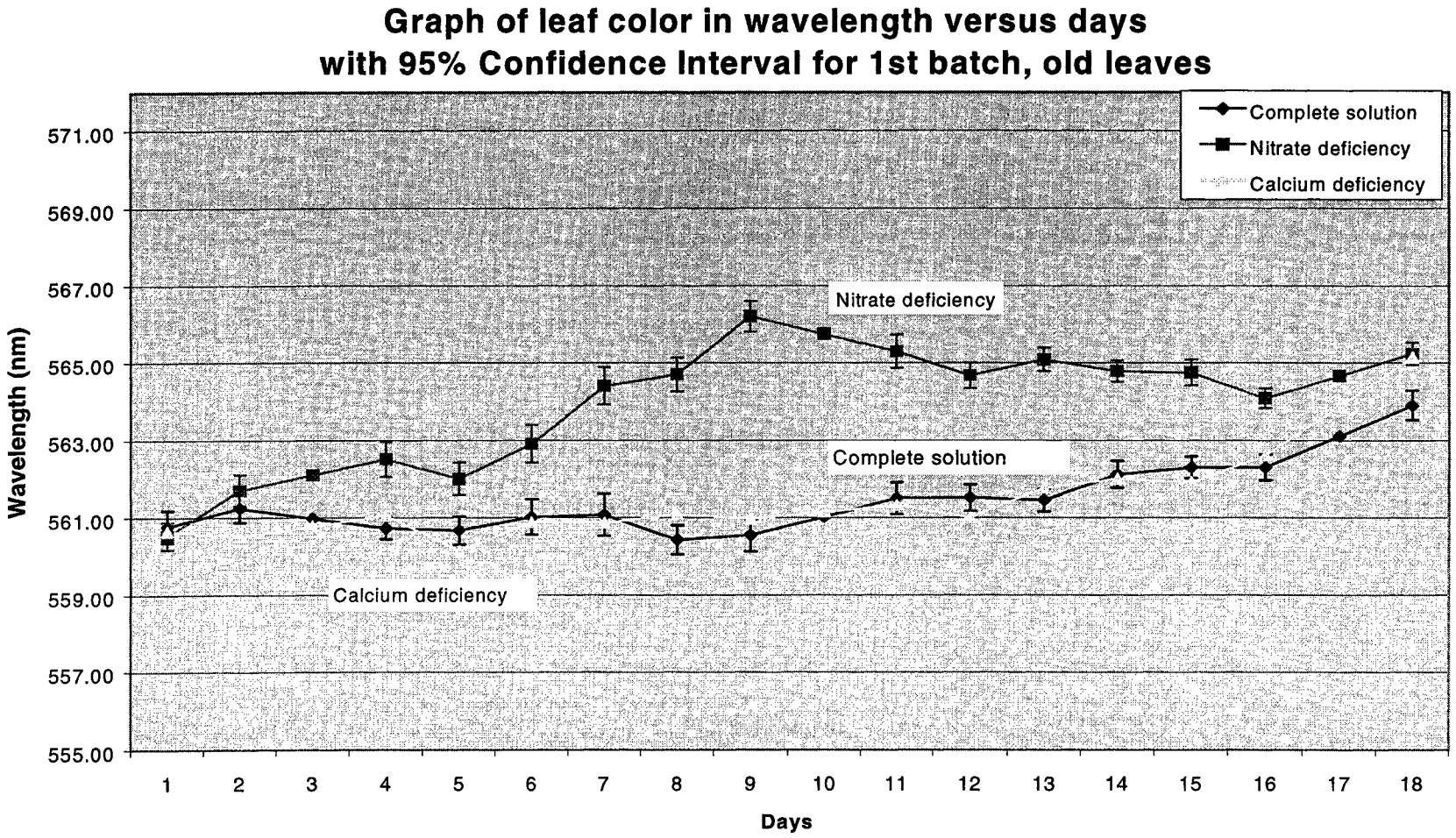


Figure 6 Graph of old leaves reflectance against number of days

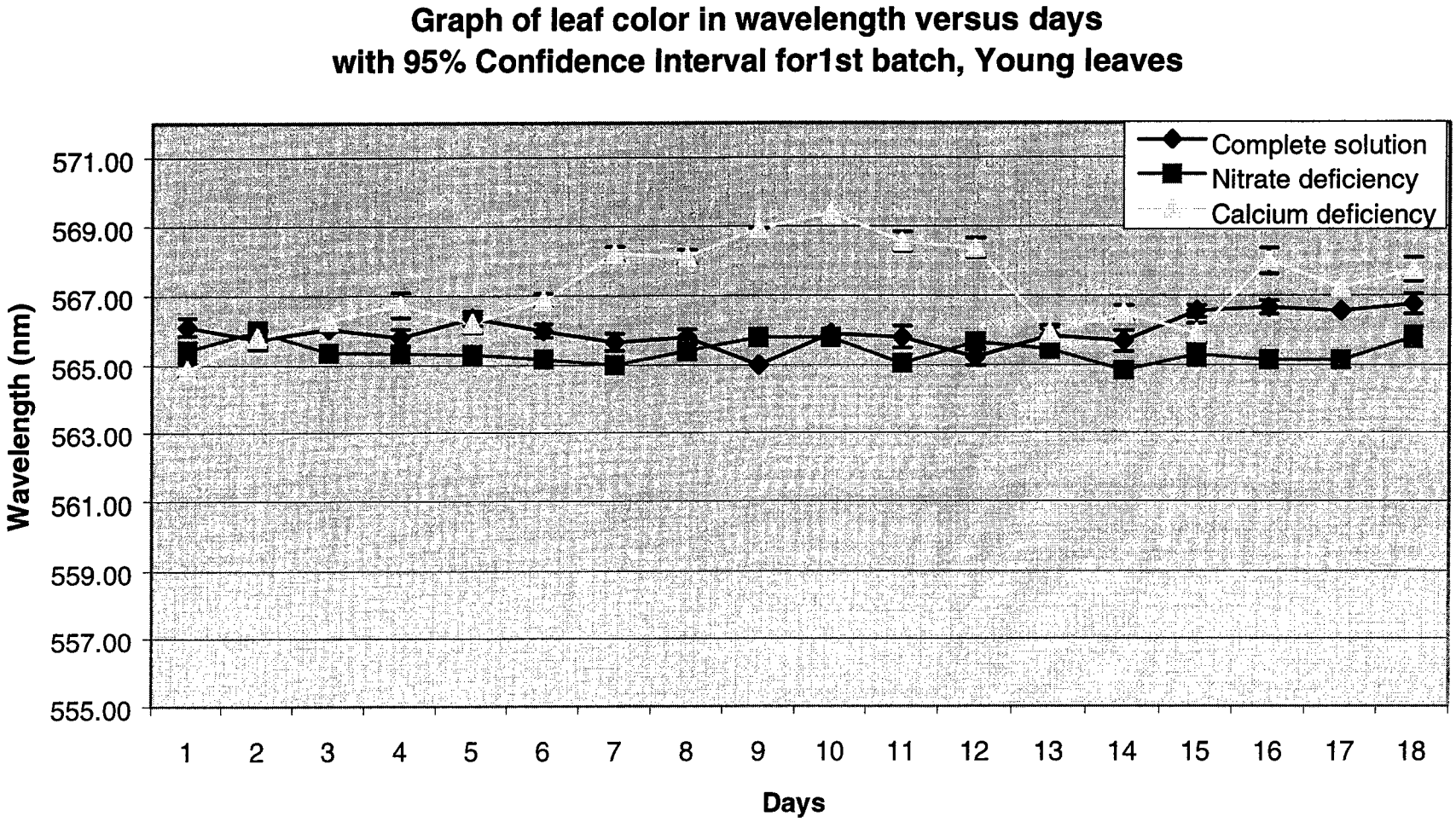


Figure 7: Graph of new leaves reflectance against number of days

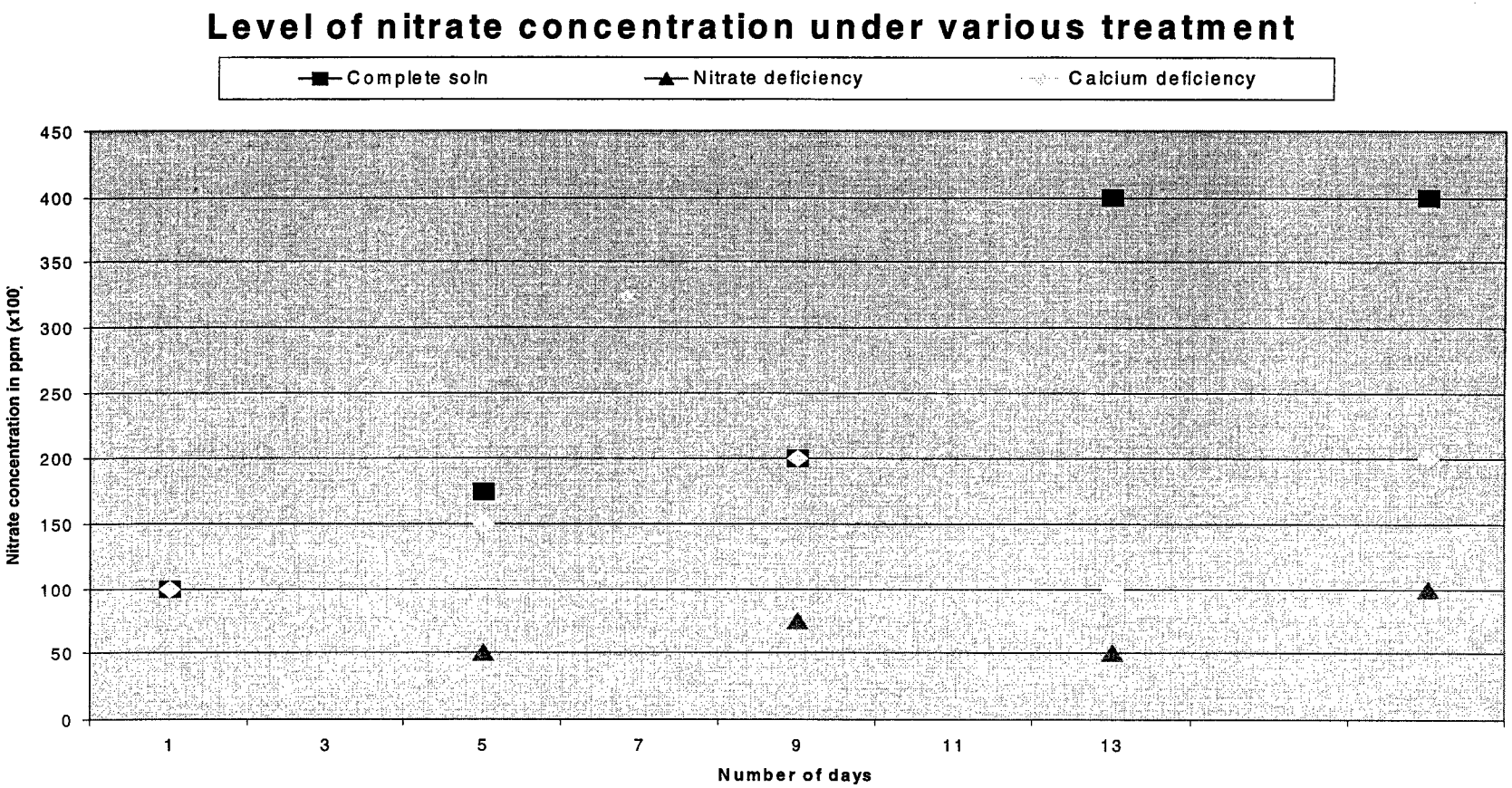


Figure 8: Graph of nitrate concentration subjected to various treatments verse days

Design of the Integrated Sensor System for Two-dimension Position Based on CCD

LIU Xiangdong* GU Yuzhang

Optical Engineering Department, Zhejiang University, Hangzhou, 310027

ABSTRACT

In this Article, a kind of design philosophy of a integrated sensor system for two-dimension position is presented. The sensing unit of this system is a two-dimension CCD (Charge-Coupled Device). Through increasing the signal-to-noise ratio of a constructed specific lightspot target and performing optical filter, the exported full video image signal of CCD which includes a group of isolated positive pulse caused by the target is formed. Through combinational logic and sequential logic processing to the full video signal, the target image's two-dimension position on the light sensitive layer of CCD can be generated. Finally, the specific lightspot target's relative two-dimension position in the visual field is gained. In addition, a typical application of the system in the computer information processing field are presented as well.

Keywords: Two-dimension Position, CCD, Full Video Signal, Specific Lightspot Target

1. INTRODUCTION

This Integrated sensor system for two-dimension position is a kind of equipment that can locate and trace a lightspot target in a two-dimension plane. The kernel of system is a photoelectric device. The following is the principle of system: First, the target is imaged on the photoelectric device's light sensitive layer through optical imaging, then the system can process the output of the photoelectric device and ascertain the target image's geometric position, finally the two-dimension coordinate of the target is educed. Generally, there are several kinds of photoelectric device available, such as CMOS device, CCD, PSD (Position Sensitive Detector), Multielement Type Si Photodiodes, etc. To realize high precise locating and tracing, the CCD that has high distinguishability is used. In the system, both line array CCD and plane array CCD can be used.

If the system uses line array CCDs, it is necessary to design a simple optical subsystem, which can decompose, focus, and image the two-dimension position of the target. Fig.1 illustrates the principle of the system consisted of two one-dimension CCDs and an optical subsystem. 1 indicates the visual field plane coordinate ststem. 2 indicates the specific target. 3 indicates a color filter used to increase the signal-to-noise ratio. 4 indicates a imaging lens. 5 indicates a reflector (its reflectance ratio is 50%), 6 indicates a meridian cylinder lens. 7 indicates an arc cylinder lens. 8 indicates an line array CCD along the meridian direction. 9 indicates an line array CCD along the arc direction. Through the optical subsystem consisted of 4, 5, 6 and 7, the specific target's two-dimension coordinate in the visual field plane is separated and focused to form two one-dimension coordinate along the meridian direction and the arc direction. The two separated images are detected by two one-dimension CCDs respectively, following the signal processing, the target's two-dimension integrated coordinate is gotten. The advantage of this design is the high detection precise. But the design has some shortages, too. First, the target must be bright to some tune. Second, the cost is comparatively high.[1]

The design of the integrated sensor system in this article is based on a two-dimension CCD to realize locating and tracing the target. The principle is: the object space field plane coordinate system is imaged on the CCD's light sensitive layer by the optical system of the CCD, then the CCD's output signal named full video signal, which contains a set of sharp positive polarity pulses caused by the specific lightspot target, is exported. Through increasing the signal-to-noise ratio of the specific optical target and processing the video signal, the sensor system can directly export the target's two-dimension coordinate in the object space field, which makes it very convenient to locate and trace the target. The kernel of the sensor system designing is to abstract the line sync signal, the field sync signal and the specific optical target's signal from the full video signal. By enhancing the chrono-characteristic distinguishability of these signals with a circuit, the locating precision can be increased. Finally, a typical application of the integrated sensor system are mentioned.

* Correspondence: Email: xdliu@moi-lab2.zju.edu.cn; Telephone: 86-571-7951187; Fax: 86-571-7951617

Part of the SPIE Conference on Advanced Photonic Sensors and Applications
Singapore • November–December 1999
SPIE Vol. 3897 • 0277-786X/99/$10.00

2. SYSTEM PRINCIPLE

2.1. Image signal

Image signal is a kind of electrical signal that is transformed from the plane image information. The image signal whose electrical level changes along with time represents the image's two important information: brightness and space. In the image signal, the electrical level represents the image's brightness. There are two types of image signal: positive polarity image signal and negative polarity image signal. The former means that the higher the electrical level, the brighter the image's corresponding part. The latter means that the lower the electrical level, the brighter the image's corresponding part. A positive polarity image signal is showed in Fig.2. It is defined that the low electrical level (75%) corresponds to the image's black part, and the high electrical level (12.5%) corresponds to the image's white part. The former is called black electrical level, the latter is called white electrical level, and the electrical level between the former and latter is called gray electrical level. In our design, the required image signal is the specific lightspot signal. The brighter the lightspot is, the higher the corresponding electrical level is. Using proper comparator, we can separate the specific lightspot signal from the image signal.

2.2. Full video image signal[2]

To make the output image signal of CCD receivable to the television signal receiver, it must be combinated with a sync signal to form the full video image signal, and the full video image signal is transmitted in one channel. Consequently, the television's requirement for two-dimension scanning is fulfilled. The sync signal includes the line sync signal and the field sync signal. To make a difference between the image signal and the sync signal, the sync signal is generally added in the image signal when the image signal's blanking appears. Moreover, the electrical level of sync signal is placed between the electrical level range where the image signal can not reach. In Fig.2, the sync pulse electrical level of full video signal stays at the bottom of sync signal (100%). So the sync signal can be separated solely, and the image signal will remain unaffected. Because of the difference between the line sync signal's pulse width and that of the field sync signal, we can use a integrating circuit to distinguish the two kinds of sync signal. The slotting in field sync signal can maintain the line sync during the period of field sync.

By analyzing the full video signal exported from the CCD, we know that it is possible to design a proper circuit to take several signals apart, which include the specific lightspot signal, the line sync signal, and the field sync signal. In line by line scanning CCD, one frame is consisted of two field, and the frame sync signal can be attained through dividing frequency of the field sync signal. To get the specific target's position coordinate, another signal named metafrequency signal is required. The distinguishability of metafrequency decides the X direction pixel resolution directly. If the metafrequency signal's periodic time is T_V, the line sync signal's periodic time is T, and the pixel number along X direction is N, then:

$$T_V = \frac{T}{N} \tag{1}$$

2.3. System construction

The specific lightspot signal, the line sync signal, and the field sync signal can be separated through processing the full video signal. Using these signals, we can construct the integrated sensor system for two-dimension position based on CCD. The system can detect the two-dimension coordinate of the specific lightspot and transmit the coordinate to the computer. Fig.3 illustrates the detailed block diagram of the system.

The specific lightspot target is focused on the CCD's light sensitive layer through the CCD' s lens. The filter lens is applied to enhance the signal-to-noise ratio of the specific lightspot target in the background. The intensity distribution of the specific lightspot target's image has character of point spread function. The diffusion area covers some pixels of different rows and columns. For example, in row m, one pixel; in row m+1, two pixels; ... ; in row n, k pixels; in row n+1, k-1 pixels; etc. Obviously, row n has the maximum pixels that are covered by the specific lightspot's image. In other words, row n runs through the center of the lightspot's image. The envelope curve of the full video signal of the specific lightspot's image in one field appears approximate normal distribution, which is showed in Fig.4. If we set a comparison threshold value in the image range of the full video signal, it is easy to filter out the specific lightspot's image signal. In one field, the first light sensitive unit on which the electrical value of the detected specific lightspot's image signal exceeds the threshold value

shows the position of the specific lightspot target. The specific lightspot image filter circuit can be normalized to a control threshold to end counting. In the design, there is fixed error between the detected position and the specific lightspot's center. The fixed error can be amended through software simulate. Using the specific lightspot's image signal, we can get the line sync signal and the field sync signal. The distinguishing metafrequency signal is generated from the pulse generator. So we can get the specific lightspot's two-dimension position coordinate in the two-dimension CCD. The method is explained as following:

The metafrequency signal is used as count signal, the specific lightspot signal is used as end count signal, and the line sync signal is used as reset signal, then these signal can control a counter to generate the X coordinate. Concretely, the counter is reset at the beginning of each line scanning, then the count value is ascended each metafrequency period, until the specific lightspot signal appears, and the count value is the X coordinate. The X coordinate count resolution lies on the distinguish value about the metafrequency signal period to the CCD's pixel drive period.

The line sync signal is used as count signal, the specific lightspot signal is used as end count signal, and the field sync signal is used as reset signal, then these signal can control a counter to generate the Y coordinate. Concretely, the counter is reset at the beginning of each field scanning, then the count value is ascended each line scanning period, until the specific lightspot signal appears, and the count value is the Y coordinate.

The higher the resolution of the specific lightspot visual field plane, the more the needed coordinate counter bits. For example, to a visual field which resolution is 1024*768, the counters of X and Y coordinate need at least 10 binary counting bits respectively (2^{10}=1024, 2^9=512). Taking it into account that the CCD visual field area is wider than the specific lightspot target area, we cascade three 4-bit binary counters for X coordinate counting and other three 4-bit binary counters for Y coordinate counting. Both of them provide 12-bit counting ability.

The coordinate synthesis circuit is based on an 8-bit PIC microprocessor produced by the Microchip Corporation. The PIC microprocessor has technical and economical advantages. One I/O port is used as the entrance of outer interrupt signal, which can receive the outer interrupt caused by the end count signal. When the outer interrupt occurs, an interrupt service program is invoked, then the program realizes the coordinate synthesizing, coding, and serial sending. Another I/O port is used as the sending end of serial signal. Other 24 I/O ports of the PIC microprocessor are used as the two-dimension coordinate data sample channel.[3]

2.4 System elementary error

2.4.1. Count error

The count error of X coordinate lies on the using metafrequency signal period. If the period of metafrenquency is shorter than the CCD's pixel drive period, the count error of X coordinate is restricted to ± 1 pixel.

The scanning method of CCD causes the count error of Y coordinate. There are two kinds of scanning method: line-by-line scanning and alternate-line scanning. For the line-by-line scanning CCD, the count value attained through normal counting is the specific lightspot image's Y coordinate.

For the alternate-line scanning CCD, each frame of the video signal is consisted of one odd field and one even field. It is random to sample the video signal, which means that the first field of one frame can be either odd field or even field. Because the specific lightspot's image covers several lines, the specific lightspot signal exists in both odd field and even field of one frame. Therefore, the specific lightspot signal in the first field of each frame will be detected, and the end count signal of Y coordinate is generated at the same time. There are four counting cases. We can calculate the count error of Y coordinate in the four cases above respectively. If we suppose the error to be E, then E=Y-V. Four counting cases are discussed as following:

If the first field of one frame is the odd field, and specific lightspot's image first appears at an even line (line V), then the first detected specific lightspot's image must exist at line V+1. Accordingly, the count value of Y coordinate is V/2+1. Then:

$$Y = 2 \times \left(\frac{V}{2} + 1 \right) = V + 2 \qquad E = Y - V = 2$$

If the first field of one frame is the odd field, and specific lightspot's image first appears at an odd line (line V), then the first detected specific lightspot's image must exist at line V. Accordingly, the count value of Y coordinate is (V+1)/2. Then:

$$Y = 2 \times \frac{V+1}{2} = V+1 \qquad E = Y - V = 1$$

If the first field of one frame is the even field, and specific lightspot's image first appears at an even line (line V), then the first detected specific lightspot's image must exist at line V. Accordingly, the count value of Y coordinate is V/2. Then:

$$Y = 2 \times \frac{V}{2} = V \qquad E = Y - V = 0$$

If the first field of one frame is the even field, and specific lightspot's image first appears at an odd line (line V), then the first detected specific lightspot's image must exist at line V+1. Accordingly, the count value of Y coordinate is (V+1)/2. Then:

$$Y = 2 \times \frac{V+1}{2} = V+1 \qquad E = Y - V = 1$$

If we subtract 1 from Y in all cases, the count error of Y coordinate is restricted to ± 1 pixel.

2.4.2. Coordinate transformation error

To apply the two-dimension position sensor system, a premise must be satisfied. The premise is that the specific lightspot target must be limited in the area that is included completely in the CCD's vision field. We suppose that the CCD's vision field belongs to coordinate system X'O'Y', the rectangle area P-S-Q-R to which the specific lightspot is restricted belongs to coordinate system XOY. Obviously, there are shift and rotate relations between the two coordinate systems because the CCD is random placed. Therefore, another elementary error is caused when transforming the position coordinate. Fig.5 shows the two coordinate systems. In coordinate system X'O'Y', the specific lightspot target's coordinate is A' (A_X',A_Y'). In coordinate system XOY, it is A(A_X,A_Y).

Before the initialization of system, we can make certain that the resolution of rectangle area P-S-Q-R is M*N, and the top left corner P is regarded as the reference point. To realize the coordinate transformation, we must know at least three vertex coordinates of the rectangle area in coordinate system X'O'Y'. If we know the coordinates such as P'(P_X',P_Y'), R'(R_X',R_Y') and S'(S_X',S_Y') via initialization of system, then the following coordinate transformation formulae can be educed according to the triangle relation.

$$A_X = \frac{M \times (A_X' - P_X') \times (S_Y' - P_Y') + N \times (A_Y' - P_Y') \times (R_Y' - P_Y')}{(S_Y' - P_Y') \times (R_X' - P_X') - (R_Y' - P_Y') \times (S_X' - P_X')} \tag{2}$$

$$A_Y = \frac{M \times (A_X' - P_X') \times (S_X' - P_X') + N \times (A_Y' - P_Y') \times (R_X' - P_X')}{(S_Y' - P_Y') \times (R_X' - P_X') - (R_Y' - P_Y') \times (S_X' - P_X')} \tag{3}$$

According to error transfer function[4], if $\varphi = F(x, y, \cdots, u)$, then:

$$\delta\varphi = \frac{\partial F}{\partial x}\delta x + \frac{\partial F}{\partial y}\delta y + \cdots + \frac{\partial F}{\partial u}\delta u$$

In coordinate system X'O'Y', if the measuring error of point P, R, S, A is ($\Delta P_X'$, $\Delta P_Y'$), ($\Delta R_X'$, $\Delta R_Y'$), ($\Delta S_X'$, $\Delta S_Y'$), and ($\Delta A_X'$, $\Delta A_Y'$), then the transfer error ΔA_X and ΔA_Y is:

$$\begin{aligned} \Delta A_X = \frac{1}{V}\{ & M(S_Y' - P_Y')\Delta A_X' + N(R_Y' - P_Y')\Delta A_Y' + [A_X(S_Y' - R_Y') - M(S_Y' - P_Y')]\Delta P_X' + [A_Y(R_Y' - S_Y') \\ & - N(R_Y' - P_Y')]\Delta P_Y' - (S_Y' - P_Y')(A_X \Delta R_X' - A_Y \Delta R_Y') + (R_Y' - P_Y')(A_X \Delta S_X' - A_Y \Delta S_Y') \} \end{aligned} \tag{4}$$

$$\Delta A_Y = \frac{1}{V}\{ M(S_X'-P_X')\Delta A_X'+N(R_X'-P_X')\Delta A_Y'+[A_X(S_X'-R_X')-M(S_X'-P_X')]\Delta P_X'+[\ A_Y(R_X'-S_X') - N(R_X'-P_X')\ ]\Delta P_Y'-(S_X'-P_X')(A_X\Delta R_X'-A_Y\Delta R_Y')+(R_X'-P_X')(A_X\Delta S_X'-A_Y\Delta S_Y')\ \} \quad (5)$$

In (4) and (5), $V = (S_Y'-P_Y')(R_X'-P_X')-(R_Y'-P_Y')(S_X'-P_X')$. (6)

According to (4) and (5), It is evident that the locating and tracing precision is related to the locating precision of P, R, S during system initialization.

2.4.3. Optical System Imaging Error

For an idealized optical system, the amplification ratio on a pair of conjugate planes is a constant. But for an actual optical system, this conclusion is exactly correct only if the visual field is small enough. When the visual field is comparatively large, the image amplification ratio changes corresponding to different position of the visual field. In our designing, a low cost 8~12mm tight shot is used. The main aberrational correction of this shot is to reduce the axial aberration to increase the imaging definition, but the radial aberration is insufficiently corrected.

In the system, the detect visual field of CCD is so large that a elementary error based on the distortion aberration is generated when detecting the position of specific lightspot target. In the center area of the visual field, the distortion aberration is chiefly based on the primary distortion, and it is scaled with the visual field's cube. In the peripheral part of the visual field, the contribution of high factorial factors to the error is gradually increased along with the enlarging of the visual field.

2.5. Experimental System Construction

According to above mentioned principle and technical line, a experimental system is constructed to actually detect the related position of the specific lightspot target. The following is the parameters of the experimental system:

The specific lightspot target: LED (Light Emitting Diode) array. Its distribution is illustrated in Fig.6;
Total number of the CCD's pixels: 542(H)*582(V);
The imaging lens' focal length: 8mm;
The experiment working distance: 2.5m;
The rectangle area: to which the specific lightspot target is restricted: 1.6m*1.2m
The preference point: The left top corner of the target rectangle area.

During the actual measurement, we use the coordinate of three points - P, R, S – to realize the initialization of the system. The actual measurement result of the integrated sensor system for two-dimension position is showed in Fig.7.

3. CONCLUSIONS

When using the integrated sensor system to detect the specific lightspot target, for the reason of positive distortion aberration, the position error is lesser in the areas around the preference point and the visual field center than in the areas away from the preference point and the visual field center. The related error is about 1% in the peripheral part of the visual field. In practice, the error is generated as a monotone function of the dimension of the visual field, therefore it can be eliminated through the method of software compensation.

4. TYPICAL APPLICATION

The photoelectric virtual blackboard system is a typical application of the 2-dimension position integrated sensor system based on CCD. In the application, operator can locate and trace the cursor in the large-screen display plane with a specific lightpen, which is equipped with a point source of light at the top. Fig.8 shows the sketch map of the photoelectric virtual blackboard system.

Along with the technical progress, Large-screen display has been the mainstream of display technique of the day. The large-screen projector holds vast marketability in both military and commercial area. It is especially suitable to the one-to-many information communication occasions such as CAI (Computer-Aided Instruction).[5,6,7] The photoelectric virtual blackboard system is based on the large-screen display technique and the interactive video processing technique. The designed specific lightpen acts as the chalk, and the point source of light at the top of the lightpen acts as the specific lightspot target to be detected by the 2-dimension position integrated sensor system. The virtual large-screen display interface acts as the blackboard, and the physical medium of the virtual interface could be a projection screen, a piece of wall, a back project display screen, etc. The operator can expediently trace on the large-screen display interface with the specific lightpen, just as using the chalk and blackboard.

In Fig.8, 1 indicates the large-screen display plane, and it's content is provided through direct project or back project by the projector 3, which is controlled by the computer 4 through the VGA video line 7. The large-screen is synchronistically displayed with the monitor 5. 2 indicates the specific lightpen, at the top of which is a controllable infrared point source of light. Using the infrared point source of light as the specific spotlight target, we can gain enough signal-to-noise ratio of the optical target in the high brightness visible light display interface. 8 indicates the 2-dimension position integrated sensor including a CCD. Optionally, 8 and 3 can be integrated to one unit. The CCD captures the position information of the specific lightpen through optical imaging, then the information is sent to the computer through the serial communication line 9 after being processed and encoded. The human-machine interact procedure is supported by the application software.

In the sensor design of the specific lightpen, the lightspot's light and dark respectively correspond to the brush-write status and the non-brush-write status. Therefore, a series of two-dimension position coordinates which represent the write trace of the specific lightpen are generated by the integrated sensor system, then the drive software supports to realize real-time locating and tracing of the cursor. The drive software is based on the Message Driving Mechanism of Microsoft Windows. To paint on the display interface, we invoke GDI (Graphic Device Interface) functions of Windows API (Application Programming Interface) in the software. By calling these API functions we can manage the GDI paint tools to paint on the display interface.

Based on the construction of above typical application system, two additional functions can be added to it. First, the specific lightpen can be made to simulate the functions of mouse. The simulative mouse instructions and the position information of the specific lightfpen can be synthetically encoded and sent to the computer, then the specific lightpen becomes an ideal substitute of the mouse. Second, if a set of pattern recognition software is worked out to recognize the write trace of the specific lightspot, the sensor system can be employed as a handwriting input equipment of the computer to input characters directly. Both of the two additional functions will make it convenient to handle the system for the operator, therefore a more advantageous interact mode for the one-to-many information communication occasions based on large-screen project display is available.[8]

5. REFERENCES

1. Liu Xiande, *CCD and Its Application Principle*, Central China University of Science and Technology Publish House, 1990
2. Luo Guixiang, Wang Liangzhong, Huang Zhuoxun, Liu Xiehe, *Principle of Television Receiver*, People Education Publishing House, 1981
3. Gu Yuzhang, Liu XiangDong, and Liu Xu, "PIC Signal-chip Microprocessor Applied in Design of Integrated 2D Position Sensor", Journal of Zhejiang University, Supplement, Vol.**33**, 1999
4. Xiao Mingyao, *Error theory and its Applications*, pp.92-93, Measuration Publishing House, 1985
5. Keiichi Kubota, Masao Imai, and Setsuo Kaneko, "High Resolution High Brightness Liquid Crystal Projector for Work Station", SPIE Vol.**2407**, pp.119-124
6. R.Gerhard.Multhaupt, G.Mahler, Light-Value Projection Displays Introduction, Displays Vol.**16**, No.1, pp.5-7, 1995
7. Gong Xianling, *Automatic Display Technology*, Metallurgy Industry Publishing House, 1989
8. Ben Shneiderman, *Design of Multimedia User Interface —— Effective Policies of Human-Machine Interface*, Shanghai Science Propagation Publishing House, 1995

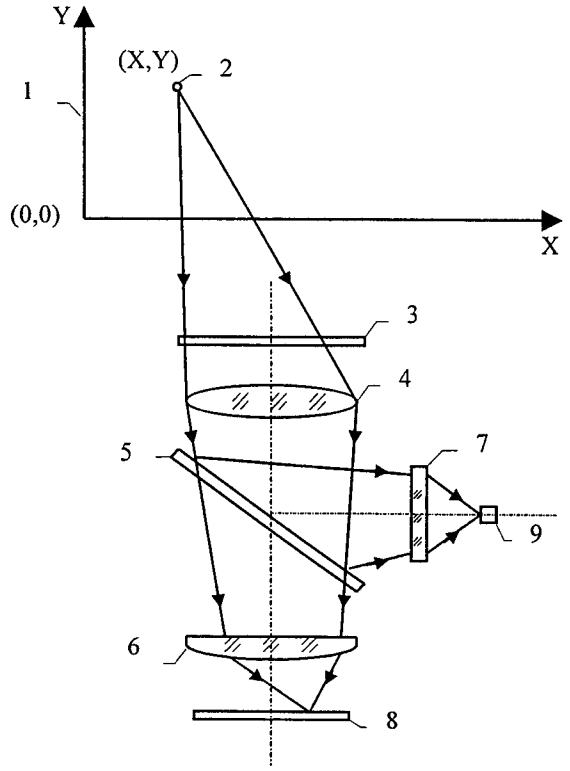


Fig.1 Two-dimension focus decompose imaging subsystem

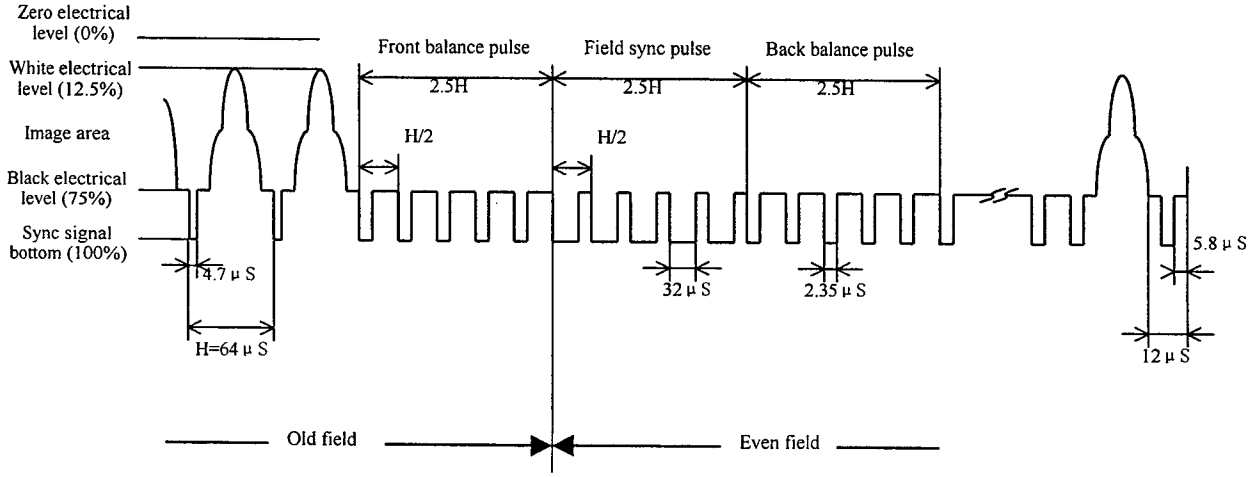


Fig.2 Positive polarity full video image signal

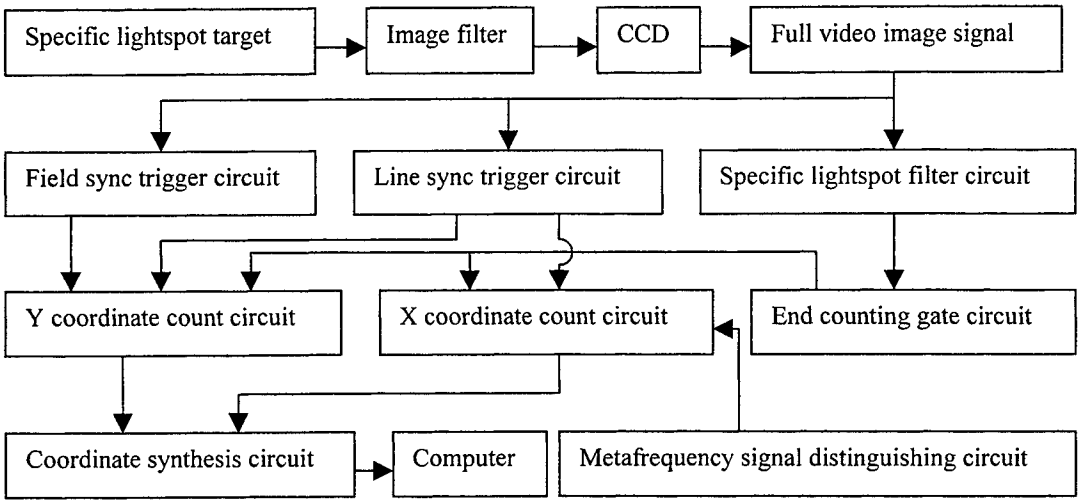


Fig.3 Block diagram of integrated sensor system

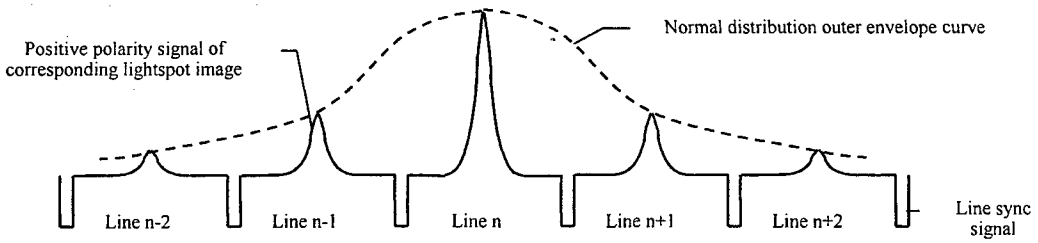


Fig.4 Positive polarity pulse space normal distribution character of lightspot image in one field

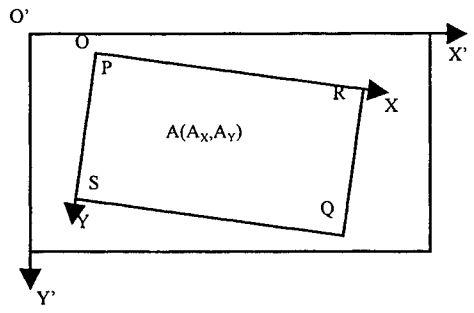


Fig.5 Position coordinate transformation between two coordinate systems

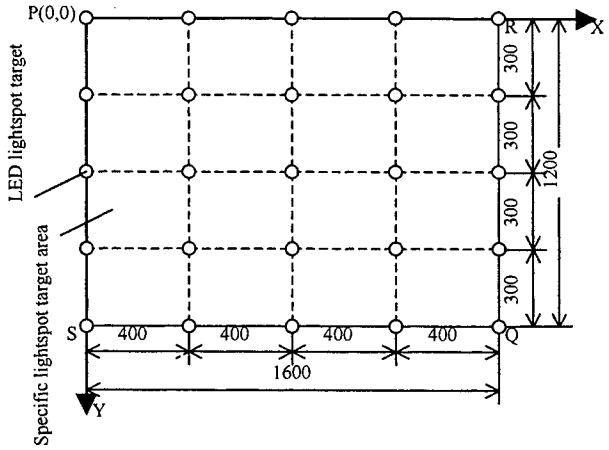


Fig.6 Design of specific lightspot target array

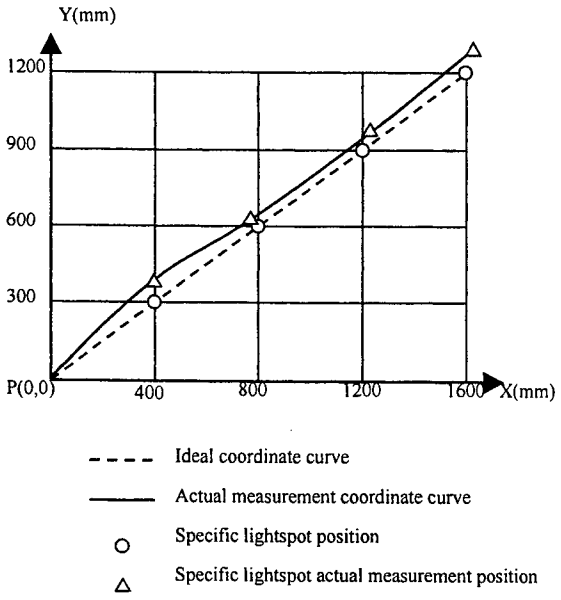


Fig.7 Result of actual measurement

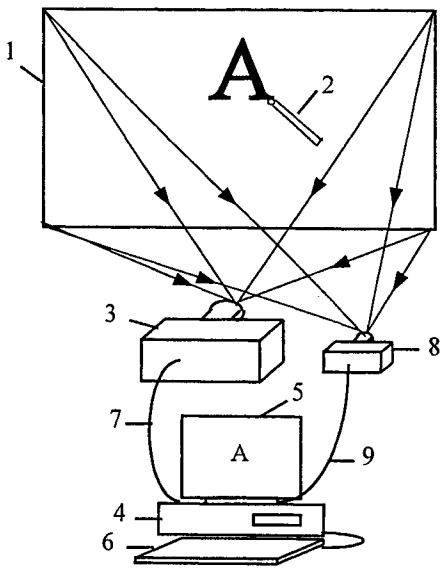


Fig.8 Sketch map of photoelectric virtual blackboard system

SESSION 5

Micro-optical Metrology

Invited Paper

Strain measurement in micrometrology

Dietmar Vogel*[a], Rolf Kuehnert[b], Bernd Michel[a]

[a]Dept. Mechanical Reliability and Micro Materials, Fraunhofer Institute IZM, Berlin, Germany

[b]K&T Measuring Systems GmbH, Chemnitz, Germany

ABSTRACT

New generations of microelectronics and microsystem devices call for the utilization of a variety of new materials and the combination of materials with a widespread of their mechanical and thermal characteristics. Related thermo-mechanical reliability issues as well as functional aims bear a big challenge to understand and to design mechanical behavior of devices and components. Finite element analysis of components and devices is widely utilized in the R&D process as a tool of predictive engineering. Unfortunately, uprising miniaturization, higher system complexity and advanced material application more often question mechanical modeling for simulation. So, a strong need exists to measure simultaneously stresses and strains on real components. At the IZM Berlin different classical kinds of laser methods have been used to determine deformation behavior of components due to thermo-mechanical load and functional performance of systems. The principal restriction of those methods with regard to spatial and measurement resolution had forced the development of new methods. In particular, microDAC a displacement and strain measurement method has been established. It bases on image processing algorithms permitting to extract strains from load state micrographs, which are picked up by SEM, optical microscopy and other high resolution imaging devices. Besides a brief review, corresponding interferometry and Micro Moiré applications of the authors to micro deformation analysis, this paper mainly focuses on the microDAC technique. The basics of the microDAC concept, its main features like measurement resolution, data presentation and coupling to finite element analysis as well as applications are presented. Different kinds of measurements on modern electronics structures, like, e.g., used in flip chip technology, chip scale packaging and optoelectronics, illustrate its capability. Finally, a discussion of possible future developments is added, which can be directed to the adoption of higher resolution imaging equipment, namely scanning force microscopy.

Keywords: displacement, strain measurement, microtechnology, electronics packaging, correlation technique, microDAC, Micro Moiré

1. INTRODUCTION

Reliability and functionality of microsystems, i.e. of small scale integrated electronic, mechanical and optical components, largely depend on their mechanical and thermal constitution. Thermo-mechanical aspects of component and system reliability become more and more important with growing miniaturization as the local physical parameters and field quantities show an increase in sensitivity due to inhomogeneities in local stresses, strains and temperature fields. Since there is usually a lack of information about the local material parameters, a pure field simulation cannot, as a rule, solve the problem. The state of the art of microsystem design more and more requires direct „coupling" between simulation tools (including e.g. FE modeling) and advanced physical experiments.

The authors have combined various laser techniques, scanning microscopy, thermography, and acoustomicroscopy etc. with advanced numerical field simulation. The input data from the experiment are directly fed into the simulation codes to get more realistic results for reliability estimation and life time prediction. The investigations have been directed towards electronic packaging, crack and fracture behavior („fracture electronics") within the interconnected regions and to other problems, e.g. occurring due to thermal „misfit" between different materials within the critical regions of microsystems[1,2].

Many problems of defective microsystems can be related to stresses that develop at various stages of microsystem processing. These problems will be more acute as microsystems become more complex in geometry and in the number of

* Correspondence: Gustav-Meyer-Allee 25, D-13355 Berlin, Germany, Email: d_vogel@izm.fhg.de, Telephone: +49-30-46403214, Fax: +49-30-46403211

Part of the SPIE Conference on Advanced Photonic Sensors and Applications
Singapore • November–December 1999
SPIE Vol. 3897 • 0277-786X/99/$10.00

compounded materials. Detailed information on the local deformation fields within critical regions becomes more and more important. Therefore, the role of advanced field measurement techniques based upon modern image processing has been considerably grown in recent years. These methods in combination with advanced simulation tools provide very powerful means for reliability assessment and design optimization as well[3-5].

As an example, Fig. 1 schematically shows some of the driving forces in microelectronics packaging industry. Desired higher packaging density, device miniaturization, higher power dissipation levels, higher clock frequencies, more severe environmental impact even for consumer products and low cost requirements have initiated the development of a variety of new packing concepts. Their realization has led to new issues in thermo-mechanical reliability. Consequently, some of the needs for future strain measurement methods can be deduced from data of Fig.1. Generally, these methods should

- provide access to microscopically small areas. E.g., strain measurement must be possible inside electrical interconnect structures, which size of less than 50 µm is defined by future chip-to-board interconnect pitches,
- be applicable to quite different materials with various surface structures gathered in advanced electronics packaging assemblies,
- allow measurement under environmental impact, like temperatures below and above room temperature, as well as under humidity and harsh environmental conditions,
- be suitable for defect detection and measurements of strains on components with interfacial delamination, cracks and voids,
- be compatible to finite element simulation tools,
- and should allow analysis for complex macro devices as well as for local parts of them.

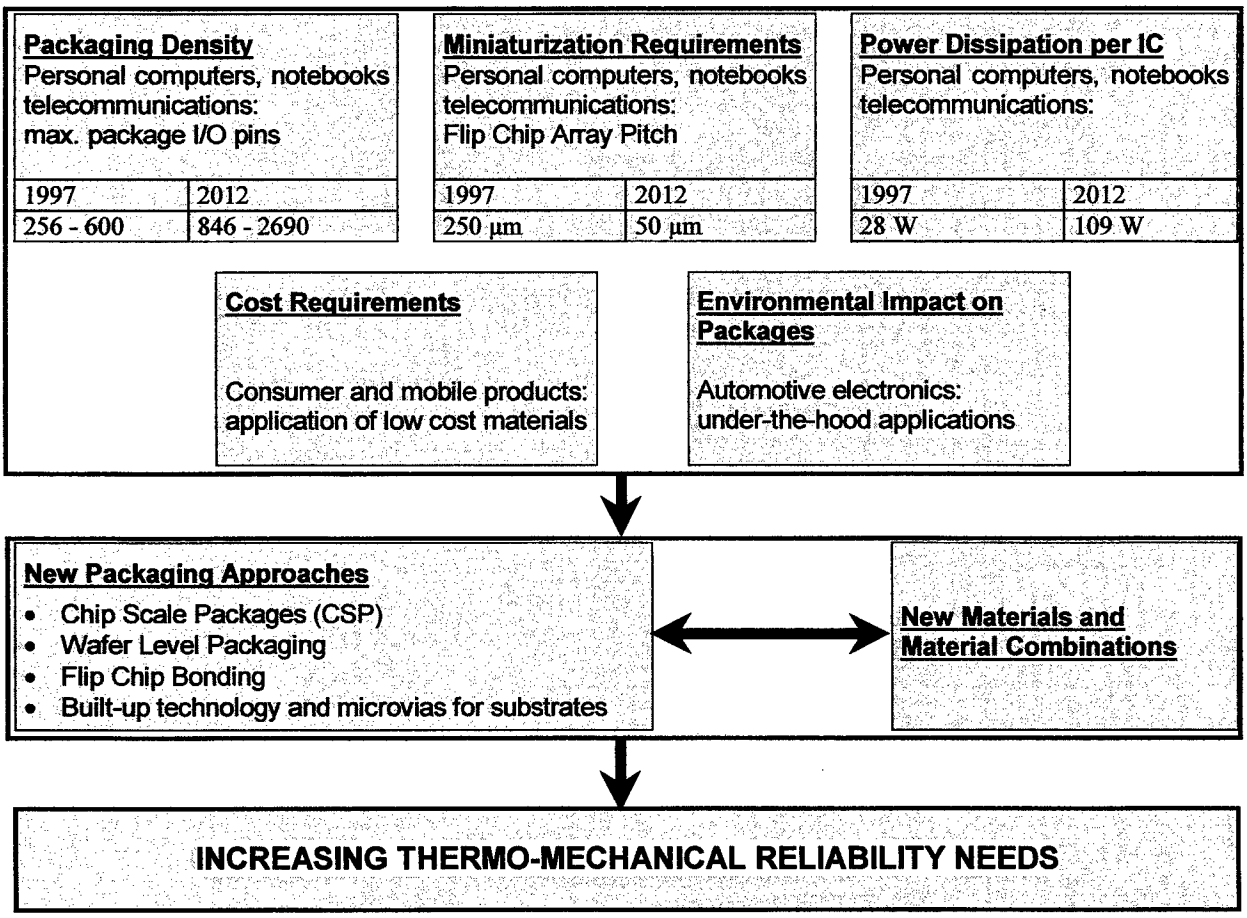


Fig. 1: Some aspects of electronics packaging developments (numbers based on "The National Technology Roadmap for Semiconductors" of the Semiconductor Industry Association (SIA)[6])

This paper mainly is dedicated to correlation based methods of strain measurement at surfaces of micro components. Nevertheless, a brief overview of other approaches is given in chapter 2. Chapter 3 and 4 refer to the application of interferometric and Micro Moiré techniques. Chapter 5 to 7 more in detail deal with the application of correlation analysis.

2. EXPERIMENTAL MEASUREMENT APPROACHES TO ANALYSE MECHANICAL STATES IN MICRO OBJECTS

Currently, for measurement of strain and stress states on real electronics devices and components only a couple of methods is being used. At present silicon piezoresistive stress sensor test chips[7], Raman spectroscopy[8], fiber sensors[9] and optical methods for strain measurement[10,11] are utilized. Among the optical methods the Moiré Interferometry is one of the most applied tools to electronic packaging, namely in the US inspired by pioneers of Moire techniques[12]. Different laser methods have been applied to measure local material characteristics like e.g. coefficients of thermal expansion (CTE)[13,14].

Rapid miniaturization of microelectronics and microtechnology structures within the last years demands a twofold improvement of existing measurement techniques. On the one hand, integrity of smaller and smaller object structures is responsible for mechanical reliability. In order to have measurement access to them, they have to be resolved optically, in some cases with respect to non-flat surfaces. Consequently, conventional light optical imaging runs out of its own capability. On the other hand, mechanical deformation behavior in a deep submicron or a nano region of displacement values has to be controlled to guarantee functional properties of components and mechanical reliability, too. Measurement accuracy of most conventional optical methods becomes insufficient.

At the IZM Berlin effort has been made to overcome the problem in two directions. Two well established optical measurement techniques have been transferred to other imaging methods not basing on optical light. Firstly, Moiré methods have been introduced to scanning electron microscopy[15]. Photolithographically created crossed fringe pattern are used as object grids. The digitally scanned electron beam serves as a reference grid. Imaging resolution as well as measurement resolution could have been improved significantly.

Secondly, speckle photographic basics and software tools have been utilized to extend in-plane deformation measurements to small sized objects non accessible before. This approach has been established last years as the microDAC measurement technique[16,17]. MicroDAC stands for **micro D**eformation **A**nalysis by means of **C**orrelation based algorithms. Mainly the high resolution of scanning electron microscopes allows to measure local strain fields inside such small structures as solder interconnects of characteristic size of 50 µm and less. Furthermore, excellent down scaling capability for displacement measurement permits to determine relative movements of microscopic parts within the nano region. A transition to highest resolution equipment, e.g. to AFM, STM or SNOM, possesses the potential to measure with subnano accuracy. First attempts for „nanoDAC" tests have been accomplished.

3. APPLICATION OF CLASSIC INTERFEROMETRIC METHODS TO A MICRO ACTUATOR DEFORMATION STUDY

Compared to MCM's (multi chip modules) in classical microelectronics, one typical development of microsystem technology (MEMS or MOEMS) is the integration of different function (sensors, actuators, signal processing) in one system or array configuration.

In the reported case silicon micromechanical arrays are designed to realize optical actuator functions in combination with microelectronic functions. Under these conditions, the mechanical characteristics of the whole array structure and also of the separate mirror component have direct influence on the performance of the microsystem.

Interferometric methods are useful experimental tools to determine the displacement fields of micromechanical arrays in dependence on the technological manufacturing process and under static as well as transient loading conditions.

Some selected results are shown in the following Figure 2. The figure contains the design of the whole scanning actuator array and of a separate mirror. Measurements have been carried out by Michelson and by Mirau interferometry. The measurement results for different steps of array manufacturing (see Michelson fringes left lower side) indicate that one-side mirror coating introduces the major warpage, while second-side coating leads to an additional mirror inclination. The same interferometric methods can be used to check the electro-statically caused mirror rotation (right upper side) as well as the mirror bending due to warming up, possible, e.g., by light absorption (left lower side).

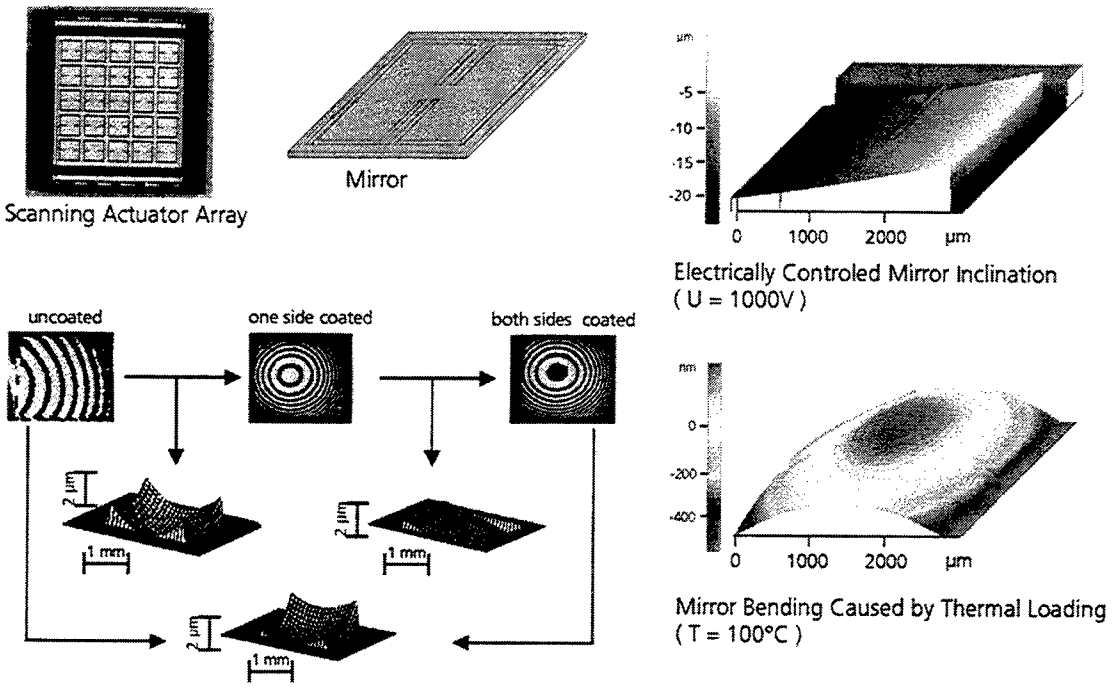


Fig. 2: Interferometric measurements on mirror elements of scanning actuator arrays (by courtesy of Volker Grosser, IZM Berlin)

4. SEM BASED MICRO MOIRÉ TECHNIQUE FOR HIGH RESOLUTION STRAIN MEASUREMENT

The transfer of optical Moiré principles into a scanning electron microscope opens the opportunity to improve spatial resolution by higher magnification factors. Strains inside smaller surface regions can be analyzed. Moreover, measurement resolution itself can be enhanced by increased pitch values of the Moiré grid.

For the measurement of strains around a crack tip in a CT crack specimen Moiré cross grids were produced photolithographically. The digitally scanned electron beam served as a reference grid. Fig. 3 shows the original cross grid on the specimen and illustrates the Moiré effect after some crack opening.

Simultaneously to the SEM measurement synthetic Moiré fringes were computed from finite element simulations for the crack tip area under load. They have been compared with the captured fringes from the experiment. So appropriate mechanical modeling could be ensured. Fig. 4 reproduces the corresponding displacement component perpendicular to the crack boundary for different load levels applied.

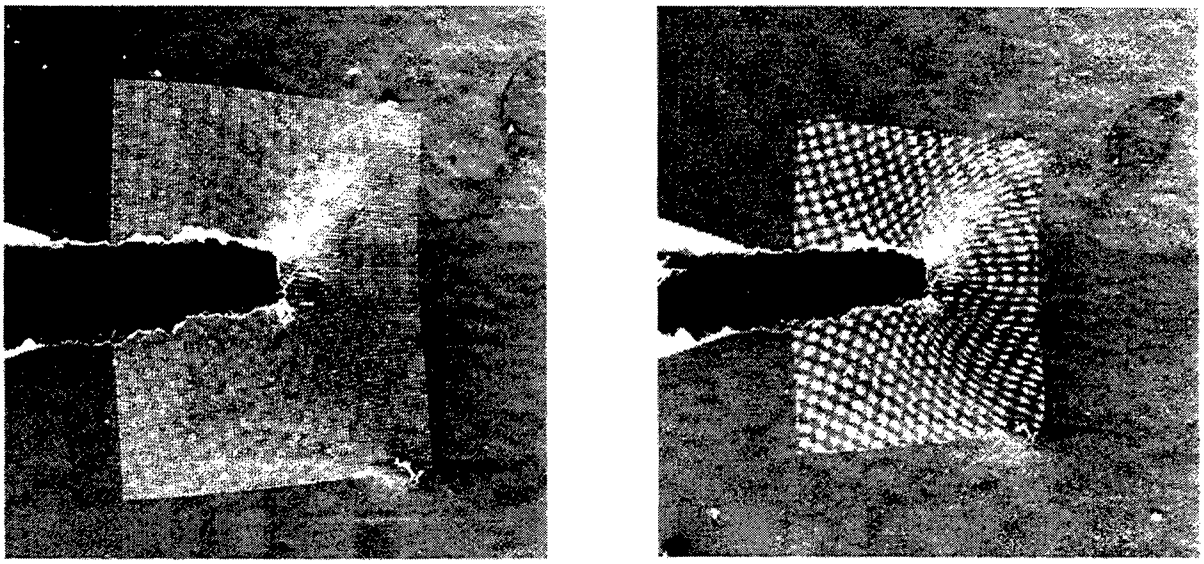


Fig. 3: Micro Moiré Measurement inside SEM (crack specimen), left - cross grid over the crack tip region, right - Moiré fringes due to grid distortion after crack opening

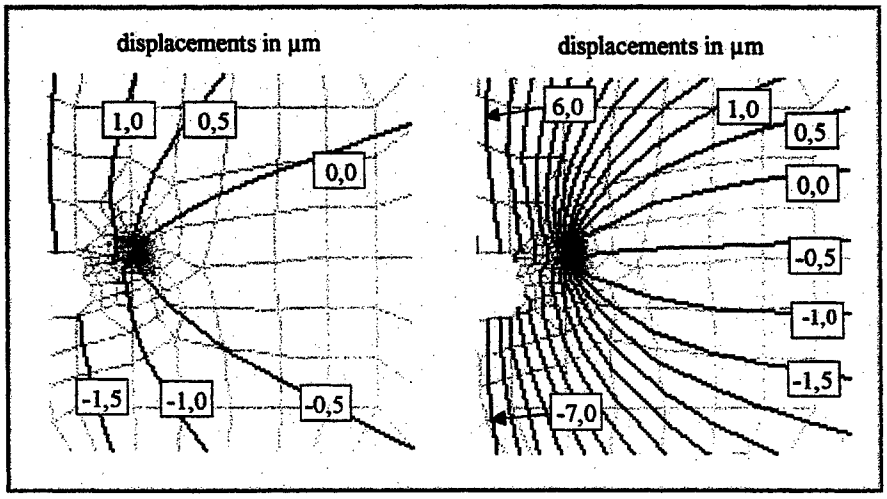


Fig. 4: Crack tip displacement field, u_y-component [in µm] for load line displacements 34,8 µm and 85,9 µm

5. STRAIN MEASUREMENT BY CORRELATION TECHNIQUES – AN APPROACH WITH UNIQUE DOWN SCALING CAPABILITY

5.1. The microDAC Concept

The basic idea of the underlying mathematical algorithms proceeds from the fact that microscopes of different kinds commonly allow to record tiny local and unique object pattern, besides of the more global object shape and structure. If these pattern are temporary of stable position and appearance in the digitized images, they can function as a local image area marker. Commonly, the microstructures even remain stable during severe thermal and/or mechanical component load and

can be recognized after load. A correlation based image processing algorithm is applied to determine a set of local pattern displacements between two object states. I.e., finally the object-side in-plane displacement fields $u_x(x,y)$ and $u_y(x,y)$ are measured, taking into account imaging magnification. Subsequent numerical derivation of displacement fields results in in-plane strain components

$$\varepsilon_{xx} = \frac{\partial u_x}{\partial x}, \varepsilon_{yy} = \frac{\partial u_y}{\partial y}, \varepsilon_{xy} = \tfrac{1}{2}\left(\frac{\partial u_x}{\partial y} + \frac{\partial u_y}{\partial x}\right) \tag{1}$$

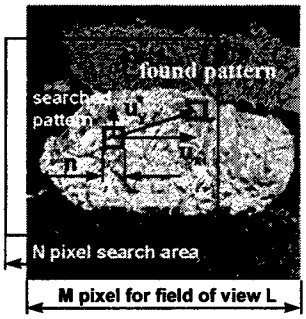
and the local rotation angle

$$\rho_{xy} = \tfrac{1}{2}\left(\frac{\partial u_x}{\partial y} - \frac{\partial u_y}{\partial x}\right) \tag{2}$$

The maximum of the set of local cross correlation coefficients

$$C(i',j') = \left(\sum_{i,j=1}^{n} I_1(i,j) I_2(i+i', j+j')\right) \left(\sum_{i,j=1}^{n} I_1^2(i,j) \sum_{i,j}^{n} I_2^2(i,j)\right)^{-1/2} \qquad \text{with} \quad |i'|, |j'| \le \frac{N-n}{2} \tag{3}$$

is taken as a measure for the pixel shift of a small image pattern with size n by n pixels (see Fig. 5). In Eq. (3) $I_1(i,j)$ and $I_2(i,j)$ describe the gray values of a pixel i,j in the images picked up from object state 1 and 2. Relative pixel addressing i', j' with regard to the searched area is used in Eq. (3). Maximum search is carried out within a interactively predefined neighbor-hood (N by N pixels) of the original pattern in state 1.



Measurement accuracy:

$\delta u_{x,y} = 0.1\ L / M$

Typical solder joint SEM micrographs : L = 100 µm, M = 1024

- $\delta u_{x,y} = 10$ nm
- $\delta\varepsilon = 1 \cdot 10^{-3}$

Fig. 5 microDAC structure tracking by digital correlation analysis

Within the past years two appropriate WINDOWS® based computer codes have been developed for efficient displacement and strain extraction from load state images. The codes are able to track local structures with a subpixel accuracy of approximately 0.1. Subpixel accuracy of 0.1 means that local displacements with a resolution of

$$\delta u_{x,y} = 0.1\ L / M \tag{4}$$

can be measured. Here, L indicates the length of the field of view for the images used, M is the number of pixels along the image edge. The lateral image resolution of measurement is a function of the local pattern size n. Non overlapping of neighboring searched pattern must be provided for independent displacement values, i.e. the lateral resolution equals

$$\delta l = L\ n / M. \tag{5}$$

In Eq. (5) n is the pattern area pixel size. For the example of a typical low resolution SEM image of solder bumps the values L = 100 µm, M = 1024 and n = 15 are applied. Therefore, a measurement resolution of $\delta u_{x,y} \approx 10$ nm and a lateral structure resolution of $\delta l \approx 1.5$ µm can be achieved.

More sophisticated is the determination of displacement derivatives because of their high sensitivity to measurement errors of single displacement values. For that reason, combined with, or in advance to derivation, displacement data smoothing in a certain neighborhood of displacement points is strongly indicated. Consequently error estimation of strain values is not performed easily. In order to get more reliable information about strain measurement accuracy real object images have been deformed artificially in a defined way. E.g., sinusoidal one-directional object stretching with displacements of

$$u_x(x,y) = u_{x0} + \delta u_x \sin(2\pi f_x x) \quad (6)$$

was applied to study smoothing effects on ε_{xx}-strain accuracy. In formula (6) the displacements $u_x(x,y)$ are altered sinusoidally with the spatial frequency f_x and the displacement amplitude δu_x over the x-direction. Measurement errors could be calculated comparing microDAC software results with the introduced in the images artificial strains. Fig. 6 illustrates such estimations for ε_{xx}-strain measurement errors in dependence on the spatial frequency f_x. Obviously, the measurement accuracy decreases with increasing frequency f_x as a result of oversmoothing the oscillating ε_{xx}-strain field. For moderate displacement gradients over the image (1.5 sinus waves over the image size) a measurement resolution in terms of strain of 1×10^{-3} can be achieved. This value is independent on the kind of image capture, i.e. independent on the image magnification, too. However, other software parameters like search pattern size n (see Fig. 5), subpixel accuracy and smoothing characteristics can influence accuracy of strains and rotation angles. Summarizing, it can be stated, that accuracy values of 1×10^{-3} are realistic for slow deformation changes over the characteristic smoothing length and an obtainable subpixel accuracy of 0.1.

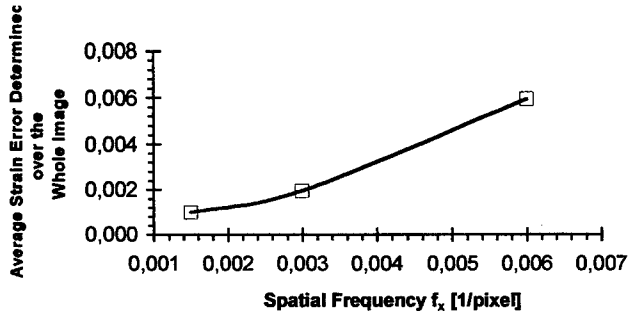


Fig. 6: Example of microDAC strain measurement accuracy as a function of strain frequency (determined for constant *strain* amplitude on simulated strain fields)

The microDAC software allows two alternative ways to extract displacement and strain data from load state images. In the first case deformation data is determined for a grid of equidistant measurement points. In the second case deformation values are found for the node points of a finite element mesh[18].

The type of measurement data available until now without using predefined finite element meshes is illustrated on the following Fig. 7 to 9. The shown chip scale package (CSP) solder joint refers to a rigid carrier CSP, which had been mounted to an organic laminate. The whole device was heated up between room temperature and approximately 125 °C. Solder deformation did take place due to the thermal mismatch between different materials, namely between silicon and the organic laminates.

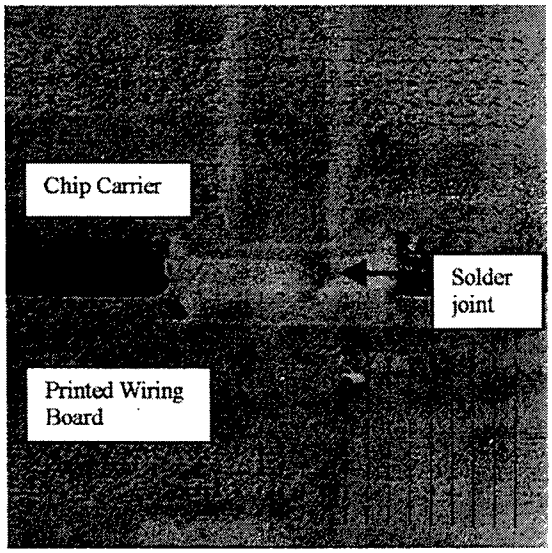
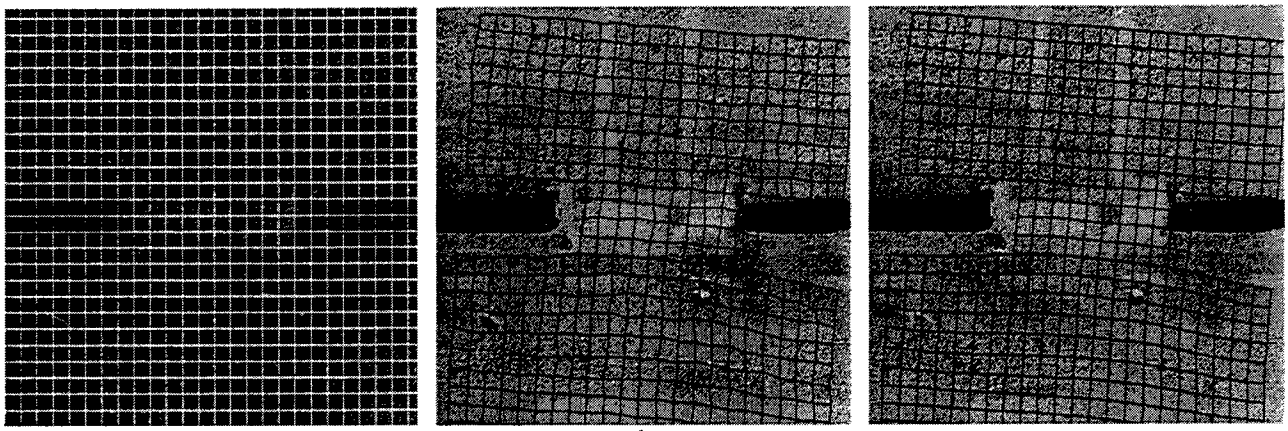


Fig. 7: Displacement vectors over the object image (solder joint of a chip scale package with rigid chip carrier)

The vector plot of Fig. 7 is one of the simplest data presentation provided. Displacement values for a single data point come immediately from the correlation analysis without any refinement unless subpixel interpolation. The same data has been used for the drawing of the deformed virtual object grid of Fig. 8b. After some additional data refinement (grid smoothing and repairing) the improved deformed grid of Fig. 8c is available. This option permits to draw low-noise grids with high displacement enlargement. The option has been established, since object warpage quite often has to be presented in a clear way. Earlier judgement upon warpage by contourlines of the local rotation angle had been proven not effective and was cumbersome for unskilled people.



a b c

Fig. 8: Deformation of the object in Fig. 7 presented by means of a deformed regular measurement grid, a - original, reference grid, b – deformed grid after image correlation, c – deformed grid after image correlation and data retrieval; grid deformations has been enlarged by a factor of 20.

Fig. 9 shows the powerful contourline presentation of displacement and strain fields, which is desired by engineers in the first place for application purposes.. Its supply calls for appropriate data smoothing, interpolation and derivation. Partially, commercially available graphic software compatible with microDAC data formats is applied.

a b c

Fig. 9: Deformation of the object in Fig. 7 presented by means of displacement and strain contourlines, a – vertical displacement component, b – vertical strain component, c – in-plane shear field

Displacement contourlines (Fig. 9a) superpose local object movements originating from rigid body displacements and rotations, strains and warpage (change of local rotation angle with coordinates). Consequently, simple interpretation of displacement maps is not possible in many cases. In the opposite, strain data (Fig.9b,c) is free of rigid body movements. Commonly, it is used for deformation analysis.

5.2. Combined microDAC Measurement and Finite Element Analysis

Commonly, measurement points are set for a regular grid overlaying the object images with selectable node distances. If desired, measurement points can also be defined from finite element (FE) input files. In this case finite element meshes have to be generated for the real component making use of component micrographs. Until now, mesh generators from ABAQUS and ALICE[19] can be utilized. The application of FE mesh nodes for microDAC meshes aims at the direct comparison between results from thermo-mechanical finite element analysis and from corresponding measurements on real components. Different uncertainties in finite element modeling make that way lucrative, although it is complex and demands close interaction between simulation and experiment.

In order to realize this concept in practice, the algorithm for strain measurement needs direct interfaces to the pre- and postprocessing of the finite element code. Some of the problems to be solved in this connection are:

- The fundamental problem of projecting a complex 3D-FE-mesh to the two-dimensional image defined by the direction of the optical axis and the field of view of the microscope.
- The center positions of searched image pattern in correlation technique must be identically with the exact positions of FE-nodes projected in the image plane. This implies that displacement measurement has to be carried out not for pixel arrays located at integer pixel positions.
- Boundary conditions regarding geometry and material limits must be chosen identically in both image correlation and finite element analysis.
- The spatial resolution of the measured displacement vectors has to fit the resolution of the numerical computation for all parts of the object.

Fig. 10 illustrates a deformation measurement on a thermo-mechanically loaded solder interconnect making use of a finite element mesh.

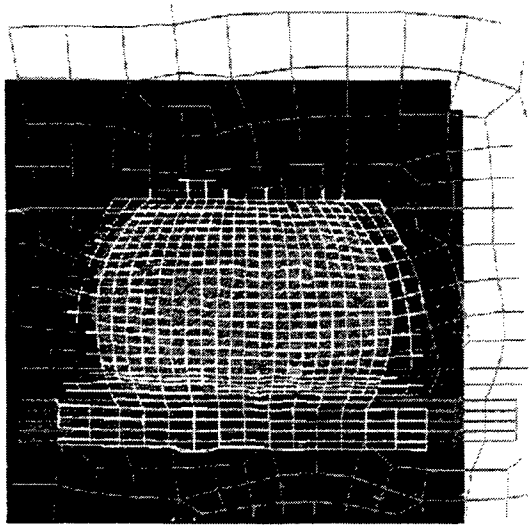


Fig. 10 Solder bump deformation measurement utilizing an ABAQUS finite element mesh,
gray mesh: part of the whole undistorted FE mesh,
white mesh: deformed FE mesh obtained from a microDAC measurement (3 times enhanced displacement values),
approx. micrograph field of view: 100 µm

6. MICRODAC APPLICATIONS

6.1 Application Areas of microDAC in Electronic Packaging

The microDAC approach basing on the measured mutual movement of local micrograph pattern of objects under load possesses advantages as:

- relatively simple measurement set-up (e.g. in comparison with Moiré techniques),
- excellent down scaling capability for displacement resolution and object size,

- relatively easy access to measurements on thermally loaded specimens,
- possibility to study fast processes using high speed cameras,
- high potential to automate the measurement process.

On the other hand, the actually most serious limitation of the method is its lower resolution corresponding to strain measurement. Accuracy of approximately 1×10^{-3} restricts the application to plastic or viscoplastic behavior of most of materials.

Taking into consideration advantages and limitations of the method, the following topics could have been selected as possible application fields:

1. Strain measurement on components and materials exposed to severe deformation. This is a common situation for most of solder or conductive polymer interconnects on modern electronic packages and leads to significant material creep. Mainly solder strain in flip chip (FC) solder bumps, in solder interconnects on chip scale packages (CSP) and in conductive adhesives of flip chips have been investigated until now.
2. High resolution displacement measurement of rigid bodies. Applications have been made to record mutual displacements of parts, which are not easily measurable with other methods, e.g. long-term drifts of laser diodes against optical fibers.
3. Warpage studies of components under load. Warpage leading to a change of the local rotation angle over the object can be determined, although measurement resolution may be not sufficient for real strain determination. These measurements have been performed for different assembled packages.
4. Measurement of coefficients of thermal expansion (CTE). Again, strain measurement accuracy seems to be insufficient to solve the problem. But considering specimens as homogeneously expanding bodies the analyzed statistics of measured strain values over the whole micrograph allows to achieve the necessary accuracy. The modified microDAC method can be applied to materials with higher CTE values, like plastics. It is favorable for materials of poor consistency, i.e. for epoxies beyond the T_g temperature and for thin layers like e.g. underfills, solder stop and adhesive materials.
5. Analysis of cracks and delaminations. Although lower elastic strains can not be measured directly, different features are accessible. So the presence and the location of delaminations and cracks can be determined, as well as crack tip locations or the crack opening amount.

Some of the listed above microDAC applications are illustrated in the following part by examples.

6.2. Measurement of Rigid Body Movements on Optoelectronic Modules

Correlation based measurement can be used to determine mutual displacements and/or rotations of components within a system. For this purpose the single components are assumed to remain rigid and undeformed. Their absolute displacement or rotation is calculated very accurately averaging displacement values over the whole visible surface. So, mutual movements can be determined with higher accuracy.

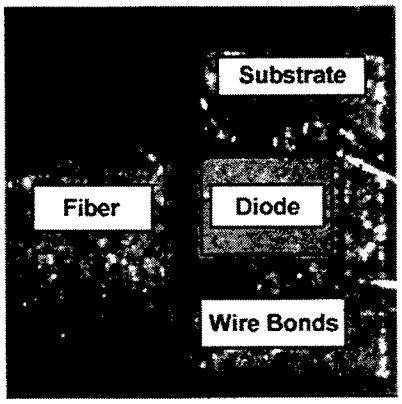


Fig. 11 Optical micrograph of optoelectronic module for rigid body displacement measurement, black frames: areas of averaging

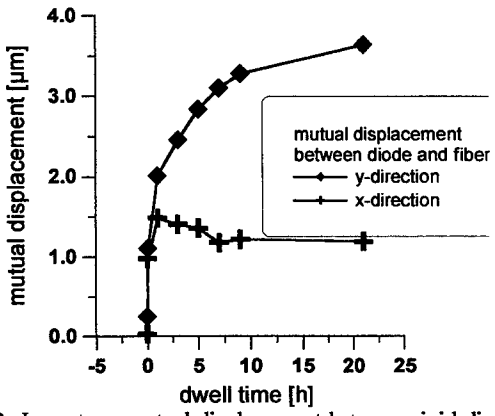


Fig. 12 Long-term mutual displacement between rigid diode and fiber, temperature: approx. 50 °C over room temperature

Fig. 11 shows a part of an optoelectronic module including a laser diode with an optical fiber. Coupling of the laser emission to the single-mode fiber is quite sensitive to the mutual position of both of components. Larger changes in position and rotation of the components to each other must be avoided by appropriate assembly to the module case. Stability has to be maintained also under long-term thermal system load. Measurement of mutual movements by conventional microscopy was impossible because of the demanded high submicron accuracy and the missing feasibility to focus the fiber and diode surface at the same time. Nevertheless, correlation algorithms tolerate some defocusing and meet the demanded accuracy. Fig. 12 demonstrates the measurement of the two in-plane components of relative displacements between the diode and the fiber. A long-term thermal drift of the diode respectively to the fiber could have been observed, which had an serious impact on the coupling quality.

6.3. Strain Measurements on Advanced Packages of the Electronics Industry

6.3.1. Solder strain measurements on flip chip bumps

Aiming at low cost electronic assemblies attempts have been made at flip chip bonding on organic substrates. In this case the silicon die is mounted face down to the substrate. Electrical contacts are realized by tiny solder bump arrays connecting the die and board pads. (see Fig. 13). A main problem occurs for this approach with regard to the large thermal mismatch between the epoxy board and the silicon die. Underfill encapsulants are used to improve the mechanical integrity of the assemblies. They have to compensate mechanical and thermal stresses between substrate and chip. Their material characteristics as thermal expansion coefficients, Young's modules, creep parameters and others are important inputs to be optimized for particular configurations.

In order to obtain information upon the influence of material properties, assembly geometry and manufacturing imperfections on the strain/stress behavior and component life times, finite element simulations and also microDAC measurements for e.g. thermal cycling are being performed. For selected cases the results are compared with each other to make sure, that adequate mechanical modeling in the FEA has been applied.

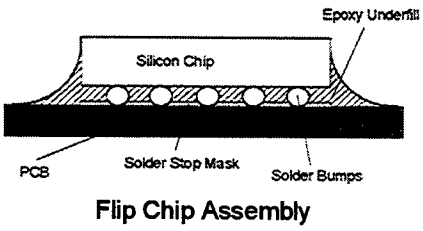


Fig. 13 Flip chip bonding (schematically)

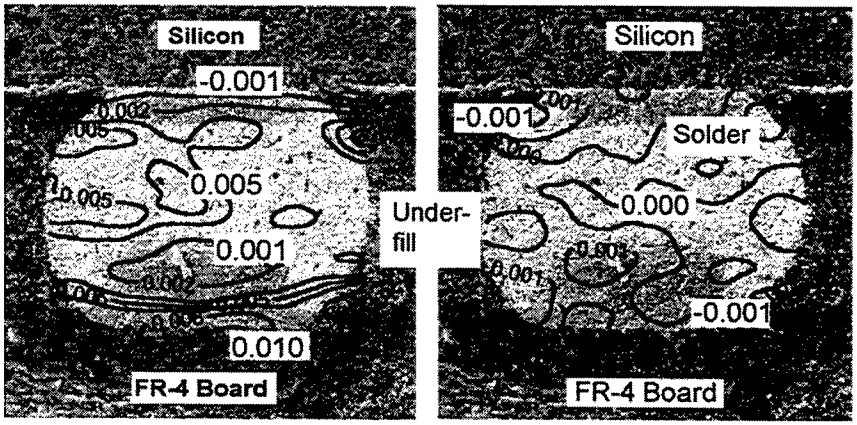
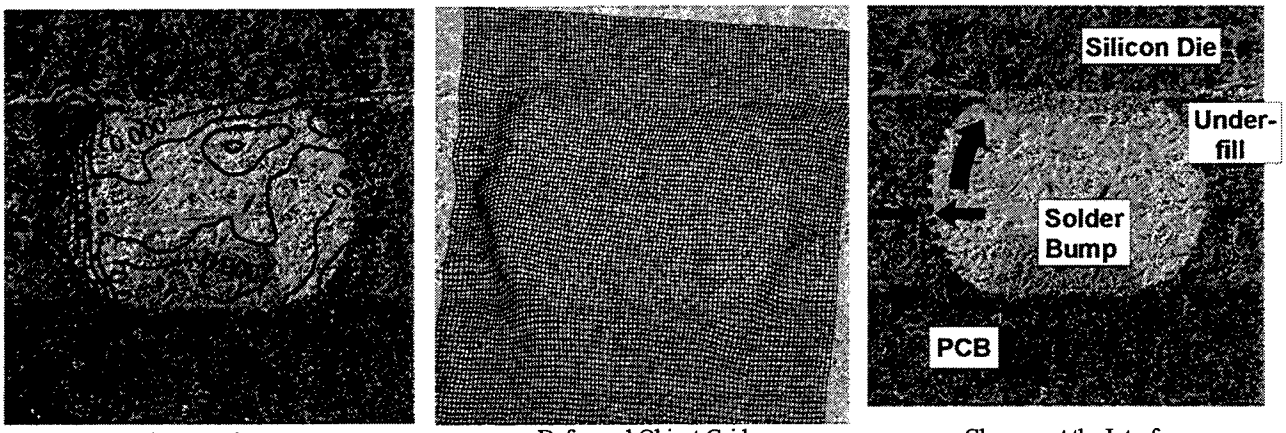


Fig. 14 PbSn solder bump strain field for assembly heating from 25°C to 125°C, strain component ε_{yy} perpendicular to board and shear strain ε_{xy}, microDAC measurement inside SEM, highly filled underfill, approx. bump size: 50 µm

Measurement specimens for microDAC applications have been cross-sectioned along a bump row and thermally loaded inside a scanning electron microscope. Measurements were done heating up the specimens from room temperature to levels above 100 °C. Fig. 14 illustrates typical solder strain measurements inside solder bumps. A maximum strain perpendicular to the board ε_{yy} of 0,5 % develops in the bump middle. It drops down to the bump borders. However, at the FR-4 PCB side an additional local strain maximum of nearby 1 % occurs. In the opposite shear of the bumps is held on a very low level less than 0.002, which is a result of underfilling and subsequent package warpage.

Fig. 15 gives an example of defects occurring on solder bumps after thermal cycling. At the left bump interface to the underfill high "pseudo" shear values appear (Fig. 15a). In fact, in Fig. 15b the distorted corresponding grid at that place

supports this result. Fig. 15c summarizes the main changes at the interface due to heating up. The bump and underfill material slide against each other and a gap between bump and underfill closes, when the assembly is heated up.



Shear Strain ε_{xy} | Deformed Object Grid | Changes at the Interface

Fig. 15 PbSn solder bump strain for assembly heating from room temperature over 100 °C, microDAC measurement inside SEM, delamination at bump/underfill interface

6.3.2. Deformation of solder balls between chip scale packages (CSP) and printed circuit boards (PCB)

Chip scale packages (CSP) are one answer to accept the challenge of miniaturization and higher packaging density. They are now becoming available on the marketplace. As the CSP should not be greater than 1.2 times the chip size the required mounting area is comparable to flip chip technology (see also paragraph 6.3.1.) because the underfill of flip chips requires some additional area. In contrast to flip chip technology the CSP should not require an technologically costly underfill process to facilitate board level assembly. Therefore the thermal mismatch between the CSP and the PCB must be handled by the package to provide sufficient reliability. Fig. 16 presents one of the developed CSP within the past years.

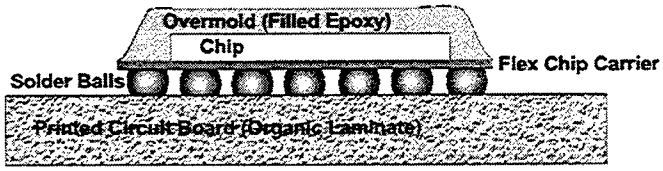


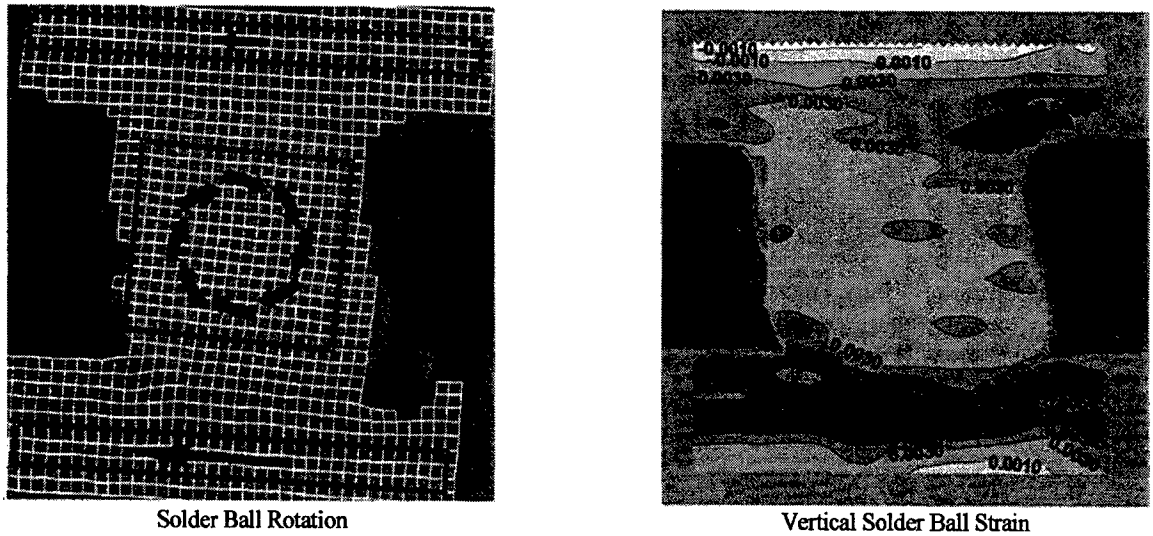
Fig. 16 Chip scale package with flex interposer (schematically)

One key to appropriate package design is the understanding of solder ball load behavior, which occurs due to different thermo-mechanical impact. Fig. 17 illustrates a microDAC deformation measurement on an outermost solder ball for a CSP with flex interposer assembled on the PCB. It could have been shown, that in this case the solder balls act similar to a "ball bearing". The large CTE difference between the silicon die and the organic laminate forces the die to move against die board (black straight arrows). The solder balls response with a overall rotation relatively to the die and board. On the other hand, strain in the direction perpendicular to the package plane is rather low.

A comparison with solder ball deformation of another package with rigid chip carrier reveals a complete different behavior. Common warpage of the rigid carrier package and the PCB[20] results in solder strains perpendicular to the package plane of about 1 % (see Fig. 9). Solder ball rotation similar to that of the CSP with flex carrier (Fig. 17) has not been observed.

6.4. Determination of Material Properties on Unconventional Specimens

Strain fields measured by optical microDAC can be utilized to determine material properties, e.g. coefficients of thermal expansion (CTE). In order to achieve sufficient strain value accuracy for subsequent CTE evaluation a homogeneous sample expansion in the microscopic field of view is presumed. Appropriate strain values are provided averaging local strain values over the whole image. CTE's are determined from a set of measured incremental strain values for different temperatures with correspondence to a reference temperature. Regression methods are applied to obtain CTE's as a function of temperature.



Solder Ball Rotation

Vertical Solder Ball Strain

Fig. 17 Outermost solder ball deformation of an assembled CSP with flex chip carrier

Strain values are calculated for two perpendicular directions, i.e. CTE's for anisotropic materials can be measured easily. Moreover, a digital image rotation algorithm allows to rotate the micrographs in the computer and figure out directions of maximum and minimum CTE without new, time consuming measurements.

The suggested method of CTE measurement is favorable, if classic dilatometric equipment cannot be applied, i.e. in the case of poor sample consistency, layered materials, anisotropic CTE behavior and for samples of microscopically small size. The method can be applied especially to plastics, e.g. to adhesives, mold compounds, encapsulants, underfills and others.

Fig. 18 shows a measurement result for a FR-4 multilayer board material. Strains in and perpendicular to board plane are plotted versus temperature. A linear regression approach has been chosen to calculate CTE's from both sides of the glass transition. For one of the in-plane CTE components for temperatures higher the glass transition temperature T_g a small shrinkage appears, which obviously is due to the fiber stress relaxation in the reinforced epoxy.

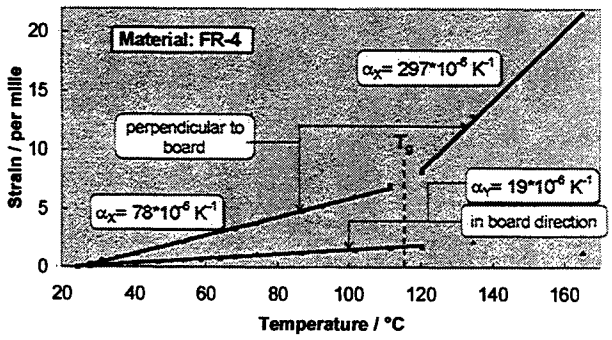


Fig. 18 CTE measurement by microDAC

7. NANODAC – A FUTURE PROSPECT OF MICRODAC

Currently microDAC measurements are being carried out on optical, laser scanning and electron scanning microscopes. Generally, higher resolution can be achieved making use of some kind of scanning force microscopes. Moreover, scanning force microscopes express different force interaction with the surfaces under investigation, i.e. captured object characteristics are not the same as in SEM or optical imaging. Thus in some cases, small image pattern suitable for

correlation analysis can be imaged if that is not possible with other equipment. E.g. sometimes sufficient topography or phase contrast in SEM may be not within reach, while SFM imaging provides the necessary surface structures. This situation would be typically for studies on very smooth object surfaces.

Fig. 19 illustrates some first attempts to utilize AFM's for high resolution correlation analysis. A ceramic sample was scanned in a constant gradient mode of the AFM. After some image refinement the mentioned AFM images were treated by the correlation software. It can be seen, that almost all measurement points had acceptable correlation quality to determine the rigid body displacement of the object in between the image captures. The result of a strain extraction from these displacement fields is presented on Fig. 20. Obviously, pseudo strains occur over the undeformed image. They reach values up to $4x10^{-3}$, well above the measurement resolution for the correlation algorithms. The reason for the high pseudo strains can be found in the still insufficient scanning stability of the scanning device used. Consequently, for future applications major effort must be made to obtain reasonable scanning and object stability during load.

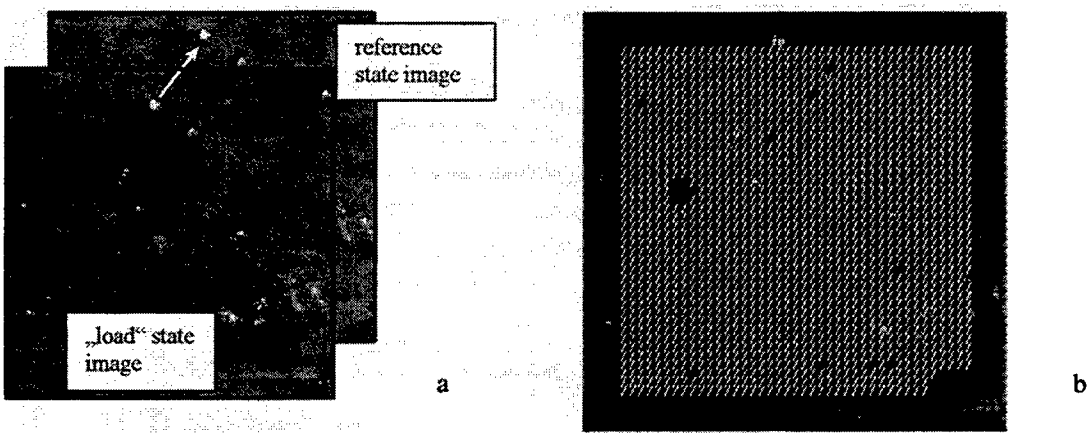


Fig. 19 Correlation analysis on AFM images, a: two object images (polished ceramics) with rigid body displacement in between, b: vector plot for displacement field (empty points: insufficient pattern correlation)

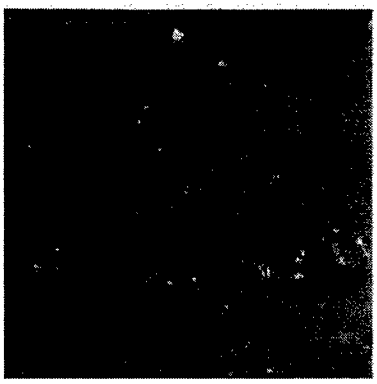


Fig. 20 Horizontal strain field computed from AFM images of Fig. 19a, maximum pseudo strains due to scanning instability approx. $4x10^{-3}$ (numerical error of strain calculation approx. $1x10^{-3}$)

8. CONCLUSIONS

The demand for spatially resolved strain measurement inside microscopic structures has been illustrated focusing on the R&D needs in micro technology and micro electronics. Different kinds of approaches to meet this goal and realized at the IZM Berlin have been presented. Among the methods are classic interferometry, microMoiré and microDAC techniques. Some emphasis has made with regard to the correlation based method microDAC. It's main principles and capability have been reviewed. Applications concerning electronic packaging, e.g. deformation analysis on optoelectronic modules, advanced packages like CSP and FC were presented. The usefulness of micro strain measurement for the determination of material properties has been shown as well. Finally, future developments were illustrated by the use of scanning force microscopy as a image source in microDAC technique.

9. REFERENCES

1. B. Michel, „Fracture electronics - concepts of fracture mechanics for reliability estimation in microelectronics and microsystem technology", Proc. Micro Materials '97 Conference, Berlin, Apr. 16-18, 1997, Berlin, ddp Goldenbogen 1997, pp. 382-389.
2. B. Michel, A. Schubert, R. Dudek, V. Grosser, „Experimental and numerical investigations of thermo-mechanically stressed micro components", Microsystem Technologies 1 (1994) 1, pp. 14-22.
3. D. Vogel, A. Schubert, W. Faust, R. Dudek, B. Michel, „MicroDAC - A novel approach to measure in situ deformation fields of microscopic scale", Microelectronics Reliability 36 (1996) 11/12, pp. 1339-1342.
4. J. Auersperg, D. Vogel, J. Simon, A. Schubert, B. Michel, „Reliability evaluation of chip scale packages by FEA and microDAC", Proc. on Design and Reliability of Solders and Solder Interconnects, TMS Annual Meeting, Orlando, Febr. 9-13, 1997, pp. 439-445.
5. B. Michel, T. Winkler, H. Reichl, „Fracture electronics - application of fracture mechanics to microelectronic systems and chip packages", Proc. 9th Int. Conf. on Fracture (ICF 9), Sydney, Apr. 1-5, 1997, in: Advances in Fracture Research (ed. B. Karihalo), vol. 6, pp. 3107-3112, Pergamon Press 1997.
6. Semiconductor Industry Association (SIA): The National Technology Roadmap for Semiconductors: Technology Needs, 1997 Edition
7. R.C. Jaeger, J.C. Suhling: Advances in stress test chips, Proc. of 1997 ASME Int. Mech. Eng. Congress and Exp., Nov. 16-21, 1997, Dallas, Vol. "Applications of Experimental Mechanics to Electronic Packaging", EEP-Vol. 22, AMD-Vol. 226, pp. 1-5.
8. I. De Wolf et al: Experimental and Analytical Evaluation of Stress Fields Generated by Solder Bump Interconnections, Proc. of Micro Mat '97, April 16-18, 1997, Berlin, pp. 261-269.
9. A.Selvarajan, A.Asundi: Photonics, fiber optic sensors and their application in smart structures, see under http://www.ntu.edu.sg/home/masundi/
10. B. Han, Y. Gou: Thermal deformation analysis of various electronic packaging products by Moiré Interferometry, J. of Electronic Packaging, Vol. 117, 1995, September, pp.185-191.
11. X. Yan, R.K. Agarwal: Underfill Selection for Flip-Chip on Organic Printed Circuit Boards, Proc. of SEM Spring Conference on Experimental and Applied Mechanics and Experimental/Numerical Mechanics in Electronic Packaging II, June 1-3, 1998, Houston, pp. 73-83.
12. D. Post, B. Han, P. Ifju: Thermal deformations in electronic packaging in "High Sensitivity Moiré", Springer-Verlag, N.Y., 1994, 331-347.
13. I. Yamaguchi, et al: Stabilized and accelerated speckle strain gauge, Optical Engineering, Vol. 32, No. 3, 1993, pp. 618-625.
14. B. Zagar, H. Weiss, B. Weiss: A high speed laser-optical correlation based strain sensor, Proc. of the XIV IMEKO World Conf. Tampere, Finland, June 1997, Vol. VIII, pp. 228-233.
15. B. Michel, R. Kühnert: Mikro-Moire-Methode und MikroDAC-Verfahren anwenden, Zeitschrift Materialprüfung, Vol. 38, No. 6, 1996.
16. D. Vogel, A. Schubert, W. Faust, R. Dudek, B. Michel: MicroDAC - A novel approach to measure in situ deformation fields of microscopic scale, Microelectronics Reliability, Vol. 36, 1996, No. 11/12, pp. 1339-1342.
17. D. Vogel, J. Auersperg, A. Schubert, B. Michel, H. Reichl: Deformation analysis on flip chip solder interconnects by microDAC, Proc. of Reliability of Solders and Solder Joints Symposium at 126th TMS Annual Meeting & Exhibition, Orlando, USA, 1997, in „Design & Reliability of Solders and Solder Interconnects", ed. by R.K. Mahidhara, TMS Publication Cat. No. 96-80433, pp. 429-438.
18. B. Michel, D. Vogel, A. Schubert, J. Auersperg, H. Reichl.: The microDAC method - a powerful means for microdeformation analysis in electronic packaging, Proc. of Symp. on „Applications of Experimental Mechanics to Electronic Packaging" at the „1997 ASME International Mechanical Engineering Congress and Exposition", Nov. 1997, Dallas, AMD-Vol. 226, EEP-Vol. 22, pp. 117-123.
19. IZM in-house code, see J. Auersperg, R. Dudek, B. Michel: Investigation of mechanical-thermal field coupling effect during quasi-static crack growth and application in micro system technology, Proceedings of ABAQUS Customer Meeting, Ulm, 28./29.09.95, pp. 138-152.
20. D. Vogel, R. Dudek, W. Faust, A. Schubert, B. Michel: Characterization of Advanced Packaging Structures by microDAC Technique, Proc. of 3rd Int. Symp. on Electronic Packaging Technology (ISEPT '98), 1998, August 17 - 22, Beijing.

Invited Paper

High-resolution stress and temperature measurements in semiconductor devices using micro-Raman spectroscopy.

Ingrid De Wolf*, Jian Chen, Mahmoud Rasras, W. Merlijn van Spengen, and Veerle Simons

IMEC, Kapeldreef 75, B-3001 Leuven, Belgium

ABSTRACT

After a short introduction on the theory and instrumentation of Raman spectroscopy, its application for local stress and temperature measurements in semiconductor devices is discussed. Examples are given for silicon isolation structures, transistors, solder bumps and back-grinding. It is shown how the resolution can be improved by using an oil immersion objective and deconvolution techniques. Different imaging modes are discussed and their resolution is compared. Examples of 1-dimensional and of 2-dimensional scans are shown.

Keywords: Raman spectroscopy, resolution, stress, temperature

1. INTRODUCTION

For a long time, Raman spectroscopy could only be found in chemistry labs of universities, where it was used to study the vibration modes of molecules of fluids and gasses. One of the first recordings of the Raman spectrum from crystalline silicon and germanium was reported only in 1967, showing weak, noisy signals[1]. But from then on Raman spectroscopy instrumentation and applications developed very fast. During the last 10 years, it found its way from university labs to microelectronics research centers and industry.

The amount of information on semiconductor devices that can be obtained using Raman spectroscopy is very large. Because Raman spectroscopy probes lattice vibrations, which are sensitive to internal as well as external perturbations, it can be used to study for example the crystal quality, crystal phase, temperature, strain, crystal orientation, composition, doping, etc. In this paper, we discuss only two of these applications: strain and temperature.

An attractive feature of Raman spectroscopy is that it makes use of laser light, and that this light can be focused through a microscope on the sample. This *micro*-Raman spectroscopy technique is very useful for the study of *micro*electronics devices. However, with the ever decreasing size of semiconductor devices into the sub-micron world, an improvement of the spatial resolution of the Raman spectroscopy technique is mandatory. The different possibilities to improve the resolution of Raman spectroscopy will also be discussed in this paper.

2. RAMAN SPECTROSCOPY

2.1 Theory

Raman scattering in crystals arises from the interaction of light with lattice vibrations. The simplest way to theoretically describe Raman spectroscopy, is by following a semi-classical approach. Vibrations of a crystal are described in terms of collective motions in the form of waves (lattice vibrations). Each possible vibration (normal mode or phonon) is characterised by a wavevector q_j, a frequency ω_j, and an amplitude Q_j (normal coordinate). The vibration amplitude, at position **r**, is given by:

$$Q_j = A_j \exp[\pm i(q_j . r - \omega_j t)] \qquad (1)$$

These lattice vibrations may cause a variation in the electrical susceptibility of the crystal.

When monochromatic light of frequency ω_i is incident on a crystal in a direction k_i, the associated electric field $\boldsymbol{E}$ will induce at position $\boldsymbol{r}$ an electric moment $\boldsymbol{P}$:

* Correspondence: E-mail: dewolfi@imec.be

Part of the SPIE Conference on Advanced Photonic Sensors and Applications
Singapore • November–December 1999
SPIE Vol. 3897 • 0277-786X/99/$10.00

$$\boldsymbol{P} = \varepsilon_o \chi \boldsymbol{E} = \varepsilon_o \chi \boldsymbol{E}_o \exp[i\,(\boldsymbol{k}_i \boldsymbol{r} - \omega_i t)] \tag{2}$$

where χ is the susceptibility tensor. If the atoms are vibrating, χ may change as a function of these vibrations. This can be expressed by expanding χ, for each mode of vibration j, in a Taylor series with respect to the normal coordinate of this vibration Qj:

$$\chi = \chi_o + \left(\frac{\partial \chi}{\partial Q_j}\right)_o Q_j + \left(\frac{\partial \chi}{\partial Q_j \partial Q_k}\right)_o Q_j Q_k + \ldots \tag{3}$$

Combining Eqs. 1 to 3 results in:

$$\mathbf{P} = \varepsilon_o \chi_o \mathbf{E}_o \exp[i(\mathbf{k}_i \mathbf{r} - \omega_i t)] + \varepsilon_o \mathbf{E}_o \left(\frac{\partial \chi}{\partial Q_j}\right)_o A_j \exp[-i(\omega_i \pm \omega_j)t] \exp[-i(\mathbf{k}_i \pm \mathbf{q}_j)\mathbf{r}] + \ldots \tag{4}$$

From which follows that the induced moment will re-radiate light, ω_s, with three different frequency components:

$$\omega_s = \omega_i \qquad \omega_s = \omega_i \pm \omega_j \tag{5}$$

The first component has a frequency equal to the one of the incident laser light, ω_i, and is called Rayleigh scattering (elastic scattering); the second and third have frequencies $\omega_i+\omega_j$ and $\omega_i - \omega_j$, and are called first order anti-Stokes and Stokes Raman scattering, respectively. Higher order terms in Eq. 4 describe second-order Raman scattering, where two phonons are involved ($\omega_i \pm 2\omega_j$), or higher order Raman scattering. Raman studies in semiconductors are mostly restricted to first order Stokes Raman scattering, which is the strongest process. First order Raman scattering can only be observed if

$$\left(\frac{\partial \chi}{\partial Q_j}\right)_o \equiv \chi_j^1 \neq 0 \tag{6}$$

which means that the susceptibility changes as a function of the lattice vibrations. Fig. 1 shows a typical Raman spectrum of the Stokes Raman spectrum from crystalline silicon. The silicon Raman peak is located at a position of 521 cm^{-1}. The sharp lines in the spectrum are plasma lines from the laser. These plasma lines have typically about the same intensity as the Raman signal. They can be used as reference when stress or temperature measurements are performed using Raman spectroscopy, as will be shown later.

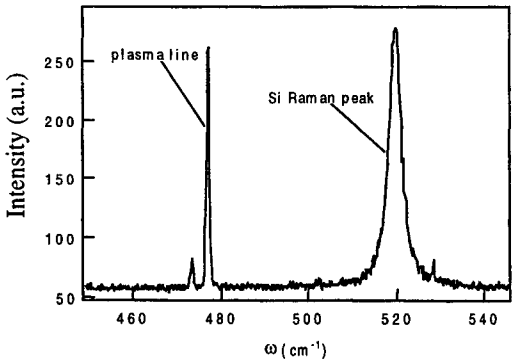


Figure 1. Typical Raman spectrum of crystalline silicon, showing the Si Raman peak and plasma lines from the laser.

2.2 Instrumentation

Raman instrumentation typically consists of a light source, focusing optics, a sample holder, collecting optics, a spectrometer, a detector, and a computer with appropriate software for data acquisition and storage. A typical large Raman spectroscopy instrument is shown schematically in Fig. 2. Because of its high intensity and linear polarization, the exciting light source normally used for Raman spectroscopy is a laser. Which laser to use depends on the application. For conventional Raman measurements in semiconductors, an Argon ion laser and/or Krypton ion laser are often used. A UV laser can be appropriate if one requires very small penetration depths[2]. In between laser and sample, different filters, lenses, pin holes and mirrors can be inserted. In the macro-Raman spectroscopy instrument, the laser light is focused on the sample through a focusing lens system (L_M). For local stress or temperature measurements, a micro-Raman spectroscopy instrument is of more interest. In such an instrument light is focused through an optical microscope, with objective lenses varying from 10x to 100x, resulting in spot sizes down to 1 μm. The classical Raman spectrometer consists of a single or double premonochromator, and a spectrometer. The premonochromator, with mostly plane holographic gratings (G), is used to eliminate the Rayleigh scattered laser line. Newer instruments make use of notch filters[3] to cut out the laser line.

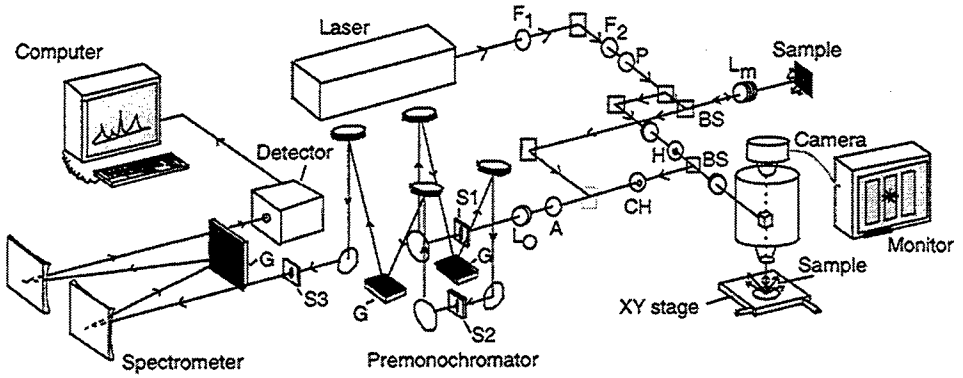


Figure 2. Scheme of a typical Raman instrument with laser, filters (F1,F2), polarizer (P), analyzer (A), beam splitters (BS), pin hole (H), confocal pin hole (CH), slits (S1-S3), and gratings (G).
The instrument has a macro stage (with macro-lens L_M) and a micro stage.

2.3 Resolution

The spatial resolution of a conventional micro-Raman spectroscopy instrument depends on the size of the probing laser spot. This size depends in turn on the wavelength of the used laser light and on the objectives of the microscope. In theory, the diameter of the focused laser beam on the sample, Φ, can be defined as the diameter at which the intensity of the spot has decreased to $1/e^2$ of its value in the center. If the laser beam has a diameter much larger than the entrance aperture of the focusing objective (which is not necessary the case in all Raman instruments) this is given by:

$$\Phi = \frac{0.88\lambda}{NA} \tag{7}$$

where λ is the wavelength of the laser light and

$$NA = n \sin \theta \tag{8}$$

is the numerical aperture of the objective lens. θ is the angle between the outer rays of the focused beam and the optical axis, and n is the refractive index of the material between lens and sample. This is in most cases air and n can be taken equal to 1. It is clear that this offers two ways to improve the spatial resolution of a Raman instrument: decreasing the wavelength, or increasing the numerical aperture.

The smallest wavelength that can be used depends on the laser, on the optics of the Raman system (mirrors and sensitivity of gratings) and on the sensitivity of the detector. The shortest wavelength of the Argon laser of our system is the 457.9 nm line. There also exist UV-Raman instruments, which have UV optics and work with a wavelength of 364 nm. A drawback there is that the advantage of the short wavelength is neutralized by the fact that the existing objective lenses for UV are not very good, the best beam size which can be obtained[2] is 0.7 μm. The main advantage of using UV-Raman spectroscopy for stress measurements is that the probing depth of the UV is very small, which makes the system much more sensitive to high stresses at the surface than conventional systems[2].

Increasing the refractive index n will also lead to a better resolution. This can for example be done by an oil immersion objective, where oil with n = 1.4 is placed between sample and objective. The theoretical values for two typical wavelengths, often used in conventional micro-Raman spectroscopy, are listed in Table 1 for two kinds of 100x lenses with different numerical aperture: a conventional lens and an oil-immersion lens.

Table 1. Diameter Φ of focused laser light. (# = number of experiments, sdev = standard deviation)

Lens	Wavelength (nm)	Φ (μm) (theory: Eq. 2)	Φ (μm) ± sdev (experiments)
100x, NA = 0.95	514	0.66	0.91 ± 0.16 (# = 16)
	457.9	0.59	0.86 ± 0.17 (# = 15)
100x oil, NA = 1.4	514	0.45	0.59 ± 0.01 (# = 8)
	457.9	0.40	0.51 ± 0.04 (# = 5)

The most commonly used objective in micro-Raman instruments is the 100x objective with a numerical aperture of about NA = 0.95. According to theory, this lens should have a beam spot diameter as small as 0.66 μm when the 514 nm line of an Argon ion laser is used (green light), and 0.59 μm for the 457.9 nm line (blue light). When using an oil-immersion objective, better results are obtained, down to 0.4 μm for the 457.9 nm line. In practice, these numbers are much larger. In order to determine the spatial resolution of our Raman system, we used a silicon wafer which was partly covered with a thin layer (30 nm) of $TiSi_2$. The sample was moved under the microscope in steps of 0.1 μm, using a computer driven XY-stage in such a way that the laser beam moved across the edge of the $TiSi_2$ film, starting on the film. At each position, where possible, a Raman signal of silicon was measured and its intensity was determined. Since the $TiSi_2$ film was not transparent for the laser light, no silicon Raman signal was obtained when the laser was completely probing the film. When crossing the edge of the film, the Si signal was detected and its intensity increased until the laser completely probed the Si. Fig. 3 shows a typical result of such an experiment. The normalized integrated intensity of the Si-Raman signal is plotted as a function of the position on the sample. A numerical approach of an integrated Gauss function[4] is fitted to the data to determine the beam diameter at $1/e^2$ of its maximal intensity. The results obtained from a number of experiments are shown in Table 1 for the two 100x lenses. It is clear that the experimental resolution does not correspond with the theoretically predicted resolution. This may be due to a bad outlining of the instrument or bad lens quality. The results for the oil-immersion objective are clearly better than the ones for the conventional 100x lens. An example of the use of an oil immersion objective for a stress measurement will be discussed in Fig. 10.

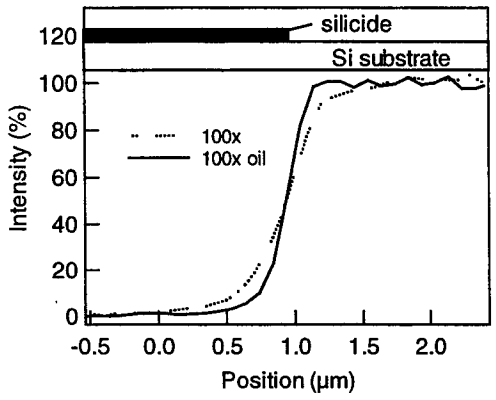


Figure 3. Raman intensity variation for a scan from the silicide to the silicon. Full line: 100x oil objective, dotted line: conventional 100x objective.

If spot sizes even smaller than 0.4 μm are required, which is certainly the case with nowadays silicon devices, near-field scanning optical microscopy (NSOM) Raman will have to be used[5-8], where the laser is focused on the sample through an optical fiber with small opening (20 to 200 nm). The Raman imaging capabilities in NSOM were shown to work for materials with high Raman scattering cross sections, such as diamond. However, the sensitivity of NSOM Raman spectroscopy is very low: for silicon a sensitivity of about 1 photon per second was reported. Raman NSOM (RNSOM) measurements of stress near a scratch in silicon were reported in reference 6. In this study, a 4 μm square area was mapped with an array of 26x21 spectra. The integration time for each spectrum was 60s. The diameter of the probe was about 150 nm. A decrease of the Si Raman frequency was observed upon approaching the scratch, indicating mechanical stresses of about 1.75 GPa in the silicon at about 2 μm from the edge of the scratch, decreasing to about 0.5 GPa at the edge. Inside the edge, amorphous Si was found. These experiments showed that high resolution Raman spectra can be obtained using NSOM. However, the technique still requires further improvement. The intensity of the Si Raman peak, as measured using RNSOM, is still small (about 150 counts in ref. 6), and therefore the peak frequency can only be determined with an error of about ± 0.2 cm^{-1}. Assuming uniaxial stress, this corresponds with a stress value of ± 90 MPa (Eq. 9). With conventional μRS, shifts as small as 0.02 cm^{-1} can be measured, indicating a 10x better sensitivity for stress[9]. This sensitivity is certainly required for the investigation of local stress in microelectronics devices. However, RNSOM mapping has been shown to work and it is certainly a very promising technique for the future.

Another reported approach is the use of a solid immersion lens (SIL)[10]. This system combines the advantage of a high refraction index material (high n in Eqs. 1,2) with near-field conditions. Because no optical fiber with narrow aperture is used, there is no direct limitation on the laser power. However, the SIL system has still some disadvantages, which make its use for Raman spectroscopy in microelectronic devices difficult. First of all, it only works well if the distance between lens and sample is small enough to have near field conditions. This is possible on flat samples, but more difficult on structured samples. For the latter, special lenses with a small tip have to be developed[11]. The SIL lenses which are at this moment available also suffer from aberration problems. Good focusing is only possible when the center of the lens is exactly above the structure which has to be measured. But also this technique is very promising for future nano-Raman spectroscopy work.

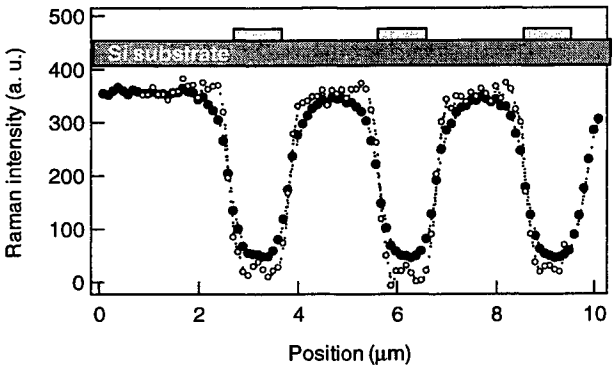


Figure 4. Intensity variation of the Si Raman signal when crossing 1 μm wide gold lines, spacing 2 μm, on Si substate, as shown schematically at the top. Full symbols: intensity obtained by fitting the spectra, open symbols: intensity obtained using deconvolution techniques on the original spectra, followed by fitting.

There is another, non-experimental approach to obtain a better spatial resolution. Fig. 4 shows (full symbols) the variation of the intensity of the Si Raman signal during a scan in steps of 0.1 μm across 1 μm wide gold lines on Si substrate. The spacing of the lines is 2 μm. This experiment was performed with a 50x objective, NA = 0.75. The focused spot diameter of this objective is 1.2 μm. The gold lines are not transparent for the laser light, so, no Raman signal can be obtained from the silicon underneath the lines. One would therefor expect that the intensity of the Raman signal drops to zero on top of the lines. This is not the case, as can be seen in Fig. 4, because the probing beam diameter is larger than 1 μm. The blurring of the line scan or image is for a large part determined by the intensity distribution of the light within the focused laser spot. So, if we know the properties of this focused laser spot, we can correct for this blurring by computational means. Indeed, standard image deconvolution techniques[12] can be used for deconvolution of the fitted value of the Raman signal. It is even more interesting to perform deconvolution on the complete Raman spectra (so not on the fitted data, but on the original spectra), because in that way a much more accurate result is obtained. The Raman spectra are in that case considered as intensity vectors in the deconvolution process. The full symbols in Fig. 4 show the original intensity data, obtained by fitting each measured Raman spectrum with a Lorentzian function. The open symbols show their deconvoluted counterpart. Deconvolution resulted in this case in sharper edges at the lines and a lower intensity on top of the lines. This is what one would expect if a beam spot with a smaller diameter would have been used. The small increase in intensity under the gold

lines in the deconvoluted result is due to the fact that the beam shape used to perform these calculations was not exactly the same as the experimental one. It is mandatory for this technique that the shape of the focused beam is very well known. Nevertheless, it is clear that this deconvolution technique, in combination with for example an oil-immersion objective, can result in a considerable improvement of the resolution of micro-Raman spectroscopy.

2.4 Raman imaging and mapping

There exist different experimental methods to obtain the Raman spectrum of a sample. One can measure the Raman spectrum of one large or very small region, or map the variation of the signal along a line or in a large predefined area. The different methods are depicted in Fig. 5.

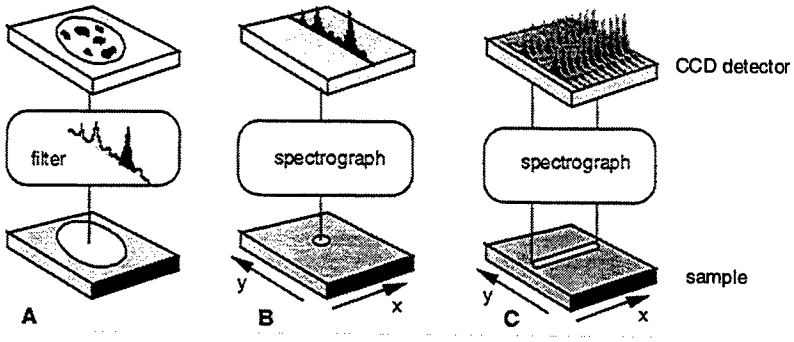


Figure 5. Raman imaging and mapping configurations:
A. Imaging of whole illuminated area.
B. Point illumination and 1D or 2D mapping through X-Y stage
C. Line illumination and 2D mapping through X-Y stage.

In the first method (A in Fig. 5), the sample is globally illuminated and a spectral filter is inserted in the scattered light beam [13,14]. This filter, which may be a narrow bandpass filter or a spectrometer with narrow slits, allows only one small spectral region, containing the Raman peak of interest, to pass to the detector. The detector then shows a 2D image of the intensity of the chosen Raman signal. This method is of interest when the spatial distribution of a certain substance has to be studied. It gives a direct, fast, two-dimensional Raman image. A disadvantage is that it does not allow confocality.
In the second method (B), a small spot of the sample is illuminated by the exciting laser. The light is collected into a spectrometer, giving one spectrum of the whole illuminated region. This method is very good to look at uniform films (macro) or at small regions (micro). Using a computer, the Raman peaks can be fitted and not only the intensity, as in method A, but also the frequency and peak shape can be studied. If the sample in method B is mounted on an XY-stage, one can scan a whole region point-by-point, and collect the Raman data for each point.
The third method, C, makes use of two moving mirrors: one to illuminate a line region on the sample and a second to project this image on the entrance slit of the spectrometer. The spectra from the different points within this region are depicted on the CCD detector in such a way that there is a one to one correlation with a certain position on the sample and the position of the spectrum on the detector. This method does allow for confocality. The advantage of this method is that it is faster than the point-by-point method, and that local heating of the sample can be avoided. This is especially important for heat-sensitive samples.
In the following, we compare method B and C on the same sample. The measurement was performed on a sample with LOCOS isolation lines with decreasing width. LOCOS isolation gives compressive stress in the lines and tensile stress at the line edges[9]. The 457.9 nm laser line was used, 100x objective.
Fig. 6a shows the result of an experiment using method C (Fig. 5). The full circles denote $\Delta\omega$ (left scale), the dotted line shows the frequency of a plasma line in the spectrum (right scale). This frequency is independent of stress. So, if variations in this frequency are observed, they are due to instability of the system. Both parameters are plotted as a function of the pixel position on the detector. It can be seen in Fig. 6a that both the frequency of the Raman line and of the plasma line increase with increasing pixel number. This is due to a slight misalignment of the detector. Indeed, if the detector is slightly turned with respect to the mirror vibrations, the Raman lines will not be parallel to the side of the detector. This problem can easily be solved by putting the detector in the correct position. It is clear from this experiment that this alignment should be controlled before performing experiments using method C. The plasma line position can easily be used for this purpose. One can also use plasma lines to correct the Raman data for this misalignment. A line is fitted through the frequency of the plasma line (full line on the figure), and the frequency of the Si Raman line is corrected using this line. The result is shown in Fig. 6b. $\Delta\omega$ is taken equal to zero in the stress free regions. A positive shift indicates compressive stress, a negative shift indicates tensile stress. This figure gives a real picture of the variation of stress ($\Delta\omega$) with position on the sample.

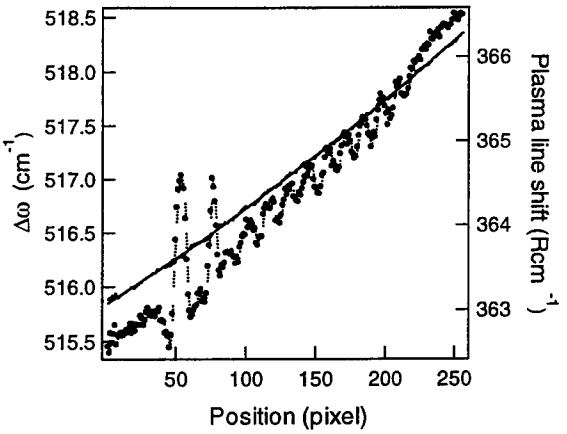


Figure 6a. Raman measurement across LOCOS isolation structures with decreasing width (3 μm down to 0.35 μm) using method C (Fig. 5). Full symbols: Raman shift (Δω), Lines: Plasma line shift and fit.

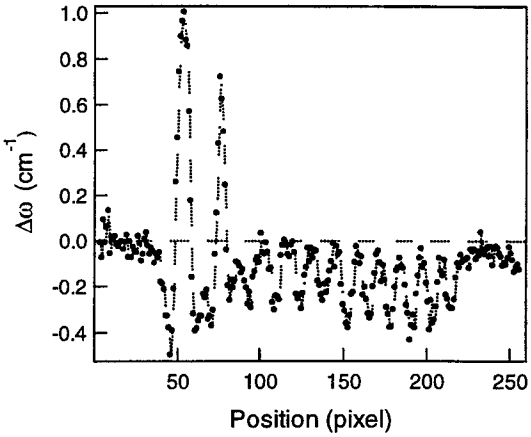


Figure 6b. Data from Fig. 6a after correction using the plasma line position.

In Fig.7, the scanning mode C (full line) is compared with the point-by-point mode B (dotted line) on a similar sample as shown in Fig. 6. Both measurement methods, B (point) and C (scan), show a similar result. Only the noise on the data is larger for the scanning method than for the point-by-point method. This is due to a smaller intensity of the Raman peak in the scanning mode for this experiment, and can be corrected for by using a higher laser power or a longer integration time.

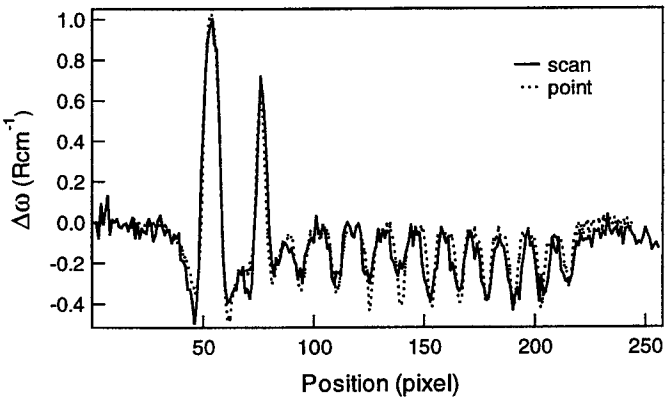


Figure 7. Comparison of method B (point) with method C (scan) (457.9 nm laser light).

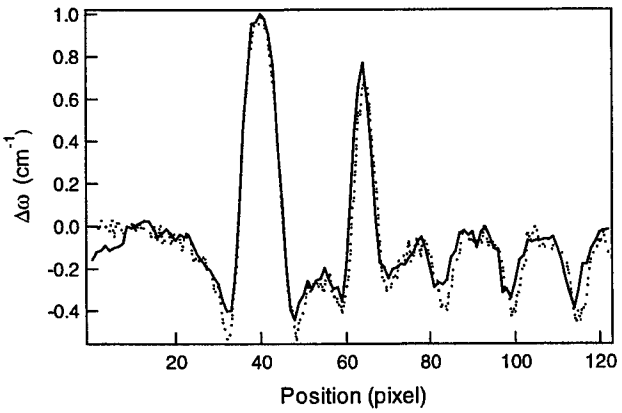


Figure 8. Comparison of the resolution obtained using method C (Fig. 5) with two different amplitudes for the first scanning mirror (457.9 nm laser light):
Full line: projection of the complete array of lines as shown in Fig. 7 on the full width of the detector.
Dotted line: projection of half of the array on the full width of the detector.

Fig. 8 compares results of two scanning mode experiments (method C in Fig. 5), but for different amplitudes of the first scanning mirror (different length of the laser line on the sample). For the full line, the image of the complete series of lines on the sample was measured and projected on the full width of the detector. For the dotted line, the image of only the first lines on the sample was measured, but projected on the same full length on the detector. It is expected that the spatial resolution is better for the latter. Only the area where both measurements overlap (first lines of the sample) is shown. Again the results correspond very well. But the smaller amplitude shows larger downshifts. This indeed indicates a better resolution.

3. MECHANICAL STRESS

Mechanical strain and stress often cause problems in microelectronics devices. These problems may be direct, such as crystal defects, film delamination and film cracking, or indirect such as effects of stress on mobility, resistance, band gap, doping distribution etc. Raman spectroscopy is the only analytical techniques that can monitor stress or strain in a non-destructive way, and with micrometer resolution.

Mechanical strain affects the frequency of the Raman peaks. By monitoring this frequency at different positions on the sample, a 'strain map' can be obtained with micrometer spatial resolution. In general, compressive stress will result in an increase of the Raman frequency, while tensile stress results in a decrease[9]. The exact relation between the frequency of each Raman peak of the material and the strain can be derived from the secular equation[9]. In general, this gives a complicated relation between the Raman frequency and all strain or stress tensor elements. Only if certain assumptions on the stress distribution in the sample are made, such as uniaxial or biaxial stress, a simple relation between Raman frequency and stress can be obtained. For example, for uniaxial stress σ along the [100] direction of Si, the stress induced shift of the Raman frequency of Si is given by:

$$\sigma(\text{MPa}) = -434\ \Delta\omega\ (\text{cm}^{-1}) \qquad (9)$$

or $$\Delta\omega\ (\text{cm}^{-1}) = -0.0023\ \sigma(\text{MPa})$$

Fig. 9 shows an experiment where external compressive (σ negative) stress was applied to a little square piece cut out of a Si wafer[15]. The shift of the Si Raman frequency from its stress free value, $\Delta\omega$, is plotted as a function of the applied stress. A linear relation is obtained, with a slope very close to the theoretically predicted one.

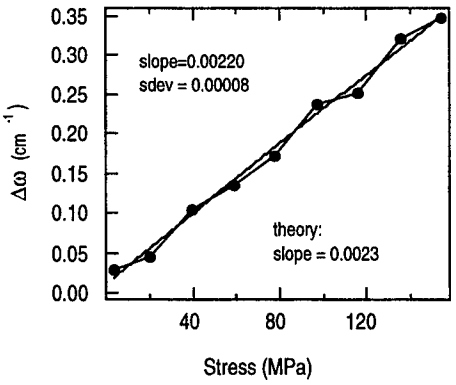


Figure 9. Shift of the Si Raman frequency as a function of uniaxial stress applied along the [100] axis (457.9 nm Argon laser).

The relation between the Raman shift and uniaxial stress along a certain crystal direction was derived experimentally for most materials which are of interest for microelectronics[16]. This relation is mostly linear or low order polynomial. In principle, Raman spectroscopy can be used to monitor the local stress for any material that shows a well defined Raman peak. This is certainly not limited to semiconductors.

An example of a local stress measurement using micro-Raman spectroscopy was already shown for the LOCOS array in Figs. 6-8. Another example is shown in Fig. 10. Here a stress measurement using the oil immersion lens is compared with

an experiment in which a conventional 100x lens is used. Si Raman spectra were measured during a scan across a sample with arrays of 1 μm wide, 150 nm thick, silicon nitride lines on silicon. Each 0.1 μm the spectrum of the silicon substrate was recorded and the frequency of the Raman peak was determined (457.9 nm laser line). Fig. 10 shows the Si Raman shift, $\Delta\omega$, from the stress free position ($\Delta\omega = 0$). The dotted line gives the shift as measured with the conventional 100x lens, the full line as measured with the oil lens. Both experiments indicate compressive stress under the nitride line ($\Delta\omega > 0$), as expected because silicon nitride itself is tensile, but we find a larger stress with the oil lens than with the conventional lens. In between the lines, tensile tress is expected. This is indeed observed with the oil lens. This stress is not resolved with the conventional 100x lens. The differences between the two experiments are due to the difference in spot size. The beam diameter for the conventional 100x lens is so large that, when focused on the nitride line, not only the compressive stress under the line but also part of the tensile stress near the edges of the line is measured. As a result, this compressive stress is underestimated. In the same way, the tensile stress in between the lines is not resolved, because the beam, when focused there, also partly probes the compressive stress region. This is not the case for the oil immersion lens, which, due to the smaller beam diameter, does resolve the local stress better.

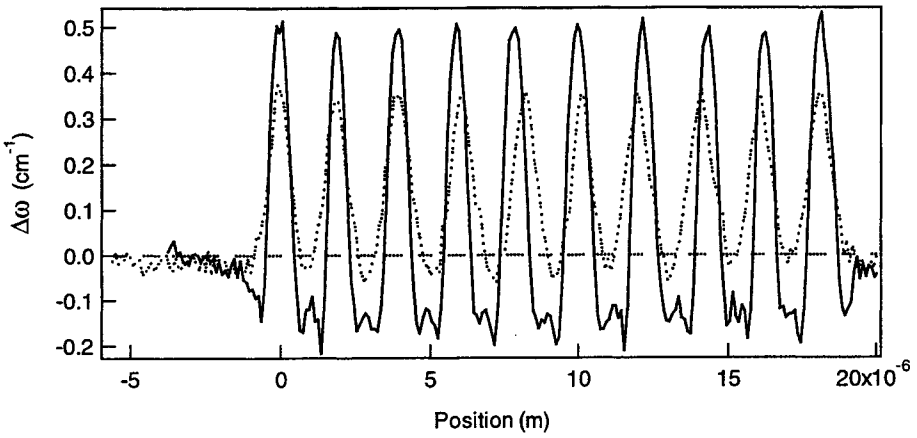


Figure 10. Measurement of mechanical stress in the silicon substrate under an array of 1 um wide silicon nitride lines with 1 μm spacing (457.9 nm Argon laser). Comparison of oil-immersion objective (full lines) with conventional objective (dotted lines).

Solder bumps allow three-dimensional chip interconnection and hence a substantially higher number of connections per unit area. The mechanical reliability remains strongly dependent on the properties of the bump material, and on the materials selected for bonding the bumps to the upper or lower levels in a device (chip or substrate board). The bump materials correspond in general to ductile alloys with a low melting point, while the materials bonding them to the other levels of the device are much stiffer (metals on ceramics). Their thermo-mechanical incompatibilities will generate stresses at the bump base, representing a potential source of mechanical failure in the device. Micro-Raman spectroscopy can be used to study these stresses. Examples are shown in the following two figures.

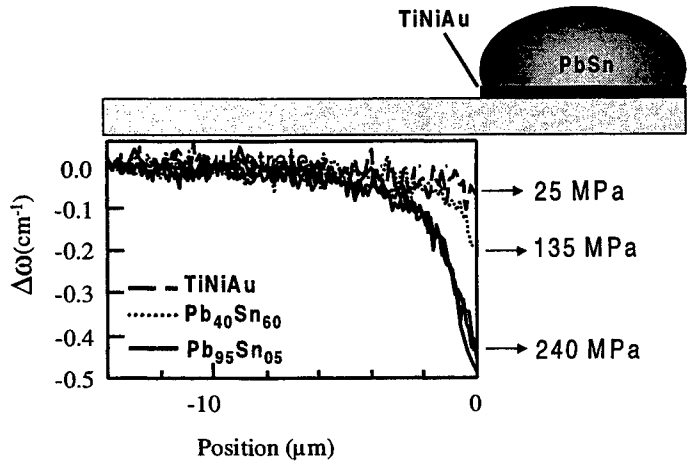


Figure 11. Measurement of stress in the silicon substrate near a solder bump (457.9 nm Argon laser).

In Fig. 11, the silicon Raman shift was monitored during a point-by-point scan starting far from the bump (position –14 μm) and moving towards the bump. Each 0.1 μm a Raman spectrum was measured. The purpose of this study was to analyse stress induced by only the TiNiAu pad and by two different bump alloys ($Pb_{40}Sn_{60}$ and $Pb_{95}Sn_5$). The results clearly show a negative shift of the Raman frequency near the bump, indicating tensile stress in the substrate. This tensile stress is very small (25 MPa) if only the TiNiAu pad is present, increases for $Pb_{40}Sn_{60}$ (about 135 MPa) and is clearly maximal for $Pb_{95}Sn_5$ (about 240 MPa). Another example in shown in Fig. 12. Here the stress in the silicon substrate near a solder bump is measured in a packaged device. The package was polished from the side, and Raman spectra were measured on the cross-section. The spectra were measured during a scan close to the Si surface, passing near to the bump/Si interface (see top Fig. 12). The measurement was difficult because some stress was introduced by the polishing. However, we clearly found a positive shift of the Raman peak just underneath the bumps, and a negative shift near the bump edges. This shows that Raman spectroscopy can also be used to investigate packaging induced stress. It is for example well possible to compare stress near bumps located at the center of the chip with the one near bumps located at the edges.

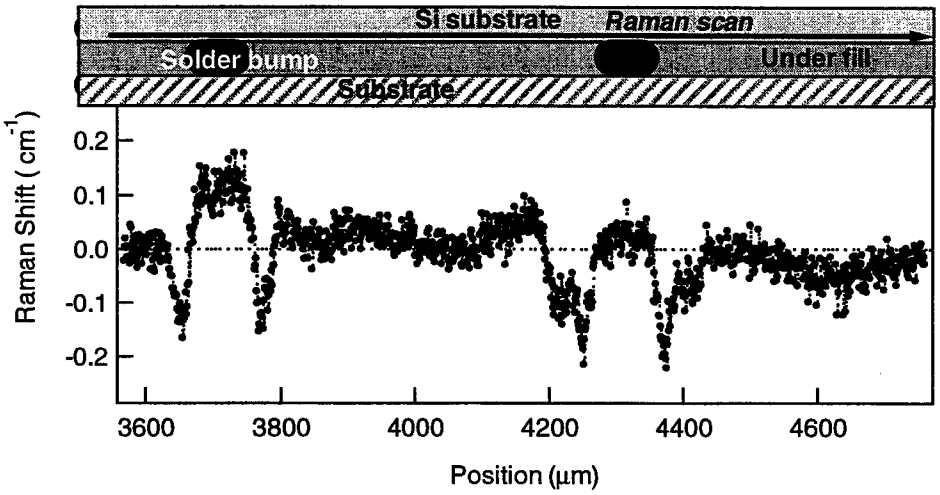


Figure 12. Measurement of the silicon Raman frequency shift on the cross-section of a packaged device. The arrow indicates the measurement position, i.e. close to the Si/under fill and bumps interface.

Raman measurements are of course also perfectly possible for the investigation of two-dimensional stress distributions. This can be done using method B or C of Fig. 5. Fig. 13 shows a 2D picture of the shift of the Si Raman frequency near a square (3 μm x 3 μm) silicon island isolated by field oxide, measured using the point-by-point method (method B of Fig. 5). The dark black regions near the sides indicate tensile stress. This tensile stress is clearly smaller near the edges. The bright region (white in the picture) inside the square indicates compressive stress. It is smaller in the center of the square (gray), but clearly larger at the edges. At the right side of the picture, another dark region is visible. This is due to the presence of another square silicon island with tensile stress along its perimeter.

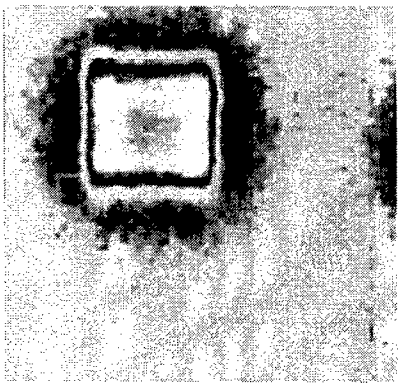


Figure 13. Demonstration of a two-dimensional measurement of the shift of the Si Raman frequency peak near a square Si island isolated by field oxide (LOCOS). Black indicates tensile stress, white indicates compressive stress.

A last experiment shows the stress introduced by back-grinding of a wafer. Fig. 14a compares the spectrum of a control crystalline Si wafer (dotted line) with a spectrum measured on a back-grinded surface (full line). The latter shows in addition to the crystalline Si peak at 521 cm^{-1} also a broad peak near 480 cm^{-1}. This peak indicates the presence of amorphous Si on the back-grinded surface. The narrow peak at 480 cm^{-1} is a plasma line of the laser. This peak was for clarity removed from the a-Si spectrum. Interesting is also that the cSi peak of the back-grinded sample is shifted towards higher frequencies than the one of the control sample. This indicates that back-grinding introduces compressive stress in the silicon. To study this further, a micro-Raman spectroscopy scan was performed starting from the bulk of the sample towards the damaged surface. To do that, a small piece of the sample was mounted under an angle on a metal holder, as shown in scheme in Fig. 14b. The sample was polished using very fine polishing paper and mounted under the microscope of the micro-Raman instrument. A scan was then performed as indicated in Fig. 14b. The open symbols in this figure show the intensty of the Si Raman peak. When crossing the edge of the sample (at the back-grinded surface) the intensity decreases. The dotted vertical line located at position about 9 μm indicates the position surface of the sample (at 50% of the maximal intensity). The full symbols show the position of the Raman peak. It clearly increases from stress free (0 cm^{-1}) to about 1.4 cm^{-1} at the surface. Assuming biaxial stress at the surface, this would correspond with a compressive stress (due to the back-grinding) of about 600 MPa ($= \sigma_{xx} + \sigma_{yy}$).

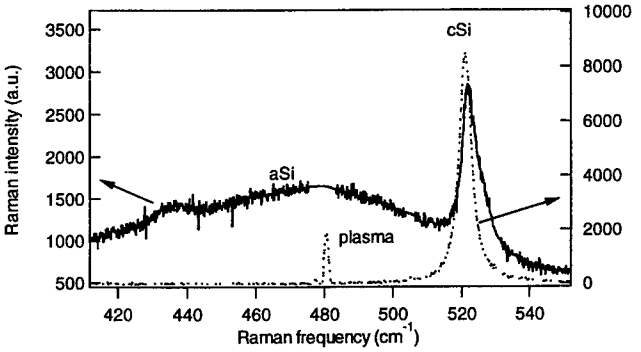


Figure 14a. Raman spectra of control cSi sample (dotted line) and back-grinded Si wafer (full line).

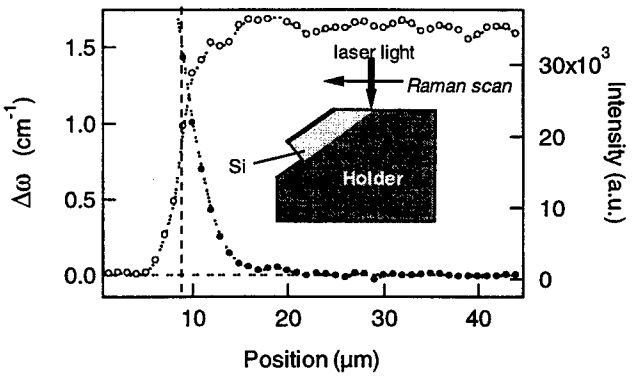


Figure 14b. Intensity (right axis) and Si Raman shift (left axis) measured during a scan from the bulk (light gray) to the back-grinded surface (black side) of a Si wafer. Experimental set-up is shown in inset.

4. TEMPERATURE

Micro-Raman spectroscopy is a sensitive probe for the local temperature in semiconductor devices. Information on the temperature can be extracted from the spectrum in three different ways: from the Stokes/anti-Stokes intensity ratio, from the frequency, and from the width of the Raman peak. The Stokes/anti-Stokes intensity ratio, I_S/I_{AS}, offers a means to directly determine the temperature, without the need for a reference measurement at a known temperature. Using the Stokes/anti-Stokes intensity ratio to derive the temperature works in general only well at relatively high temperatures, because of the small intensity of the anti-Stokes Raman peak at low temperature (up to 400 K). The typical error on the temperature estimation is 5 to 10 K. The temperature dependence of the Raman frequency, ω, offers often a more accurate and faster way to control the temperature of semiconductor devices. Indeed, in the temperature range of interest for most devices (room temperature to 400 K), the relation between ω and T can be assumed to be linear for most semiconductors. For Si, the relation between the Raman peak position and the temperature change is given by[16]:

$$\Delta\omega\,(\mathrm{cm}^{-1}) = -0.025\,\Delta T\,(\mathrm{K}) \qquad (10)$$

So, an increase in temperature will result in a decrease of the Raman peak frequency. With this method, an accuracy as good as 1 K can be obtained. The disadvantage is that one calibration experiment, at a known temperature, is required. The dependence of the peak width on the temperature is in general not used in practice.

Figure 15 shows the frequency shift of the Si Raman peak, $\Delta\omega$, as a function of the output power of the laser used to excite the sample. The 514 nm light of an Argon laser was focused through a 100x objective on the sample. The power on the sample was about 10 times smaller than the output power of the laser. Note that even for very small values of the output power, a shift of the Si peak is detected, indicating a laser induced heating of the sample. Using the value given in Eq. 10 for Si, the local heating of the sample can be deduced as a function of the laser power.

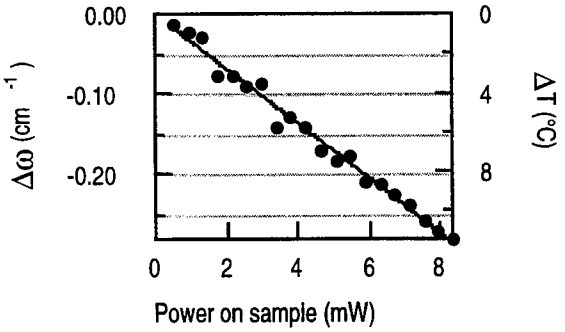


Figure 15. Temperature shift (ΔT) and Raman shift ($\Delta\omega$) in Si due to changes in the power of the focused laser beam.

This temperature sensitivity of the Raman spectrum can be used to probe the local temperature in semiconductor devices, for example along the width of the gate of a MOSFET during operation of the device. Such experiments were first reported by the group of Abstreiter[17]. The following figures show an example of an experiment we performed on a a 50 μm wide grounded-gate nMOS device during avalanche and snap-back conditions[18]. Such devices are used for ESD protection. The objective of this experiment was to determine the local temperature in the nMOS device during the different operation stages from avalanche to snap-back, and its possible relation with light emitted by this device. A layout of the investigated device is shown in Fig. 16. Photon Emission Microscopy (PEM) was used to monitor the current-dependence of the light emission. The PEM results obtained at different currents are shown sequentially in Fig. 18a. The device has in these emission photos the same orientation as in the layout shown in Fig. 16. At 1 mA, localized emission spots during avalanche breakdown can be observed at the device corners. Avalanche breakdown is expected to occur first at positions with highest electrical field, which is indeed at the corners[19]. At 5 mA, light is visible along the whole width of the gate, indicating that avalanche breakdown occurs in this whole region. At 10 mA, snapback triggers but only at the right side of the device. Only at higher currents snapback occurs across the whole width of the device. This non-uniform triggering of ESD protection devices was discussed previously by Russ et al.[19]. They also showed that this non-uniform snap-back can jump from one side of the device to the other side, which was referred to as the 'ping-pong' effect.

Micro-Raman spectroscopy experiments were performed on the same gg-nMOS, using the 457.9 nm line of the Argon laser. The Raman spectrum of silicon was measured each 0.1 μm during a scan along the width of the gate (50 μm, see Fig. 16). When no voltage was applied, a negative Raman shift of about -0.3 cm^{-1} was observed along the entire width of the gate, as shown in Fig. 17 (0 mA). This indicates that there is about 130 MPa tensile stress near the gate (assuming uniaxial stress, Eq. 9). When applying a constant current to the device, the measured Raman frequency did decrease, as can be seen in Fig. 17 for 10 mA. In order to distinguish between stress and temperature, the Raman data measured at zero bias were subtracted from the data measured at non-zero biases. The results are shown in Fig. 18b. It is clear that for increasing currents, the local temperature also increases ($\Delta\omega$ becomes more negative). The local temperature, calculated from the Raman shift using Eq. 10, is shown on the right axis. We did not observe a clear correlation between these Raman results and the PEM pictures at low currents (< 5 mA). This is partly due to the generation of electron-hole pairs near the gate by the laser light used to excite the Raman spectra. At 5 mA (7.2 V), the temperature seems to be uniform across the device. In this case there was probably avalanche across the whole width of the device (picture of 5 mA in Fig. 18a). At 10 mA (7.16 V) the device is in snapback. The overall temperature along the gate increases, with a maximum at the left side. This could correspond with a local triggering of the device. For 26 mA (7V) this maximum moves from left to right. This can be explained by the so called ping-pong effect. At 30 mA (7V) the maximum temperature is again located at the left side of the gate. For higher currents (75 mA (7.7 V) and 100 mA (8V)) the device is completely triggered (as for the 50 mA and 100 mA pictures in Fig. 3). The temperature profile also becomes more uniform. Sharp temperature peaks become visible at the edges of the gate. In the 100 mA picture of Fig. 3, one can observe an upward extension of the emitting region at the edges of the gate. They could explain the temperature peaks. These Raman results show a clear correlation between the local temperature along the gate width and the PEM images. They also demonstrate that Raman spectroscopy is useful for in-situ local temperature measurements.

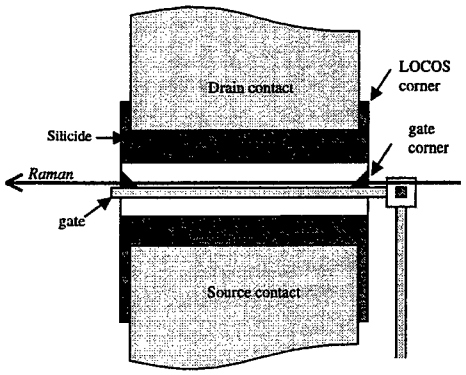


Figure 16. Layout of a gg-nMOS.

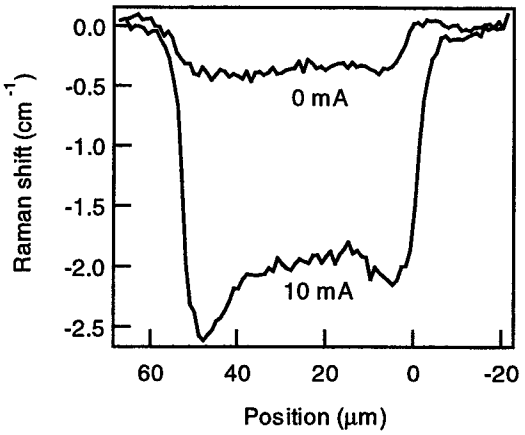


Figure 17. Raman frequency shift from the stress free value (left axis) for a gg-nMOS device (same as shown in Figs. 3 and 4) under 0 mA and 10 mA bias. The dotted line denotes intensity of the Raman peak for 0 mA (right axis). The arrow denotes a local minimum in the intensity at the gate-source metal contact.

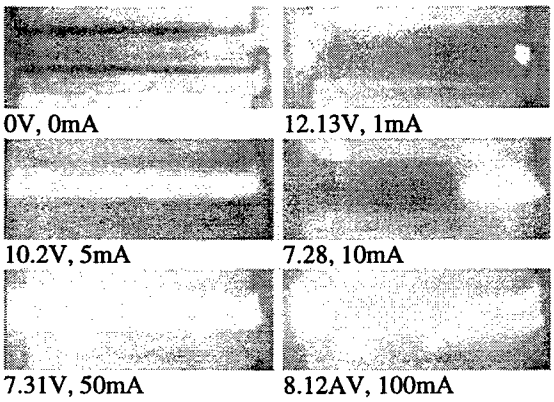


Figure 18a. PEM images of a 50 μm wide gg-nMOS for different applied currents.

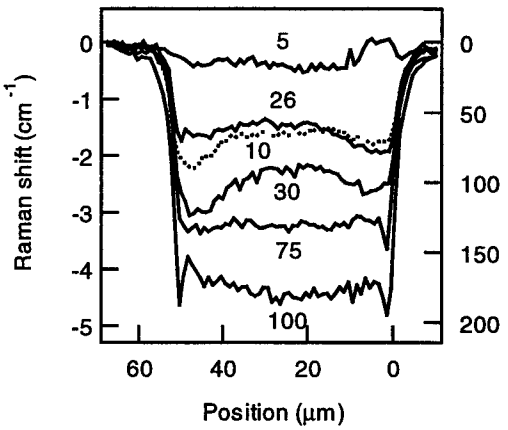


Figure 18b. Raman frequency shift (left axis) and corresponding increase in lattice temperature (right axis) of a 50 μm wide gg-nMOS (same device as in Fig. 3) for different applied currents (as indicated on the graphs, in mA).

5. CONCLUSIONS

In this paper it was demonstrated that micro-Raman spectroscopy is a very interesting technique to study the local mechanical stress and temperature in silicon devices. The spatial resolution of a conventional instrument is about 1 μm. This can be improved in a simple way by using an oil-immersion objective and by applying convolution techniques to the Raman data. The sensitivity for stress is about 10 MPa. in silicon. It is the only technique that provides information on stress in a non-destructive way, with reasonable resolution and without the need for special sample preparation. It was also shown that micro-Raman spectroscopy is useful for local temperature measurements.

ACKNOWLEDGEMENTS

Part of this work was funded by the EC project 'Nostradamus' (SMT4-CT95-2024 DG12-RSMT).

REFERENCES

1. J.H. Parker Jr., D.W. Feldman and M. Ashkin, *Physical Review*, **155** (3), 712-714, 1967.
2. K.F. Dombrowski, B. Dietrich, I. De Wolf, R. Rooyackers and G. Badenes. Proc. 29th European Solid-State Device Research Conference (ESSDERC'99), Ed. H. E. Maes, R. P. Mertens, G. Declerck and H. Grunbacher (Editions Frontieres), pp. 196-199, 1999
3. M. Delhaye and E. Da Silva. *Publication of ISA (Dilor-Jobin Yvon-Spex) instruments.*
4. M. Abramowitz, *Handbook of mathematical functions*, 9th ed. , Dover publications, New York 1972
5. W. Duncan, *J. Vac. Sci. Tech. A* **14**(3), pp.1914-1918, 1995
6. S. Webster, D.N. Batchelder, and D.A. Smith, *Appl. Phys. Lett.* **72**(12), pp. 1478-1480, 1998
7. J. Grausem, B. Humbert, A. Burneau and J. Oswalt, *Appl. Phys. Lett.* **70**(13), pp. 1671-1673, 1997
8. D. Zeisel., B. Dutoit, V. Deckert. T. Roth and R. Zenobi, *Anal. Chem.* **69**, pp. 749-754, 1997
9. I. De Wolf, *Semicond. Sci. Technol.* **11**, 19, 1996
10. C.D. Poweleit, A. Gunther, S. Goodnick and J. Menéndez, *Appl. Phys. Lett.* **73**(16), pp. 2275-2277, 1998
11. L.P. Ghislain and V.B. Elings, *Appl. Phys. Lett.* **72**(22), pp. 2779-2781, 1998
12. F. Van der Heijden, *Image based measurement systems, object recognition and parameter estimation*, John Wiley & Sons Ltd., Chichester, 1994
13. M. Delhaye and P. Dhamelincourt, *J. Raman Spectr.* **3**, pp. 33-43, 1995
14. M. Delhaye, P. Dhamelincourt and E. Da Silva. In Proc. of the XVIII International Congress on Raman Spectroscopy (ICORS), Grenoble, 1975
15. M. Ignat, Personal communications.
16. I. De Wolf, J. Jimenez, J.-P. Landesman, C. Frigeri, P. Braun, E. Da Silva and E. Calvet, *Catalogue of optical and physical parameters*. Catalogue composed in the frame of the EC contract Nostradamus (SNT4-CT95-2024 DG12), Publication of the Euroepan Commission, EUR 18595, 1999.
17. R. Ostermeir, K. Brunner, G. Abstreiter and W. Weber, *IEEE Transactions on Electron Devices*, 39 (4), pp. 858-863, 1992.
18. M. Rasras, I. De Wolf, G. Groeseneken, J.Chen, K. Bock and H. E. Maes. To be published in the proceeding of ISTFA '99, Nov. 1999, Santa Clara.
19. C. Russ, K. Bock, M. Rasras, I. De Wolf, G. Groeseneken, H. E. Maes, EOS/ESD conference proceedings, 1998.

Full Field Stress/Birefringence Analysis with a Polarizing Microscope

Anand Asundi, Bing Zhao

School of Mechanical & Production Engineering
Nanyang Technological University
Nanyang Avenue, Singapore 639798

ABSTRACT

An automatic method of full-field stress measurement using a transmitting micro-polariscope is proposed. A compact optical transmitting polarizing microscope with white light source is rebuilt by developing a loading and recording system, in order to perform a tension test. Both isoclinics and isochromatics are measured in real-time with phase shifting technique. A new simple algorithm for isochromatics is proposed. It is found to be suitable for the current micro-polariscope.

Keywords: Photoelasticity, polarizing microscope, phase shifting, isochomatic, isoclinic, full-field measurement.

1. INTRODUCTION

Mechanical properties of materials such as polymers depend on the phenomena that occur on a microscopic level. For example, dimensions of inclusions in polymer composites are often below 200 to 300 μm, and the regions with concentrations of stresses around inclusions have a similar range[1]. The micro phenomena such as generation and propagation of microcracks and shear bands often play a decisive role in overall mechanical performance. The optical microscope is a ubiquitous tool in micro-mechanics, is often used to determine micro-deformation. Among the various optical microscopes, polarizing microscope is used to determine the birefringence of minerals and polymers. For example, Pawlak and Galeski[1] used a micro-polariscope to measure full-field characteristic parameters in polymer composites with inclusions. The authors also analyzed the influence of optical elements of microscope on the state of polarization and retardation of the light. Additionally, the polarizing microscope is used in frozen stress analysis[2]. Fessler et al.[3] have developed a scanning micro-polariscope for automatic stress analysis. It should be emphasized that all these micro-photoelastic systems are based on pointwise analysis.

Phase-shifting technique as a full-field fringe detection technique has been applied to most optical techniques, such as interferometry, holography, moiré and speckle. It possesses an outstanding advantage: it is a punctual analysis method, as the measurement on each pixel is carried out independently, and without the influence of the neighboring pixels. Thus the phase-shifting technique is a better approach for the micro-measurement where high spatial resolution is needed. In references 4~7, some works on the applications of phase-shifting technique in photoelasticity are reviewed. To our knowledge, works on photoelastic phase-shifting technique by using micropolariscope have not been reported. It should be emphasized that it is difficult to adopt the developed photoelastic phase-shifting technique to polarizing microscope, due to hardware restrictions, especially for isochromatics measurement.

The present paper realizes a full-field isochromatic fringe measurement with phase-shifting technique. A real-time tensile test under a Leitz polarizing transmitting microscope is demonstrated. In section 2, the experimental set-up is introduced. In section 3, the background of phase shifting technique and its application on photoelasticity is presented. A simple 2-phase-step algorithm is proposed. Some questions are answered: why some of previous phase shifting algorithms for isochromatics can not be used for our microscopic measurement; how to get the dark and bright field data of pure isochromatics from a plane polariscope etc.

2. EXPERIMENTAL SYSTEM

The optical system consists of a Leitz-Orthoplan II Pol polarizing microscope (Fig. 1a), an epoxy resin specimen and an image acquisition & processing system. The transmitting microscope includes a white light source, a filter, a field and aperture diaphragm, an objective, a tilting compensator, two quarter wave plates (QWP), a polarizer and an analyzer as shown schematically in Fig. 1b, and a pair of eyepieces shown in Fig. 1a. The analyzer can be rotated with a precision of

Part of the SPIE Conference on Advanced Photonic Sensors and Applications
Singapore • November–December 1999
SPIE Vol. 3897 • 0277-786X/99/$10.00

0.1°. The upper (output) QWP and tilting compensator can only be used one at a time, as they are inserted in the same place of the microscope. The bottom QWP can be rotated through 90°, however the upper QWP cannot be rotated. The specimen is placed between the condenser and the objective of the microscope. Three types of objectives 6.3/0.2, 10/0.25, 25/0.50 and 40/0.65 (Magnification / Numerical Aperture) are used. A 10 × eyepiece can be used for visual observation. The image recording and processing system is composed of a CCD Mintron MTV - 1802CB, image grabber MVP 512× 512× 8 bits, a Pentium personal computer and software. The CCD is directly connected to the microscope eyepiece tube, without using lens. The loading system that was specially designed for this microscope consists of a loading frame and a two-axis translation stage. Fig. 2 shows a picture of the loading frame. The frame is fixed onto the translation stage which in turn is fixed on to the mechanical stage of the microscope. The specimen to be tested can be translated in two directions and can rotate with the loading frame. So, we can easily examine the different areas of the specimen. The specimen must be transparent or translucent, and is bonded onto the loading frame. Rotation of the gyro-stat loads the specimen in tension.

3. PHOTOELASTIC PHASE-SHIFTING TECHNIQUE

3.1 Photoelastic phase-shifting measurement

The phase-shifting technique for photoelasticity was first introduced by Hecker and Morche[10] in 1986. In 1991, Patterson et al.[11] and Barone[12] extended Hecker and Morche's approach by using a circular polariscope to realize a phase shifting simultaneously for isoclinic and isochromatic fringe measurement. In 1992, Sarma et al.[13] reported a methodology using a plane polariscope to extract both isoclinic and isochromatic phase fields. The method is based on 3 phase-step patterns obtained by rotating the analyzer. Later, in 1993, Asundi[14] extended the tardy method of compensation from a pointwise approach to evaluate the fractional fringe order on all points lying on the isoclinic line. In 1995, Zhao et al.[9] proposed a method using a circular polariscope to determine full-field isochromatics. Two dark and bright field patterns of circular polariscope are used in Eq. (4). In 1996, Buckberry and Towers[4] proposed a 3-phase-step algorithm based on a plane polariscope. Comparing the three algorithms for isochromatics, i.e., Sarma et al.'s[13] algorithm, Zhao et al.'s[9] algorithm, and the algorithm of Buckberry and Towers[4], we find that the algorithms possess some similar properties. For example, all of them are of arc-cosine form; No QWP is used as phase shifter, and phase shifting is realized by rotating analyzer or both analyzer and polarizer; However, they are sensitive to the background/stray light.

It is easy to measure isoclinic fringe with phase-shifting technique[4-7, 9-13] with a crossed plane polariscope (see Fig. 3(a)). However, it is relatively difficult to apply phase-shifting technique to isochromatic fringe measurement, as the shift of isochromatic fringe phase will also provoke the shift of isoclinic fringe phase in most case. The simplest way of using phase-shifting technique for isochromatic parameter measurement is only using the bright and dark fields in a circular polariscope[4,9]. The circular polariscope is illustrated in Fig. 3(b). The chromatic fringe pattern of the dark field can be treated as the result of 180° phase shifting of that of bright field. The bright (I_1) and dark (I_2) field can be expressed by following Equation:

$$\begin{cases} I_1 = I_0 \cos^2(\varphi) = I_0(1+\cos\varphi) \\ I_2 = I_0 \sin^2(\varphi) = I_0(1-\cos\varphi) \end{cases} \tag{1}$$

where I_0 is the total intensity transmitted by the polariscope, θ, isoclinic parameter, is angle of the first principle stress with respect to the polarizer axis, it represents the primary principle direction; φ is the isochromatic value of optical phase retardation. We proposed the following algorithm to calculate the chromatic fringe order[9]:

$$\cos\varphi = \frac{I_1 - I_2}{I_1 + I_2} \tag{2}$$

This algorithm has been used for isochomatics measurement and will be extended to micropolariscope in present paper.

3.2 Phase Shifting with Micropolariscope

The measurement is carried out with a plane polariscope, i.e., all QWPs are out of the optical path of the microscope (see Fig. 3). The important step is to get a "pure" isochromatic fringe pattern from a plane polariscope using phase shifting technique. Generally, the "pure" isochromatic fringe pattern needs a circular polariscope. We do not use a circular polariscope under the microscope, because:

- QWP is only accurate for a specific wavelength, so a micropolarscope without QWPs is inherently less error prone[4];
- QWP will cause an attenuation of the maximum intensity in dark field, and an increase in the minimum intensity in the bright field, for a circular polariscope in monochromatic and white light[15];
- Avoid the problem caused by using two polariscopes for isoclinics and isochromatic, i.e., plane and circular polariscope.

The fringe intensity from a parallel plane polariscope (bright field) can be expressed as

$$I_{11} = I_0\left[1 - \sin^2(2\theta)\sin^2(\varphi/2)\right] \tag{3}$$

If the isoclinic phase is shifted 45°, which is achieved by simultaneously rotating the polarizer and analyzer by 45°, a new fringe pattern is obtained:

$$\begin{aligned} I_{12} &= I_0\left[1 - \sin^2(2\theta + 90^0)\sin^2(\varphi/2)\right] \\ &= I_0\left[1 - \cos^2(2\theta)\sin^2(\varphi/2)\right] \end{aligned} \tag{4}$$

The intensity field for a crossed plane polariscope (dark field, see Fig. 3(a)) can be written as:

$$I_{21} = I_0 \sin^2(2\theta)\sin^2(\varphi/2) \tag{5}$$

Similarly, if the polarizer and the analyzer are rotated by 45°, we have

$$I_{22} = I_0 \cos^2(2\theta)\sin^2(\varphi/2) \tag{6}$$

Adding together Eqs (3) and (4), we have:

$$I_1 = I_{11} + I_{12} = I_0 \cos^2(\varphi/2) = I_0(1 + \cos\varphi) \tag{7}$$

The above equation is of same result of bright field of a circular polariscope. Adding together Eqs (5) and (6), we get the pure isochromatics which is equal to that of dark field of a circular polariscope:

$$I_2 = I_{21} + I_{22} = I_0 \sin^2(\varphi/2) = I_0(1 - \cos\varphi) \tag{8}$$

The above procedure shows clearly that how to extract the isochromatics from the mixing fringe patterns of isoclinics and isochromatics by using phase shifting technique. Buckberry and Towers[4] used this method to extract the "pure" isochromatics of dark field from a crossed plane polariscope, i.e., Eq. (8). Fig. 4 shows an example of obtaining pure isochromatics from a crossed plane micropolariscope with phase-shifting technique. The specimen is a rectangular epoxy resin plate with a crack. The objective magnification is 40×. Fig. 4(a) and (b) are photographies of the two dephased fringe patterns. Fig. 4(c) shows the isochromatics distribution around the crack.

Combining Eqs (2) and (7,8), we obtain a new algorithm for isochromatics:

$$\cos\varphi = \frac{I_1 - I_2}{I_1 + I_2} = \frac{I_{11} + I_{12} - I_{21} - I_{22}}{I_{11} + I_{12} + I_{21} + I_{22}} \tag{9}$$

This phase shifting technique possesses several useful advantages. First, the method is of simple form. Four patterns are needed and the manipulation of phase shifter is simple. Second, the calculation of isochromatics is independent of the isoclinics. That is to say, the error of calculated isoclinics will not influence the calculation of isochromatics. Thirdly, this algorithm can decrease the noise influence, as both the denominator and the numerator are the average of 4 patterns.

Fig. 5 shows a test result of isochromatics by using above equation. A rectangular specimen with size of 70×8×1.5 is examined, and the measurement is carried out in real time. Fig. 5(a) is the result of $I_{11} + I_{12}$, i.e., isochromatics of bright field, and 5(b) is the result of $I_{21} + I_{22}$, i.e., isochromatics of dark field. In Fig. 5(c), the isochromatics phase distribution around the crack is displayed in 16 pseudo-colors.

As with Sarma et al.[11], and Buckberry and Towers[4]'s allgorithms, the above approach does not take into account the effect of background/stray light. To resolve this problem the unstressed model is captured with the polarizer and analyzer in the crossed (or parallel) position. The intensity values of this image should be subtracted from all other images captured after loading before any processing is done.

4. DISCUSSION AND CONCLUSION

A new full-field micro-photoelastic system has been described. By developing and rebuilding an ordinary transmitting polarizing microscope, a system for full-field micro stress measurement with high resolution and sensitivity can be realized. This gives the polarizing microscope a new important use: real time mechanical test and photoelastic measurement. This measurement system possesses a higher automation and processing speed than existing pointwise system

To achieve high spatial resolution micro-mechanics measurement, a combination of microscopy with phase shifting technique is adapted to photoelastic measurement. A new phase shifting method for isochromatics is proposed. This phase shifting technique possesses several useful advantages. This algorithm is not limited to be used to micropolariscope measurement, but can be used in all the photoelasticity and birefringence measurement.

Further work on phase unwrapping is needed for improving the automation and precision of the measurement.

ACKNOLOGEMENTS

The support of Ministry of Education and Nanyang Technological University through research grant MLC1/97 is acknowledged.

REFERENCES

1. Pawlak, A. Galeski, "Measurements of characteristic parameters in polymer composites by means of a micropolariscope," *Optical Engineering*, **34**(12), pp. 3398-3404, 1995.
2. D.J. White, A.S. Tickner, "A transmission microscope polariscope for photoelastic analysis," *J. Scientific Instruments* (*J. Phys E*), **1**, p. 649, 1968.
3. H. Fessler, R.E. Marston, E. Ollerton, "A micropolariscope for automatic stress analysis," *The Journal of Strain Analysis for Engineering Design*, **22**, pp. 25-35, 1987.
4. C. Buckberry, D. Towers, "New approaches to the full-field analysis of photoelastic stress patterns," *Optics and Lasers in Engineering*, **24**, pp. 415-428, 1996.
5. K. Ramesh, V. Ganapathy, "Phase-shifting methodologies in photoelastic analysis - The application of Jones calculus," *Journal of Strain Analysis*, **31**(6), pp. 423-432, 1996.
6. J. Carazo-Alvarez, S.J. Haake, and E.A. Patterson, "Completely automated photoelastic fringe analysis," *Optics and Lasers in Engineering*, **21**, pp. 133-149, 1994.
7. K. Ramesh, S.K. Mangal, "Data acquisition techniques in digital photoelasticity: a review," *Optics and Lasers in Engineering*, **30**(1), pp. 53-75, 1998.
8. P. Hariharan, B.F. Oreb, T. Eiju, "Digital phase-shifting interferometry: a simple error-compensating phase calculation algorithm," Chapter 14, *Optical Shop Testing*, edited by D. Malacara, John Wiley & Sons, Inc, New York, 1992.
9. B. Zhao, S.D. Guo, L.C. Zhang, R.H. Fang, "Computer aided photoelastic measurement," *Proceeding of 8th national conference on experimental mechanics*, Harbin Institute of Technology Publish, Harbin, China, pp.318-321, 1995.
10. Hecker, B. Morche, "Computer-aided measurement of relative retardations in plane photoelasticity," in *Experimental Stress Analysis*, edited by H. Wieringa, Martinus Nijhoff Publishers, Dordrecht, The Netherlands, pp. 535-542, 1986.

11. E.A. Patterson, Z.F. Wang, "Towards full-field automated photoelastic analysis of complex components," *Strain*, **29**(1), pp. 49-53, 1991.
12. S. Barone, E.A. Patterson, "Full-field separation of principal stresses by combined thermo- and phoroelasticity," *Experimental Mechanics*, **36**(4), pp. 318-324, 1996.
13. AVSSSR. Sarma, S.A. Pillai, G. Subramanian, T.K. Varadan, "Computerized image processing for whole-field determination of isoclinics and isochromatics," *Experimental Mechanics*, **32**, pp. 24-29, 1992.
14. A. Asundi, "Phase shifting in photoelasticity," *Experimental Techniques*, **17**, pp. 19-23, 1993.
15. A. Ajovalasit, S. Barone, G. Petrucci, "Automated photoelasticity in white light: influence of quarter-wave plates," *Journal of Strain Analysis*, **30**(1), pp. 29-34, 1995.

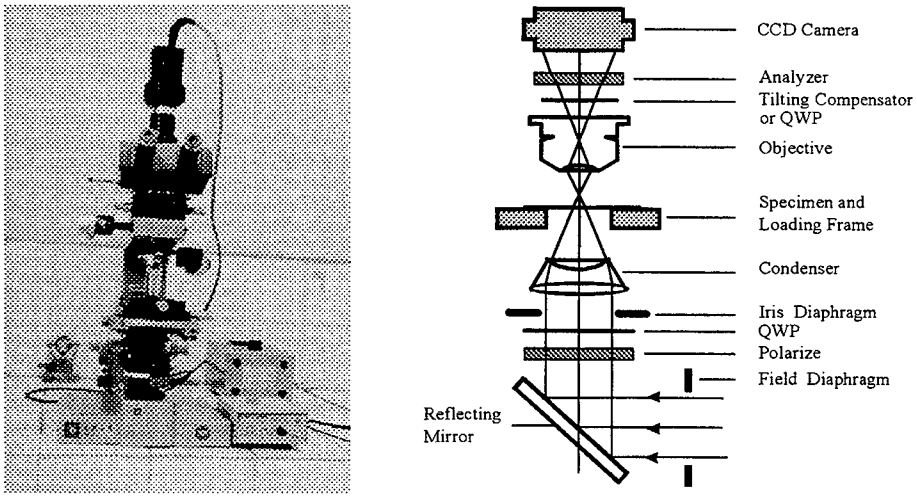


Fig. 1 Fig. 1 Sketch of optical system and principle.
(a) Photo of Leitz-Orthoplan II Pol polarizing microscope;
(b) Optical components of the measurement system, QWP is quarter wave plate.

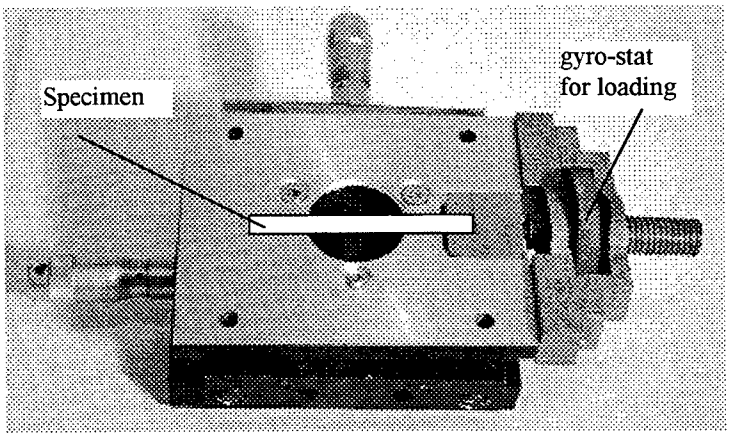


Fig.2 Photography of loading frame.

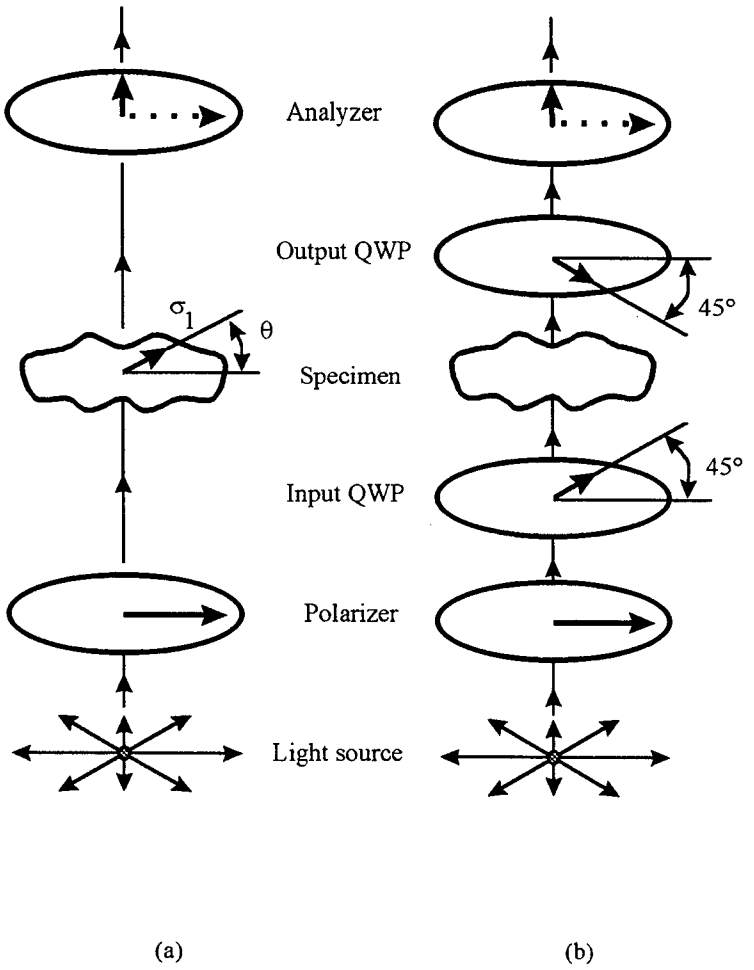
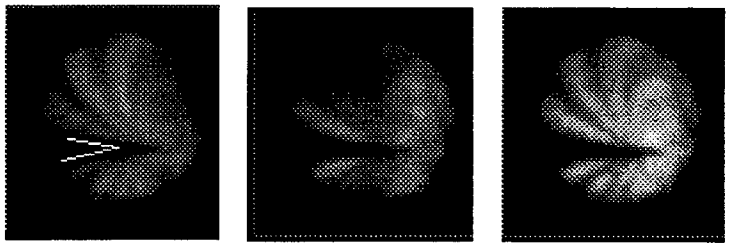
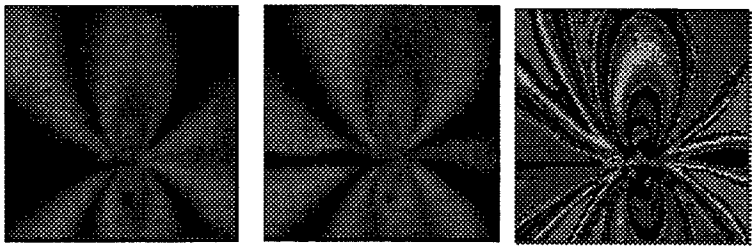


Fig. 3 (a) Plane polariscope; (b) Circular polariscope. QWP is Quarter wave plate. The arrow in QWPs represent the fast axis direction. σ_1 is primary principal stress, and θ is isoclinic angle.



(a) (b) (c)

Fig. 4 Example of obtainning pure isochromatics from a crossed plane micropolariscope with phase-sh technique. (a) fringe pattern from the crossed plane polariscope; (b) fringe pattern from the crossed polariscope, but the analyzer and polarizer is rotated by 45°. (c) pure isochromatics which is the digital superpo of (a) and (b)



(a) (b) (c)

Fig. 5 Isochromatics measurement example. (a) isochromatics of bright field, result of the superposition of two dephased patterns from a crossed and a parallel plane polariscope; (b) isochromatics of dark field, result of the superposition of two dephased patterns from a parallel plane polariscope; (c) Isochromatics phase distribution around a crack expressed by 16 pseudo -colors. It is phase shifting result with (a) and (b).

Grid Method for Strain Measurement in Electronic Packaging Using Optical, Electronic, and Atomic Force Microscope

Bing Zhao, Anand Asundi, Kim Eng Oh

School of Mechanical & Production Engineering
Nanyang Technological University
Nanyang Avenue, Singapore 639798

ABSTRACT

The grid method is combined with microscopy for micro deformation measurement. High density grids of 1200 lines/mm are transferred to electronic packaging specimens at high temperature. The residual stress distribution is recorded on the grid after it was cooled to room temperature. The grid pattern is magnified and digitally recorded with optical microscope (OM), scanning electronic microscope (SEM), and atomic force microscope (AFM) respectively. The recorded grid patterns are processed with Fourier transform technique. The spatial resolution degradation due to Fourier transform is discussed. It is found that the upper limit of spatial resolution for displacement measurement is half the grid pitch for grid method. Nevertheless the spatial resolution which can be attained is in micron to sub-micron region even with this degradation. A comparison on the effect of three microscopy techniques is also carried out.

Keywords: Grid method, Fourier transform, electronic packaging, spatial resolution, optical microscope, scanning electronic microscope, atomic force microscope, thermal stress.

1. INTRODUCTION

Thermal strain is a major cause of electronic packaging failure. The strain results from the mismatch of the thermal expansions of the various constituent materials. As the components and structures are made smaller, the thermal gradients increase and the strain concentrations become more serious. In order to solve the thermal strain problem of packaging, three analysis methods, analytical, numerical and experimental methods, are widely used to reveal the packaging structure response to thermal loads. Analytical and numerical techniques, especially the latter can handle a wider class of problems. However there are limitations for these two techniques, as the complexities of the problems increase. For example, boundary condition cannot always be reliable prescribed. This problem is usually resolved via experimental technique. Furthermore, the validation of analytical model and numerical result is always required to provide confidence in their findings. This is only achieved by experimental techniques.

Among many advanced photomechanical tools for thermal deformation measurement, moiré interferometry has been used extensively for in-situ measurement[1,2,3]. It provides whole field contour maps of displacement with subwavelength sensitivity. However, with the present moiré interferometry, the spatial resolution of measurement leaves much to be desired. Although microscopy has been used in micro moiré interferometer[1,4,5]. In this paper, a microscopic grid method using optical microscope (OM), scanning electronic microscope (SEM), and atomic force microscope (AFM), and its application to thermal deformation of electronic packaging are presented. Compared to microscopic moiré interferometry, spatial resolution for displacement measurement is in micron or submicron order. This high spatial resolution value can not be attained with current moiré interferometry technique. Furthermore, the sensitivity of proposed microscopic grid technique for displacement measurement is also comparable with moiré interferometry.

2. GRID METHOD WITH FOURIER TRANSFORM

Grid methods are widely used in mechanics and engineering for displacement/strain measurement. If the grid line carries the same spatial information as a moiré fringe, then grid technique provides more spatial information than moiré techniques, as grid line (spot) density is higher than moiré fringe density. This is the main advantage of grid technique.

2.1 Fourier transform approach

Part of the SPIE Conference on Advanced Photonic Sensors and Applications
Singapore • November–December 1999
SPIE Vol. 3897 • 0277-786X/99/$10.00

When a grid is applied on the surface of a test specimen, it will be deformed as the specimen is loaded. For analysis of deformation, it is necessary to determine the differences between the images of grid pattern before and after loading[6]. The key step of grid method is to extract the displacement information from the grid image. The often-used technique is Fourier transform method which is widely employed in fringe analysis. The application of this method on grid/moiré strain analysis is explained in detail in literature[7,8]. It is thus briefly summarized here by considering 1-D case as an example (see Fig. 1).

The intensity of 1-D grid (actually a grating) line can be expressed as:

$$g(x,y) = a(x,y) + b(x,y) \times \cos\left[2\pi f_0 x + \phi(x,y)\right] \tag{1}$$

where the phase $\phi(x, y)$ is proportional to the required displacement $u(x, y)$:

$$\phi(x,y) = 2\pi f_0 u(x,y) \tag{2}$$

f_0 is the grid frequency, and $a(x, y)$, $b(x, y)$ are the mean and modulated intensity of the grid lines respectively. It should be noted that the higher order harmonics are neglected in Eq. (1). The Fourier transform function along the sensitivity direction (x) of grid can be expressed as:

$$G(f,y) = A(f,y) + C(f - f_0, y) + C^*(f + f_0, y) \tag{3}$$

where * denotes the complex conjugate. Because spatial variations of $a(x, y)$, $b(x, y)$ and $\phi(x, y)$ are slow compared with f_0, the function $G(f, y)$ will be trimodal function with peaks at $-f_0$, f_0 and the origin. The function $C(f-f_0, y)$ can be isolated using a filter centered at f_0. The carrier frequency can be removed by shifting $C(f-f_0, y)$ by f_0 to the origin to give $C(f, y)$. The inverse transform of $C(f, y)$ with respect to f yields

$$C(x,y) = 0.5b(x,y) \times \exp\left[i\phi(x,y)\right] \tag{4}$$

From which the phase $\phi(x, y)$ can be deduced as:

$$\phi(x,y) = \tan^{-1}\left\{\mathrm{Re}\left[C(x,y)\right] / \mathrm{Im}\left[C(x,y)\right]\right. \tag{5}$$

The phase thus obtained is wrapped into the range $(-\pi, \pi)$ and has to be corrected using an appropriate phase-unwrapping algorithm. The strain measurement procedure using grid technique with Fourier transform involves the following steps:

- Record the grid images before loading,
- Calculate the phase field from recorded grid image with above mentioned Fourier transform method,
- Record the grid images after loading, and then calculating the phase field,
- Calculate the displacement field by subtracting the two phase fields,
- Calculate the strain field from the displacement field.

2.2 Spatial resolution degradation due to Fourier transform

As mentioned in introduction of this paper, a deformation measurement with high spatial resolution is increasingly important in electronic packaging testing. Spatial resolution is defined as the distance between the detection points[9], more precisely, the minimum spatial unit needed to perform a measurement. It is because of this low spatial resolution that the conventional strain gauges can not be used in a micromechanical measurement, despite its high strain sensitivity and precision is high. This high precision strain value is actually obtained over a sensing area of about 1~10 mm^2. For different measurands, the spatial resolutions of measurement are different. For example, the spatial resolution for displacement will be higher than that for strain measurement if the strain value is obtained from the differentiation of displacement value. Another important point is that the spatial resolution for 2-D displacement field measurement is not the resolution of test (optical) system. It is also associated with detecting system including post image processing system and method, such as filtering operation, Fourier transform etc.

A grid can record the full spatial deformation information, but the deformation information appears as modulation of a carrier pattern. Moiré method automatically eliminates the carrier pattern, and the moiré fringe pattern directly represents the deformation information. In grid method, post image processing method, often Fourier transform method is used to eliminate the carrier information and to extract the deformation information. The problem which always exists is how to restore this deformation information precisely and with minimum loss of spatial resolution. As a spatial filter technique, Fourier transform method will certainly degrade the spatial resolution of measurement system. Fig. 2 illustrates this degradation. Among all the Fourier spectra of grid signal shown in this figure, only +1 spectrum is used in inverse Fourier transform to get the measured phase information. In Fig. 2, L, the bandwidth of the filter limits the spatial resolution of Fourier transform. All the signal with frequency beyond this filter bandwidth is filtered out, and can not be

restored. Therefore, $1/L$ is defined as the spatial resolution of Fourier transform. As L must be less than $2f_0$, otherwise 0 and +2 spectra are included in the filter. This means that the upper-limit of the spatial resolution is $1/(2f_0) = 0.5$ pitch of grid. For a digital grid image, grid pitch must be large than 2 pixels. Thus, the upper-limit of spatial resolution is one pixel. This means that sub-pixel interpolation is not reliable if Fourier transform grid method is used.

Let's further discuss the spatial resolution problem through an example. Fig. 3(a, b) shows an AFM grid image and its spectrum with size of 488×395 (pixels). The grid frequency $f_{0X} = 59$ pixels in the spectrum figure, i.e., the distance between 0 and +1(or -1) order spectrum is 59 pixels. In order to show the influence of the filter bandwidth on spatial resolution, two mismatch moiré patterns were obtained with different filter bandwidth, shown in Figs. 3(c) and 3(e). The method and procedure for obtaining mismatch moiré is clearly described in reference 8. In order to obtain the mismatch moiré in Fig. 3(c), ±1-order spectra in X direction are shifted 53 pixels towards origin, and the shifted spectrum is shown in Fig. 3(d). The corresponding filter bandwidth is 12 pixels. The spatial resolution is therefore about 59/12 = 4.91 pitches of grid, or 4.1 micron. The filter bandwidth corresponding to Fig. 3(e) is 60 pixels. The shifted spectra are shown in Fig. 3(f). The spatial resolution is about 0.98 pitches of grid, or 0.82 micron. Comparing Fig. 3(c) and 3(e), more details of the fringe pattern can be observed in Fig. 3(e). Fig. 4(a) and 4(b) display two line profiles along Y = 150 pixel from the two sub-figures 3(c, e). The curve in Fig. 4(a) is smoother, implying lower spatial resolution.

3. EXPERIMENT AND RESULT ANALYSIS

The specimen is a BGA (Ball Grid Array) Flip Chip. A high frequency diffraction grid (1200 lines/mm) is replicated on the specimen at high temperature. After cooling, the grid records the thermal-induced deformation, which is imaged via a microscope. Through an off-line image processing, the recorded deformation information can be restored. It should be noted that all the grid images presented in this paper were taken with different types of microscope but with the same specimen.

3.1 Specimen preparing

Grid replication and sample preparation is a critical part for the grid measurement technique. The detailed procedure is described in references 4 and 5. The main techniques are simply reviewed as follows:

- sawing of the chip specimens into halves with a diamond wire cutter,
- grinding of the specimens with a grinding machine,
- pre-baking of the specimens for 4 hours at 125^{0}C,
- Heating the specimen to the required temperature,
- Replicating the grid onto the specimen at 100^{0}C.

3.2 AFM grid image processing

Fig. 5(a) displays a grid images using a Dimension 3000 (Digital Instrument) AFM. The image was taken from the same specimen but different part then shown in Fig. 3(a). The size of both image is 48×31 μm^2 (488×395 pixel^2), but Fig. 5(a) is magnified for visualization. It is apparent that the grid in Fig. 5(a) has larger average deformation than that in Fig. 3(a). Fig. 5(b) shows the spectrum of the grid in Fig. 5(a). Comparing the spectrum in Figs. 3(b) and 5(b), the relative deformation of two grating can bee obtained. f_{0X}, the grid frequency in X direction varies from 59 pixels (in Fig. 3(b)) to 58 pixels (in Fig. 5(b)), This means an average relative strain of (58-59)/59 = -1.7%. Similarly in Y direction, the grid frequency, f_{0Y} varies from 37 pixels to 39 pixels, and the average relative strain is 5.4%. The comparison of two spectrum patterns provides an average relative deformation information. While the comparison of two a mismatch moiré pattern will let us know the full field deformation information. Fig. 6 displays a mismatch moiré pattern in X direction. The filter window and the shift distance are the same with that of Fig. 3(c). The difference of two fringe patterns clearly demonstrates the relative deformation.

3.3 SEM grid image processing

Fig. 7 shows two grid images obtained using a SEM (Cambridge S360), at 6000 × magnification. The size of the images are 12.5×18.3 μm^2 (207×311 pixel^2). The image in Fig. 7(b) is taken after the sample stage is translated 29 μm in X direction and 11 μm in Y direction. Although the distance of this translation is short, but the difference between the

deformation of the two grid image is still visible. In order to further visualize the heterogeneous deformation in the different place of the packaging, the four corresponding mismatch moiré patterns are displayed in Fig. 8. The filter bandwidth is 17 pixels and the shift distance is 8 pixels.

3.4 Optical microscopic grid image processing

The pitch of grid frequently used in moiré interfereometry is 0.83 μm (1200 lines/mm). In order to resolve this grid, the spatial resolution of the used microscope should be at least 0.42 μm. This value is near the upper diffraction-limited resolution of an optical microscope. Thus, it is not easy to get such high frequency grid image with an optical microscope. However, there is an important advantage of using optical microscope, in that it is easy to connect with a mechanical (heating) loading system.

Fig. 9(a, b) display an optical microscopic grid image and its spectrum. The image size is 243×203 pixel^2. The grid image contrast is not high, as high magnification is required. Due to the low (numerical) spatial frequency that is about 0.5 lines/pixel, the ±1-order spectra are far from the 0-order spectrum. This high carrier frequency will make it easier to obtain mismatch moiré. Fig. 9(c, f) shows the mismatch moiré in two directions. For moiré in X direction (Fig. 9(c)), the filter bandwidth is 29 pixels and the shift distance is 109 pixels. For moiré in Y direction (Fig. 9(f)), the filter bandwidth is 27 pixels and the shift distance is 92 pixels. The contrast of the mismatch moiré fringes is much better than that of the original grid pattern.

4. DISCUSSION AND CONCLUSION

A series of electronic packaging chip tests with grid method are performed under optical microscope, AFM and SEM. The grid images representing the thermal-induced deformation are processed through Fourier transform method. Some conclusive remarks can be drawn from these tests.

- This microscopic grid method with Fourier transform technique provides for localized point-measurement but full-field data. The spatial resolution for displacement measurement is in micron or submicron order. Therefore, this method is suitable for strain concentration analysis.
- The Fourier transform method is used as a digital image processing technique. What is of concern is the numerical frequency of grid image which is often in unit of lines/pixel, not the absolute grid frequency.
- With Fourier spectrum method, deformation field can be visualized by mismatch moiré.

Comparing the three microscopes, it is difficult to use optical microscope for high frequency grid image, and the image quality is not good for processing, but it is relative easy to connect a mechanical or heating test. AFM and SEM provide a very good quality image, and are suitable for high frequency grid measurement. SEM with mechanical loading system or heating system may be a good compromise for the further study of mechanical behavior of electronic packaging with grid method.

ACKNOLOGEMENTS

The support of Ministry of Education and Nanyang Technological University through research grant MLC1/97 is acknowledged.

REFERENCES

1. D. Post, B. Han, and P. Lfju, *High sensitivity Moiré: experimental analysis for mechanical and materials*, Springer-Verlag, New York, 1994.
2. A. Voloshin, "Analysis of the thermal loading on electronics packages by enhanced moiré interferometry," Chap. 8 in *Thermal stress and strain in microelectronics packaging*, edited by John H. Lau. Van Nostrand Reinhold, New York, pp. 272-304, 1993.
3. Y. Guo, S. Liu, "Development in optical methods for reliability analysis in electronic packaging applications," *J. of Electronic packaging*. **10**. pp. 186-193, 1998.
4. K.M. Chua, "Moiré interferometry in electronic packaging," *Final year project report*, Nanyang Technological University, Singapore, 1998.

5. K.E. Oh, "Development and application of Moiré interferometric strain sensor," *Internal report of School of Mechanical and Production Engineering*, Nanyang Technological University, Singapore, 1999.
6. P.J. Sevenhuijseen, "The photonical, Pure grid method" *Optics and Lasers in Engineering*, **18**(3), pp.173-194, 1993.
7. Y. Morimoto, Y. Seguchi, "Application of moiré analysis of strain using Fourier transform," *Optical Engineering*, **27**(8), pp. 650-656, 1988.
8. Y. Morimoto, Y. Seguchi, and T. Higashi, "Two-dimensional moiré method and grid method using Fourier transform," *Experimental Mechanics*, **29**(4), pp399-404, 1989.
9. T.M. Kreis, "Computer aided evaluation of fringe patterns," *Optics and Lasers in Engineering*, **19**(4-5), pp. 221-240,1993.

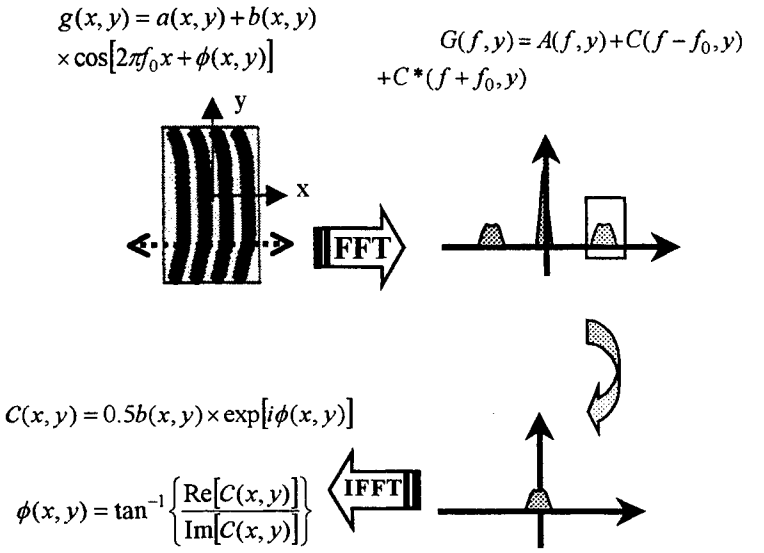


Fig. 1 Illustration of Fourier transform method

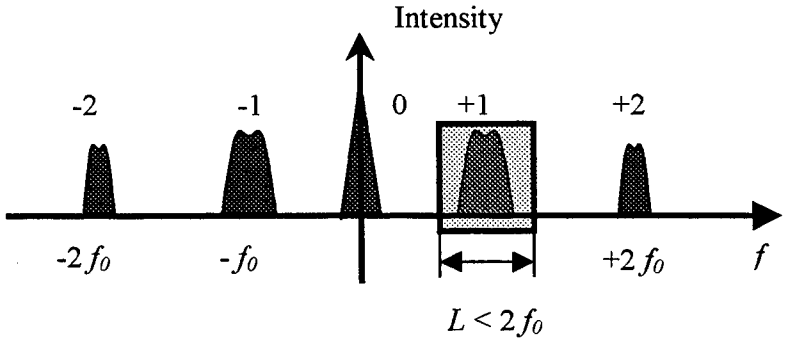


Fig. 2 Illustration the relationship between filter bandwidth and spatial resolution. L is the filter bandwidth, $1/L$ is deefined as the spatial resolution for phase or displacement measurement

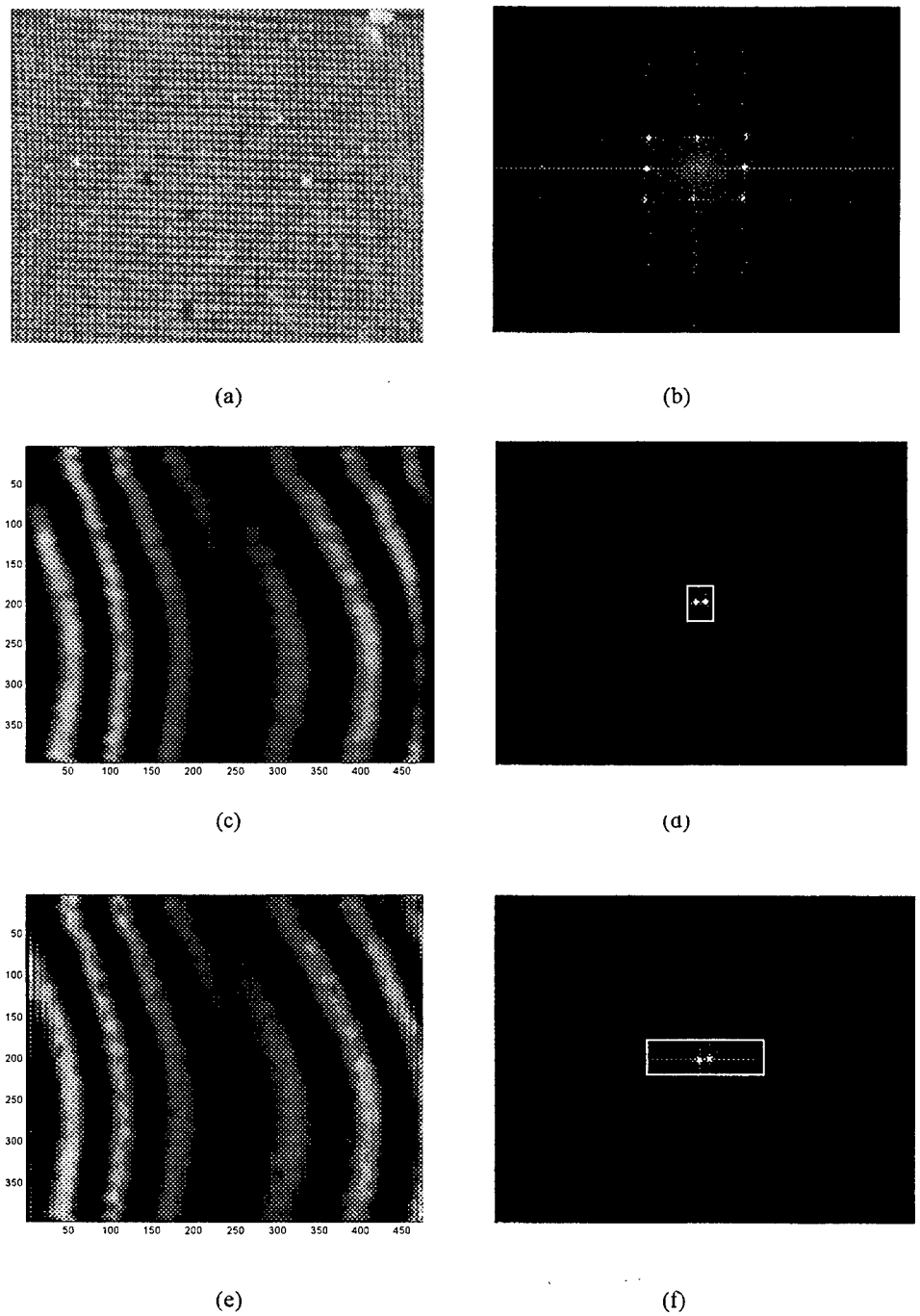
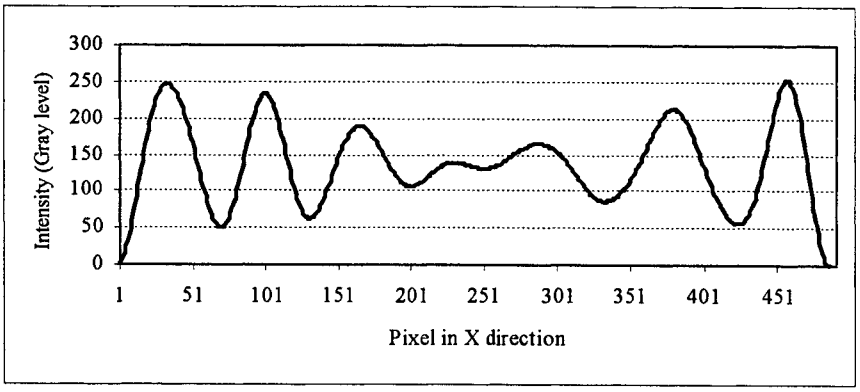
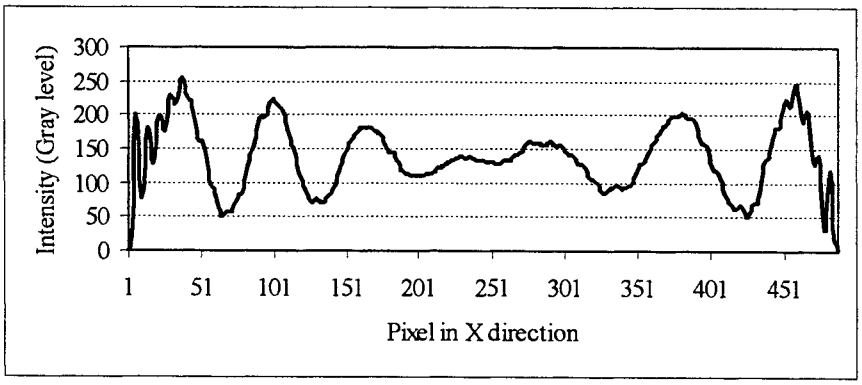


Fig. 3 Example of showing the influence of filter bandwidth on spatial resolution; (a) AFM grid image with size of 488×395 pixel^2 (48×31μm^2); (b) its spectrum, +1 order spectrum frequency f_{0x} = 59 pixels; (c) Mismatch moiré by shifting ±1 spectra 53 pixel toward origin; (d) spectra of moiré (c) with filter bandwidth of 12 pixels; (e) The same with (c) but filter bandwidth of 60 pixel; (f) spectra of (e).

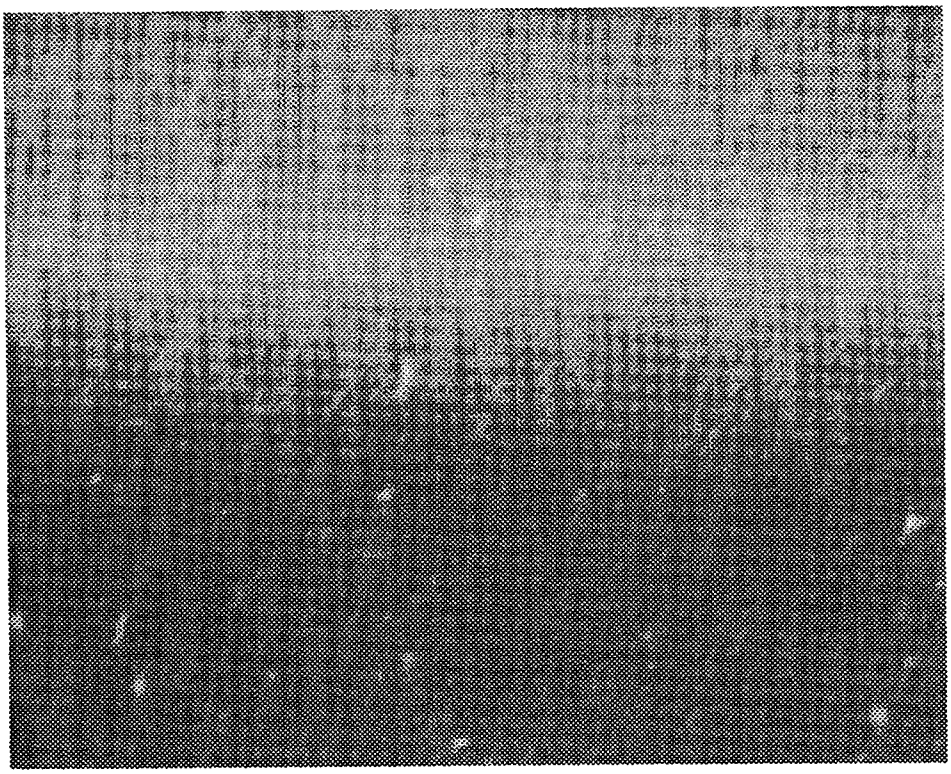


(a)

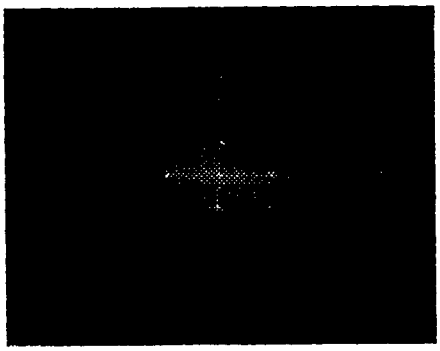


(b)

Fig. 4 Line profiles of mismatch moiré for further illustrating the influence of filter bandwidth on spatial resolution. (a) corresponding to line at Y= 150 pixel in Fig. 3(c); (b) corresponding to Fig. 3(e).



(a)



(b)

Fig. 5 AFM grid image and its spectrum taken with same magnification and from the same specimen with the grid image in Fig. 3(a). (a) grid image (b) spectrum

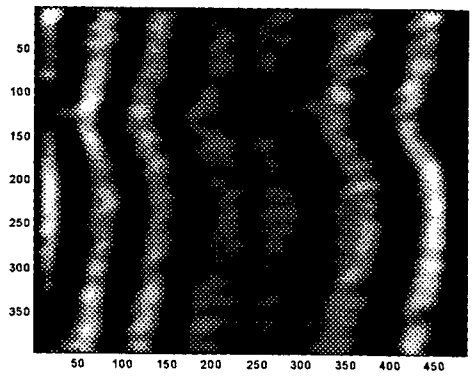
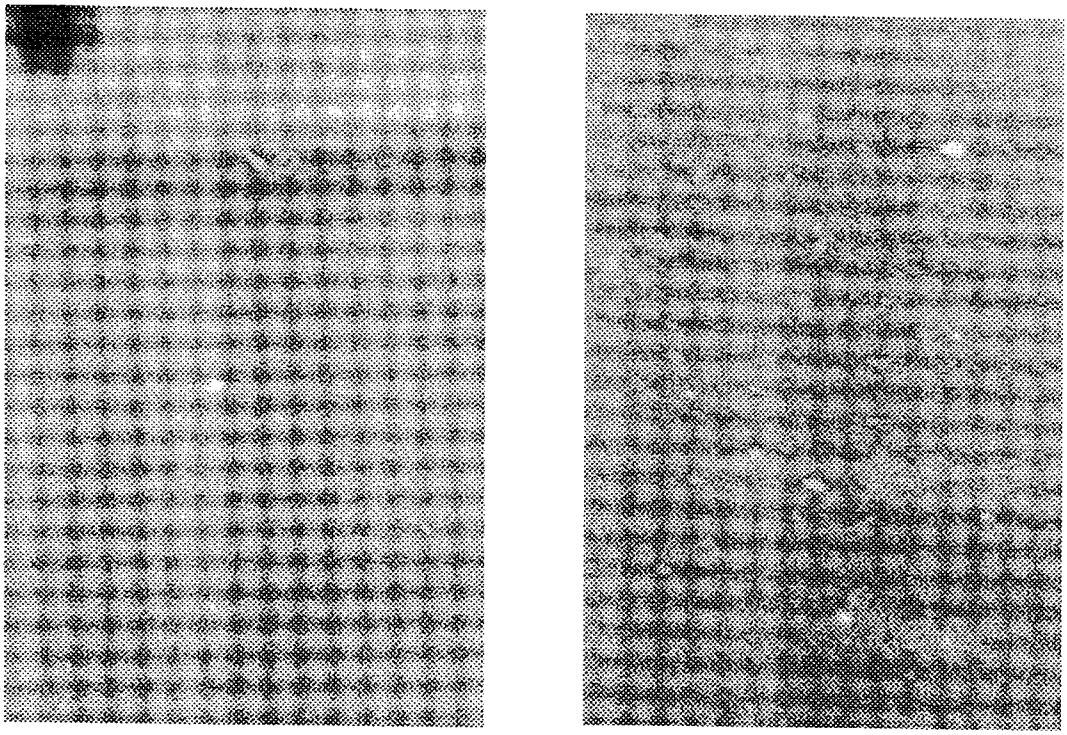
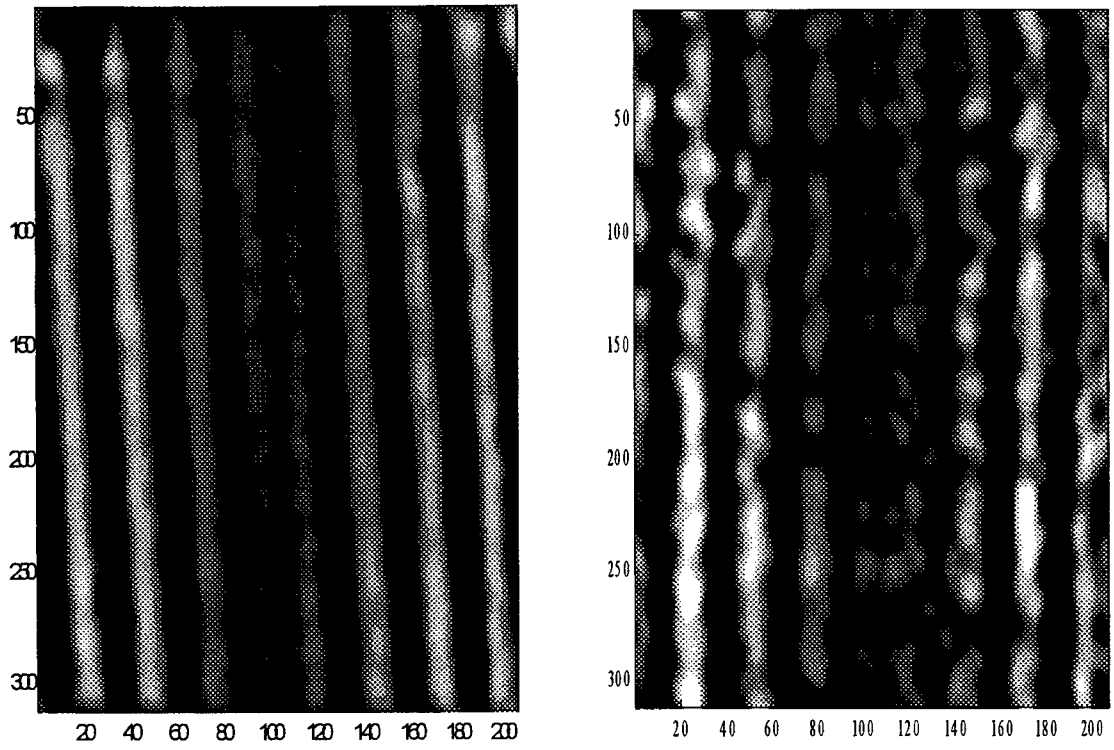


Fig. 6 Mismatch moiré generated from the grid image in Fig. 5(a). The filter window and the shift distance of ±1 order spectra are the same with that of in Fig. 3(a).



(a) (b)

Fig. 7 SEM grid images with size of 12.5×18.3 μm^2 (207×311 $pixel^2$), and a magnification factor of 6000. Image (b) was taken after specimen was translated with a distance of 3 μm in X direction and 15 μm in Y direction.



(a) (b)

Fig. 8 Mismatch moiré patterns corresponding to the two subfigures in Fig. 7. The filter bandwidth is 17 pixels, and ±1 order spectra are shifted 8 pixels toward the origin.

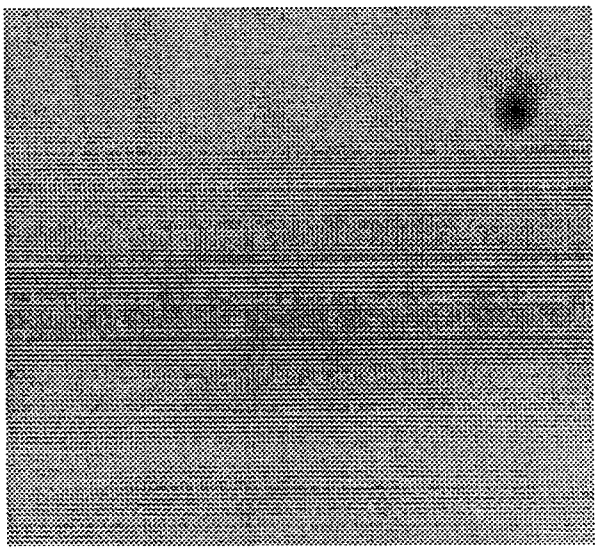
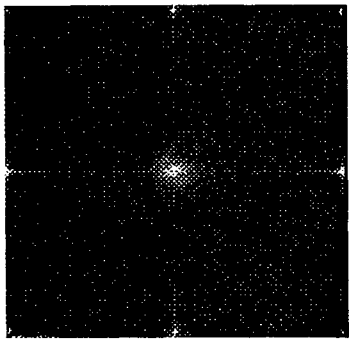
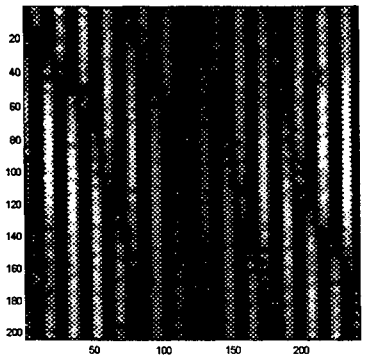
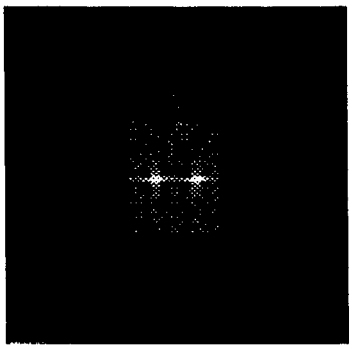


Fig. 8(a)



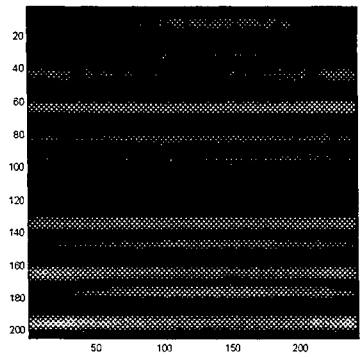
(b)

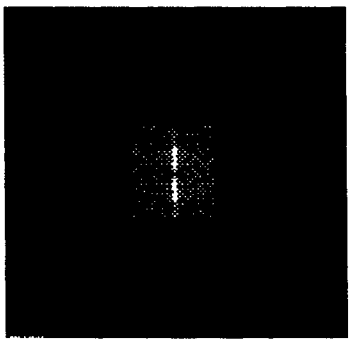




(c)

(d)





(e)

(f)

Fig. 9 Optical microscope grid image and its processing result. (a) grid image with size 243×203 pixel^2. One pixel represents half pitch; (b) spectra of grid in (a), (c) Mismatch moiré in X direction, by shifting ±1 order spectra (in X direction) 109 pixel toward the origin, the filter bandwidth is 29 pixels. (d) the spectra corresponding to (c); (e) Mismatch moiré in Y direction, by shifting ±1 order spectra (in Y direction) 92 pixel toward the origin, the filter bandwidth is 27 pixels. (f) the spectra corresponding to (e).

SESSION 6

Vibration Sensors

Non-destructive diagnostics of layered structures: advanced signal analysis algorithms applied to vibrometric data.

Prof. Enrico P. Tomasini[a], Prof. Francesco Piazza[b], Dr. Enrico Esposito[a],
Dr. Monica Possanzini[b].

[a]Dipartimento di Meccanica, Universita' di Ancona.
Via Brecce Bianche
60131 Ancona - ITALY
Phone: +39-71-2204487
Fax: +39-71-2204813
e-mail: tomasini@mehp1.unian.it
[b]Dipartimento di Elettronica e Automatica, Universita' di Ancona.
Via Brecce Bianche
60131 Ancona - ITALY
Phone: +39-71-2204453
e-mail: upf@eealab.unian.it

ABSTRACT

A Scanning Laser Doppler Vibrometer (SLDV) employs Doppler frequency shift of a probe laser beam and interferometric mixing with a reference beam to measure target velocity. This technique offers high sensitivity, high spatial resolution and very low if any intrusivity.

At present all commercial instruments include FFT based analysers, while some of them also allow RMS and lock in capabilities. The possibility of saving and eventually retrieving and elaborating signal time histories is sometimes included in the software or offered as an option.

Aim of this work is to present a new approach to the problem, where the main focus is on saving time histories that can be post processed in any desired way.
Basically the system consists of a commercial laser head controlled by a LabVIEW Virtual Instrument (VI) and a Matlab Graphical User Interface (GUI) capable of signal processing and results presentation.

The VI allows the user to set all measurement parameters (trigger, sample length, sampling frequency, input/output filtering, spatial resolution), while the GUI permits different frequency analysis methods (eigenvalues based algorithms, amplitude calculation by Prony method, standard and rank reduced parametric models, FFT and RMS), and data presentation through 2D and 3D graphs. Moreover data maps may be modified applying different grades of smoothing, average value subtraction, image or data zooming.

More noticeable advantages are high-resolution frequency analysis, low noise spectra, automatic resonance frequencies identification, small data files and partial time transient analysis.

Several tests were performed on different artificially defected layered structures (frescoes, Byzantine icons and composite materials samples) and a full set of actual responses will be presented.

Keywords: laser scanning vibrometer, non invasive diagnostics, layered structures, signal analysis, spectral analysis methods.

Part of the SPIE Conference on Advanced Photonic Sensors and Applications
Singapore • November–December 1999
SPIE Vol. 3897 • 0277-786X/99/$10.00

1. SLDV INSTRUMENTATION

1.1 Introduction

Vibration measurements are nowadays becoming of increasing interest for a large number of scientific and industrial applications. In this contest laser vibrometry is opening new possibilities with respect to traditional contact (e.g. accelerometers, strain gages) and non-contact (e.g. triangulation or reflection sensor, proximity) techniques. The standard Laser Doppler Vibrometer (LDV) is a non-contact velocity transducer working on the principle of measuring the Doppler frequency shift of a laser beam scattered from a moving target by means of an interferometer. The on-board electronics converts the Doppler signal to an analog voltage proportional to the instantaneous velocity of the target.
The combination of an interferometer with two moving mirrors driven by galvanometric actuators makes it possible to direct the laser beam to the desired measurement points. Such an instrument, named the Scanning Laser Doppler Vibrometer[1], can quickly perform a series of velocity measurements on a grid of points over the structure under test.
These techniques are effectively used in structural dynamic testing, biological and clinical diagnostics, fluid-structure interaction, on-line monitoring of industrial plants, acoustics, and fault detection, to quote only a few. Furthermore, the coupling of laser vibrometers and scanning systems seems to open up new possibilities, e.g. in the field of measurements in tracking mode on moving objects[2]. A novel application regards the employment of the SLDV for non-invasive and non-contact diagnostics of works of art[3,4].

1.2 Non-Destructive Laser Doppler Vibrometers: Measurement Principle and Methods.

When a coherent radiation of frequency f, emitted by a source, interacts with a reflecting or diffusing surface moving target at velocity u, the radiation observed by the source (backscatter configuration) is affected by a Doppler shift (f_D) proportional to the surface velocity. Such frequency shift is given by[5]:

$$f_D = \frac{2u}{\lambda}\cos\theta \tag{1}$$

being $u\cdot\cos\theta$ the velocity component along the laser beam.

If a He-Ne ($\lambda = 632.8$ nm) laser is used for velocity measurements the frequency $f = 4.7\cdot10^{14}$ Hz is shifted of about $3.16\cdot10^{3}$ Hz per 1 mm/s. In order to measure such relatively small frequency shifts and to have a reference in the phase assessment the interferometric approach can be used. The most common configurations employ either the Michelson or the Mach-Zender interferometers, the former employing usually electronic mixing to discriminate velocity direction, the latter acousto-optic devices (Bragg cell)[6]. A schematic of the instrument is shown in Fig.1, with the two options in gray boxes.

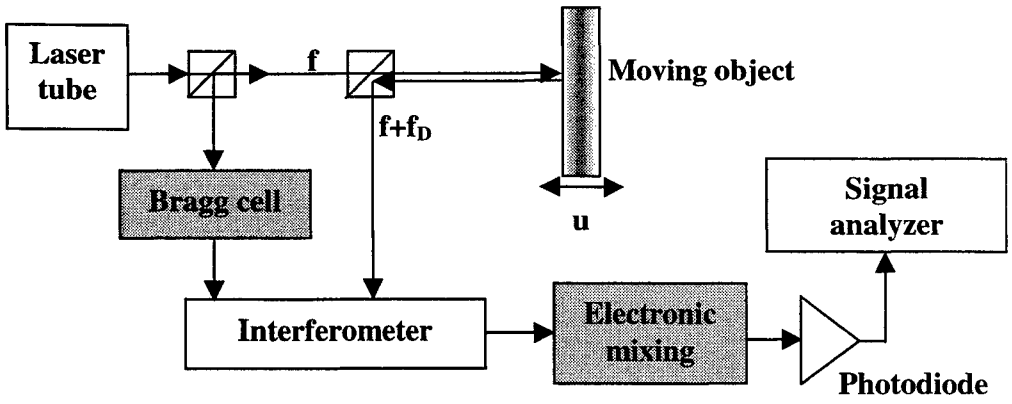


Fig. 1 Schematic of an LDV.

By scanning the laser beam across an object one can observe the frequency content of the velocity signal that will show significant differences in presence of an hidden defect, e.g. a void, with respect to an integral sample or to an integral area of the same defected sample.
In the first case one has a structure that closely resembles a cavity backed plate that when excited will resonate at some specific frequencies and generally will have greater average amplitude.

2. PROPOSED SYSTEM

2.1 Physical Implementation

Standard commercial systems are usually organized as shown in Fig.2. The system interferometer is contained in the laser head and data (demodulated or not) are sent to a control unit that also drives the mirror actuators. Vibrational data in the form of frequency spectra are then processed and presented as 2D or 3D images; measured spectra are also stored in usually large files (tens of Mb). The possibility of saving time histories is often an optional for these systems.

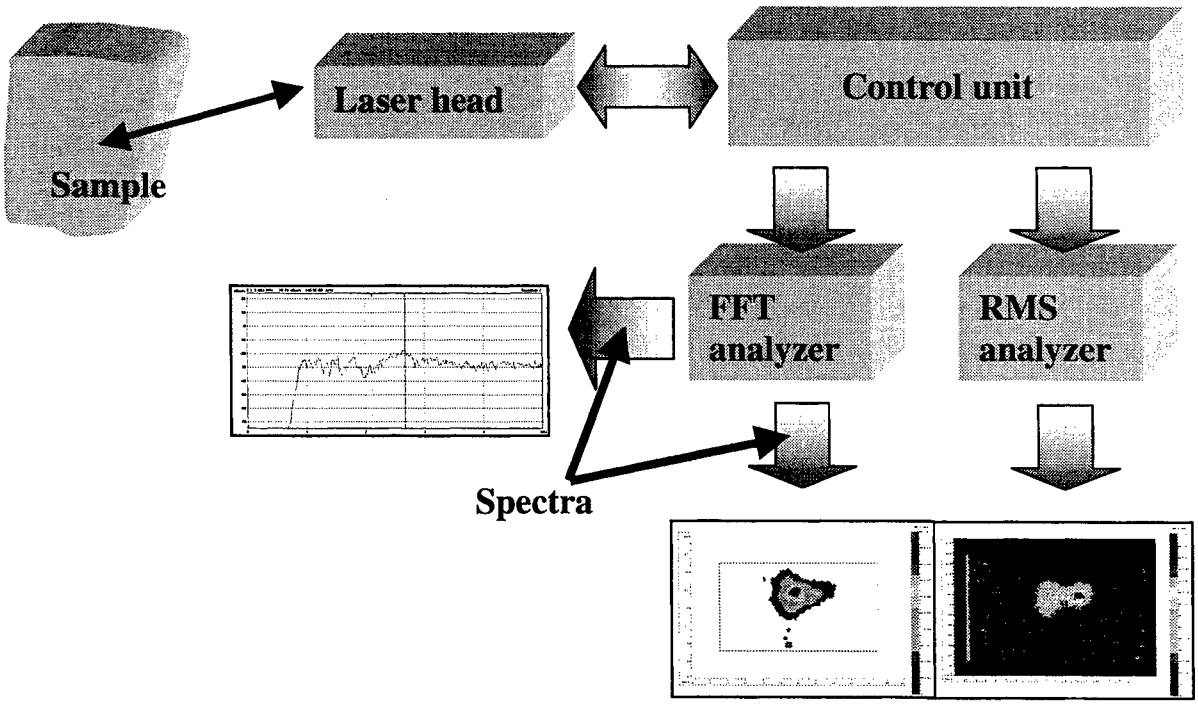


Fig. 2 Typical layout of a commercial SLDV.

The proposed system (see Fig.3) uses a much more flexible approach: a LabVIEW virtual instrument substitutes the control unit and all computations are executed through a MATLAB script and moreover time histories are stored and elaborated to extract data maps. The simultaneous use of LabVIEW and MATLAB allows for a possible continue upgrading of the system that can implement any new algorithm by just adding a new MATLAB module. Any commercial laser head may be used by just adjusting the parameters of the LabVIEW panel and other minor modifications.

2.2 Spectral Estimators

Basically the implemented software allows performing two different operations on acquired data: the user may directly calculate the PSD (Power Spectral Density) of the time series or use a spectral estimator of the same input signal.
There exist different methods of spectral estimation, each one of them producing result spectra with different characteristics. In most practical situations signal frequency content is not known a priori and this makes difficult the choice of the best method among those normally available.

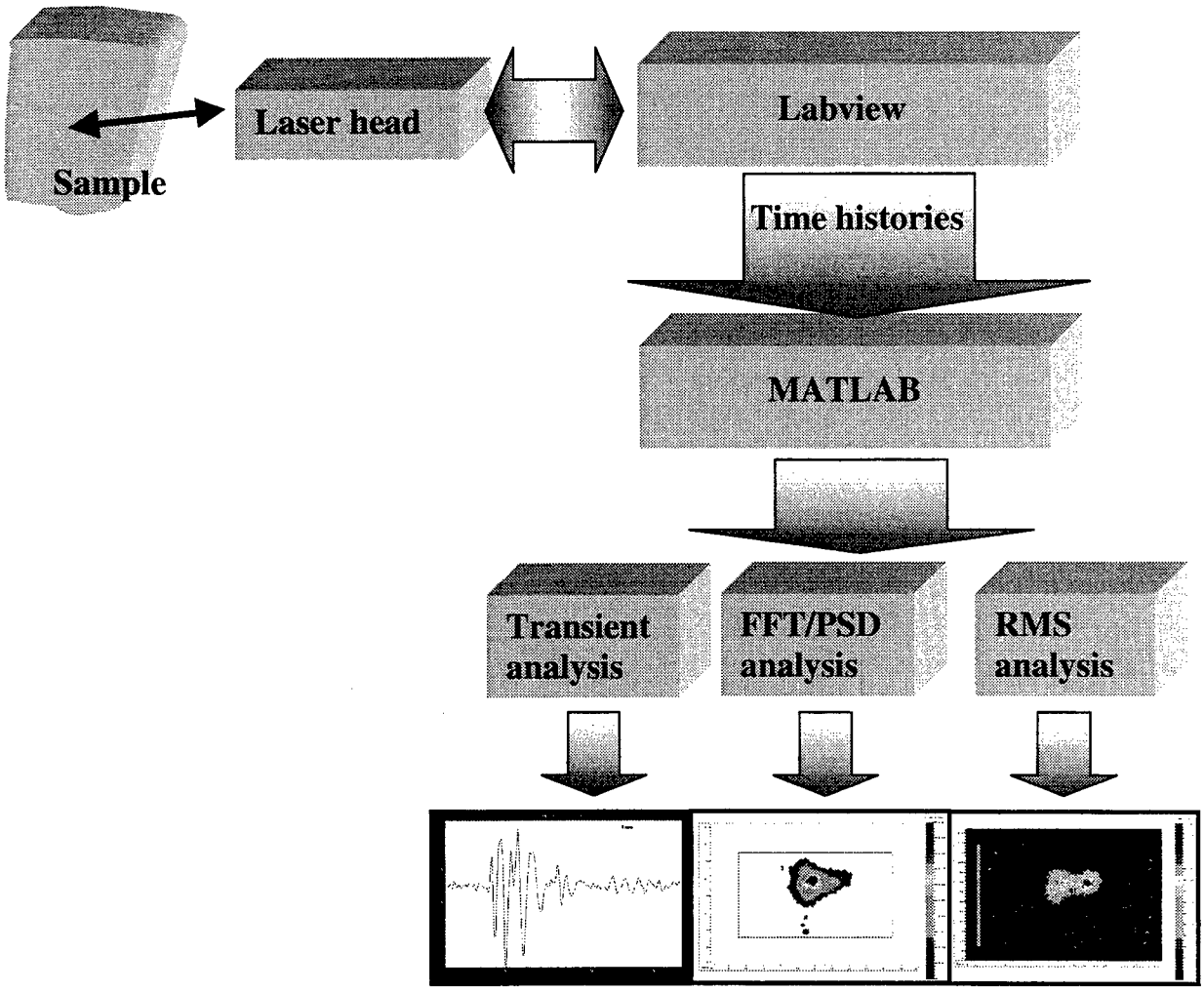


Fig. 3 Layout of proposed SLDV system.

In our case attention was focused on improving spectral estimation of very short data records sampled from very noisy signals. These limits were chosen considering that worst cases SNRs are in the order of 10÷15 dB and data are collected on possibly vast grids, leading to large data files.

Our applications require for spectral estimator with high frequency resolution to distinguish close coupled resonance modes, high robustness, and high reliability, even working with short data streams. A selection of possible options led to implement one spectral calculator, and four spectral estimators.

2.2.1 Spectral FFT calculator (Periodogram)

The Periodogram is part of the classical methods based on FFT to perform spectral analysis. It can be considered a "direct" method, because it calculates spectra from time histories neglecting autocorrelation estimation.

To calculate the signal PSD by a standard FFT algorithm, time records are acquired for a finite time slot depending on system capabilities and operator choices then sampled and processed[7]. The basic standard formula for spectra calculation is reported here below,

$$Pxx = \frac{T}{N} \left| \sum_{n=0}^{N-1} x[n] \exp(-j2\pi fnT) \right|^2 \tag{2}$$

where T is the sampling interval, x(n) the sampled data and N the number of samples.

An adverse effect of signal truncation is represented by the so-called *spectral leakage*; this leads to the appearance of unwanted secondary lobes and enlargement of principal lobes in the estimated spectrum. Amplitude errors and frequency peaks masking thus follow. The usual remedy is the use of *time windows* that avoid abrupt signal termination by introducing different sorts of data smoothing and edge tapering; the most widely used is the Hann (acknowledged as Hanning) window due to ease of implementation and narrow principal lobes.
Classical techniques for PSD calculation are particularly robust and may be employed with a wide range of stationary signals, while other methods are able to deal only with a limited range of signals. Moreover frequency peaks amplitudes in PSD are linearly proportional to the power of sinusoids present in the signal; with other methods this is not generally true and amplitudes are estimated less precisely.
Leakage and low spectral resolution, limited to 1/NT represent disadvantages of such method, while the only way of increasing statistical stability is to reduce frequency resolution even more. Estimation methods allow to break this limit and allow high resolution and good statistical stability at the same time.

2.2.2 Parametric Models

PSDs are normally calculated either as the Fourier transform of the signal autocorrelation (ACS) or by direct methods; *parametric models* are based on the assumption that time series may be represented by a model and so its PSD is not a function of the ACS but of the model parameters.

We have seen that using standard PSD calculation methods we introduce data distortion and spectra are obtained on the assumption of zero value samples outside the applied window. This assumption is clearly unrealistic and to remedy this fact parametric methods use only the acquired data to extract an estimation of the PSD; doing so windows are no more necessary and the obtained spectra are less affected by distortion. The difficulty of the method resides in choosing the right model and in its ability to reconstruct data from a few samples.

The model class employed in our work assumes white noise input and rational transfer functions; in this class we find three possible models, called AR, ARMA and MA[7].

The basic discrete time series approximating many processes both deterministic and statistical may be represented as the equation of a complex coefficients linear filter,

$$x[n] = -\sum_{k=1}^{p} a[k]x[n-k] + \sum_{k=0}^{q} b[k]u[n-k] = \sum_{k=0}^{\infty} h[k]u[n-k] \tag{3}$$

where x[n] is the output of the casual ARMA filter, h[k] the filter impulse response and u[k] the input sequence. Generally b[0]=1 and u[n] is a white noise sequence with zero average value and variance ρ_w.

The filter transfer function is represented as H(z) = B(z)/A(z), where

$$A(z) = 1 + \sum_{k=1}^{p} a(k) z^{-k} \qquad B(z) = 1 + \sum_{k=1}^{q} b(k) z^{-k} \tag{4}$$

and finally the PSD is calculated as

$$P_{ARMA}(f) = T\rho_w |B(f)/A(f)|^2, \text{ using } z = e^{j2\pi fT}. \tag{5}$$

The AR case is derived by using b[k]= 0 $\forall k \neq 0$,

$$P_{AR}(f) = T\rho_w |1/A(f)|^2, \tag{6}$$

or

$$P_{AR} = \frac{T\rho_W}{\vec{e}_p^{\,H}(f)\vec{a}\vec{a}^{\,H}\vec{e}_p(f)}, \tag{7}$$

where $\mathbf{e}_p(f)$ is the vector containing the $e^{j\omega T}$ complex sinusoids and **a** is the vector of the AR model parameters.

Usually AR models are employed instead of ARMA or MA due to their lower computational cost and sharper peaks of found resonances. The number of estimated parameters is defined also as the *model order* and this determines the compromise between resolution and variance of estimated spectra.

To implement the AR models there exist different methods[7], the most efficient being in our case the so-called *modified covariance* method (FBLP). This method is part of the larger family of the *linear prediction methods*, and is based on the minimization of forward and backward squared errors; it has been proved the best solution if the signal contains sinusoidal components. This type of signal is commonly encountered in our measurements, where resonance peaks are superimposed to a relatively important noise floor.
The user may also select other algorithms, namely the Yule-Walker[7] and Burg[7] but in most of our applications they usually show lower resolution.

As already mentioned PSD obtained by estimators do not show lobes as those calculated by FFT, but in the former case there exist some limitations: choosing to low an order gives rise to poorly defined spectra while exaggerating the modes number may introduce "ghost" peaks in the frequency spectra. Moreover if the signal contains white noise and sinusoids calculated spectra depend on signal SNR and particularly resolution gets lower as the SNR diminishes.

2.2.3 Eigenvalues Models

These techniques employ eigenvalues analysis of autocorrelation or data matrices[7, 8, 9] as those already considered for AR methods. These techniques usually show a better resolution and resonant frequencies estimation than the latter ones, especially with low SNRs.

The basic idea behind these techniques is to divide the aforementioned matrices in two vector subspaces, one containing the eigenvectors of signals subspace the other the eigenvectors of the noise subspace. It must be pointed out that the results are not real spectra because signal power is not maintained and the autocorrelation sequence is not obtainable transforming the spectrum by an FFT.

This estimator class comprises:

1. algorithms for classification of multiple signals (MUSIC)[7];
2. the FBLP algorithm modified by Kumaresan and Tufts[8, 9]; this algorithm has also been modified by Cadzow and Wilkes[10].

The former is particularly good for signal containing sinusoids and white noise but the ultimate choice is the latter because it works very well if the signal is white noise containing complex exponentials or sinusoids even with very low SNRs.
Typical fresco samples painted in different colors and presenting different grades of surface finish have been investigated and a 15 dB average value for the SNR has been computed; maximum values of 28 dB have been obtained on bright yellows or reds with well polished surfaces. On dark and dirty surfaces the SNR can get as low as 0 dB.

Because these methods do not give correct amplitude values for resonance peaks, we added the possibility to use the Prony method[7], which gives much more precise approximations. To extract these amplitude values the samples series is approximated by a linear combination of exponentials:

$$\hat{x}[n] = \sum_{k=1}^{p} A_k \exp[(\alpha_k + j2\pi f_k)(n-1)T + j\Theta_k] \quad , \tag{8}$$

for $1 \leq n \leq N$,

where:

A_k is exponential amplitude;
α_k is damping factor (s^{-1});
Θ_k is the initial phase expressed in radians.

2.3 LabVIEW

LabVIEW is a well-known software that implements so-called Virtual Instruments (VI) by programming an acquisition board through a PC. Data may be acquired and also modified, stored and displayed in many different ways. What the user see is a panel on his computer screen and by simple mouse clicks this instrument is controlled like a real instrument; the main advantage of this approach is flexibility and any implemented VI may be used as a part of a bigger instrument.
Programming is done using a graphic language based on icons and C written routines may be incorporated as part of the diagram.
In the picture below we show two panels of our acquisition VI. Every measurement parameter may be set to the desired value and software or hardware triggering may be used to synchronize acquisitions.

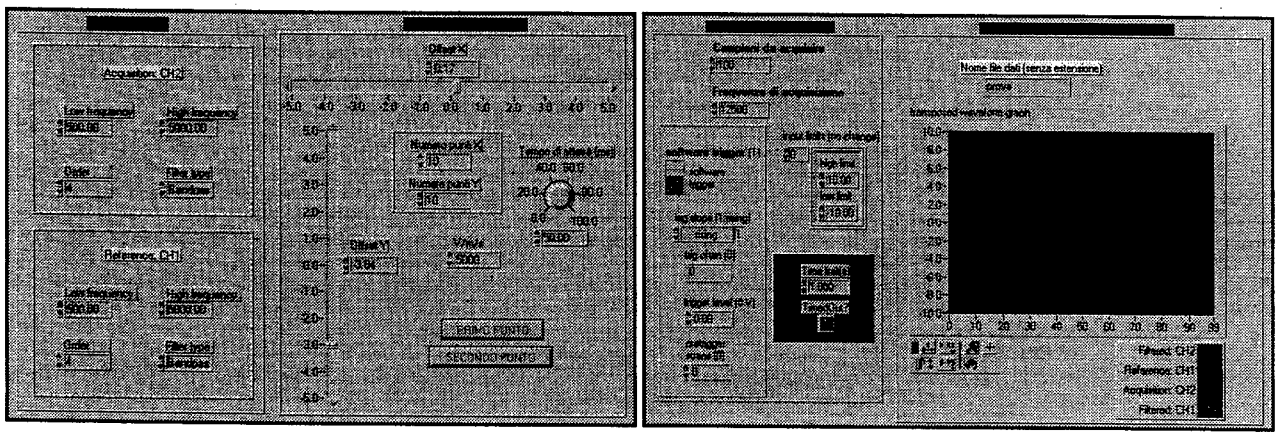


Fig. 4. Two panels of the implemented LabVIEW VI.

Acquired data are transferred as time histories to a MATLAB script that performs all spectral calculations.
An advantage of our implementation is that data storage requires much less disk space than standard instrumentation, typically one tenth, keeping all acquisition parameters equal; for example a 1000 points acquisition on a square grid would require 2 Mb in the former case and 20 Mb in the latter. Saving data in a binary format reduces the need to 7 Mb but data cannot be transferred to external spreadsheet anymore. Measurement time is reduced accordingly, passing from about half an hour to five minutes.

2.4 MATLAB

The proposed MATLAB script implements all spectra calculations described so and moreover may operate on saved time histories to perform transient analysis.
Different MATLAB libraries are employed and five GUIs (Graphical User Interface) allow the user to process data; they are organized as instrumentation panels to give a uniform graphical feel to the software.

Graphic filters have been added to allow different types and grades of data interpolation and smoothing; this feature is particularly useful when operating with noisy signals like in our case.

3. EXAMPLE

To illustrate the implemented software a measurement session on a fresco sample has been implemented. In Fig.5 a sketch of the sample is shown. It is made of two layers of plaster with a calcareous 45 x 45 cm stone as substrate. The surface is painted with different colors to simulate real ancient pigments. All measurements have been made using a regularly spaced 900 points grid, a data record of 200 samples per point, acoustic white noise excitation by a professional horn loudspeaker bandpass filtered from 0.5 to 5 kHz.

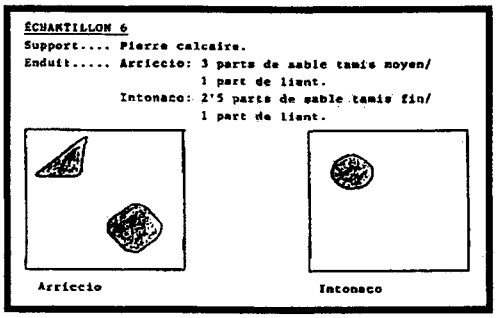


Fig. 5 Fresco sample.

Following an established procedure, at first a map of the point by point value of surface average velocity (RMS value) is traced. In Fig.6 (a) the acquired rough data are shown, while in (b) data have been passed through a median and cubic filter to increase spatial resolution and eliminate isolated high amplitude peaks; defects are clearly evident (to the left of the white 1 and 2 digits). Defect number 1 is closer to the surface than defect 2 and by vibration theory should vibrate more than the latter; a 3D presentation is very effective in showing that measured amplitude values follow the foreseen trend. A profile tracing option may be activated so that just defects edges are traced (see Fig.7). Spectral analysis may be started independently but by clicking on the RMS map the user may view instantly the local spectrum and individuate resonance peaks manually (Fig.8). There is also a MAX function that makes this automatically selecting as many peaks as desired.

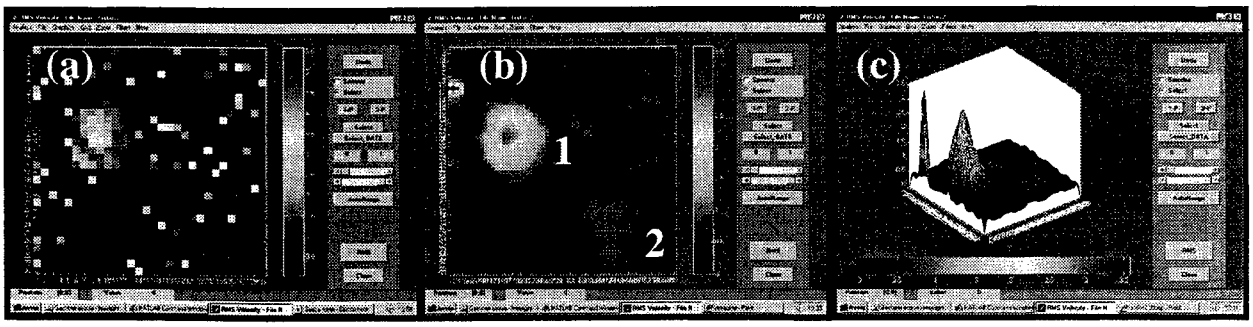


Fig. 6 RMS map of fresco sample (a) rough acquired data (b) smoothed data (c) 3D rendition.

Once activated the spectral analysis menu different kinds of algorithms may be used to calculate and display spectra. Each option has a set of parameters to be set, in particular when employing parametric models, the user is required to fix the model order (AR model) or peaks number (Kumaresan). In Fig.9 we show the result obtained on the same point of the sample selecting respectively the periodogram (a), a parametric model estimator (b) and the Kumaresan algorithm (c). All three methods are able to select the first important peak at about 925 Hz, but Kumaresan calculation defines it with superior resolution.

Once decided which frequency to plot, the user may use the same graphical capabilities described for the RMS analysis; as an example in Fig.10 a 3D rendition is showed, along with its "shaded" version which gives a more pleasant view of the same data.
For comparison in Fig.11 the same sample has been analyzed using a well known commercial system which gave the same value for the principal resonance frequency (925 Hz).

The last menu is dedicated to transient analysis and it is a unique feature of this software when compared to commercial products.

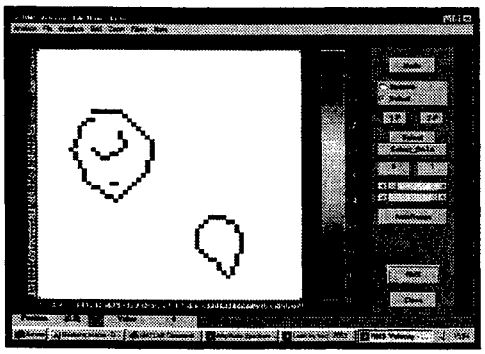


Fig. 7 Defects edge tracing.

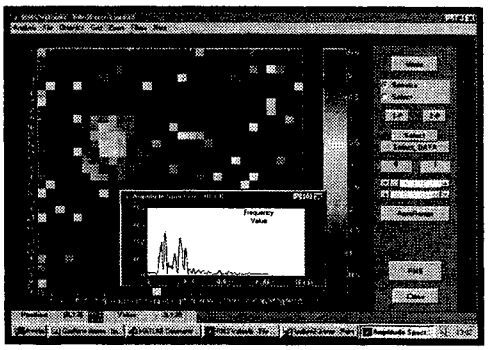


Fig. 8 RMS map with selected point spectra displayed.

The user selects a point on the sample and clicks on the TIME button to visualize the measured time history; selecting it entirely gives the same results of the previous RMS or spectral analysis but cutting it in "slices" of arbitrary length allows the operator to perform partial time analysis. This is quite a useful feature to study, for example, the frequency content of the attack, steady state and decay of a signal. The same time slice is used for all measured point so 2D or 3D maps may be traced.

To obtain more fruitful data this time we employed 1.5 ms pulses instead of steady white noise excitation. In Fig.12 (a) a sample time history is showed and in Fig.12 (b), (c) and (d) RMS maps are traced; the first map refers to the first 3 ms of signal, while the second map to the following 3 ms; last map considers the 12-16 ms time interval. These maps give a clear visualization of time and space distribution of vibrational energy and could be used to create a GIF or AVI animation of surface motion under acoustic stimulation.

4. CONCLUSIONS

In this work a commercial Scanning Laser Doppler Vibrometer scheme has been compared to an implementation based on data acquisition boards controlled by LabVIEW VIs and MATLAB scripts to acquire and manipulate data. Spectral estimators have been implemented as possible alternatives to spectra calculation by standard FFT and time histories rather than spectra storing has been realised.

More noticeable advantages are high-resolution frequency analysis, low noise spectra, automatic resonance frequencies identification, small data files and partial time transient analysis.

Another marked advantage is the possibility to upgrade the measurement system by adding MATLAB modules or adjusting Vis parameters.

5. ACKNOWLEDGEMENTS

The European Commission-Standard Measurement & Testing and BRITE/EURAM programmes under contracts "Laserart", "Impact" and "Valse" has financially supported most of the activity.

The contribution of the following bodies in the field of art conservation is acknowledged:

Laboratoire de Recherche des Monuments Historiques (LRMH), Champs-sur-Marne, France, for co-ordinating the construction of the fresco samples (Dr. M. Stefanaggi).

Mr. Eudald Guillamet and Mr. Javier Chillida, professional restorers, Zaragoza, Spain, for actually making the fresco samples.

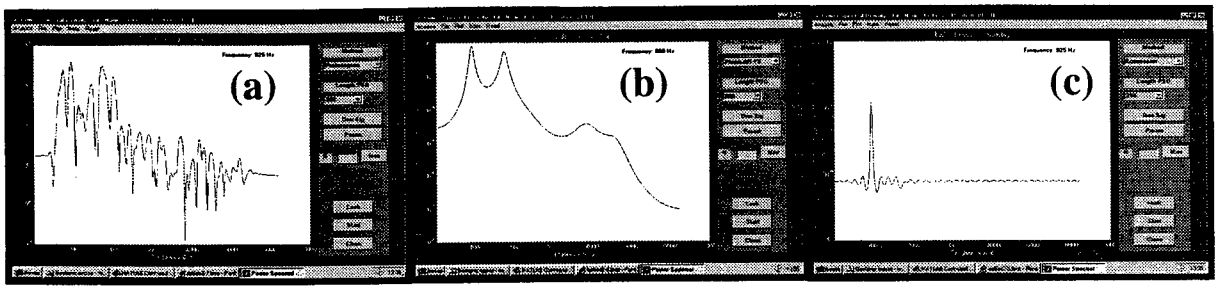


Fig. 9 Frequency spectrum of fresco sample calculated by FFT (a) parametric model (b) Kumaresan (c).

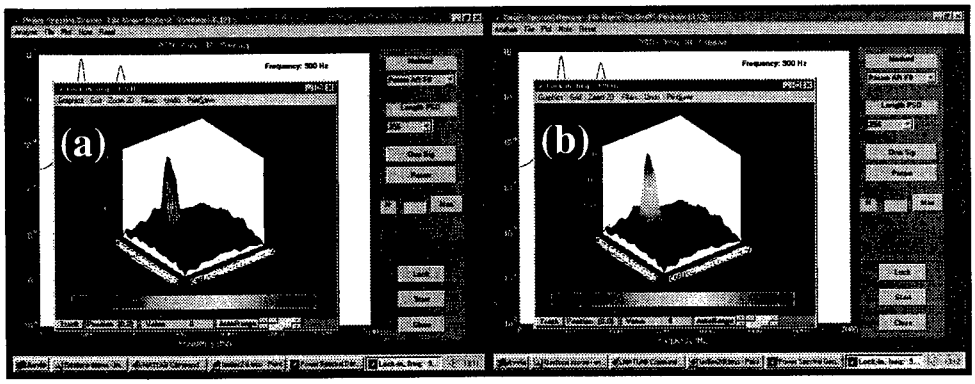


Fig. 10 3D map of sample vibration at 925 Hz (a) not shaded (b) shaded.

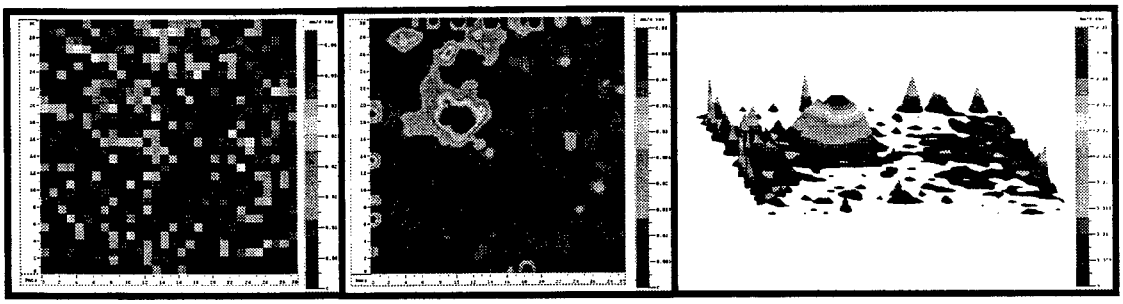


Fig. 11 Fresco sample analyzed by a commercial scanning vibrometer.

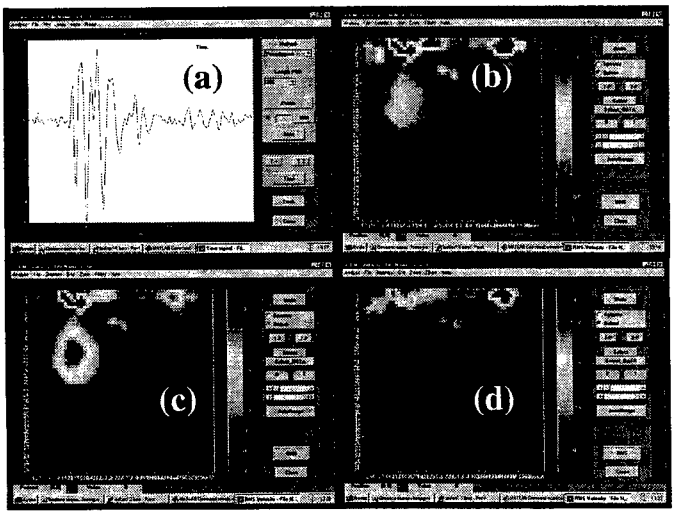


Fig. 12 Transient analysis of fresco sample.

6. REFERENCES

1. D. E. Oliver, "Scanning Laser Vibrometers as Tools for Vibration Measurement and Analysis", *Test Engineering & Management*, pp. 18-21, 1991.
2. P. Castellini, and C. Santolini, "Vibration measurements on naval propeller rotating in water", *Proc. of the 2nd International Conference on Vibration Measurements by Laser Techniques: Advances and Applications*, Ancona, Italy, SPIE Vol. 2868, pp. 186-194, 1996.
3. P. Castellini, E. Esposito, N. Paone, and E. P. Tomasini "Conservation of frescoes paintings and icons: noninvasive measurement of damage by a laser scanning vibrometer", *Proc. of the SPIE Int. Symp. on Nondestructive Evaluation Techniques for Aging Infrastructure & Manufacturing*, San Antonio, Texas, SPIE Vol. 3396, 1998.
4. P. Castellini, E. Esposito, N. Paone, and E. P. Tomasini, "Non-invasive measurements of damage of frescoes paintings and icon by Laser Scanning Vibrometer: experimental results on artificial samples and real works of art ", *Proc. of the 3rd International Conference on Vibration Measurements by Laser Techniques: Advances and Applications*, Ancona, Italy, SPIE Vol. 3411, 439-448, 1998.
5. M. Born, and E. Wolf, *Principles of Optics*, Pergamon Press, Oxford, 1980.
6. L. E. Drain, *The Laser Doppler Technique*, John Wiley & Sons, Chichester, 1980.
7. S. Lawrence Marple Jr., *Digital spectral analysis with applications*, Prentice Hall, Englewood Cliffs (NJ), 1987.
8. D. W. Tufts, R. Kumaresan, "Estimation of frequencies of multiple sinusoids: making linear prediction perform like maximum likelihood", *Proceedings of the IEEE*, **70**, n. 9, September 1982.
9. R. Kumaresan, D. W. Tufts, "Estimating the parameters of exponentially damped sinusoids and pole-zeroes modeling in noise", *IEEE Trans. On Acoustics, Speech and Signal Processing*, **ASSP-30**, n. 6, December 1982
10. J. A. Cadzow and D. M. Wilkes, "Enhanced Sinusoidal and Exponential Data Modeling", in *SVD and Signal Processing II*, ed. by R. Vaccaro, Elsevier Publishers, Amsterdam, 1991.

Torsional vibrations: a laser vibrometry approach

G.M. Revel and E.P. Tomasini

Dipartimento di Meccanica, Università degli Studi di Ancona
Via Brecce Bianche, I-60100, Ancona, ITALY
tel. +39-71-2204441, fax +39-71-2204801
e-mail: revel@mehp1.cineca.it

ABSTRACT

Due to the importance of torsional vibrations in mechanical rotating structures, increasing interest is nowadays devoted to the development of experimental techniques for torsional modal analysis. The rotational laser Doppler vibrometer seems to offer large potentials in this field, as it allows the analysis of torsional vibrations also under operational conditions with simple and non-contacting set-up.
In this work the applicability of the rotational vibrometer has been assessed on a case study (a camshaft), with the main aim of detecting possible interfering inputs. In particular it is shown that the presence of translational or bending motions can introduce uncertainty in the results and in the evaluation of torsional modal parameters. Some suggestions for the reduction of flexural mode effects are given.

Keywords: measurement, laser vibrometry, torsional vibrations, uncertainty.

1. INTRODUCTION

Torsional vibrations in shafts and in rotating objects often represent a significant component of the mechanical stress acting on a structure; furthermore, they are one of the major causes of rotation irregularities and several failures in rotating machinery are generated by vibration problems. Because of that, torsional modal analysis is becoming an important issue on which researchers are spending their efforts [1].
The main problems in the determination of torsional resonances and damping are connected to correct excitation and measurement of rotational vibrations. In fact, the common techniques usually seem to be sensitive also to bending components and other effects, which introduce uncertainty in the results. A suitable torsional hammer was presented by Li et al. [2], which has the capability of inducing a pure torque on a rotating shaft. This system can be helpful in the determination of torsional frequency response functions, as it avoids bending loads usually present when excitation is performed by applying a tangential force.
Concerning the measurement, several techniques can be used, but in most cases they require difficult and intrusive installations, as in the case of variable reluctance pick-up/toothed gear sensors or encoders. Furthermore, in several applications these systems can have inadequate performances, as in the field of active control [3], because of their low resolution and of the noise introduced in the demodulation process. Accelerometers and strain gages, opportunely disposed, can supply information on torsional components, but they are intrusive, they present limitations for "on-line" applications and they can be sensitive also to bending motions.
In recent years rotational laser Doppler vibrometers, which allow the measurement of rotation speed up to 10000 rpm in a frequency band up to 10 kHz, have been introduced [4,5]. These sensors permit the analysis of torsional vibrations also under operational conditions with simple and non-contacting set-up. Their application seems to be useful to extend the applicability of torsional modal analysis [6], but, depending on the specific experiments approached, their performances must

Part of the SPIE Conference on Advanced Photonic Sensors and Applications
Singapore • November–December 1999
SPIE Vol. 3897 • 0277-786X/99/$10.00

be accurately checked. In fact the vibrometers, like all the electro-optical instrumentation, are inherently sensitive to several interfering inputs.
In this work, a rotational laser Doppler vibrometer has been used to determine torsional modal properties of a camshaft in steady conditions. The vibrometer's accuracy for torsional modal analysis is discussed, in particular concerning the system sensitivity to interfering inputs due to bending vibrations.

2. MEASUREMENT PRINCIPLE OF THE ROTATIONAL VIBROMETER

Rotational laser Doppler vibrometers are based on two interferometers with parallel measurement beams used to determine the velocity components (in the direction of the beams) of two different points of the rotating shaft [5]. From the knowledge of these velocity components, it is straightforward to extract the instantaneous value of the rotational speed. In fact, each point on the perimeter of a rotating part of any shape has a tangential velocity v_t depending on the rotational radius R and on the angular velocity ω. As shown in what follows, it is possible to determine the angular velocity ω by measuring two parallel, translational velocity components (Figure 1).

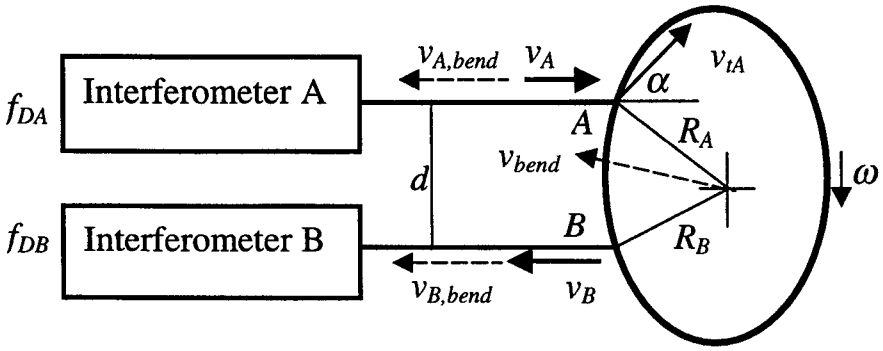


Figure 1. Measurement principle of the rotational vibrometer.

Two interferometers with parallel measurement beams with separation d are used and the velocity components v_A and v_B in the direction of beams are acquired.
Due to vectorial break down, the following formulas apply:

$$\begin{aligned} v_A &= v_{tA}\cos\alpha = \omega R_A \cos\alpha \\ v_B &= v_{tB}\cos\beta = \omega R_B \cos\beta \end{aligned} \tag{1}$$

The velocity components in the direction of the measurement beam produce a Doppler frequencies f_{DA} and f_{DB} in the back scattered beams:

$$\begin{aligned} f_{DA} &= \frac{2\cdot v_A}{\lambda} = \frac{2\cdot\omega\cdot R_A\cdot\cos\alpha}{\lambda} \\ f_{DB} &= \frac{2\cdot v_B}{\lambda} = \frac{2\cdot\omega\cdot R_B\cdot\cos\beta}{\lambda} \end{aligned} \tag{2}$$

The geometrical relationship for the distance d between the measurement beams and the angles α and β at given radii R_A and R_B is given by:

$$d = R_A\cos\alpha + R_B\cos\beta \tag{3}$$

and thus the following formula is obtained for the Doppler frequency shift

$$f = f_{DA} + f_{DB} = \frac{2 \cdot d \cdot \omega}{\lambda} \tag{4}$$

The system works in such a way as to have Doppler frequency shifts f_{DA} and f_{DB} with opposite sign when the velocity vectors have the same versus. The resulting Doppler frequency depends only on the system parameters d and λ and the angular velocity ω and therefore it can be used to measure rotational velocity.
Typical velocity range is from –7000 to 11000 rpm and until 1000 rad/s. The frequency range in vibrational angle measurement is from 1 Hz to 10 kHz. The typical calibration accuracy is of about ± 0.5%. In theory, an additional translational motion (in any direction), overlapping both velocities v_A and v_B, has no influence on the resulting Doppler frequency. Therefore this technique seems to be basically insensitive to bending motions of the rotating object. The only measurement requirement is that the laser beams plane must be perpendicular to the system's rotation axis. No particular surface preparation is usually required.

3. EXPERIMENTAL FACILITIES AND TESTS

In order to asses the applicability of laser rotational vibrometer for torsional modal analysis, a camshaft of a FIAT TIPO 1.4 (Figure 2) was used as a case study. The camshaft has a length of 424 mm, a weight of 2.484 Kg, 8 cams and 4 bearings with a diameter of 24 mm.

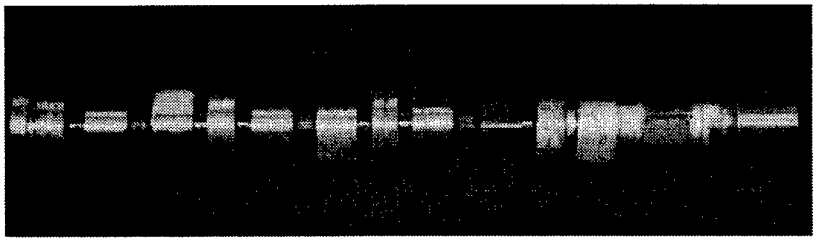


Figure 2. The tested camshaft.

The camshaft was vertically fixed at a heavy base (weight 60 Kg), as shown in Figure 3, and a shaker was employed to excite the camshaft vibrations. The stinger from the shaker was connected to a bar fixed to the free end of the shaft, in such a way as to apply the force tangentially with respect to the camshaft rotation axis. In practice the bar constitutes the arm of the force, which was disposed perpendicular to laser beams direction. In this configuration, both torsional and bending vibrations are excited, as desired, since the aim of the work was also to analyse interfering inputs due to flexural motions. A burst random signal was used to drive the shaker and a load cell acquired the force signal. The sensitivity of the rotational vibrometer was set at 10 °/s/V. The velocity signal was acquired (measuring on the bearings) without any filtering and processed using a modal analysis software.
In order to monitor the bending of the camshaft, a traditional modal analysis was performed also using an axial vibrometer measuring in the same points and with the same excitation.

4. ANALYSIS AND DISCUSSION OF RESULTS

The comparison of the results of the two different modal analyses (torsional and flexural) allows to evaluate what the rotational vibrometer is measuring and the influence of bending modes. In Figure 4 (a) and (b) the sum of the frequency response functions for the two analyses are shown. This function has been chosen for comparison, since it highlights the global modes of the structure under investigation.

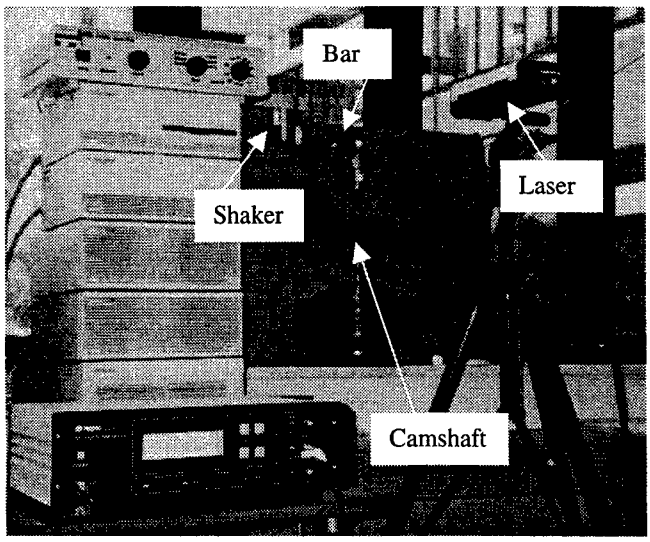


Figure 3. The experimental set-up.

In this case the problem is to verify if the modes measured by the rotational vibrometer are real torsional modes or not, in order to assess the reliability of this technique for modal analysis. From Figure 4 it appears evident that some modes are common to both the analyses. These modes are probably bending modes, since the axial vibrometer, measuring in ideal conditions, can not be sensitive to torsional vibrations. Thus, the other modes, marked with a circle in Figure 4 (a), are torsional.

An indication of the presence of some inconsistencies in the results is given also by the auto-MAC (Modal Assurance Criterion) values for the torsional modal analysis, reported in Figure 5 for the frequency range between 0 and 4000 Hz. Usually, high auto-MAC values out of the diagonal indicate problems due to low spatial resolution, which can lead to aliasing phenomena in the geometrical description of the high frequency modes. In the present case, it is worth noting that high values have been found in the correlation of the low frequency modes, where the most important bending modes are excited. Thus, these modes seem to be observed by the vibrometer in a similar way with respect to the first torsional modes.

Therefore, the analysis of results shows that in particular conditions the torsional laser vibrometer can be sensitive also to bending modes. In order to describe the possible causes of that, we can suppose to have a generic additional bending motion of the shaft (Figure 1) with velocity v_{bend}. This motion will generate an additional Doppler shift in both the interferometers proportional to the velocity components ($v_{A.bend}$ and $v_{B.bend}$) induced by the bending vibration in the laser beams directions. Therefore, the resulting Doppler shift will be:

$$f = f_{DA} + f_{DB} + (f_{A.bend} + f_{B.bend}) = 2/\lambda[(v_A + v_B) + (v_{A.bend} + v_{B.bend})]. \quad (5)$$

In ideal conditions, when the vibrometer is correctly installed and its arms are perfectly balanced, since the system works in such a way as to have Doppler frequency shifts at the interferometer outputs with opposite sign when the velocity vectors have the same versus, $f_{A.bend}$ and $f_{B.bend}$ are equal, but with opposite sign. Thus, the effect of the generic additional bending modes will be null.

On the contrary, if for some reasons the two velocity $v_{A.bend}$ and $v_{B.bend}$ are seen to be different, the term $(f_{A.bend} + f_{B.bend})$ will no longer be zero. Therefore, the rotational vibrometer becomes sensitive also to bending modes. Possible causes of not perfect balancing between $f_{A.bend}$ and $f_{B.bend}$ could be the following:

a) not perfect balancing of the measurement conditions for the two arms: this can be due to optical (e.g. the local reflection conditions are different for the two spots; the two beams are disposed highly asymmetrically with respect to the rotation axis), electrical or hardware (e.g. differences in the two photodetectors) causes;
b) not perfect parallelism between the two measurement beams;
c) $v_{A.bend}$ and $v_{B.bend}$ are not perfectly equal (e.g. the structure or the mode has a particular complex shape);
d) the laser beams plane is not perpendicular to the rotation axis: this will cause the two beams to measure on different sections of the shaft, where probably different bending velocity are present.

These problems are minimised when the vibrometer is measuring on a rotating shaft. In fact, also if speckle noise increases at the fundamental rotation frequency with respect to the case of modal analysis in steady conditions, the desired output f_{DA} + f_{DB} is larger (the higher the velocity ω, the larger the Doppler shift $f_{DA} + f_{DB}$). Therefore, the effect of bending will have a lower relative influence. On the contrary, in the test performed, the shafts vibrates and rotates on a steady condition and thus the terms $(f_{DA} + f_{DB})$ and $(f_{A,bend} + f_{B,bend})$ are comparable. In this situation, because of the contribution of the above factors, the frequencies of the larger bending modes are recorded in the rotational velocity signal.

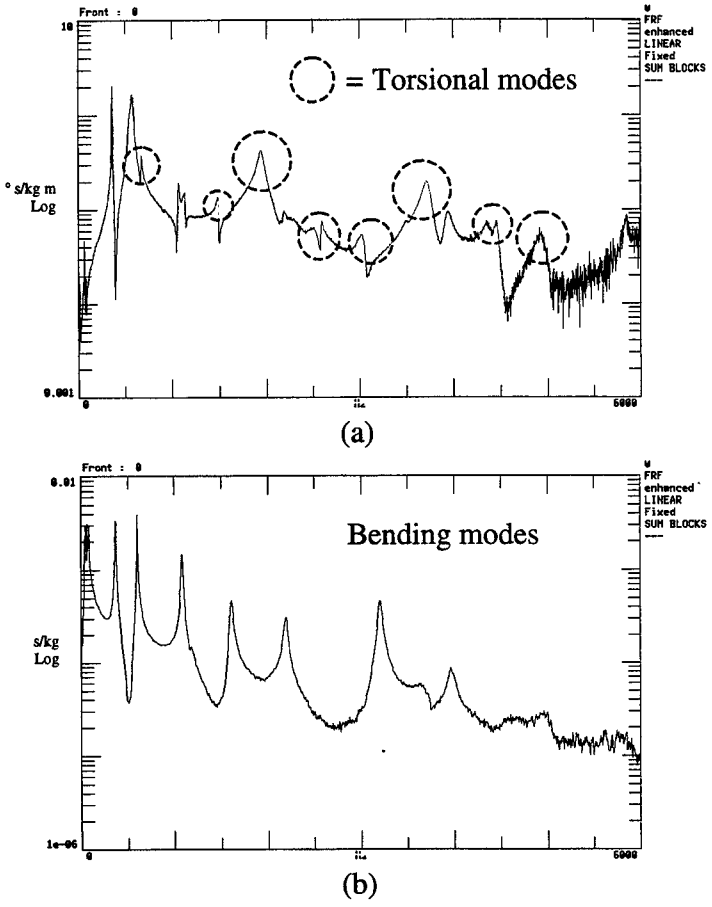


Figure 4. FRF's sum for the modal analysis performed with the torsional vibrometer (a) and with the axial vibrometer (b) respectively.

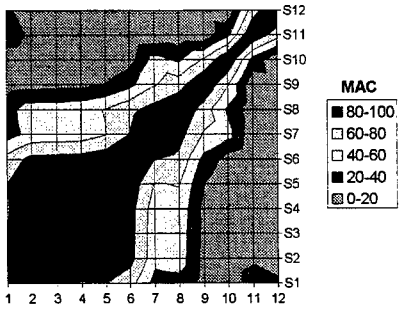


Figure 5. AutoMAC matrix (0–4000 Hz) for the torsional modal anlysis.

5. CONCLUSIONS

In this work a rotational laser Doppler vibrometer has been tested for experimental torsional modal analysis of a camshaft. The vibrometer accuracy is discussed, in particular concerning the system sensitivity to interfering inputs due to additional bending vibrations, which could be excited during the test. Some suggestions for the reduction of these undesired interfering inputs could be obtained from the analysis of the presented results. In particular:

- arrange the experimental set-up in order to have good balancing and symmetry for the two laser beams and check that the signal level for the two interferometer is equal;
- excite the structure in such a way as to reduce flexural motions (e.g. using the torsional hammer by Li et al. [2]);
- some indicators (e.g. auto-MAC) can be used to understand if some inconsistencies are present in the results;
- monitor bending vibrations using an additional sensor.

REFERENCES

1) B.J. Stone, "Torsional vibration – its importance, excitation, measurement and modelling", *Proceedings of the Experimental and Theoretical Mechanics Conference*, ETM93, pp 182-190, Bandung, Indonesia, 1993.

2) H. Li, R.D. Entwistle, and B.J. Stone, "Impact excitation of torsional vibration", *The International Journal of Analytical and Experimental Modal Analysis*, **9** (3), pp 163-174,1994.

3) I. Ali, R. Wynn, "Sensing torsional vibration for active control", *Proceedings of the 15th International Modal Analysis Conference*, pp 1483-1487, Orlando, 1997.

4) P.G. Eastwood, N.A. Haliwell, "Laser-based measurement of torsional vibration", *SPIE Optics in Engineering Measurement*, Vol. 599, pp 358-366, 1985.

5) A.C. Lewin, R. Volkmar, S. Georg, "New concept for interferometric measuremnt of rotational vibrations", *Proceedings of the 1st International Conference on Vibration Measurements by Laser Technqiues*, SPIE Vol. 2358, pp 24-36, Ancona, Italy, 1994.

6) S. Seidlitz, "Using modal analysis to verify a torsional mass elastic model", *Proceedings of the 16th International Modal Analysis Conference*, pp 1505-1509, Santa Barbara, 1998.

SESSION 7

Current/Voltage Sensors

Problems of Optical Fiber Voltage Sensor in Practical Usage

Chongqing Wu*, Prof. SPIE member, Yongjun Wang, Xiangjun Xin
Northern Jiaotong University Inst. of Lightwave Tech
Beijing (100044) China

Abstract

Optical Fiber Voltage Sensor (OVS) has potential applications in power supply network, but upon to now the OVS can't be industrially used in practice. This paper introduces some problems we met in development of practical bulk-modulation OVS using BGO ($Bi_4Ge_3O_{12}$) crystal. The optical source must be pre-modulated so that to detected the average optical power easily. The measure error (linear error and phase shift), nonlinear distortion, sensitivity produced by deviation of optical element are discussed. Analysis and tests show, that a small collimating error of optical elements will lead to decrease the sensitivity and increase phase shift. So long as the angle α between the polarizer and $\lambda/4$ wave plate is equal to $\pi/4$, the location of the polarizer for detection and BGO is not very important and the maximum sensitivity can be achieved when the angle γ between the detected polarizer and BGO is equal to $\pi/4$. But if $\alpha \neq \pi/4$ and $\gamma \neq \pi/4$, not only the sensitivity will be decreased, but also the distortion will be increased, depending on the angle β between BGO and $\lambda/4$ wave plate. So we use a quartz plate to hold $\lambda/4$ wave plate in order to obtain good circular polarization light.

In order to eliminate the effect of random fluctuation of average optical power, some data process must be needed. We use 80C196KC single chip CPU to calculate the instantaneous value, and the effective value of measured voltage can be displayed in this transducer. We use silicon gel insulator to introduce high voltage to sensor. Some calculation and test results are given in this paper.

Keywords: Optical Fiber Voltage Sensor (OVS), BGO ($Bi_4Ge_3O_{12}$) crystal , measure error, sensitivity, data process

1.Introduction

Optical fiber Voltage Sensor (OVS) has potential applications in power supply network because it has many advantages. For example it can completely isolate power network and test circuits, because of isolation characteristics of optical fiber, and can avoid the dangers of shorts of second circuit of potential transformer (PT). It can also decrease the harmonic vibration of power network produced by transition process because of inductanceless of OVS, etc. Many countries developed various kinds of OVS, for example, all optical fiber OVS[1][2], optical waveguide OVS[3][4][5], etc. But so far the most stable and relative one is the bulk-modulation OVS in practical usage.

The bulk-modulation OVS is based on electrooptical effect of electrooptical crystal. BSO ($Bi_{12}SiO_{20}$), BGO ($Bi_4Ge_3O_{12}$), LN ($LiNbO_3$) electrooptical crystals are in common use. Because BGO has higher half-wave voltage and stable characteristics, it's the most commonly adopted one. Fig. 1 shows construction of bulk-modulation optical fiber transducer and Fig. 2 shows the construction of OVS.

BGO crystal is classified as cubic system, when test voltage is not put on it, it's an isotropic propagation medium. But when test voltage is put on its (110) crystal plane, it changes into biaxial crystal caused by Pockels electrooptical effect. The index difference between double principal axes is proportional to the test voltage put on this crystal. When two optical beams, which polarization directions coincide with the two principal axes, respectively, incident into the crystal from its ($\bar{1}10$) crystal plane, the phase difference ϕ between two beams can be expressed as

$$\phi = \frac{2\pi n_0^3 r_{41} U l}{d\lambda} = \frac{\pi U}{U_\pi} \tag{1}$$

were λ is wavelength in vacuum, n_0 is index of crystal without modulation, γ_{41} is electrooptical coefficient, l is crystal length along with optical propagation direction, U is test voltage, d is crystal length along with electric field, and $U_\pi = \lambda d / 2n_0^3 r_{41} l$ is half-wave voltage, which is only depending on crystal and incident optical wavelength. From formula (1) we can see, that ϕ is proportional to test voltage, so we can determine test voltage by measuring ϕ.

*Correspondence: Email: cqwu@center.njtu.edu.cn; Telephone: (86)63240046

Part of the SPIE Conference on Advanced Photonic Sensors and Applications
Singapore • November–December 1999
SPIE Vol. 3897 • 0277-786X/99/$10.00

Above principle have proved by many experiments, but when using it for practice some problems were met. This paper will discuss these problems: measure error (linear error and phase shift), nonlinear distortion, sensitivity depending on deviation of optical elements etc. Finally introduce the data process in order to eliminate the influence of random fluctuation of average optical power. Some calculation and test results are given.

2.Problems caused by collimating error of optical elements of OVS

Analysis and tests show that assemble and package of optical elements must be very careful, a small collimating error will lead to decease the sensitivity, and increase phase shift and nonlinear distortion. We assume that the polarization direction of incident beam after passing through input polarizor is coincide with x axis, the output beam of sensor can be written as (see Fig.2)

$$\begin{bmatrix} E_x \\ E_y \end{bmatrix} = M_A M_3 M_c M_2 M_w M_1 \begin{bmatrix} E_0 \\ 0 \end{bmatrix} \tag{2}$$

were $M_A = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$, $M_c = \begin{bmatrix} 1 & 0 \\ 0 & e^{-j\phi} \end{bmatrix}$, $M_W = \begin{bmatrix} 1 & 0 \\ 0 & -j \end{bmatrix}$

they are transmission matrixes of output polarizor, crystal and $\lambda/4$ wave plate, respectively.

And (see Fig.3) $M_3 = \begin{bmatrix} \cos\gamma & \sin\gamma \\ -\sin\gamma & \cos\gamma \end{bmatrix}$, $M_2 = \begin{bmatrix} \cos\beta & \sin\beta \\ -\sin\beta & \cos\beta \end{bmatrix}$, $M_1 = \begin{bmatrix} \cos\alpha & \sin\alpha \\ -\sin\alpha & \cos\alpha \end{bmatrix}$

they are transform matrixes caused by coordinate transformation from crystal, $\lambda/4$ wave plate and input polarizor to output polarizor, crystal and wave plate, respectively.

Thus we can get

$$\begin{bmatrix} E_x \\ E_y \end{bmatrix} = E_0 \begin{bmatrix} \cos\gamma(\cos\beta\cos\alpha + j\sin\beta\sin\alpha) + e^{-j\phi}\sin\gamma(-\sin\beta\cos\alpha + j\cos\beta\sin\alpha) \\ 0 \end{bmatrix} \tag{3}$$

and the output optical power I is

$$\begin{aligned} I = I_0(&\cos^2\gamma\cos^2\beta\cos^2\alpha + \sin^2\gamma\sin^2\beta\cos^2\alpha + \sin^2\gamma\cos^2\beta\sin^2\alpha \\ &+ \cos^2\gamma\sin^2\beta\sin^2\alpha + \frac{1}{2}\sin\phi\sin 2\gamma\sin 2\alpha - \frac{1}{2}\cos\phi\sin 2\gamma\sin 2\beta\cos 2\alpha) \end{aligned} \tag{4}$$

From this formula, we can see that:

(1)If we divide this sensor into two parts: input polarizor and $\lambda/4$ wave plate combine one part, and BGO crystal and output polarizor another part. So long as the angle α between input polarizor and wave plate is equal to $\pi/4$, the relative location of two parts (angle β) is not very important, in this case,

$$I = I_0(1 + \sin 2\gamma\sin\phi)/2 \tag{5}$$

If the angle γ between BGO crystal and output polarizor is also equal to $\pi/4$, the maximum sensitivity can be achieved, and

$$I = I_0(1 \pm \phi)/2 \tag{6}$$

In this case, because the test voltage is sine signal $U = U_0\sin\omega t$, we ought to consider the error of effective value as measuring sensitivity, not of instantaneous value, thus

$$V = V_0(1 \pm \sin(m\sin\omega t)) = V_0(1 \pm 2\sum_{n=0}^{\infty} J_{2n+1}(m)\sin(2n+1)\omega t) \tag{7}$$

were $V_0 = RI_0/2$, R is the responsivity of PIN, $m = \pi U_0/U_\pi$ is called as modulation degree, $J_{2n+1}(m)$ is $2n+1$ order Bessel function. There are two kinds of distortion: effective value error and nonlinear distortion. The effective value error can be expressed as

$$\Delta = \frac{J_1(m) - m}{m} = \frac{J_1(m)}{m} - 1 \tag{8}$$

due to $m < 1$, thus

$$\Delta \approx 0.624(m/3)^2 \tag{9}$$

If demand $\Delta < 1\%$ for practical usage, must $m < 0.39$ or $m < 0.4$. For our sensor, $U_\pi = 37\text{kV}$, so the linear region of effective value is 0~3300V.

The nonlinear distortion can be expressed by nonlinear distortion coefficient Γ_0, here it is

$$\Gamma_0 = \sum_{n=0}^{\infty} J_{2n+1}^2(m) / J_1^2(m) \tag{10}$$

substituting $m = 0.4$, one can calculate $\Gamma_0 = 26\times10^{-6}$. This result shows that in this case nonlinear distortion is not important.

The sensitivity is also depending on noise N, including quantum noise, dark current noise, amplifier noise and other noise. If the acceptable signal-noise-ratio (SNR) for average optical power without test voltage modulation is $\eta_0 = V_0^2/N$, then for an alternating effective value the SNR is $\eta = 2J_1^2(m)\eta_0$, it means that sensitivity of alternating effective value is $2J_1^2(m)$ times of average optical power.

(2)If $\alpha \neq \pi/4$, we assume $\gamma = \pi/4$ for simplification, in this case output optical power is not only depending on α, but also on β, and

$$I = I_0(1 + \sin 2\alpha \sin\phi - \sin 2\beta \cos 2\alpha \cos\phi) \tag{11}$$

From formula (11), can get

$$V_{ave} = V_0[1 - \sin 2\beta \cos 2\alpha J_0(m)] \tag{12}$$

$$v_1(t) = 2V_0 \sin 2\alpha J_1(m) \sin\omega t \tag{13}$$

$$v_{hi}(t) = 2V_0\left[\sin 2\alpha \sum_{n=1}^{+\infty} J_{2n+1}(m)\sin(2n+1)\omega t - \sin 2\beta \cos 2\alpha \sum_{n=1}^{+\infty} J_{2n}(m)\cos 2n\omega t\right] \tag{14}$$

were V_{ave}, $v_1(t)$, and $v_{hi}(t)$ are average, base harmonic and high order harmonic, respectively. The SNR for this case is decreased to

$$\eta_1' = \eta_1 \cos^2 2\theta \tag{15}$$

were θ is the deviation of α angle from $\pi/4$. So we can see that α is the most important for the sensitivity. In order to maintain $\alpha = \pi/4$, we hold a $\lambda/4$ wave plate by using of a quartz plate to obtain good circular polarization light before regulating this sensor.

The linear region of effective value is not changed with variation of angle α, but the nonlinear distortion is. In this case, nonlinear distortion coefficient Γ_1 is

$$\Gamma_1 = \left[\sin^2 2\alpha \sum_{n=1}^{50} m^{4n+2} / 2^{(4n+2)!}((2n+1)!)^2 + \sin^2 2\beta \cos^2 2\alpha \sum_{n=1}^{50} m^{4n} / 2^{4n}(2n!)^2\right] / m^2 \sin^2 2\alpha \tag{16}$$

Substituting relevant parameter, can get

$$\Gamma_1/\Gamma_0 = 2.7 \tag{17}$$

Although the nonlinear distortion is still very small, but it increase 2.7 times than further case.

(3)Phase shift between test voltage and output signal is an important problem, but the reason has not been found. The experiments show that at 50 Hz the shift can be less than 1° depending on collimating, if the collimating is not very good, the phase shift sometimes maybe achieve 3-5° . Phase shift is decreased with increase of frequency, and at 80 Hz the shift is about zero. With increasing frequency continuously, the shift will change into negative.

3.Pre-modulation and data process

We know, in this sensor there are many noise, and the dark current noise and quantum noise are the most important of all. They change with temperature, influencing output average power lever of PIN, making unstability of working point. But the fluctuation of dark current or direct current working point varies very slowly, so we use an isolation capacity to prevent this noise from entering next amplifier. For this reason, we premodulate the LED to make the optical signal incident into sensor is an alternating signal. A simple circuit, for example, a voltage controlled oscillator CD4046, can complete this purpose. Due to the frequency of test voltage is about 50 Hz, we let the premodulation speed-ratio be equal to 500 kHz, it can't affect test bandwidth.

In order to eliminate the effect of random fluctuation of average optical power, some data process must be needed. There are many reasons to produce random fluctuation of average optical power, for example, random fluctuation of optical source, random variation of attenuation of optical elements and connection loss etc. In BGO crystal there is another important noise, natural birefringence. The affect of natural birefringence may be considered as an additional phase shift, because there are many kinds of natural birefringence (total amount is *N*), it can be written as

$$I = \frac{I_0}{2}\left[1+\left(\phi+\sum_{n=1}^{N}\delta_{nl}\sin 2\theta_n\right)\right] \tag{18}$$

were $\delta_{nl} = \dfrac{k_0 n_0 \Delta\varepsilon_{nl} l}{2\varepsilon_0}$ is phase shift of n-th natural birefringence, θ_n is direction angle of n-th natural birefringence, ϕ is normal phase shift caused by test voltage modulation. Because δ_{nl} varies with temperature slowly, it also lead to random fluctuation of average optical power. Thus

$$\phi = \frac{2I}{I_0\left[1+\sum_{n=1}^{N}\delta_{nl}\sin 2\theta_n\right]} = \frac{2I}{I_{ave}} \tag{19}$$

So the first purpose of data process is calculating instantaneous value using formula (19). We use 80C196KC single chip CPU to calculate this value. Fig. 4 is the block scheme of data process. The most major difficulty of using 80C196KC is the processing speed. The oscillator frequency of 80C196KC is 16 MHz, and the machine cycle is two times of oscillator. An instruction have 2~3 machine cycles, and to ensure for effectively reading and writing, must insert 3~5 machine cycles. Thus, it takes 0.6~1 μs to perform an instruction. If for one cycle of test voltage the 100 sample points is needed (the bandwidth of test voltage is less than 5 kHz), the time space between two sample points is about 200 μs. Subtracting sample time and A/D transformation time (about 18 μs), the time to perform program is only about 180 μs, so the program is not over 200 instructions. Otherwise, if the program is memorized in outer EPROM, it will take additional time to read instruction from outer EPROM. So all the instructions and data are memorized in internal registers of 80C196KC.

The effective value of test voltage can also be calculated in this data process unit. The effective value can be expressed as

$$U = \sqrt{\frac{1}{T}\sum_{0}^{N} f^2(k\Delta t)\Delta t} \tag{20}$$

were $f(t)$ is the signal after A/D transformation of test voltage, Δt is time space of sample, T is the cycle of test voltage, N is amount of sample points in one cycle. Because the cycle varies slightly, N is not a constant, so using formula (20) is very difficult.

We use the effective value of non-cyclical signal to substitute effective value of cyclical signal, that is

$$U = \lim_{T\to\infty}\sqrt{\frac{1}{T}\int_0^T f^2(t)\mathrm{d}t} = \sqrt{\frac{1}{N}\sum_{0}^{N} f^2(t)} \tag{21}$$

So long as the amount of sample points is very bigger, the variation of cycle of test voltage slightly affects the calculation result. Thus the calculation is greatly simplified.

4.Summary

We have developed the Optical fiber Voltage Transducer, which will be applied in 35kV power supply network of electrical railway. It is combined by OVS sensor, data process apparatus, and introducing set for test voltage, see Fig.5. We use silicon gel insulator to introduce high voltage to sensor. Tests show that this kind of OVS can be satisfied for practical usage of 35kV power supply network.

References

1. L. Fabiny et al, "High Resolution Fiber-optic Low-Frequency Voltage Sensor Based On the electrostrictive effect" IEEE *Photonic Technology Letters*, Vol. **5**(8), pp.952-953, 1993.
2. Chia-Nan Chang et al, "Temperature and Voltage sensors by Mode-Mode Interference in Polarization-Maintaining Optical Fibers, " *J. of the Chinese institute of Engineers*, Vol. **18**(5), pp.625-632, 1995.
3. Nicolas A.F.Jeager et al, "High Voltage Sensor Employing an Integrated Optics Mach-Zehnder Interferometer in Conjunctive with a Capacitive Divider," *J. of Light-Wave Tech.* Vol. 7(2), pp.229-234, 1989.

4. Sang-Shin Lee et al, "Integrated Optical High Voltage Sensor Based on a Polymeric Y-Branch Digital Switch", IEEE *Photonics technology Letters*, Vol. 8(7), pp.921-923, 1996.
5. S.A.Kingsley and S.Siron, "Parallel-Plate Integrated Optic High Voltage Sensor," *Electronic Letters*, Vol. 31(13), pp.1096-1097 1995.

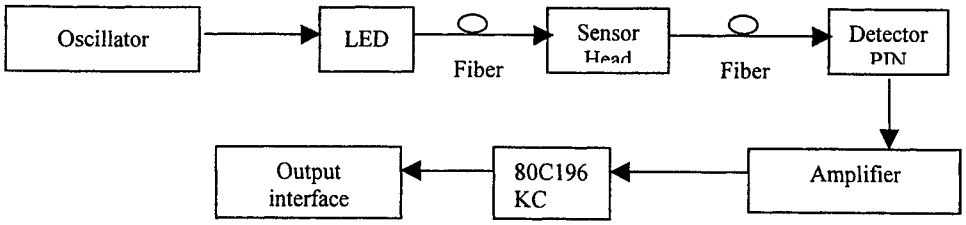


Fig.1 construction of bulk-modulation optical fiber transducer

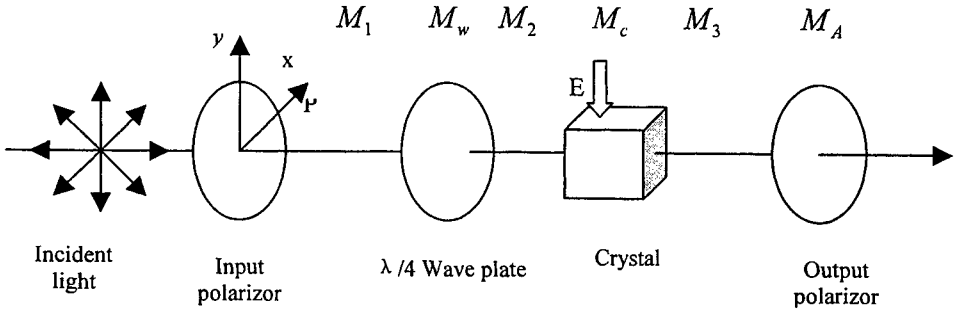


Fig.2 construction of OVS

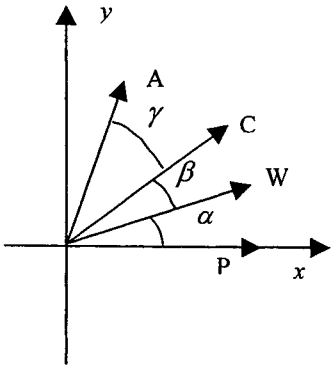


Fig.3 Angles between elements
P: fast axis of input polarizor
W: fast axis of wave plate
C: principal axis of crystal
A: fast axis of output polarizor

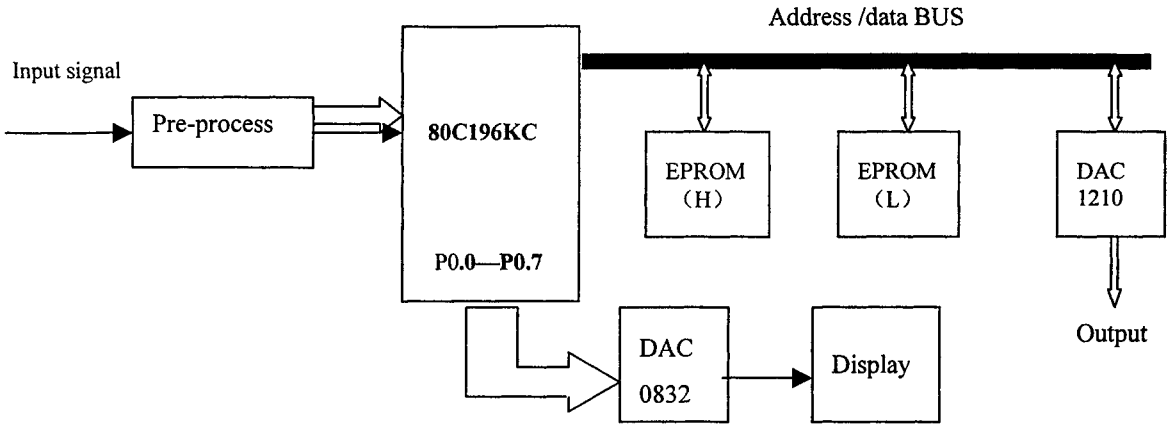


Fig.4 block scheme of data process

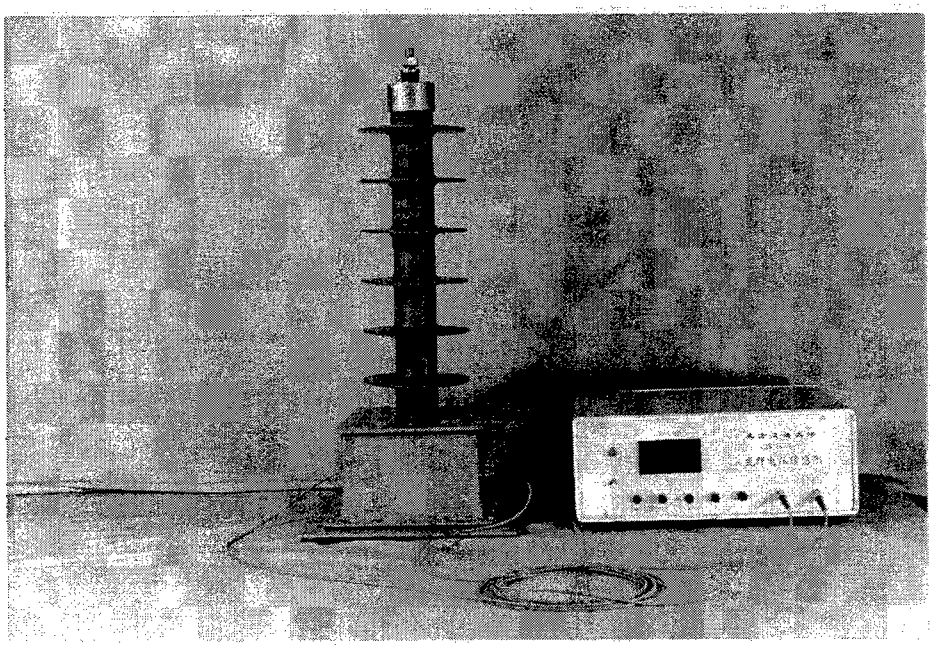


Fig.5 Optical fiber Voltage Transducer used for 35kV power network

An optical current transducer with the compensation for the polarization and the temperature disturbance

Deming Liu*, Jingyi He, Dexiu Huang

Dept. of Optoelectronic Engineering, Huazhong Univ. of Sci. & Tech., P. R. China

ABSTRACT

This paper presents a novel bulk type of optical current transducer (OCT), in which a single-mode fiber to multi-mode fiber coupling configuration is adopt to reduce the optical loss and to restrain the random mode-coupling noise in a conventional multi-mode fiber to multi-mode fiber coupling structure. A special design is used to automatically compensate the changes of the State of Polarization (SOP) of the optical fiber. Experiment shows that about 10 times of precision improvement can be achieved compared with conventional sensing element without SOP compensation. The OCT system employs microcomputer together with corresponding software to accomplish signal processing. The overall performance of the OCT system is tested experimentally. The results show that the system has good long-term stability — the precision during 100 hours is 0.3%, and temperature compensation can considerably improve the temperature characteristic of the system.

Keywords: Optical current transducer, Fiber-optic sensor, Polarization compensation, Temperature compensation

1. INTRODUCTION

OCT technology is finding an increasing interest in the electric industry, especially in the very high voltage application situation where the excellent high voltage isolation is of great importance[1-3]. The all-dielectric construction of OCT provides many significant advantages over conventional electrical CT: well isolation performance, no danger of explosion, no saturation under fault current conditions, no ferroresonance effects and no hysteresis effects, immunity to electromagnetic interference, small volume, light weight and potential low cost, and so on.

The bulk-type OCT is recognized as the most promising structure to be commercially used. However there are still several obstacles for the stable operation of the OCT, including the influence of the random disturbance of the state of polarization (SOP), the coupling between guide modes in the multi-mode fiber and the temperature on the sensitivity of the OCT system. In this paper a new configuration for the bulk-type OCT with the compensation for the polarization- and the temperature-induced fluctuation was presented. The single-mode fiber was used as the transmission line so as to reduce the coupling loss and the mode-coupling noise. A polarization-insensitive configuration was specially designed for eliminating the random disturbance of SOP of the single-mode fiber. At the same time a novel temperature-sensing scheme was used to compensate the temperature fluctuation. The fabricated OCT system employed a microcomputer together with the corresponding

* Correspondence: Email: dmliu@mail.hust.edu.cn; Telephone/Fax: 0086-27-87543355

Part of the SPIE Conference on Advanced Photonic Sensors and Applications
Singapore • November–December 1999
SPIE Vol. 3897 • 0277-786X/99/$10.00

software to process the signal data, supplying for the function of in-situ monitoring the electric current value and the waveform and storing the data for further analysis. A detail depiction of the OCT system as well as the measurement and analysis of the performances were also given in this paper. The results showed that a 10 times of precision improvement could be achieved compared with the conventional OCT sensing system. The new OCT system has the excellent long-term stability and reliability.

2. THEORY

In most cases, the operation principle of the OCTs is based upon the "Faraday magneto-optic effect": the polarization plane of a linearly polarized light propagating through a magneto-optical material is rotated under the influence of the magnetic field generated by the electrical current to be measured. This can be described as following:

$$\phi = V \int_l \boldsymbol{B}\, d\boldsymbol{l} \ , \tag{1}$$

where ϕ is the Faraday rotation angle, V is the Verdet constant (about $10^{-5} \sim 10^{-6}$ rad/A [4]) of the magneto-optic material, B is the magnetic flux density along the optical path, and l is the optical path. When the optical path forms a closed loop, Eq. (1) becomes:

$$\phi = V \oint \boldsymbol{B}\, d\boldsymbol{l} \ . \tag{2}$$

If the light encircles the current carrying conductor for N turns, according to Ampere's law, Eq. (2) can also be written as:

$$\phi = \mu V N I \ , \tag{3}$$

where μ is the permeability of the magneto-optic material, I is the electrical current to be measured. So the current value can be attained if the Faraday rotation angle ϕ is detected.

The sensor head of OCT is generally composed of an input fiber, an input graded-index (GRIN) lens for the collimation of the light, a magneto-optic glass circuit for the rotation of the SOP, an output GRIN lens for the convergence of the light and an output fiber. In a conventional OCT the coupling configuration is multi-mode fiber to multi-mode fiber. There is huge coupling loss because of the large distance between the input fiber and the output fiber, reaching about 22 dB (including 6dB polarizer loss). The random coupling between the discrete modes results in a noise to the system. Moreover, there is also considerable polarization noise in the multi-mode fibers. In order to overcome this problem N. E. Fisher and D. A. Jackson suggested a common-mode optical noise-rejection scheme to restrain the polarization noise[5], which was involved with a complex structure with one up-lead plus two down-lead fibers.

The scheme of the sensing element we developed is shown in Fig. 1. The light beam emitted from the light source is launched into the optical fiber 1, which is a single-mode fiber integrated with a GRIN lens at the rear end. The light beam emitted from the GRIN lens is collimated, and is then divided into two orthogonally polarized light beams by a polarizing beam splitter (PBS). The polarization planes of the two linearly polarized light beams are rotated respectively when they transmit through the magneto-optic glass circuit. After they pass through a $\lambda / 2$ waveplate, each of them is divided into two orthogonally polarized light beams by another PBS. So there are four linearly polarized light beams in the PBS, and two orthogonally polarized beams which derive from the two different light beams before the PBS are combined with each other when they pass through the PBS, and then launched into the optical fiber 2 (multi-mode fiber). The final output light from

the fiber 2 is detected by a PIN detector, which generates the electrical signal for further processing.

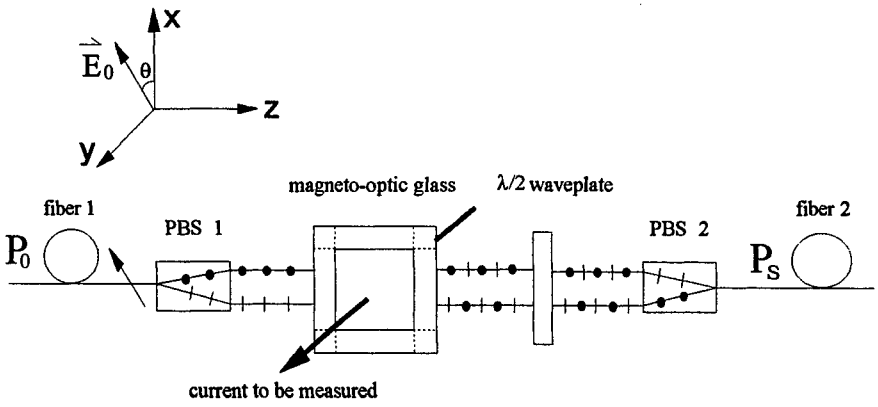


Fig. 1 Schematic diagram of the sensing element with SOP compensation

Since a light beam of random polarization state can be considered as being composed of many linearly polarized beams, we suppose that the light beam incident PBS 1 is linearly polarized, and the azimuth of its polarization plane is θ to x axis. Using Jones matrix, it can be described as following:

$$\vec{E}_0 = A\begin{bmatrix}\cos\theta \\ \sin\theta\end{bmatrix}. \tag{4}$$

After it passes through PBS 1, it becomes two linearly polarized beams:

$$\begin{aligned}\vec{E}_1^{(1)} &= A\sin\theta\begin{bmatrix}0\\1\end{bmatrix} \\ \vec{E}_2^{(1)} &= A\cos\theta\begin{bmatrix}1\\0\end{bmatrix}\end{aligned}. \tag{5}$$

The SOP of them after transmitting through the magneto-optic glass circuit and the $\lambda / 2$ waveplate successively is denoted by $\vec{E}_{1,2}^{(2)}$ and $\vec{E}_{1,2}^{(3)}$ respectively. If the Faraday rotation angle of the SOP of the lights is ϕ while transmitting through the magneto-optic glass, the transmission matrix (denoted by G_1) of this course can be expressed as:

$$G_1 = \begin{bmatrix}\cos\phi & -\sin\phi \\ \sin\phi & \cos\phi\end{bmatrix}, \tag{6}$$

then, for the two beams modulated by the electrical current we get:

$$\begin{aligned}\vec{E}_1^{(2)} &= G_1\vec{E}_1^{(1)} = A\sin\theta\begin{bmatrix}-\sin\phi\\ \cos\phi\end{bmatrix} \\ \vec{E}_2^{(2)} &= G_1\vec{E}_2^{(1)} = A\cos\theta\begin{bmatrix}\cos\phi\\ \sin\phi\end{bmatrix}\end{aligned}. \tag{7}$$

In order to get a linear response with the Faraday rotation angle ϕ, a $\lambda / 2$ waveplate is used. When the angle between the fast axis of the waveplate and x axis is α, the transmission matrix of the $\lambda / 2$ waveplate is:

$$G_2 = \begin{bmatrix} -i\cos 2\alpha & -i\sin 2\alpha \\ -i\sin 2\alpha & i\cos 2\alpha \end{bmatrix}. \tag{8}$$

When the two beams pass through the λ / 2 waveplate, they could be written as:

$$\begin{aligned} \vec{E}_1^{(3)} &= G_2 \vec{E}_1^{(2)} = iA\sin\theta \begin{bmatrix} \sin(\phi - 2\alpha) \\ \cos(\phi - 2\alpha) \end{bmatrix} \\ \vec{E}_2^{(3)} &= G_2 \vec{E}_2^{(2)} = iA\cos\theta \begin{bmatrix} -\cos(\phi - 2\alpha) \\ \sin(\phi - 2\alpha) \end{bmatrix} \end{aligned} \tag{9}$$

The second PBS combines the x component and y component of the two beams respectively, and forms the final output light:

$$\vec{E}_s = \vec{x}\left[iA\sin\theta\sin(\phi - 2\alpha)\right] + \vec{y}\left[iA\cos\theta\sin(\phi - 2\alpha)\right]. \tag{10}$$

Its intensity can be calculated as following:

$$\begin{aligned} P_s &= \left|iA\sin\theta\sin(\phi - 2\alpha)\right|^2 + \left|iA\cos\theta\sin(\phi - 2\alpha)\right|^2 \\ &= A^2\sin^2(\phi - 2\alpha) = P_0\sin^2(\phi - 2\alpha) \\ &= \frac{P_0}{2} - \frac{P_0}{2}\cos\left[2(\phi - 2\alpha)\right] \end{aligned} \tag{11}$$

It is obvious that the parameter P_S is independent to θ, that is, the SOP of the light beam of the input fiber doesn't affect the output signal, and so the SOP fluctuation of the light source and the stochastic SOP disturbance in the input fiber would not cause error in current detection. We call this technique SOP compensation. Optical fiber 2 is just used to transmit light, and doesn't need SOP compensation. This intensity signal is detected by a photodiode and then converted into electrical signal for further processing. In the end, an output independent to the light intensity is attained:

$$S_{out} = \cos\left[2(\phi - 2\alpha)\right]. \tag{12}$$

Set α=22.5°, when $\phi \ll 1$ we have:

$$S_{out} = \sin 2\phi \approx 2\phi \tag{13}$$

In reference to Eq. (3), we see that we can get a signal directly proportional to the electrical current to be measured.

3. EXPERIMENTS AND ANALYSIS

A sensor head based on the above theorem was practically fabricated in the laboratory and tested. The testing system is shown in Fig. 2. A laser diode (LD) works as the light source, and its output light is coupled into a polarizing controller after it transmits through optical fiber 1. The light out of the polarizing controller is coupled into the sensing element with SOP compensation (including the magneto-optical material, two PBSs, and a λ / 2 waveplate, referring to Fig. 1). The current generator is used to generate a 0~600 An AC current which modulated the light passes through the glass. The modulated light emerging from the sensor head is then detected by a PIN detector and converted into an electrical signal, which is then amplified and sent to a microcomputer for further processing.

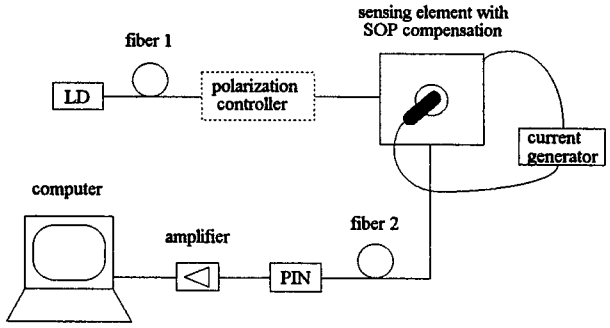


Fig. 2 Schematic diagram of the testing system

3. 1. Coupling loss for the sensor head

In the new configuration, the coupling loss is reduced dramatically. First, the two polarization components of the input light are used which results in a reduction of the 3 dB polarizer loss; then for a coupling from a single-mode fiber(8μm diameter) to a multi-mode fiber (62.5μm diameter), the coupling loss for a coupling distance 120mm is reduced to less then 3 dB(compared to a 16 dB loss for the coupling of multi-mode to multi-mode fibers). The total optical loss for the sensor head Is finally reduced to 6.1 dB. This reduction of the coupling loss brings at least four advantages:

1) the enhanced optical intensity increases the signal-to noise ratio and reduces the difficulty of the receiving electrical circuit;
2) the coupling between the input and output fibers is more stable and makes the noise reduced further;
3) the polarization noise is eliminated;
4) the configuration is adapted to the longer coupling distance used for the ultra-large current.

3. 2. Polarizing compensation

In order to examine the polarizing characteristic of OCT, we adjust the polarizing controller to change the SOP of the incident light, and monitor the measuring result of OCT at the same time. A polarizing sensitivity *PS* is defined to describe the polarization performance of OCT, as the following formula:

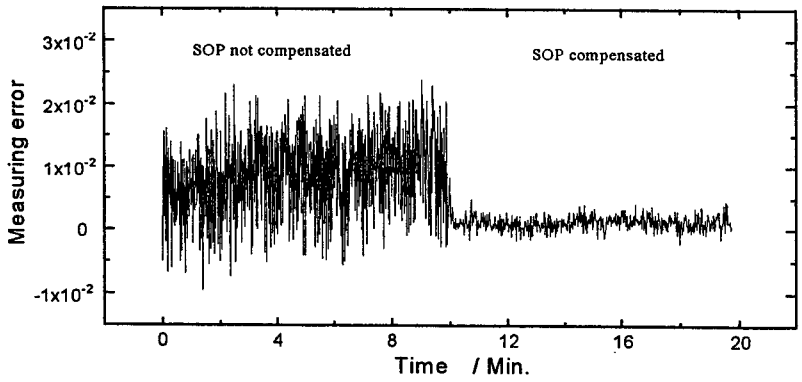


Fig. 3 Improvement in precision by SOP compensation (current:200A)

$$PS = -10\log\left[1 - \left|\frac{P_{max} - P_{min}}{P_{max} + P_{min}}\right|\right], \tag{14}$$

where P_{max} and P_{min} are the maximum and the minimum intensity of the output light signal when the state of the polarizing controller is changed. We got a polarizing sensitivity of only 0.05dB when SOP compensation is adopted. For comparison, we also recorded the measuring performance at the current of 200A when SOP compensation is introduced and not introduced. The result in Fig. 3 shows that the measuring precision is improved about 10 times for the SOP compensated OCT compared with conventional OCT without SOP compensation.

3. 3. Temperature compensation

The variation of the environment temperature has great influence on the performance of OCT system. The main influence focuses on the sensing material. On one hand, temperature variation changes the Verdet constant of the magneto-optic glass; on the other hand, temperature variation also changes the linear birefringence of the magneto-optic material, and hence changes the effective Verdet constant or causes deterioration of the SOP of the light beam. All these will cause fluctuation of the output signal value and decrease in the measuring sensitivity. Furthermore, temperature influence on the operational point of the electrical circuit and the excursion of the light source wavelength also adds up to the measuring error.

Considering the universality and complexity of the temperature influence, we develop a scheme of temperature compensation through combining the hardware and the software. In the optical circuit a fiber-optic temperature sensor (FOTS) is inserted to get the temperature information. For this FOTS, a dual-wavelength detecting technique is adopted, as shown in Fig.4. Two light beams with the wavelength, λ_1 and λ_2, are coupled through a wavelength-division-multiplexer (WDM) into the input fiber. While both of the light beams pass through the temperature-sensitive wafer inserted in the optical circuit of the OCT sensor head, one light beam (with λ_1) is sensitive to the temperature and the another one (with λ_2) insensitive to the temperature. The former is used to detect the temperature and the latter is used as the sensor light of OCT and also the reference light of FOTS at same time. With this configuration the current signal and the temperature signal could be transmitted through the same optical fiber way, that makes the system more compact and reliable.

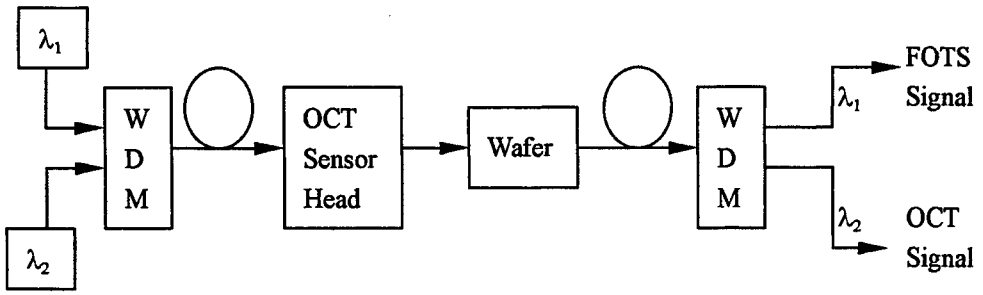


Fig. 4 Scheme of the FOTS for the OCT sensor head

In signal processing, we don't use the calculated current value I_c as the final measuring result, instead, we multiply I_c by a temperature coefficient $\alpha(T, I_c)$ and use the product as the final result. That is:

$$I = \alpha(T, I_c)\, I_c \tag{15}$$

The temperature coefficient $\alpha(T, I_c)$ is relative to two parameters: temperature values T and calculated current value I_c, and determined by experiment data. According to the range of the electrical current to be measured, we can properly set the spacing of the current value (e.g. 1A) and the spacing of the temperature value (e.g. 1°C) for calibration in experiments. And after comprehensive experiments we can get a temperature coefficient matrix $\alpha(T, I_c)$ which is relative to temperature value T and calculated current value I_c. When the matrix is stored in the computer, the software can look up the matrix according to the environment temperature detected by FOTS and the calculated current values responded by OCT, and then work out the final measuring current value.

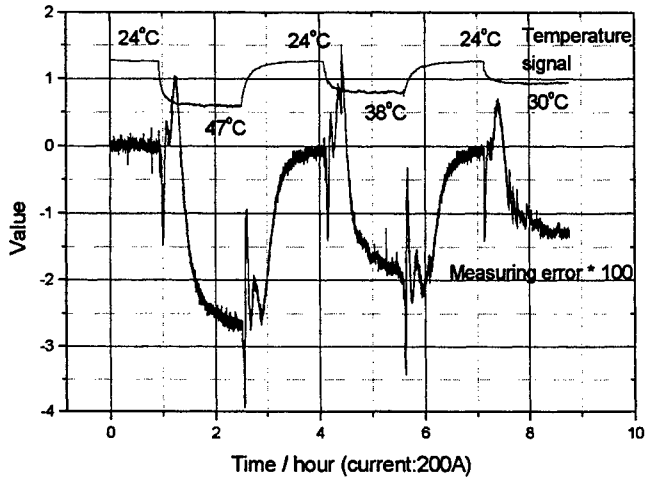


Fig. 5 Temperature performance (current: 200A)

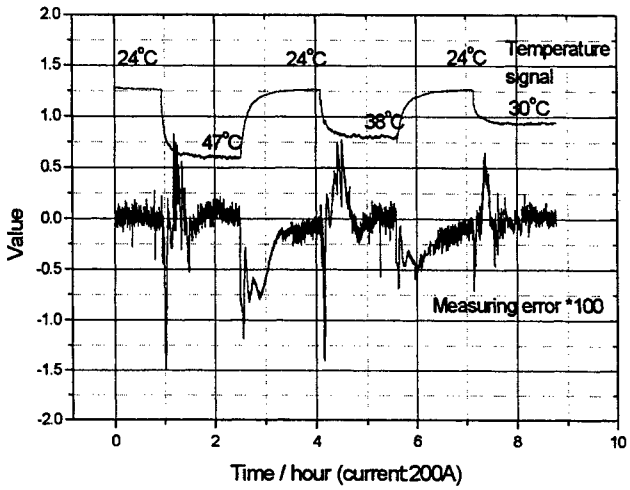


Fig.6 Temperature compensation (current: 200A)

Fig. 5 and Fig. 6 show an improvement of the performance of OCT system with temperature compensation. It is easily found that: (1) When the temperature changes, the measuring error also has considerable undulation. And when the temperature becomes stable, the measuring error also tends to be stable. But they are not synchronized. (2) The measuring error has permanent relation with the temperature value, which can be proved by the constant value of the measuring error at the temperature of 24°C during the whole course of the experiment. (3) The absolute value of the measuring error is bigger when the temperature is higher.

The above results, especially point (2) and point (3) which give out the relation between the measuring error and the temperature value, prove the feasibility of the scheme of temperature compensation through software we discussed above. Based upon the results Fig. 5 shows, we get the temperature coefficients at the temperature of 24°C, 30°C, 38°C and 47°C respectively when the current is 200A. The test result when the temperature compensation is introduced is shown in Fig.6. By comparing Fig. 5 with Fig. 6 we can see that temperature compensation has effectively cancelled out the influence of the temperature variation. Because we just set the temperature coefficients at the certain temperature in this experiment, the measuring error during the changing course of the temperature is not desirable. But it is reasonable to expect that the measuring precision at these temperature values be improved when more temperature coefficients are determined.

3. 4. Measurement of the long-term stability

In the OCT system an AC/DC measuring method [6] is used to eliminate the influence of the light intensity fluctuation (referring to Eq. 11 and 12). Nevertheless experiments show that the influence can't be neglected. To stabilize the light

source power, we introduced optical power feedback and temperature-control circuit. Under this condition we made a measurement on the long-term performance of OCT for 100 hours at the current of 200 A. The result illustrated in Fig. 7 shows that the system has very well long-term stability — the measuring error during 100 hours is confined within 0.3%.

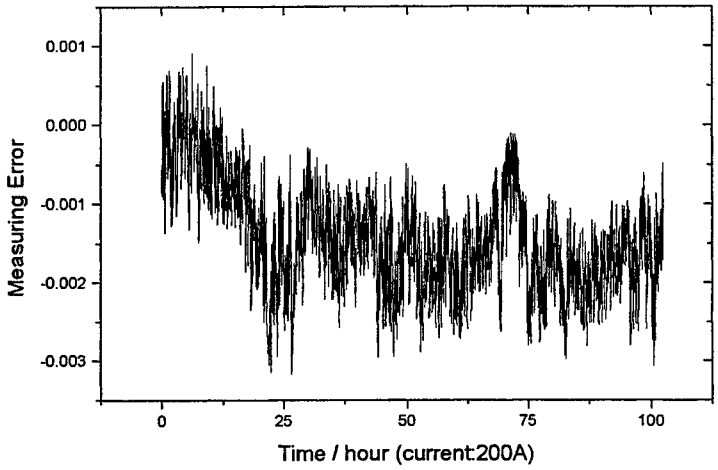


Fig. 7 100-hour performance of the OCT system (current:200A)

4. CONCLUSION

A new type of OCT system with unique structure of the sensing element and perfect data processing has been demonstrated. The new configuration of the sensing element realizes the SOP and temperature compensation, which greatly improve the measuring precision. Using computer and software to accomplish these compensations also proved to be effective. And the whole system is credible for long-term operation.

ACKNOWLEDGMENTS

Authors should be greatly appreciated the funding support from the national foundation for transient-century young talents of Education Ministry of China and from the Henan New Optoelectric Co (China).

REFERENCES

1. O.H. Modi, Kumar V. Jagadeesh, P. Sankaran, "Optical fiber sensor for non-sinusoidal current", *IEEE Journal of Research,* **1**(1), pp.79-83, 1997.
2. G.R. Jones, G.Li.J.W. Spencer, R.A.Aspey, and M.G. Kong, "Faraday current sensing employing chromatic modulation", *Optics Communi.,* **145**(1), pp.203-212, 1998.
3. Y. N. Ning, Z. P. Wang, A. W. Palmer, K. T. Grattan, and D. A. Jackson, "Recent progress in optical current sensing techniques", *Rev. Sci. Instrum.*, **66** (5), pp.3097-3111, 1995.
4. G.A. Woolsey, N.E. Fisher, and D.A. Jackson, "Tuning a bulk-glass optical current sensor by controlling conditions external to its reflecting surfaces", *Sensors and Actuators, A:Physical*,**63**(1), pp.27-32, 1997.
5. N.E. Fisher, and D.A. Jackson, "A common-mode optical noise-rejection scheme for an extrinsic Faraday current sensor", *Mesas. Sci. Technol.*, **7**, pp.796-800, 1996.
6. Emerging Technologies Working Group *et al.*, "Optical Current Transducers for power systems: A review", *IEEE Trans. On Power Delivery*, **9**(4), 1994.

The Compensation for the Effect of Incident Light Polarizing State in OCT

Bi Xijun[a], Chen Haiqing[a], Yan Guoping[b]
[a]Dept. of Optoeletronics Eng., HUST, Wuhan 430074, China
[b]Dept. of Electronic and information Eng.,HUST,Wuhan 430074,China

ABSTRACT

Optical current sensing has manifold advantages over the conventional device. It is of great benefit to develop optical current transducer. Due to the inherent birefringence of optical fiber, fiber-optic current transducer has lower sensitivity than bulk glass optical current transducer Therefore, in research the bulk glass OCT is preferred to fiber optic current transducer.. Because multimode optical fiber has larger core diameter and aperture diameter, the couple coefficient of light source to multimode optical fiber is much larger than that to the single mode optical fiber. So multimode optical fiber is adopted as the medium to carry the optical signal in bulk glass OCT. However, the polarization state of optical signal will change randomly because a multimode optical fiber is not polarization-maintaining. As a result, the intensity of the optical signal will change randomly. The sensitivity of the transducer will be suffered. And it makes the design of electrical equipment difficult. A way to solute this problem is discussed. And an inventive optical current configuration is described. Incident light is divided into two orthogonal polarized components through splitters. After a serial of processing, as dephasing etc, the intensity of output light can be proved to be independent on the polarization state of the incident light. Thus, the random variation of the intensity of the optical signal is eliminated theoretically.

Key words : Faraday effect, optical current transducer(OCT), bulk glass OCT, birefringence, polarized light, state of polarization, fluctuation of polarization state, compensation for the fluctuation of polarization state, optic-fiber, magneto-optic glass

1. INTRODUCTION

The optical current transducer (OCT) is superior to the conventional current transducer (CT) in a lot of perspectives. With the development of the technology of the optical measurement of current, more and more effective OCT types are to be developed to substitute for the conventional CT devices. While it is not the only way to convert a current into an optical effect, the Faraday effect, or magneto-optic effect, is used in most OCT types. According to the different kinds of sensing material used in sensor configuration, OCT can be classified into different types, for example, fiber-optic current transducer, bulk glass optical current transducer. Of these different types, bulk glass OCT that uses bulk glass in the configuration of the sensor head is chosen as the subject of research in most practical research work owing to three characteristics it has. One is that bulk glass has a larger verdet constant. Another is that Faraday effect is strong with bulk glass. And finally, it is not subject to the influence of inherent birefringence and bending induced linear birefringence of the optic sensor head and the birefringence induced by both temperature and stress is relatively minor.

OCT designs may be of different kinds. But all OCT devices invariably use optical elements in the configuration of the sensors. Detection of current is based on the optical features. Thus, optic fiber is always used as the medium to carry the optical signal. However, the optic fiber is not polarization-maintaining, that is it is in the condition of random fluctuation of polarization state. Therefore, the transmission of the light signal along the fiber will be interfered with the random fluctuation of the polarization state. This effect, superposed by the rotation of the polarization plane resulted from the Faraday effect, has greatly reduced the measuring sensitivity.

The present paper is going to discuss the effect that the random fluctuation of the polarization state has on detection and measurement. Furthermore, in an effort to eliminate the influence, an inventive optical device configuration is designed. In this design, compensation is made for the effect of incident light polarization state and has eliminated the influence of random fluctuation in theory to achieve higher accuracy in measurement.

Part of the SPIE Conference on Advanced Photonic Sensors and Applications
Singapore • November–December 1999
SPIE Vol. 3897 • 0277-786X/99/$10.00

2. BASIC THEORY AND CONFIGURATION OF OCT

Faraday effect: when linearly polarized light passes through a magnetic field, the plane of polarization of polarized light will be rotated under the action of the magnetic field parallel to the transmission direction of the light, and the rotation angle θis determined by the following formula:

$$\theta = V \int Bdl, \tag{1}$$

where V is the verdet constant of magneto-optic material, B the magnetic field intensity, l the path length in the glass over which the field interacts with the light. When the light path forms a closed circulate path, the above formula can be changed into:

$$\theta = V \oint Bdl, \tag{2}$$

which through Ampere's circulate law reduces to:

$$\theta = \mu NVI, \tag{3}$$

where N is the number of turns the light makes around the conductor, I the current,μthe permeability of the material. From this formula we could see that the value of current is determined by the measurement of the rotation angle.
The basic configuration of OCT is illustrated by Fig 1:

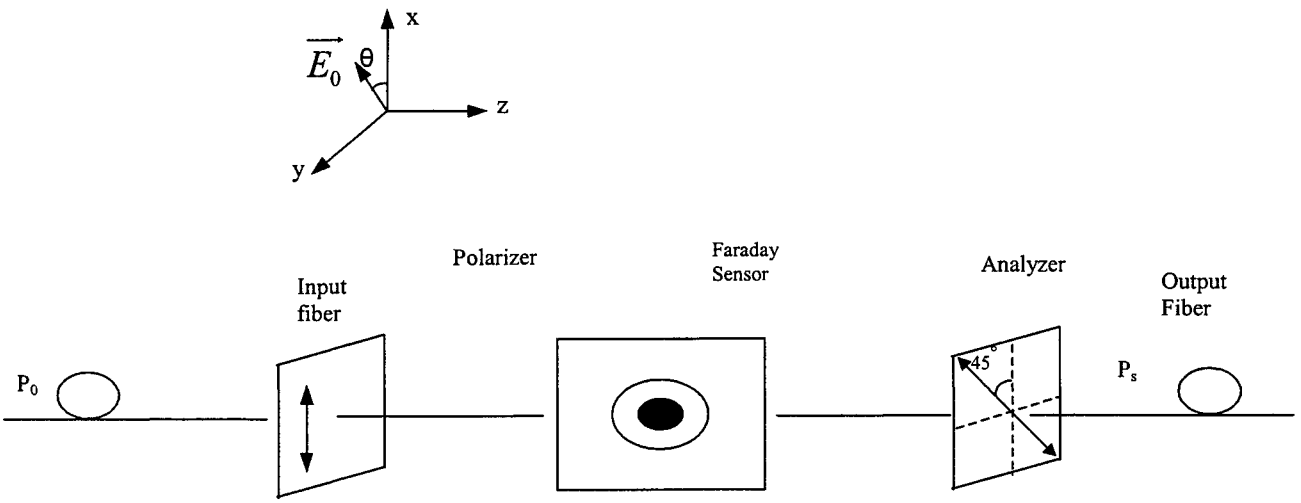


Fig.1. Configuration of conventional OCT

Light P_i sent out from LED is propagated to the sensor head through optic fiber. The incident light will then be polarized by the polarizer. Then the polarized light will be rotated for an angleθwhen it passes through the magneto-optical material by Faraday effect. The rotated light will then pass through an analyzer before it is transmitted to the receiving apparatus through optic fiber. The process can be described by the following equation:

$$P = P_i \cos^2(\varphi - \theta), \tag{4}$$

where,
P_o : input light intensity
ψ : relative angle between polarizer and analyzer
θ: rotation angle in Faraday Sensor

3. THE INFLUENCE OF THE FLUCTUATION OF POLARIZATION STATE

As for the OCT illustrated by Fig. 1, the accuracy of the value of current is dependent on the accuracy of the measured value of the rotation angleθ. However, in practice there exist many factors that can reduce the accuracy in the measurement of the angle. The fluctuation of the polarization state presents a difficult problem for solution. Since optic fiber, used as transmission medium is indispensable in nearly all OCT devices, the influence of the fluctuation of polarization state on the measuring operation is thus inevitable.

On one hand, the fluctuation of the polarization state has great influence on the measuring of the rotation angle. On the other hand, we have to increase the intensity of P_i since the higher the intensity is, the easier the signal light will be received, and the more accurate the detection will be. Correspondingly, LD is adopted as the light source and multimode optic fiber is used as transmission medium in OCT devices to enhance the light intensity. However, though the step index optic fiber with a large core diameter and large aperture diameter is certainly more practical in operation and can gather more light intensity, the influence of fluctuation of the polarization state on the polarized light propagated through the optic fiber path is far more stronger. In conclusion, the fluctuation of the polarization state inherent of the transmission optic fiber constitutes an important aspect of the system detecting error. The following discussion is focused on an analysis of the influence it has on system measuring.

Since any polarized light can be taken as the superposition of linearly polarized light, therefore, for convenience's sake, only the linearly polarized light incident to a polarizer from a certain angle is considered here. Suppose the included angle between the plane of polarization and x axes isθ. Then the linear light can be described by Jones matrix:

$$\overrightarrow{E_o} = \begin{bmatrix} \cos\theta \\ \sin\theta \end{bmatrix}, \tag{5}$$

where A is the amplitude of the optic field, and the input light intensity $P_0=A^2$. Suppose the transmission axes of the polarizer is along the x axes. The polarized output light can be described as:

$$\overrightarrow{E_i} = A\cos\theta \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \tag{6}$$

Due to Faraday effect, the plane of polarization of the linearly polarized light will be rotated to an angle Ψ when it passes the Faraday sensor. The transmission matrix of this process is given by:

$$G_1 = \begin{bmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{bmatrix}, \tag{7}$$

Suppose the transmission axes of the analyzer is set at 45^0 to the x axes. The transmission matrix can be given by:

$$G_2 = \frac{1}{2}\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}, \tag{8}$$

Therefore, the output light passing through the analyzer can be described as:

$$\begin{aligned} \vec{\mathbf{E}} &= G_1 G_2 \vec{E} \\ &= \frac{\mathrm{A}}{2}\cos\theta \begin{bmatrix} \cos\phi + \sin\phi \\ \cos\phi + \sin\phi \end{bmatrix}, \end{aligned} \tag{9}$$

And the light intensity is :

$$P_s = \left|\frac{A}{2}\cos\theta\right|^2 \left|(\cos\phi+\sin\phi)^2 + (\cos\phi+\sin\phi)^2\right|$$
$$= \frac{A^2}{2}\cos^2(1+\sin 2\phi) \quad , \qquad (10)$$
$$= \frac{P_0}{2}\cos^2\theta + \frac{P_o}{2}\cos^2\theta\sin 2\phi$$

This equation shows that signals relative to the Faraday rotation angle Φ can not be distinguished after filtration if θ is a random variant (corresponding to the random fluctuation of the polarization plane of the input light). Thus the current intensity can not be effectively detected.

4. COMPENSATION FOR THE EFFECT OF POLARIZING STATE FLUCTUATION

How to eliminate the influence of this random fluctuation. Solution to this problem is the key to more accurate measurement in OCT system. Though complicated depolarization measures can be taken to reduce the influence to a certain degree, the fluctuation can not be reduced completely and influence on the measuring system can not be ignored. In stead of depolarization measures, an inventive optical design is advanced in this paper in an effort to solve the problem effectively. The optical configuration of the design is illustrated below:

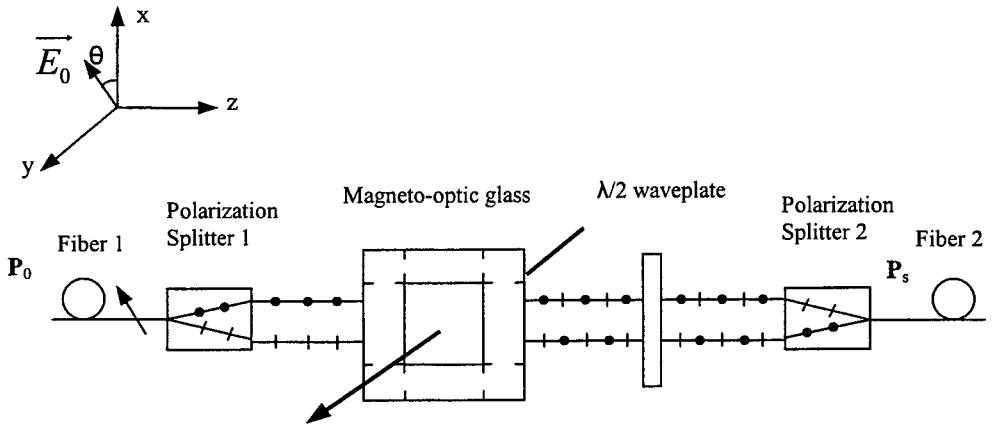


Fig.2. A new optical sensor design and compensation made for the effect of polarizing state

Compared to the conventional OCT devices, two polarizing splitters are used to substitute for the polarizer and analyzer in the new device. In addition, aλ/2 wave plate is placed before the second splitter. The transmission process consists of the following steps: first, the input light lead in through optic fiber 1is divided into two beams of orthogonal linearly polarized light when passing through the first polarizing splitter; then, the plane of polarization is rotated when the linearly polarized light passes through the magneto-optical glass sensor head. Finally, after proper phase calibration by the λ/2 wave plate, the two beams of linear light is synthesized into one beam by the second polarizing splitter and led out through optic fiber 2 for detection.

Still take the linearly polarized light described by equation (5) as the subject for the following discussion. The beam of light is decomposed into two beams of orthogonal light by the first polarizing splitter, which can be described respectively as:

$$\overrightarrow{E_1}^{(1)} = A\sin\theta\begin{bmatrix}0\\1\end{bmatrix}, \tag{11}$$

$$\overrightarrow{E_1}^{(2)} = A\cos\theta\begin{bmatrix}0\\1\end{bmatrix}, \tag{12}$$

The polarization state of the two beams of light after passing through the magnetic glass and then the λ/2 wave plate can be described respectively by $E_{1,2}^{(1)}$ and $E_{1,2}^{(2)}$. The transmission matrix of the magneto-optic glass is given by G_1 (see equation (7)). Then the output light can be described as:

$$\overrightarrow{E_1}^{(2)} = G_1\overrightarrow{E_1}^{(1)} = A\sin\theta\begin{bmatrix}-\sin\phi\\\cos\phi\end{bmatrix}, \tag{13}$$

$$\overrightarrow{E_2}^{(2)} = G_1\overrightarrow{E_2}^{(1)} = A\cos\theta\begin{bmatrix}\cos\phi\\\sin\phi\end{bmatrix}, \tag{14}$$

Suppose the faster axes of λ/2 wave plate is set atαto x axes. Then the transmission matrix of the wave plate is:

$$G_2 = \begin{bmatrix}-i\cos 2\alpha & -i\sin 2\alpha\\-i\sin 2\alpha & i\cos 2\alpha\end{bmatrix}, \tag{15}$$

And the state of polarization of the two light beams after passing through theλ/2 wave plate is described as:

$$\overrightarrow{E_1}^{(3)} = G_2\overrightarrow{E_1}^{(2)} = iA\sin\theta\begin{bmatrix}\sin(\phi-2\alpha)\\\cos(\phi-2\alpha)\end{bmatrix}, \tag{16}$$

$$\overrightarrow{E_2}^{(3)} = G_2\overrightarrow{E_2}^{(2)} = iA\cos\theta\begin{bmatrix}-\cos(\phi-2\alpha)\\\sin(\phi-2\alpha)\end{bmatrix}, \tag{17}$$

Finally, the light beams pass through the second polarizing splitter and the x component and the y component are synthesized into one beam, the output signal light, which can be given by:

$$\overrightarrow{E_s} = \vec{x}[iA\sin\theta\sin(\phi-2\alpha)] + \vec{y}[iA\cos\theta\sin(\phi-2\alpha)], \tag{18}$$

And the light intensity is:

$$\begin{aligned}P_s &= |iA\sin\theta\sin(\phi-2\alpha)|^2 + |iA\cos\theta\sin(\phi-2\alpha)|^2\\&= \mathrm{A}^2\sin^2(\phi-2\alpha)\\&= \mathrm{P_o}\sin^2(\phi-2\alpha)\\&= \frac{\mathrm{P_o}}{2} - \frac{P_o}{2}\cos[2(\phi-2\alpha)]\end{aligned}, \tag{19}$$

Equation (19) shows that the intensity of the output signal light is independent on rotation angleθ. That is that the influence of the fluctuation of the polarization state of the input light in optic fiber 1 is eliminated. After the input light has gone through a serial of processes, the output signal turns out to be independent on the input light intensity P_i :

$$S_{out} = \cos[2(\phi-2\alpha)], \tag{20}$$

Suppose $\alpha=22.5^0$. When $\Phi<<1$, we can get:

$$S_{out} = \sin 2\phi = 2\phi \,, \tag{21}$$

The above equation shows that angle Φ is in linear relationship to the current to be tested. Thus the linear detection of the system is realized.

5.EXPERIMENTAL RESULTS

The configuration of the testing system adopted in the experiment is illustrated by Fig.3. LD sends out the source light signal. The source light is then coupled into polarization controller through optic fiber 1. The linearly polarized incident light from the polarization controller then enters the sensor unit (including the magneto-optic glass, two polarizing splitters, and a$\lambda/2$ wave plate, as illustrated in Fig.2, where compensation for the effect of polarization state is made. The output signal from the sensor unit is coupled into PIN detector through optic fiber 2. The detected signal then gets amplified and input into computer for processing and analysis. The major system function parameters to be tested include polarizing sensitivity and detecting sensitivity of the system.

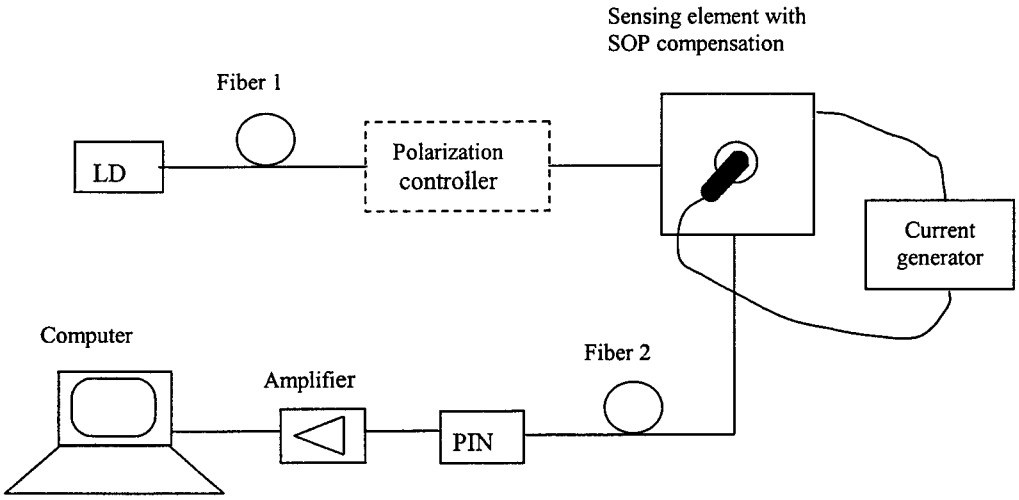


Fig.3. Experimental device configuration

The polarizing sensitivity of the system is defined by:

$$PS = -\log\left[1 - \left|\frac{P_{\max} - P_{\min}}{P_{\max} + P_{\min}}\right|\right], \tag{22}$$

The polarization state of the linearly polarized light incident to the sensor unit can be changed by adjusting the polarization controller. And the light intensity of the light in the sensor unit will alter correspondingly. P_{max} stands for the maximum value of output light intensity while P_{min} stands for the minimum value of the intensity. The result gained from the experiment illustrates that the polarization sensitivity is reduced to 0.15db with compensation made for the state of polarization.
In the testing of the detecting sensitivity of the system, device illustrated by Fig. 3 is adopted except that one component, the polarization controller should be removed (the part shown by the dotted lines).

The detecting sensitivity can be defined as:

$$DS = \frac{I_d - I_o}{I_o}, \tag{23}$$

where, I_d is detected current value and I_0 is the actual current value.

Ordinary bulk OCT systems can only resolve to polarization control within the sensor head to reduce the variation induced by inner reflection. However, the practicality of this method is compromised since the fluctuation of the polarization state of the transmission optic fiber will result in the variation of the output signal. Compared to these ordinary systems, the inventive design advanced in this paper has several advantages. In this system, the quality of output signal is greatly improved and the detecting sensitivity is increased for several times. Besides, the stability of the system is also improved. The comparative results are illustrated below:

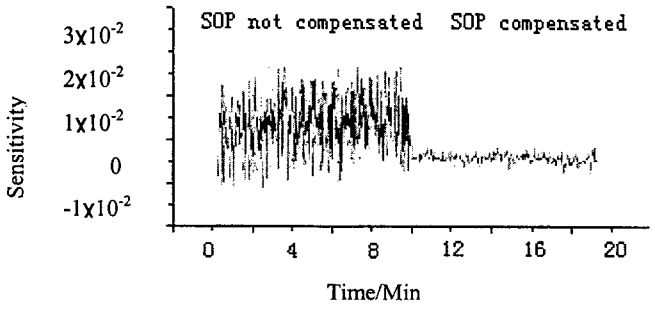


Fig. 4. Experimental results when compensation is made for the effect of polarizing state fluctuation

6. CONCLUSION

In conclusions, this paper has given a comprehensive description of the influence of the fluctuation of the polarization state of the incident light in conventional OCT systems. An inventive optical configuration design is advanced to compensate for this effect. Theoretical analysis is made to support the new design. And experiments have proved the effectiveness of this new configuration in reducing the influence of the fluctuation of the polarization state of the transmission optic fiber and that the new device is much more improved in OCT functions.

ACKNOWLEDGMENTS

Owing thanks to these follow workers who are working in the same lab with the author and have aided in accomplishing this paper. They are: Yanghai Pi, Zuwang Zhou. And especial appreciation is owed to the director of the lab, Dexiou Huang, and follower worker Jingyi He.

REFERENCES

1. R.H. Stolen and E.H. Turner, "Faraday rotation in highly birefringent fibers", Appl . Opt., 19, pp. 842-845, 15 March 1980.
2. M. Kanoi, G. Takahashi, T. Sato, M. Higaki, E. Mori, and K. Okumura, "Optical voltage and current measuring system for electric power systems", IEEE Trans. on Pwr. Del., vol. PWRD-1, pp. 91-97, Jan. 1986.
3. P.R. Forman and F. C. Jahoda, "Linear birefringence effects on fiber-optic currene sensors", App. Opt., vol. 27, pp. 3o88-3o96, 1988.
4. E.A. Ulmer, Jr, "High accuracy Faraday rotation measurements" , OSA/IEEE 1988 Technical Digest of Optical Fiber Sensors Topical Meeting, January 27-29, 1988, New Orleans, LA, pp. 288-291.
5. N.F. Borelli, "Faraday rotation in glasses," J. of Chem. Phys., vol. 41, pp 3289-3293, 1 Dec. 1964.

6. A.A. Jaecklin and M. Lietz, "Elimination of disturbing birefringence effects on Faraday rotation:, Applies Optics, vol. 11, 617-621, March 1972.
7. S.C. Rashleigh and R. Ulrich, "Magneto-optic current sensing with birefringent fibers", Appl. Phys. Lett., vol. 34, pp. 768-770, June 1979.
8. Y.yamagata,T,Oshi,H.Sakurai, "Development of Optical Current Transformers and Application to Fault Location Systems for Substations," IEEE Trans.Power Deliverry,Vol.8,No3,1993
9. T.Mitsui,K.hosoe,H.Usami,S.miyamoto, " Dedelopment of Fiber-Optic Voltage Sensors and Magnetic Field Sensors,"IEEE?PES 1986 Summer Meeting,"86SM442-8,1986
10. A.Papp,H.Harms, "Applied Optics,"Vol.19,No.22,3729,1980
11. Y.W.Bibby,D.C.Larnson,Tyagi,ang L.C.Bobb, "Proceedings on the 8th Optical Fiber Sensors Conference,"1992,p.161
12. H.katsukawa,H.Ishikawa,H.Okajima,andT.W.cease, "IEEE Transactions on Power Delivery,"Vol.11,No2,702,1996
13. W.J.Tabor and F.S.chen, "Electromagnetic propadation through materials possessing both Faraday rotation and birefringence :experiments with Ytterbium Orthoferrite,"J.Appl.Phys,40.pp2760-2765,1969
14. D.C.Erickson, "The use of fiber optics for communications,measurement,and control within highvoltage substations,"IEEETrans.on Pwr.Appr.andSyst.,vol.PAS-99,pp.1057-1065,May/June 1980
15. S.Satio,J.Hamasaki,Y.Fujii,K.Yokoyama,and Y.Ohno. "Development of the laser current transformer for exrrahigh-voltage power transmissin lines,"IEEE j.Quant.Elec.,vol QE-3,pp.589-597,Nov.1967.
16. S.C.Rashleigh and R.Ulrich, "Magneto-optic current sensing with birefringent fibers',Appl.Phys.Left.,vol.34,pp.768-770,June 1979.

SESSION 8

Three-Dimensional Inspection

Invited Paper

New Light Sources and Sensors for Active Optical 3D-Inspection

Wolfgang Osten*, Werner Jüptner

BIAS, Bremer Institut für Angewandte Strahltechnik
Klagenfurter Str. 2, D-28359 Bremen, Germany

ABSTRACT

The implementation of active processing strategies in optical 3D-inspection needs the availability of flexible hardware solutions. The system components illumination and sensor/detector are actively involved in the processing chain by a feedback loop that is controlled by the evaluation process. Therefore this article deals with new light sources and sensors which appeared recently on the market and can be applied successfully for the implementation of active processing principles. Some applications where such new components are used to implement an active measurement strategy are presented.

Keywords: active optical metrology, 3D-inspection, light valve technology, CMOS-sensor, optical shape measurement, near field optics

1. INTRODUCTION

'Look but don't touch' is the guideline for the development of new technologies for industrial inspection. In this sense modern techniques of optical metrology such as holographic interferometry, speckle metrology, fringe projection and moiré are important prerequisites for the derivation of informations about the structure of objects and their behaviour under typical operational or environmental conditions. Recent implementations in optical range and displacement sensors are the basis for numerous applications in civil engineering, industrial inspection, traffic control, remote sensing and medicine[1]. The advantages of these techniques are obvious: fast and noncontact interaction with the object under investigation, fieldwise working mode, adaptive sensitivity and last but not least the human-like representation of the results as images. However, the information content of the generated images is very high but in general not sufficient yet to permit the reconstruction of reliable data from one single shot. Consequently the current trend in image evaluation is to consider more complex situations like image sequences where the information content of every single frame is systematically changed by the manipulation of selected parameters which are relevant for the image formation process.

Following the current trend in image analysis more flexibility in the analysis strategy is obtained by combining the classical data driven bottom-up strategy, Fig. 1a, with the so-called expectation driven top-down strategy[2], Fig. 1b. The first strategy has proved as very efficient but extra effort must be paid to obtain a high image quality and in most cases a-priori knowledge has to be added by operator interaction. The image formation process is considered as a rather fixed/passive data source and is not actively involved in the evaluation process. In contrast to this the second strategy includes the image formation as an active component in the evaluation process. Dependent on the complexity of the problem and the state of evaluation new data sets are actively produced by driving a feedback loop between the system components which are responsible for data generation (light sources, sensors and actuators) and those which are responsible for data processing and analysis (computers). Support from other sensors at different positions, recordings from different time instances or the exploitation of different physical sensor principles, i.e multisensory data, are considered in that concept. The strategy which supervises the image analysis is connected to and controlled by the information gathered by sensors and by a knowledge base.

For this new way of thinking with respect to the solution of vision tasks a paradigm has already arisen, known under the various names active/purposive/animate, etc. ..., vision[3]. In the following we use the term *active vision* because it describes the new quality best. There is an important difference between the active and the classical passive approach. When we work in passive mode, a set of images is given which have to be processed with the algorithms we are going to develop. On the other hand, when we work in active mode we do not want prerecorded data since we include the image acquisition as an component of equal importance into the complete evaluation process. Active vision offers a new approach for building intelligent and more flexible systems especially for optical metrology[4]. Such an active system is not only able as just to „see". It is designed to do something, i.e. to make an action, which is anything that changes the state of the system and/or the environment. Similar

* Correspondence: wolfgang@uni-bremen.de

Part of the SPIE Conference on Advanced Photonic Sensors and Applications
Singapore • November–December 1999
SPIE Vol. 3897 • 0277-786X/99/$10.00

to human perception as an active way to explore a natural scene by changing the view or the focus, the role of the observer and observation system, respectively, is defined as an active component for gathering data. Some examples for active approaches will be discussed more detailed later.

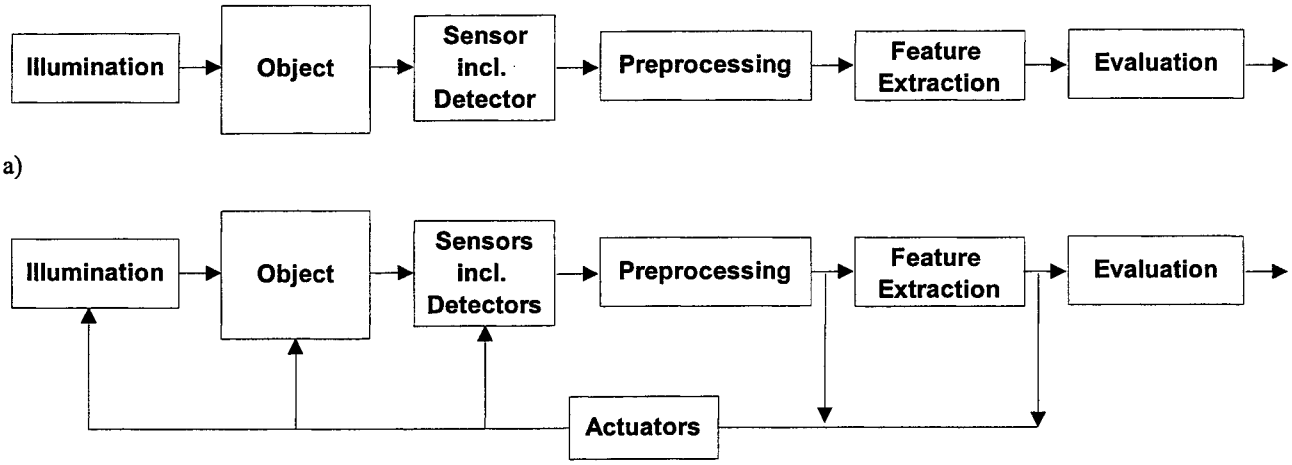


Fig. 1: Different strategies for image processing: a) classical data driven bottom-up strategy, b) combination with the expectation driven top-down strategy

In computer vision and optical metrology sometimes the difference between a so called active and passive technique is only made with respect to the way of working with the light, i.e. illumination. For instance in shape measurement the technique shape from shading[5] is classified as an passive technique whereas structured illumination or fringe projection[6] is called as active[7]. The difference between both approaches is that in the first case the shape is recovered from the brightness at each point of an image usually illuminated by one light source with the same lighting throughout the surface and in the second case the structure of the lighting is modulated systematically throughout the surface of the object. In our understanding the other way of object illumination by using structured light delivers not more as the possibility to realize the triangulation principle in a fieldwise manner by addressing each image point with a mod2π phase value. But a lot of problems still remain, especially with respect to the ambiguity of the phase reconstruction[8]. Here the application of more sophisticated procedures such as absolute phase measurement[9] or wavelength scanning[10] is necessary with respect to a unique solution. These methods bring a certain amount of activity into the system since the role of the observer is active by capturing systematically more information as a single view can bring. However, active vision/metrology is not just the use of multiple frames. The modern understanding of being active means that the image acquisition process is controlled and thus constraints that facilitate the recovery of information about the state of the object under test (3D-shape, 3D-displacements, subsurface defects, ...) are introduced. In practice such approaches use an active controlled feedback loop, Fig. 1b, between the components that capture the image data and those that evaluate them.

However, the implementation of active processing strategies needs the availability of flexible hardware solutions. As shown in Fig. 1a, the system components illumination and sensor/detector are actively involved in the processing chain by a feedback loop that is controlled by the evaluation process. Therefore this article deals with new light sources and sensors which appeared recently on the market and can be applied successfully for the implementation of active processing principles. In the next chapter the main technical components for the active manipulation of image properties in optical metrology are discussed and some examples of new technological solutions are given. Finally some measurement problems with an active solution strategy are discussed.

2. COMPONENTS OF AN OPTICAL INSPECTION SYSTEM

The sinusoidal nature of light, its ability to interfere and the precise knowledge of its wavelength make it possible to measure several metrological quantities such as distances and displacements with high accuracy. However, the high frequency of light results in the difficulty that the primary quantity to be measured, i.e. the phase of the light wave, cannot be observed directly. All quadratic sensors (CCD, CMOS, photo plates) are only able to record intensities. Therefore one makes use of a trick: the transformation of phase changes into observable intensity values as the basic principle of interferometry and holography. But this apparent advantage must be paid expensively since the phase is coded now in sinusoidal intensity fluctuations, the so called interference fringes. Consequently the task to be solved in fringe analysis can be defined as the conversion of the fringe pattern into a continuous phase map taking into account the quasi sinusoidal nature of the intensity distribution.

The output information obtained from the 2D intensity sensor (usually a CCD camera) in the optical measuring setup has its general form given by the intensity distribution:

$$I_m(x,y;t) = \sum_{n=0}^{\infty} a_n(x,y) * \cos[n2\pi(f_{ox}x + f_{oy}y + \nu_o t) + n\alpha_m(t) + \Phi_n(x,y)] \quad \text{with } m = 1 \ldots M \qquad (1)$$

where $a_n(x,y)$ and $\Phi_n(x,y)$ are the amplitude and the phase of the nth-order harmonic of the signal, f_{ox}, f_{oy} are the fundamental spatial frequencies, ν_o is the fundamental temporal frequency, α is the phase shift parameter and m is an index to distinguish between the different phase shifted images. The measurand is coded in the phase of the first harmonic of the signal ($\Phi_1(x,y)$) and is always screened by the background given by $a_0(x,y)$ and the modulation which is represented by $a_1(x,y)$. The same is valid for incoherent fringe projection using periodic intensity distributions which are generated with special transparencies or electronic light valves.

The measured intensity distribution can be systematically influenced by several components of the set-up. Four main components can be distinguished, Fig. 2: the illumination components (light source), the components responsible for the acquisition of the light fields (detector), the components for the manipulation of the light fields such as phase shifting elements or light modulators (actuator) and the components for the processing of the images and the presentation of the results (processors). Table 1 gives a survey about some principles and the relevant parameters which can be controlled[11].

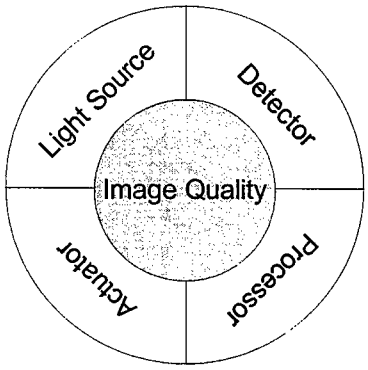


Fig. 2: Main components to tune/manipulate the quality and the properties of the acquired fringe patterns

Several new light sources, detectors and actuators are on the market now which allow to control actively the illumination of the object and acquisition of the light fields after interaction with the object. Some examples are described briefly in this papragraph.

2.1 VERTICAL CAVITY SURFACE-EMITTING LASER (VCSEL)

Semiconductor diode lasers play an growing role in optical metrology. Some of their advantages in comparison to conventional laser sources are their very small size, low power consumption, wavelength tunability and low costs. The overwhelming majority of commercial diode lasers are horizontal-cavity lasers with an optical output beam emitting from the edge of the laser chip. They are often referred to as edge-emitting lasers (EELs). The vertical cavity surface-emitting laser (VCSEL) represents a new family of semiconductor devices with a wide range of applications that have only just started. A VCSEL's cavity is perpendicular to the wafer plane, Fig. 3, and thus the optical beam is guided in the vertical direction. The

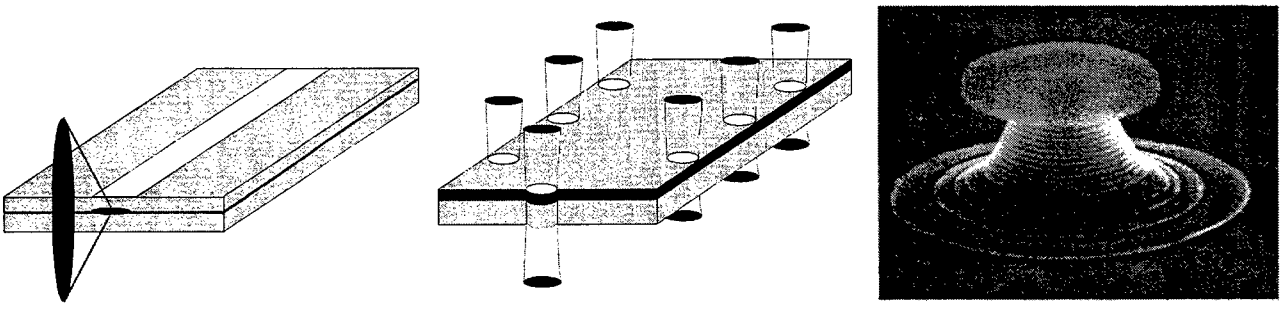
output beam has a circular rather than the elliptical profile of an EEL. The radically symmetric Gaussian far field means that the device is easier to couple to fibre or optical elements than the edge emitter. Further advantages are:

- single, longitudinal mode optical output,
- very small size,
- low power consumption,
- wavelength tunability free from mode hopping,
- chance for fabrication in two-dimensional arrays.

Component	Principle	Example	Controllable	Parameter
Light Source	Coherent low coherent	gas laser solid state laser diode laser (VCSEL) white light source	active material current temperature	output power wavelength degree of coherence
Detector	electronic (solid state) photo-chemical photo-refractive photo-biological	CCD, CMOS silver halide $LiNbO_3$ bacterio rhodopsin	material amplification temperature	sensitivity (wavelength, intensity)
Actuator	electro-mechanical electro-optical acusto-optical	PZT, DMD LCD, OASLM, EASLM, fiber optics	voltage light intensity temperature length	position transparency phase fringe profile
Processor	Electronic Optical Hybrid	frame grabber 4f-processor software		phase difference displacement coordinates patterns

Table 1: Components for the generation, acquisition, manipulation and processing of light fields

The VCSEL presents a real challenge not only for telecom but also for metrology. It turns out that the VCSEL is ideal for wavelength engineering. It has an ultrashort cavity length - more than two orders of magnitude shorter than an EEL - and thus typically has only one Fabry-Perot (longitudinal) mode (within the laser gain bandwidth), which determines the lasing wavelength[12]. Thus, by varying the cavity length slightly, the lasing wavelength can be varied accordingly. This presents an excellent an unique opportunity for engineering multivavelength laser arrays and wavelength-tunable lasers with high application potential in optical metrolog too. There are different approaches to vary the VCSEL cavity length. One solution is to create a structure where the VCSEL top mirror may be translated via a micromechanical structure[13].

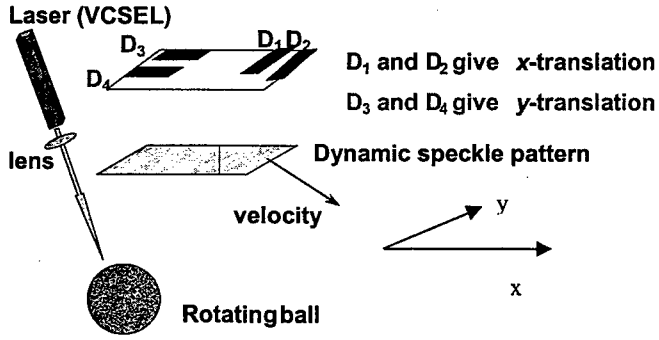


a) b) c)

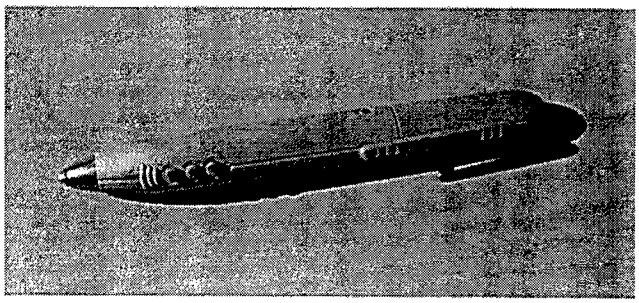
Fig. 3: Schematics of an EEL (a) and VCSEL array (b); Scanning electron micrograph of a VCSEL[12]

An interesting application of VCSEL for the detecting of rotations of a reflective ball by speckle interferometry was recently published[14]. The primary goal for the set-up shown in Fig. 4a is to realize a simple and compact concept that can be used for PC-cursor control, yet having an adequate precision for cursor control. A slightly diverging beam from a VCSEL mounted on a PC-board with a diameter of 6.8 mm illuminates a reflective ball (Ø 4 mm) placed at a distance of 15 mm. Light scattered off the rotating ball produces a speckled field, which sweeps across two pairs of elongated photo detectors arranged

perpendicular to each other. The four photo detector signals are digitized and fed into an ASIC, which determines the rotation of the ball in two directions including the sign of the rotation.



a) Principle of the optical set-up



b) The complete pen unit with the ball in the pen tip

Fig. 4: Application of VCSEL for the detection of rotations of a reflective ball by speckle interferometry[14]

2.2 COMPLEMENTARY-METAL-OXIDE-SEMICONDUCTOR (CMOS) IMAGE SENSORS

Smart image sensors based on complementary metal-oxide semiconductor (CMOS) technology are opening a broad range of applications that charge-coupled-device (CDD) technology cannot fulfil. In contrast to CCD a CMOS-Sensor consists of light sensitive diodes and enables a permanent conversion of the photo current into an output voltage. Consequently a CMOS-sensor can be regarded as a matrix of light meters. This new sensor technology brings several advantages to imaging applications:

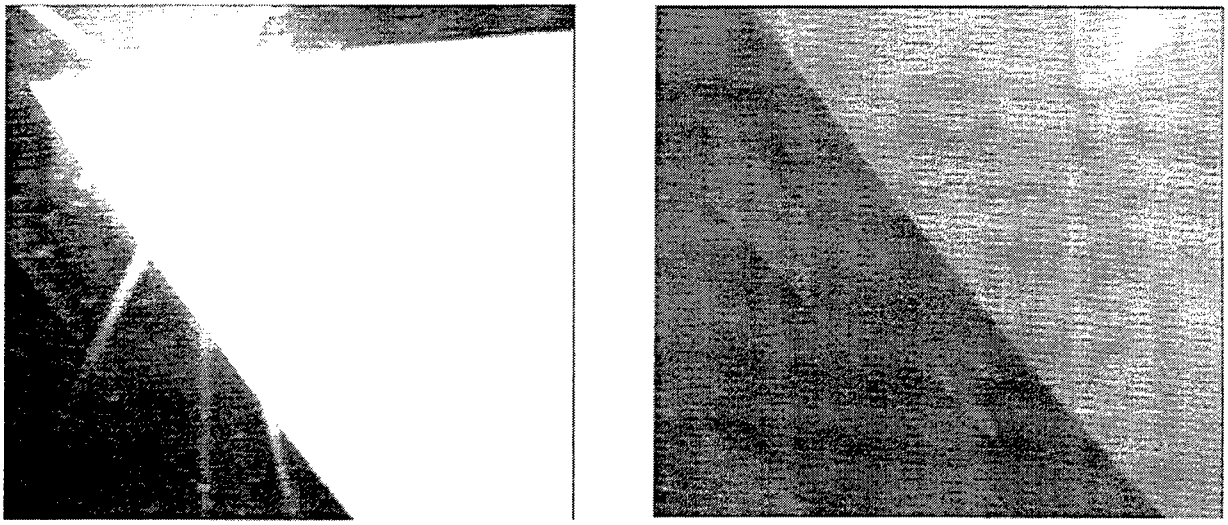
- ability to fabricate image sensors and processing technology onto a single chip (camera on a chip architecture),
- additional pixel functionality possible (active pixel sensor (APS) technology),
- flexible readout device, i.e. random pixel access,
- high dynamics of 120dB ... 140dB , logarithmic characteristics (in comparison: CCD 50dB ... 70dB, human eye about 200 dB),
- marginal blooming,
- low power consumption.

CMOS detectors generally provide fast access and better accuracy for measurements as a result of their direct access readout capability. The high dynamics delivers an improved image quality for such environments where strongly varying intensities and high contrasts, respectively, occur, Fig. 5. Unlike interline-transfer CCDs that need careful shielding in the pixel to reduce smear and have concomitant low fill factor, CMOS APS devices are immune to smear and have much larger effective fill factors - typically 30%-40%[15]. For instance FASTCOM Techn. offers the MVS-135, a complete image processing system, which relies on state-of.the-art CMOS image sensors and embedded DSP technology. The camera integrates in a single casing all the hardware and software needed to capture and process visual information in real time. The sensor is a CMOS image sensor with 512x512 square pixels, has a dynamics of 12odB and allows for random pixel access. The last allows in real time which pixel should ne read. The effective pixel rate of 2.4 MHz makes it possible to read more than 235 frames/second for a window of 100x100 pixels.

2.3 DIGITAL MICROMIRROR DEVICES (DMD)

During the past decade, new technologies in image modulators have made electronic projection systems more comfortable and widely available for a variety of applications. A digital micromirror device (DMD) is an array of microscopic mirrors built over a CMOS SRAM cell, fabricated with semiconductor manufacturing techniques. The DMD chip is approximately 2x1,5 cm. Each mirror, 16µm square, represents one pixel within the displayed image. Commercially available DMD have a resolution up to 1024x768 mirrors, or 786.432 mirror-pixels per chip. Mirrors on the array are independently modulated form the *on-* (or reflect) position to the *off-* (or light-dump) position by applying an address voltage. The amount of time a mirror remains in the on- and off-position determines the brightness of the projected pixel. Consequently grey value images can be projected. For color projection the light enters a prism and is devided and filtered into each of the three color primaries. A three-chip DMD light engine uses three static dichroic filters to separate the primary colors without any time devision

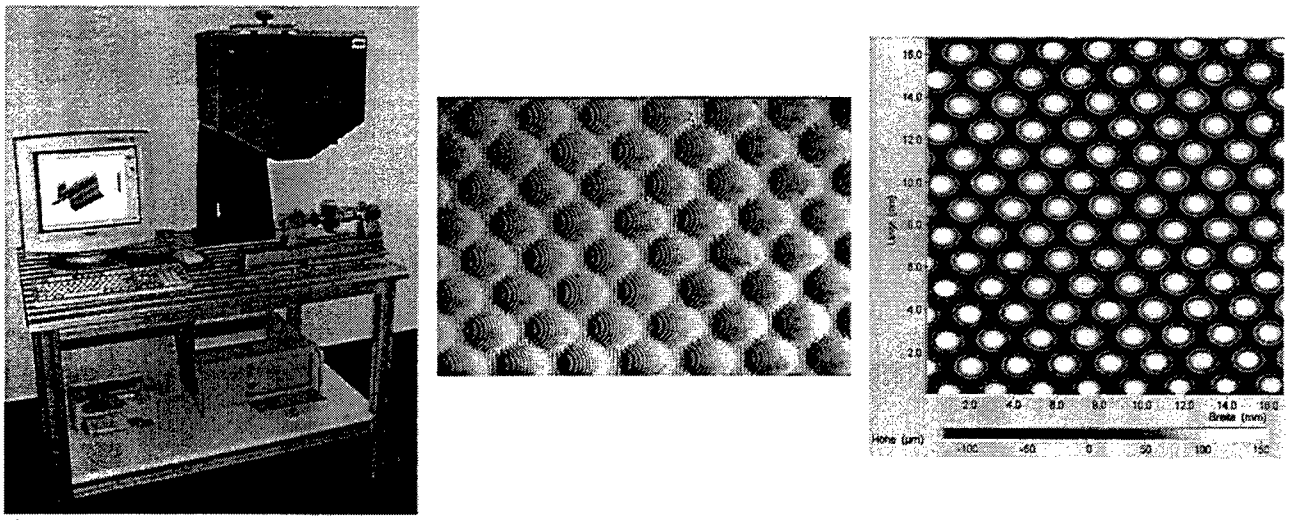
multiplexing. Each DMD chip within the projection system creates a monochrome image for each color primary: red, green, blue. Signal information is presented to the DMD in a 10-bit digital format to achieve full color depth. Each DMD creates up to 1024 shades of grey. The densely arranged mirrors provide an on-screen fill factor of greater than 90% with a high intrinsic reflectivity greater than 90%.



a) b)

Fig. 5: Image captured with CCD-camera (a) and CMOS-camera (b)[16]

The use of DMD for shape measurement using digital fringe projection is a real challenge. Several application and commercialized systems are already known[17,18]. Fig. 6 shows for example the 3D-measurement system MicroCAD from GF Meßtechnik[19] and the contour measurement of a plastic surface.



a) b) c)

Fig. 6: DMD Shape-Measurement System MicroCAD from GF Meßtechnik (a), plastic surface illuminated with projected fringes using a DM (b), color encoded height map of the plastic surface (c)[19]

2.4 LIQUID CRYSTAL DISPLAY (LCD) PROJECTION UNITS

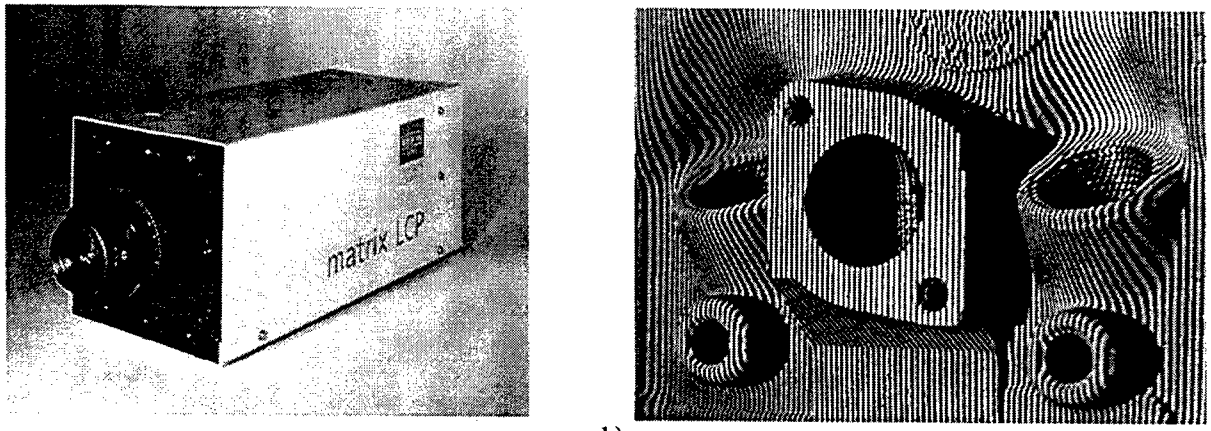
The elongated molecules of liquid crystals LC tend to have ordered orientations that are alterated when the material is subjected to mechanical or electric forces. Because of their anisotropic nature, LC can be arranged to serve as wave retarders

or polarization rotators. In the presence of an electric field, their molecular orientation is modified, so that their effect on polarized light is altered. Therefore LC can be used as electrically controlled optical wave retarders, spatial light modulators, and switches[20].

As active element for fringe projection a free programmable LCD projector "matrix LCP" was developed as a joint venture of BIAS, JENOPTIK and VEW. At the heart of the projector is a modulator with a TFT-LCD having a spatial resolution of 832x624 pixels and specifically designed for applications requiring a flexible dynamic pattern generation. Polycrystalline silicon thin-layer transistors have been selected to provide grids of high resolution. These LCD spatial light modulator (SLM) can be used for configuring intensity distributions of any desired shape. The projector can thus be matched to the particular metrology task, test object and test procedure in a perfect manner. Every individual pixel can be controlled digitally by 256 different grey values. In this way, the projection pattern can be adapted by local variation of the line frequency, the pattern type and the intensity until a best possible evaluation is achieved. An important feature of the panel is also the high contrast (>0,9) that can be observed. Fig. 7 shows the LCD projector and a technical component illuminated with projected straight lines.

In contrast to DMD that have as well a high fill factor as a high intrinsic reflectivity, transmissive LCD projection panels have pathways for control traces between each pixel. This creates a panel that typically uses as much as 50% of its surface area for the creation of images and transmission of light. More than half of the luminance from the light source is restricted or blocked by the control trace pathways, LCD panel, and required polarizers. Therefore recent developments use C-MOS based reflective liquid crystal light valves[21]. In contrast to transmissive LCDs that use horizontal liquid-crystal alignment, reflective liquid-crystal light valves use an active-matrix LCD with a vertical liquid crystal alignment, providing a constant aperture ratio that allows reflective LCD projectors to achieve higher-brightness and higher resolution, while the vertically aligned crystal allows higher contrast and faster response time.

Although LCD are used mainly for intensity modulation, they can be used for phase modulation of the incident light field too. Liquid crystal variable phase shifter for an uniform phase modulation of the incident wave are commercially available[22] and are used e.g. for phase shifting interferometry. However, high-resolved pixelized phase modulators enabeling a controlled and reproducable modulation of the field in the region of $0...2\pi$ are not available yet. Such devices would open a broad variety of metrological applications by active wavefron control.



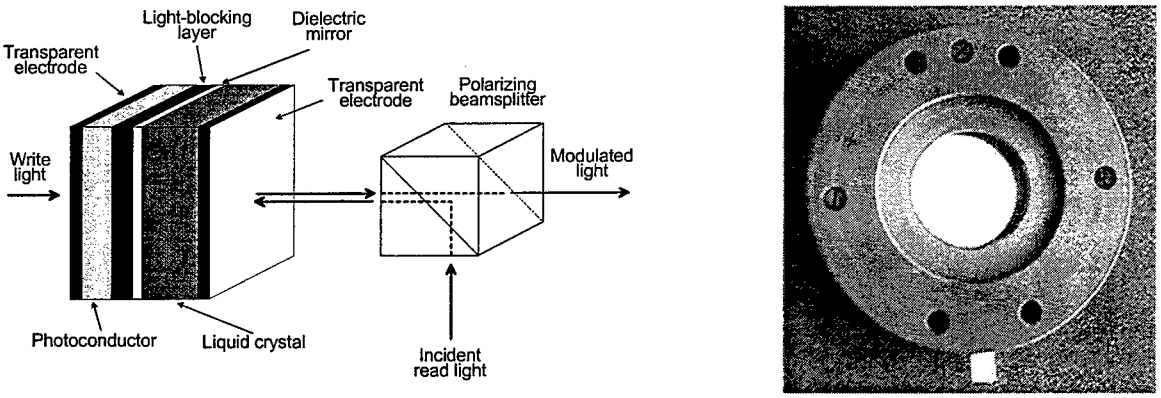
a) b)

Fig. 7: Matrix LCP made by JENOPTIK: a) LCP, b) projected fringes on a technical component

2.5 OPTICALLY ADDRESSED SPATIAL LIGHT MODULATORS (OASLM)

Mosed LCDs are addressed electrically. However, LC based optically addressed spatial light modulators (OASLM) are attractive for applications involving image and optical data processing. Light with an intensity distribution $I_W(x,y)$, the so-called „write" image , is converted by an optoelectronic sensor into a distribution of electric field E(x,y), which controls the reflectance R(x,y) of a LC cell operated in the reflective mode. Another optical wave of uniform intensity is reflected from the device and creates the „read" image $I(x,y) \propto R(x,y)$. Thus the „read" image is controlled by the „write" image. If the „write" image is carried by incoherent light, and the „read" image is formed by coherent light, the device serves as a spatial incoherent-to-coherent light converter. The wavelengths of the „write" and „read" beams need not be the same. A scheme of a

LC-OASLM is shown in Fig. 8a. This device made by Hughes is essentially a capacitor with two low-reflectance transparent electrodes with a number of thin layers of materials between. There are two principal layers: the liquid-crystal, which is responsible for the modulation of the „read"light; and the photoconductor layer, which is responsible for sensing the „write" light distribution and converting it into an electric-field distribution. These two layers are separated by a dielectric mirror, which reflects the „read" light, and a light blocking dielectric material, which prevents the „write" light from reaching the „read" side of the device. The polarizers are placed externally[20].



a) b)

Fig. 8: Liquid-crystal OASLM: a) Principle[20], b) Example produced by JENOPTIK[22]

OASLM are commercially available in different versions. One version with an active area of 30mm is offered by JENOPTIK[22], Fig 8b. The modulator consists of a thin-film sandwich structure with photosemiconductor, liquid crystal, dielectric mirror and transparent electrodes on glass substrates. The voltage applied to the electrodes is devided depending on the illumination intensity between photo-semiconductor and liquid crystal layers, enabling 2D-image modulation of the refractive index of the liquid crystal. The specification considers write light spectral sensitivity between 550nm and 700nm, a read-out wavelength between 450nm and 630nm and a switch time in the region of 20ms ... 50ms.

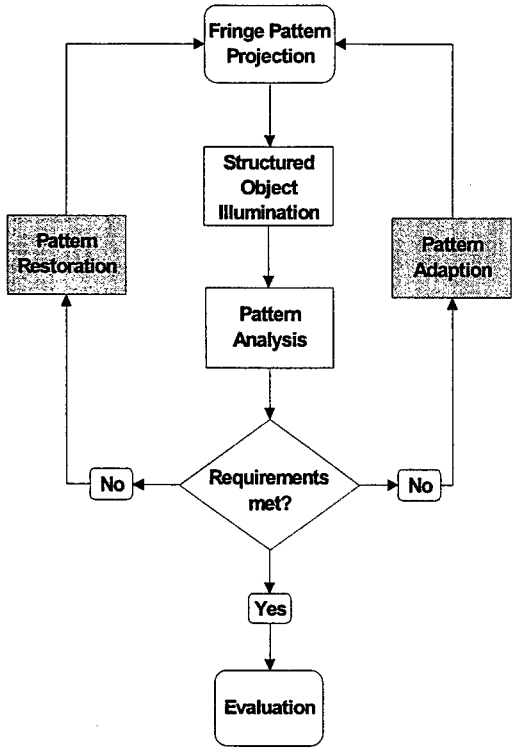
3. APPLICATION OF MODERN OPTICAL COMPONENTS FOR ACTIVE METROLOGY

3.1 INVERSE FRINGE PROJECTION FOR SHAPE CONTROL

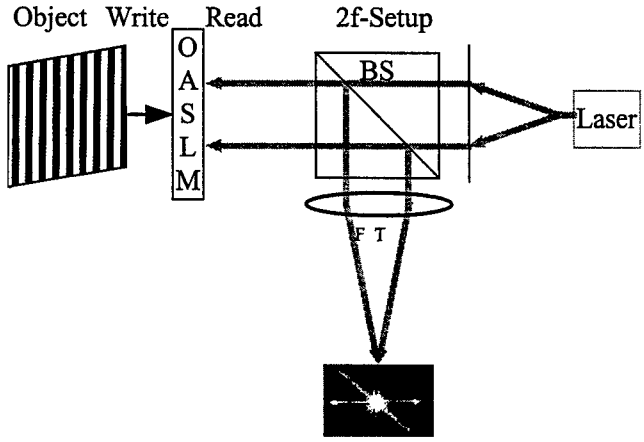
Fringe projection is an established technique for the measurement and inspection of shape of technical components[23]. Several methods are implemented meanwhile which allow not only the precise measurement of coordinates using modern phase measuring techniques but also their transformation into computer systems for the generation of a consistent CAD description of complex 3D-objects[24]. Here a new technology for the shape inspection of technical components is described. This technology is also based on fringe projection but uses an active fringe control for both the improvement of the fringe pattern with respect to saturated and nonresolved fringes[25] and its adaptation to different objects with respect to a considerable shortening of the inspection time. To ensure the projection of as well arbitrary as precise sinusoidal fringe patterns the free programmable LCD projector with digital pixel control as described in chapter 2.4 and the OASLM from JENOPTIK, chapter 2.5, are used.

The objective of pattern restoration is to make the fringe projection technique more robust against relevant disturbances such as too high fringe-frequency or bright spots owing to specular reflections. The procedure is based on a result dependent projection in a closed-loop control system, Fig. 9a. Similar to holographic nondestructive testing (HNDT) the observed fringe patterns offers the possibility for a qualitative inspection of the shape of the object. This can be done on the basis of a comparison between that pattern which is produced by the sample and those produced by a master piece. To make this comparison more effective the same technology is used in an inverse way to deliver object adapted projection data for the inspection process (pattern adaptation), Fig. 10. This inverse pattern is computed either by using the real master piece and a special developed projection method or by solving the inverse problem with the stored CAD-data and coordinates of the masterpiece, respectively. Having a test object that completely fits in shape with the master piece the inverse projection delivers an equidistant and non-distorted fringe pattern on the observation side, Fig. 10c. Every faulty object area causes distortions relatively to this known model and can be consequently detected very fast by means of suitable correlation techniques. Especially the simplicity of the master pattern (straight and uniform fringes) offers the possibility of very simple

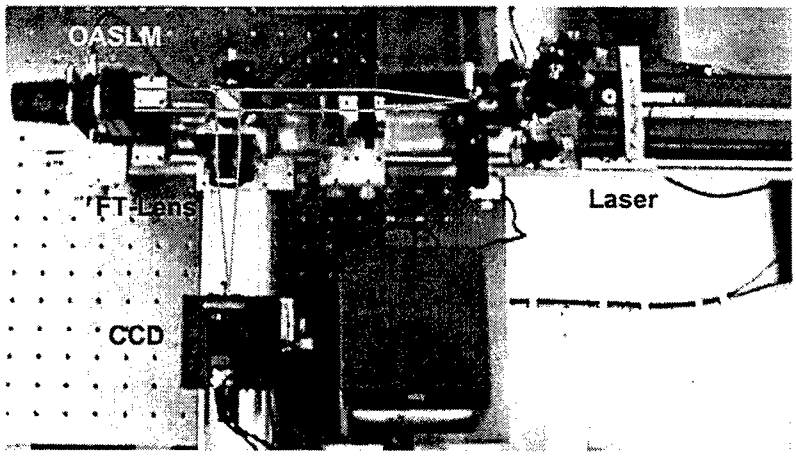
filters and an optical implementation of the correlation process using an OASLM for the spatial incoherent-to-coherent light conversion, Fig. 9b,c.



a) Feedback loop for active pattern restoration and pattern adaption (inverse projection)

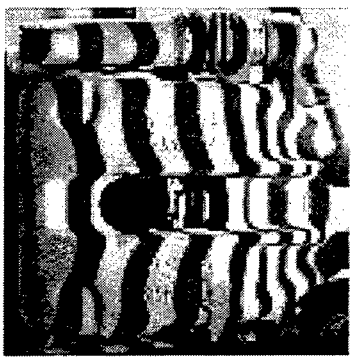


b) Optical correlator using an OASLM

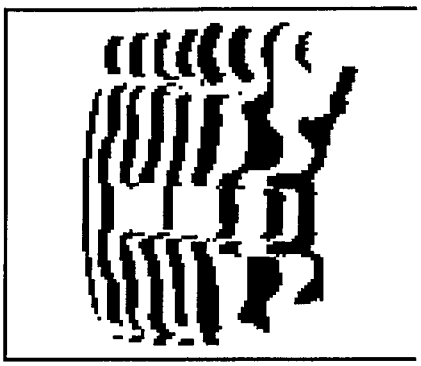


c) Optical setup of the correlation unit

Fig. 9: Shape control with an inverse projected-fringe-technique and an optical correlator using an OASLM for spatial incoherent-to-coherent light conversion



a) Car brake saddle covered with projected straight fringes

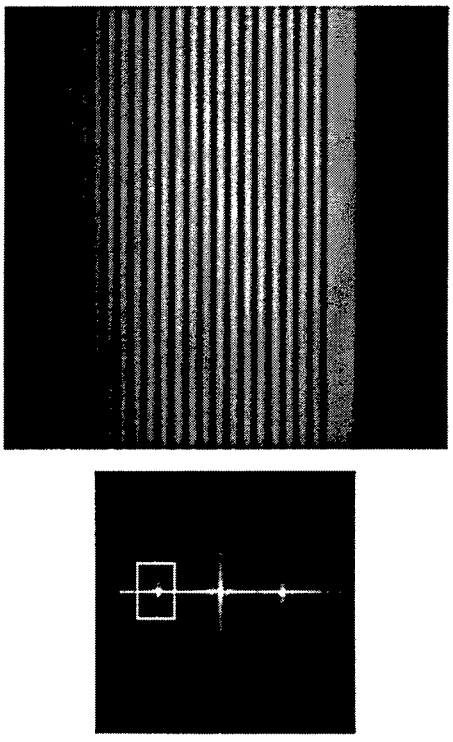


b) Computed inverse fringe pattern

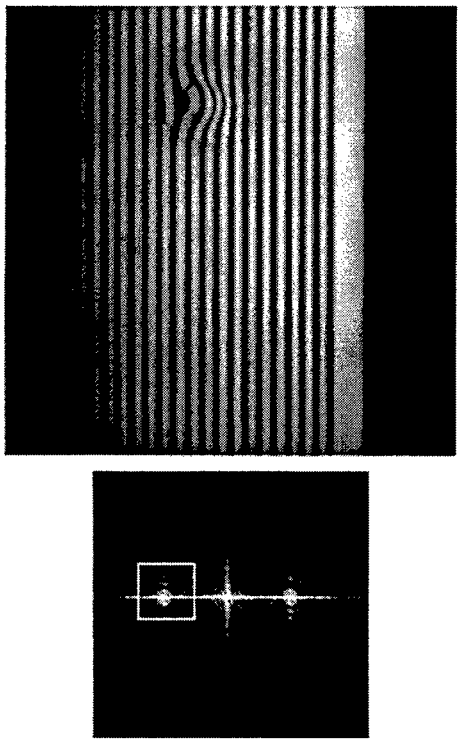
c) Result of the projection of the inverse pattern

Fig. 10: Inverse fringe projection technique

The local deviation from straightness within the analyzed fringe pattern can be detected very fast as a broadened 1st order peak in the Fourier spectrum, Fig. 11.



a) Cylinder without defect



b) Cylinder with defect

Fig. 11: Recognition of deviations from straightness of fringes as broadened 1st order peak in the Fourier spectrum

3.2 TESTING OF SNOM TIPS BY MEANS OF COMBINED FAR-FIELD AND NEAR-FIELD EVALUATION

Nowadays many tasks in physics, biology, medicine and technology deal with analyzing nanometer sized objects and structures such as photomasks and compact optoelectronic devices, nanopowder particles, nanosubstrates, tools for micro and laser surgery, polymers, biomolecules, etc. A significant member of scanning microscopy methods is the scanning near-field optical microscopy (SNOM) which becomes more and more preferable because of its variety of possibilities for optical, spectroscopic and topographical measurements. The working principle behind near-field optical imaging involves

illuminating a specimen through a sub-wavelength sized aperture while keeping the specimen within the near-field regime of the source. This microscopy may supply additional physical data about the nanosized matter.

However, the quality of the optical fiber tip is of decisive importance. Because the production process of pulled and coated glass fiber tips is still highly empirical and full of defects, a technique would be useful to determine the tips' quality before they are shipped to the user or mounted in the microscope. Consequently this part of the paper deals with an easy and fast full field method for the characterization of common 633nm glass fiber SNOM tips. This method is based on solving the reverse task of the image formation process[26]. Size and shape as well as disturbances at the aperture can be recognized by means of evaluating the far field distribution of the emitted intensity and phase which are recorded by a CCD target. A numerical model[26] is used which solves the reverse task that allows to draw conclusions from the measured intensity and phase distributions to the shape of the tip itself. A scheme of the inspection procedure is shown in Fig. 12.

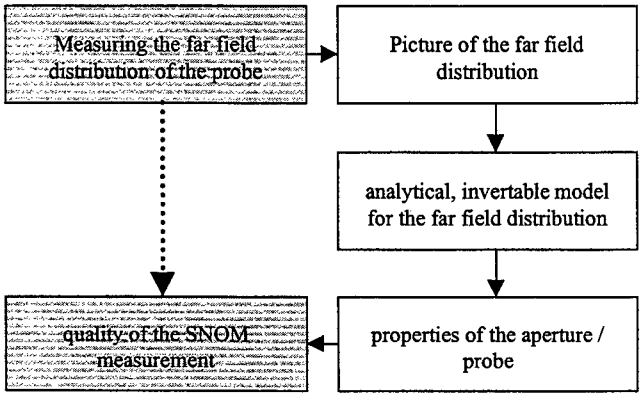


Fig. 12: Scheme of the evaluation process

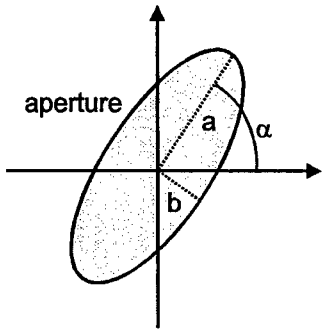


Fig. 13: Scheme of the aperture parameters to be reconstructed from the far field distribution: half axes a and b. Angle of inclination α.

The main aim of the experiment is the investigation of parameters of the aperture geometry by a full field approach. These parameters are:

- Shape of the aperture (circular, elliptical or rectangular)
- Sizes of the half axes of the aperture (see Fig. 13; parameters a and b)
- Angle of inclination of the elliptical/rectangular aperture (see Fig. 13; parameters α).

The experimental part of the procedure of parameter determination mainly consists of the recording of the light emitted from the probe if laser light (632nm) is coupled into the glass fiber from the other side. After the recording the inverse reconstruction algorithm[26] is used to calculate the appropriate parameters of the aperture.

The intensity of the light emitted from the tip is very low. Without using special (expensive) cooled CCDs or complicated photo chemical substrates it is necessary to collect as much light as possible on a small area on the CCD but on the other hand the structure of the far field distribution must remain visible. This means that as few optical components shall be positioned between tip and recording media as possible so that no internal reflections, absorbing and other losses decrease the intensity. The next problem is the angle of observation. As described in the theory section the intensity emitted in high angles is most important for the reconstruction of the aperture. Using high-aperture objectives (NA>0,95) one can register beams not over 65°..70°. This is not sufficient for a full reconstruction of the aperture geometry because tips with apertures<λ/2 emit light in high angles.

Several approaches were tested[26]. Here we concentrate on the application of spherical mirrors for the registration of the far-field distribution to overcome the problems with limited angular detection range, distortions and weakening, Fig. 14. Difficulties arise from the imaging properties of the mirror. There is no common focus and some of the rays cross the optical axis and interfere with other rays. So there is no simple way for the reconstruction of the angular intensity distribution by just one picture. At least seven pictures are necessary with varying distances *a* and *d3*. With each of the pictures one angular part of the intensity distribution can be achieved.

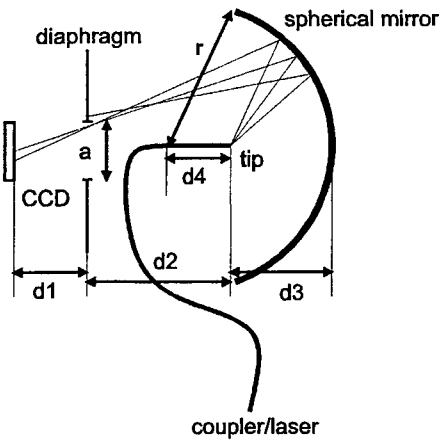


Fig. 14: Scheme of a setup for far field registration using a spherical mirror for collecting the light.

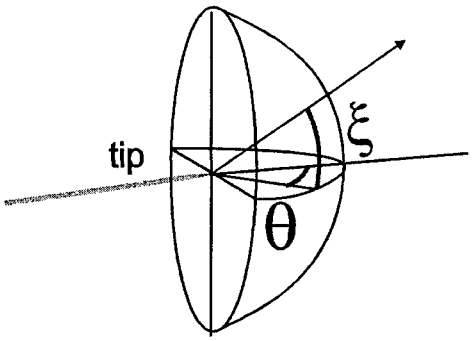


Fig. 15: Polar coordinate system

The following pictures are the basis for the reconstruction made in polar coordinates, Fig. 15.. The final image will consist of four parts, Fig. 16, which cope different angular part. The plots in Fig. 16 are in polar coordinates. For each part there is a special transformation dealing with intensity correction and distortion correction. The angular range is from 0° to 60° in these pictures but an angle of 90° can be achieved without difficulty. It depends just on the size and curvature of the mirror.

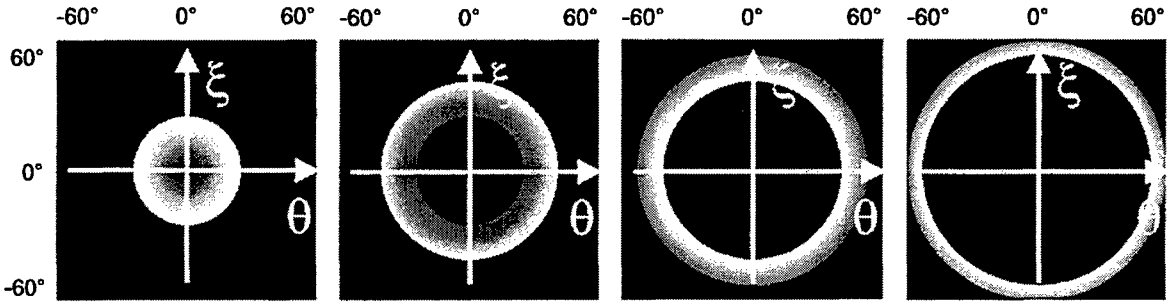


Fig. 16: Transformation elements for the reconstruction of the far field registration. Four steps are necessary to reconstruct a range from 0° to 60°.

Two different tips have been examined, Fig. 17. The first one was a SNOM tip which was used before so that its aperture was not perfect any more. The second was the end of a broken glass fiber tip so one can see a lot of interferences and disturbances. This second picture is good for adjustment and testing of the setup and the algorithms. The plots are in polar coordinates and show the angular distribution in the different directions in space. The plots in Fig. 17 can be used as input for the inverse procedure. But some filtering (removing camera noise, interferences) and subtracting of intensity offsets is necessary. The technique using a spherical mirror seems most promising for the far field characterization although it demands not only a single shot with the camera but seven images and some manipulation of the setup parameters.

The picture shown in Fig. 18b is used as the input for the inverse procedure. An circular aperture diameter of about 240nm (a=120nm, b=120nm, α=0°) could be reconstructed. Of course this aperture size is too big for SNOM applications but this experiment can show the principle of the evaluation. The measured far field distribution was used for the verification of the algorithms and the setup. The tip examined was used and slightly broken as REM pictures showed later.

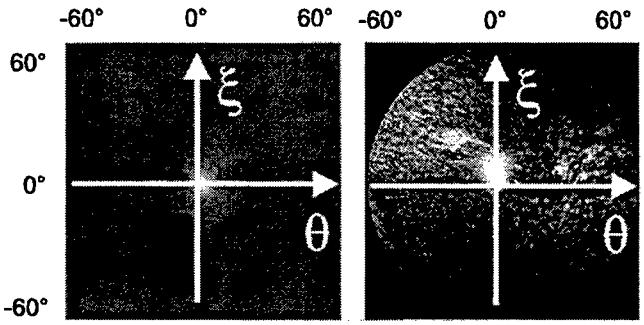


Fig. 17: Far field distribution of a (used) SNOM tip and a broken glass fiber tip

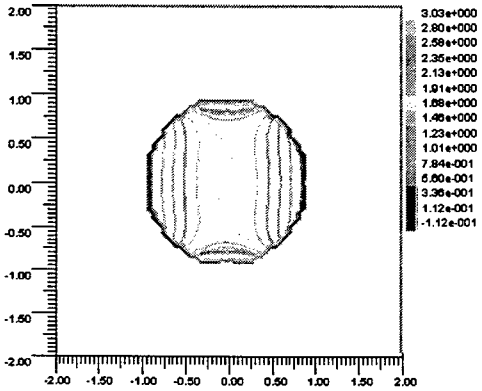


Fig. 18a: Simulated far field distribution for a SNOM tip in polar coordinates (coordinates are in sin(θ) and sin(ξ))

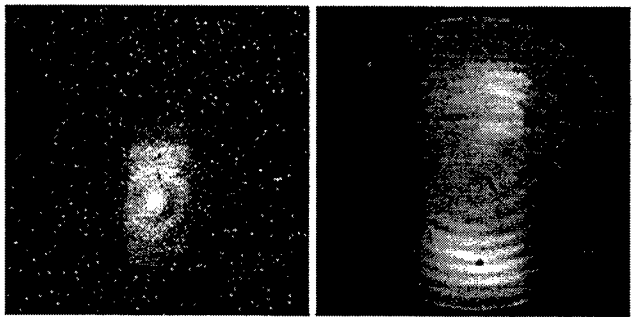


Fig. 18b: Left: experimentally recorded far field distribution on a plane CCD chip. (Original image. No intensity correction. No interference elimination). Right: Original image after transformation into polar coordinates.

4. CONCLUSIONS

In this paper several new system components for the implementation of active processing strategies in optical inspection were discussed. To them belong Vertical Cavity Surface-Emitting Laser (VCSEL), Complementary-Metal-Oxide-Semiconductor (CMOS) image sensors, Digital Micromirror Devices (DMD), Liquid Crystal Display (LCD) projection units and Optically Addressed Spatial Light Modulators (OASLM). Based on these new devices several modern applications such as adaptive fringe projection for shape control and aperture inspection of SNOM tips as an example for the solution of inverse problems were discussed showing the advantages of an active processing strategy in optical metrology.

5. ACKNOWLEDGEMENTS

This work was supported by the Federal Ministry of Research and Technology (BMBF) with grant No. 13N7104/8 and the German Research Foundation (DFD) with grant No. JU 142/31. The authors greatly appreciate the help of M. Kalms and S. Seebacher.

6. LITERATURE

1. P.K. Rastogi (Ed.), *Optical Measurement Techniques and Applications*, Boston & London, Artech House, 1997
2. C.-E. Liedtke, „Intelligent Approaches in Image Analysis," In: M. Kujawinska, G. Brown, and M. Takeda (Eds.), *Laser Interferometry IX: Techniques and Analysis*, Proc. SPIE, Vol. 3478, pp. 2-10, Bellingham, 1998
3. J. Aloimonos, „Reply, What I have Learned," Image Understanding, Vol. 60, No. 1, pp. 74-85, 1994
4. W. Osten, „Active Optical Metrology - A Definition by Examples," In: M. Kujawinska, G. Brown, and M. Takeda (Eds.), *Laser Interferometry IX: Techniques and Analysis*, Proc. SPIE, Vol. 3478, pp. 11-25, 1998

5. B.K.P. Horn, „Understanding Image Intensities," Artificial Intelligence. Vol. 8, 1977
6. V. Srinivasan, H.C. Liu, and M. Halioua, „Automated phase-measuring profilometry of 3-D diffuse objects," Applied Optics, Vol. 23, pp. 3105-3108, 1984
7. P.J. Besl P, „Active Optical Range Imaging Sensors," In: J.L.C. Sanz (Ed.), *Adavances in Machine Vision*, pp. 1-63, New York, Springer Verlag, 1989
8. W. Osten, W. Nadeborn, and P. Andrä, „General Hierarchical Approach in Absolute Phase Measurement," In: M. Kujawinska, G. Brown, and M. Takeda (Eds.), *Laser Interferometry VIII: Techniques and Analysis*, Proc. SPIE Vol. 2860, pp. 2-13, 1996
9. W. Nadeborn, P. Andrä, and W. Osten, „A Robust Procedure for Absolute Phase Measurement," Optics & Lasers in Engineering, Vol. 24, 245-260, 1996
10. M. Takeda , and H. Yamamoto, „Fourier Transform Speckle Profilometry: Three-Dimensional Shape Measurements of Diffuse Objects with Large Height Steps and/or Spatially Isolated Surfaces," Applied Optics, Vol. 33, pp. 7829-7837, 1994
11. W. Osten, and M. Kujawinska, „Active phase measuring metrology," In: P.K. Rastogi (Ed.): Metrology 2000, Elsevier (to be published)
12. C.J. Chang-Hasnain, „Report about VCSEL," Optics and Photonics News, May 1998, pp. 35-39
13. M.Y. Li et al.,"Top-emitting micromechanical VCSEL with a 31,6 nm tuning range," IEEE Photonics Technology Lett. 10(1), pp. 18-20, 1998
14. S.G. Hanson, R.S. Hansen, and B.H. Hansen, „Simple low-cost optical system for measuring 2D rotation of reflective ball," Proc. SPIE Vol. 3824, 1999, to be published
15. E. R. Fossum, „ CMOS digital cameras need new figures of merit," Laser Focus World, April 1999, pp. 101-108
16. DaimlerChrysler Aerospace, Newsletter RST Rostock 1/1999
17. G. Frankowski, „The ODS 800 - a new projection unit for optical metrology," In: W. Jüptner, and W. Osten (Eds.), Proc. Fringe'97, Akademie Verlag Berlin, pp. 533-539, 1997
18. P.S. Huang, Q. Hu, F. Jin, and F.-P. Chiang, „Color-encoded fringe projection and phase shifting for 3-D surface contouring," Proc. SPIE Vol. 3407, pp. 477-482, 1998
19. GF-Meßtechnik, „Product documentation for DMD shape measurement system MicroCAD," GFM GmbH, Teltow, Germany, 1999
20. B.E.A. Saleh, and M.C. Teich, Fundamentals of Photonics. John Wiley & Sons Inc. New York 1991
21. B. Carlson, „New turns on light valve technology," Advanced Imaging, May 1998, pp.62-63
22. JENOPTIK LOS Product brief, „Liquid crystal modulators," Jena, May 1999
23. V. Srinivasan, H.C. Liu and M. Halioua, "Automated phase-measuring profilometry of 3-D diffuse objects," Appl. Opt. 23 (1984), 3105-3108
24. W. Osten, „The Application of Optical Shape Measurement for the Nondestructive Evaluation of Complex Objects," Opt. Eng. 39(2000)1, (to be published)
25. M. Kalms, W. Osten, and W. Jüptner, „Inverse projected fringe technique with automatic pattern adaption using a programmable spatial light modulator," Proc. SPIE Vol. 3407(1998), pp. 483-489
26. S. Seebacher, W. Osten, W. Jüptner, V.P. Veiko, and N.B. Vosnessenski, „Determination og geometric properties of SNOM tips by means of a combined far-field and near-field evaluation,". Proc. SPIE Vol. 3740, OICOSN'99, pp. 312-322, 1999

Automatic measurement equipment for quality control in the loop slitting department

Stephan Hußmann*[1], Wolfram Kleuver[2], Bernd Günther[2], Jörg Gröneweller[2], Holger Rath[2]

[1]Universität-Gesamthochschule Siegen, Zentrum für Sensorsysteme (ZESS)
Paul-Bonatz-Straße 9-11,57068 Siegen, Germany
[2]aicoss GmbH, Birlenbacher Str. 14, 57078 Siegen, Germany

ABSTRACT

With this publication an automatic measurement equipment for quality control in the loop slitting department is presented. The sensor is a stand-alone system. A Digital Signal Processor (DSP), a Field Programmable Gate Array (FPGA) and a Microcontroller (MC) collect and process the data of the linear CCD-sensors.

The basic elements of the sensor are three CCD linear sensors and their associated LED light sources. The LED's are used for the transillumination of the sensor system. A multiplexer connects one active linear sensor at a time to the CPU. The switching between the linear sensors is automated by the DSP.

Furthermore this paper gives some detailed information about the accuracy, resolution and the measuring range of the sensor. One industrial application is listed at the end of this paper.

Keywords: Thickness measurement; industrial automation; optical design, loop slitting department, band material

1 INTRODUCTION

This sensor system has been developed for the foam fabricating industry. There was a need of a sensor that measures the foam thickness up to 0.01 mm correctly and saves the processed measurements during the whole production online. The customer needs the acquired data for a later documentation and evaluation in the quality control.

Other measuring methods don't work satisfactory. Triangulation sensors don't work because of pores in the foam material. The depth of the pores doesn't deliver the correct thickness value of the foam material. Ultrasonic sensors have drop outs because of the work noises of the plant. Radiometric sensors are harmful and therefore not applicable.

The used 'range of shadow' method is harmless and independent of the foam pores or the work noises. But the mechanical construction is more complicated as by the other measuring methods. Two strong stepping motors and a solid mechanical construction are required.

* S.H. (correspondence): Email: hussmann@zess.uni-siegen.de; WWW: http://www.zess.uni-siegen.de/private/hussmann;
Telephone: ++49-271 / 740-2449; Fax: ++49-271 / 740-2336
W.K.: Email: kleuver@aicoss.de
Telephone: ++49-271 / 77309-50; Fax: ++49-271 / 77309-59
B.G.: Email: guenther@aicoss.de
Telephone: ++49-271 / 77309-350; Fax: ++49-271 / 77309-390
J.G.: Email: groeneweller@aicoss.de
Telephone: ++49-271 / 77309-43; Fax: ++49-271 / 77309-49
H.R.: Email: rath@aicoss.de
Telephone: ++49-271 / 77309-54; Fax: ++49-271 / 77309-59

Part of the SPIE Conference on Advanced Photonic Sensors and Applications
Singapore • November–December 1999
SPIE Vol. 3897 • 0277-786X/99/$10.00

2 SYSTEM DESCRIPTION

2.1 Measurement principle

The sensor system consist of three CCD linear sensors with 3000 pixels and their associated illumination units. The illumination units deliver a almost parallel light beam. The receivers measure a continues value of the material thickness (range of shadow) over the complete cutting length at three different positions over the slitted foam material. The position of the outer sensors can be adjusted by means of a hand crank in a range from 250mm-2100mm absolute width. Therefore the thickness of different widths of the foam material can be measured. The middle sensor has a fixed position.

For starting the measuring cycle the user has to start the automatic calibration routine. During the automatic calibration no foam is allowed to be on the lift. The stepping motors put the upper edge of the foam lift in the belonging height under the light beam after entering the actual foam material thickness. Therefore the upper edge of the foam is always in the optimal measure position. The still measurable thickness divergence can reach up to the half of the light beam height. The automatic calibration has to start again only if the customer takes another foam material or change the cutting thickness of the foam material.

The principle of measurement for one linear sensor is shown in Figure 1.

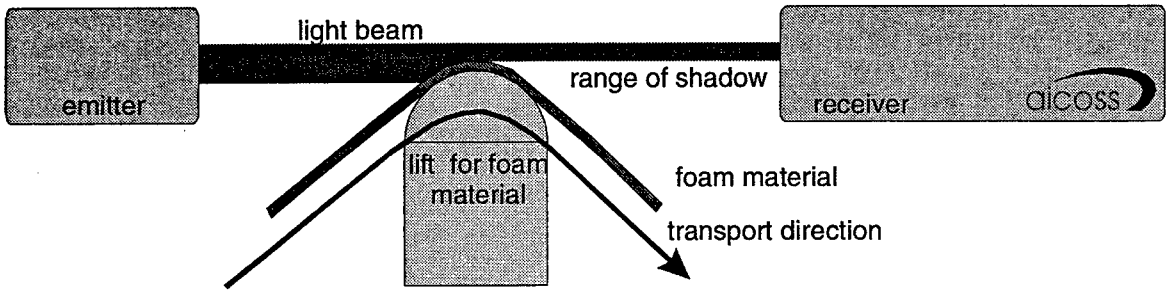


Figure 1: Principle of measurement

2.2 Functional block diagramm

The hardware consists of a DSP, a FPGA and a 8-Bit microcontroller. The FPGA delivers the clock signals for the three linear CCD-sensors. The DSP controls the timing and sends the preprocessed sensor data to the microcontroller. This hardware is used in our other sensors as well [1], [2], [4].

The multiplexer is controlled by the DSP. It connects the linear sensors serially to the CPU. In other applications we have used this method as well [3],[5],[6],[7].

The data is displayed online on the screen of the user terminal. During the operation, the user can overview separately the cutting thickness for all three tracks on the screen and intervene, if necessary. The processed thickness data is send to a PC via the ControllerAreaNetwork (CAN) bus and after finishing a cutting length (coil length) the PC prints out a quality card. That quality card can be sent with each individual coil to the customer. The PC also controls the stepping motors of the foam lift.

The functional block diagram as shown in Figure 2 displays the principal structure of the sensor.

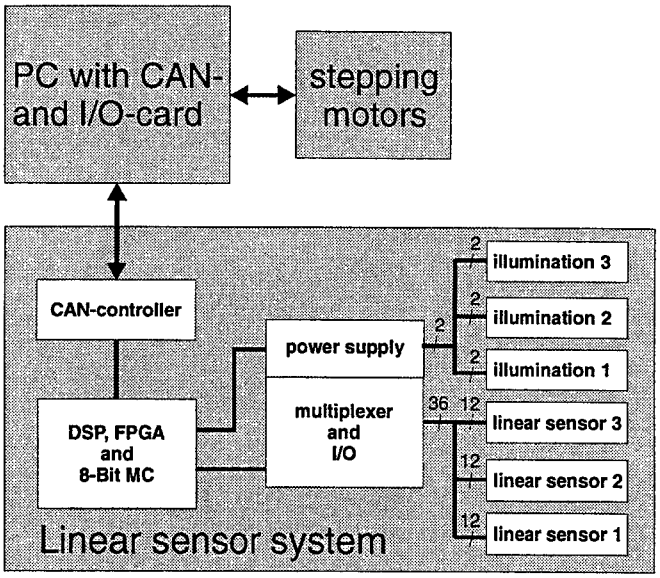


Figure 2: Functional block diagram of the sensor system

2.3 Specification of the sensor system

Table 1 lists the specifications of the sensor system. This specifications are not only limited to the foam fabricating industries. It can be used in the textile-, paper-, plastic-, metal- ... fabricating industry as well.

model	AI-S 100
range of application	measuring the thickness of foam, metal, textile, paper etc.
measuring range	0-10 mm
measuring rate	min. 200 Hz
accuracy	± 0,01mm
resolution (geometrical)	7 µm
interface	CAN-bus
system of protection (IEC 529)	IP64
power supply	24 VDC ± 30 %
current consumption	max. 1 A
light receiver	three 3000 pixel CCD-linear sensors
light transmitter	three red LED's (635nm)
package geometry transmitter	ca. 28 * 28 * 60mm
package geometry receiver	ca. 28 * 28 * 162mm
working temperature	0°- 50° C
storing temperature	-20° - 70°C

Table 1: Specifications of the high resolution linear sensor system

3 RESULTS

Figure 3 shows a sample of a quality card. In the upper right corner the customers setting values are printed. He has to input the nominal thickness of the foam and the required upper and lower tolerance limit. Additionally he inputs the distance between the outer two sensors of the sensor system (width of tracks). The value has to be read off the hand crank display.

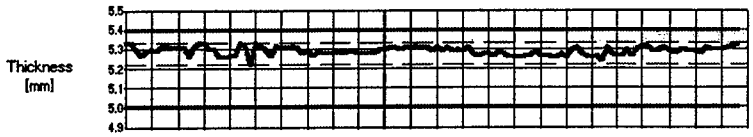
aicoss GmbH
Birlenbacher Str. 14
D-57078 Siegen

aicoss

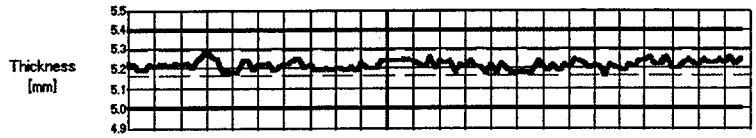
Foam thickness measurement - Quality assurance

Client order:	123456	**Nominal thickness [mm]:**	5,2
Internal order num.:	123456	**Upper tolerance limit [mm]:**	0,2
Block number:	123456	**Lower tolerance limit [mm]:**	-0,2
Role:	1	**Width of tracks [mm]:**	1830
Client sign:	TEST	**Material code:**	AB 123456
Running number:	1		

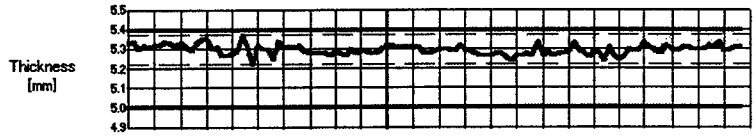
Left track



Middle track



Right track



Measurement value recording interval / process control:	1 sec
For the graphics: Measurement values per interval:	5
Time difference between the displayed values:	1 sec

Maximal thickness [mm]:

Left track:	5,37	Middle track:	5,29	Right track:	5,37

Minimal thickness [mm]:

Left track:	5,22	Middle track:	5,17	Right track:	5,22

aicoss Date: 03.09.99 Time: 16:49

Figure 3: Sample of a quality card

The three diagrams display the complete measure cycle of a foam coil. Every sensor has its own diagram. The measurement value recording interval amounts one second. That means every second all three thickness values are stored onto the hard disk.

In the three diagrams the time difference between two thickness value is one second. Per interval five values are displayed. The upper and lower limit of the foam thickness is plotted as two bold vertical lines. The measured foam thickness has to be between that lines to fulfill the quality criterion of the production.

In the lower part of the quality card the maximal and minimal thickness of every measure value is displayed. Again there is a separation between the three sensors. These values are used for the quality control of the produced product as well.

The QM-DIN* EN ISO 9000-9004 applies to the quality management of a enterprise. Those norms are no indicator for a high product quality. But the organization of the production by means of those norms meet the requirements for a constant high product quality. At the moment it isn't imperative by law of having a QM-certification. But the international competition requires that certification. It is an important marketing instrument.

The improvement of the quality control in the production process requires new types of sensor systems - especially for inspection tasks.

The manufacturer needs the developed sensor system to control the quality in the loop slitting department. By means of the sensor system he can prove his customers that the slitted foam has a constant thickness.

4 CONCLUSION

- The developed sensor fulfills its requirements. It measures the thickness of the foam material in a range from 0–10mm with an accuracy of ± 0,01mm online.
- Pores in the foam material don't influence the measured foam thickness.
- Due to the hand crank the thickness of different foam widths can be measured.
- The sensor works independent of the color and the material of the foam.
- With the generated quality assurance the manufacturer fulfills parts of the QM-DIN EN ISO 9000-9004
- The use of the sensor is not only limited onto the foam fabricating industry. It can be used for all coiling applications to measure the thickness of the material. The sheet metal-, textile-, paper- or plastic producing industries are named for example.

5 REALIZATION

The prototype of the automatic measurement equipment for quality control has been already integrated into the loop sliting department and works fine. It is in operation service since December 1998.

Figure 4 shows one transmitter/receiver unit of the sensor system in detail. The foam lift and the cutted foam is displayed as well. Below the foam lift is the uncutted foam block.

*DIN: German institute of standardization

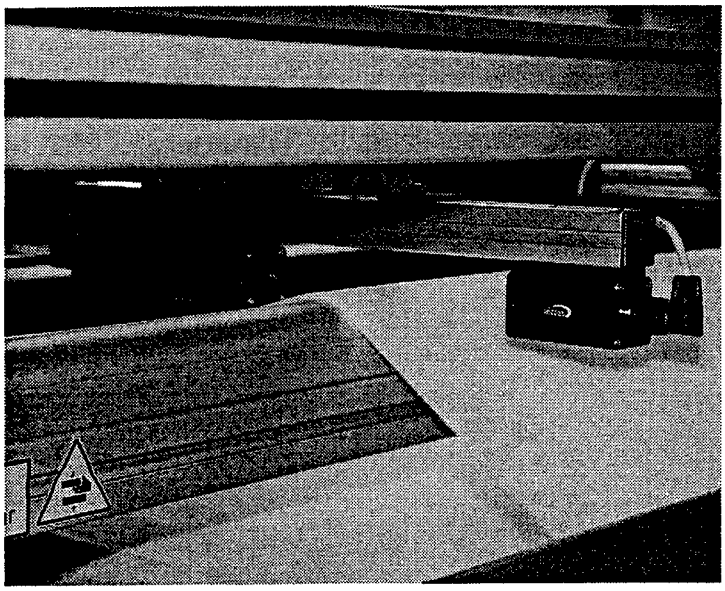


Figure 4: One of the three transmitter/receiver units in detail

The sensor system was developed at the center for sensor systems (ZESS) in cooperation with the company aicoss GmbH. Figure 5 shows the sensor system in its industrial environment.

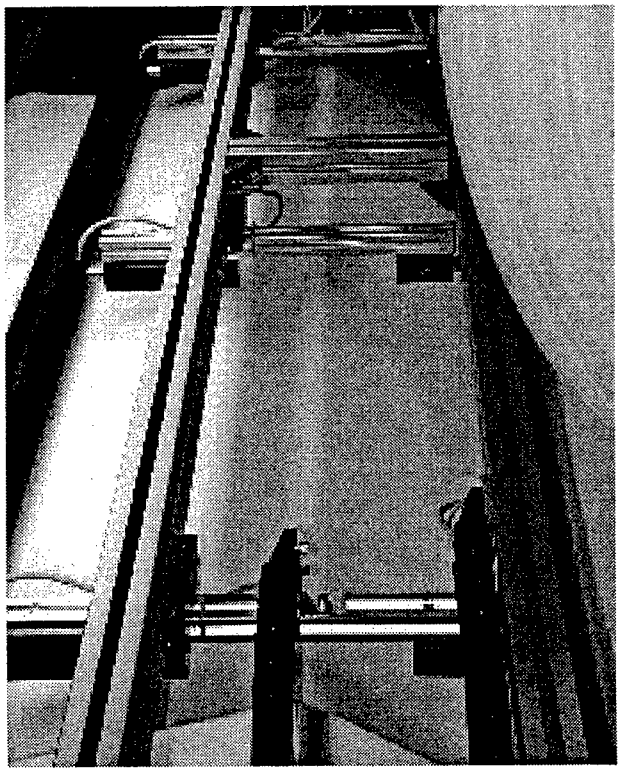


Figure 5: Three sensors on different positions over the slitted foam material

6 ACKNOWLEDGMENTS

The authors are grateful to the center for sensor systems (Zentrum für Sensorsysteme, ZESS) at the university of Siegen for the financial and technical support.

7 REFERENCES

1. W. Kleuver, M. Fuss, "Triangulation and surface inspection in the automatic inspection of drillings and hollows", *Machine Vision Applications in Industrial Inspection IV*, SPIE Vol. 2665, pp. 167-176, 1996
2. W. Kleuver, J. Weber, D. Justen and K. Hartmann, "Fast inspection of radius, wall thickness, eccentricity and ovality at the end of high precision tubes: TRUDE 60-28 in the production", *Sensors, Sensor Systems and Sensor Data Processing*, SPIE Vol. 3100, pp. 340-348, 1997
3. S. Hußmann, W. Kleuver "Coillängen-Meßgerät", professional magazine: *Bänder-Bleche-Rohre*, Vogel Verlag Würzburg, p. 52-54, May 1998
4. W. Kleuver, "Berührungslose Entfernungsmessung in der Qualitätssicherung mit Triangulationssensoren", dissertation at the university of Siegen, Center for Sensor Systems (ZESS), ISBN 3-8265-4110-3, July 1998
5. S. Hußmann, W. Kleuver, "High precision triangulation sensor for residual shorts measurement of coiling materials", *Optics and Optoelectronics: Lasers and Applications*, O. P. Nijhawan, A. K. Gupta, A. K. Musla and K. Singh, Volume1, pp. 267-270, Narosa Publishing House, ISBN 81-7319-285-5, 1998
6. S. Hußmann, W. Kleuver and J. Gröneweller, "High resolution linear sensor system for sheet metal width measurement", *Optical Measurement Systems for Industrial Inspection: Industrial Applications II*, SPIE Vol. 3824, 1999
7. S. Hußmann, W. Kleuver, "Coil identification algorithm inside a high precision triangulation sensor for residual shorts measurement of coiling materials ", *Sensors and Controls for Intelligent Machining II: Sensors and Applications*, SPIE Vol. 3832, 1999

Absolute thickness measurement using automatic fractional fringe order method

M. Tsai, R. Tian, H. Huang, M. Itoh and T. Yatagai
Institute of Applied Physics, University of Tsukuba
Tennoudai 1-1, Tsukuba, Ibaraki 305-8573, Japan

Abstract

The use of a broad-band interferometer and its fringe analysis method are presented in which the fractional fringe order method for measuring the absolute thickness. An approximate integer part of the fringe order is estimated by mechanical measurements, and the fractional part is determined by interferometric fringe pattern analysis. The fringe patterns are obtained with a Michelson interferometer by illumination of several selected wavelengths, respectively. The use of the fractional fringe order method can determine accurately more than 2π phase jumps. The influence of wavelength and approximate integer part of fringe order on the measurement accuracy is discussed.

Keywords: absolute measurement, fractional fringe order method, excess fraction, Phase stepping method, block gauge, 2π ambiguity

1. INTRODUCTION

Optical interferometry such as spatial and temporal heterodyne methods[1,2] and phase stepping[1-6] are being applied to quantitative measurement of two-dimensional continuous surface profiles. However, when the height of surface discontinuities is greater than a wavelength λ absolute measurement is difficult because the phase retrieved from the fringe pattern includes ambiguities of 2π phase. Beard-band interferometers have been introduced to overcome this difficulty. This technique is also called the coherence probe or the coherence radar method.

Many approaches have been proposed to deal with the problem of 2π ambiguity. These techniques include using a longitudinal fringe, a lateral fringe, and both a longitudinal and a lateral fringe.[7-16] Using the longitudinal fringe method, Dresel *et al.* proposed a coherence radar method for sensing the three-dimensional profile of rough surfaces.[11] Lee and Strand traced the peak point of a white-light interferogram of which the optical path difference is zero while they moved a reference mirror along the optical axis to characterize surfaces using a coherence scanning microscope,[12] and Itoh *et al.* used this technique to measure a 3-D shape using a Twyman-Green interferometer.[13] Using the lateral fringe method, Takeda *et al.* proposed a Fourier transform technique for analyzing the spatial fringe pattern.[14] Li *et al.* used the equivalent wavelength of two frequencies on one grating to measure an unwrapping phase profile.[15] Utilizing simultaneous measurement of a longitudinal and lateral interferometric fringe, Ikonen *et al.* used a white light source to mark the peak difference distance of interferogram and one stabilized laser to determine phase shifts between the reference flat surface and the block gauge surface , respectively.[16]

Several fringe analysis algorithms have also been proposed to study the problem of 2π ambiguity. Macy applied a sinusoid fitting method that needed no locating of fringe centers or assignment of fringe orders to determine two-dimensional wave-front topography phases.[17] Takeda *et al.* proposed a spatial frequency-multiplex Fourier transform technique combined with the Gushov-Solodkin phase unwrapping algorithm to measure three dimensional shape objects with large height discontinuities.[18] Lewis used the phase-stepping technique and multiple-wavelength interferometry to measure the length of mechanical length standards.[19] We also combined the fractional fringe order method with the phase shift method to scan surface

Part of the SPIE Conference on Advanced Photonic Sensors and Applications
Singapore • November–December 1999
SPIE Vol. 3897 • 0277-786X/99/$10.00

profiles with a wide gap.[20]

We propose a new method that combines the Fourier transform with the fractional fringe order method to measure a block gauge thickness that exceeds the wavelength. The absolute thickness is expressed by a real number multiplied by the half-wavelength used. We use the Fourier transform method to determine the fractional part of the real number and the fractional fringe order method[21-23] to determine its integral part from several interferograms of different wavelengths.

2. FRINGES IN BROAD-BAND INTERFEROMETRY

2.1 Fringes in broad-band interferometry

The interferometric fringes in a mono-chromatic interferometer is written by

$$I(x,y,\delta)=i(\sigma)(1+\cos(2\pi\sigma(\Phi(x,y)+\delta))) \tag{1}$$

where $i(\sigma)$ is the spectrum of the source as a function of the spectroscopic wavelength number σ, $\Phi(x,y)$ the optical path difference between the object beam and the reference beam, which is related to the shape of a testing object, and δ the initial valve of optical path length. For the broad-band light source whose wave number is from σ_1 to σ_2., the intensity of the fringes is given by the integrating Eq.(1) from σ_1 to σ_2.

$$I(x,y,\delta)=\int_{\sigma_1}^{\sigma_2} i(\sigma)[1+\cos(2\pi\sigma(\Phi(x,y)+\delta))]d\sigma \tag{2}$$

For the narrow-band case, the fringes of Eq.(2) are analyzed by the phase stepping method or the Fourier transform method. In the case when the band width $2\Delta\sigma$ of the light source is relatively narrow and $i(\sigma)$ is constant with in the band, the fringe intensity is given by

$$\begin{aligned} I(x,y,\delta) &= \int_{\sigma_0-\Delta\sigma}^{\sigma_0+\Delta\sigma} i(\sigma)[1+\cos(2\pi\sigma(\Phi(x,y)+\delta))]d\sigma \\ &= 2i(\sigma_0)\Delta\sigma + 2i(\sigma_0)\frac{\sin(2\pi\Delta\sigma(\Phi(x,y)+\delta)}{2\pi(\Phi(x,y)+\delta)}\cos(2\pi\sigma_0(\Phi(x,y)+\delta)) \\ &= a_0 + a_1\cos\delta' + a_2\sin\delta' \end{aligned} \tag{3}$$

where

$$a_0 = 2i(\sigma_0)\Delta\sigma$$

$$a_1 = 2i(\sigma_0)\frac{\sin(2\pi\Delta\sigma(\Phi(x,y)+\delta)}{2\pi(\Phi(x,y)+\delta)}\cos(2\pi\sigma_0\Phi(x,y))$$

$$a_2 = -2i(\sigma_0)\frac{\sin(2\pi\Delta\sigma(\Phi(x,y)+\delta)}{2\pi(\Phi(x,y)+\delta)}\sin(2\pi\sigma_0\Phi(x,y))$$

$$\delta' = 2\pi\sigma_0\delta$$

For a small bandwidth of σ_0, the term $\frac{\sin(2\pi\Delta\sigma(\Phi(x,y)+\delta)}{2\pi(\Phi(x,y)+\delta)}$ is considered to be constant, and therefore the phase stepping method could be used for this case. The phase is obtained by

$$\Phi(x,y)=\frac{1}{2\pi\sigma_0}\tan^{-1}\frac{a_1}{a_2}=\frac{1}{2\pi\sigma_0}\tan^{-1}\frac{\sum I_i\sin\delta_i'}{\sum I_i\cos\delta_i'} \tag{4}$$

where

$$\delta_i'=\frac{2\pi i}{N} \qquad (i = 0, 1, \ldots, N-1)$$

In the phase analysis method mentioned above, the 2π phase ambiguity problem should be solved.

2.2 Fractional fringe order method

Consider the problem of determining the gap height L of a step between two parallel surfaces. The fringe order equation for height L is simply given by

$$L = (n+\varepsilon)\frac{\bar{\lambda}}{2}, \tag{5}$$

where $\bar{\lambda}$ is a central wavelength, n is an integer, and ε is the excess fraction of the difference between the fringe orders on two parallel surfaces of a sample, that is to say, L is as large as $(n+\varepsilon)$ times the period of $\bar{\lambda}/2$. The excess fraction of ε_j for different wavelengths of $\bar{\lambda}_j$, where $j = 1,2,3,.....m.$ and $m \geq 2$, respectively, is determined by phase shifting interferometry as described in Section 2.1. For an approximated value of the measured height L_0, integer n_1 is estimated for $\bar{\lambda}_1$ as

$$n_1 = \left[\frac{L_0}{(\bar{\lambda}_1/2)}\right], \tag{6}$$

where $[\alpha]$ is an operator that takes an integer part of a real number α. The approximated value L_0 is known by a mechanical measurement or *a priori* information to limit the range of calculation process and therefore to save computing time. By using n_1 and the measured excess fraction of ε_1, we obtain the first estimated height

$$L_1 = (n_1+\varepsilon_1)\frac{\bar{\lambda}_1}{2}. \tag{7}$$

Using this L_1, we can obtain

$$\varepsilon_j' = \frac{L_1}{(\bar{\lambda}_j/2)} - n_j, \tag{8}$$

where $n_j = \lfloor L_1/(\bar{\lambda}_j/2) \rfloor$ and $j = 2,3,4,.....,m$. If ε_j' agrees with ε_j in the error range $\Delta\varepsilon$ $(j = 2,3,4,.....,m)$, L_1 is estimated to be an absolute height of L. When ε_j' does not agree with ε_j, n_1 is increased or decreased by 1, and similar procedures are repeated to obtain the k-th estimated height L_k until ε_j' agrees with ε_j $(j = 2,3,4,.....,m)$. In our algorithm the value L_0 is scanned among a range, therefore, this *a priori* knowledge is not critical. The initial value of $\Delta\varepsilon$ is 0.02. If there is no solution or more the one solutions, we increase or decrease $\Delta\varepsilon$ a small amount, and the procedure is repeated until a best fitted solution is obtained. The selection of wavelengths is arbitrary. In order to maximize the unambiguous detection range of the method, the selected wavelengths should not be integer multipliers of each other.[24]

3. EXPERIMENTAL VERIFICATION

To verify the broad-band interferometric fringe analysis, a wing of a dragonfly is measured.In the experiment, the Twyman-Grenn interferometer
was used, and a collimated brad-band light tungsten lamp as the light source. The reference mirror is mounted on pizeo-electric transducer that can be precisely moved by applying voltages and controlled by a computer. A sample is attached on the mirror of measurement stand. A cooled CCD camera with 1024x1024 resolution pixels and 14-bit accuracy was used to capture the fringe patterns. A color filter is used to pick up the special interferograms related with the central wavelength of the filter. In the experiments, four-fringe patterns were recorded consecutively with a phase shift of $\pi/2$. Several sets of such four-fringe patterns that correspond to the individual central wavelengths of the filters were saved in the memory for analysis. . Interferograms and the surface shape analyzed are shown in Fig. 1.

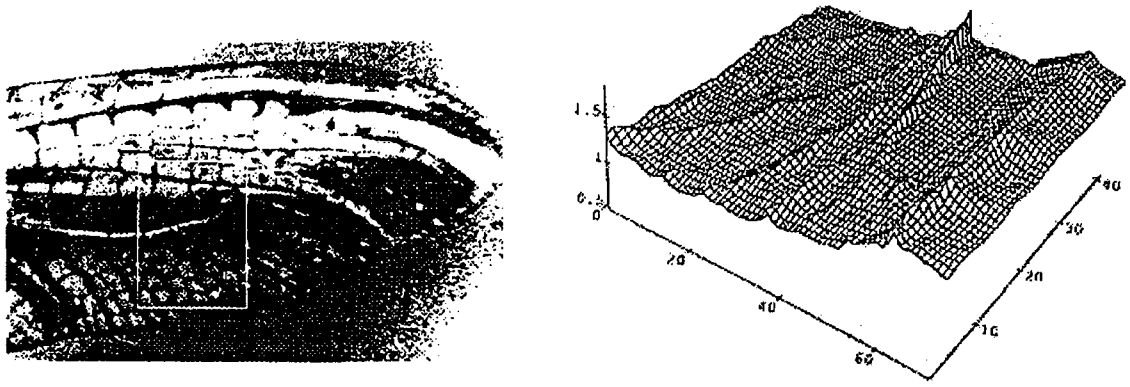


Fig. 1 Interferogram of a wing of a dragonfly and its 3-D shape.

To verify the fractional fringe order method, we have measured a test sample. A poled polymer film of about 5 μ m in thickness was doped on the upper half of a glass substrate. A thin aluminum layer covered the whole surface. Three sets of four-fringe patterns were recorded consecutively with central wavelengths of the filters of 549.8nm, 650.2nm and 700.0nm. The filter had a full width of half maximum (FWHM) of 40nm. Fig. 2 shows one of fringe patterns for wavelength 700.0nm. In the analysis, we cut out the interesting small part of 64 x 64 pixels from the whole 512 x 512 images as shown in Fig.2. According to the algorithm, we obtained a gap height of 4.897 ± 0.003 μ m. A sectional shape of the measured sample is shown in Fig. 3. The shapes of the upper and lower surfaces are also measured in the error range of $\Delta\varepsilon = 0.07$.

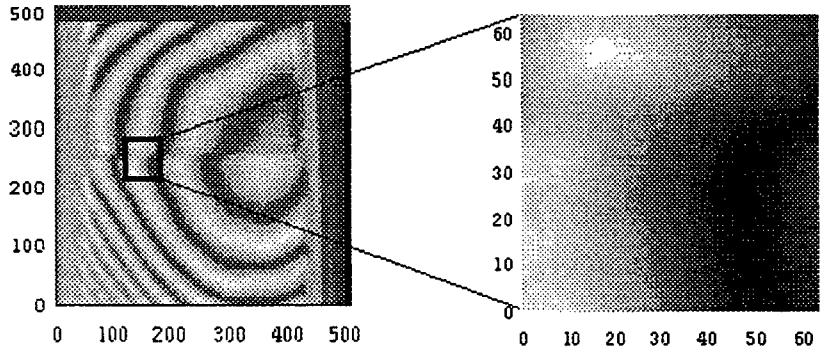


Fig. 2 Interferogram and its analyzed area.

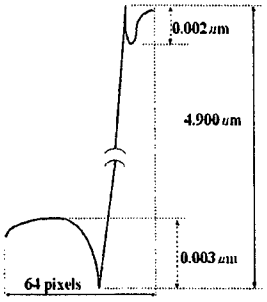


Fig.3 Sectional shape of the object.

4. CONCLUSION

In this paper, a novel technique in broad band interferometry is discussed. A general fringe formation for a wide-band interferogram is described. The use of fractional fringe order method is critical in absolute distance or thickness measurement. Two examples of measurement for test sample are presented.

5. REFERENCES

1. K. Creath: *Interferogram Analysis: Digital Fringe Pattern Measurement Techniques*, eds D. W. Robinson and G. T. Reid (Institute of Physics Publishing Bristol and Philadelphia, 1993) p. 94.
2. M. Kujawinska: *Interferogram Analysis: Digital Fringe Pattern Measurement Techniques*, eds D. W. Robinson and G. T. Reid (Institute of Physics Publishing Bristol and Philadelphia, 1993) p. 141.
3. K. Creath: *Progress in Optics*, ed. E. Wolf (North-Holland Elsevier Science Publishers, Amsterdam, 1988) Vol XXVI, p. 349.
4. V. Srinivasan, H. C. Liu and M. Halioua: Appl. Opt. **23** (1984) 3105.
5. K. Hibino, B. F. Oreb, D. I. Farrant and K. G. Larkin: J. Opt. Soc. Am. A **14** (1997) 918.
6. J. H. Bruning: *Optical Shop Testing*, ed. D. Malacara (John Wiley and Sons, Inc., NewYork,1978) p. 409.
7. P. A. Flournoy, R. W. McClure and G. Wyntjes: Appl. Opt. **11** (1972) 1907.
8. P. de Groot and L. Deck: Opt. Lett. **18** (1993) 1462.
9. L. Deck and P. de Groot: Appl. Opt. **33** (1994) 7334.
10. K. Matsui and S. Kawata: Proc. SPIE **1720** (1992) 124.
11. T. Dresel, G. Häusler and H. Venzke: Appl. Opt. **31** (1992) 919.
12. B. S. Lee and T. C. Strand: Appl. Opt. **29** (1990) 3784.
13. M. Itoh, R. Tian and T. Yatagai: Proc. SPIE **1755** (1992) 24.
14. M. Takeda, H. Ina and S. Kobayashi: J. Opt. Soc. Am. **72** (1982) 156.
15. J. Li, H. Su and X. Su: Appl. Opt. **36** (1997) 277.
16. E. Ikonen, J. Kauppinen, T. Korkolainen, J.Luukkainen and K. Riski: Appl. Opt. **30** (1991) 4477.
17. W. W. Macy, Jr.: Appl. Opt. **22** (1983) 3898.
18. M. Takeda, Q. Gu, M. Kinoshita, H. Takai and Y. Takahashi: Appl. Opt. **36** (1997) 5347.
19. A. J. Lewis: *Appl. Opt. Digest, Appl. Opt. & Opto-Elec. Conf.*, **17-20** *Sep. 1990, Nottingham, UK*, ed. J. C. Dainty (1990) p. 269.
20. M. Tsai, M. Itoh, T. Yatagai:*Technical Digest of International Workshop on Interferometry Satellite Meeting, Japan* (1996) p. 11.
21. C. Candler: *Modern Interferometers* (Hilger and Watts, Glasgow, London, 1951) p. 218.
22. W. H. Steel: *Interferometry*, eds A. Herzenberg and J. M. Ziman (Great Britain, Cambridge, U. P., London, 1967) p. 137.
23. P. Hariharan: *Optical Interferometry* (Academic Press, New York, 1985) p. 119.
24. C. R. Tilford: Appl. Opt. **16** (1977) 1857.

SESSION 9

Confocal Microscopy

Invited Paper

Separate measurement of geometrical thickness and refractive index by an interference confocal microscope

Takashi Fukano*[a,b] and Ichirou Yamaguchi[a]

[a] Optical Engineering Laboratory, RIKEN (The Institute of Physical and Chemical Research),
2-1 Hirosawa, Wako, Saitama 351-0198, Japan

[b]Graduate School of Science and Engineering, Saitama University,
255 Shimo-Okubo, Urawa, Saitama 338-8570, Japan

ABSTRACT

In this paper we first survey non-contacting and separate measurement of thickness and refractive index of a transparent plate or film. Then we describe the separate measurement of refractive index and geometrical thickness of multiple layers which uses a combination of confocal microscopy and low-coherence interferometry. The measurement has been accelerated by replacing the low-coherence interferometry by wavelength-scanning heterodyne interferometry using a laser diode. Finally, accuracy of the measurement has been experimentally studied and display of a cross-sectional image of a stack of glass plates has been demonstrated.

Keywords: Interferometry, microscopy, thickness measurement, refractive index measurement, wavelength tuning

1. INTRODUCTION

Thickness measurement of transparent films and plates of glass and plastics are important in optical engineering. Recently, demands for quality control of thickness in production process are rapidly increasing, and hence non-contact, non-destructive, and quick measurement are especially desired. For this purpose, low-coherence interferometry (LCI) [1,2], microscopic spectrophotometry [3], and confocal microscopy [4,5] have been proposed. Each of these methods, however, delivers not geometrical thickness but optical thickness that is equal to the product of the geometrical thickness and the refractive index.

The refractive index is also important in quality control of optical components. Imperfect control of uniformity in material composition and temperature in production process leads to the refractive index inhomogeneity. Moreover, the residual stress arising from non-uniform cooling causes birefringence. Therefore, separate measurement of refractive index and thickness are required. Recently, the needs for geometrical thickness measurement of multiple layers and films composed of transparent or scattering media as well as inspection of inhomogeneous films have also increased.

In this paper, we first survey the previous methods applicable to a single layer. Second, we explain the principle and experimental setup of the method that uses a combination of confocal microscopy and low-coherence interferometry. Then we propose an acceleration of the measurement by replacement of the low-coherence interferometry by wavelength-scanning heterodyne interferometry. Finally, we present some experimental results for measurement of a single glass sheet and cross-section of stacked glass plates.

2. PREVIOUS METHODS

For separation between refractive index and geometrical thickness, we can employ a combination of two independent measurement methods. Flournoy *et al.* proposed a method based on the low-coherence interferometry as shown in Fig.1(a) [6]. A light beam from a broad-band source is incident on the specimen. The optical path difference between the lights reflected at the front and the rear surfaces of the specimen is measured by using an external Michelson interferometer. The interference signal shows the maximum when the optical path difference of the interferometer is equal to that introduced by the specimen. The

*Correspondence: Email: tenchan@optsun.riken.go.jp; WWW: http://optsun.riken.go.jp/index.html; Telephone: +81 48 462 1111 (ex. 3243); Fax: +81 48 462 4653

Part of the SPIE Conference on Advanced Photonic Sensors and Applications
Singapore • November–December 1999
SPIE Vol. 3897 • 0277-786X/99/$10.00

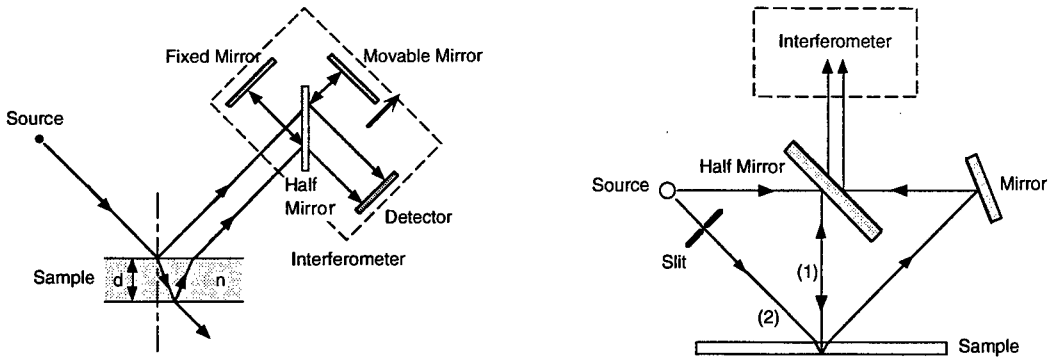


Fig.1 (a) Schematic diagram of measuring optical thickness by low-coherence interferometry. (b) Extended arrangement for separate measurement of refractive index and geometrical thickness [6].

separation between refractive index and geometrical thickness is performed by measuring optical path differences at the two different incident angles, for example, normal and oblique incidence upon the film as shown in Fig.1(b).

Another method using low-coherence interferometry was proposed by Sorin and Gray [7]. A sample and a reference mirror RM are positioned in the sample arm of a low-coherence Michelson interferometer using optical fibers as shown in Fig.2. By moving the reference mirror, the change in the optical path length after removal of the sample is measured. The optical thickness of the sample is measured from the intervals between the fringe contrast peaks corresponding to the front and the rear surfaces. From the two quantities thus obtained, the refractive index and geometrical thickness can be separated. In this method, we need scan the reference mirror twice. Recently, an improved version has been proposed [8,9]. The difference is that the sample is placed on a reference mirror of an ordinary Michelson interferometer to form a step. Here, only single scanning of the reference mirror is needed. The separate measurement of refractive index and geometrical thickness of a crystalline quartz has been realized [9].

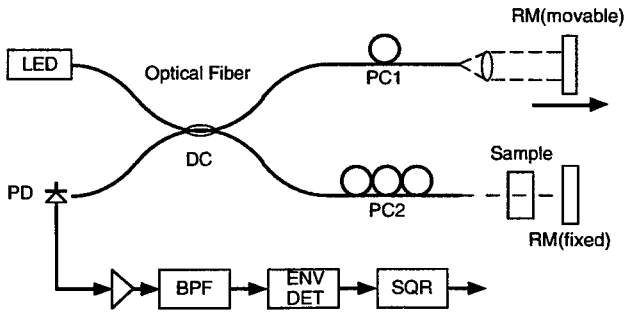


Fig.2 Fiber optic low-coherence interferometry for separate measurement of refractive index and geometrical thickness. Measurement are performed by scanning reference mirror twice with and without sample [7].

We also proposed a method that does not use low-coherence interferometry [5]. As shown in Fig.3, a *p*-polarized laser beam is incident upon a plane parallel sample and we measured the interferometric signals arising from multiple reflection at both surfaces of the sample by varying the incident angle. The refractive index is determined from the Brewster's angle, where the detector output reaches the minimum. The optical thickness is derived from the period of interferometric signals. Accordingly, the geometrical thickness is calculated from both the measured refractive index and optical thickness.

A method based on geometrical optics has also been presented by El-Zaiat [10]. A narrow laser beam is incident on the glass plate, and then two beams reflected, from the front and the rear surfaces are projected on a distant screen as shown in Fig.4. The separations between the two reflected beams that are measured at two incident angles lead to the refractive index and geometrical thickness. Although this method is very simple, the resultant accuracies of refractive index and geometrical thickness for a 1-mm thickness sample are as high as 10^{-3} and 10 μm, respectively.

All these methods, however, are applicable only to a single layer sample with uniform thickness and refractive index. Methods

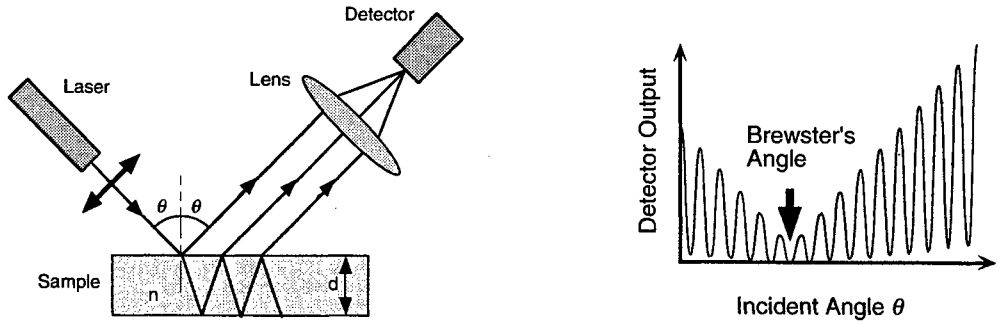


Fig.3 (a) Measurement of refractive index and geometrical thickness using detection of Brewster's angle and period of the fringe [5]. (b) Example of detector output versus incident angle.

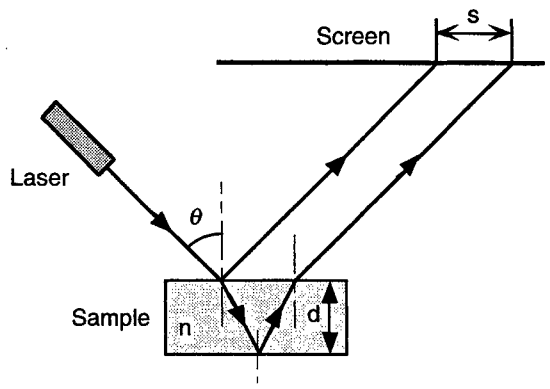


Fig.4 Separate measurement of refractive index and geometrical thickness by detecting separation of two reflected beams caused by front and rear reflections [10].

that can also be used for multiple layers have been recently proposed by us [11] and another group [12,13], independently. They can also be applied to a non-uniform sample because a convergent beam is incident on it. We describe their principles in detail below.

3. Method Using Confocal Microscopy and Low-Coherence Interferometry

3.1. Apparatus

We first proposed a separate measurement of refractive index and geometrical thickness that is based on combination of confocal microscopy and low-coherence interferometry. As shown in Fig. 5, we use a laser diode operating below the threshold current as the broad-band point source required for both low-coherence interferometry and confocal microscopy. The emitted light is collimated and divided into two arms. One illuminates a sample through an objective lens and the other is used as a reference light. For the reference mirror we use a cat's eye arrangement composed of the plane mirror, while the same objective lens as that inserted in the sample arm. This setup compensates for distortion of interference signals due to wavelength dispersion. The light beams returned from the sample and the reference cat's eye are superimposed and collected onto the pinhole. The detected light is recorded by the computer. Both the sample and the cat's eye are mounted on separate translation stages.

3.2. Separation of the refractive indices and geometrical thicknesses

Here, we explain the principle of the separate measurement of the refractive indices and geometrical thicknesses. Figure 6 shows the schematic diagrams of the principle with related confocal signal and interferograms. When the focal position of the convergent beam incident on a sample coincides with the front surface of the m-th layers as shown in Fig.6(a), the detected

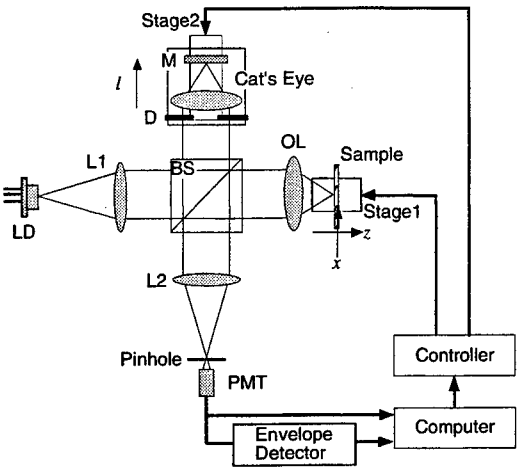


Fig.5 Experimental apparatus for separation of refractive indices and geometrical thicknesses[14]. LD: laser diode (LED operation), L's: lenses, BS: beam splitter, OL: objective lens, M: mirror, D: diaphragm, PMT: photo-multiplier tube.

signal from the confocal microscope becomes maximum as shown in Fig.6(c). This sample position is represented by z_m. With this sample position the reference mirror of the interferometer is then moved. The interference signal peak appears at l_m when the optical path difference of the interferometer is zero. We repeat the same procedure for the rear surface of the m-th layer as shown in Fig.6(b). At the next confocal peak of z_{m+1}, the interference signal peak appears at a different position l_{m+1}. The interval Δz_m between the confocal peak positions corresponds to the sample displacement needed for focus readjustment, while the difference Δl_m between interference peak positions stands for the increment of optical path length for focus readjustment. We can derive the geometrical thickness d_m and the refractive index n_m of the m-th layer from the two measured values Δz_m and Δl_m in the following way. Owing to the focal shift from the front surface to the rear one of the m-th layer the marginal ray is subject to parallel shift away from the optical axis as illustrated in Fig.6(d). For simplicity, this figure represents the case in which the objective is moved to the sample. Using the Snell's law, we can write down the following two equations concerning the shift of the marginal ray between the adjacent confocal peaks,

$$d_m \tan\theta_m = \Delta z_m \tan\theta_0 , \tag{1}$$

$$n_m \sin\theta_m = \sin\theta_0 = \mathrm{NA} \qquad (m = 0,1,2\cdots m,\cdots) , \tag{2}$$

where θ_0 and θ_m are an incident angle at the external surface and a refraction angle at the front surface of the m-th layer, respectively, and NA represents the numerical aperture of the objective lens. On the other hand, for the shifts of the interference signal peaks, we have relationship

$$\Delta l_m = n_m d_m - \Delta z_m \qquad (m = 0,1,2\cdots m,\cdots), \tag{3}$$

which means the sum of the decrement of the optical path length Δz_m in air and the increment of the optical path length $n_m \times d_m$ of the m-th layer. These three equations lead to the refractive index n_m and geometrical thickness d_m for the m-th layer [11] such as

$$n_m^2 = \frac{1}{2}\left[\mathrm{NA}^2 + \sqrt{\mathrm{NA}^4 + 4\left(1-\mathrm{NA}^2\right)\left(1+\frac{\Delta l_m}{\Delta z_m}\right)^2}\right], \tag{4}$$

and

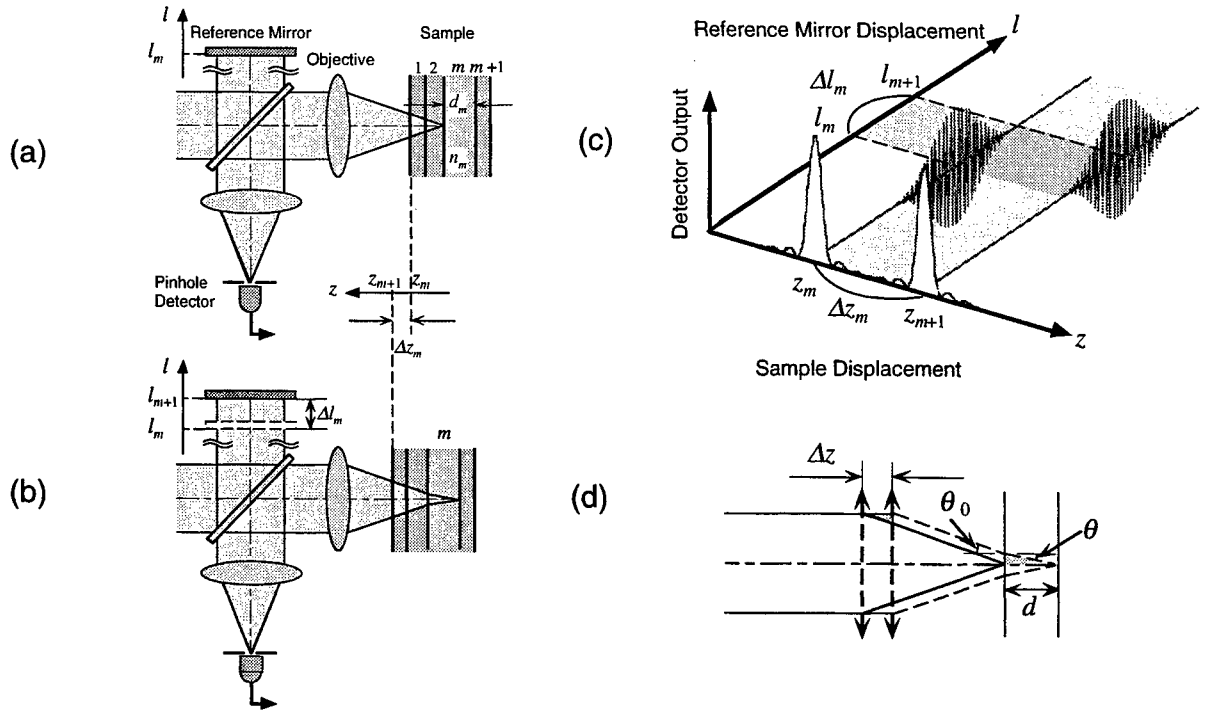


Fig.6 Principle of the separate measurement of refractive indices and geometrical thicknesses of multiple layers. Focusing shift from the front surface (a) to the rear surface (b) of the m-th layer of the sample requires the displacement of the reference mirror of the low-coherence interferometer. (c) Schematic drawing of the detected signals. Signals along the z-axis shows the confocal signal obtained from sample scanning and the localized interference signals along the l-axis shows the interferogram obtained by the low-coherence interferometer. (d) Shift of the marginal ray caused by the focal shift.

$$d_m = \frac{\Delta l_m + \Delta z_m}{n_m} \ . \tag{5}$$

If we repeat the same procedures for other interfaces, we can obtain the refractive index and the geometrical thickness separately for each layer. Using this principle and the apparatus, we have already reported the measurement of several kinds of glass plates and the 13-layered sample [11]. We also applied the method to measurement of a liquid crystal cell and birefringent epoxy-rubber sheet subject to stress [14]. The latter specimen was laterally scanned to yield distribution of geometrical thickness and birefringence.

3.3. Issues of low-coherence interferometry

The above method exhibits three main issues: (a) long measurement time needed for mechanical scanning of both the sample and the reference mirror, (b) the influence of sample dispersion from using a broadband spectral source, and (c) low available light power from a polychromatic point source. In particular, the need for long measurement time becomes critical in case of multiple layers and distribution measurements in lateral directions. These issues are caused mainly by the procedure of low-coherence interferometry. To overcome these drawbacks, we proposed a new method in which the low-coherence interferometry is replaced by wavelength-scanning heterodyne interferometry to be described below [15].

4. METHOD USING CONFOCAL MICROSCOPY AND WAVELENGTH-SCANNING INTERFEROMETRY

4.1. Wavelength-scanning interferometry

In this subsection, we explain the principle of the wavelength-scanning interferometry (WSI). We assume a Michelson interferometer with unbalanced arms as shown in Fig.7. With the wave-number k of the source and the optical path difference D, which is equal to the difference of the two arms lengths l_1 - l_2, the detected intensity is represented by

$$I(k, D) = I_0 + \Delta I \cos kD \ , \tag{6}$$

where I_0 and ΔI are the average and the modulation of the intensity, respectively. The phase of the interference signal is given by

$$\phi = kD \ , \tag{7}$$

which is directly proportional to the OPD. The relationship between the phase and OPD is shown in Fig.7(b) for the wave-number of k_1 and k_2. If the wave-number of the source is continuously tuned from k_1 to k_2, the slope of the line changes. By measuring the phase shift $\Delta\phi$ which is directly proportional to the OPD, the OPD is directly determined by

$$D = \frac{\Delta\phi}{\Delta k} \ . \tag{8}$$

The advantages of the WSI are: (a) absolute measurement of distances that exceeds half an optical wavelength and (b) quick measurement without any mechanical moving part.

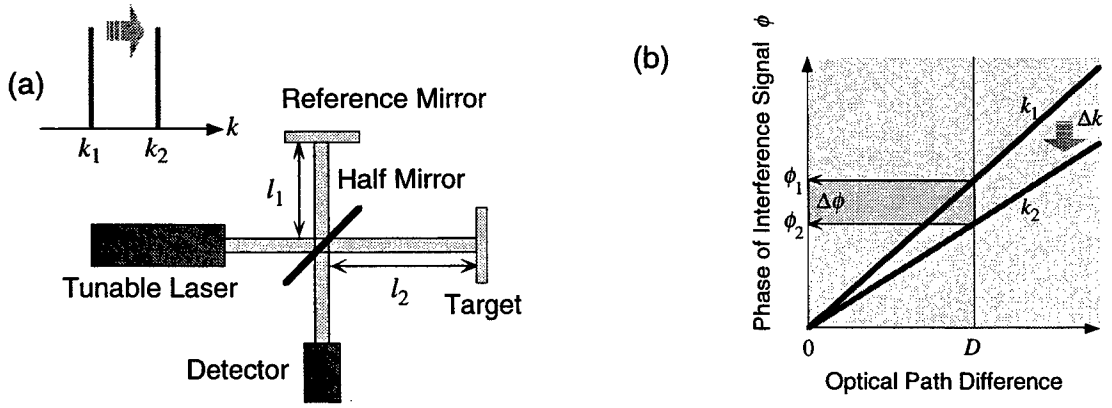


Fig.7 Basic configuration of the wavelength-scanning interferometer. (b) The relationship between the optical path difference and the phase of the interference signal.

4.1.1. Wavelength scanning heterodyne interferometry (WSHI)

In the WSI, if the phase variation exceeds 2π, the interference signal varies periodically as the wavelength is scanned. In this case, the fringe-counting technique or frequency analysis by the Fourier transform method can be used to measure the phase shift. These processings of interference signals have been used for surface profilometry of steps and diffuse surfaces [16-21] and

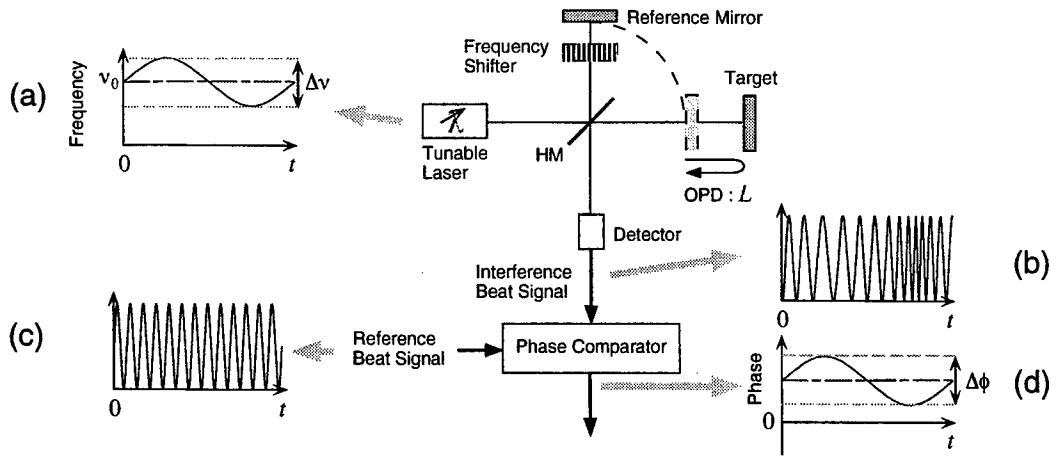


Fig.8 Schematic diagram of wavelength-scanning heterodyne interferometer and signal examples. (a) The frequency of the tunable laser. (b) Detected interference beat signal. (c) Reference beat signal for phase comparison. (d) Phase demodulated signal.

for tomography [22,23]. However, if we use a source with a small wavelength shift such as single-mode laser diode, the phase change is also small. In such case, to attain high resolution, it is necessary to measure the phase shift in high resolution. For this purpose, we use the WSI in combination with the optical heterodyne technique, so-called wavelength-scanning heterodyne interferometer (WSHI) [24,25].

Figure 8 shows the schematic diagram of the wavelength-scanning heterodyne interferometer. By means of sinusoidal modulation of the injection current, the frequency of the tunable laser is subject to sinusoidal modulation around the central frequency ν_0 with the modulation width $\Delta\nu$ as shown in Fig.8(a). For heterodyne detection, the frequency shifter such as an acousto-optic modulator is inserted in one arm. The detected interference beat signal is phase modulated as shown in Fig.8(b). Using the phase comparator, this phase modulation is demodulated by comparison with a reference beat signal (Fig.8(c)) generated in an appropriate way. The amplitude of the output signal of the phase comparator is linearly proportional to the OPD. Therefore, we can determine the OPD by Eq.(8). As the most important advantage of this method, the optical path difference can be measured in real time without any mechanical movement and post-processing of the signals such as frequency analysis.

4.2. Experimental Setup

4.2.1. Optical System

Figure 9 shows an optical setup and its controlling system used for separate measurement of refractive indices and geometrical thicknesses. The optical system is a combination of the wavelength-scanning heterodyne interferometer and the confocal microscope. When the objective lens is focused onto an interface of the sample, the system acts as the WSHI. For a tunable laser, we use a single-mode laser diode (LD, Toshiba Model TOLD-9140) whose optical frequency is modulated sinusoidally at 1 kHz by its injection current modulation driven by a function generator (FG). The temperature of the LD is stabilized at 25.00±0.01°C. For heterodyne detection, the divergent beam from the LD is collimated and enters an optical frequency shifter module (FS; HOYA Model S-210) that is composed of two acousto-optic modulators (AOM1,2) driven at 80.0 and 80.1 MHz, respectively. The light beams leaving the frequency shifter are coaxial and have mutually orthogonal linear polarization states, with a frequency difference of 100 kHz. The beams are expanded and led to the polarization interferometer. One beam transmitted through a polarizing beam splitter illuminates a sample, through a quarter-wave plate and an objective lens (OL; Olympus MDPlan; NA, 0.25; WD, 7 mm). The other is used as a reference beam. Light beams returned from the sample and the reference mirror are superimposed again by the polarizing beam splitter and collected onto the pinhole through a polarizer. A photomultiplier tube positioned behind the pinhole detects the interference light.

4.2.2. Signal Processing Means

The beat signal detected by the PMT is subject to both amplitude and phase modulation in synchronization with the LD

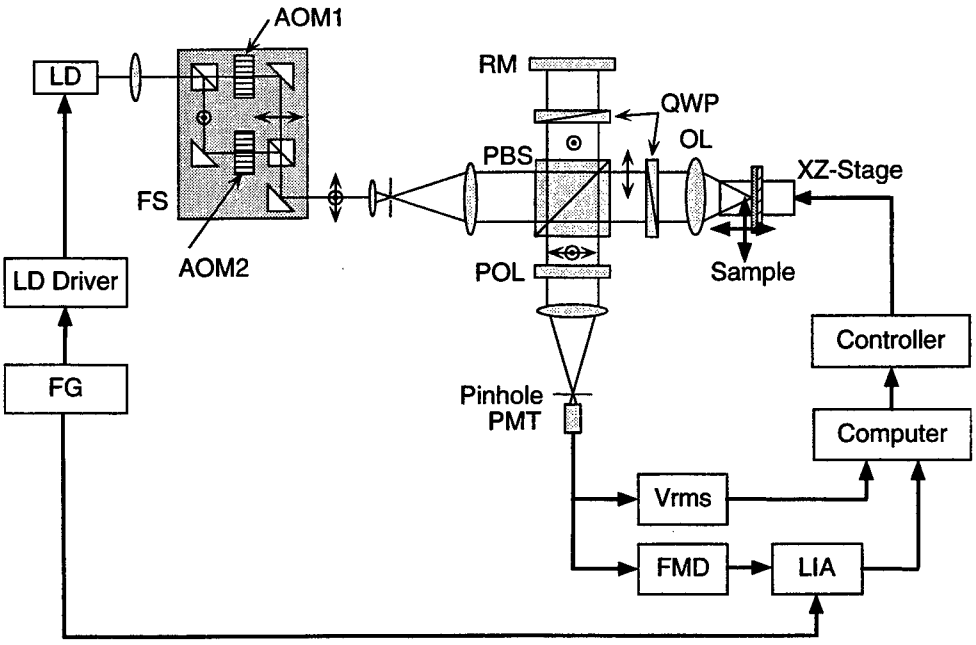


Fig.9 Experimental configuration for separation of the refractive index and geometrical thickness. LD, laser diode; FS, frequency shifter; AOM, acousto-optic modulator; PBS, polarizing beam splitter; QWP, quarter-wave plate; OL, objective lens; RM, reference mirror; POL, polarizer; PMT, photomultiplier tube; FG, function generator; FMD, frequency modulation-demodulator; LIA, lock-in amplifier.

modulation signal. Amplitude modulation is caused by the sinusoidally intensity modulation of LD. The amplitude of the beat signal also varies with the sample movement introduced for confocal microscopy. We need to separate the phase modulation signal which corresponds to the WSHI signal and the amplitude modulation which represents the confocal signal. To extract the confocal signal, we measure the amplitude by means of a voltmeter (Analog Devices, AD736) after high-pass filtering. Output signals of the voltmeter is stored in the computer.

To measure the WSHI signal, we use a FM-demodulator based on a phased-locked loop (NE565). First, the beat signal is debiased and converted to square-wave for suppressing amplitude variation. Second, the signal is led to the FM-demodulator whose output voltage is linearly proportional to the frequency deviation from its free-running frequency. By delivering its output to a lock-in amplifier (LIA), we measure both amplitude and phase of the output signal having the same frequency as the LD modulation with the reference signal from the function generator. These measured values are then transferred to the computer.

The sample is mounted on a two-dimensional (x-z) translation stage which is controlled by the computer. For separate measurement of refractive index and geometrical thickness, we measure the confocal signal and the WSHI signal simultaneously for each sample position along the optical axis (z-axis). For distribution measurement, the same procedure is repeated after the sample position is moved in the lateral direction.

4.3. RESULTS AND DISCUSSIONS

4.3.1. OPD measurement

To apply the WSHI, we need to know the amount of the frequency shift of the LD in the running state. However, quick measurement of the optical frequency is difficult by a wavelength meter. We estimate the value by measuring the relation between the phase shift of the interference signal and the OPD by moving a plane mirror in the interferometer without the objective lens. Figure 10 shows the measured amplitude and phase of the output signal of the FM-demodulator. The amplitude is linearly dependent on the mirror displacement. A point of the zero amplitude corresponds to the zero OPD. Across this point, the phase is inverted by 180 degrees. Therefore, from both the amplitude and the phase variations, we can discriminate the sign of the OPD. However, the linearity of the amplitude breaks at the mirror position exceeding 18.5 mm. In this region, the FM-

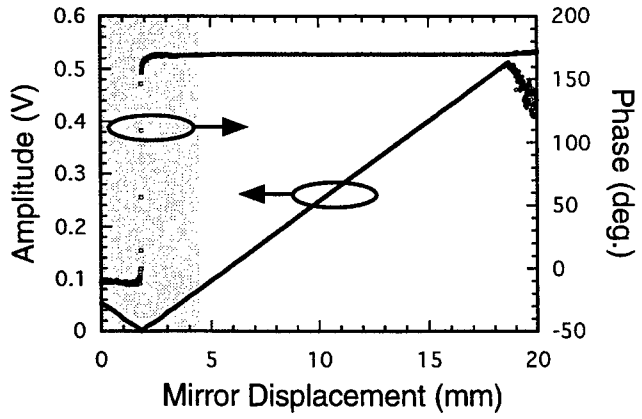


Fig.10 Measured amplitude (circles) and phase (squares) of the output signal of the FM demodulator against mirror displacement. The shaded area shows the measurable range using previous phase meter.

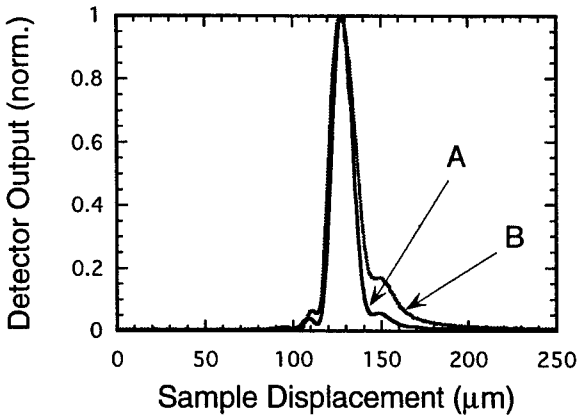


Fig.11 Axial responses of confocal microscope obtained from the mirror displacement. (A) Obtained from the ordinary confocal setup. (B) Obtained from the proposed setup.

demodulator cannot respond to the speed of the phase change of the beat signal. Thus, maximum range of OPD to be measured is ±34 mm.

On the other hand, shaded zone in Fig.10 represents measurable range when the phase comparator is used. This is limited by that the phase comparator is able to measure angle up to 180 degree. In this case, the maximum range to be measured is only 5.4 mm for the OPD including the sign. Hence one of the advantage of using the FM-demodulator is the extension of the measurable OPD range. Another advantage is no need of the reference signal. This means that the initial phase adjustment that is required in the case of the phase comparator is not necessary. Therefore, the measurement with continuous movement of the sample is possible. The resolution estimated from the standard deviation over five repeated measurements was approximately 10 μm.

4.3.2. Axial response

Figure 11 shows the axial response of the confocal microscope. Both the amplitude of the beat signal resulting from the proposed method and the intensity without the reference light resulting from the ordinary confocal microscope are displayed. The FWHM of each graph is equal to 17 μm and 13 μm, respectively. The broadening of the former is explained by considering that the former is based on the amplitude profile instead of the intensity profile. This can be verified by the fact that the square of the former coincided with the latter. The reason of the asymmetry of the amplitude graph is the wave-front aberration caused by acousto-optic modulators. However, the accuracy of the peak position determination is about 1 μm for both graphs.

4.3.3. Separate measurement of refractive index and geometrical thickness

To demonstrate the separate measurement of the refractive index and geometrical thickness, we employed several glass plates. As typical examples of signals, Fig. 12 shows the observed signals from the glass plates (BK7, 2-mm thickness). The confocal signal (Fig.12(a)) results from the sample translation. The peak positions 1 and 2 represent focusing states at the front and the rear surfaces, respectively. The WSHI signal, namely, the amplitude of the output signal of the FM demodulator is plotted in Fig.12(b). Both the confocal signal and the WSHI signal are obtained simultaneously from single scanning of the sample. The WSHI signal is diminished outside the confocal peaks because no interference signal appears there owing to extinction of the light returned from the sample. The derivation of refractive index and geometrical thickness is performed in the following way. First, we measure the interval Δz of the two confocal peaks. Second, from the FM demodulator output, the corresponding optical path length change Δl is measured with reference to the OPD-calibration results as shown in Fig.10. Finally, the refractive index and geometrical thickness are calculated by using Eqs. (4) and (5), respectively.

Table 1 summarizes the results from measurement of other samples. The standard deviations over five repeated measurements are also shown. In the reference values, the refractive indices at the LD wavelength of 690 nm were calculated from Cauchy's equation, whereas the geometrical thicknesses were measured with a micrometer gauge. For each sample the reference values

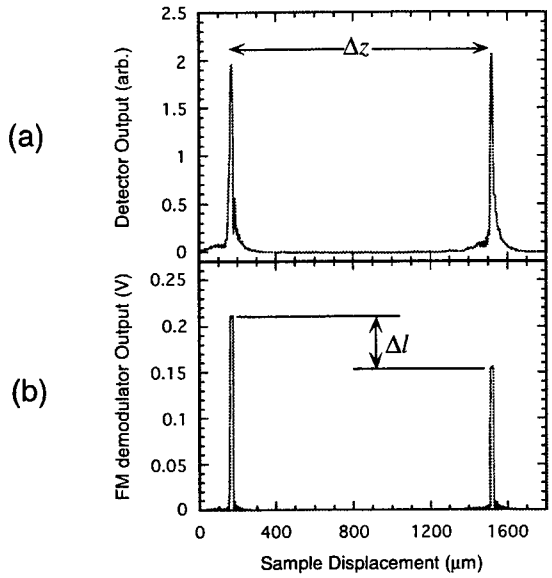


Fig.12 (a) Confocal signals and (b) FM demodulator output obtained from sample scanning.

Table 1 Measured results for the refractive index and the geometrical thickness with reference values.

Sample	Refractive Index		Thickness (μm)	
	n (meas.)	*n* (catalog)	*d* (meas.)	*d* (catalog)
Slide Glass	1.509±0.010	1.523[a]	1116±7	1092[b]
	1.517±0.003		1103±2	
BK7 Glass	1.507±0.009	1.513	1092±8	1084[b]
	1.513±0.002		1097±2	
Fused Silica Glass	1.468±0.014	1.455	1043±10	1020[b]
	1.454±0.002		1031±3	

[a]Measured at 588 nm. [b]Measured using a micrometer-gauge.

Measured using the previous method (confocal + low-coherence interferometer)

and the results from the previous method using low-coherence interferometry are also cited for comparison. Agreement between the measured and the reference values is within an error of 2.1 %. However, the dispersion is ten times larger than that obtained with our previous method using low-coherence interferometry. The reason for this is that the accuracy of resultant refractive index and geometrical thickness values is limited by the measurement accuracy of the change in OPD Δl. At present, the accuracy of Δl is 10 μm in the proposed method and 1 μm in the previous method. These values are determined by the WSHI and LCI, respectively. In contrast, the accuracy of Δz obtained from confocal microscopy is 1 μm in both methods. The difference of the measurement accuracy of Δl causes the larger dispersion seen in the current results. Hence, to improve the accuracy, the extension of the wavelength tuning range is necessary.

4.3.4. Reconstruction of geometrical cross-section

For distribution measurement that has been greatly accelerated by the above method, we prepared a wedge sample as shown in Fig.13. It is composed of the two sheets of glasses cemented with the same sheet that is inserted for a spacer at one end. The angle of the apex is less than 1 degree. We performed the separate measurement of the refractive index and the geometrical thickness at various positions of the sample. The sample was laterally shifted at a step of 1 mm over a distance of 100 mm. At

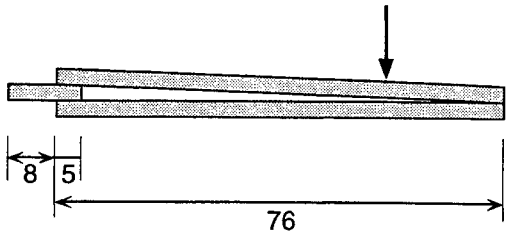


Fig.13 Cross-sectional composition of the prepared wedge sample (unit : mm). The geometrical thickness of of the glass plates is about 1100 mm.

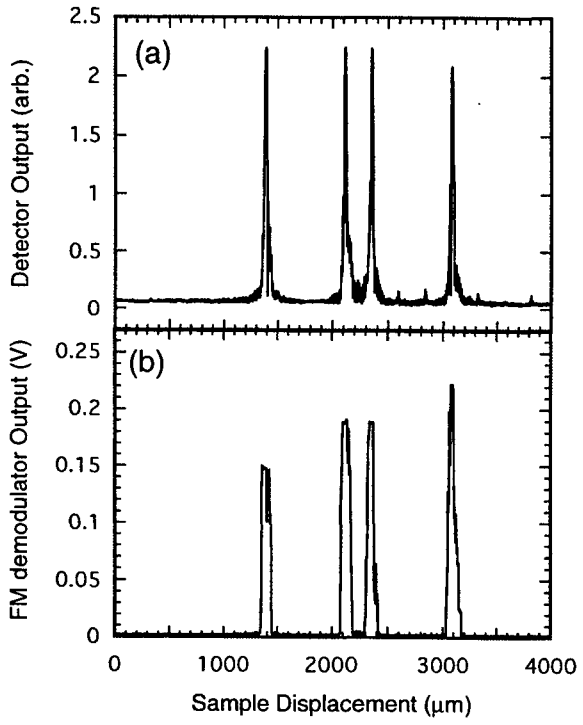


Fig.14 (a) Confocal signals and (b) corresponding FM demodulator output obtained from the point indicated in Fig.13

each lateral position, the confocal signal and the WSHI signal were measured simultaneously at 4000 points with an interval of 1 μm. As a typical example of signals, both the confocal signal and the WSHI one are shown in Figs.14 (a) and (b), respectively. These signals were acquired at the lateral position denoted by the arrow in Fig. 13. Each confocal peak appeared when the focal point coincided with each of the sample interface. The corresponding output of the FM-demodulator shows the OPD between the optical path from the BS to the focal point and that of the reference arm. No change of the output of the FM-demodulator for the focal shift across the air layer is observed because no optical path length change takes place in this layer. Figure 15 shows a two dimensional distribution of confocal signals. Brightness of each pixel is coded according to height of confocal signal. The bright lines correspond to the confocal peaks, that is an optical cross-sectional image. This brightness is attenuated along the sample depth by internal refraction. Since the lights reflected from the interface of glass and spacer are weak, the cross-section for glass-spacer-glass represented by the arrow in Fig. 15, looks like a single layer.

Next, reconstructed cross-sectional image is shown in Fig.16. It was obtained by the following steps. In the first step, we calculate the refractive index and geometrical thickness from the intervals of confocal peaks and corresponding OPD for each

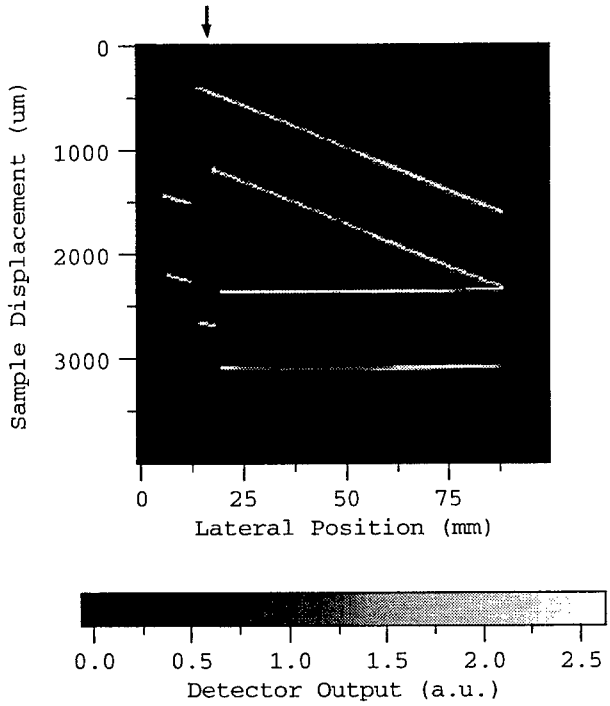


Fig.15 Cross-sectional image obtained from the confocal signals. The bright spot corresponds to the focusing at the interfaces of the specimen.

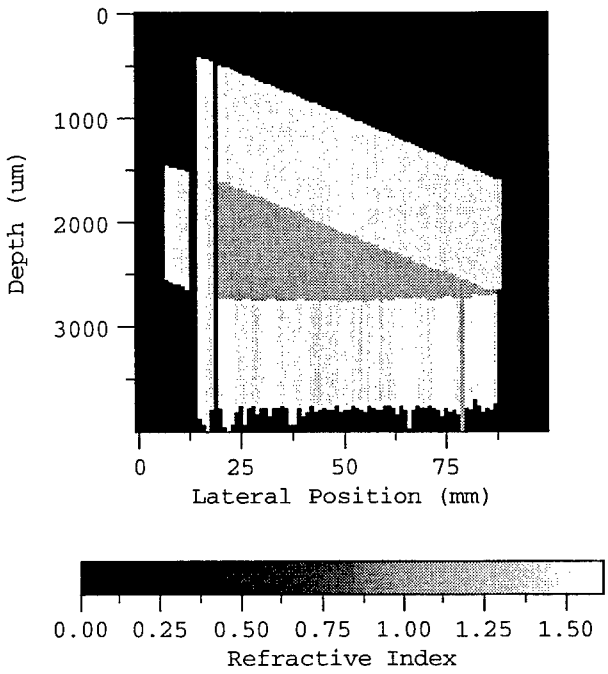


Fig.16 Reconstructed cross-sectional image.

lateral position. In the second step, the position of the front surface was derived from the confocal peak positions, which corresponds to the surface profile of the sample on the objective side. In the third step, the shape of the second surface was obtained from addition of the measured geometrical thickness to the front surface profile. Repeating the same procedure leads to the display of cross-sectional image. In this case, the cross-section means not the optical but the geometrical cross-section. In addition, each layer at each lateral position in Fig.16 is grey coded according to the refractive indices. The coded values represent the difference from 1, that is, the index of air. Hence, the refractive index of the background is put to be zero. The profile of the deepest surface exhibits more irregularity. The reasons for this might be, first, lens aberration caused by transmission through the front layers and, second, the influence of the light reflected from other surfaces. At the both edges, the calculated refractive indices showed singular values. The reason for this is that one of the signals needed for calculation of the refractive index and geometrical thickness could not be deduced as seen from Fig.15. Since the present system works as both a surface profiler using a confocal microscope and a sensor for the refractive index and geometrical thickness for transparent specimens, we can obtain the geometrical cross-sectional image.

6. CONCLUSION

After reviewing several methods for separate measurement of refractive index and geometrical thickness of transparent plate, we proposed a new quick measurement scheme using a combination of confocal microscopy and wavelength-scanning heterodyne interferometry. It requires only single axial movement of a specimen by introduction of the FM-demodulation technique applied to wavelength-scanning interferometry. Basic performances in detection of confocal peak position and the optical path difference have been examined. We performed separate measurement with glass sheets. Finally, by introducing lateral scanning of samples we also could display geometrical cross-section of multi-layered objects.

REFERENCES

1. T. Tsuruta and Y. Ichiyama, "Accurate measurement of lens thickness by using white-light fringes", *Proc. ICO Conf. Opt. Methods in Sci. and Ind. Meas.* pp. 369-372, 1974 .

2. B. L. Danielson and C. Y. Boisrobert, "Absolute optical ranging using low coherence interferometry", *Appl. Opt.* **30** pp. 2975-2979, 1991.
3. H. E. Bennet, J. M. Bennett, *Physics of Thin Films* 4 (edited by G. Hass, R. E. Thun, Academic Press), pp. 1-96 , 1967 .
4. C. J. R. Sheppard, T. J. Connolly, J. Lee, and C. J. Cogswell, "Confocal imaging of a stratified medium", *Appl. Opt.* **33** pp. 631-640, 1994.
5. T. Fukano and I. Yamaguchi, "Measurement of layer thickness by a laser confocal microscope", *Proc. of 15th Meeting on Lightwave Sensing Technology* pp. 91-98, 1995 .
6. P. A. Flournoy, R. W. McClure, and G. Wyntjes, "White-light interferometric thickness gauge", *Appl. Opt.* **11**, pp. 1907-1915, 1972.
7. W. V. Sorin and D. F. Gray, "Simultaneous thickness and group index measurement using optical low-coherence reflectometry", *IEEE Photo. Tech. Lett.* **4**, pp. 105-107, 1992.
8. G. J. Tearney, M. E. Brezinski, J. F. Southern, B. E. Bouma, M. R. Hee, and J. G. Fujimoto, "Determination of the refractive index of highly scattering human tissue by optical coherence tomography", *Opt. Lett.* **20**, pp. 2258-2260, 1995 .
9. N. Sakai, S. Ohkubo, A. Takayanagi, and N. Umeda, "Simultaneous measurement of thickness and refractive index of crystalline quartz using low coherence interferometry", *Proc. of 20th Meeting of Japan Society for Laser Microscopy*, pp. 59-65, 1997.
10. S. Y. El-Zaiat, "Measuring the thickness and refractive index of a thick transparent plate by an unexpanded laser beam", *Opt. Laser Tech.* **29**, pp. 63-65, 1997.
11. T. Fukano and I. Yamaguchi, "Simultaneous measurement of thicknesses and refractive indices of multiple layers by a low-coherence confocal interference microscope", *Opt. Lett.* **21**, pp. 1942-1944, 1996.
12. M. Ohmi, T. Shiraishi, H. Tajiri, and M. Haruna, "Simultaneous measurement of refractive index and thickness of transparent plates by low coherence interferometry", *Opt. Rev.* **4**, pp. 507-515, 1997.
13. M. Haruna, M. Ohmi, T. Mitsuyama, H. Tajiri, H. Maruyama and M. Hashimoto, "Simultaneous measurement of the phase and group indices and the thickness of transparent plates by low-coherence interferometry", *Opt. Lett.* **23**, pp. 966-968, 1998.
14. I. Yamaguchi and T. Fukano, "Measurement of thicknesses and refractive indices by low-coherence interferometric microscope", *Optics and Optoelectronics* Vol. 2 (edited by O. P. Nijhawan, A. K. Gupta, A. K. Musla, and K. Singh), pp. 1081-1093, 1998.
15. T. Fukano and I. Yamaguchi, "Separation of measurement of the refractive index and the geometrical thickness by use of a wavelength-scanning interferometer with a confocal microscope", *Appl. Opt.* **38**, pp. 4065-4073, 1999.
16. M. Takeda and H. Yamamoto, "Fourier-transform speckle profilometry: three-dimensional shape measurements of diffuse objects with large height steps and/or spatially isolated surfaces," *Appl. Opt.* **33**, pp. 7829-7837, 1994.
17. T. H. Barnes, T. Eiju and K. Matsuda, "Rough surface profile measurement using Spekle Optical Frequency Domain Reflectometry with an external cavity tunable diode laser," *Appl. Opt.* **103**, pp. 93-100, 1996.
18. S. Kuwamura and I. Yamaguchi, "Wavelength scanning profilometry for real-time surface shape measurement," *Appl. Opt.* **36**, pp. 4473-4482, 1997.
19. H. J. Tiziani, B. Franze and P. Haible, "Wavelength-shift speckle interferometry for absolute profilometry using a mode-hop free external cavity diode laser," *J. Mod. Opt.* **44**, pp. 1485-1496, 1997.
20. M. Suematsu and M. Takeda, "Wavelength-shift interferometry for distance measurements using the Fourier transform technique for fringe analysis," *Appl. Opt.* **30**, pp. 4046-4055, 1991.
21. I. Yamaguchi, Akihiro Yamamoto and S. Kuwamura, "Speckle decorrelation in surface profilometry by wavelength scanning interferometry" *Appl. Opt.* **37**, pp. 6721-6728, 1998.
22. F. Lexer, C. K. Hitzenberger, A. F. Fercher, and M. Kulhavy, "Wavelength-tuning interferometry of intraocular distances," *Appl. Opt.* **36**, pp. 6548-6553, 1997.
23. S. R. Chinn, E. A. Swanson and J. G. Fujimoto, "Optical coherence tomography using a frequency-tunable optical source," *Opt. Lett.* **22**, pp. 340-342, 1997.
24. H. Kikuta, K. Iwata and R. Nagata, "Distance measurement by the wavelength shift of laser diode light," *Appl. Opt.* **25**, pp. 2976-2980, 1986.
25. E. Fischer, E. Dalhoff, S. Heim, U. Hofbauerr and H. Tiziani, "Absolute interferometric distance measurement using a FM-demodulation technique," *Appl. Opt.* **34**, pp. 5589-5594, 1995.

Beam shaping with component of cone axicons in confocal system

Weijian TIAN*[a], Zhihua DING[b], Zhengkang BAO[b], Zukang LU[b], Zhengdong LIU[c]

[a] Institute of Technical Physics, Zhejiang University of Technology, Hangzhou, Zhejiang, 310014, P.R.China

[b] State Key Lab of Modern Optical Instrumentation, Zhejiang University, Hangzhou, Zhejiang, 310027, P.R.China

[c] Modern Optics Research Center, Zhejiang University, Hangzhou, Zhejiang, 310027, P.R.China

ABSTRACT

In this paper, a novel beam shaping configuration for confocal imaging system is presented, which relies on the reversal effect of axicons-component on along with the Gaussian distribution of laser beam. The Gaussian peak of incident light coming into the one side of such component at the central part, may go out from the other side of component near the peripheral part, conversely, the skirt of Gaussian beam coming into the peripheral part may go out from the central part. Thus, the beam shaping is obtained with the component of cone axicons and therefore the superresolution can be realized in a confocal imaging system. Meanwhile, the energy loss is much diminished and the optical power deserted hitherto is recovered.

Keywords: Beam shaping, confocal imaging system, component of cone axicons, Gaussian distribution, optical superresolution, spatial filters, weighted skirt, annular pupil function, effective numerical aperture, diminishing energy loss

1. INTRODUCTION

The Gaussian distribution of a laser beam is not always the most optimal distribution to utilize in applications. It is required and useful with a uniform illumination in many laser applications, which include holography and optical processing, semiconductor processing, laser fusion and laser projection printing, optical data processing and optical gauging etc. However, the Gaussian intensity distribution does not provide uniform illumination and suffer loss of optical power when the beam is expanded to obtain uniform illumination. For these reasons, it is necessary to manipulate the intensity distribution of laser beam, i.e. beam shaping [1, 2, 3].

*. Correspondence: E-mail: tianwj@moi-lab2.zju.edu.cn; Telephone: (0086) 571 832 0370(or 795 1194); Fax: (0086) 571 795 1617

The beam shaping of detecting light is also important in a confocal imaging system, because it can directly affect some main characteristics of confocal system like lateral resolution, depth discrimination and energy efficiency etc. Various type of spatial filters such as annular filter, leaky annular filter, double-point filter, double-slit filter, and even "Frieden" filter were implemented for optimizing detecting beam properties[4, 5, 6.] . Unfortunately, owing to inserting of spatial filters, a lot of the incident energy is wasted due to absorbing of the passive filters and diffraction outside the specified Rayleigh circle. Although great efforts were taken to find the diffraction pattern and corresponding pupil function, with the property that the maximum possible fraction of the total energy in the pattern is contained within the specified Rayleigh limit, there still exists the problem of energy wasting due to the passive absorbing filters. Furthermore, not all these spatial filters are easy to fabricate with high accuracy and therefore nevertheless degrade the pursued results.

In this paper, a novel beam shaping configuration is presented, which relies on the effect of axicons on along with the distribution of Gaussian beam itself. Beam shaping is obtained without any spatial filter, and hence the energy loss is diminished and the optical power deserted hitherto is recovered. Because the ultimate aim of beam shaping in a confocal imaging system is the realization of optical superresolution, the confocal experimental system with the component is also discussed briefly.

2. PRINCIPLE

2.1. The Intensity distribution of laser beam

In general, the profile of fundamental (TEM_{00}) mode laser beam is in a Gaussian distribution form. Although it bears a resemblance to plane wave in some aspects, Gaussian beam is an uneven wave. Gaussian intensity distribution is not a uniform profile, and its phase front is also not a plane. For a rotationally symmetric system, mathematically Gaussian beam distribution on the plane perpendicular to the optical axis is given by amplitude $A(\rho)$

$$A(\rho) = W \exp(-\varepsilon^2 \rho^2), \tag{1}$$

where W is a normalized factor and ρ is the normalized radial coordinate of objective aperture. $\varepsilon^2 = \frac{1}{2\sigma^2}$, σ denotes the distance between the central point of maximum intensity (I_m) and the position of $I_m \cdot e^{-1/2}$.

According to the deduction of Reference 7. , the focal spot intensity I_0 related to ε can be expressed as

$$I_0 = C\left[1 - \exp\left(-\varepsilon^2/2\right)\right]/\varepsilon^2, \tag{2}$$

where C is a constant. The relationship curve given in Eq. (2) is shown in Fig.1.When the value of ε^2 is very small, the focal spot intensity I_0 is weak because the most part of incident light is filtered out by the objective aperture. When the value of ε^2 is very large, the I_0 is also weak because the objective numerical aperture is not utilized in the most effective case.

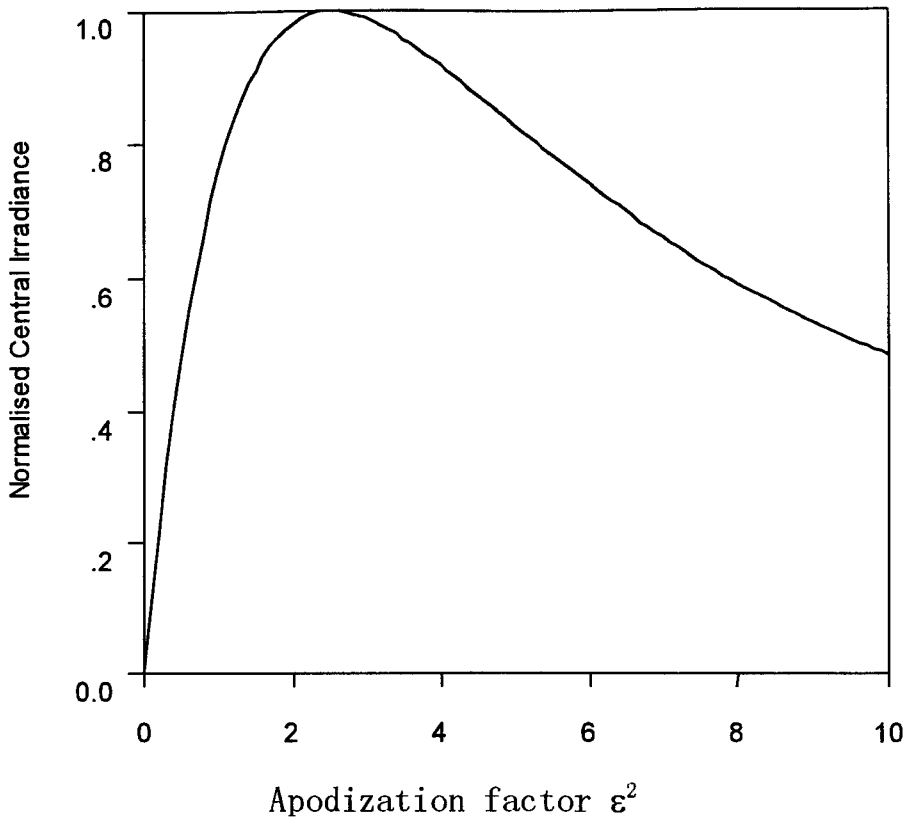


Fig.1 Central irradiance of a focused spot as a function of apodization factor ε^2.

2.2. The property of component for beam transformation

In order to use the incident light energy as much as possible, the property of component of cone axicons, which is as an energy redistribution element, is investigated in the process of beam shaping.

The component as shown in Fig.2 plays a key role in the manipulating process. There are the same wedge angles within the two axicons and is an appreciate distance between them. The peak of incident Gaussian beam coming into the central part of first cone axicon, may go out from near the peripheral part of second cone axicon, and the skirt of Gaussian beam coming into the peripheral part goes out from the central part.

The relation between the radius R of input ray and the radius r of output ray is

$$R + r = \rho_0 . \tag{3}$$

Here ρ_0 is the radius for the two cone axicons, ρ_0=D/2.

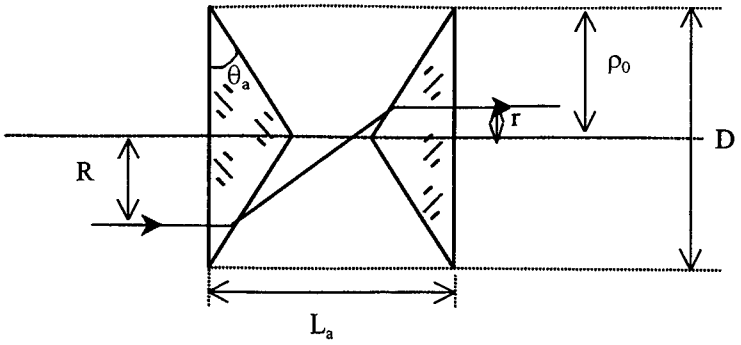


Fig.2 Component of two cone-axicons.

For realizing rays inversion expressed as Eq.(3), the geometrical parameters are required to satisfy the relationship as

$$L_a / \rho_0 = tg\theta_a + ctg[\sin^{-1}(n\sin\theta_a) - \theta_a], \tag{4}$$

where n is the refractive index of cone axicons. Assume that the absorbed and reflected energy with the component can be neglected, so that the energy of input ring belt (from R_1 to R_2) should be equal to the one of output ring belt (from r_1 to r_2) on the law of conservation of energy

$$\int_{R_1}^{R_2} 2\pi R f_i(R) dR = \int_{r_1}^{r_2} 2\pi r f_o(r) dr, \tag{5}$$

$f_i(R)$ and $f_o(r)$ are the intensity distribution functions on the input plane and the output plane, respectively. The energy relationship between input pupil and output pupil is shown as in Fig.3.

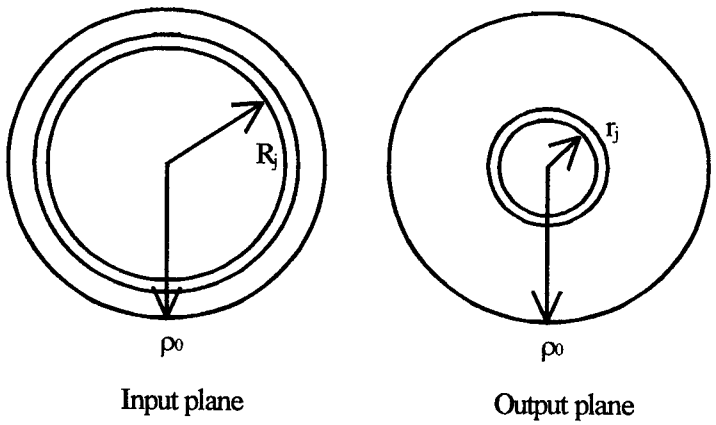


Fig.3 Mapping of energy in annulus of input plane into annulus of the output plane.

If the transmissivity function of component is defined as

$$\eta = f_o(r) \Big/ f_i(R), \tag{6}$$

and Eq.(5) is integrated in the $R_1 \rightarrow R_2$ limit, then

$$\eta = \frac{R}{r} = (\rho_o - r) \Big/ r = R \Big/ (\rho_o - R). \tag{7}$$

The curve corresponding to Eq.(7) is shown as in Fig.4.

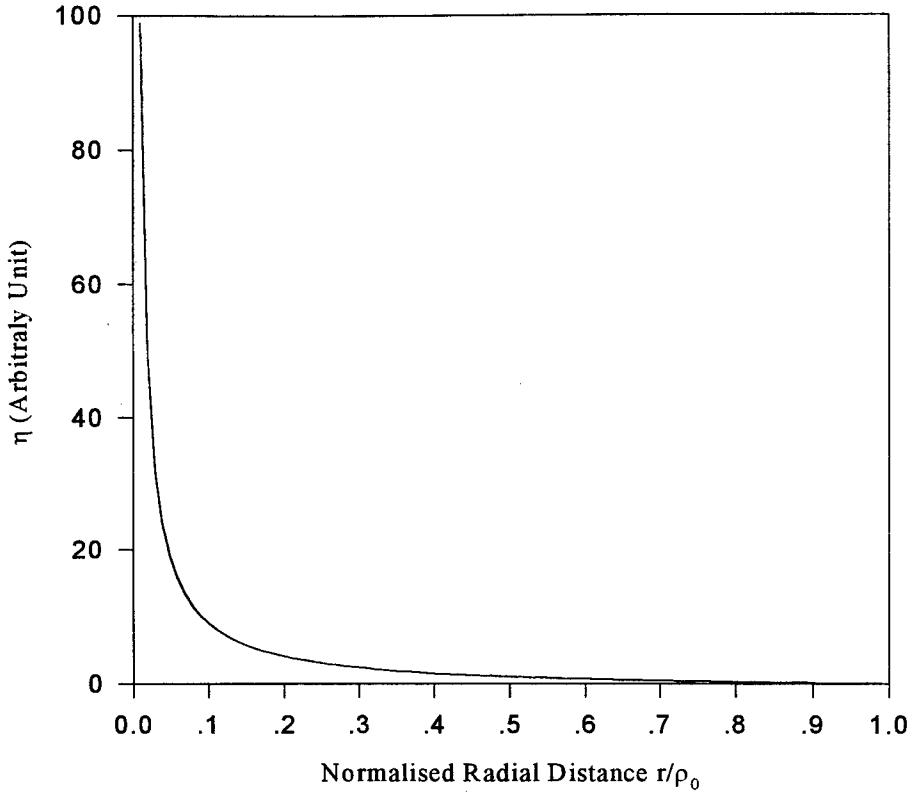


Fig.4 Radial distribution of transmission modulation coefficient through a component.

3. RESULT AND DISCUSSION

If the incident beam of fundamental mode is a Gaussian distribution expressed in Eq. (1), the intensity distribution after passing through the component, becomes as

$$f(\rho') = \eta f_G(\rho) = K[(\rho_o - \rho') / \ \rho'] \exp[-\varepsilon^2 (\rho_o - \rho')^2], \tag{8}$$

where ρ' denotes the normalized radial coordinate on the objective plane after the component (see Fig.6).

The curves of intensity distribution along the radial before and after through the component are shown as in Fig.5. It is quite evident that the weighted center of intensity distribution has turned into the weighted skirt. Therefore, the beam shaping is obtained with the component of cone axicons. Meanwhile, all of energy can almost pass trough the optical system because no beam ray is cut off.

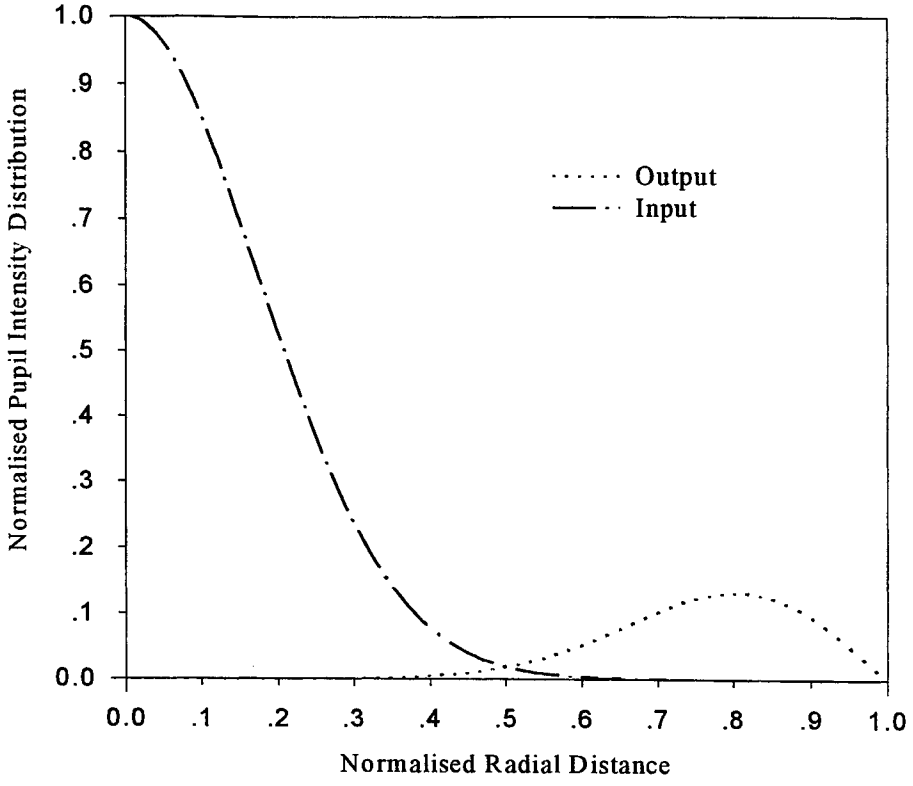


Fig5 Input and output irradiance distribution for TEM_{00} mode Gaussian beam through the component of two cone axicons.

Fig.6 illustrates the primary optical concept of component for realizing beam transformation and superresolution.

It is seen that the phase redistribution does not occur with the component. Therefore, the effective pupil function for the optical configuration as shown in Fig.6 can be expressed as

$$P(\rho') = \sqrt{f(\rho')}. \tag{9}$$

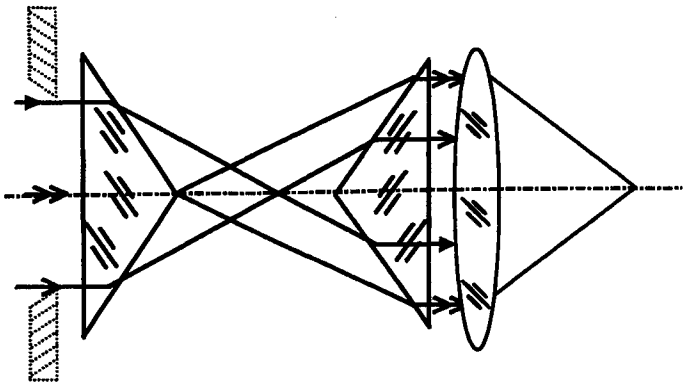


Fig.6 Schematic diagram for superresolution through the optical configuration with the component.

The pupil function with the weighted skirt can improve the effective numerical aperture of objective lens. As a result after beam shaping, the optical superresolution of confocal imaging can be realized without any spatial filter.

If the different size of pinhole mask placed in front of the component is used as shown in Fig.6 (dashed lines), it is obvious that can result in the effect of annular lens with the central obstruction of different radius. However, because of the skirt obstruction for the Gaussian beam in this case, unlike the central obstruction in general, the energy loss is much diminished. Meantime, such a component of cone axicons is more effective in inverting the distribution of incident beam than the double-concave-cone lens, thus less in longitudinal dimension.

The problem of side lobes introduced by the pupil function with the weighted skirt, can be resolved in a confocal system. The confocal experimental system with the component of cone axicons for energy redistribution is shown as in Fig.7. In this confocal system, used the redistribution component instead of the pupil filter, the main advantage is to overcome the drawback of energy loss and the difficulty to fabricate filters accurately.

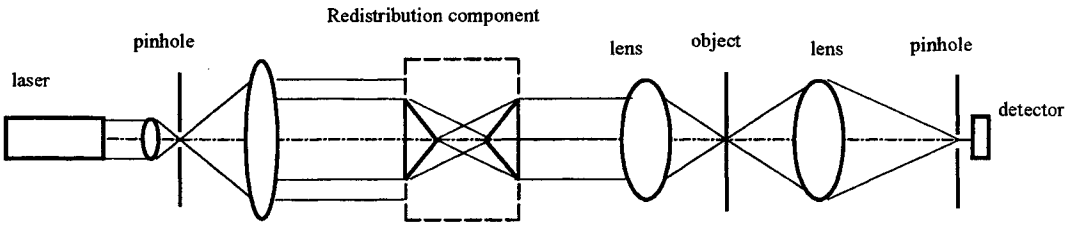


Fig.7 Confocal experimental system with the energy redistribution component of two cone-axicons.

What is more important is that, by extending the distance between two axicons and enlarging the aperture of second axicon correspondingly, annular illumination with different obstructions is formed, which in reality acts as an apodized annular filter. Finally, if the second cone axicon is replaced by a diffraction optical element fabricated in combination with the succeeding refractive focusing lens, an achromatic system is realized because these two elements have opposite chromatic aberration which can compensate for each other.

4. CONCLUSION

It is well known that the optical superresolution is the one of important and unique properties of confocal system. The investigation in the 70' had indicated that transverse resolution in confocal microscopy is 1.4 times as large as that in a conventional microscope. Consequently, a number of researchers analyzed the theoretic principles on the optical superresolution in confocal imaging, and some others made efforts to realize the superresolution in practice imaging system as mentioned in the Section 1, which improved the resolution by utilizing various type of pupil filters.

Compared with the pupil filter approaches, the main development with the component of cone axicons is that, it is possible to realize the superresolution after beam shaping, while diminishing energy loss and recovering hitherto deserted optical power.

The research work reported here is supported by the National Natural Science Foundation of China (No. 69808002) and the Opening Project of the State Key Lab of Modern Optical Instrumentation (No. LMOI-9815).

REFERENCES

1. W. B. Veldkamp, "Laser beam profile shaping with interlaced binary diffraction gratings," *Appl. Opt.* **21**, pp.3209-3212, 1982.
2. F. S. Roux, "Intensity distribution transformation for rotationally symmetric beam shaping," *Appl. Opt.* **30**, pp.529-536, 1991.
3. C. C. Aleksoff, K. K. Ellis, and B. D. Neagle, "Holographic conversion of a Gaussian beam to a near-field uniform beam," *Opt. Eng.* **30**, pp.537-543, 1991.
4. C. J. R. Sheppard and M. Gu, "Improvement of axical resolution in confocal microscopy using annular pupil," *Opt. Comm.* **84**, pp.7-13, 1991.
5. C. K. Sieracki, E. W. Hansen, "A leaky annular pupil for improved lateral resolution in confocal fluorescence microscopy," *Proc. SPIE* **2184**, pp.120-126, 1994.
6. Z. Ding, G. Wang, M. Gu, Z. Wang and Z. Fan, "Superresolution using an apodization film in a confocal setup," *Appl. Opt.* **36**, pp. 360-363, 1997.
7. A. B. Marchant, *Optical Recording*, pp.101-110, Addison-Wesley Publishing Company, New York, 1990.

SESSION 10

MOEMS

Analysis of Mach-Zehnder Interferometric micro-opto-electro-mechanical (MOEM) Pressure Sensor

T.Srinivas , Prasant Kumar Pattnaik, T.Badri Narayana, A.Selvarajan
Department of Electrical Communication Engineering
Indian Institute of Science, Bangalore-560 012, INDIA.
Tel: +91 80 3092852; Fax: +91 80 3340563
Email: srinu@ece.iisc.ernet.in

ABSTRACT

Combination of Integrated Optics and micro-machining technologies offer immense potential for sensor applications. Small mechanical deformations can often produce considerable changes in optical properties of devices resulting in drastically improved sensitivities. Here we propose and analyze a novel pressure sensor consisting of integrated optic Mach-Zehnder interferometer whose sensing arm is fabricated on a silicon micro-machined diaphragm. The analysis consists of determining the changes in optical output corresponding to the diaphragm deflections due to impressed pressure. Dynamical equations of motion are solved and resulting displacement fields are related to refractive index and optical path length changes of the Mach-Zehnder interferometer. Results can be used to obtain the changes in sensitivity due to changes in path length and refractive index variations. The analysis can easily be applied to other MOEM sensor devices like those consisting of micro-machined vibrating cantilevers and bridges controlling optical waveguides, directional couplers or multi-mode-multi-waveguide structures.

Keywords: Integrated Optics, micromachined diaphragm, Mach-Zehnder Interferometer

1. INTRODUCTION

Micro-opto-electro-mechanical (MOEM) devices and systems, based on the principles of integrated optics and micromachining technology on Silicon have immense potential for sensor applications[1]. Several MOEM devices such as mirrors, beam splitters, lenses and switches, have been reported [2]. Micro-mechanical components like diaphragms can be used to support optical waveguiding structures by forming an additional optical layer. Thus optical device properties can be influenced by the mechanical properties. Employing optical techniques have important advantages such as functionality, large bandwidth and higher sensitivity. Miniaturization brings in advantages like small size, lightweight, ruggedness, batch fabrication, free from alignment problems and so on.

Pressure sensing using micromachined structures have utilized the changes induced in either the resistive or capacitive properties of the electro-mechanical structure by the impressed pressure [3]. Integrated optical pressure sensors can utilize the changes to either the amplitude, phase, refractive index profile, optical path length, or polarization of the lightwave by the external pressure. Here we consider a pressure sensor incorporating an integrated optical Mach-Zehnder interferometer formed on a Silicon diaphragm. First the configuration and fabrication aspects are described in the next section. Analysis of the mechanical properties of the rectangular diaphragm and integrated optical Mach-Zehnder interferometer are given in section 3. . Section 4 consists of results of simulation and their interpretation. Concluding section discusses several issues involved in the study of this MOEM pressure sensor.

2 CONFIGURATION

The structure of MOEM Pressure sensor under consideration is shown is shown in Fig.1. A square diaphragm is micromachined in the <100> silicon wafer. The top view shows the optical Mach-zehnder interferometric structure. The mechanical diaphragm etched in the Si is just below the upper arm of the interferometer. Crossection of the device across the sensing arm is shown in Fig. 1(b). First a square diaphragm is etched in a silicon wafer, over which several layers are formed for waveguidance. First the top of the Si containing the diaphragm is oxidized to from a thick layer of SiO_2 which acts as the cladding layer for integrated optic waveguides. A uniform layer of high refractive index Si_3N_4 is deposited over the SiO_2 which confines the light in the vertical direction. Another layer of SiO_2 formed over this layer and etched in the

Part of the SPIE Conference on Advanced Photonic Sensors and Applications
Singapore • November–December 1999
SPIE Vol. 3897 • 0277-786X/99/$10.00

form of a Mach-Zehnder interferometer forms a ridge over the guiding Si_3N_4 layer resulting in confinement in both the directions in the crossection. Air envelopes the SiO_2 ridge layer. Typical structural parameters of the device chosen are[4] : diaphragm thickness 50 micron., diaphragm size 1mm side, SiO2 supporting layer thickness 1.3 micron: Si3N4 layer thickness 0.45 micron, SiO_2 overlay layer 0.2 micron with ridge height 1.2 micron. The ridge with is 5 micron. Refractive indices are : SiO_2 1.47, Si_3N_4 1.9. The dimensions have been chosen to make the waveguide single mode at 1300 nm.

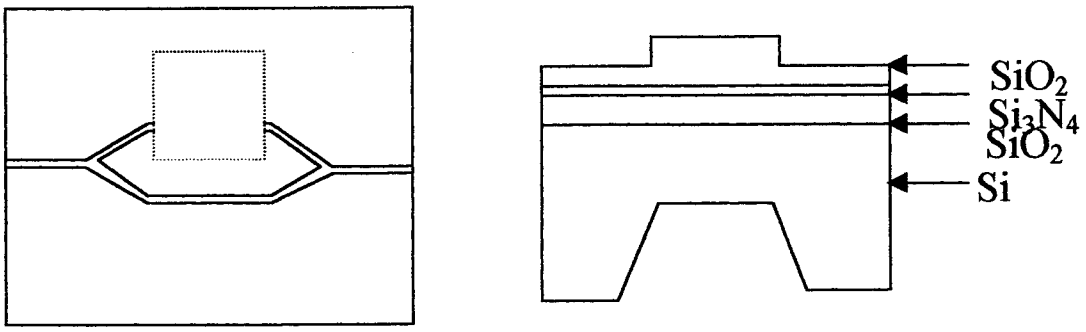


Fig 1. Structure of MOEM Pressure sensor (a) Top view (b) Crossection

3. ANALYSIS

1. Mechanical Analysis

The general equation of the deflection of a diaphragm under an applied hydrostatic pressure with the assumption that its deflection is much smaller compared to its thickness, is given by [5]

$$D\frac{\partial^4\omega}{\partial x^4}+2H\frac{\partial^4\omega}{\partial x^2\partial y^2}+D\frac{\partial^4\omega}{\partial y^4}=P \qquad 1(a)$$

With boundary conditions

$$\omega|_{x=\pm a/2,y=\pm a/2}=0\ ,\ \frac{\partial\omega}{\partial x}|_{x=\pm a/2}=0\ ,\ \frac{\partial\omega}{\partial y}|_{x=\pm a/2}=0 \qquad 1(b)$$

Where

$$D=\frac{Eh^3}{12},H=D+2D_{xy},D_{xy}=\frac{Gh^3}{12} \qquad 1(c)$$

E= Young's Modulus, G= Shear Modulus, h= diaphragm thickness, P= applied pressure

The above equation can be solved[6] by superposition of three components; one ω_0, the deflection of a simply supported diaphragm, the others (ω_x,ω_y) , the deflection caused by the moments distributed along the edges subjected to the clamped boundary conditions. The complete solution can be written as

$$\omega(x,y)=\omega_0(x,y)+\omega_x(x,y)+\omega_y(x,y) \qquad (2)$$

Changing the variables to dimensionless form

$$W=\frac{D\omega}{Pa^4}, X=\frac{x}{a}, Y=\frac{y}{a} \tag{3}$$

Hence the deflection in dimensionless form can be expressed as

$$W(x,y)=W_0(x,y)+W_x(x,y)+W_y(x,y) \tag{4}$$

All the three component can be solved and the solution appears in the form

$$W_0=\frac{4}{\pi^5}\sum_{\substack{m=odd\\ m>0}}^{\infty}\frac{(-1)^{(m-1)/2}\cos(m\pi X)}{m^5}\left[\begin{matrix}1-R(m)\cos(D_m\pi Y)\cosh(D_+m\pi Y)\\ +S(m)\sin(D_m\pi Y)\sinh(D_+m\pi Y)\end{matrix}\right] \tag{5}$$

Where,

$$R(m)=\frac{\cos(D_\alpha_m)\cosh(D_+\alpha_m)+\dfrac{H}{\sqrt{D^2-H^2}}\sin(D_\alpha_m)\sinh(D_+\alpha_m)}{\cos^2(D_\alpha_m)\cosh^2(D_+\alpha_m)+\sin^2(D_\alpha_m)\sinh^2(D_+\alpha_m)} \tag{6}$$

$$S(m)=\frac{\dfrac{H}{\sqrt{D^2-H^2}}\cos(D_\alpha_m)\cosh(D_+\alpha_m)-\sin(D_\alpha_m)\sinh(D_+\alpha_m)}{\cos^2(D_\alpha_m)\cosh^2(D_+\alpha_m)+\sin^2(D_\alpha_m)\sinh^2(D_+\alpha_m)} \tag{7}$$

Where,

$$\alpha_m=\frac{m\pi}{2}, D_-=\sqrt{\frac{D-H}{2D}}, D_+=\sqrt{\frac{D+H}{2D}} \tag{8}$$

W_x, W_y are given by

$$W_x=\frac{-1}{\pi^2}\sum_{\substack{m=odd\\ m>0}}^{\infty}\frac{(-1)^{\frac{(m-1)}{2}}E_m\cos(m\pi Y)}{m^2A(m)}\left[\begin{matrix}\sin(D_m\pi X)\sinh(D_+m\pi X)-\\ \tan(D_\alpha_m)\tanh(D_+\alpha_m)\cos(D_m\pi X)\cosh(D_+m\pi X)\end{matrix}\right] \tag{9}$$

$$W_y=\frac{-1}{\pi^2}\sum_{\substack{m=odd\\ m>0}}^{\infty}\frac{(-1)^{\frac{(m-1)}{2}}E_m\cos(m\pi X)}{m^2A(m)}\left[\begin{matrix}\sin(D_m\pi Y)\sinh(D_+m\pi Y)-\\ \tan n(D_\alpha_m)\tanh(D_+\alpha_m)\cos(D_m\pi Y)\cosh(D_+m\pi Y)\end{matrix}\right] \tag{10}$$

Where,

$$A(m)=\left(\frac{H}{D}+\frac{\sqrt{D^2-H^2}}{D}\tan n(D_\alpha_m)\tanh(D_+\alpha_m)\right)\sin(D_\alpha_m)\sinh(D_+\alpha_m)+$$
$$\left(\frac{\sqrt{D^2-H^2}}{D}-\frac{H}{D}\tan n(D_\alpha_m)\tanh(D_+\alpha_m)\right)\cos(D_\alpha_m)\cosh(D_+\alpha_m) \tag{11}$$

and E_m can be calculated by

for i= odd and i>0 (12)

The stress components can be calculated from[7]

$$\sigma_{xx} = \frac{-h}{2(1-\nu^2)}\left[E\frac{\partial^2\omega}{\partial x^2} + E\nu\frac{\partial^2\omega}{\partial y^2}\right]$$

$$\sigma_{yy} = \frac{-h}{2(1-\nu^2)}\left[E\frac{\partial^2\omega}{\partial y^2} + E\nu\frac{\partial^2\omega}{\partial x^2}\right] \quad (13)$$

$$\sigma_{xy} = hG\frac{\partial^2\omega}{\partial x\partial y}$$

3.2 Optical Analysis

When a coherent light beam of intensity I_0 coupled from a laser diode to a Mach-Zehnder Interferometer, light intensity is divided equally at the first Y-junction. The sensing arm of the interferometer is at the edge of the diaphragm. The light in the sensing arm experiences a phase shift due to the applied pressure on the diaphragm. The output light intensity is given by[8]

$$I = \frac{I_0}{2}(1+\cos(\Delta\Phi)) \quad (14)$$

The phase change in the sensing arm can occur due to two reasons. First one Photo-elastic effect. When pressure is applied, it induces a stress , which in turn causes the change in the refractive index of the waveguide. Second one is the elongation of length of the waveguide. When pressure applied, the diaphragm gets displaced and hence the length of the sensing waveguide increases. This causes the phase change of the light in the guide.

$$\Delta\Phi = \beta\Delta L + \int_L \delta\beta dL \quad 15\ (a)$$

where L is the length of the sensing arm and β is the propagation constant.

$$\Delta\Phi = \Delta\Phi^{PE} + \Delta\Phi^{D} \quad 15(b)$$

$\Delta\Phi^{PE}$ is due to phot-elastic effect and $\Delta\Phi^{D}$ is due to diaphragm deflection.

3.3 Opto-Mechanical Coupling

The change in refractive index for TE mode is given by

$$\Delta n_{yy} = C_2\sigma_{xx} + C_1\sigma_{yy} \quad (16)$$

and for the TM mode

$$\Delta n_{zz} = C_2\sigma_{xx} + C_2\sigma_{yy} \quad (17)$$

Where σ_{xx}, σ_{yy} are the stress components and C_1, C_2 are the stress optic coefficients. The stress components with $\sigma_{zx}, \sigma_{zy}, \sigma_{zz}$ are neglected because of the assumption that diaphragm is thin.Using the coupled mode formulation, the change in the phase due to dielectric perturbation is given by

$$\Delta\Phi^{PE}{}_{TE,TM} = \frac{2\pi}{\lambda}\frac{n}{n_{eff}}\int \Delta n_{TE,TM}(x,y)dx \quad (18)$$

3.4 Phase change due to change in length

Because of deflection of the diaphragm, the difference between the length of arc along the deflection and the original length L, for small deflections can be approximated as

$$\Delta L = \int_0^{a/2} \left(\frac{\partial \omega}{\partial x} \right)^2 dx \tag{19}$$

Hence the phase shift due to this change in length is given by

$$\Delta \Phi^D = \frac{2\pi}{\lambda} n_{eff} \Delta L \tag{20}$$

Equations (15), (18) and (20) give total change in phase.
Once the phase change is known, the change in optical intensity can be calculated using equation 14. The sensitivity of the sensor can be put as the change in the optical phase to the applied pressure.

$$S = \frac{\partial \Delta \Phi}{\partial P} \tag{21}$$

4. RESULTS AND DISCUSSION

Computations have been performed for a typical device structure. The mechanical properties have been computed assuming that the optical layer is thin compared with the SiO_2 diaphragm and the deflections are small compared with the diaphragm thickness. The diaphragm thickness is 50 micron. The photo-elastic coefficients of SiO_2 are $C_1 = -0.65 \times 10^{(-12)}$/Pa, $C_2 = -4.22 \times 10^{(-12)}$/Pa. Fig. 2 shows the diaphragm deflection, normalized with respect to pressure, diaphragm size, and the D-coefficient. The deflection is maximum at the center as expected. However the stresses are maximum at the edges (Fig. 3 and 4). Since the waveguide refractive index depends on the stress distributions, the sensing arm of the interferometer is located on the edge of the diaphragm. Fig. 5 show the change of refractive index profile due to photo elastic effect for TE and TM modes. The propagation is considered along x-axis of the diaphragm, which is in x-y plane. The observe that the normalized refractive index change is much larger for TM mode (dny(x,y)) than for TE mode (dnx(x,y).

Fig. 5(b) and Fig. 5(d) show the refractive index variations at the location of the sensing arm. Computation of optical path length change due to the displacement is shown in Fig. 6. We observe that phase changes due to the refractive index variations are more significant than the phase change due to optical path length change. The sensitivity can be obtained by differentiating the phase change with respect to the impressed pressure. Sensitivity is independent of pressure in case of photo-elastic effect and linearly dependent on pressure in case of deflection.

5. CONCLUSION

MOEM pressure sensing devices have improved characteristics. Here a Mach-Zehnder interferometric waveguiding structure on a silicon diaphragm is considered and analyzed to obtain the optical phase changes from the mechanical deflections due to impressed hydrostatic pressure. The phase changes are converted to intensity by the interferometer. By integrating over the optical path length we find that the refractive index changes have more significant effect on the final output and so the sensitivity. The optical wave guiding properties can be simulated using the refractive index profile in a beam propagation algorithm. The approach presented can be extended and improved in several ways such as considering a rectangular diaphragm instead of a square diaphragm, considering the diaphragm as a composite multilayer structure including various optical layers, etc.

6. REFERENCES

1. M.Tabib Azar and G.Behelm ,"Modern trends in microstructures and integrated optics for communication, sensing and signal processing", *Optical Engineering* **36(5)**, pp. 1307-1318 , 1997.
2. M.C. Wu, "Micromachining for Optical and Optoelectronics Systems", *Proceedings of the IEEE* ,**85 (11)**, 1997.
3. L. Rosengren, Jan Soderkvist and L. Smith , "Micromachined sensor structures with linear capacitive responses", *Sensors and Actuators*, **31** pp.200-205 1992.
4. H. Porte, V. Gorel, S. Kiryenko, J. Goedgebuer, W. Daniau and P. Blind, "Imbalanced Mach-Zehnder Interferometer integrated in Micromechined Silicon substrate for Pressure Sensor", *Journal. of Light wave Technology,* **17(12),** pp.229-233 1999.
5. S. Timoshenko and S. Woinowsky-Krieger ,"Theory of Plates and Shells" ,McGraw-Hill ,NewYork, 1959.
6. T.Y. Bin and R. S. Huang, "CAPSS: A Thin Diaphragm Capacitive Pressure Sensor" *Sensors and Actuators* , **11,** pp. 1-22,1987.
7. S. K. Clark and K.D. Wise, "Pressure Sensitivity in anisotropically etched thin-diaphragm Pressure Sensors" *IEEE Transactions on Electron Devices,* **ED-26,** pp. 1887-1896,1979.
8. K. Benaissa and A. Nathan, "IC Compatible Optomechanical Pressure Sensors Using Mach-Zehnder Interferometry", *IEEE Transactions on Electron Devices,* **43(9),** 1996.

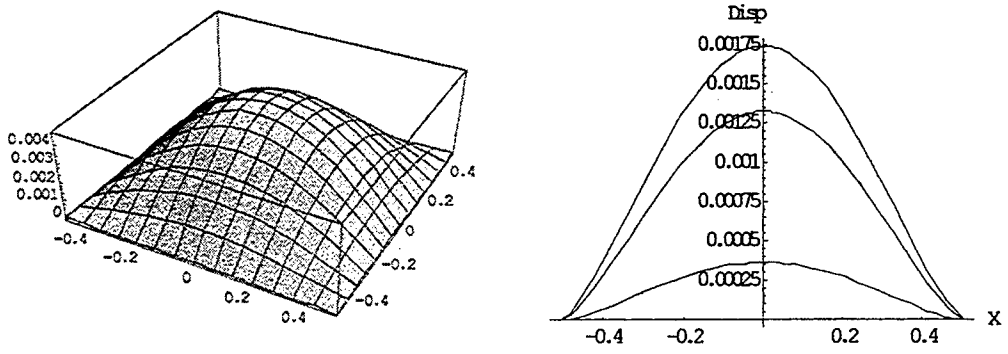


Fig. 2 Normalized . Displacement (a) W(X,Y) (b) W(X,Y) for different locations Y of sensing arm

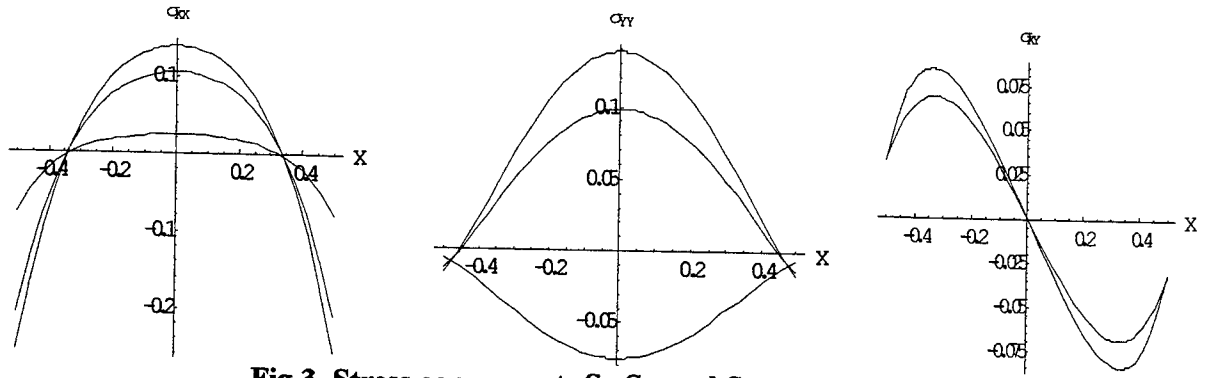


Fig.3 Stress components Sx,Sy and Sxy at y=0.4,0.2,0

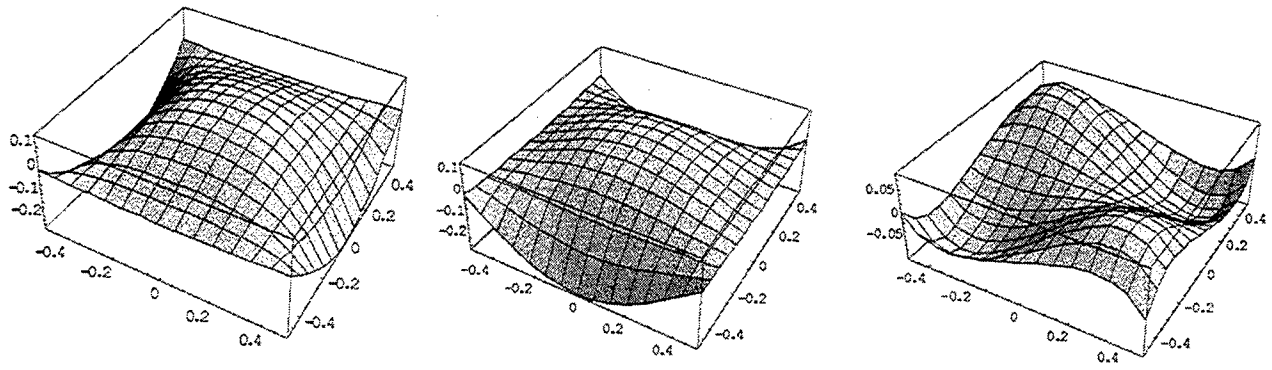


Fig4 Stress Components S_{xx},S_{yy} and S_{xy} (Normalised)

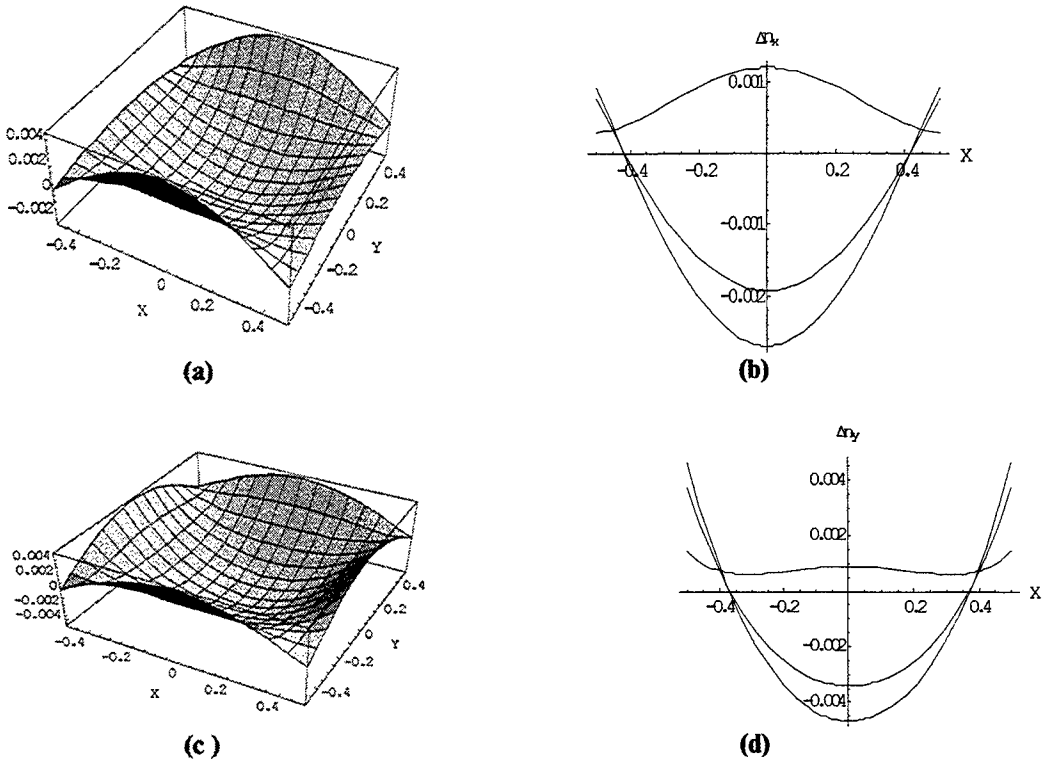


Fig. 5 Refractive Index changes (a) and (b) for TE Mode and (c) and (d) for TM Mode, (b) and (d) for Y=0, 0.2, 0.4

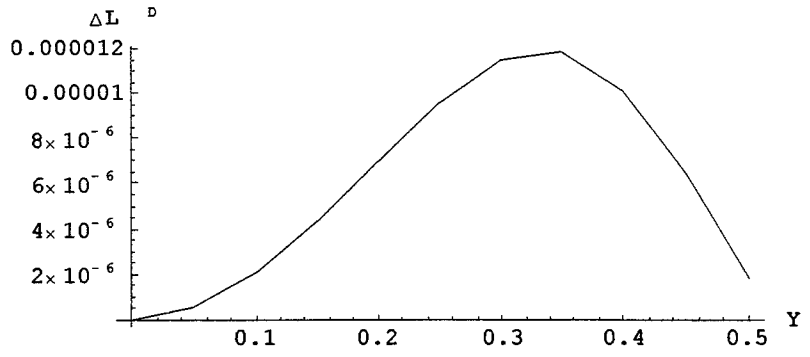


Fig. 6 Change in the Optical Path length due to displacement of the diaphragm , variation with position Yof the interferometric arm

The Development of a Pyroelectric Thin-film Infrared Sensor by Micro-opto-electro-mechanical System (MOEMS) Technology

Jyh-Jier Ho[a], Y.K. Fang[b], M.C. Hsieh[b], C.Y. Chen[a], S.F. Ting[b], S.M. Lin[a] and K.H.Wu[c]

[a] Dept. of Electronics Eng., Fortune Institute of Tech., Chi-Shan Town, Kaohsiung, Taiwan, R.O.C.

[b] VLSI Tech. Lab., Dept. of Electrical Eng., National Cheng Kung Univ., Tainan, Taiwan, R.O.C.

[c] Dept. of Electrical Eng., Nan-Tai Univ. of Technology, Yung-Kang City, Tainan, Taiwan, R.O.C.

ABSTRACT

An infrared (IR) sensor with the lead-titanate ($PbTiO_3$) thin-film and thermal isolation improvement structure using the technology of micro-opto-electro-mechanical systems (MOEMS) has been designed, fabricated and developed. In this paper, both numerical analysis of the static operation mechanism such as the effects of IR light power on the the depletion layer width and voltage drop across the thin $PbTiO_3$ film, and the dynamic response of a thin-film pyroelectric sensor to sinusoidal modulated radiation are reported and compared to the experimental results. The fitting is quite well. The major IR-sensing part on the cantilever beam of the developed sensor consists of a 50-nm $PbTiO_3$ layer deposited by RF sputtering, and a gold (Au) layer evaporated as an IR radiation absorber. With active cantilever dimensions of $200\times100\times5$ μm^3 formed by etching processes, the cantilever structure exhibits a much superior performances to that of a traditional IR-sensing bulk structure under the 800-μW incident optical light with wavelength of 970 nm.

Keywords: IR sensor, MOEMS, $PbTiO_3$ thin-film, RF sputtering, etching process..

1. Introduction

The pyroelectric thin film sensor is a thermal transducer based on the pyroelectric effect. A pyroelectric infrared (IR) sensor shows some advantages compared with the other types of IR sensors, because the pyroelectric can be operated even at room temperature [1], has little wavelength dependence of the response over a wide IR range and a fast response[2]. Among the pyroelectric materials, such as triglicine sulfate (TGS) and its isomers, the lead-titanate ($PbTiO_3$) possesses some excellent characteristics, i.e., large pyroelectric coefficient, small dielectric constant, high Curie temperature and comparatively large pyroelectric figure of merit, Fm_v ($= \gamma/\varepsilon_\gamma C_v$, where C_v is a specific heat). Traditionally, a pyroelectric IR sensor with $PbTiO_3$ film has been reported with a metal-ferroelectric $PbTiO_3$ - semiconductors (MFSS) structure [3-5]. In which, the $PbTiO_3$ film is attached directly to the Si-substrate, the thermal energy absorbed in the $PbTiO_3$ film will be lost partly to the substrate, thus limiting the current gain of the sensor.

Thus, in design a high sensitivity IR sensor with a pyroelectric thin film, the thumb rule is to put the sensing element be thermally isolated to its substrate, so that the photo-current can be fully response to the incident IR light [4]. We have proposed an improved structure with a cantilever beam by the micromachining technology, in which the high thermal isolation from its surrounding is obtained and leads high performances, such as current gain, current responsivity and dynamic response.

Part of the SPIE Conference on Advanced Photonic Sensors and Applications
Singapore • November–December 1999
SPIE Vol. 3897 • 0277-786X/99/$10.00

2. Design and Fabrication of the Pyroelectric IR Sensor

The cantilever-beam structure used (as shown in Fig. 1) in this analysis consists of a composite sensing element / substrate / air-layer system. In the composite system, the substrate thickness (t_s), which is controlled by the etching technique, plays a key role on the thermal transport equilibrium between the sensor and its surrounding, thus affecting the dynamic response of the sensor seriously. Therefore, in the analysis the effect of t_s-value on the dynamic response has been considered firstly.

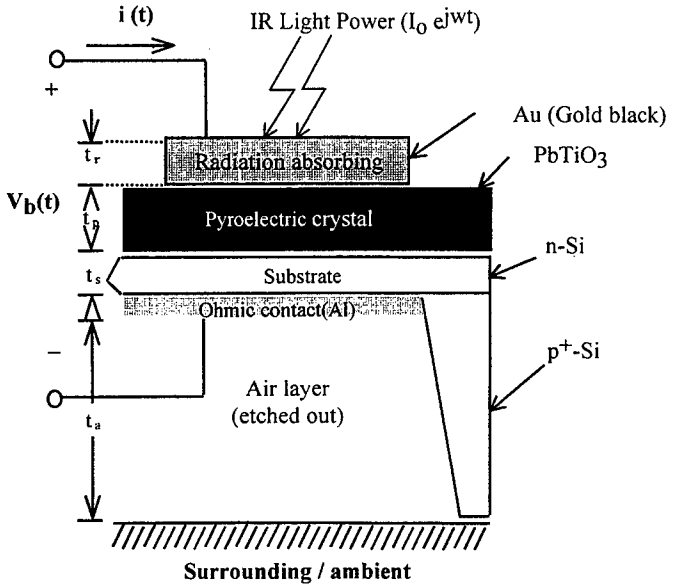


Fig.1 The diagrams of the IR-sensor with a cantilever beam improved structure shows a schematic illustrating the various layers compromising the structure.

For static analysis (t=0), the terminal voltage (V_b), the voltage drop (V_p) and the charge (Q_s) on the $PbTiO_3$ layer in Fig.1 can be written as

$$V_b = -\phi_{ms} + V_p + V_s \quad (1)$$

$$V_p = (Q_s+Q_f)/C_p = (d_p/\varepsilon_p)(Q_s+Q_f) \quad (2)$$

$$Q_s = qN_D W_s = (2qN_D\varepsilon_s V_s)^{1/2} \quad (3)$$

where ϕ_{ms} is 0.82 eV, work function difference between Au and Si; V_s is the voltage drop across the surface depletion region under the $PbTiO_3$ layer; Q_f equals to 1.6×10^{-8} coul/cm^2, effective oxide charge within $PbTiO_3$ layer; d_p equals to 50 nm which is $PbTiO_3$ film thickness; ε_p is dielectric constant of $PbTiO_3$ film which is dependent with IR light power; ε_s equals to $11.9\varepsilon_o$ which is dielectric constant of Si ($\varepsilon_o = 8.854\times10^{-14}$ F/cm); N_D equals to1.2×10^{15} cm^{-3} which is the doping level of the n-type epi-layer.

Substituting Q_s in Eq.(3) into Eq.(2) and combining Eq.(1), we obtain V_b:

$$V_p = (d_p/\varepsilon_p)[\,(2qN_D\varepsilon_s V_s)^{1/2}+Q_f] = (V_b + \phi_{ms}) - [(V_p\,\varepsilon_p / d_p)]^2/(2qN_D\varepsilon_s) \quad (4)$$

With some manipulation in Eq.(4), V_p can be expressed as a positive solution to the second order equation, i.e.,

$$[V_p]^2 + A[V_p] + B = 0 \quad (5a)$$

$$\Rightarrow V_p = [-A \pm (A^2-4B)^{1/2}]/2 \quad (5b)$$

where $A = (d_p/\varepsilon_p)^2(2\, qN_D\varepsilon_s) - 2(d_p/\varepsilon_p)Q_f$ (5c)

$$B = (d_p/\varepsilon_p)^2(Q_f)^2 - [(V + \phi_{ms})\,(d_p/\varepsilon_p)^2(2qN_D\varepsilon_s)] \quad (5d)$$

According to Eq.(1), the macroscopic electric field (E_p) in the $PbTiO_3$ film will increase as the polarization (P) reduces due to the IR light radiation. So, as shown in Fig.1, the tunneling current (I_t) through the $PbTiO_3$ film will increase with the E_p based on the following equation,

$$I_t \sim [V_p]^2\exp\{[-4(2m^*)^{1/2}(q\phi_B)^{3/2}]/(3qhE_p/2\pi)\} = C[V_p]^2\exp[D/V_p] \quad (6)$$

where m^* is effective mass; ϕ_B is barrier height; h is Planck constant, V_p equals to E_pd_p with d_p (=50 nm), C (= 1.1875×10^5 $\mu A/V^2$) and D (= 0.5 V) constants determined from the experimental data.

For the dynamic analysis, assumed that the radiation energy (= $I_o\, e^{j\omega t}$), modulated sinusoidally by a mechanical chopper at an angular frequency, ω, falls at the front surface of the radiation absorbing layer (Au). The radiation energy is absorbed by the Au-layer, converted to thermal energy that isolated by air-surroundings, and then induced a pyroelectric current (i_p). Thus, a one-dimensional heat conduction equation is employed and the lateral heat flow is neglected due to three lateral sides were etched out and filled with air. The thin electrical conducting electrode on the back surface of the substrate is also neglected on the assumption that it has negligible thermal conductance and thermal capacitance. Thus, the heat transfer in the IR sensor is described by the heat conduction equation given by

$$\alpha_i[\partial^2T_i(x,t) / \partial x^2] = \partial T_i(x,t) / \partial t \quad (7)$$

where $T_i(x,t)$ is the excess temperature above the ambient temperature in the i-th layer at a distance x measured from the upper surface of each layer and α_i is the thermal diffusivity of the i-th layer. The α_i is equal to k_i/C_{Vi} where k_i and C_{Vi} are thermal conductivity and volume specific heat of the i-th layer, respectively. The subscript i symbolizes the following subscripts r, p, s, a and o for the radiation absorbing layer, pyroelectric layer, substrate, air (etched out) layer and surrounding (ambient) layer, respectively.

By using separate variable method, we get the general solution ($T_i(x,t)$) and its heat flow ($f_i(x,t)$) to the heat conduction Eq.(7) for each of the four individual layers as

$$T_i(x,t) = T_i(x)\, e^{j\omega t} \quad (8a)$$

$$T_i(x) = A_i\, sh\,(u_ix) + B_i\, ch\,(u_ix) \quad (8b)$$

$$f_i(x,t) = -k_i \frac{\partial T_i(x,t)}{\partial x}\Big|_{x=0} = -k_iu_iT_i(0,t) \quad (9)$$

where the u_i is defined as $(1+j)[\omega/(2k_i)]^{1/2}$, which k_i is the thermal diffusivity of the i-th layer. In the pyroelectric thin film, $T_i(x)$ and $f_i(x,t)$ are the temperature distribution and the heat flow, respectively, which are governed by the heat conduction equation and by the boundary conditions as

$$T_i = T_{i+1} \quad (10a)$$

$$f_i = f_{i+1} \quad (10b)$$

Eqs (10a) and (10b) mean that the temperature (T) and heat flow (f) at the interfaces are continuous. Thus, the relations of T_o and f_o in the ambient can be written as a matrix type [6]

$$\begin{pmatrix} T_0 \\ f_0 \end{pmatrix} = \begin{pmatrix} a_{11}^a a_{12}^a \\ a_{21}^a a_{22}^a \end{pmatrix} \begin{pmatrix} a_{11}^s a_{12}^s \\ a_{21}^s a_{22}^s \end{pmatrix} \begin{pmatrix} a_{11}^p a_{12}^p \\ a_{21}^p a_{22}^p \end{pmatrix} \begin{pmatrix} a_{11}^s a_{12}^s \\ a_{21}^s a_{22}^s \end{pmatrix} \begin{pmatrix} T_r(0) \\ f_r(0) \end{pmatrix} = \begin{pmatrix} s_{11} s_{12} \\ s_{21} s_{22} \end{pmatrix} \begin{pmatrix} T_r(0) \\ f_r(0) \end{pmatrix} = \bar{S} \begin{pmatrix} T_r(0) \\ f_r(0) \end{pmatrix} \tag{11}$$

Setting det ($\bar{S}$) = 1 and f_o = 0 in Eq.(5), temperature ($T_r(0)$) and heat flow ($f_r(0)$) of the top radiation absorbing layer can be solved by the matrix method, i.e.,

$$\begin{pmatrix} T_r(0) \\ f_r(0) \end{pmatrix} = \bar{S}^{-1} \begin{pmatrix} T_0 \\ f_0 \end{pmatrix} = \begin{pmatrix} s_{22} - s_{21} \\ -s_{12} s_{11} \end{pmatrix} \begin{pmatrix} T_0 \\ 0 \end{pmatrix} \tag{12}$$

Meanwhile, the heat flow of the top radiative absorbing layer can be described by [6]

$$f_r(0,t) = -k_r \frac{\partial T_r(x,t)}{\partial x}\Big|_{x=0} = \eta I_0 - g[T_0 - T_r(0)] \tag{13}$$

where η is the fraction of the incident radiation power that is converted into heat; I (= $I_o e^{j\omega t}$) is the radiation intensity per unit area, and g (= $4\sigma T_o^3$) is the radiation loss conductance per unit area. At T_o = 300 K, the value of g is 6.1×10^{-4} W/cm^2 K using the value of the Stefan-Boltzmann constant $\sigma = 5.67\times10^{-12}$ W/cm^2 K [7].

Substituting Eq.(12) into Eq.(13) and obtaining the general solution ($T_p(x)$) of heat conduction Eq.(7) in the pyroelectric layer as

$$T_p(x) = \{[T_{21}/k_p u_p]\, sh\,(u_p x) + ch\,(u_p x)\} T_o = \{[T_{21}/k_p u_p] sh(u_p x) + ch(u_p x)\}\{[\eta I_o]/[g(1-s_{22})-s_{12}]\} \tag{14}$$

where the transform matrix, $\bar{T}$, at the pyroelectric layer, can be denoted as

$$\bar{T} = \begin{pmatrix} T_{11} T_{12} \\ T_{21} T_{22} \end{pmatrix} = \begin{pmatrix} a_{11}^a a_{12}^a \\ a_{21}^a a_{22}^a \end{pmatrix} \begin{pmatrix} a_{11}^s a_{12}^s \\ a_{21}^s a_{22}^s \end{pmatrix} \begin{pmatrix} a_{11}^p a_{12}^p \\ a_{21}^p a_{22}^p \end{pmatrix} \tag{15}$$

The pyroelectric current ($i_p(t)$) and the current responsivity (R_I) in the IR sensor can be defined and expressed as [9]

$$i_p(t) = \frac{p_e}{t_p} \frac{\partial}{\partial t} \int_0^{t_p} T_p(x,t)\, dx = \frac{jwe^{jwt} p_e}{t_p} \frac{\eta I_o}{[g(1-s_{22})-s_{12}]} \{\frac{T_{21}}{k_p u_p^2} ch(u_p t_p) + \frac{T_{22}}{u_p} sh(u_p t_p) - \frac{T_{21}}{k_p u_p^2}\} \tag{16}$$

and $$R_I = \left|\frac{i_p(t)}{\eta I_o e^{jwt}}\right| = \frac{wp_e}{t_p} \frac{1}{[g(1-s_{22})-s_{12}]} \{\frac{T_{21}}{k_p u_p^2} ch(u_p t_p) + \frac{T_{22}}{u_p} sh(u_p t_p) - \frac{T_{21}}{k_p u_p^2}\} \tag{17}$$

where p_e and t_p are the pyroelectric coefficient and the thickness of the pyroelectric film. Meanwhile, the Si substrate thickness and the pyroelectric coefficient (γ) of $PbTiO_3$ are 5 μm and 9.5×10^{-8} C/cm^2 °K [10], respectively.

As shown in Fig.1, the sample was prepared on a <100>-oriented, n/p$^+$ epitaxial Si-substrate with resistivity of 3-4 Ω cm and thickness of 4-5 μm in the n-layer. At first, the optimal cantilever beam with dimensions of 200×100×5 μm^3 was photo-lithographed and fabricated by the anisotropic etchant mixture, i.e., an ethylene diamine (120 ml) / pyrocatechnol (20 g) / water (60 ml) (EDP) solution with the etch rate of 72 μm/hr at 100 °C. Proper layout will form a gap around the cantilever beam, giving high thermal isolation from the surrounding substrate. Next, the optimal $PbTiO_3$ film thickness (50 nm) was deposited by a RF sputtering system [7,10] to form the active IR-sensing region on the etched cantilever beam. The target was a powder mixture of $PbTiO_3$ made from a sintered mixture of Pb_3O_4 and TiO_2. Excess Pb_3O_4 of 10 wt% was added to the target to compensate for the Pb deficiency in the deposited film. Then, gold (Au), a heat-absorption material, was evaporated upon the $PbTiO_3$ film to assist the absorption of the incident IR radiation. Finally, Al was evaporated on the back of the etched beam and sintered as an anode ohmic contact electrode.

3. Results & Discussions

The simulated cantilever and traditional types of V_p and V_s versus IR light power curves can be found as shown in Fig.2, in which the V_p increases with the IR light power but the V_s decreases. This is due to the fact that the ε_p decreases with the IR light power and hence the V_p increased based on Eqs.(5), and due to the increase of the V_p, the V_s will decrease with the IR light power according Eq.(1). In comparing both sensor types, we find that the sensor with cantilever beam not only gets a higher V_s values but also increases more significant with increasing IR power than that of a traditional type. This attributes to the advantage of cantilever beam structure, since three sides and bottom of the beam have been etched out, hence the cantilever beam possesses a smaller effective absorption dimension and better isolation to the surrounding of the beam. Thus, the voltage raising responded to the regression of polarization in $PbTiO_3$ film under IR light incident will be faster.

In addition, if the found V_p in Eqs.(5) is substituted into Eq.(6) to find the tunneling (terminal) current I_t, i.e., the photo current of optical sensor under IR irradiation we can construct the simulated tunneling current versus IR light power curve as shown in Fig.3, in which the I_t increases with the IR light power. This is owing to that the V_p increases with the IR light power and hence the I_t based on Eq.(8). Therefore, it also demonstrates the fact that the conductivity of the diode increases with the IR light power. Again, the thermal isolation improvement, which leads a higher tunneling current phenomenon, can be found in the simulated I_t values under various IR light power shown in Fig.3. Meanwhile, the experimental measurements are also expressed in the figure, which shows the fit is very good.

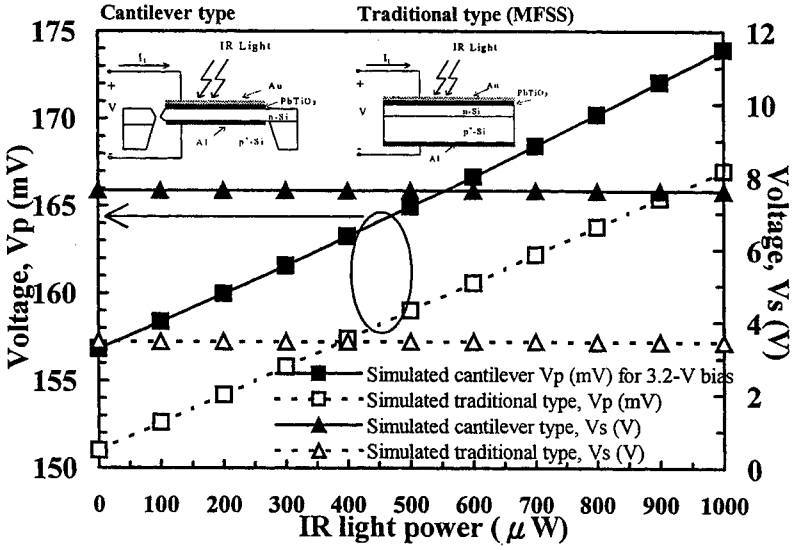


Fig.2 Effects of IR light power on the voltage drop V_p across the $PbTiO_3$ layer (left), the voltage drop V_s across the surface depletion region under the $PbTiO_3$ layer (right), at the terminal voltage (V) of 3.2 V.

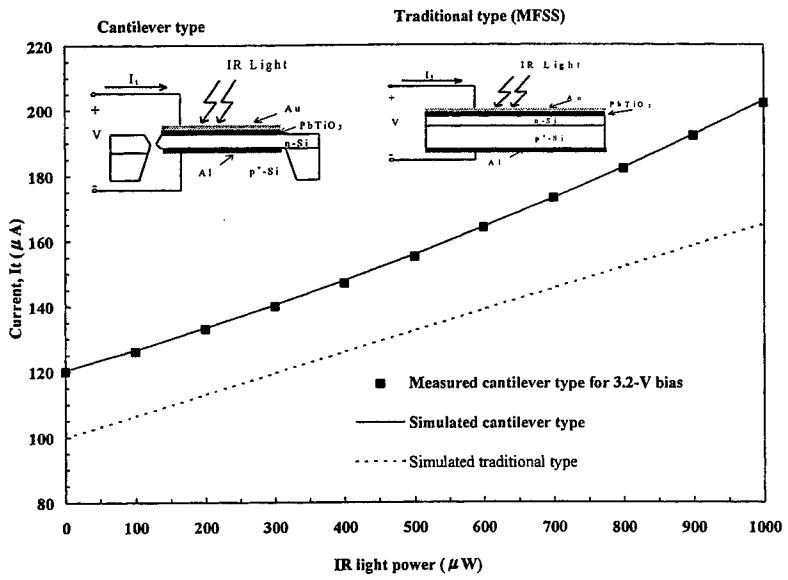


Fig.3 Effects of IR light power on the tunneling current (I_t) at the input bias of 3.2V and comparison of the simulated and measured values of the cantilever improved structure with the traditional structure.

Fig.4 shows the simulated curves of current responsivity versus frequency with the thickness of silicon substrate (t_s) as the parameter. The thickness of the gold black (Au) radiation-absorbing layer (t_r) and the pyroeletric $PbTiO_3$ crystal (t_p) are assumed to be 0.5 and 0.05 μm, respectively. The current responsivity falls off at about 9×10^5 Hz with decreasing frequency. Physically, this is due to the fact that at frequencies below 9×10^5 Hz the heat is able to conduct to the bottom surface of the $PbTiO_3$ crystal and flows out of the crystal into the substrate. At the lower frequencies (less than about 3000 Hz), the current responsivity is enhanced with the decreasing substrate thickness (t_s), because of the less heat loss in the substrate. That is, the thinner the substrate, the higher the responsivity. However, when the substrate thickness less than 5 μm, say t_s = 2 μm, it appears the sharp discontinuity at about 0.2 Hz, and there is a very weak and fragile substrate for practical application. Thus, we select substrate thickness (t_s = 5 μm) as the optimal IR sensor.

In addition, if the found t_s-value in Eq.(15) is substituted into Eqs.(16) and (17) to find the photo (terminal) current $i_p(t)$, i.e., the photo current of optical sensor under IR irradiation (I_o = 800 μW), and the current responsivity (R_I) (has been defined as $| i_p(t)/[\eta I_o e^{jwt}] |$ in Eq.(17)), we can plot the simulated current responsivity versus frequency curve as shown in Fig.5, in which the R_I increases as compared with the traditional type (t_s = 350 μm). This attributes to the advantage of cantilever beam structure, since three sides and bottom of beam substrate have been etched out, hence the cantilever beam possesses a smaller effective absorption dimension and better isolation to the surrounding of the beam. Meanwhile, the experimental measurements are also expressed in the figure, which shows again the fitting is very good.

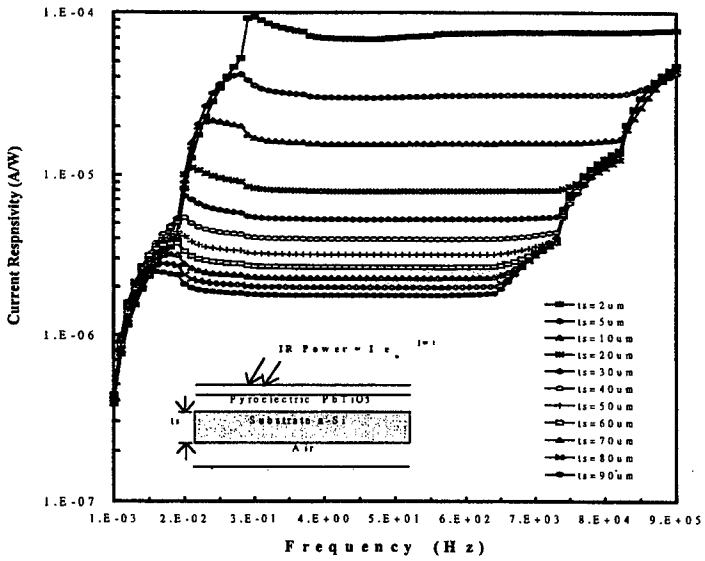


Fig.4 Effects of the substrate thickness (t_s) values on current responsivity versus frequency in the four-layer-model analysis. $t_r = 0.5$ μm, $t_p = 0.05$ μm, $t_s + t_a = 350$ μm and for various substrate layer thickness: $t_s = 2$ μm, 5 μm, 10 μm, 20 μm, 30 μm, 40 μm, 50 μm, 60 μm, 70 μm, 80 μm and 90 μm from top to bottom.

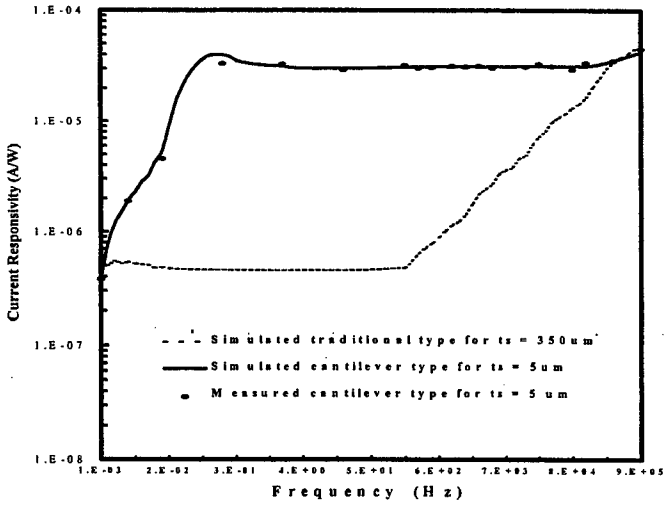


Fig.5 The current responsivity versus frequency of the traditional type (t_s = 350 μm, t_a = 0 μm) and the developed structure improvement under values of the optimal parameters used in theoretical calculation listed in Table I. These theoretical calculations are also compared with the experimental results under 800-μW IR radiation and 5-V biasing voltage.

4. Acknowledgements

The authors acknowledge financial support from *The National Science Council of R.O.C.* under contract No. NSC 89-2218-E-268-001.

5. References

1. J.H. Putley: *The Pyroelectric Detector*, **5**, Semiconductors and Semimetals (Academic, New York), pp.259-285, 1970.
2. M. Okuyama, K. Ohtani, T. Ueda and Y. Hamakawa, "Si-monolithic integrated pyroelectric infrared sensor made of $PbTiO_3$ thin films", *Int. J. Infrared Millm. Waves*, vol. **6**, p. 71, 1985.
3. R. Takayama, Y. Tomita, K. Iijima and I. Ueda, "Pyroelectric linear array infrared sensors made of c-axis-oriented La-modified $PbTiO_3$ thin films": *Jpn. J. Appl. Phys.*, vol. **63**, p.5868, 1988.
4. M. Okuyama, Y. Toshihiro and Y. Hamakawa, "Pyroelectric infrared-CCD image sensor using $LiTaO_3$", *Sensors and Actuators*, vol. **16**, p. 263, 1989.
5. Keizo Sakurai, "Matsushita uses $PbTiO_3$ thin film for high response speed pyroelectric IR imager", *NIKKEI Electronics Asia*, p. 36, Jan. 1993.
6. H.S. Carslaw and J.C. Jaeger, *Conduction of Heat in Solids*, Clarendon Press, Oxford, Chap. 3, 1959.
7. F.Y. Chen, Y.K. Fang, C.Y. Hsu and J.R. Chen, "Time response analysis of a pyroelectric detector", *Ferroelectrics*, Vol. 200, pp. 257-268, 1997.
8. A.G. Chynoweth, "Dynamic method for measuring the pyroelectric effect with special reference to barium titanate", *J. Appl. Phys.*, Vol. 27, pp.78-84, 1955.
9. Chian-ping Ye, Takashi Tamagawa and D.L. Polla, "Experimental studies on primary and secondary pyroelectric effects in $Pb(Zr_xTi_{1-x})O_3$ $PbTiO_3$ andZnO thin films", *J. Appl. Phys.*, Vol. 70(10), pp. 5538-5543, 1991.
10. Jyh-Jier Ho, Y.K. Fang, K.H. Wu, W.T. Hsieh, C.W. Chu, C.R. Huang, M.S. Ju and C.P. Chang, "A High Sensitivity Lead-titanate ($PbTiO_3$) Pyroelectric Thin-film Infrared Sensor with Temperature Isolation Improvement Structure", *IEEE Electron Device Letters*, Vol. 19 (6), p.189, 1998.

Microelectromechanical fiber optical switches

A. Q. Liu[a], Franck Chollet[a], A. Asundi[b], and Jianmin Miao[b]

[a]MEMS Technology, Institute of Materials Research & Engineering
Blk S7, Level 3, National University of Singapore
Singapore 119260
[b]School of Mechanical and Production Engineering
Nanyang Technological University
Singapore 639798

ABSTRACT

This paper reports the design, fabrication, and testing of a Micro-opto-mechanical grating/mirror switch driven by electrostatic actuator for fiber-optic switching applications. It consists of two bounded silicon wafers. One input fibre with a hemispherical lens at its end and three photodetectors are fabricated on the upper wafer. A movable platform with two gratings and one mirror are fabricated on the lower wafer. When the movable platform is at a certain position, the input beam can be split into three beams by the gratings. Details of the grating switching design, theoretical analysis, fabrication and experimental results are presented in this paper.

Keywords: Micro-Opto-Electro-Mechanical Systems (MOEMs), optics switch, micromirror, micrograting, multiplexers, and fiber-optics communication.

1. INTRODUCTION

One attractive feature of optics especially for communication and information processing is its ability to transmit large amounts of information in parallel. Fiber optics switches are used to reconfigure the network for example the high speed digital fiber-optics network. An optical network or a hybrid electronic/optical network needs various types of optics switches. The rapid growth of fiber-based local area networks has created a large demand for low-loss, low-cost, reliable, and mass-producible fiber optics switches.

The significant advantage of using optical MEMS (Microelectromechanical Systems) for opto-mechanical switches is its small size, low cost, and high reliability and stability[1, 2]. Micro-opto-mechanical gratings are very attractive for implementing optical multiplexers, as reflection grating is used to disperse incident light into angular directions corresponding to the incident light's wavelength and the construction properties of the grating. Variable blaze gratings (VBGs) have adjustable blaze angles with a fast time response (less than 10 ms) are suitable to be designed as $1 \times N$ fiber optics switches using MEMS technology. It provides the important capability of high density digital transmission communication with a bandwidth as high as 2.5 Gbps. Diffraction grating without any movable mechanical part has been used to produce a diffraction spot pattern for optical interconnection[3]. The type of interconnection is mainly used for computing network. Two diffraction gratings of other type used as beam switch have been reported[4, 5].

This paper proposes a Micro-opto-mechanical grating/mirror switch. It is a potential application device to be used as spatial terminal switches for optical local network, and a single-wavelength system like a passive telephony optical network. Firstly, the design of a grating/mirror switch with one-input and three-outputs used as a miniaturized fibre-optics network terminal switch is described. The design includes grating structure, mechanical actuators and electronic signal processing.

The main effort is concentrated on improving the diffraction efficiency of the interested diffraction order. Then, the simulation results of the grating diffraction efficiency is shown. Lastly, the fabrication process is explained and experimental results are compared with theoretical analysis.

Correspondence: Email: aq-liu@imre.org.sg; Telephone: (65) 8748161; Fax: (65) 8720785

Part of the SPIE Conference on Advanced Photonic Sensors and Applications
Singapore • November–December 1999
SPIE Vol. 3897 • 0277-786X/99/$10.00

2. GENERAL CONSIDERATIONS OF OPTICAL GRATING

The grating interference orders result when light reflected off each grating element combines constructively in the far field. Normally, the diffraction grating is divided into reflective flat grating, blazed grating and concave grating.

2.1 Flat Gratings

The flat reflective grating is to produce constructive interference. The locations of constructive interface regions are sequentially numbered as interference orders 0, 1,..., m. The light reflected off grating elements adds destructively in the region between orders. Because the phase of light arriving at the location of interference order is dependent on the wavelength, the location of the orders depends on the wavelength of the incident light. The direction of a particular order, θ_m resulting from normally incident light is given by

$$\theta_m = Arc\sin(\frac{m\lambda}{P}) \tag{1}$$

where m is the interference order number, λ is the wavelength of the source, and P is the period of the grating. The spatial intensity distribution is

$$I(\Psi) = cons \times \left\{ L \cdot k \cdot \frac{\sin(\pi N\Psi)}{N\sin(\pi\Psi)} \cdot \frac{\sin(\pi k\Psi)}{\pi k\Psi} \right\}^2 \tag{2}$$

where N is total number of grating lines, w is width of the reflection part, k is aperture factor ($k = w/P$), L is the total width of the grating ($L = NP$), and $\Psi = P(\sin\theta_i + \sin\theta_m)/\lambda$.

The grating equation (2) describes both the formation of spectrum and the formation of diffracted orders. Some insight into the way in which the various orders are formed may be gained by considering the vector sum in each case. Actually the vector sum goes round one complete revolution for the first order, two for the second, three for the third, etc. Since the second half is suppressed this gives rise to an amplitude of ½ in the zero order, ½π in the first, 0 in the second (and all even orders) and $1/6\pi$ in the third (and $1/2m\pi$ in all odd orders). In this particular case the efficiency in the zero order is ¼, in the first $¼\pi^2$, in the third $1/36\pi^2$. However, if we control the variation of optical path across each individual groove, then any order can be chosen. Fig. 1 is the curve of the diffraction intensity I (without considering the constant factor L and k). The effect of the finite width reflection line gives out a filtering window. When $k = 0.4$ and $N = 6$, namely -2, -1, 0, +1 and +2 diffraction orders are visible only. The other diffraction orders are missing. When $\Psi = m$, there is a bright fringe and I reaches its maximum. The simple condition is $-\frac{2P}{\lambda} \le m \le \frac{2P}{\lambda}$ due to $-2 \le \sin\theta_i + \sin\theta_m \le 2$, that is, the pitch of grating has influence on the visible diffraction fringes.

2.2. Tunable Gratings

In grating design, when the grating pitch P, the incident beam wavelength λ and the angle θ_i are fixed, the diffraction efficiency of the different orders is determined by the grating line profile. The major consideration for the design of an individual grating element is to ensure a minimum power in +2 order so that the most of input power can be distributed in 0 and ±1 orders uniformly. The value of y will vary depending on the position of the upper Poly 2 grating. Three positions, which span a full period, are indicated for the following three values of γ: (a) γ = 0 μm, (b) γ = 2 μm and

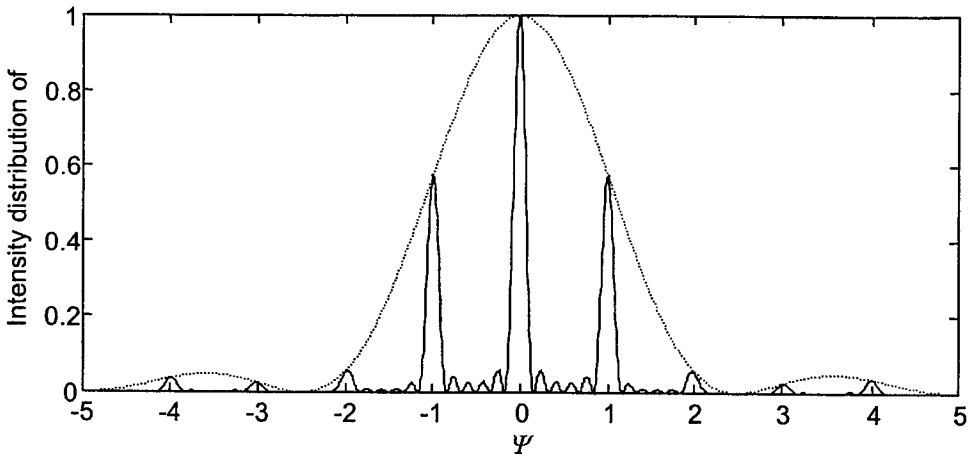


Figure 1. The diffraction intensity distribution of the finite width reflection lines (k=0.4, N=6)

(c) $\gamma = 6$ μm. The Poly 1 grating suspended by spring flexures can be pulled down electrostatically. The grating intensity distribution of the diffracted beam is then dominated by two parameters: the fill factor w/P, and the depth h. Diffraction efficiency is symmetrical about fill factor $w/P = 0.5$ in the range of ($0 < w/P < 1$).

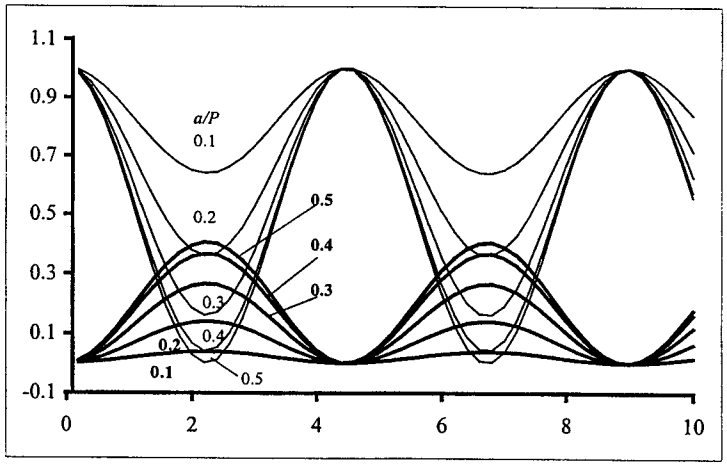


Figure 2. Diffraction efficiency versus the grating groove depth h

Fig. 2 shows the simulation results of the relative optical power P of the diffracted beams versus the grating groove depth h with respect to different values of the fill factor w/P. The grating parameter in the vicinity are: $w/P = 0.4$, and $h = 1.8$ μm for grating 1. In order to facilitate the fabrication process, the final parameters of grating are chosen as $w = 1.5$μm, which indicates $w/P = 0.375$, and $h = 1.5$μm. The grating design is considered compatible to use the polysilicon fabrication process and design rule. The corresponding relative power distribution is 29% for 0 order, 26% for ±1 orders and is less that 4% for +2 order.

2.3. Variable Blaze Gratings

Blazing is contouring individual grating elements and can be used to improve the ratio of light diffracted in a specific direction to the incident light, increase the intensity of light diffracted into a specific interference order. Variable blaze gratings (VBGs) is operated by adjusting the blaze angle of each slat so that the specular reflection of the incident light

matches a particular interference order. For normal incident light, the blaze angle ϕ_m required to select a particular interference order m is given by

$$\theta_i - \phi_m = \theta_m + \phi_m \tag{3}$$

where the blaze angle is measured with respect to a plane parallel to the face of the grating. For example VBCs blazed grating parameters are designed as the pitch $P = 8\mu m$, tilt angle $\phi = 6.9^{\circ}$, total grating lines $N = 6$, and the incident light wavelength $\lambda = 1300nm$, the incident angle $\theta_i = 0$ (normal incident). The main energy is blazed to the 3rd diffraction order and $k \approx \cos\phi = 0.99$, and the diffraction intensity distribution is shown as Fig. 3. It is clear that most of the energy is focused on the only one fringe (+3th diffraction order), all the other diffraction orders are not visible. Micromachined VBGs is used to improve the frequency sensitivity of these spectral analyzers. The VBGs have higher diffraction efficiencies than typical transmission grating, and the high diffraction efficiency increases the strength of diffracted light on photodetector array elements of optics switches.

3. OPTICS SWITCH DESIGN

Optical fibers are an attractive medium for signal transmission in optical communication as well as in optical metrology. There is an increasing need for higher versatility of fiber-optic systems with special requirements for reconfiguration, variable access to one or N channels. These requirements can be met by an implementation of fiber-optics switches of various types. For each particular application there are well-defined parameters to be met by the switch with respect to: the number of input and output channels, insertion loss, cross-talk, switching speed, temperature, vibration stability. The principle of optics switching scheme uses micro-optical components that are translated laterally to optical beams in order to accomplish beam deflection. A number of elements or combination

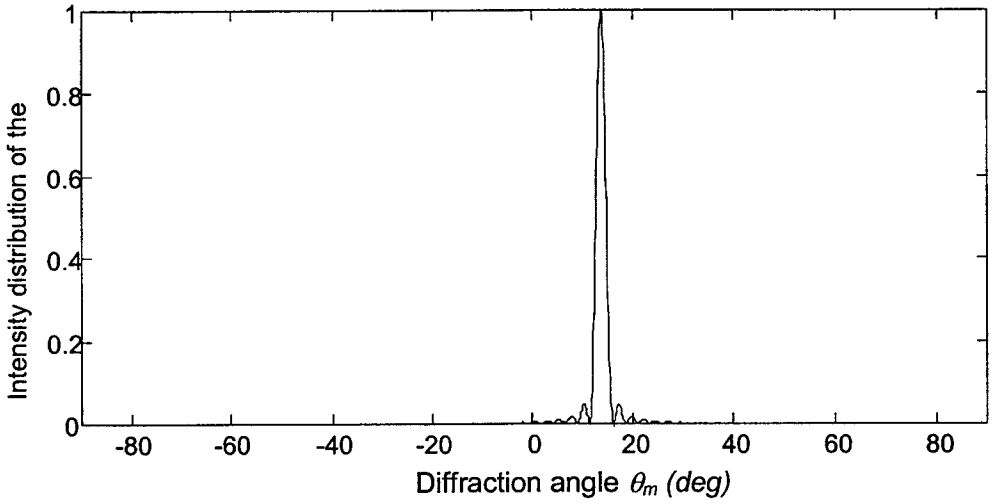


Figure 3. The intensity distribution of a blazed grating along the diffraction angle (3rd order)

of elements such as mirrors, prisms, and gratings, can deflect incoming beams. As agile switching is required for most of the applications, displacements have to be guaranteed by the mechanical actuators. As a separation of the single beams is achieved by deflecting beams, the angular overlap of the deflected beams, which subsequently enter different output channels, has to be small. For a Gaussian beam with a waist radius ω_0, a wavelength λ and a numerical aperture θ_0 given by $\theta_0 = \lambda/\pi\,\omega_0$, the overlap η of a beam deflected by an angle with an undeflected beam with waist radius ω_0 is given by [6-8]

$$\eta = \exp\left(-\frac{\theta^2}{\theta_0^2}\right) \tag{4}$$

Therefore, the cross-talk χ of the beam is given by

$$\chi = 10\lg\left[\exp\left(-\frac{\theta^2}{\theta_0^2}\right)\right] \tag{5}$$

In communication networks cross-talk below –50dB is required. The deflection angle θ has to exceed 3.4 θ_0.

3.1. Optical Grating/Mirror Switch

The advantages of the concepts are basically due to the discrete or continuous steering ability of the various concepts. Discrete beam steering is accompanied by low required positioning accuracy with respect to the switching element that decreases the requirements to the actuator. The micrograting and micromirror can construct 1 × N switches to steer an incoming collimated beam continuously and two-dimensionally. The switch route signals from one channel to a number of output channels with varying numbers of output channels depending on the concrete situation. The 1 × N switches can be used to build up complex networks as well. A diffraction grating can diffract a monochromatic incident beams into several orders. The diffraction angles and diffraction efficiency, i.e. the ratio of light diffracted can be changed by structure parameters of the grating and the wavelength of the incident beam in a specific order to the incident light. The optimal values of them will be determined by the diffraction equation as follows:

$$m\lambda = P(\sin\theta_i + \sin\theta_m), \qquad m = 0, \pm 1, \pm 2, \pm 3, \ldots, n \tag{6}$$

where m is the diffraction order, θ_m is the diffraction angle, P is the grating pitch, and λ is the incident beam wavelength. The value of λ is taken as 1300nm, and the value of P is taken as 4 μm in favor of fabrication. When the angle α equals to 23°, the diffraction beams of 0, ±1, +2 order can exist, and the +2 order beam count-strokes back to the direction of the incident beam. Thus this useless beam is automatically made inactive. We don't need to suppress it by absorbing or masking.

3.2. Device Configurations

A gratings/mirrors switch can be designed as shown in Fig. 4. The proposed 1 × 3 grating/mirror switch consists of two wafer switching platforms. One input fibre with hemispherical end and one 3-PIN photodetector are mounted on the upper wafer platform. The hemispherical end acts as a lens to focus the laser source into a 15μm spot. In the bottom wafer platform, two reflective gratings with size of 30 × 50 μm^2 are etched and coated with gold. A mirror of 30 × 50 μm^2 is placed between these two gratings. As the diameter of the diffraction beam at the upper platform is about 100 μm, the distance between two adjacent photo-detectors should be set larger than 200 μm to prevent the overlap and interference of the diffraction beams from the adjacent diffraction orders. The exact distance between two photodetectors is actually determined by the incident angle α of laser beam and the distance H between upper and bottom wafer platform. Further consideration is required to determine the positions of all the three photodetectors. Our initial choice of α is 20°~25° and H is 100~200 μm.

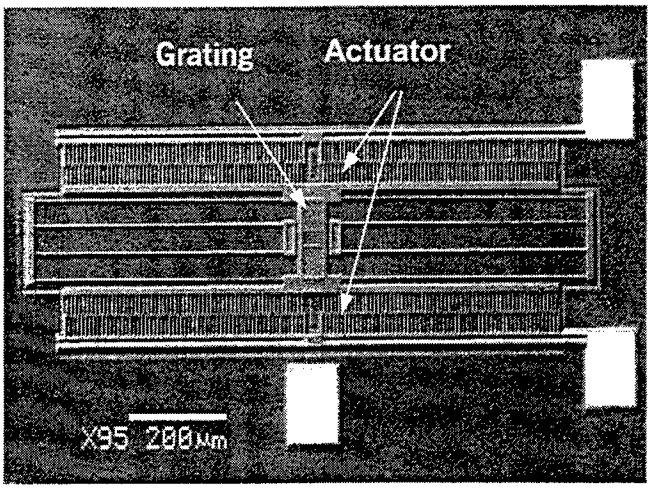


Figure 4. SEM photograph of grating/mirror platform is driven by comb-driver actuators

The central of gratings and mirrors is supported by four suspended beams and driven by two comb-drivers. As shown in Fig 4, the supporting beams have a U shape and are connected to the frame close to its center. The incident beam is focused on the surface of one grating element. It is equally transmitted to all three-output photodetectors 1, 2, and 3 located at the upper wafer platform. The central of gratings and mirrors is driven to the working position from –25μm to +25μm large displacement. When the switching platform is displaced +25 μm driven by the comb-driver, the input laser beam is incident onto the mirror. It is well-known that the zero order diffraction angle is equal to the corresponding reflection angle, and the reflected beam reaches photodetector Output 2.

4. FABRICATION PROCESS AND EXPERIMENTAL RESULTS

The fabrication of the grating/mirrors uses three-layer-polysilicon and wafer to wafer bonding process. The gratings, the mirror and the actuator are fabricated on to the bottom wafer using the 3-layer-polysilicon process as shown in Fig. 5. The two wafers are integrated using wafer to wafer bonding process. The gratings and mirror are constructed using surface micromachined polysilicon process. MUMPs offers three patternable layers of polysilicon, and two sacrificial layers of doped glass on a base layer of silicon nitride. A top layer of gold is used as the reflective surface for the grating. All polysilicon depositions are made by low pressure chemical vapour deposition (LPCVD) with fine-grained polycrystalline silicon at 580°C. The fabrication of the grating and the mirror, including the electrostatic comb-drive, requires two depositions of polysilicon. The original polysilicon layer provides a voltage reference plane and electrical interconnection, while the remaining two polysilicon layers form the mechanical structures. The polysilicon is referred to as POLY 0, and the first and second mechanical polysilicon films are referred to as POLY 1 and POLY 2 respectively. To begin, a silicon substrate is coated with dielectric isolation films of LPCVD silicon-rich nitride over thermal oxide as a blanket starting point. The reason for the blanket isolation films is to ensure that proper electrical isolation is achieved between electrically active parts of the comb-drive structures. The first patterned layer, which uses mask #1, is the electrical interconnect and shield polysilicon of POLY 0. The electrostatic comb-drive and suspending anchor areas and the stiction-reduction dimple molds are patterned into the first thick sacrificial glass layer deposited after POLY 0, with masks #2 and #3, respectively. The subsequent polysilicon film to be deposited fills in the anchor and dimples on the otherwise flat underside of the polysilicon for stiction reduction. At two intermediate steps, namely after POLY 1 and POLY 2 depositions, suitable post-deposition annealing must be done to ensure that the polysilicon mechanical films do not display undesirable internal stress that would distort the structure upon the final release etch. As a result, the surface micromachining technique offers a new approach to implement the grating, mirrors, and driving actuators. After fabrication, the device is released by removing sacrificial glass layers in a bath of 49% hydrofluoric acid for 3.5 min followed by a short rinse in deionized water. The two wafers can be bonded together along with active devices such as optics fibre and photodetectors onto a silicon optics switch chip.

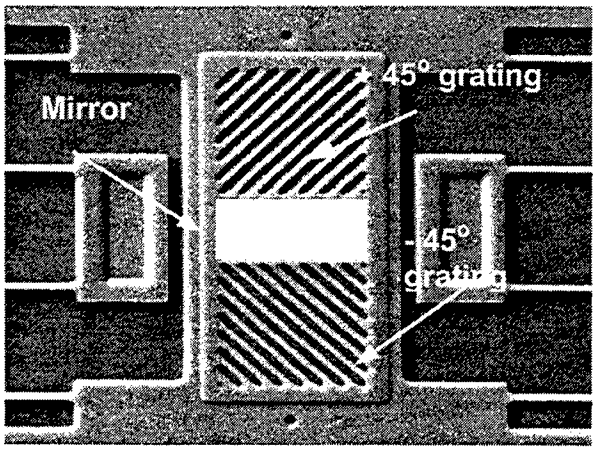


Figure 5. SEM photograph of grating and mirror switch

The grating switch experimental set-up is shown in Fig. 6. The experimental system set up consists of a diode laser, a zoom beam expander, laser focusing lens, 2-D translation and rotation stage and a photodetector. Measured reflectivities were obtained in air with a He-Ne laser operating at a wavelength of 632.8 nm, and are in good agreement with theoretical results. The grating switch chip is fixed on the 2-D translation stage. The lens is able to focus the expanded collimated beam to 15 μm. The rotation stage is used to control the incident angle of the input laser beam. The measured relative diffraction powers are: 21% for 0 order, 15% for +1 order, 16% for +1 order, and 5% for +2 order that the experimental results are in good agreement with theoretical analysis.

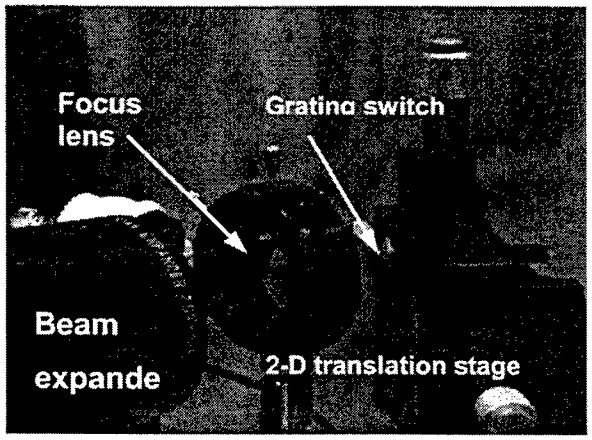


Figure 6. Grating/mirror switch experimental set-up

CONCLUSIONS

The design and fabrication of a micro-opto-mechanical grating switch using three-layer polysilicon surface micromachining polysilicon process is reported in this paper. Based on theoretical diffraction analysis and grating parameters optimization, grating line profile is determined as: $w = 1.5\mu m$ ($w/P = 0.375$) and $h = 1.5\mu m$. The corresponding relative power distribution in percentage of input beam power is 29 % for 0 order, and 26 % for ±1 orders. The prototype of the 1 × 3 terminal switch is successfully fabricated. Further work on analyzing the influence of convergent beam on grating diffraction and on different fabrication processes of gratings with different profiles will enable the grating/mirror switch to be widely used in communication networks.

REFERENCES

1. H. Toshiyoshi and H. Fujita, "Optical crossconnection by silicon micromachined torsion mirrors", *Digest of IEEE/LEOS 1996 Summer Topical Meeting*, pp. 63-64, Keystone, Co, August, 5-9, 1996.
2. S. S. Lee and M. C. Wu, "Surface-micromachined free-space fiber optic switches with integrated microactuators for optical fiber communication", *International Solid-State Sensors and Actuators Conference – Transducers'97*, pp. 85-88, Chicago, IL. June 16-19, 1997.
3. H. S. Hinton, *An introduction to photonic switching fabrics*, Plenum Press, New York, 1993.
4. T. Glaser, S. Schroter, H. Barlelt, "Beam switching with binary single-order diffractive gratings", *Optics Letters*, **23**, No. 24, 1998.
5. D. B. Payne, *WDM techniques in broadband optical networks in Optical technology and wideband local networks*, The Royal Society, London, 1989.
6. J. W. Goodman, *Introduction to Fourier optics*, McGraw-Hill, New York, 1998.
7. D. A. Gremaux, N. C. Gallagher, "Limits of scalar diffraction theory for conducting gratings", *Applied Optics*, **32**, No. 11, 1948-1953, 1993
8. S. Glockner, R. Goring, and B. Gotz, "Multichannel fiber-optics switches based on MOEM systems," *Proceedings of SPIE, Microelectronics Structures and MEMS for Optical Processing III, Vol. 3226, Austin, USA, 29-30 September, 1997*, pp76-85.

Comparison of grating structures in optical fibers and integrated optics for MEMS applications

Natarajan S. Ramaswamy, V.M. Murukeshan and Anand Krishna Asundi

School of Mechanical & Production Engineering
Nanyang Technological University
Nanyang Avenue, Singapore 639798

ABSTRACT

With the emergence of new trends in optical communication and sensing systems, the importance and need for novel concepts such as diffractive gratings that can be implemented on a microscale has become a thrust area of research. In this paper, we compare various grating structures and their suitability for MEMS (Micro Electro-Mechanical System) applications. The design approach, analysis and integration to optical systems are discussed in the context of both fiber and planar waveguide technology. Also, the differences and similarities among these methodologies using silicon micromachining are brought out. Parameters of interest, such as efficiency of micro-grating structures and ambient effects that influence the performance of the devices/systems are also discussed.

Keywords: MEMS, gratings, FBG, integrated optics, sensors, optical switching, wave guide

1. INTRODUCTION

A host of applications for micro-machined components has been in vogue these days. Micro-gratings that include diffraction gratings and similar strutcures, spatial light modulators using cantilevers and paddles and micro-lens arrays are some of them. The reason for the growth potential of these silicon micromachined devices as components is due to the fact that this technique is best suited to the production of arrays of devices/ components. These devices play an important role in sensor arrays which involve many sensing elements providing a single sensing function. Also, the inherent precision of the technique is ideal for alignment, for example, of individual pairs of optical wave-guides within a relatively large single device. Gratings inside an optical fiber are one of the grating structures which will bee considered with tremendous application in the sensor technology. The other grating structures of interests fall in the design and fabrication of free space optics for micro- rotation, switching and sensing. Grating technique is largely used in the field of metrology, integrated optics, adaptive optics (for space applications), accousto-optics (NDT), optical communication and optical information processing.

Micromachining exploits photolithographic techniques to fabricate structures with extremely precise dimensional tolerances. A wide variety of micromachining techniques has become available with the capabilities to fabricate structures of ever increasing mechanical complexity. The dimensional tolerances are immediately compatible with optical sub-assemblies and the potential for batch production of substantial numbers of identical structures at very low cost is very attractive. This project examines the role of micro-machined structures in the specific context of grating structures both in planar waveguide and optical fiber sensor technology for MEMS applications.

The technique which offers particular potential in approaching these sort of structure assemblies and stability problems is Silicon Micro-machining, and it is the possibilities of the combination of fiber optics and silicon micro-machining which is the principal objective. In this review paper, we compare various grating structures and their suitability for MEMS (Micro Electro-Mechanical System) applications. The design approach, analysis and integration to optical systems are discussed with respect to both optical fiber

Part of the SPIE Conference on Advanced Photonic Sensors and Applications
Singapore • November–December 1999
SPIE Vol. 3897 • 0277-786X/99/$10.00

and planar wave-guide technology. Further, parameters of interest, such as diffraction efficiency of the fabricated micro-grating structure, ambient effects that influence the performance of the devices/ systems are being considered in the analysis. Emphasize is on the individual applications of different gratings for telecommunications, data transfer, micro laser fabrication, signal processing and instrumentation, and sensors. Keeping these in mind, the main objectives of the paper can be divided as (a) design and fabrication of gratings in planar waveguide and fiber structures and (b) their integration for specific applications

2. DESIGN & FABRICATION OF GRATINGS

2.1 Design& fabrication of gratings in optical fibers

The Fiber Bragg Grating is a longitudinal periodic variation of the index of refraction in the core of a single-mode optical fiber (Fig. 1). The fiber Bragg reflectors are written onto the core using a high intensity UV laser source. This phenomenon of refractive index change in the core is due to the photosensitivity of optical fiber (usually confined only to the core) which was first discovered in 1978 [1].

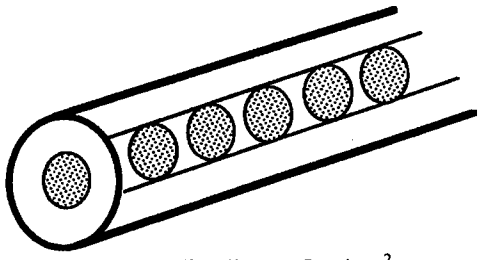


Fig. 1 Fiber Bragg Grating [2]

2.2 Techniques to produce fiber bragg gratings

Optical fibers can be made photosensitive by doping the core of the fiber with germanium, phosphorous, or boron during the fabrication of the preform. Fibers can also be sensitized by loading with hydrogen under temperature and/or pressure — chamber pressurized with hydrogen [3].

The photosensitive optical fibers used were of germanium-doped core silica fiber manufactured by SpecTran Specialty Optics Company. When illuminated with UV light in the region of 248nm, a wavelength corresponding to the absorption band, or 'colour centre', in the Ge/S_2 bonding in the glass is absorbed breaking the bonds causing electrons to move into colour centre traps in the silica structure (Fig. 2). This modifies the absorption characteristics of the glass[4]. This results in the index of the glass being shifted at wavelengths removed from the absorption region through the Krammers-Kronig relationship [5].

To fabricate the FBGs, the outer protective coating of the photosensitive fibers are first removed along the length where the FBG is desired. The FBG is then fabricated using the +1/-1 order phase masks, as shown in Fig. 2 (Phase Mask Method (near field)). The schematic drawing of "Excimer laser proximity mask aligner for fiber grating writing" can be found in [6]. The UV light source used is a 248nm KrF Excimer laser (Lambda Physiks model: COMPex205). Based on the +1/-1 order principle, the phase mask is optimized to diffract the UV light equally into the +1 and -1 orders. Self-interference between these two orders creates an interference pattern into the core of the photosensitive fiber with a distance equivalent to half the pitch of the phase mask. The zero order and higher order noise is highly suppressed by the phase mask. The details regarding the typical pitch of mask gratings can be inferred from the following Table. 1.

Table .1 Details Of Phase Mask

Bragg Wavelength(nm)	1535	1545	1550
Period On Mask(nm)	1061	1068	1071
Grating Size(mm x mm)	10 x 3	10 x 3	25 x 3
Zero Order(%)	4.8	4.2	0.79

The germanium-doped fiber selected in the experiments reduces the magnitude of the short wavelength loss frequently seen in Fiber Bragg Gratings and its similar numerical aperture (0.11±0.02) and mode field diameter (10.5±1.5μm @1550nm) provides minimal splice loss when spliced to standard single-mode fiber (averaging splice loss is about 0.05dB). Furthermore, high reflectivity gratings (at least 20dB) can be easily made in less than two minutes [7]. Figure 3 shows the spectrum of a Fiber Bragg Grating made using this method used in this research.

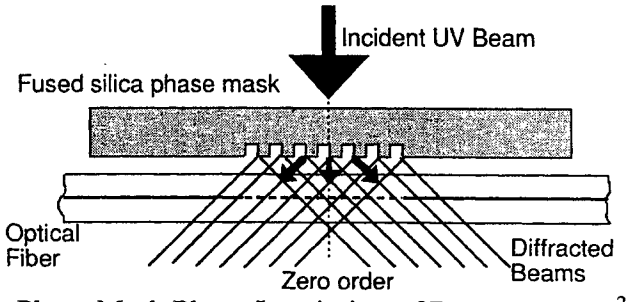


Fig. 2 Phase Mask Photo-Imprinting of Bragg Grating [2]

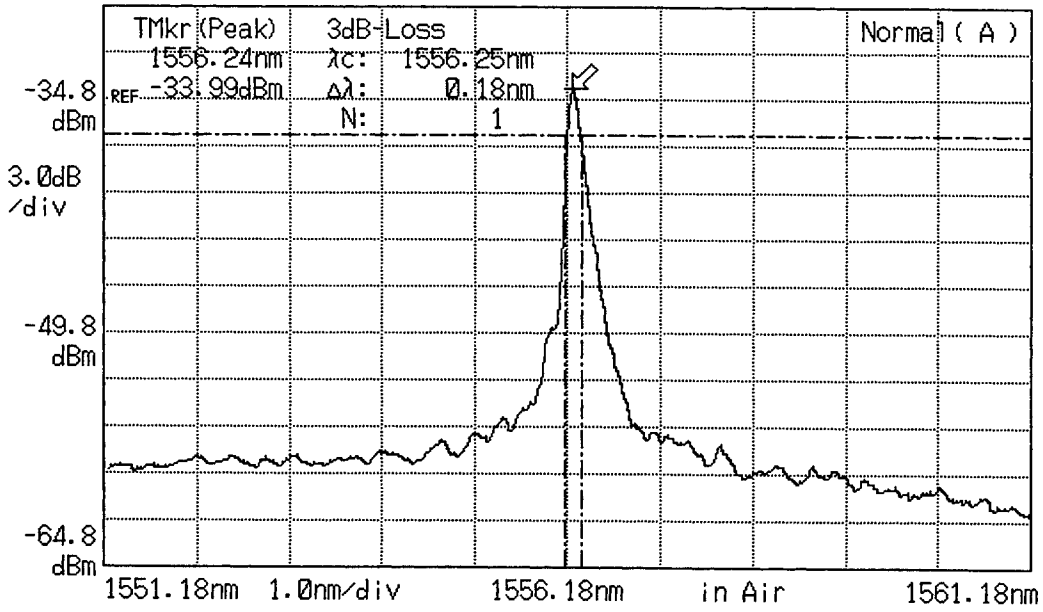


Fig. 3. Spectrum of a Fiber Bragg Grating [23]

Other methods to fabricate Fiber Bragg Gratings include split-beam interferometer [8], point-by-point method or a phase mask (far field method) to form the fringes [9,10,11]. For the split-beam interferometer and phase mask methods, the UV exposure can either be delivered as a series of moderate intensity pulses or in a single high intensity pulse at power approximating the onset of surface damage for silica[12,13].
For the split-beam interferometer technique, there are three different configurations:

1) Transverse Holographic Method (Fig. 4) [2,8]

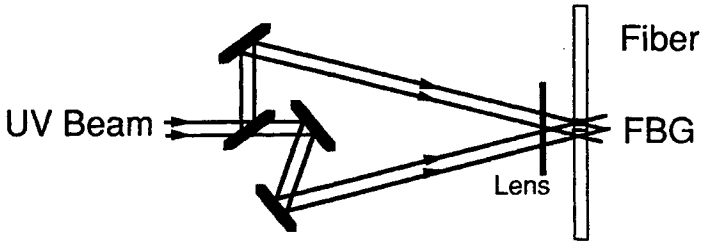


Fig. 4 Transverse Holographic Method

2) Right Angle Prism Method (Fig.5) [2,14]

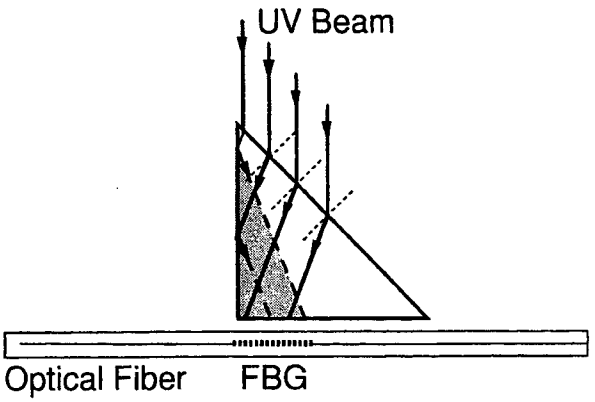


Fig. 5. Right Angle Prism Method

3) Roof Top Prism Method (Fig. 6) [2,15].

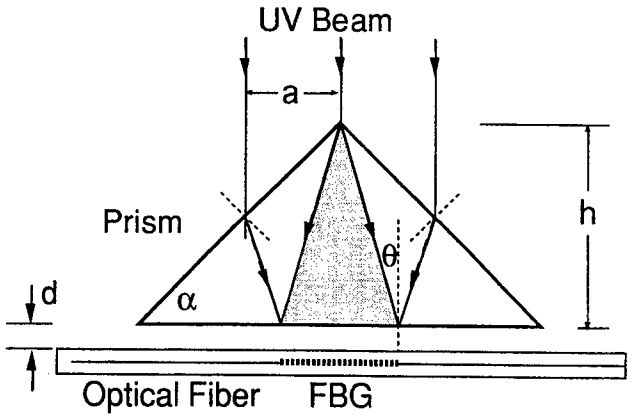


Fig. 6 Roof Top Prism Method

The Relationship between the Prism Angle α and the Bragg Wavelength is definite and this defines the Bragg wavelength.

4) As for the point-by-point method, the individual gratings are written one at a time by the incident UV beam while the optical fiber is transversing (Fig. 7)[2, 16].

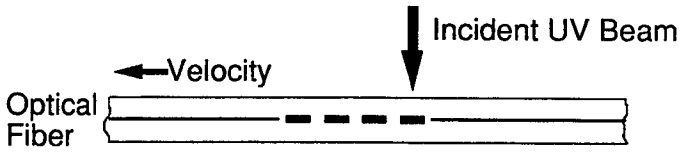


Fig. 7 Point-By-Point Method

5) For the Phase Mask Method, in addition to the near field Phase Mask Method, there is the far field Phase Mask Method (Fig. 8) [2,14]

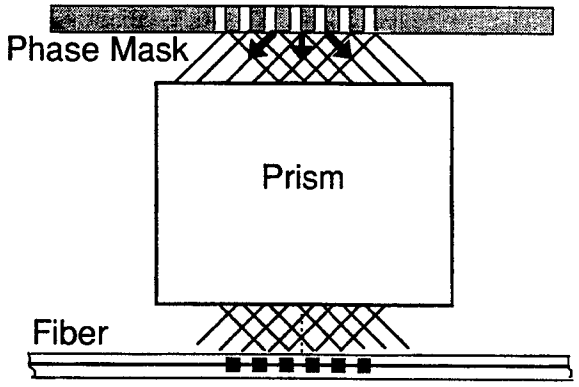


Fig. 8 Phase Mask Method (Far Field)

Comparing the techniques, the Holographic method is versatile but needs a high coherence source. Both the Right Angle Prism and the Roof Top Prism methods are stable. For the Point-by-Point method, though it is time consuming, it is highly precise. Finally, for the Phase Mask method, its advantage is that it only needs a short coherence for its source.

2.3 Types of gratings in optical fibers

Based on the different expoosure condditions different types of gratings has been identified in optical fibers [16,17]. They can be classified as Type-I, Type II and Type III gratings.

The gratings formed under exposure to continuous wave (CW) or weak multiple pulses are known as Type-I gratings. In this case the change in the index is quite uniform all across thw core and the gratings exhibit a nearly ideal transmission spectrum. Such gratings can be written normally on germano-silicate fibers.
Thee gratings belongs to Type-I obtain due to normal photorefractive effect. In certain cases two distinct cases of gratings are formed due to the optical break-down at the core-cladding interface. Such gratings are known as Type-II gratings possess two distinct kinds of gratings with values below and above a threshold limit. Type-III gratings are mainly associated with high numerical aperture (NA) fibers.

3. MAIN APPLICATIONS

3.1 Switching applications in the context of OPTICAL MEMS Technology

Recently, there has been a rapid surge in the growth of fiber optic network[18-20]. This has prompted scientists and engineers to design and develop fast, high performance broadband local area networks (LANs). Methods for producing optical components at low cost and in huge quantities have thus become suitable approaches to meet the trend. In this regard, optical switches, which forms the most essential component in LANs, have been receiving wide attention.

Micromachining using the MEMS technique can be adopted for designing optical switches that are compatible with the fiber optic networks. Broadly speaking, the switch can be realized by two methods -

(a) Free space approach
(b) Waveguide approach

As the name suggests, in free space approach, switching is takes place in air using some form of actuation. The most favored methods are based on electrostatic or magnetic actuators. A variety of mechanical design methodologies for optical switching can thus be conceived using the above approaches. The basic advantages of free space optical switches include good extinction ratio and high reliability.

Possible approaches include exploiting the surface tension properties of suitable index matching liquids at the crossing points of a switch, or the use of Bragg gratings. In the Bragg grating technique, the input light can be modulated to separate into two beams -one reflection and the other a transmission, based on the Bragg wavelength. The Bragg grating approach has an edge over other methods in terms of its low insertion loss. The inherent advantage of the waveguide optical switch is that it subdues electromagnetic noise, which is a leading cause of worry in micromachines.
The biggest advantage of using fiber bragg gratings based networking and switching is their high speed of (~terrabits) [18].

In the wave-guide approach, the switching action is brought about in the optical path, usually a medium other than air. Planar optical structures that are compatible with fibers form the optical path, called the wave-guide. The switching can be brought about using optically powered actuators. Tiny laser diode transfers optical energy to a micro pump that then operates to produce the necessary switching action. A detailed account on the design and applications of gratings in Integrated Optic devices will be presented.

3.2 Use of fiber Bragg gratings for sensing applications

The use of fiber gratings as sensing elements has been very well established and being used for a variety of sensing applications exploiting their inherent properties[21-25]. As explained in the previous sections, the FBG has a longitudinal periodic variation of the refractive index in the single mode fiber core. Normally, such intra-core gratings are written using high intensity UV laser source. Wavelength of the light reflected at this fiber depends on the gratings spacing, i.e. spacing between the refractive index variations. Light travelling inside the fiber core will be scattered in the regions of varying refractive indices. The scattered light will generally be out of phase and thus, tend to cancel, except when the light reflected from each perturbation in the gratings adds up constructively. Thus, forms the back-reflected light with a center wavelength known as Bragg wavelength. The condition to be satisfied to get such a Bragg reflected wavelength is known as Bragg condition and is given by[2,22]

$\lambda_B=2nD$,

where λ_B is the reflected Bragg wavelength, n is the average refractive index of the fiber core, and D is the grating spacing. The spacing of the periodic variation of refractive index, which is a function of strain, will be elastically deformed with respect to the external strain on the fiber. Also, due to the strain-optic effect the average refractive index will be changed. These two affect the Bragg wavelength to be shifted for an applied strain.

This can be expressed as

$$\frac{\Delta\lambda}{\lambda}=K\varepsilon \qquad (1)$$

where K is a constant and ε is the applied strain.

The reflected light is a narrowband of the source bandwidth and the reflected bandwidth depends on several parameters, especially the grating length, but is usually about 0.05 to 0.3nm in most fiber optic sensor applications. When subjected to perturbations (for example strain or temperature change), the Bragg wavelength will be shifted[2,7,22]. The input spectrum, transmitted and reflected signal is shown in Fig. 9.

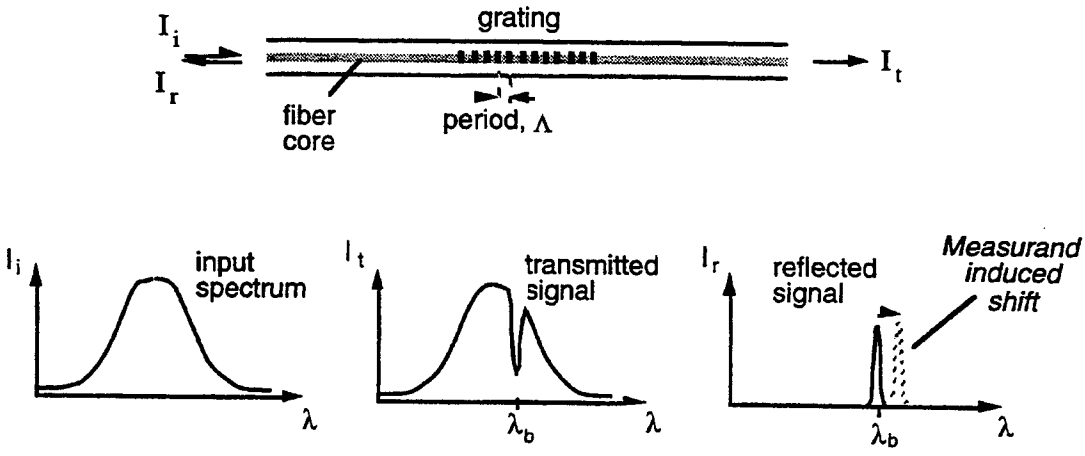


Fig. 9 Basic Bragg Grating Based Sensor System

The two primary measurands that are of interest are the effect on the Bragg wavelength due to strain and temperature. When subjected to strain, both the physical elongation of the optical fiber (and corresponding the change of the grating pitch) and the change of the refractive index of the fiber due to photoelasticity affects the shift in the Bragg wavelength and is given by

$$\Delta\lambda_B = (1-p_e)\,\lambda_B\varepsilon \tag{2}$$

where ε is the applied strain, and p_e is an effective photoelastic coefficient term given by:

$$p_e = (n^2/2)(P_{12}-\nu(P_{11}+P_{12})) \tag{3}$$

where the P_{ij} coefficients are the Pockel's coefficients of the strain-optic tensor and ν is the Poisson's ratio. The temperature response of the Bragg wavelength is due to two primary effects: the change of refractive index of the glass due to temperature change and the thermal expansion of the glass (corresponding to the change in the period of the grating). For silica fibers, the former has a more dominant effect, accounting for about 95% of the observed shift in the Bragg wavelength. The Bragg wavelength shift is given by

$$\Delta\lambda_B = \left(\left(\frac{\delta\Lambda}{\delta T}\right)/\Lambda + \left(\frac{\delta n}{\delta T}\right)/n\right)\lambda_B\Delta T \tag{4}$$

where the first term of the equation is the shift of Bragg wavelength due to thermal expansion and the second term the refractive index change due to temperature change.

In other words, the spectral shift of the Bragg shift corresponds to an applied strain or change in temperature. This principle is applied for sensing various parameters by way of direct or indirect analysis. Recently the FBG sensor system has been extensively used for various sensing applications especially for the cure monitoring of smart composites [22,25]. The performance of the embedded FBG sensor smart composite specimens under three-point and four-point bending conditions is also being investigated. These micro fiber gratings has a definite advantage that it is wavelength encoded shifts in the spectrum is the interrogation parameter and is uniquely associated with each grating. Strain or temperature change on the grating is reflected as shift in the reflection or transmission spectrum. Thee limit of this hence is due to the fact that they are sensitive to both strain and temperature and a single measurement of the Bragg wavelength shift cannot distinguish between their effects. Anyway several techniques are being proposed to discriminate these effects[25]. This sensing methodology with micro gratings in fiber finds tremendous applications in variety of military, aerospace and civil engineering applications.

4. ROLE OF GRATINGS IN IO DEVICES

Integrated optical traducers with on-chip variables along with gratings and other structures has been found suitable for realizing micro-optical switches and sensors. Different configurations have come-up based on tapered wave guides and chirped grating couplers[26]. Also, the study on the concepts for realizing the TIOS (Totally Integrated Optical Sensors)[27] , the role of grating structures in such IOs for switching and networking applications[28] will be discussed in the presentation.

5. CONCLUSION

This paper reviews the recent developments in the gratings technology in the context of optical MEMS or micro-optic devices for various sensing and switching applications. A detailed account on bragg gratings in fiber has been reviewed here and the gratings in IO context will be dealt while presenting the paper.

6. ACKNOWLEDGEMENTS

The authors acknowledge the financial grant from MoE/NTU through grant MLC1/97.

REFERENCES

1) Hill, K.O., Fuji, Y., Johnson, D.C., and Kawasaki, B.S. Photosensitivity In Optical Waveguides: Application To Reflection Filter Fabrication. *Appl. Phys. Lett.*, **32**, pp. 647-649, 1978.

2) Brown, D.A., and Zhang, Q. Optical Fiber Bragg Gratings: Principles And Applications. Fiber Optic Sensors Video Series. Bellingham, Wash: SPIE., 1995
3) Juma, S. Bragg Gratings Boost Data Transmission Rates. *Laser Focus World*, pp. S5-S9, 1996.
4) Hand, D.P., and Russell, R.S.J. Photoinduced Refractive-Index Changes In Germanosilicate Fibers. *Optics Lett.*, **15**, pp. 102., 1990
5) Kersey, A.D., Davis, M.A., Berkoff, T.A., Bellemore, D.G., Koo, K.P., and Jones, R.T. Progress Towards The Development Of Practical Fiber Bragg Grating Instrumentation Systems. Proc. SPIE, 2839, pp. 40-63, 1996.
6) PS2000 Excimer Laser Projector Handbook And Operating Instructions, Exitech Limited, 1998.
7) Photosensitive Fiber, Spectran Specialty Optics Company, Ct, USA, 1996.
8) Meltz, G., Morey, W.W., and Glenn, W.H. Formation Of Bragg Gratings In Optical Fibers By A Transverse Holographic Method. *Optics. Lett.*, **14**, pp. 823., 1989
9) Meltz, G. Overview Of Fiber Grating-Based Sensors. Proc. SPIE, 2838, pp. 2-22, 1996.
10) Hill, K.O. et al. . Bragg Gratings Fabricated In Monomode Photosensitive Optical Fiber By UV Exposure Through Phase Mask. *Appl. Phys. Lett.*, pp. 1035, 1993.
11) Anderson, D.Z., Mizrahi, V., Erogan, T., and White, A.E. Production Of In-Fiber Grating Using A Diffractive Element. Elect. Lett., **26**, pp. 730-732, 1990.
12) Askins, C.G., Tsai, T., Williams, G.M., and Friebele, E.J. Stepped-Wavelength Optical-Fiber Bragg Grating Arrays. *Opt. Lett.*, **19(2)**, pp. 147-149, 1994.
13) Putnam, M.A., Askins, C.G., Williams, G.M., Friebele, E.J., Baskansky, M., and Reintjes, J. Single Pulse Fabrication Of Fiber Bragg Gratings Using A Phase-Conjugated Krf Excimer Laser. Proc. SPIE, 2444, pp. 403-409, 1995.
14) Kashyap, R., Armitage, J.R., Campbell, R.J., Williams, D.L., Maxwell, G.D., Ainslie, B.J., and Millar, C.A. Light Sensitive Optical Fibres And Planar Waveguides. *BT Technology Journal*, **11(2),** pp. 150-159, 1993.
15) Zhang, Q., Brown, D.A., Reinhart, L., and Morse, T.F. A Simple Prism Scheme For Fabricating Bragg Gratings In Optical Fibers. *Optics Letters*, 19(23), 1994.
16) Malo, B., Hill, K.O., Bilodeau, F., Johnson, D.C., and Albert, J. Point-By-Point Fabrication Of Micro-Bragg Gratings In Photosensitive Fibre Using Single Excimer Pulse Refractive Index Modification Techniques. *Elect. Lett.*, **29(18),** pp. 1668-1669, 1993.
17) A. Selvarajan, Bragg grating sensors for smart structure applications, Proc. International Conf. On Smart materials, structures and systems, pp. 174-181,1999.
18) R. Winhardin All – optical neetworks may onee day form national backbone: An inteerview with R. Ramaswamyy and C. Qiao, OE Reports, Aug.1999.
19) L. Zhang., Sugden. K ., Bennin. I. And Molony. A., Wide stop band chirped fiber Moire grating transmission filters, Electronics Letters, Vol. 31, No. 6, 477-478
20) T. Galser., Schrotter. S., and Barlet. H., Beam switching with binary single order diffractive gratings, Opti. Letts., Vol. 23, No.24, 1998.
21) B. Fabeineek, K. Fitcheew, S. Myken and A. Houghton. Optical network Research and Deveelopment in European Community programs: From RACE to ACTS. IEEE Communications magazine pp. 50-75, 1997.
22) P.A. Crossby, G.R. Powell, G. F. Fernado, D.N. Waters, C.M. France and R. C. Spooncer. (1997). A Comparative Study Of Optical Fiber Cure Monitoring Methods, Proc. SPIE, 3042, pp.141-153.
23) Murukeshan, V.M., Chan, P.Y., Ong, L.S., and Seah, L.K. Cure Monitoring Of Smart Composites Using Fiber Bragg Grating (FBG) Based Embedded Sensors. (In Sensors & Actuators).
24) Murukeshan, V.M., Chan, P.Y., Ong, L.S., and Asundi, A. Performance Evaluation Of Fiber Polarimetric And Intra-Core FBG Sensors As Self-Monitoring NDT Devices For Smart Structure Applications. (Presented elsewhere- in press).
25) Murukeshan, V.M., Chan, P.Y., and Ong, L.S. Intra-Core Fiber Bragg Gratings For Strain Measurement In Embedded Composite Structures. (Published elsewhere).
26) R.E. Kunz and L. U. Kempen Miniature integrated optical sensors. SPE Vol. 2068, pp. 69-86, 1998.
27) A. Doyle and B. D. MacCraith. Fabrication of sl-gel based planar waveguide/grating coupler platforms for use as optical chemical sensors. SPIE Vol. 3105, pp.61-70, 1997.
28) M. Renaud and F. Masetti. Network and system concepts for packet switching. IEEE communications Magazine. Pp. 96-102, 1997.

SESSION 11

Micrometrology

Testing Method for Microprofile of Optical Supersmooth Surface①

Jianbai Li, Shaorong Xiao*, Xiaoyun Li,

Aihan Ying, Xiaoling Zhang, Anqing Zhuo

Jiangxi Academy of Sciences, Nanchang 330029, Jiangxi, China;

* Nanchang University, Nanchang 330047, Jiangxi, China.

ABSTRACT

In this paper, some advantage on testing method for optical supersmooth surface using Atomic Force Microscope (AFM) are presented, and compared with classical interference method. Important applications of AFM for optical technique are introduced in the paper. Some testing results on microprofile and roughness of optical supersmooth surface using AFM method are listed, and it has indicated that result isn't obtained with interference method.

Keywords: micro-metrology, optical testing, microprofile, AFM, supersmooth surface.

1.INTRODUCTION

Recently developing of optical technology, especially used of short wave length and high-intensity laser, the quality for morphology and microprofile of optical elements is required to higher level. For example, wave length of soft x-ray is about 10~20 nm. According to surface roughness of optical elements being 1/50~1/100 wave length of used ray, then surface roughness in x-ray optical system may be about 0.1~0.5 nm RMS[1].

Atomic Force Microscope (AFM) is one of some Scanning Probe Microscope (SPM). This instrumentation is available to use for testing surface microprofile and roughness of above mentioned nonconduction with nanometer or subnanometer accuracy. AFM is consist of scanning probe of Si_3N_4; laser lever system from semiconductor laser; position sensitive detector(PSD) with four zone; micromoving stage of a piezoelectrical tube(PZT) scanned X,Y plane and controlled the feedback of Z axis; inspecting and controlling circuits and microcomputer hardware and software etc. On some details of this instrumentation reader can see ref. [2].

Comparing with another method, AFM method has both good vertical and lateral resolution, about 0.01 nm and 0.1 nm respectively. The classical interferometer method, including recent common-path interference profiler, has a poor lateral resolution which is low 10^3~10^4 comparing with vertical one[3]. AFM can find nanometer-grade micro-defect at surface microprofile, but classical interferometer method can't. Therefore, the testing result using AFM may be true microprofile of optical supersmooth surface, but interference method only get a statistical mean

①The item of Chinese National and Jiangxi Province Scientific Fund

Part of the SPIE Conference on Advanced Photonic Sensors and Applications
Singapore • November–December 1999
SPIE Vol. 3897 • 0277-786X/99/$10.00

"profile" ,which isn't true microprofile of the surface. So that. Authors think that AFM method has many advantages which is available for evaluating microprofile and micro-defect of optical supersmooth surface

2.MEASURIMENT RESULT AND ANALYSIS

Some measuriment results which listed in presented paper are used by Nanoscope III instrument made in Digital Instruments Inc. Testing data used this AFM instrument listed in ref. [4],[5] are in agreement with theoretical value of known structural constand of Au(111)and mica surface. These tested results shown that it is available for measuriment microprofile and roughness of optical supersmooth surface using this instrument. The scanning probe using in presented testing is NP normal model made in DI Inc. too. Arm length of the probe cantilever is 200 μm, force constant is 0.12 Nm^{-1}. The head of probe is normal four-pyrimad with length of bottom side 4 μm and angle of pyrimadal top 70° . Testing results listed in this paper are gained at following conditions: Some samples were imaged with the AFM in contact mode, in air at room-temperature. All images were recorded in height mode, and processed to remove low-frequency noise.

2.1. Testing on true morphology and microprofile of optical supersmooth surface.

True morphology and microprofile of optical supersmooth surface is consist of many peaks and valleys with nanometer or subnanometer grade, and their vertical and lateral dimension have nanometer or subnanometer grade too. Therefore, only AFM method is available for testing true microprofile of supersmooth surface, as it has both good vertical and lateral resolution with subnanometer grade, but another methed (e.g. inteferometer method) hasn't.

As example, testing result listed at fig. 1 is true morphology of supersmooth surface of a silicon wafer with zone 1 ×1 μm at X,Y direction respectively. Z-direction is vertical to the surface. Rising and falling hight of microprofile in Z-direction is represented by gray density which the lighter gray is the taller of the hight. The gray density scale listed at right-up side in fig. 1. The bigger white peak positioned 0.50 μm in X-axis is nanometer-grade micro-defect of the supersmooth surface, and rising hight about 5 nm. Recently, highest roughness standard is 12 nm RMS in Chinese and International Standard of Roughness[6][7]. Testing results listed in this paper, contained above mentioned micro-defect of supersmooth surface, have been beyoud highest grade of roughness in these standards.

2.2. Evaluating microstructure and calculating roughness etc. before and behind coating for supersmooth surface.

The surface roughness and microprofile for substrate of multilayered reflective mirror may be below 0.1~1.0 nm RMS. Coated multilayer for this surface, it is as best as possible to keep nanometer or subnanometer roughness and microprofile. Therefore, it is available to test this surface quality by AFM method before and behind coating. Two-dimension density (Z-direction hight) picture and sectional drawing in indicating direction of surface microprofile picture before and behind coating are represented fig. 2 and fig. 3 respectively.

Sampling between two black allows about 300 nm in fig. 2(b) and fig. 3(b) section analysis, some paraments on

roughness and microprofile of supersmooth surface at two tested wafers before and behind coationg are calculated in table 1.

surface	RMS Rq	arithmetric mean Ra	p-v value R_{max}	height of Z direction R_Z
polishing	0.319	0.239	1.54	0.86
coated	3.265	2.319	11.525	7.286

Table 1. Roughness paraments, unit: nm

Reviewing form fig. 2, fig. 3 and table 1, surface roughness of two samples before and behind coating have reached nanometer or subnanometer grade, which is greatly beyoud highest degree of Chinese and International Standard (12 nm). The surface microstructure behind coating is appeared many ellipse crystalloid grains, and surface roughness is low than before coating as technology factor of coating.

2.3. Evaluating polished graduated plate

The graduated plate precisely polished in level has highest class roughness in commonly optical elements. Tested graduated plates presented in the paper are provided by Jiangxi Optical Instrument Inc. Reviewing and evaluating polished surface of two samples are listed fig. 4 and fig. 5 respectively. From surface gray density graph in fig.4 you can see, residual profile peaks in surface of a graduated plate are cut as polished ,forming many platforms with nearly level and smooth .The picture of polished platforms apparently are reviewed in section analysis of fig. 4 too. But picture of polished platform in another fig. 5 apparently aren't reviewed.It is becouse of the latter is polished with long tine and formed good platform, but the former is polished with shorter time and formed worse platform as polishing technology factor. Thus ,it can be seen that polished level between tested samples is difference due to evaluating using more accurate AFM method, in spite of two graduated plates are seen as good polished elements evaluating using common optical technology, and reached 2.661 nm and 1.116 nm respectively, beyond highest class of roughness standard (12 nm).

This reseach has important significance for improving optical polishing technology and gaining supersmooth polishing surface.

3.CONCLUSION

Some valuable conclusions have been gained and listed to following using AFM method for testing optical and supersmooth surface, but isn't gained by another method, e.g. common-path interference profiler: For example, the microstructure of polished and coating of optical and supersmooth surface; nanometer or subnanometer-grade defect of the surface; profile section in indicating direction of surface morphology etc. These are very important and significance for polishing and evaluating of optical elements in high-intensity laser and

soft x-ray systems, and it is new development on testing method of surface micrprofile and micro-morphology.

ACKNOWLEDGMENTS

Authors represented acknowledgment for Chirese Key Lab. of Applied Optics and Younger Lab. CAS on Nanometer Sci .& Tech. as their help in experimerts.

REFERENCES

[1] Junping Zhang, Jianlin Cao ,et al., ACTA OPTICA SINICA ,Vol.15,No.8, pp 1088-1093, Aug.1995.

[2] Jianbai Li, et al., Jour. Optoelectronics Laser Vol.9, Supp, pp 272-274, Aug. 1998.

[3] G. Rasigni, et al., J. Opt. Soc. Am. Vol.73, No.2, pp 222-233, Feb. 1983.

[4] DENG Wenli, YANG Daben ,et al, SCIENCE IN CHINA (Series B), 39 (3) , pp 225-234, 1996.

[5] Tiam Fang, Wang Chen, Bai Chunli, et al, Acta Phys.-Chim. Vol.14,No.4, pp 293-298, Apr. 1998.

[6] International Standard ISO 468-1982(E), International Org. for stand., Printed in Switzerland. pp1-3.1982.

[7] P.R.China National Standard , GB/T 1031-1995, Published in Chinese Standard Publishing House. Beijing, pp1-4,1996.

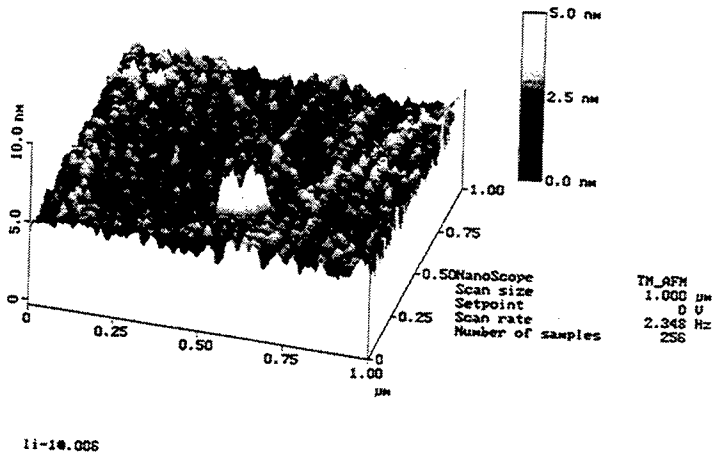


Fig . 1

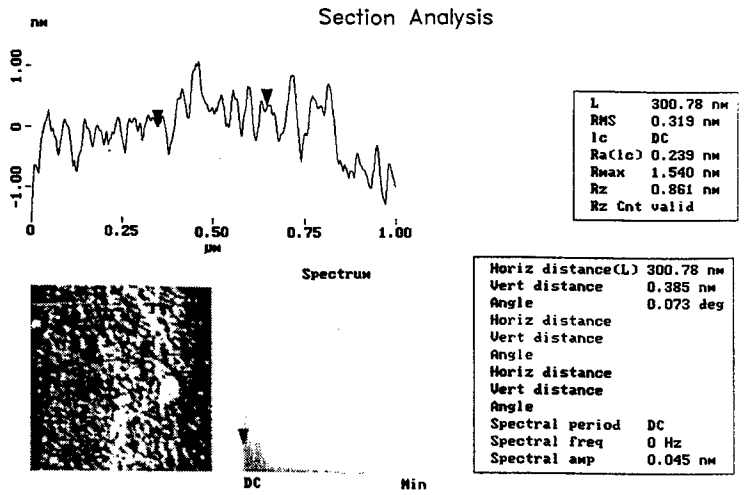

Section Analysis
nm
1.00
0
-1.00
0
0.25
0.50
0.75
1.00
μm
L 300.78 nm
RMS 0.319 nm
lc DC
Ra(lc) 0.239 nm
Rmax 1.540 nm
Rz 0.861 nm
Rz Cnt valid
Spectrum
Horiz distance(L) 300.78 nm
Vert distance 0.385 nm
Angle 0.073 deg
Horiz distance
Vert distance
Angle
Horiz distance
Vert distance
Angle
Spectral period DC
Spectral freq 0 Hz
Spectral amp 0.045 nm
DC
Min

Fig . 2

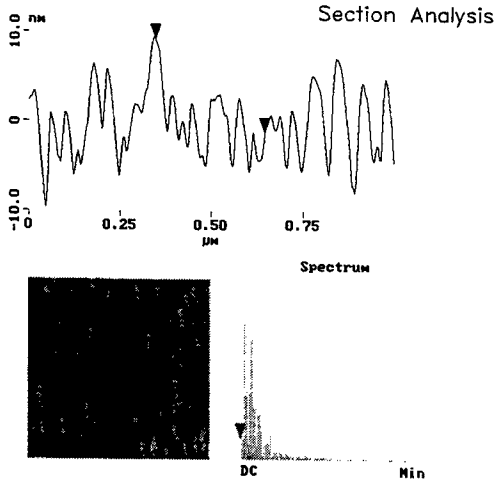

Section Analysis
10.0 nm
0
-10.0
0
0.25
0.50
0.75
μm
Spectrum
DC
Min

Fig . 3

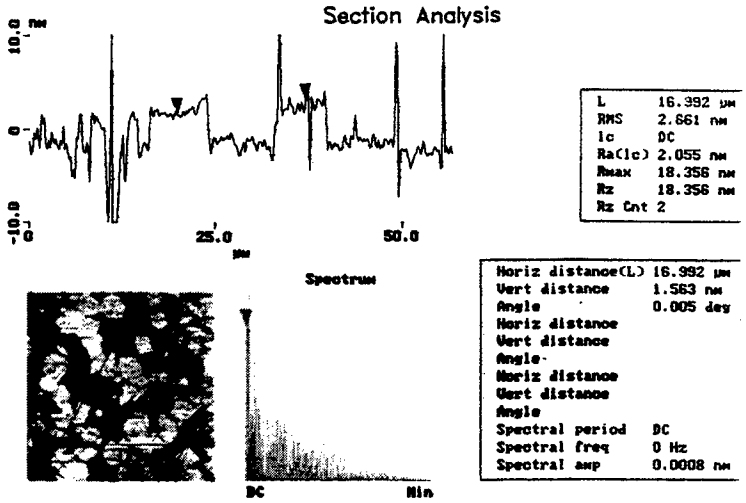

Section Analysis
10.0 nm
0
-10.0
0
25.0
50.0
μm
L 16.392 μm
RMS 2.661 nm
lc DC
Ra(lc) 2.055 nm
Rmax 18.358 nm
Rz 18.356 nm
Rz Cnt 2
Spectrum
Horiz distance(L) 16.392 μm
Vert distance 1.563 nm
Angle 0.005 deg
Horiz distance
Vert distance
Angle
Horiz distance
Vert distance
Angle
Spectral period DC
Spectral freq 0 Hz
Spectral amp 0.0008 nm
DC
Min

Fig . 4

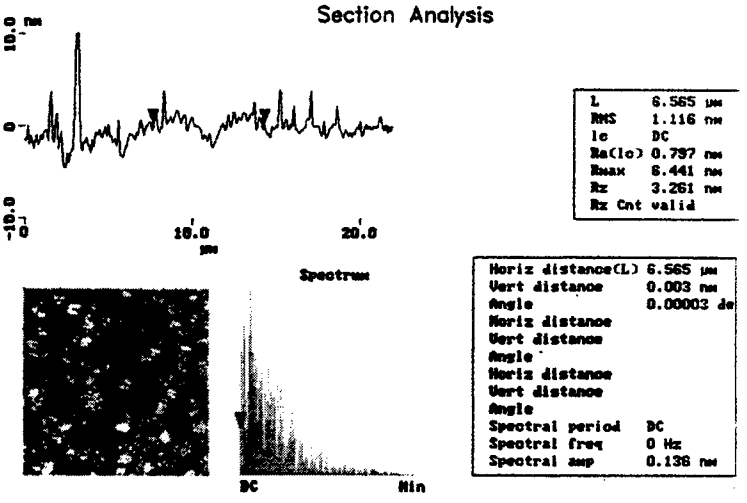

Section Analysis
10.0 nm
0
-10.0
0
10.0
20.0
μm
L 6.565 μm
RMS 1.116 nm
lc DC
Ra(lc) 0.797 nm
Rmax 6.441 nm
Rz 3.261 nm
Rz Cnt valid
Spectrum
Horiz distance(L) 6.565 μm
Vert distance 0.003 nm
Angle 0.00003 de
Horiz distance
Vert distance
Angle
Horiz distance
Vert distance
Angle
Spectral period DC
Spectral freq 0 Hz
Spectral amp 0.136 nm
DC
Min

Fig . 5

Thermal and mechanical properties of micromaterials using laseroptical strain sensors

Markus Anwander*, Alfred Hadrboletz*, Brigitte Weiss* and Bernhard Zagar**

*) Institute of Materials Physics, University of Vienna, Austria
Strudlhofg.4, A - 1090 Vienna,

**) Institute of Electrical Measurement and Signalprocessing
Technical University of Graz, A-8010 Graz ,Kopernikusg.24

ABSTRACT

There is an increasing necessity to record the deformation characteristics of microelements due to their increasing demand for engineering applications being used in the automotive industry, micromachines, special sensors etc. The data required are either thermal or mechanical such as Young's moduli, stress-strain values, creep-, fatigue- and fracture data .
In this investigation two non-contacting laseroptical strain sensors are being used to determine deformation data in combination with a special designed microtensile machine. The laseroptical strain sensors are applied to determine non-contacting strain values including a laser interferometric system with a high spatial resolution and a laseroptical speckle correlation method with high strain resolution.
The following materials data are presented and discussed:
Thermal expansion data of small metallic composites and $BaTi0_3$ specimens
Elastic-, plastic and fracture properties of metallic (rolled and electrodeposited Cu and Ni) foils with various thicknesses ranging from a few micrometers up to 200 μm. A pronounced thickness effect was detected and is discussed.
Fatigue and crack propagation properties of metallic foils with varying thicknesses (few micrometers up to several 100 μm) showing unexpected crack growth behavior as function of foil thickness which is explained with different stress states.
The applicability and limitations of the laser sensors are being discussed.

Keywords: laserinterferometer, laserspeckle correlation, non-contacting strain sensors, micromaterials, size effect, thin foils, thermal expansion values, mechanical data

1. INTRODUCTION

To increase the reliability of microelectronic parts it is frequently required to determine mechanical and thermal properties of the used materials in its practical dimensions (e.g. as thin foils and films, thin wires, membranes).This holds true for the development of special materials for conductor path, integrated circuits, special sensors, micromachines beeing used e.g. in the automotive and aircraft industry[1,2]. Due to thermal cycling of these microelectronic parts fracture may be induced. To reduce the fracture sensitivity an optimization and matching of various material properties such as thermal expansion coefficients, Young's moduli and other mechanical data are required. Such data can be found in[2].
The most important materials used in microdevices are listed in[2] and consist mostly of conventional metals, semiconductors, ceramics, composites etc. However, due to the known "size effect", the mechanical data cannot be taken from standard macrospecimens[3].
It is well known that mechanical testing of parts of microelectronic systems being small in all dimensions is a difficult task. Because they require a high degree of accuracy for the measurement of stress and strain and furthermore specimen preparation and handling cause additional experimental efforts.Several different mechanical testing techniques have been developed over the years[4]. The following testing procedures have been frequently used: The nanoindentation method, biaxial bulge testing, bending techniques and the uniaxial tensile testing of free standing foils. A critical evaluation of these methods is presented in[4]. Since the interpretation of the tensile test is simple, various investigators prefer this testing technique, although it is known that specimen loading, handling and alignement require considerable experimental experience. For the determination of strain non-contacting recording systems have to be applied, since strain measurements from cross-head movement are not accurate enough and conventional strain testing devices cannot be applied to specimens with small dimensions.

Part of the SPIE Conference on Advanced Photonic Sensors and Applications
Singapore • November–December 1999
SPIE Vol. 3897 • 0277-786X/99/$10.00

In this investigation two different non-contacting optical systems to monitor strain are presented. One is a laserinterferometer[5,6] with high spatial resolution and the other system resembles a non-contacting laser extensiometer with high strain resolution. The latter is based on the digital laserspeckle correlation technique[7,8]. In combination with a specially designed microtensile testing machine the following data were obtained:
Elastic-, plastic and fracture properties of metallic foils with varying thicknesses (electrodeposited and rolled Cu and Ni) have been used to study size effects. Furthermore using a specially designed testing technique fatigue crack growth properties have been determined to obtain data for the so called "fracture micromechanics" sometimes called fracture electronics (see Michel[2]). For a better understanding of the deformation damaging processes (e.g. interaction of microstructural features (dislocations, grain boundaries) with the advancing crack tip special SEM techniques in combination with electron channeling imaging contrast (ECCI) are being introduced[9].
Furthermore thermal expansion data of selected micromaterials have been acquired using the presented high resolution laserinterferometer. In the summary and conclusions the applicability and the limitations of the laser sensors are presented.

2. LASEROPTICAL STRAIN SENSORS

2.1 Laserinterferometer with nanometer resolution

The basic optical principle underlying the laserinterferometric system is the modified Young's double slit experiment, which was first used by Sharpe[10] as a strain sensor. In Fig. 1 the optical set-up is presented schematically. The specimen surface is marked by two Vickers microhardness indentations which are separated by a distance of approximately 100 μm essentially forming the baselength of this interferometer. These indentations are illuminated by a beam of laser light (He-Ne laser with λ = 632 nm) giving rise to two diffraction pattern interfering in the far field. These interference patterns are projected onto the CCD sensor by mirrors. The change in distance between the two indentations caused by thermal or mechanical loading is interferometrically measured by changes of the interference patterns, which were analyzed in the signal processing unit[5]. It is well known that almost in all loading systems rigid body motion and out-of-plane movements of the specimen occur. To minimize these effects the originally set-up presented by Sharpe[10] was rearranged (see Fig.1) and the lightpath between laser and CCD sensor was reversed. Therefore two shutters direct two beams of laser light alternately under equal angles of about 42° onto the indentations thus producing two interference fringe systems in a plane parallel to the specimen surface. This arrangement allows to minimize the effect of the out-of-plane as well as to average the effect of the in-plane movement. A detailed description of this system is given in[5].
Finally the relative displacement is presented by the following equation

$$\delta s = \frac{(\Delta m_{upper} - \Delta m_{lower})\lambda}{2 \sin \phi} \qquad (1)$$

with
δs ... change of distance between the indentations
Δm ... relative fringe motion (fraction of the fringe spacing)
λ ... wavelength of the laserlight employed
Φ ... angle between direction of illumination and of observation being 42°

The determined displacement resolution is about 2 nm , (this value is smaller than 1/100 of the wavelength used) due to the interpolation between adjacent fringes. With a base length of 100 μm the strain resolution is in the order of 2.10^{-5}. Using such a small base length it is possible to determine local strain values of small specimens with sizes of approximately 2mm x 2mm x 2mm .

2.2. Laserspeckle correlation method (LSC)

Laser speckle techniques can be used as a non-contacting optical method to determine strain which was originally introduced by[12] and later optimized by[7]. Laserspeckles form when coherent light is reflected off an optical rough surface. For practical purposes the surfaces of most materials are very rough on the scale of the wavelength of visible light, thus they yield visible speckles if illuminated. They are formed by the interference of dephased but coherent wavelets emanating from different microscopic elements.

Given constant illumination and observation geometry, a particular subjective speckle pattern which is characteristic for a certain (imaged) surface element moves in a well defined way when that surface element is displaced. This behavior permits to determine relative displacement of surface elements of material specimens before and after loading by simply

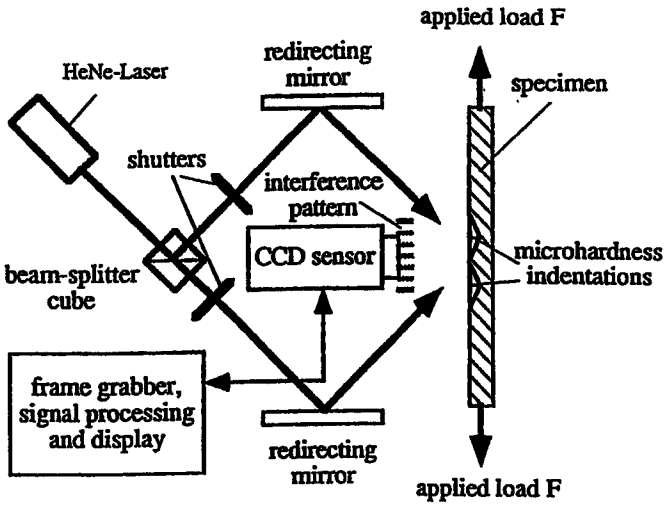


Figure 1: Principle optical set-up of the interferometric laser strain sensor

tracking the movement of the speckles associated with it. Usually this tracking is evaluated by different kinds of pattern matching algorithms. In the present study a digital laser speckle correlation technique is used. The resulting displacement vector is derived from the location of the maximum of the digitally calculated two-dimensional cross-correlation estimate (for details see[7,8]).

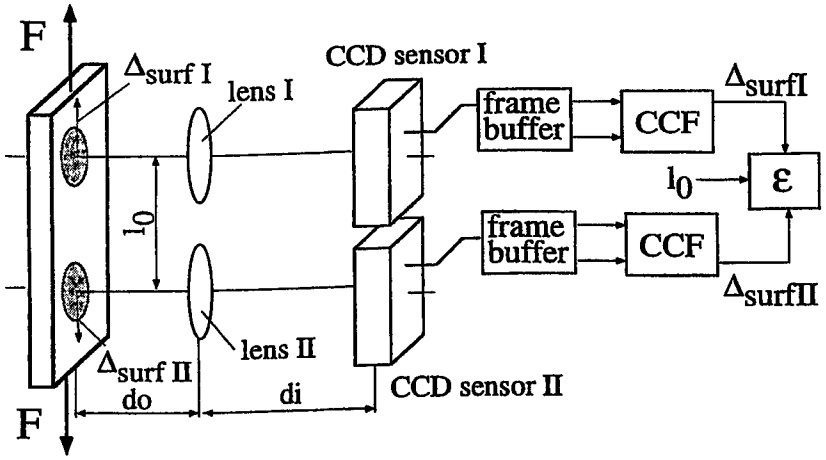


Figure 2: The optical arrangement of the laser speckle correlation system

The optical arrangement to determine strain within a specimen is presented in Fig. 2 schematically.
In order to determine strain ε of the specimen two more or less distant surface elements are tracked. By determining the surface element displacements Δ surf, taking their spatial difference Δl and dividing by a selectable baselength l_o any strain value can be determined as follows:

$$\varepsilon = \frac{\Delta l}{l_0} = \frac{(\Delta surfII - \Delta surfI)}{l_0} \qquad (2)$$

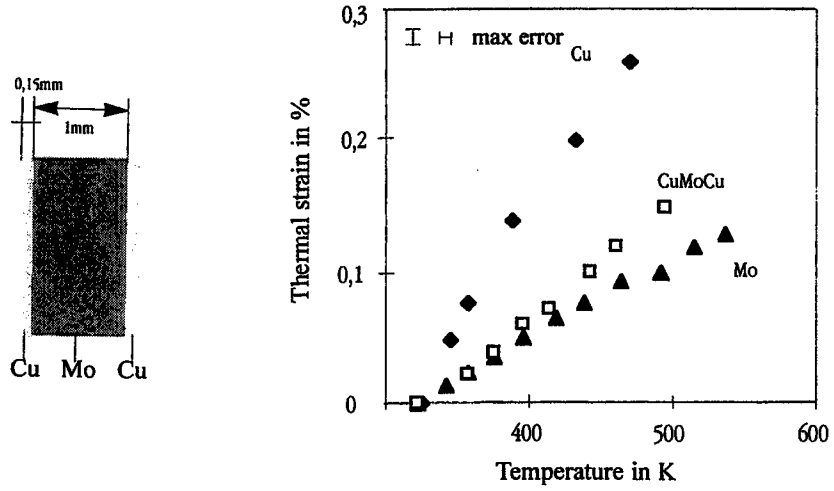
The technical arrangement of the LSC set-up consists of an illuminating system and two displacement recording systems.The illumination consists of two collimated laser diodes with max power output of 15 mW and a wavelength of

668 nm providing beams each of 3 mm in diameter. The displacement recording is performed by two lenses (separated by the base length l_0) with a selectable focal length and a standard CCD camera feeding their signals into a PC-based frame grabber, where the signal processing occurs.
For a selected baselength of 20 mm a strain resolution of 2.10^{-5} can be achieved. Since loading of specimens may cause plastic deformation, phase transformation etc. distorsions of the speckle patterns occur resulting in decorrelation effects. To overcome such effects a repetitive reinitialization of the image aquisition system is performed. Therefore Δl is taken as a sum of individual displacement increments. Details are described in[8]. The laser speckle correlation method in combination with a mechanical testing device is suitable to determine the strain of two-dimensional structures such as foils. No surface preparation is required.

3. APPLICATION OF LASEROPTICAL STRAIN SENSORS

3.1. Thermal expansion

It was demonstrated in a preceding investigation that the laserinterferometer is especially suitable to measure local coefficients of linear thermal expansion of microspecimens with arbitrary geometries[11].
In the present study the thermal expansion behavior of a three-layered structure (Cu-Mo-Cu) was investigated. The tested specimen consisted of a 1mm Mo sheet covered with a thin layer (0,15mm) of Cu on each surface.The dimensions are presented in Fig.3a. Firstly the thermal strain of the entire specimen was determined up to 500 K by introducing microindentations into the Cu layer. Then the copper layers were removed by etching and the thermal expansion was measured of the remaining Mo only. These data are given in Fig.3b. The Cu data were determined from a separate Cu specimen. From these experiments the linear thermal expansion coefficient of the composite was evaluated and yields a value of $\alpha = (9{,}8 \pm 0{,}9) \times 10^{-6}K^{-1}$. If α is calculated from the ratio of the thicknesses of the layers a value of $\alpha = (8{,}7 \pm 0{,}6) \times 10^{-6}K^{-1}$ is obtained.



a. specimen geometry b. thermal strain data
Figure 3: Thermal strain data of a three layered structure consisting of Cu-Mo-Cu

Due to calculations, using the Stoney equation[13,14] taking into account differences of both Young's modulus and the thermal expansion coefficients of the two materials, indicate that layer thicknesses > 8% influences α . This may explain the difference between experimental and predicted thermal expansion coefficient.
Another application of the laserinterferometer was to determine phase transformations associated with volume changes.The material investigated was a ferroelectric polycrystalline small piece of $BaTi0_3$. The determined α-values from several experiments are plotted versus temperature in Fig.4. An evaluation of these data shows a singularity for a temperature near 410K. This may be explained by a well documented phase transformation from a tetragonal phase to a cubic structure[15] Furthermore the data indicate also different values of α for the two different phases. The comparatively small scatter of the thermal expansion curves indicate the high reproducibility of the measurements obtained with the laser system.

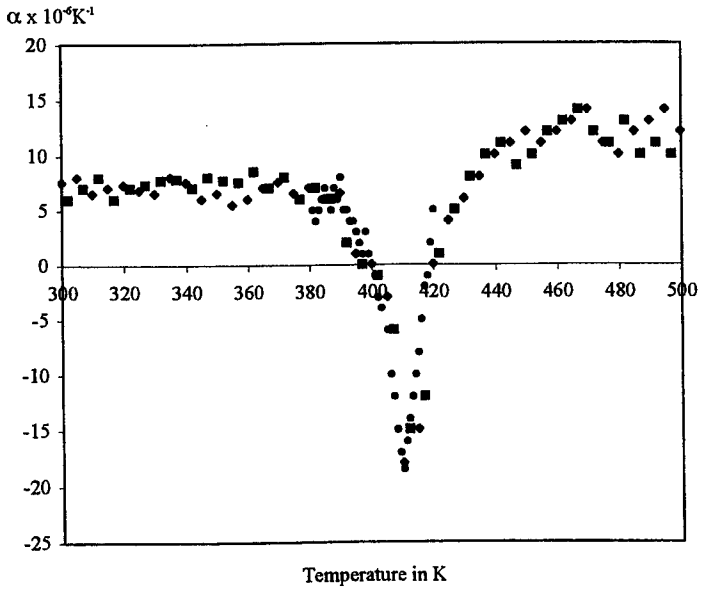


Figure 4: Linear thermal expansion coefficient of $BaTiO_3$

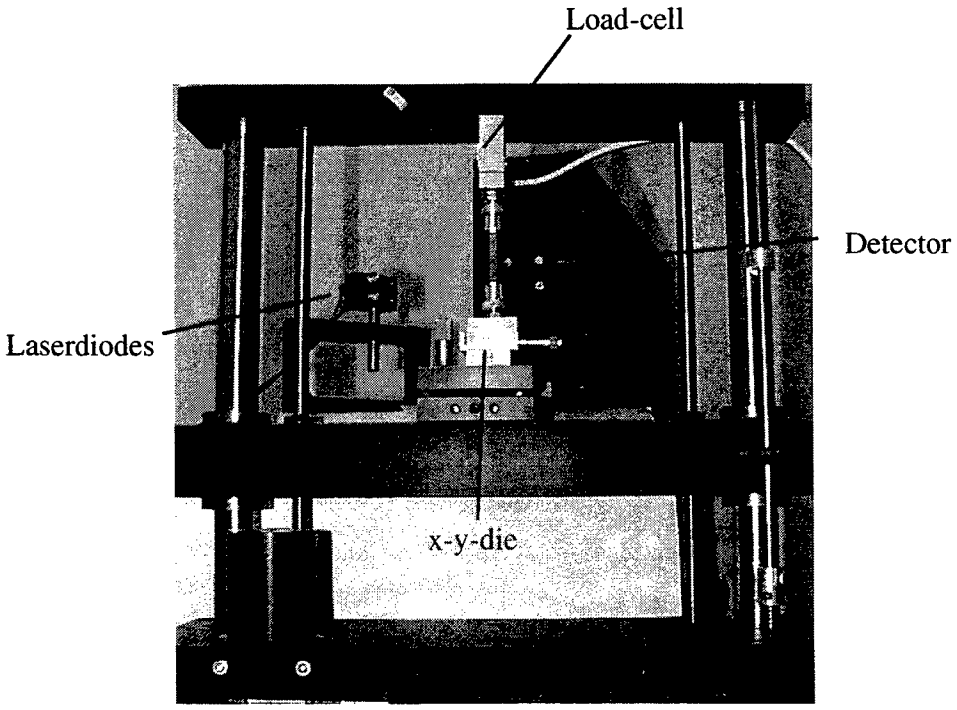


Figure 5: Microtensile testing machine and laser strain sensor

3.2. Elastic, plastic and fracture behavior

As already indicated in chapter 2.2 the LSC method is suitable to determine strain of thin elements in a non-contacting way. Thus to study the influence of the thickness of metallic thin foils on mechanical properties, the so called "size effect", the laser-speckle correlation technique was applied. As shown in Fig.5 the testing system consists of a specially designed microtensile testing machine (loading between 10 N – 200N) in combination with the LSC technique. Special emphasis was focused on the precise alignement of the foil using a x-y die and special designed grips. The very thin foils were glued to the

grips.To avoid the effect of detrimental vibrations the entire testing device was mounted on an optical table.The tensile testing was performed at room temperature with a deformation rate of 2%/min.The designed data acquisition system allows to obtain stress-strain curves in real time.

In this investigation rolled and electrodeposited Cu foils with a nominal purity of 99,95% are presented with thicknesses of 9, 19, 35, 78, 100, 125, 250 μm and 35, 105 μm, respectively. The rolled Cu foils were heat treated in vacuum at a temperature of 973 K for 2h resulting in grain sizes ranging between 15 – 35 μm determined from the intercept method (not taking into account twin boundaries). This means a single grain layer for foil thicknesses up to 35 μm, a double grain layer for a thickness of 78 μm and a multiple grain layer exceeding thicknesses of about 100 μm. The electrodeposited Cu foils were tested in the as received condition and exhibited extremely small grain sizes ranging from <1 μm to 5 μm (see insert in Fig. 8). The dimensions of the tested specimens were 10mm x 40mm, the preparation procedure being described in[16]. The rolled Cu specimens exhibit a pronounced cube texture with a twin texture depending on the manufacturing process.

For recrystallized rolled Cu foils the stress-strain curves are presented in Fig. 6a and the fracture strain as function of the foil thickness in Fig. 6b.

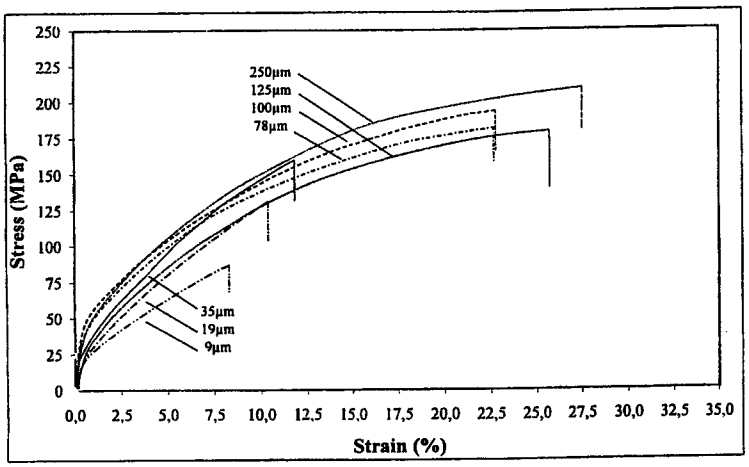


Figure 6a: Stress-strain curves of Cu-foils with varying thicknesses

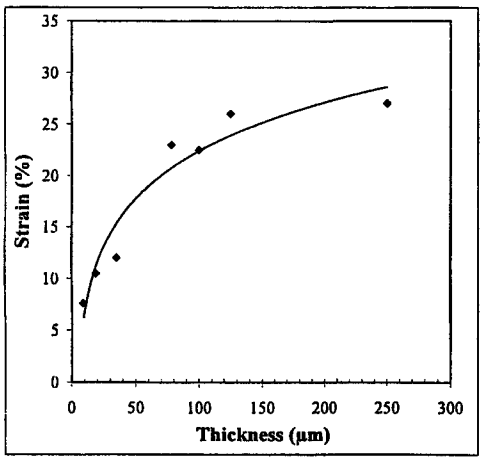
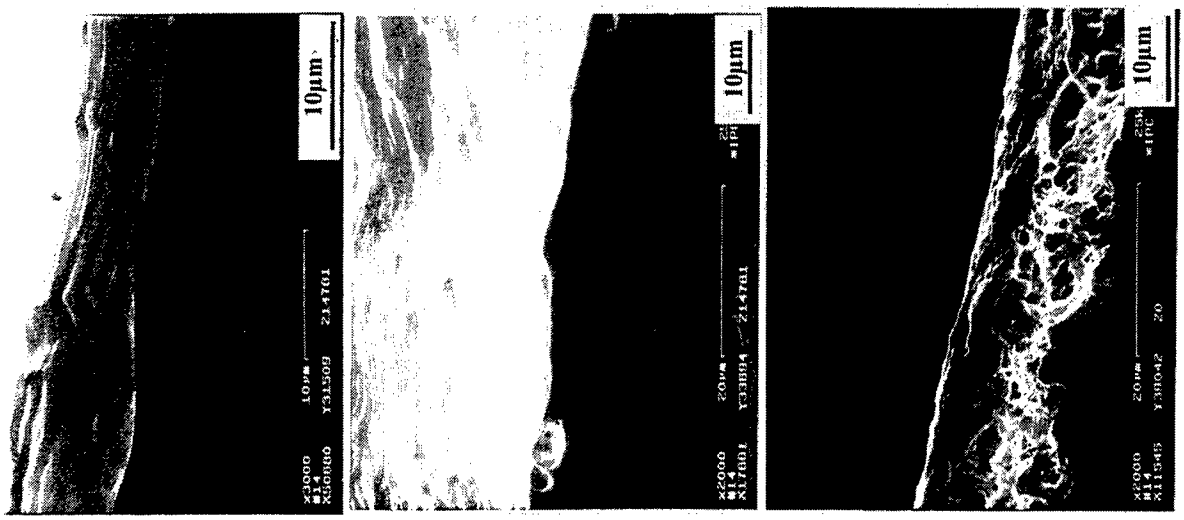


Figure 6b: Fracture strain as function of foil thickness



35 μm rolled and recrystallized

250 μm

35 μm electrodeposited (as received)

Figure 7: Fracture topography of ruptured Cu-foils with various thicknesses

From Fig. 6a a similar hardening behavior can be observed almost independent of foil thickness. In contrast the fracture strain is strongly dependent on the foil thickness as shown in Fig. 6b, decreasing with decreasing thickness. The fracture strain of a bulk specimen with a thickness of 6mm indicate a high value of fracture strain of > 50%.
This size effect may be attributed to the limited number of activated gliding systems in dependence of the ratio of grain size to foil thickness. To support these experimental findings the fracture topography has been investigated as shown in Fig. 7.
For the recrystallized foils two types of fracture topography after rupturing appear. For foils with thicknesses < 100 μm ductile failure with typical knife edge rupture without voids and dimples can be observed e.g. as shown for a 35 μm foil in Fig.7. With increasing foil thickness dimples and voids can be observed more frequently as shown for the 250 μm foil. These dimples are very small (about 5μm) (see Fig. 7).

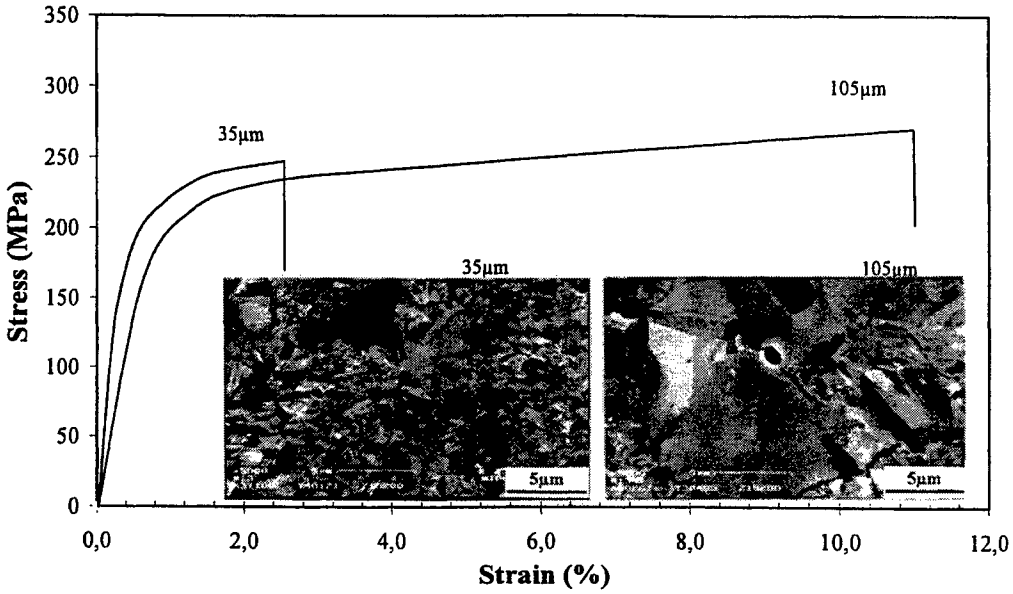


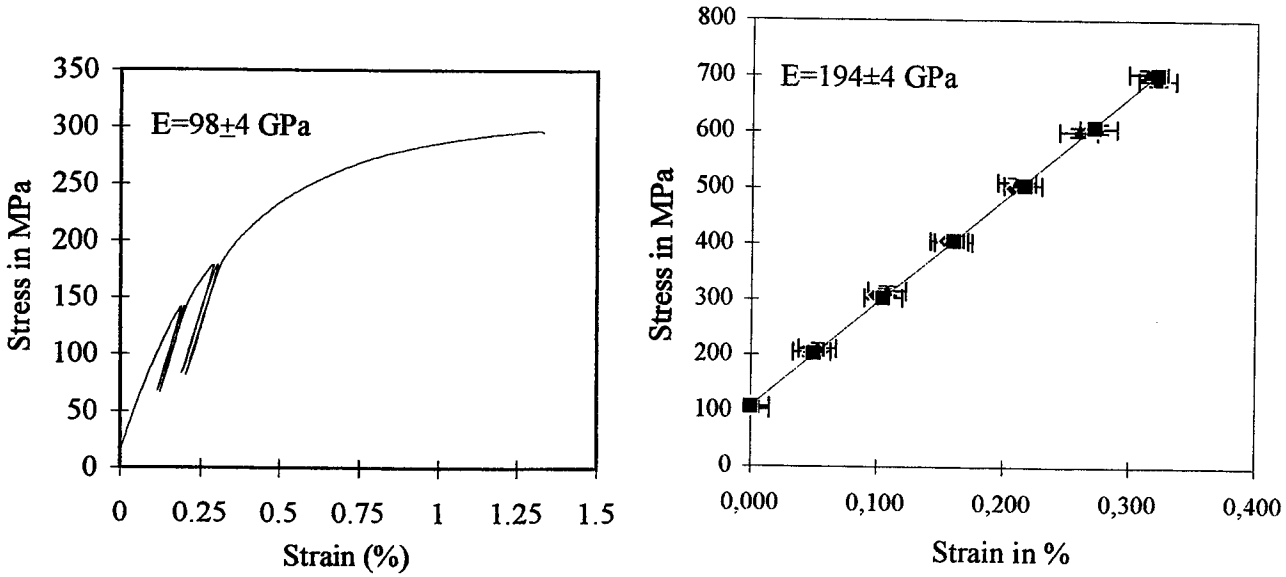
Figure 8: Stress-strain curves of electrodeposited Cu-foils with two thicknesses

Fig. 8 exhibits the stress- strain curves for electrodeposited Cu-foils indicating almost identical curves almost independent of foil thickness possibly due to the high dislocation density being introduced by the special manufacturing process. Variations in the fracture strain may be attributed to small differences in the distribution of the fine grain structure (see Fig. 8). The corresponding fracture surfaces show strong crystallographic appearance as shown in Fig. 7. A detailed explanation of the size effect will be presented in[17].
Because of the high strain resolution the laser speckle technique is especially suited for the determination of the Young's modulus of thin foils, however, great care has to be taken on the alignement and gripping systems.
The values are determined by a loading – unloading technique as shown for a 17 μm thick electrodepos. foil, see Fig. 9a, a detailed study of the influence of changes in the testing procedure on the value of Young's modulus is given in[17]. Almost parallel lines in the stress- strain diagram were obtained showing a high degree of strain resolution, resulting in Young's moduli values of 98 GPa with an error of measurement smaller than 5%.
The LSC system was also applied to determine the Young's modulus of a 3,5 μm thick electrodeposited Ni foil (see Fig. 9b). The data points resemble repeated tests, the evaluation supports good reproducibility and results in a value of 194 ± 4 GPa.

3.3. Fatigue properties

It is known that microdevices consist of layers with different thermal expansion coefficients and may be frequently exposed to thermal cycling and/or mechanical vibrations. These loading procedures may cause premature failure reducing the reliability. Therefore it is necessary to study the fatigue and fracture processes, (as already strongly emphasized by Michel[2], in more detail in terms of micromechanics.
In a first attempt the fatigue behavior of metallic foils with different thicknesses was studied (see also[18]). For this purpose a special test set-up has been designed, which allows to investigate the crack growth behavior of free standing thin foils.

Symmetrical push-pull loading of such thin foils was performed using constant strain amplitudes and a testing frequency of 20 kHz. Some results are presented in Fig. 10 showing crack growth curves of recrystallized Cu-foils (crack-length a versus number of loading cycles) demonstrating the influence of the thickness. For the purpose of comparison a crack growth curve of bulk material (6 mm thick strip) is added.



a stress-strain curve of electrodeposited Cu (t = 17 µm)

b stress-strain curve of an electrodeposited Ni-foil (t = 3,5 µm)

Figure 9: Determination of the Young's moduli from loading-unloading tests

For all foils up to 100 µm it was shown that the crack propagation rate decreases with increasing number of cycles. Compared to the behavior of bulk material including a foil with a thickness of 200 µm the curvature of the presented curves is opposite. In addition the crack growth curves are characterized by the occurrence of intermediate saturation values for thicknesses up to 100 µm. This anomalous behavior is investigated in more details for a 100 µm thick foil as an example (see also[18]).

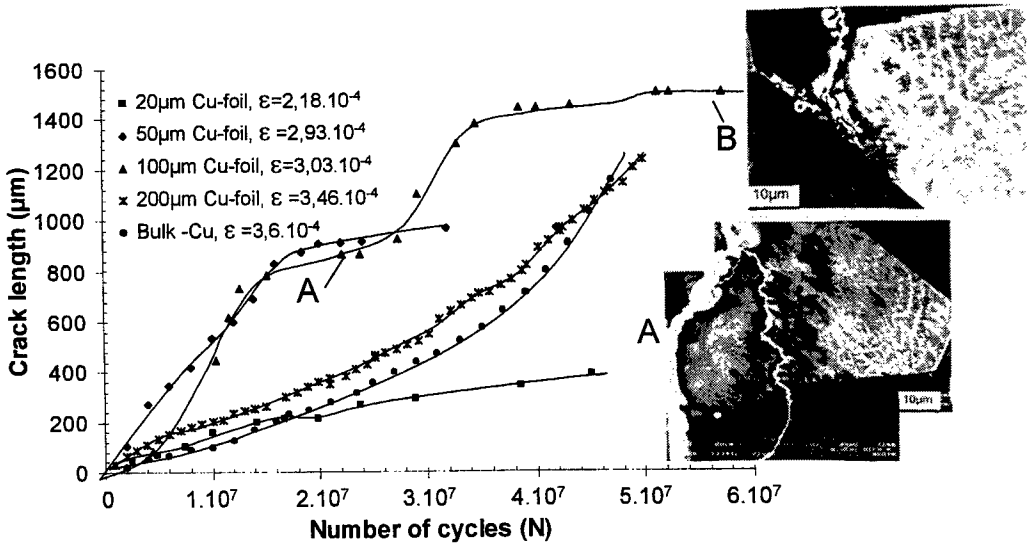


Figure 10 Crack growth curves of Cu-foils with different thicknesses

As presented in the inserts of Fig. 10 from ECCI-micrographs[9,18] the intermediate crack arrest may be attributed to a strong interaction of the growing crack with grain-boundaries (A) and the onset of crack bifurcation (B). This behavior may be explained by the assumption of a plane stress state for foils thinner than 100 μm and a state of plane strain for foils with thicknesses exceeding 200 μm. For bulk material the entire crack propagation is governed by plane strain.
From the findings presented above it is obvious that crack growth in foils is characterized by the transition from plane stress to plane strain with increasing thickness.
It is emphasized that the ECCI technique allows to reveal the global dislocation structure in the Cu-foils after cyclic loading, from which the kind of local plastic deformation and the interaction of the growing crack with the prevalent microstructure may be deduced. This knowledge should result in an improved designing of micro systems.

4. SUMMARY AND CONCLUSIONS

i. The presented laserinterferometer is characterized by the relative high spatial resolution of about 100 μm. Thus this system is applicable to determine thermal expansion data from small pieces of arbitrary geometry of various materials. If the laserinterferometer is applied for ceramics and polymers due to their reduced reflectivity and deformability higher laser power or the application of thin reflective layers is required. As shown in the experimental part the thickness of such layers should not exceed 8% of the matrix material in order to reduce detrimental influences on the experimental result.
To avoid the simple but locally damaging indentation technique it is suggested to apply evaporated structures f.i two gold lines[19] onto the investigated material in order to obtain interference pattern.

ii. The laser based non-contacting strain sensor is due to the simple optical arrangement easily adaptable to various testing machines and characterized by high resolution in the order of 2.10^{-5}. This system allows to determine the thermal and mechanical response of thin foils and films of most materials, without any special surface preparation. Elastic properties such as Young's modulus – necessary for the characterization of thin foils- have been successfully measured with high accuracy using a special microtensile machine in combination with the LSC- based sensor. Since the base length is at present relatively large, structures with minimum size of about 15 mm al least in one dimension are required. Investigations can be performed over a wide range of testing temperatures corresponding to real test conditions of the microelectronic devices.

iii. Both laser sensors can be applied for the determination of the linear thermal expansion coefficients and for the detection of volume dependent phase transitions in various materials including piezoelectric and magnetorestrictive components. The error of measurement of the linear thermal expansion coefficient is in the order of $1.10^{-6}K^{-1}$. These methods may be extended to the investigation of multilayered structures.

iv. A strong size effect could be detected especially concerning the fracture strain of thin metallic foils.This was mainly attributed to the ratio of grain size to foil thickness for values in the order of unity. Similar reasons are responsible for the observed specific fracture behavior. The data obtained from the tension test allow to design microcomponents subjected to complicated stress- strain states. Using a newly designed test set-up fatigue data of free-standing thin metallic foils under symmetrical loading conditions have been obtained. Also a strong size effect was detected, which can be explained with the corresponding stress state. In addition the ECCI technique is feasible to record the interaction of the microstructure with the crack resulting in a better understanding in order to optimize the properties of the used materials. The fatigue data obtained allow to establish microfracture criteria which are necessary to design reliable microelectronic components. This method can also be applied for the study of the fatigue behavior of multilayered structures.

ACKNOWLEDGEMENT

The authors thank the Austrian Bundesministerium für Wissenschaft und Verkehr (COST 510) and Fonds zur Förderung der wissenschaftlichen Forschung, Wien for financial support.

REFERENCES

1. Micromat 95, Conf. Proc. Micromat 95, B. Michel and T. Winkler, Deutscher Verband für Materialforschung und -prüfung, Berlin 1995.
2. Micromat 97, Conf. Proc. Micromat 97, B. Michel and T. Winkler, Deutscher Verband für Materialforschung und -prüfung, Berlin 1997.
3. E. Arzt, „Size effects in materials due to microstructural and dimensional constraints: A Comparative Review",

Acta Mater. **46**, pp. 5611-5626, 1998.
4. F. Brotzen, „Mechanical testing of thin films, *Int. Materials Reviews*, **39**, pp. 24-45, 1994.
5. B. Zagar, „A laser-interferometer measuring displacement with nanometer resolution", *IEEE Trans. Instrumentation and Measurement*, **43**, pp. 332-340, 1994.
6. B. Zagar, H. Weiss, M. Anwander, B. Weiss, and R. Stickler, „A laser interferometric system for local strain measurements with nanometer resolution", Int. Symposium *Local Strain and Temperature Measurements in Non-uniform Fields at Elevated Temperatures*, J. Ziebs et al., Woodhead Publ. Lim., Cambridge U.K., 1996.
7. H. Manser, „Laser-Speckle-Korrelationsverfahren zur Verschiebungsmessung bei hohen Probentemperaturen, Ph.D. Thesis, Technical University Graz, Austria, 1992.
8. M. Anwander, B. Zagar, B. Weiss, and H. Weiss, „Non-contacting strain measurements at high temperatures by digital-laser-speckle technique", *Experimental Mechanics*, in print.
9. D. Chen, D. Melisova, B. Weiss, and R. Stickler, „The electron channeling contrast technique applied to the characterization of dislocation structures in the vicinity of a fatigue crack", *Fat.and Fract. of Eng. Mat. and Struct.*, **20**, pp. 1551-1561, 1997.
10. W.N. Sharpe, „Applications of the interferometric strain/displacement gauge", *Optical Engineering*, **21**, pp. 483-488, 1982.
11. H. Weiß, B. Zagar, R. Stickler, B. Weiss, and M. Anwander, „Laser techniques for non-contacting strain measurements at high temperatures", *Final Report COST 510*, H. Ehrlich, and T. Valente, European Comunities, Brüssel, pp. 165-178, 1998.
12. I. Yamaguchi, „A laser-speckle strain gauge", *J. Phys. E: Sci. Instrum.*, **14**, pp.1270-1273, 1981.
13. S. Brezina, „Anwendung eines Laserinterferometers zur berührungslosen Messung der lokalen Wärmeausdehnung von verschiedenen Materialien", *Master Thesis*, University of Vienna, Vienna, Austria, 1996.
14. Y.L. Shen and S. Suresh, „Elastoplastic deformation of multilayered materials during thermal cycling", *J. Mater. Res.*, **10**, pp.1200-1215, 1995.
15. T. Mitsui, I. Tatsuzaki, E. Nakamura, „An introduction to the physics of ferroelectrics ...", I. Lefkowitz and G.W. Taylor, Gordon and Breach, New York, 1976.
16. C. Dinhobl, M. Klein, M. Anwander, B. Weiss, and H.G. Brokmeier, „Die Bestimmung der elastischen und plastischen Eigenschaften dünner Metallfolien mit laseroptischen Methoden", *Band 10, Symp. 13: Werkstoffprüfung*, W.J. Muster et al., Wiley-VCH, pp. 247-255, 1999.
17. M. Klein, A. Hadrboletz, G. Khatibi, and B. Weiss, „The size-effect on mechanical properties of thin metallic foils", to be published.
18. A. Hadrboletz, G. Khatibi, and B. Weiss, „The „size-effect" on the fatigue and fracture properties of thin metallic foils", *Proc. of EUROMAT*, 1999, in print.

Simultaneous Strain and Displacement Measurement in Micromechanics

Bing Zhao, Anand Asundi

School of Mechanical & Production Engineering
Nanyang Technological University
Nanyang Avenue, Singapore 639798

ABSTRACT

In this paper, two quantitative methods to measure micro-deformation using grating microscopy are proposed, a grating diffraction method and a full-field grating phase shift method. A compact optical transmitting microscope with white light source is reconstructed by developing a loading and recording system. For direct strain measurement, a grating diffraction method is presented. With the help of a Bertrand lens, the Fourier spectrum of the grating is formed on the CCD sensor plane with high image quality. Software for precise, fast and automatical determination of the diffraction spot centroids is developed. Local strains are thus measured with high resolution. For the deformation measurement, a phase-shifting grating microscope method with high sensitivity and spatial resolution is proposed. Phase shifting is based on the slab refraction effect and is realized via a tilting compensator. The system possesses a high spatial resolution (1 μm), and a displacement precision of 0.1 to 0.03 μm. The combination of the grating diffraction method and the phase shifting method in the same test provides simultaneous measurement of strain and displacement, thus demonstrating that the grating techniques are viable in the microscope environment.

Keywords: Grating diffraction, phase-shifting, optical microscope, Bertrand lens, tilting compensator, displacement and strain.

1. INTRODUCTION

Optical microscope is an extremely useful tool in microstructural observation. By combining microscopy with optical deformation measurement techniques such as the speckle and grating methods, a potential system to measure micro-deformation of materials is possible. In the present paper, two quantitative methods for deformation measurement using grating microscope are proposed.

The first method involves realizing a grating diffraction technique[1~5] in a traditional transmitting optical microscope for direct strain measurement. Compared to other optical techniques such as moiré and holography, grating diffraction technique avoids the difficulty in fringe pattern interpretation associated with most optical techniques. This technique also possesses several advantages[4]: (i) no master grid is needed; (ii) the sign of deformation is known; (iii) system integration is simple. For micromechanics and microengineering studies, it is necessary to develop measurement techniques under microscopes. In this paper, we synthesize optical microscopy and grating diffraction method, to develop a compact strain microscope. It possess several important features: real-time loading and measurement, variable sensitivity, field of view (examined area) can be easily changed, as the specimen can rotate and move in X and Y direction, the set-up is compact, as optical components are easily adjustable; image quality is very good, the diffraction pattern approximates to a binary image.

The second method deals with grating phase-shifting technique used in the same microscope as a tool to measure deformations with high displacement sensitivity and microscopic spatial resolution. Most of optical technique requires interpretation of fringe pattern to deduce the measurand. Phase-shifting [6~9] is a powerful fringe analysis tool. In this paper, a phase shifting grating microscope technique is proposed to measure deformations with high displacement sensitivity and microscopic spatial resolution. The proposed technique images a grating printed or bonded on the object under the microscope at magnifications up to 100× (objective magnification). Analysis of the deformed grating of the object under load is accomplished using the powerful phase shifting approach method. This requires recording of three or more shifted images of the grating in steps of fractions of the pitch. The traditional way for grating or moiré phase shifting measurement is realized using a precise mechanical translation device to shift the grating. For gratings of 10 μm pitch, the gratings have to be shifted in steps of 2.5 μm for a phase shift of 90°. Space restriction prevents the use of the conventional phase shifter under a compact microscope. Thus there is no support for phase shifting techniques using a conventional compact microscope. A tilting compensator, an inherent part of a polarizing microscope, is used as a phase shifter, enabling the grating image to be easily and precisely shifted optically. It should be emphasized that for this grating method used in a conventional microscope, the upper limit of spatial resolution for displacement measurement approaches the diffraction-

Part of the SPIE Conference on Advanced Photonic Sensors and Applications
Singapore • November–December 1999
SPIE Vol. 3897 • 0277-786X/99/$10.00

limited resolution. This value is much higher than that of micro-moiré interferometry regardless of microscope and carrier method used.

Further advantages of a conventional microscope over a long distance microscope or video microscope are:

The resolution power of conventional microscope is generally higher than that of normal commercial long distance and video microscopes. Thus, high density grating can be used, and both the spatial resolution and sensitivity of measurement can be improved.

In current work, the conventional microscope is used both for image magnification and as a diffractometer. Long distance microscope and video microscope can only be used for image magnification.

The combination of the grating diffraction method and the phase shifting method demonstrate originality, and the system provides benefits of both methods. With grating diffraction method, strain is directly measured, as there is no influence of rigid body displacement. With grating phase shifting method, full field displacement is measured with high spatial resolution. Both methods can be used for the same specimen. The transformation from diffraction pattern to grating image can be easily realized by adjusting the Bertrand lens. The current work demonstrates that the grating techniques are viable in the microscope environment.

2. EXPERIMENTAL SYSTEM

The optical system consists of a Leitz-Orthoplan II Pol polarizing microscope (Fig. 1a), a specimen grating and an image acquisition as well as processing system. The transmitting microscope includes a white light source, an interference filter, a field and aperture diaphragm, an objective, a tilting compensator, and a Bertrand lens as shown schematically in Fig. 1b, and a pair of eyepieces shown in Fig. 1a. A photograph of the tilting compensator is shown in Fig. 2. The specimen and the grating are placed between the condenser and the objective of the microscope. The density of the grating used ranges from 40-200 l/mm. Three types of objectives 6.3/0.2, 10/0.25, 25/0.50 (Magnification / Numerical Aperture) are typically used. A $10\times$ eyepiece can be used for visual observation.

The image is recorded using a CCD camera coupled to 512× 512× 8 bit frame grabber. Software for image capture and processing is developed. The CCD is directly connected to the microscope eyepiece tube after removing the eyepiece.

The loading system consists of a loading frame and a two-axis translation stage. Fig. 3 shows the photo of the loading frame. The frame is fixed onto the translation stage, and as a whole fixed on the mechanical stage of the microscope. The specimen to be tested can be translated in two directions and can rotate with the loading frame. So, we can easily examine the different area of the specimen. The specimen must be transparent or demi-transparent, and is fixed onto the loading frame. Rotation of the gyro-stat generates tensile loading to the specimen.

3. GRATING DIFFRACTION METHOD

3.1 Microscope Image

The principle of strain measurement described here is based on the grating diffraction properties. Typically, this is realized using a laser beam to illuminate a diffraction grating giving rise to a line of diffraction dots, as illustrated in Fig. 4. Using an optical microscope where the image of the object is examined, its Fourier spectrum is not normally obtained. According to the microscopic image-forming principle (Abbe's theory[10]) the transition from the object to the image involves two stages. First the grating-like object when illuminated gives rise to a Fraunhofer diffraction pattern in the back focal plane of the objective, where every point can be considered to be a center of a coherent secondary disturbance. The light waves that proceed from these secondary sources will then interfere with each other and give rise to the image of the object in the image plane of the objective. Thus the diffraction pattern is formed in the back plane and image of the grating is formed at the image plane. Hence, in order to get a video image of the grating diffraction pattern through a microscope, the back focal plane of the objective should be extended to the CCD sensor plane. Otherwise the grating image is obtained. This can be realized using a relay lens or a Bertrand lens. The grating diffraction in this way is Fraunhofer diffraction and is illustrated in Fig. 4.

The important step in the measurement is to get the grating diffraction pattern. In our test, this was done by using the Bertrand lens. This lens is generally placed between eyepiece and objective and is often used to observe the interference or polarizing pattern. With this lens, one can see clearly the image formed in the back focal plane of the objective. On gradually adjusting the focusing screw of this lens which controls the up-and-down movement of the lens, the transformation from the grating image to the grating diffraction pattern is observed through the eyepiece or the video monitor.

3.2 Strain Measurement

The grating diffraction method uses only one specimen grating, no master grating is needed. The grating is bonded on the surface of the specimen, and deforms with the specimen as the specimen is loaded. Consider a thin monochromatic collimated beam that is normal to a linear grating plane, illuminating a point on a specimen grating, the diffraction is governed by the equation

$$P\sin\theta_n = n\lambda \qquad n = 0, \pm1, \pm2,... \tag{1}$$

where P is the pitch of grating, n is the diffraction order, λ is the light wavelength and θ is the diffraction angle (see Fig. 4). After the grating is deformed, the pitch changes to P^*, resulting in variation in θ according to equation (1). Thus, the strain along the X direction (see Fig. 4) at the illuminated point on the grating can be calculated by

$$\varepsilon_{xx} = (P^* - P)/P \approx (X_n^* - X_n)\cos\theta_n / X_n \tag{2}$$

Where X_n and X_n^* are the undeformed and deformed diffraction spot centroid of the nth diffraction order. In our experiment, the density of the grating is less than 200 lines/mm. If only the first diffraction order is considered, ie., $n = 1$, the angle (θ_1) of the diffracted beam is small. For example, for green light, $\lambda = 500$ nm, grating pitch $P = 1/200$ mm, $\theta_1 = 5.73°$, and ($\cos\theta_1$) $= 0.995 \approx 1$. So Eqn. (2) can be simplified as:

$$\varepsilon_{xx} \approx (X_1^* - X_1)/X_1 \tag{3}$$

If a cross grating is used instead of a linear grating, all the three components of in-plane strain, ε_{xx}, ε_{yy}, ε_{xy} can be similarly obtained.

3.3 Image Processing and Displacement Sensitivity

There is an advantage in using a microscope to realize the grating diffraction test, i.e., the recorded image possess a good quality. This is mainly because:

- no ground-glass screen is required, the diffraction spots are directly projected onto the CCD sensor plane;
- the microscope with high quality components provides us with a sharp precise optical image.

The obtained diffraction image is approximated to a binary image. This makes it easier to perform image processing. The precision of strain measurement depends on the determination of centroid of the diffraction spots. Several algorithms for calculating the spot centroid with subpixel precision have been developed and are reviewed by West et Clarke[11]. After comparing these algorithms, all were found equally suitable in the case. So the simplest algorithm[2, 4] is chosen:

$$X = \sum_{i=1}^{N_i}\sum_{j=1}^{N_j} g(i,j)i \Big/ \sum_{i=1}^{N_i}\sum_{j=1}^{N_j} g(i,j) \tag{4}$$

Here, X is the centroid coordinate to be determined, g is the recorded grey value at point (i, j). When the intensity $g(i, j)$ is less than a certain threshold intensity value, $g(i, j)$ is put to zero. N_l and N_c are the sizes of rectangular window which surrounds the diffraction spot. A software that can automatically detect the diffraction spot centroids and then calculate the strain is developed.

Following equation (3), the strain sensitivity can be expressed as

$$\sigma_\varepsilon \approx \sqrt{2}\sigma_x / X_1 \tag{5}$$

Where σ_X is the sensitivity of spot centroid determination for recorded spot pattern with equation (4). One way of increasing the strain measurement sensitivity is to decrease the value of σ_X. This approach is limited by the spatial resolution of CCD camera and the algorithm for determining the diffraction spot centroid. Another way is to increase X_1, i.e., increasing the distance between two spots of ±1 orders of the recorded pattern. This is generally realized by using a grating with a higher density, or by increasing the distance between grating and the image-forming screen.

For strain microscope measurement, several methods can be used to increase the distance between two diffraction spots of ±1 orders. One way is to increase the distance between the grating and CCD sensor by increasing the length of the tube connecting the eyepiece. Other ways, such as using the objective with lower power, using high-density grating can also increase the strain measurement sensitivity. There is a limitation to using high density grating. A higher density grating will

result in a larger diffraction angle. If this angle is too large, ±1 diffraction beam can not enter the objective. From our experience, grating frequency of 40 ~ 200 lines/mm is a good compromise.

It was reported that the precision of the spot centroid determination would be 0.1 pixel[2,12] with algorithm described by Eqn. (4). For our test, the sensitivity of spot centroid determination for recorded spot pattern with equation (4) is estimated as σ_X = 0.08 pixel. This was determined as follows. First, a series of diffraction patterns in same test state where each pattern is actually an average result of 5 images are recorded. Second, the diffraction spot centroids are calculated using equation (4). Then, the standard deviation of the series of centroid values are determined. This standard deviation value is considered as the spot centroid determination sensitivity. Maximum X_I can not beyond 512/2 = 256 pixels, thus the strain sensitivity is 442 με.

3.4 Test Example

In order to verify the experimental procedure and check the measurement precision, we carried out a test using the microscope translation stage to stretch a grating film specimen, which is bonded onto the stage (see Fig. 3). This specimen including a plastic substrate film and pure grating film is thin, so it is easily loaded by using the micrometer drive, and we can control the displacement with a precision of 5 μm. This precision is not high, but the travel distance of the stage is long. We can calculate the average strain by using the measured overall displacement. The comparison between results of this average strain and the strain measured by diffraction method shows the two groups of results are in general agreement. An example is demonstrated in Fig. 5, where two diffraction patterns of an 80 l/mm cross grating before and after loading are shown. The specimen has an overall displacement of 0.6 mm, and the calculated average strain is 12000 με. The white line in this figure is for reference. After the deformation, the lower three diffraction spots show a noticeable shift.

4. PHASE SHIFTING GRATING METHOD

4.1. Slab Tilting for Phase Shifting

An optical beam transversing an inclined optical slab, will be shifted horizontally due to refraction, as illustrated in Fig. 6. If the slab, with thickness d, is inclined at an angle of α to the horizontal, beam AB and CD represent the incident and refracted beams respectively, The length of CE is the horizontal shift due to refraction and β is the refraction angle.
The refractive index of the slab is

$$n = \sin\alpha / \sin\beta \tag{6}$$

The length of CE is thus

$$|CE| = |BC|\sin(\alpha - \beta) = d\sin(\alpha - \beta)/\cos\beta \tag{7}$$

Noting that α and β are less than 90°, a nonlinear relationship between the shifted length CE and α is obtained:

$$|CE| = d\left[\sin(\alpha) - 0.5\sin(2\alpha)\Big/\sqrt{n^2 - (\sin\alpha)^2}\right] \tag{8}$$

The optical slab in Fig. 7 can rotate around a horizontal axis that is parallel to the horizontally placed grating lines. As shown in Fig. 6, AB represents the illuminating beam and CD is the refracted beam. Following the above analysis, the image of the grating will be shifted a distance CE from Eq (7) which is equal to a phase shift of the grating fringe signal. The equivalent phase-shift is:

$$\delta = 2\pi|CE|/P^* = 2\pi dK(\alpha)/P^* = const \times K(\alpha) \tag{9}$$

Where P^* is the pitch after magnification by the objective, and K is a function of rotation angle α and refractive index n, and is expressed by

$$K(\alpha) = \sin\alpha - 0.5\sin(2\alpha)\Big/\sqrt{n^2 - (\sin\alpha)^2} \tag{10}$$

The phase shifter, i.e., the optical slab, is a tilting compensator (see Fig. 2) in the microscope. This type of compensator is generally used for birefringence measurement in a polarizing microscope. It consists of a slab of a uniaxial crystalline medium. It is mounted so that the slab can be tilted via a graduated wheel. The crystal plate in this compensator can rotate with a precision of 0.05°, over a range of ±32°.

It is difficult to know the exact value of P^* as the magnification of Bertrand lens is not known. Thus, it is difficult to determine the phase shift amount δ. Among numerous phase-shifting algorithms, Carré[8] and Hariharan[9] algorithms do not require the value of the first δ. In this paper, the latter is chosen. 5 phase shifted fringe patterns are required:

$$I_i = I_a + I_b \cos(\varphi + \delta_i), \quad i = 1, 2, ..., 5 \tag{11}$$

Where I_i represents the ith fringe pattern, I_a is background intensity, I_b is modulated intensity, φ is the phase to be measured. These four parameters are functions of designated point coordinate (X, Y). δ_i is the phase shift of the ith pattern which can be expressed as

$$\delta_i = (i-3)\delta_0, \quad i = 1, 2, ..., 5 \tag{12}$$

The phase can be derived as

$$\tan\varphi = 2(\sin\delta_0)(I_2 - I_4)/(2I_3 - I_5 - I_1) \tag{13}$$

Where δ_0 satisfies

$$\cos\delta_0 = 0.5(I_5 - I_1)/(I_4 - I_2) \tag{14}$$

It is clear that the measured phase can determined without knowing the phase shift δ_0, provided δ_i is proportional to δ_0 as in Eq. (12). For example, a group of α_i, the inclined angle of the refraction slab (see Fig. 6), and its corresponding phase shift δ_i is tabulated:

i	1	2	3	4	5
α_i(°)	-9.93	-5	0	5	9.93
δ_{ti}(°)	$-2\delta_0$	$-\delta_0$	0	δ_0	$2\delta_0$

The measurement procedure is as follows:

- Adjust the microscope system, record 5 phase-shifted grating patterns, then calculate initial phase field $\varphi_s(X, Y)$;
- Load the specimen, record the grating patterns and obtain the final phase field $\varphi_m(X, Y)$;
- Unwrap the phase field to be measured $\varphi = (\varphi_m - \varphi_s)$, and calculate the displacement as

$$u = P \times \varphi/(2\pi) \tag{15}$$

Here, P is the real grating pitch, in Eq. (9), P^* indicating the pitch of grating magnified by the objective.

4.2 System Specifications

The spatial resolution of displacement is restricted by optical system and CCD recording system. In order to detect fine grating lines with the CCD, the grating pitch must greater than 2 pixels, thus the CCD's resolution is defined as 2 pixels. For the CCD used, the sensor pixel size is about 10 μm. As the image formed on the CCD sensor plane is magnified by an objective, the system resolution is 20 μm divided by the objective magnification factor. Typical resolution values for three typical objectives:

Objective magnification	25	40	100
Resolution (μm)	0.80	0.50	0.20

Furthermore, the limitation of resolving power of objective must be considered. The following formula gives the distance between two object points which a microscope (or objective) can just resolve[10], termed the resolution of the microscope.

$$Res = (0.61 \sim 0.82)\lambda / NA \tag{16}$$

Where λ is the wavelength in vacuum, NA the numerical aperture of the objective used. The multiplying factor 0.61 is for incoherent light, and 0.81 for coherent light. In present case, the light used is partially coherent. So, the actual factor is between 0.61 and 0.82. For white light, λ as 0.55 μm, and the resolution for the three objectives:

Objective magnification	25	40	100
NA	0.5	0.65	1.3
Resolution (μm)	0.67~0.90	0.52~0.69	0.26~0.35

The lower-limit values are chosen as the spatial resolution for displacement measurement, that is 0.90 μm, 0.69 μm, and 0.35 μm respectively for 25×, 40× and 100× objectives.

Eq. (9) gives a non-linear relationship between phase-shift angle δ_i and the rotation angle of refractive slab α_i. This relationship can be further simplified when α_i is small, for example, less than 20°.

$$\delta_i = 2\pi dK(\alpha_i)/P \approx 2\pi d\alpha_i(n-1)/(nP) \tag{17}$$

Thus the phase shift error due to an error in the rotation of the refraction slab rotation can be expressed as:

$$\Delta\delta_i \approx 2\pi d(\Delta\alpha_i)(n-1)/(nP) \tag{18}$$

As depicted in section 4.1, the rotation of the refraction slab can be controlled to within 0.05° Based on a slab refractive index of n = 1.37859, slab thickness of d = 1.52mm, and grating pitch multiplied by objective magnification (P^*), the calculated

phase shift error is small. For example, for a 100 lines/mm grating and a 40× objective, $P = 4$ mm, the error is less than 0.33°. Comparing this with the fringe intensity noise, the error is negligible.

The main error is the fringe intensity noise. We assume that the standard deviation (SD) of this noise is $\sigma = 3$ gray levels. This estimation is reasonable and conservative, and can be verified by a simple test, wherein two fringe patterns in the same state but at different times are recorded and subtracted. The result obtained is a noise pattern, from which the SD of noise can be calculated, Dividing it by $\sqrt{2}$, we obtain the approximate SD of fringe intensity noise and found it to be less than 2.

The calculated phase error due to the noise reaches its maximum when δ_0 is 90°. In this case, the algorithm in Eq. (13) becomes the well-known 5-buckets error-compensating algorithm]. The phase error of this algorithm due to the noise is phase dependent, and its maximum SD value can be estimated as[13]

$$\sigma_{\Delta\varphi} = (1+\beta)/(\eta\sqrt{N}SNR)$$

Where $\beta = 1/7$, $\eta = 4\sqrt{2/35}$, $N = 5$ is the fringe pattern number. $SNR = I_b/\sigma$ is the signal noise ratio. Suppose the modulated fringe intensity varied from 64 to its maximum 128 grays, the correspondent SNR varies from 21.3 to 42.7. Thus, the maximum phase error due to noise is between 0.013 and 0.025 grd, or we say the phase error is less than 1.4 degree. Finally, we take the SD of the phase error as $\sigma_{\Delta\varphi} = 2$ degree, and get the displacement measurement precision as follows:

Grating density (L/mm)	40	80	100	150	200
Displacement precision (μm)	0.139	0.069	0.056	0.037	0.028

4.3. Test Example

In order to verify the experimental procedure and check the measurement precision, a test was conducted using the microscope translation stage to stretch a grating film specimen, which is bonded onto the stage. As mentioned in section 3.4, this film including a plastic substrate is very thin, so it is easily loaded by using the micrometer drive, and we can precisely control the displacement. This displacement read from the micrometer is the overall displacement of the specimen. However, the displacement field measured with the above mentioned grating phase-shift method covers only small part of the specimen length, so it is difficult to compare the localized displacement obtained with the grating phase-shift method and the overall displacement. Actually, we compare the average strain calculated based on the specimen's overall displacement, i.e., micrometer displacement divided by the length of specimen, and the strain calculated from the experimentally measured local displacement.

Fig. 8 shows a test result. The first 5 photographs are the phase-shifted grating line patterns before loading, and the last sub-figure shows the wrapped phase map. The grating density is 100 lines/mm, and a 40× objective is used. The grating line patterns are consecutively recorded with the CCD camera each time after the grating image is shifted by turning the tilting compensator, and these grating images are stored in the extended memory of the computer. The difference among the 5 patterns is that the phase of the grating signal in each pattern is shifted. The phase distribution is then calculated using Eq. (13) from the 5 stored grating images. The corresponding unwrapped phase map representing a displacement field with an area of 50×50 μm^2 is displayed in Fig. 9 in gray levels. This displacement map corresponds to a specimen with no-load and is referred to as the virtual displacement field. After loading, the deformed gratings with added phase shifts are recorded, and the final phase (or displacement) map is obtained. The subtraction of the final and the initial phase maps gives the phase (displacement) field due to the load alone. Fig. 10 gives a comparison of the average strain calculated based on the specimen's overall displacement, and the strain calculated from the experimentally measured local displacement. 10 tests were conducted with 1000 micro-strain increment of average strain, with the average strain for first test as 1000 micro-strain. We see the two groups of results are in general agreement.

5. DISCUSSION AND CONCLUSION

A new method of grating diffraction for determining the plane strain has been described. By developing and reconstructing a traditional microscope, a direct measurement of strain is developed. This gives the microscope a new important usage. This measurement system possesses a higher automatization and processing speed. To achieve high spatial resolution displacement measurement, a combination of microscopy with phase shift technique is adapted to grating measurement. A new phase shift method using optical slab refraction effect is suggested. The manipulation of the phase shifter is very simple, but with high precision. Applications of this phase shift method to various grating/ moiré measurement are proposed that can make the system compact but high precision. Through low cost development and reconstruction of a traditional microscope, a field displacement measurements has been realized. This provides a novel use for the microscope. This system is originally

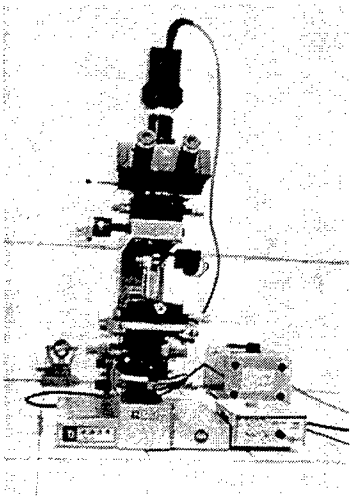
designed for investigating the micro-mechanical and photo-mechanical behavior of plastic film. One example is studying the interaction of micro-holes of a specimen film. In addition to film specimen, other non-film polymer specimens can also be tested. These tests can be either static or of creep. The system can only be used on transparent or demi-transparent specimen. The performed experiment can be uniaxial tension/compression test, either static test. If a reflecting light microscope is used, opaque specimens could be measured, as the principle and the technique presented in this paper is applicable to this microscope.

ACKNOWLEDGEMENTS

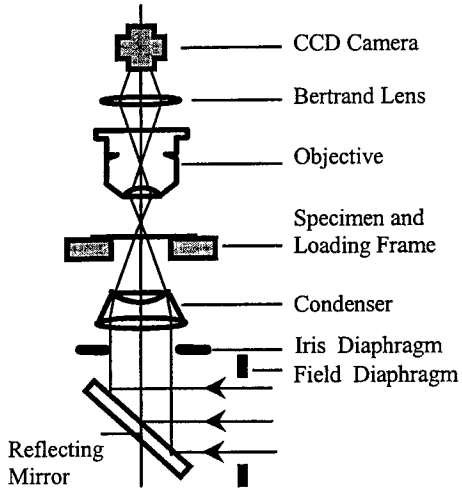
The support of Ministry of Education and Nanyang Technological University through research grant MLC1/97 is acknowledged.

REFERENCES

1. J. F. Bell, "Determination of dynamic plastic strain through the use of diffraction gratings," *J. of Appl. Phys.*, **27**, pp. 1109-1113, 1956.
2. J. C. Moulder J. F. Cardenas-Garcia, "Two-dimensional strain analysis using a video optical diffractometer," *Experimental Technique*, September/October, pp. 11-16, 1993.
3. P. J. Sevenhuijsen, J.S. Sirkis, and F. Bremand, "Current trends in obtaining deformation data from grids," *Experimental Technique*, May/June, 22-26, 1993.
4. F. Bremand, J. C. Dupre and A. Lagarde, "Non-contact and non-disturbing local strain measurement methods, I. Principle," *Eur. J. Mech., A/Solids*, **11**, pp. 349-366, 1992.
5. A. Asundi, "Developments in the moiré interferometric strain sensor," *Proc. SPIE*, Holographic interferometry and speckle metrology, Vol. 1163, pp. 63-68, 1989.
6. J. E. Greivenkamp, J.H. Bruning, "Phase shifting interferometers," Chapter 14, *Optical Shop Testing*, edited by D. Malacara, John Wiley & Sons, Inc, New York, pp501-598, 1992.
7. R.S. Sirohi, M.P. Kotthiyal, "Heterodyne and phase shifting interferometry," Chapter 6, *Optical Components, Systems, and Measurement Techniques*, Marcel dekker, inc, New York, pp.219-246, 1991.
8. P. Carré, "Installation et utilisation du comparateur photoelctrique et interferentiel du bureau International des poids et mesures," Métrologie **2**, pp.13-23, 1966.
9. P. Hariharan, B.F. Oreb, and T. Eiju, "Digital phase-shifting interferometry: a simple error-compensating phase calculating algorithm," Appl. Opt. 26, pp.2504-2506, 1987.
10. M. Born, E. Wolf, "*Principles of Optics, electromagnetic theory of propagation, interference and diffraction of light*," 6th ed. Pergamon Press, 1983.
11. G.A.W. West, T.A. Clarke, "A survey and examination of subpixel measurement technique," *Proc. SPIE,* Close-Range photogrammetry meets machine vision, Vol. 1395, pp.456-463, 1990.
12. J.C. Dupre, M. Cottron and A. Lagarde, "Grating interrogations: from small to large strain measurement," *Experimental Mechanics*, pp. 153-158, 1995.
13. Y. Surrel, "Additive noise effect in digital phase detection," Appl. Opt. **36**, pp. 271-276, 1997.



(a)



(b)

Fig. 1 Sketch of optical system and principle. (a). Photo of Leitz-Orthoplan II Pol polarizing microscope; (b). Optical components of the measurement system, optical paths and images of

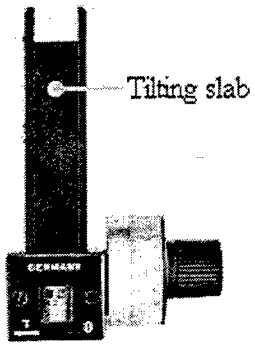


Fig.2 Photograph of tilting compensator

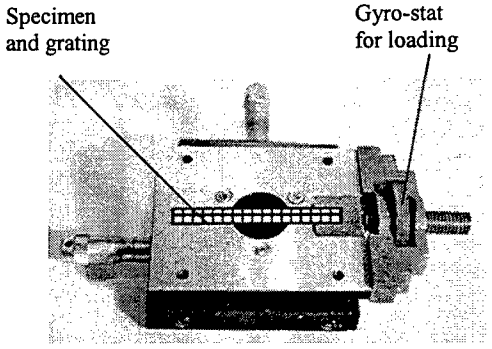


Fig.3 loading frame

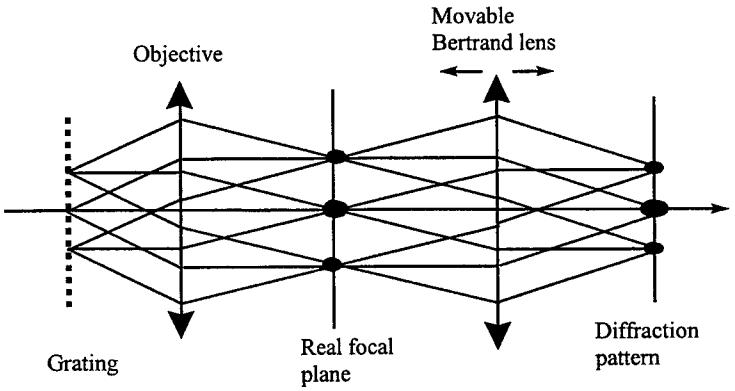


Fig. 4 Microscopic grating diffraction illustration of image forming from objective lens to Bertrand lens. The diffraction spots are formed in the back focal plane of the equivalent lens which includes an objective and a relay lens.

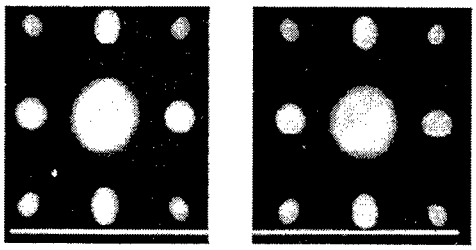


Fig. 5 Diffraction patterns of an 80 l/mm cross grating before and after loading. The white line is a reference one. After the deformation, diffraction spots in last row are shifted towards this reference line. The two black lines indicate the centroid positions of middle spot in last row.

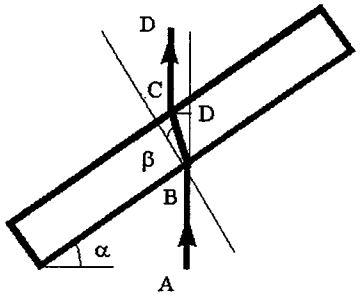


Fig. 6 Optical slab and refraction

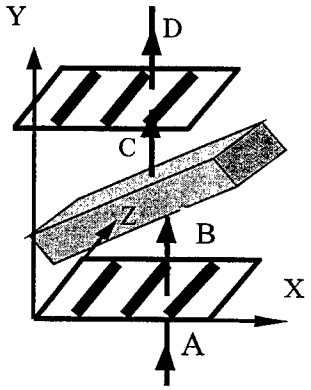


Fig.7 Illustration of a grating shifting by an optical slab

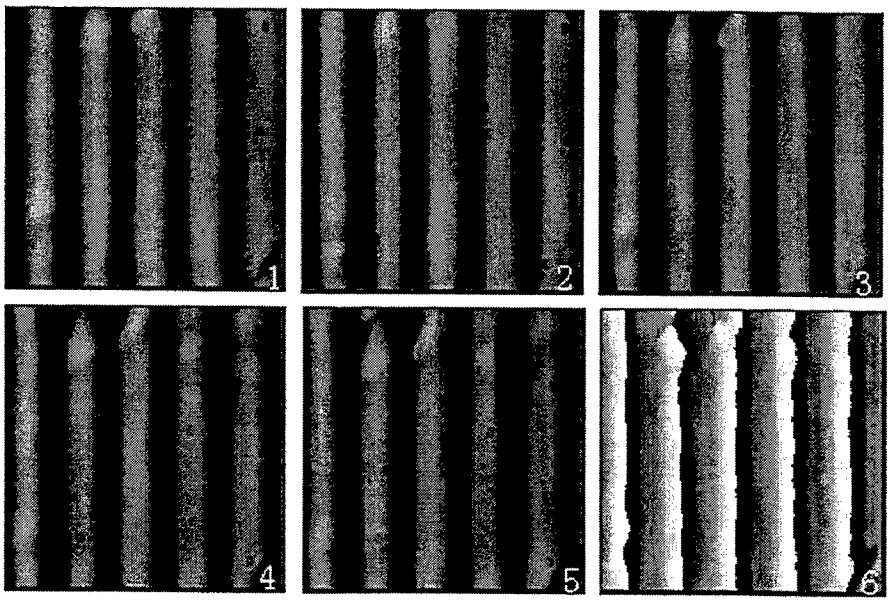


Fig. 8 Test example, phase shifted grating line patterns before loading and calculated wrapped phase map. The pixel value (gray level) in the grating line patterns represents the grating line intensity, and the pixel value in the phase map (photograph 6) represents the phase of grating signal. The difference between the first 5 photographs is they are relatively shifted from right to left. The displayed field is an area of 50×50 μm^2.

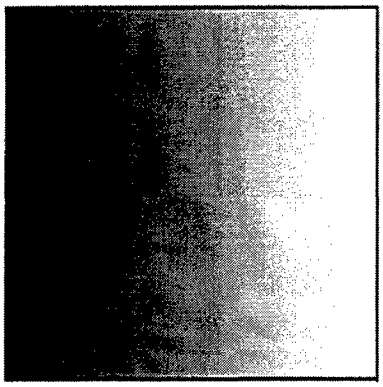


Fig. 9 Test example, unwrapped phase map representing a displacement field displayed in gray levels.

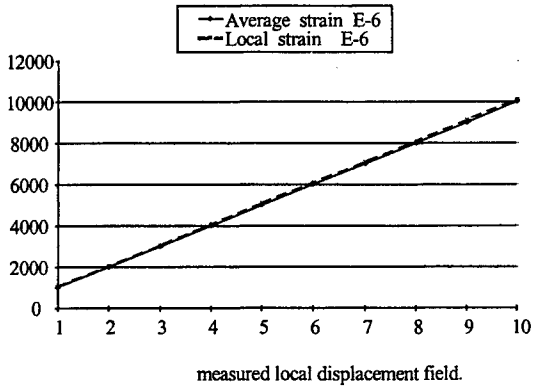


Fig. 10 A comparison between results of the average strain calculated with the total displacement of the specimen and the local strain calculated with measured local displacement field.

Usage of polarization for high-accuracy micro-metrology sensors

M. Totzeck, H. Jacobsen, H. J. Tiziani
Institut für Technische Optik, Universität Stuttgart, D-70569 Stuttgart, Germany

ABSTRACT

The general idea of the presented investigations is to use the polarization of the electromagnetic field in high-resolution optical microscopy to get information about sub-wavelength details of topographical structures. The main application is the localization of vertical edges. For structures in non-magnetic materials, polarization effects are caused by the different boundary conditions for the tangential and normal electric field components. Using rigorous numerical simulations we show that two physical-optics model describe the polarization dependent images of vertical edges and sub-wavelength structures: a boundary diffraction wave originating from the tip of the edge and waveguiding effects, respectively. We report on two experimental approaches for polarization usage: a) comparison of interference microscopy images for s- and p-polarization, and b) polarization interferometry. The former is capable of „synthetical microscopy". Measurements of well defined structures (calibrated with a scanning force microscope) are compared with rigorous numerical simulations.

Keywords: Metrology, polarization microscopy, CD measurements, boundary conditions, numerical simulation

1. INTRODUCTION

Optical inspection of surfaces has the well known advantages of a minimum interaction with the object and high measurement rates. All this at comparatively low technical expense. The not less well known main problem is the Rayleigh resolution limit, i.e. limitation of the resolution to the half wavelength. This is crucial for the microstructures considered in ths contribution (Fig. 1). However, the reproducibility of optical measurements is much better than the resolution: in principle it is possible to localize the position of an isolated edge according to threshold- or extreme-value criteria[1] with <1nm precision. Although a precision of 3 nm is state of the art, deviations of up to 100 nm to the true edge position are observed.[2] The reason for the decreased accuracy is the assumption of a reflection (or transmission) according to geometrical optics neglecting physical-optics effects. The electromagnetic boundary conditions provide a polarization dependent edge scattering resulting in polarization dependent threshold values. If edges are closely neighbored, multiple scattering occurs, resulting in an even increased polarization dependence. These polarization dependence must be known and understood in order to correct microscopical measurements, i.e. to calibrate the sensor. To this end, physical-optics models are very useful.

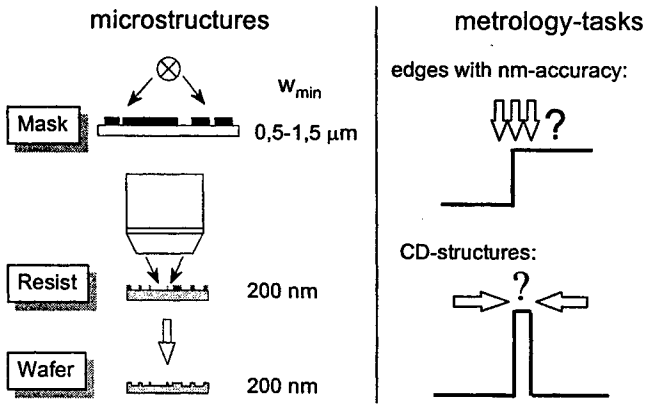


Fig. 1: Objects and tasks of optical micrometrology

Microstructures of main interest are photomasks, resist on silicon and etched wafers with a critical dimension of now <200 nm (Fig. 1a). For the present investigation, we restrict ourselves to pure topographical isolated edges or line structures down to sub-wavelength width (Fig. 1b).

From a physical optics point of view, usage of the polarization in micro-metrology means to perform measurements of structure dimensions that take two different boundary conditions as a basis. The advantages are obvious. For example:

Correspondence: MT: totzeck@.ito.uni-stuttgart.de ,HJ: jacobsen@ito311.ito.uni-stuttgart.de HT: tiziani@ito311.ito.uni-stuttgart.de, phone: ++49.711.685.6074, fax: ++49.711.685.6586, www http://uni-stuttgart.de/UNIuser/ito/

Part of the SPIE Conference on Advanced Photonic Sensors and Applications
Singapore • November–December 1999
SPIE Vol. 3897 • 0277-786X/99/$10.00

a) increased accuracy of threshold-value measurements, for instance by finding an optimum polarization

b) simultaneous imaging and detection of regions where the boundary conditions differ. The signal strength depends on the amount of the difference which itself depends on edge-material and –angle. This enables a discrimination of a damaged or soiled edge from an edge at the wrong position.

c) structure characterization where no spatial information is available, as for sub-wavelength gratings.[3]

In a restricted sense, polarization usage means the generation of contrast from the polarization difference, i.e. from the fact that the x-polarized image $E_x(x,y)$ does not equal the y-polarized image $E_y(x,y)$.[18] This could be done, for instance, by the crossed polarizer configuration[4,5] (image = $|E_x^2 + E_y^2|$) or polarization interferometry[6] (image = $E_x E_y^*$). However, several more contrast-mechanisms (i.e. functions of E_x and E_y) are possible.[19] Which should be applied for a given structure and illumination and what systematical errors should we expect? In our contribution we try to present a method to answer these questions.

The organization of the paper is as follows: In the next section we explain the measurement of complex fields with interference microscopy. Then we present measured and computed polarization effects in section 3 and discuss their physical optics origin in terms of boundary diffraction waves and waveguiding effects in section 4. An interferometrical method to use polarization effects for enhanced edge detection are presented in section 5.

2. COMPLEX FIELD MEASUREMENT WITH INTERFERENCE MICROSCOPY

For a systematical investigation of the influence of polarization on the image of line-structures we employed a Linnik-type interference microscope that is equipped with a linear polarizer in the illuminating section. We will show that it enables a measurement of the diagonal elements of the Jones-matrix of the object. This provides a direct measure for the polarization effect due to the structure. Furthermore, we can perform "synthetical microscopy", which means that we can take the complex field as basis to perform typical microscopic operations, like focusing, polarization contrast, or dark-field contrast, subsequently in the computer.

2.1 Interference microscopy set-up

The interference microscope is based on a modified LEICA-Orthoplan 100. The optical system is sketched in Fig. 2. A mercury high-pressure lamp (HBO50) with subsequent interference filter (λ=549nm, $\Delta\lambda$=20nm) polarizer and adjustable aperture stop provides an approximately monochromatic, linearly polarized, partially coherent illumination with adjustable coherence parameter (ratio of illumination NA and imaging NA) between 0.2 and 0.95. For the present investigations, a coherence parameter of 0.2 was used.

The images of the investigated structure and a plane reference mirror, provided by two paired microscope lenses (100x, NA=0.9), are superposed by a beamsplitter onto a CCD-camera chip with 8.2 x 9.4 µm-pixel pitch (Hitachi KP160). A second magnifying lens-system in front of the CCD-camera yields an additional 4x-magnification with the consequence that the system operates in the regime of empty magnification with an object distance of 21 nm between adjacent camera-pixels. The image is sampled with 10 bit resolution using a MV1000 frame-grabber. With proper adjustment, high-contrast, straight interference fringes can be observed even for white light illumination.

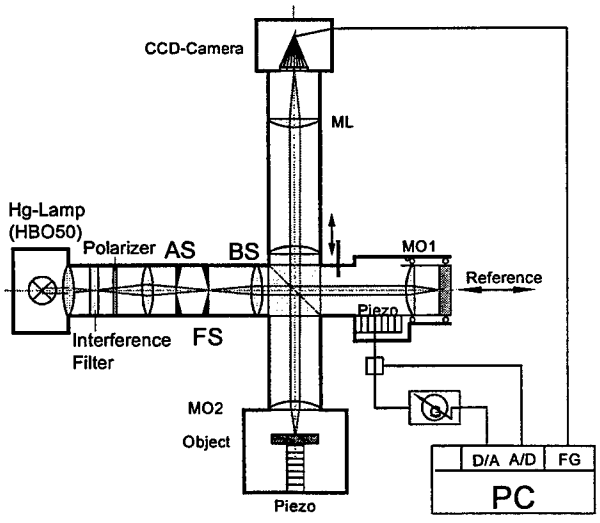


Fig. 2: Linnik-type interference microscope (MO1, MO2: paired microscope lenses, ML: magnifying lens system, AS: aperture-stop, FS: field-stop)

The noise is reduced by two means: 1) 10 frames are averaged for each interferogram. 2) Because of the high magnification the interferograms may be low-pass filtered considerably without affecting the lateral resolution. With λ=549nm, M=400 and 9µm pixel-size, NA_{Ill}=0.25, NA_{Img}=0.9, the cut-off frequency of $1.1/\lambda$ in the object plane is equivalent to a cut-off frequency of 24 pixel^{-1} in the image plane. Therefore, each interferogram is low-pass filtered using a Gaussian filter with a standard deviation of 2 pixel.

2.2 Interferogram-evaluation for vector fields

The interferogram on the CCD chip stems from the vector field $\mathbf{E_O}$ due to the object and a phase shifted reference field $\mathbf{E_R}$ according to (to simplify the notation, the spatial dependence is not written)

$$I_{\Delta\varphi} = \left|\mathbf{E_O} + \mathbf{E_R}\exp(i\Delta\varphi)\right|^2 \tag{1}$$

that becomes in the usual form

$$\mathrm{I}_{\Delta\varphi} = \mathrm{I}_g\left(1+\gamma\cos(\phi+\Delta\varphi)\right) \tag{2}$$

with phase ϕ, intensity I_g, and contrast γ

$$\begin{aligned} I_g &= \left|\mathbf{E_O}\right|^2 + \left|\mathbf{E_R}\right|^2 \\ \gamma &= \frac{2\left|\mathbf{E_O}\cdot\mathbf{E_R}\right|}{I_g} \\ \phi &= \arg\left(\mathbf{E_O}\cdot\mathbf{E_R^*}\right) \end{aligned} \tag{3}$$

The measurement is performed according to the eight-step method of phase-shift interferometry[7]: 8 images I_j (j=1-8) are captured where the j'th component is phase-shifted by $\beta_j = j\pi/2$. Phase ϕ, intensity $\mathrm{I}_0 = \mathrm{I}_g$, and contrast γ are obtained by

$$\tan\phi(x,y) = \frac{Z}{N} \tag{4}$$

and

$$\gamma(x,y) = \frac{\sqrt{Z^2+N^2}}{32I_0} \tag{5}$$

with

$$\begin{aligned} Z &= 5I_2 - 15I_4 + 11I_6 - I_8 \\ N &= I_1 - 11I_3 + 15I_5 - 5I_7 \\ I_0 &= \frac{1}{8}\sum_{j=1}^{8} I_j \end{aligned} \tag{6}$$

The phase-shift is performed either by moving the reference lens and mirror simultaneously (the reference mirror is always in focus) with a piezo or by defocusing the object using a piezo-driven table (Physik-Instrumente). The z-position of the piezo-driven table is actively stabilized using capacitive sensors with sub-nm resolution. We used for the present investigations the phase shift of the reference because for an object-defocus the well known aperture factor of interference microscopy[8] has to taken into account.

According to (3), the scalar, complex quantity $\mathbf{E_O}\mathbf{E_R}^*$ is obtained from

$$\mathbf{E_0}\cdot\mathbf{E_R^*} = 0.5\gamma\mathrm{I}_g\exp(\mathrm{i}\phi) \tag{7}$$

If we ensure normally incident, plane, linearly polarized reference wave along the x- and y-direction, i.e.

$$\mathbf{E_x^R} = \begin{pmatrix}1\\0\end{pmatrix} \quad \text{and} \quad \mathbf{E_y^R} = \begin{pmatrix}0\\1\end{pmatrix} \tag{8}$$

we get from Eq. (4) the complex field components $\mathbf{E_x^O}$ and $\mathbf{E_y^O}$.

Demanding a linearly polarized reference wave implies an optics without depolarizing elements. However, we used an interference microscope without particular strain-free lenses and without a non-depolarizing beam-splitter. Experiments have

shown, that the main source for polarization-anisotropy in the set-up is the beam-splitter. Due to its construction a beam-splitter has two orthogonal eigenpolarizations along which no depolarization occurs. We have chosen these directions as the x- and y-axis of the global coordinate system.

For each structure, two measurements are performed, with an incident x-polarized and an incident y-polarized wave, respectively.

The intensity $I=|E|^2$ of the object wave can be measured separately by inserting a stop into the reference arm. Combined with Equ. (3) we may also obtain the complex object wave from

$$E = I^{1/2}\exp(\phi(x,y)) \tag{9}$$

This straight forward expression has the advantage of being independent from intensity-variations of the reference. But it may lead to inconsistencies as will be shown below.

Eq. (7) and (9) are equivalent if no depolarization occurs at the structure. This is the case, for instance, for a line structure and an incident wave polarized parallel or perpendicular to its edges. In our coordinate system, the configuration has a diagonal Jones-matrix if the edges are oriented along the x- and y-axis

$$\mathbf{E_0} = \begin{pmatrix} J_{xx} & 0 \\ 0 & J_{yy} \end{pmatrix}\mathbf{E_i} \quad \text{with} \quad \mathbf{E_i} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \text{ or } \mathbf{E_i} = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \tag{10}$$

For a depolarizing structure, i.e. a configuration with a Jones-matrix with non-diagonal elements, we get

$$\mathbf{E_0} = \begin{pmatrix} J_{xx} & J_{xy} \\ J_{yx} & J_{yy} \end{pmatrix}\mathbf{E_i} = \begin{pmatrix} J_{xx} \\ J_{yx} \end{pmatrix}E_x^i + \begin{pmatrix} J_{xy} \\ J_{yy} \end{pmatrix}E_y^i \tag{11}$$

Assuming a perfectly reflected reference wave we set $\mathbf{E_R}=\mathbf{E_i}^*$ and get

$$\mathbf{E_0}\cdot\mathbf{E_R} = J_{xx}\left|E_x^i\right|^2 + J_{yy}\left|E_y^i\right|^2 + J_{xy}E_y^iE_x^{i*} + J_{yx}E_x^iE_y^{i*} \tag{12}$$

which becomes J_{xx} for a x-polarized and J_{yy} for a y-polarized incident wave, i.e. we measure separately the diagonal elements of the Jones-matrix even if non-diagonal elements are present. But in. Eq. (9) the intensity obtained with shut reference arm, includes the non-diagonal-elements

$$I_O = \left(\left|J_{xx}\right|^2 + \left|J_{yx}\right|^2\right)\left|E_x^i\right|^2 + \left(\left|J_{yx}\right|^2 + \left|J_{yy}\right|^2\right)\left|E_y^i\right|^2 \tag{13}$$

While the phase ϕ according to Eqs. (3) and (4) does not.

2.3 Example for synthetical microscopy: a-posteriori focusing

Consider a trench of 121 nm depth and 710 nm width in a Silicon substrate. For a linearly polarized, perpendicularly incident plane wave of 549nm wavelength, the reflected, complex field is imaged according to section 2.1 and 2.2. The obtained amplitude and phase are shown in Fig. 3.

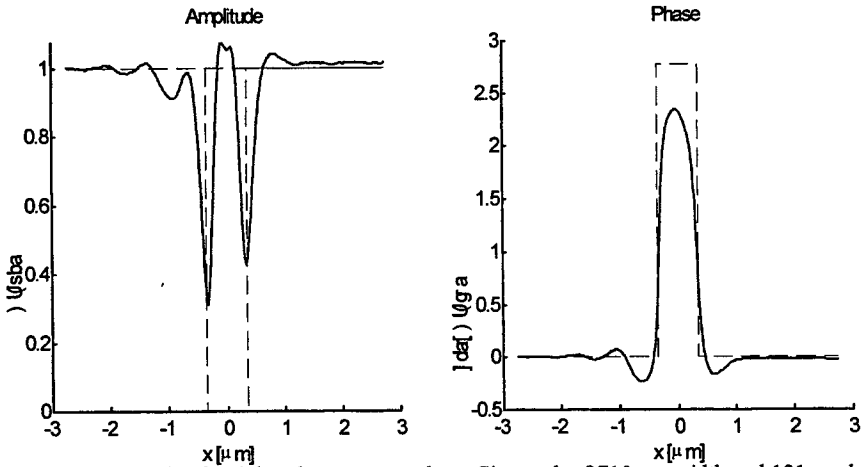


Fig. 3: Left: amplitude, right: phase measured at a Si trench of 710 nm width and 121nm depth in Si.

The incident wave is not perfectly coherent, because the finite half-width of the applied

interference filter ($\Delta\lambda$ = 20 nm) and the non-zero coherence parameter. With an illumination NA of 0.25, the coherence parameter becomes 0.28. Assuming a coherent field $U(x,y,z)$ we may compute a defocus of Δz from a focusposition $z=z_0$ according to

$$U(x,y,z_0+\Delta z) = F^{-1}\left\{F\left\{U(x,y,z_0\right\}\exp\left[i\left(k_0^2-k_x^2-k_y^2\right)^2\Delta z\right]\right\} \tag{14}$$

where F{} denotes the Fouriertransform and $F^{-1}\{\}$ its inverse. k_0 is the free space propagation constant and k_x and k_y are the spatial frequency components in x- and y-direction, respectively. Aberrations of the optical system could be included into Equ. (13), but this was not done here.

Fig. 4 shows a comparison of a computed intensity focus scan (left) and a measured one (right).

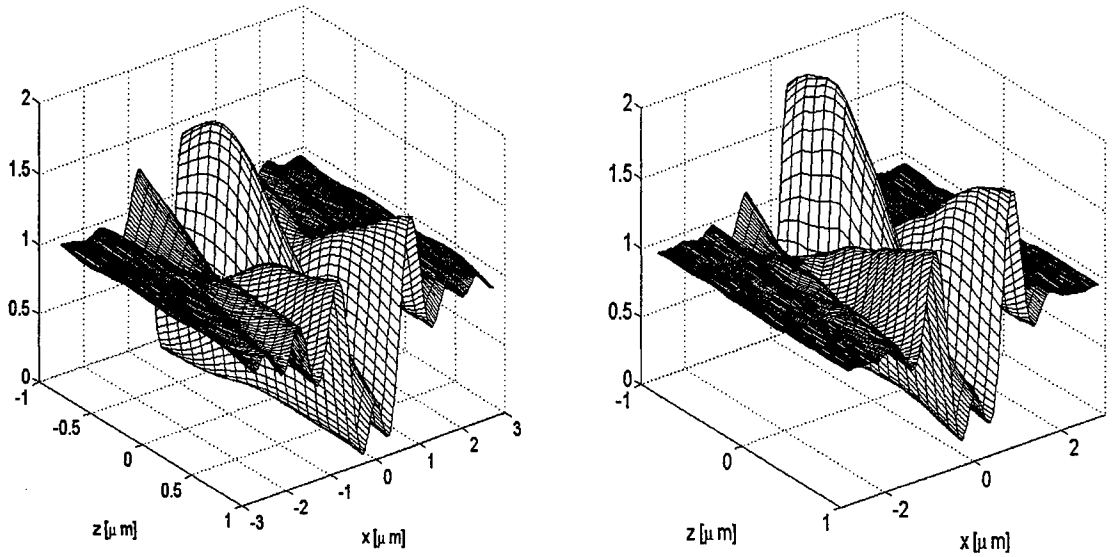


Fig. 4: Comparison of synthetic (left) and measured defocus (right) at a Si trench of 710 nm width and 121nm depth in Si

The agreement is satisfactorily. The main discrepancy is observed in the central peak for negative defocus (here: above the structure). Notice the higher diffraction lobes in the synthetic defocus. This stems from the fact that the mesurement is performed with a partially coherent wave field while the computed defocus according to Eq. (14) assumes a perfectly coherent wave field.

3. POLARIZATION EFFECTS IN COHERENT MICROSCOPY

Using the field measuring capability of interference microscopy and a rigorous numerical simulation of the imaging process we will give in this section an account of polarization effects observed in the images of topographical line structures. In particular we will investigate wide structures and small lines of sub-wavelength width. The structures are oriented parallel to the y-axis. They are etched into silicon that has an ellipsometrically determined refractive index of n=3.8+i*0.3 at λ=633nm. The refractive index of bulk Si is according to Ref. 9 is n=3.882+i*0.019 at λ=633nm, i.e. the real part agrees well but the imaginary part is significantly higher in our measurement. At our measurement-wavelength of 549nm, Ref. 9 gives a value of n=4.18+i*0.04 for the refractive index of Si. Having the increased imaginary part at HeNe wavelengths we assume that the imaginary part for our samples is also higher at λ=549nm. Therefore we used a refractive index of n=4.2+i*0.3 for a comparison with the interference microscopy measurements.

The focus position for the investigated small structures is adjusted visually so that the intensity above the structure shows two closely neighbored, sharp minima and a „coherent ringing" that is at most close to the structure edges. This is not a strict but a sufficient focus criterion for the present investigation.

For a rigorous computation of the near-fields and coherent images of semiconductor structures, we developed a procedure based on the rigorous coupled wave theory (RCWA[10]) and Hopkins image formation theory using the effective source[11].

3.1 Wide trench

The first considered structure is a trench of 910 nm width and 121 nm depth. Both values are determined by scanning force microscopy. Fig. 5 shows amplitude and phase of the measured field components (circles) and the rigorous simulations (lines). For comparison the geometrical optics phase shift is included into the phase images.

Evidently, the modulation of amplitude and phase is stronger if the electric field vector is oriented perpendicular to the structure edges (TM-polarization) than in TE-polarization where it is oriented parallel to the edges. The edges are parallel to the y-axis. From a metrological point of view TM polarization seems to be better suited for edge localization of the present structure than TE-polarization. Despite this difference the agreement of the phase half-width and also of the amplitude minima with the actual structure edges is quite well.

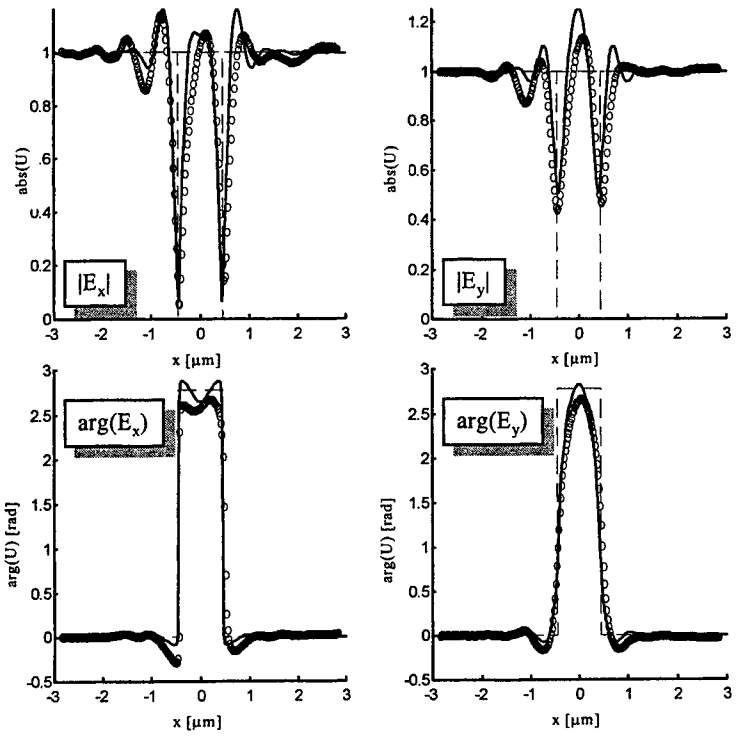


Fig. 5 Amplitude (top) and Phase (bottom (o: measurement, -: computation) of a Si trench of 910 nm width and 121nm depth in Si. left: TM-Pol., right: TE-Pol

Unfortunately, the conclusions from the above paragraph are not a general behavior: it depends on the structure depth as is demonstrated in Fig. 6, where a rigorous computation of the amplitude of edge images for steps of various heights (0-500 nm) is shown. The height is increasing from top to bottom.

The focus position is fixed to the top of the edge. The minimum is indicated by a solid line. It shows oscillations in accordance with the phase shift introduced by the step. However, the oscillations are more pronounced and more regularly in TM than in TE polarization. The difference becomes stronger with increasing illumination NA. Furthermore, the minimum is shifted away from the edge. Therefore, the minimum detection method for edge localization provides problems with increasing structure height.

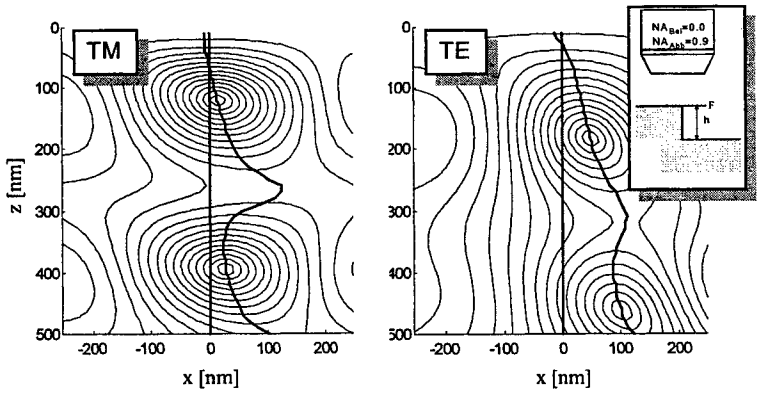


Fig. 6 Thickness-dependence of minimum-position at an Si-edge

3.2 Signal strength of sub-λ line

As is well known, the difference of TE and TM polarized images increases with decreasing line width. The effect for complex images is demonstrated in Fig. 7. The experimental parameters are identical to Fig. 5 but now the object is a trench of only 210 nm width. The observed amplitude modulation is distinctively stronger in TM than in TE polarization. I.e. the scattering cross section is higher in TM polarization. In TM polarization, two minima are still discernible with a distance of approximately 200nm, while they have already merged in TE polarization. An interesting phenomenon is observed in the phase image: The sign of the phase and therefore the apparent topography is opposite for both polarizations.

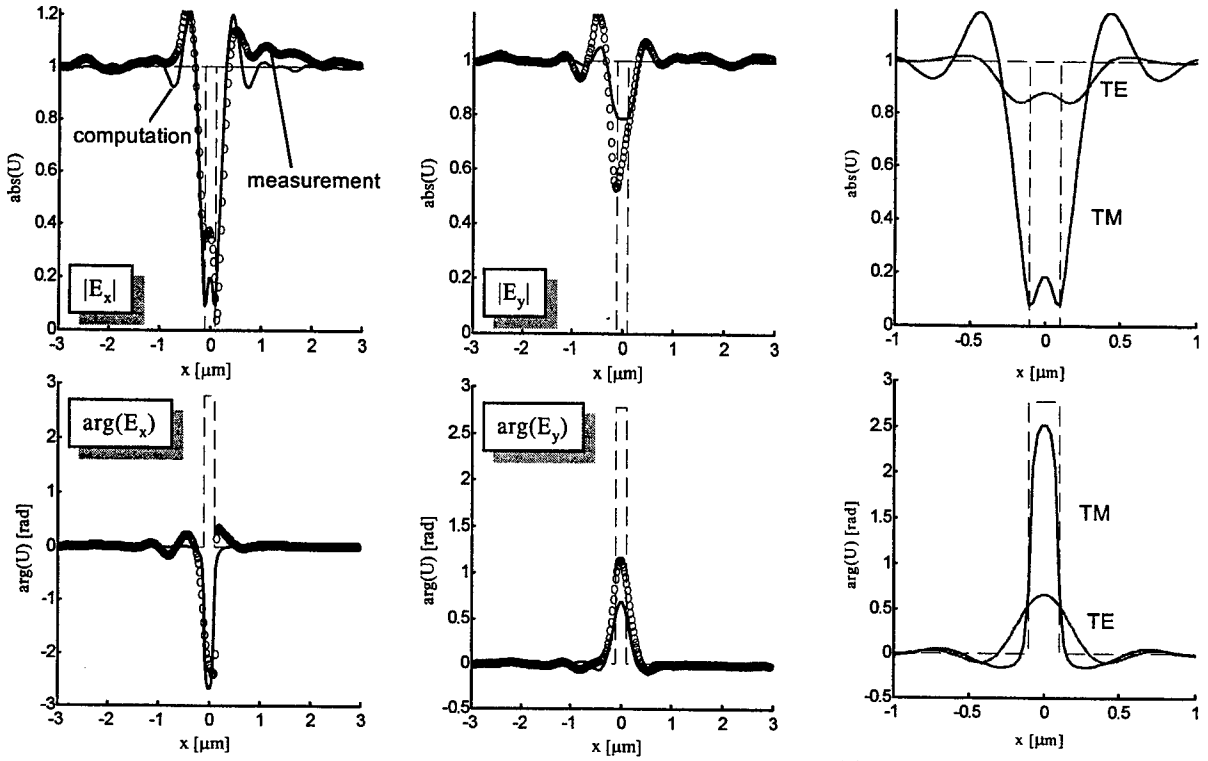


Fig. 7 Amplitude (top) and Phase (bottom) of a Si trench of 210 nm width and 121nm depth in Si. left: TM-Pol., right: TE-Pol

Fig. 8 Computed field of Fig. 7 with 120nm defocus.

The reason for the latter behavior is the close proximity of two phase-singularities[12] to the focus plane in TM-polarization. Phase-singularities are points in a wave-field where the intensity is zero with the consequence that the phase is not defined. Crossing the plane where the singularities have formed results in an approximately inverted phase-image. This is a consequence of the general property of phase singularities that crossing them yields a π-phaseshift. Here, a slight negative defocus crosses the singularity plane and removes the effect (Fig. 8).

Concluding this section we discuss what happens if we actually reverse the structure, i.e. if we image a bar of 210 nm width and 121 nm height. A comparison of both cases is shown in Fig. 9. Because the structure was not available, only computations are shown. The images have changed strongly: The general modulation is stronger and the amplitude images have become more alike. This is also true for the phase, although the sign is still reversed. Now, the half width of the phase-image as well as the minima in the amplitude image yield a comparatively large deviation from the actual edge position, i.e. they are less valid indicators for the true linewidth.

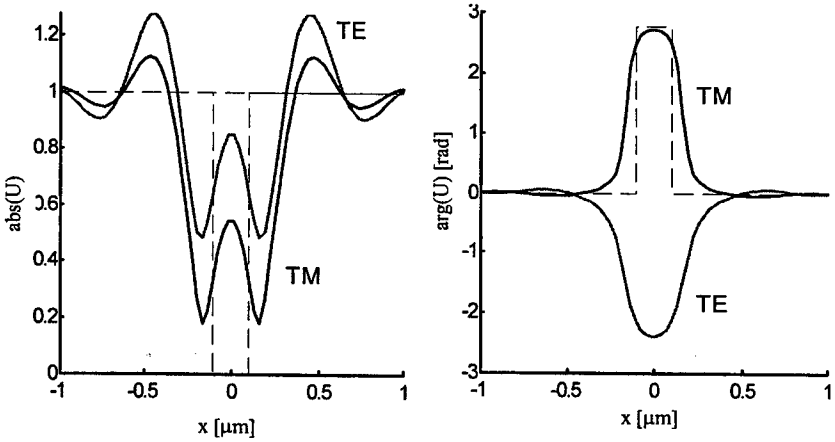


Fig. 9 Computed amplitude (left) and Phase (right) of a Si-bar of 210 nm width and 121nm height on Si for TE and TM polarization

4. PHYSICAL-OPTICS MODELS

In a rather simplified manner we may say that the phase image of a pure topographical structure represents the topography, the amplitude image the amount of scattering beyond the imaging NA and the polarization effect the difference of boundary conditions for both orthogonal polarizations. These are most pronounced for edges and inclined sections.

To use the polarization effects for resolution enhancement or at least a better interpretation of the images, their origin should be understood in terms of simple physical-optics models. Of course it is possible, as was demonstrated in the preceding section, to compute the images rigorously and to include by this means vector aspects into the analysis. But to our opinion this is more useful for systematical investigations and eventually providing a kind of look-up table (polarization effect vs. structure parameter). For the development of sensor concepts using these effects a more „operational understanding" is necessary. In the following two subsections we discuss rather simple models to explain the observed effects

4.1 Boundary diffraction wave

The concept of the boundary diffraction wave dates back to T. Young. The basic idea is to consider diffraction (U_{diff}) as a deviation from the geometrical optics propagation (U_g). This is due to a boundary diffraction wave[13] U_{BDW} originating from the rim of the aperture or in our case from the structure edges.

$$U_{diff} = U_g + U_{BDW} \qquad (15)$$

U_{BDW} is discontinuous to compensate for the discontinuity of the geometrical optics field. Polarization dependent boundary diffraction waves from diffraction by perfect conductors are well known.[14] This model was used by Kimura and Wilson[5] to understand the contrast mechanism of confocal microscopy between crossed polarizers („polarization dark-field").

Because for perpendicular incidence and vertical edge walls the geometrical optics fields are identical for both polarizations, this concept is useful to describe the difference of differently polarized images. Fig. 10 shows the difference of the rigorously computed near-field and the geometrical optics near-field for both linear polarizations and two step heights in silicon. The geometrical optics field was computed according to the thin film matrix theory (TFMT[15]).

In Fig. 10 the magnitude of the real part of the electric near-field is plotted. It represents the field amplitude at a certain moment. It visualizes by this means both amplitude and phase of the field.

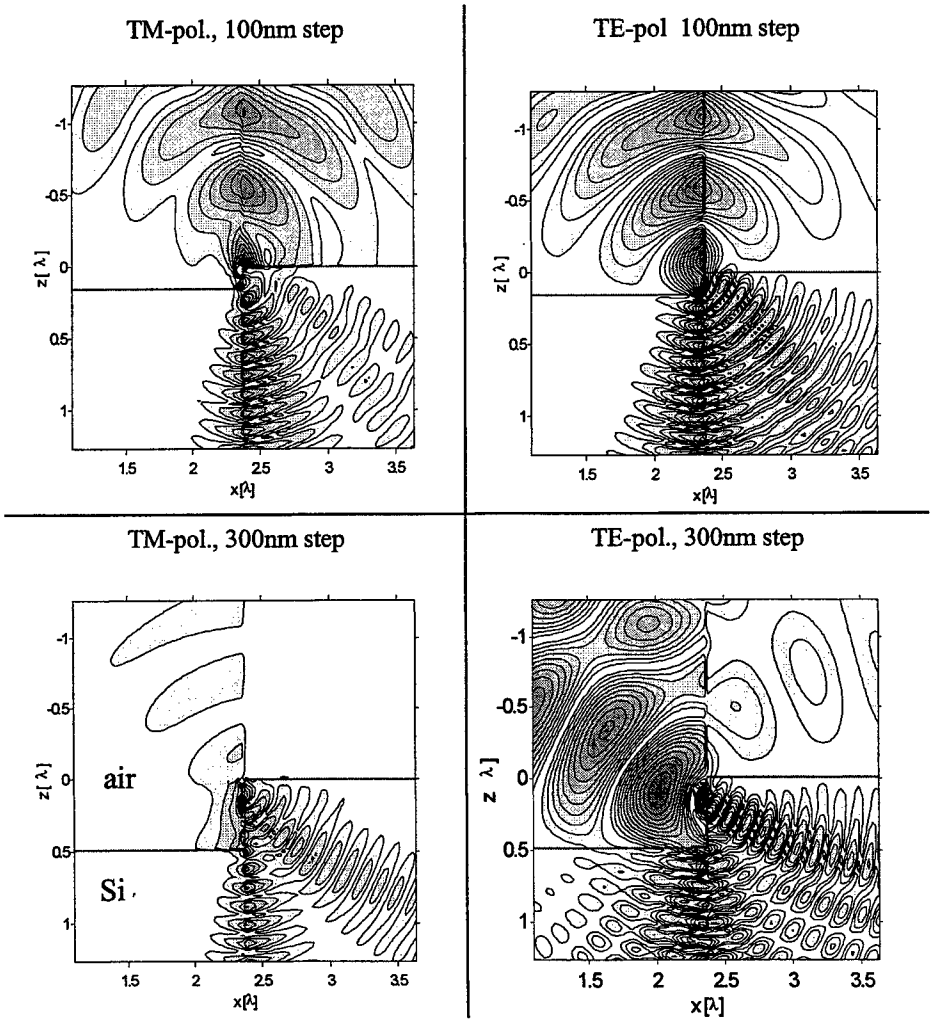


Fig. 10 Nearfield (magnitude of real part of electric field) of the boundary diffraction wave (U_{diff}-U_g) at steps of different heights in Si, λ=549nm.. Identical color-maps.

If the concept of the boundary wave is valid, we should observe a cylindrical wave originating from the structure edge. This is approximately true, but the wave is more asymmetrically in TE- than in TM-polarization, in particular for the 300nm step. There the reflected boundary diffraction wave is very weak in TM- but quite strong in TE-polarization. To these effects we attribute the observation that the assumption of a geometrical-optics structure reflection provides better agreement with measured images in TM than in TE polarization.

4.2 Waveguiding

As was demonstrated in section 3.2, sub-wavelength bars and trenches of identical width and height resp. thickness are imaged rather differently. These effects may be understood quantitatively by investigation of the corresponding structure nearfields. A qualitative explanation relies on waveguiding effects. In this section we present preliminary results of an investigation of this topic.

The main idea is to consider the sub-wavelength structures as short waveguides situated on (or in) the substrate. Due to its upper and lower boundaries such a structure forms a resonator that is excited by the incident wave. As is well known the excitation is polarization dependent. In particular for trenches of width w in a perfect conductor there is for excitation in TE-polarization (electric field parallel to the edges) a cut-off (free-space) wavelength of $\lambda=2w$ because the tangential electric field has to vanish at the walls.

For a perpendicular electric field (TM-Polarization) there is no cut-off frequency because the electric field component is oriented normal to the structure edges. Hence, the lowest order mode may penetrate into the groove for TM polarization while it is evanescent and therefore strongly attenuated for TE-polarization. Such a model was already used qualitatively by Marx and Psaltis for explanation of the polarization interferometrical detection of DVD pits.[16]

The same argument applies approximately for a strong dielectricum as is demonstrated in Fig. 11. There, the amplitude of electric nearfield for TM-polarization (left) and TE-polarization (right) is shown for a perpendicularly incident plane wave and a trench of 0.16λ width and 0.4λ depth in Si (λ=500 nm, n=4.18+i*0.3). The different penetration depth is clearly visible. However, for a weaker dielectrica like glass the effect is considerably weaker.

For bars on a substrate the argument may be applied in a modified form. Now the structure is considered as a short film-waveguide consisting of a high-index film in a low-index medium. It is excited by the incident wave. Responsible for the polarization effects are now their different mode-wavelengths. These are easily obtained from standard textbooks[17] for a film of thickness d and refractive index n_f in a medium with refractive index n_0 by solution of

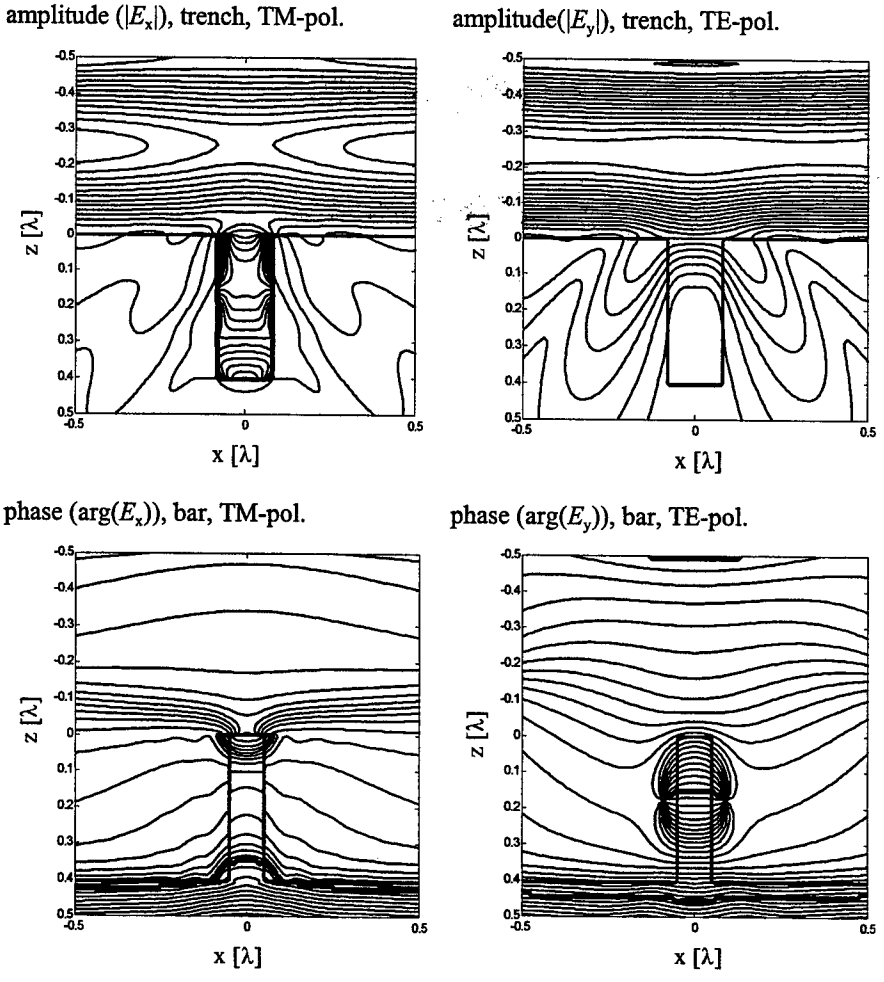


Fig. 11 Field penetration into a trench in Si (top, amplitude isolines), phase of Si-bar on Si (bottom, phase-isolines).

$$\tan(u - m\pi) = \frac{2uv}{q^2u^2 - q^{-2}v^2} \qquad (16)$$

with the mode-number m, $u = \left((n_f k_0^2) - \beta^2\right)^{1/2} d$,

$v = \left(\beta^2 - (n_0 k_0)^2\right)^{1/2} d$, $k_0 = 2\pi\lambda^{-1}$, $q = n_0/n_f$ and the mode-wavelength $\lambda_M = 2\pi/\beta$. The mode-diagram is plotted in Fig. 12. The thin line shows the relative mode-wavelength for TE-polarization and the thick-line for TM-polarization.

For a given mode, the mode-wavelength decreases faster with increasing width for TE polarization, providing a comparatively stronger phase-shift.

Summing up: the waveguiding effects are responsible for the observed opposite behavior of sub-λ bars and trenches in imaging. Bars provide a stronger phase-shift in TE-polarization because of the lower mode-wavelength. Trenches yield a stronger phaseshift for TM-polarization because of the higher penetration depth.

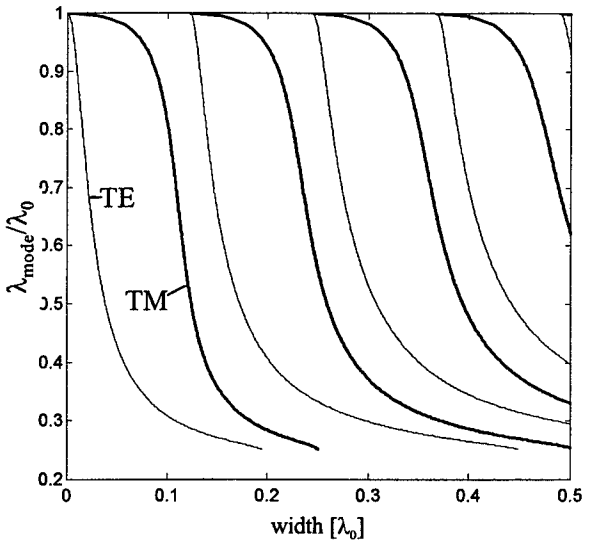


Fig. 12 Mode-diagram for Si film-waveguides at λ=500nm

5. POLARIZATION INTERFEROMETRY

According to the preceding sections, the relative phase-shift between the TE and TM polarized components is an important polarization effect in particular for sub-wavelength structures. It can be measured directly by polarization-interferometry , i.e. instead of superposing the fields reflected by the object and an external reference as in conventional interference-microscopy, we interfere the TE- and the TM-polarized image of the same structure. Phase and contrast are measured using temporal phase shifting. Of course, we loose the height information but, as we will show, we gain an increased edge sensitivity. Furthermore, the method is less sensitive to vibrations because it constitutes basically a common-path interferometer.

The experimental set-up for polarization interferometry is based on a Leica DMR microscope (Fig. 13) with an attached laser-illumination and a detection unit mounted on the photo-tubus. The non-polarizing components form a reflection-type microscope with a two-step magnifying system. The illumination consists of a laser-diode of wavelength λ=635nm (Sanyo DL3038023, 3mW) that is collimated and focused with f/D=4 onto a rotating ground-glass. The laser-diode is driven near its threshold to reduce the temporal coherence. The spatial coherence is adjustable via the diameter of the illuminated spot on the ground glass. The ground glass is imaged into the aperture-stop of the Köhler illumination system of the microscope. Due to experimental constraints it was necessary to insert the polarizer λ/4 plate combination between laser-diode and focus lens. This somewhat unusual configuration worked well, most likely because we applied a small roughness ground-glass with an NA of approximately 0.5. To estimate the depolarization we measured the contrast of polarization-interferograms without and with the ground-glass. The contrast differed by less than 1%.

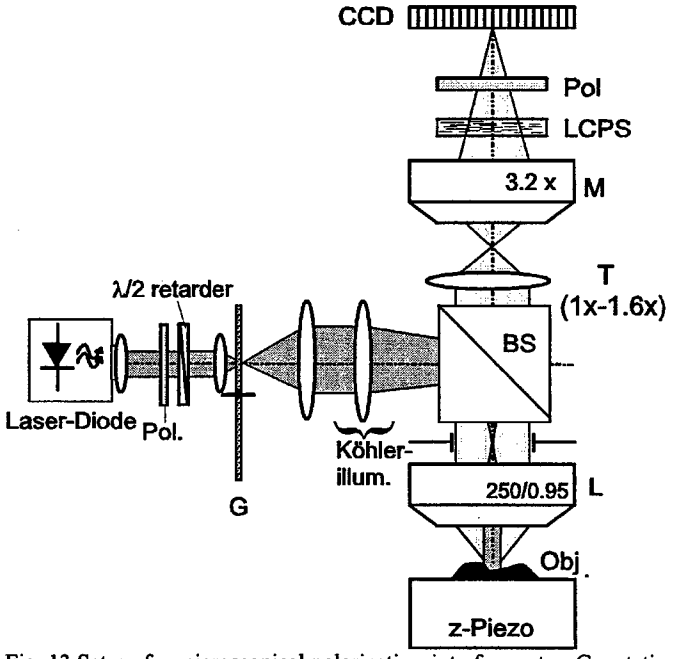


Fig. 13 Set-up for microscopical polarization-interferometry. G=rotating ground.glass, L= microscope lens, T=tube-lens, M additional magnification, BS: non-polarizing beam-splitter, LCPS: liquid crystal phase-shifter, Pol: polarizer

The image of the structure due to a 250/0.95 lens and tube lens (magnification: M_T=1.0, 1.25, 1.6) is projected by a 3.2× magnifying lens *L* (NA=0.15) onto a CCD-Chip (Hitachi KP160). The overall magnification amounts at M_T=1.25 to 1000, i.e. a single 9.2×8.4 μm^2 pixel on the CCD corresponds to a 9.2nm × 8.4nm pixel in the object space. The high magnification is necessary to resolve the observed steep edge images. Additonally it allows a strong noise-reduction (c.f. section 2.1) The focus-position can be adjusted with nm-accuracy by a capacitive-sensor controlled piezo translation-stage (Physik Instrumente).

The electrooptic compensator consists of a liquid crystal phase-shifter (LCPS), i.e. a planar nematic liquid-crystal cell between two transparent ITO electrodes (Jenoptik). The maximum phase-shift is ≈6π at λ=635nm. The TE and TM field-components interfere at the polarizer behind the LCPS. The resulting interferogram on the CCD chip is transferred via a 10bit frame-grabber to the computer. Image acquisition and interferogram evaluation are identical to the above described conventional interference microscopy, as described in section 2.1.

A measurement example is shown in Fig. 14. The illumination-NA is 0.4. In particular the contrast image indicates distortions rather sensitively, compared with the intensity image.

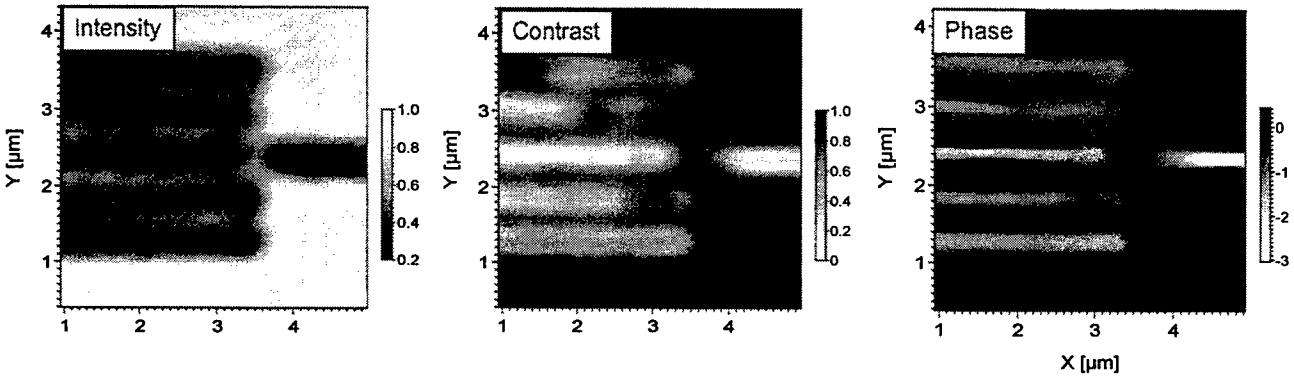


Fig. 14 Intensity (a), contrast(b), and phase-image (c) of 5 trenches of 120nm depth and 600nm pitch in Si.

6. CONCLUSIONS

Interference microscopy at different orthogonal polarizations with temporal phase shifting can be used for a systematical investigation of polarization effects observed at linear microstructures. We have shown that the evaluation provides the diagonal elements of the Jones-matrix. In the absence of non-diagonal elements, subsequent defocus and image-contrast generation is possible according to a „synthetical microscopy". The agreement with rigorous numerical simulations according to the RCWA is satisfactorily. Using near-field computations, we have shown that for steps and narrow trenches in Si, TM-polarization provides a considerably better agreement with geometrical optics than TE-polarization. Polarization interferometry measures directly the phase-shift between the TE- and TM-polarized fields. By this means it provides a sensitive method for linewidth-measurements of topographical microstructures.

REFERENCES

1. D.Nyyssonen, "Linewidth measurement with an optical microscope: the effect of operating conditions on the image profile", *Appl. Opt.* **16**, pp. 2223-2230 (1977).
2. K.-P. Schröder, W. Mirandé, H. Geuther, C. Herrmann, "In quest of nm accuracy: supporting optical metrology by rigorous diffraction theory and AFM topography", *Opt. Commun.* **115**, pp. 568-576 (1995).
3. P. Lalanne, P. Pichon, E. Chavel, E. Cambril, H. Launois, "Interferometrical characterization of subwavelength lamellar gratings", *Appl. Opt.* **38**, pp. 4980-4984 (1999).
4. R. Barakat, "Optical linewidth measurements using a polarized microscope with crossed polarizers", *Appl. Opt.* **29**, pp. 5038-5039 (1990).
5. S. Kimura, , T. Wilson, "Confocal scanning dark-field polarization microscopy", *Appl. Opt.* **33**, pp. 1274-1278 (1994).

6. M. Totzeck, H.J. Tiziani, "Phase-shifting polarization-interferometry for microstructure linewidth measurements", *Opt. Lett.* **24**, 294-296 (1999).
7. J. Schmit, K. Creath, "Window function influence on phase error in phase-shifting algorithms", *Appl. Opt.* **35**, pp. 5642-5649 (1996).
8. K. Creath, "Calibration of numerical aperture effects in interferometric microscopy objectives", *Appl. Opt.* **28**, pp. 3333-3338 (1989).
9. E.D.Palik, *Handbook of Optical Constants of Solids*, Academic Press 1985, 547ff.
10. M.G. Moharam, E.B. Grann, D.A. Pommet, T.K. Gaylord, "Formulation for stable and efficient implementation of the rigorous coupled-wave analysis of binary gratings", *J. Opt. Soc. Am. A* **12**, pp. 1068-1076 (1995)
11. H. H. Hopkins, "Image Formation with partially coherent light", *Photogr. Sci. and Eng.* **21**, pp.114-123 (1977).
12. M. Totzeck, H.J. Tiziani, "Phase-singularities in 2D diffraction fields and interference-microscopy", *Opt. Commun.* **138**, pp. 365-382 (1997).
13. M. Born, E. Wolf, *Principles of Optics*, Pergamon Press 1980, 6th Ed. pp. 449f.
14. G. A. Deschamps, J. Boersma, S.-W. Lee, "Three-dimensional half-plane diffraction: Exact solution and testing of uniform theories", *IEEE Transact. Ant. & Prop.* **AP-32**, pp. 264-271 (1984).
15. J. Chilwell, I. Hodgkinson, "Thin-films field-transfer matrix theory of planar multilayer waveguides and reflection from prism-loaded waveguides", *J. Opt. Soc. Am. A* **1**, pp. 742-753 (1984)
16. D.S. Marx, D. Psaltis, "Polarization quadrature measurement of subwavelength diffracting structures", *Appl. Opt.* **36**, pp. 6434-6440 (1997).
17. Textbooks on waveguide theory, for instance: H. G. Unger, *Optische Nachrichtentechnik I*, Huethig 1993 (in German).
18. The term „image" is used in a rather general sense here. It is a collection of spatially dependent data f(x,y) derived from a microscopical measurement. Without prefix it denotes the complex, electric field component in the image plane. Particular field quantities or combinations of them are characterized by the appropriate prefix: Phase-image, itensity-image, contrast-image etc..
19. T.R. Corle, G.S. Kino, *Confocal Scanning Optical Microscopy* (Academic, San Diego, 1996) pp. 286f.

High-precision measurement scheme for half-wave voltage of Y-tap MIOC

Yuanhong Yang Weixu Zhang Jing Ma
The 5th Research Division, Beijing University of Aeronautics & Astronautics
PO BOX 85, 100083, Beijing, China

ABSTRACT

A high-precision measure scheme for half-wave voltage V_π of Y-tap MIOC is proposed. This scheme is based on Sagnac interferometer and a saw-tooth wave with special period is used. With this scheme, the temperature characteristic of V_π of a type of MIOC is studied experimentally. The result shows that its half-wave voltage is temperature-dependent. In the range of -10 to +55 , the varying value is 0.162V, the temperature coefficient is 662 PPM/°C and the variation is linear and repetitive. Experimental study and detail discussion demonstrate that high accuracy can be achieved with this scheme and it is very suitable to be used in studying and calibrating V_π of the modulator used in Close-loop FOG in-site.

Key words: half-wave voltage, temperature coefficient, integrated optical circuit; fiber-optic gyroscope

1 INTRODUCTION

Y-tape MIOC (multifunction integrated optical circuit, or Y-tape multifunction integrated electro-optical phase modulator) is one of the key components of close-loop fiber optic gyroscope (FOG). It take the roles of mode filter (polarizer), splitting and combining light (directional coupler) and phase modulator in FOG optical path. The basic configuration is shown in Fig1, where wave-guide is fabricated in $LiNbO_3$ with annealing proton exchanged (APE) technology, while electrode with coating

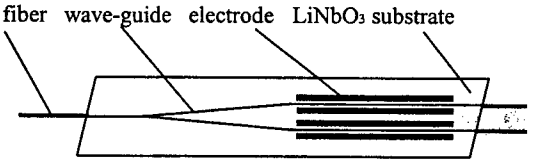


Fig 1 Configuration of Y-tap modulator used in FOG

and light input/output with polarization-maintain(PM) optical fiber. For an electro-optic component, V_π is a very important parameter which variety will bring nonlinearity to the output of close-loop FOG directly. So, It is very significant to measure it and define its variation model accurately. There are a few measure schemes[1][2] reported now. But they are all based on polarization interference or Mach-Zehnder interferometer and can not achieve high accuracy. To satisfy the demand of high-precision measuring and modeling, a scheme based on Sagnac interferometer is proposed. The experimental study and detail discussion was made meanwhile.

2 PRINCIPLE

Figure 2 is the measurement scheme diagram for V_π of MIOC, where SLD is superluminescent diode; DC is directional coupler. As shown in figure 2, SLD, DC, detector, fiber coil , and the MIOC tested compose a Sagnac interferometer. When the interferometer is working, the output of detector will be:

$$v_o = K(1+\cos\phi) \qquad 1$$

where, K is the coefficient related with incident optical power on detector and gain of detector unit. φ is nonreciprocal phase shift and it is zero without nonreciprocal modulation or noise. The variation of V_o due to φ is shown as wave-form (a) in figure 3. When a saw-tooth wave-form with amplitude Vs and period 4τ(τis the time demanded for light transmit across the

Part of the SPIE Conference on Advanced Photonic Sensors and Applications
Singapore • November–December 1999
SPIE Vol. 3897 • 0277-786X/99/$10.00

fiber coil) is added to one of the two modulation arms, the nonreciprocal phase shift will be occur because the modulation arm is at nonreciprocal position and the clockwise and counter clockwise light go through the arm and be modulated at different time due to the time delay τ induced by fiber coil delay line. This can be equivalent to the phase modulation wave-form of counter propagating lights are shifted in time axis with each other. The time difference shifted is τ and the nonreciprocal phase can be calculated using following equation[3]:

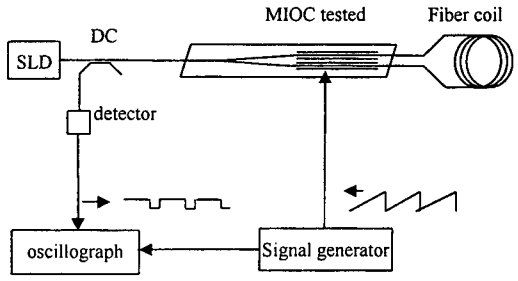


Fig2 Measurement scheme diagram for V_π

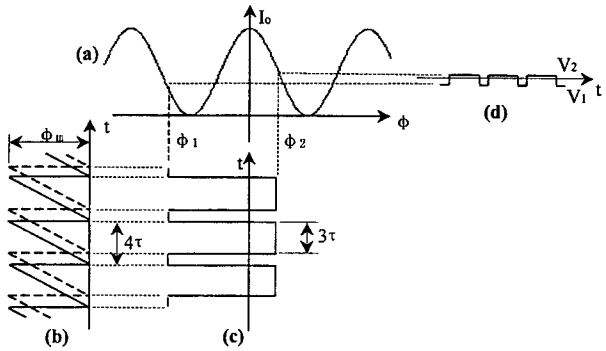


Fig3 Wave-form during measuring

$$\phi = \phi_m(t) - \phi_m(t-\tau) \qquad 2$$

where, $\phi_m(t)$ is the modulation phase wave, which is saw-tooth wave here. In figure 3 waveform (b), the solid lines represent $\phi_m(t)$, while the dash lines represent $\phi_m(t-\tau)$. Because the half-wave voltage to be measured is V_π, the modulation coefficient Kp and the slope Ks of saw-tooth wave added are respectively:

$$K_p = \pi / V_\pi \qquad (3)$$

$$K_S = V_S / 4\tau \qquad (4)$$

During a modulation period, when 0<t<τ, the nonreciprocal phase is:

$$\phi_1 = -K_p \cdot K_S \cdot 3\tau = -\frac{3\pi V_s}{4V_\pi} \qquad (5)$$

the detector output is

$$V_1 = K[1 + \cos(\phi_1)] \qquad (6)$$

when τ<t<4τ the nonreciprocal phase is:

$$\phi_2 = K_p \cdot K_S \cdot \tau = \frac{\pi V_s}{4V_\pi} \qquad (7)$$

the detector output is

$$V_2 = K[1 + \cos(\phi_2)] \qquad (8)$$

This result can be seen in figure 3 waveform (b) (c)and (d) directly. In order to make equation $V_1 = V_2$ satisfied, combining equation (5)~(8), the first non-zero solution of V_π can be got:

$$V_\pi = V_S / 2 \qquad 9$$

This means, when $V_1 = V_2$, the value of V_π can be calculated with equation (9) because Vs is set artificially and known. So, the measurement procedure can be summarized as:

(1) Mounting the testing setup according to figure 2;
(2) Set the period of saw-tooth as 4τ;
(3) Increasing the voltage Vs from low to high and monitor the variation of V_1 and V_2;
(4) When $V_1 = V_2$, the V_π is half of Vs.

When $V_1 = V_2$, $\phi_1 = -3\pi/2$, $\phi_2 = \pi/2$, V_1, V_2 is very sensitive with the variation of Vs and their difference is:

$$\Delta V = V_2 - V_1 = 2K\sin(\frac{\pi V_S}{4V_\pi})\sin(\frac{\pi V_S}{2V_\pi}) \qquad 10$$

The value of ΔV is used to determine whether equation (9) is satisfied. when $V_s \approx 2V_\pi$, assume $V_S + \Delta V_S = 2V_\pi$, where ΔV_S is error item. because $\Delta V_S / 2V_\pi$ is very small, the equation $\sin(\pi/2 - \Delta V_S / 4V_\pi) = 1$ is satisfied approximately. equation (10) can be:

$$\Delta V = 2K\sin(\frac{\Delta V_S}{2V_\pi}) \qquad 11$$

This equation shows that the value of ΔV_S is determined by the detection accuracy of ΔV. Because the gain is very high, ΔV is very sensitive with ΔV_S too. In fact, this situation is very similar to the signal processing in close-loop FOG[3] and the accuracy of signal detection in close-loop FOG can be used to estimate the accuracy of ΔV_S. In close-loop FOG, $1\,°/h$ accuracy can be achieved easily and it corresponds to 10^{-6} rad phase detection accuracy approximately. For equation (11), this means $\Delta V_S / (2V_\pi) = 10^{-6}$. V_π is only a few voltages ΔV_S will be much smaller. This demonstrates that this scheme have the potential of achieving high accuracy.

3 EXPERIMENT

A type of MIOC has been measured with this scheme. The experimental setup is shown in figure 2, where the length of fiber coil is 1Km, the type of oscillograph and signal generator is TDS220 and HP33120A respectively. They are all digital,

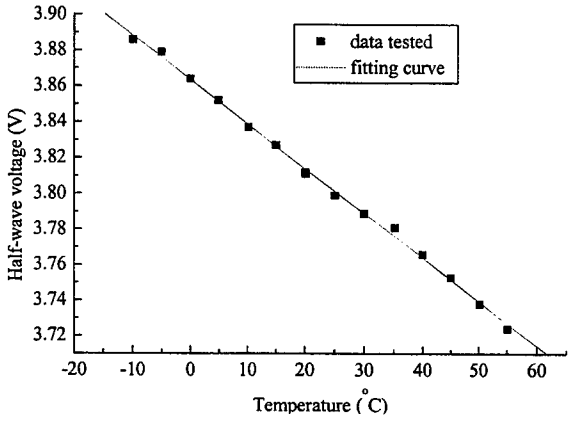


Fig4 curve of half-wave voltage vs temperature

TDS220 is used to monitor the varying of V_1, V_2 ; while HP33120A for saw-tooth wave generating and giving the value of Vs, because its read-out accuracy is 1mV, The system measurement accuracy is 1mV too with this setup. In the experimental procedure, the MIOC is only mounted in a temperature control box, V_π is measured at different temperature point, the result is shown in figure 4. This show V_π is temperature dependent. The varying value is 0.162V in the range of -10 to +55 and the temperature coefficient is 662 PPM/ . The result corresponds with the report in literature [4]. After least square polynomial fit with the data tested, the variation is found linear and the result of many times experiments prove it repetitive too.

4 DISCUSSION

(1)From equation (5)~(9) and (11), we know it is the 2π phase difference between ϕ_1 and ϕ_2 that make equation (9) satisfied, meanwhile $V_1 = V_2$, $\Delta V = 0$ and all of above occur in a modulation period - about few ten μs. The measuring can be thought as instantaneous. So, any nonreciprocal phase shift induced by other factor, such as environmental temperature

varying, fiber coil sensing rotation, etc., will not affect the measuring because these sort of varying is low frequency, the variation of ϕ_1 and ϕ_2 will be equal during a modulation period. With the same reason, the measuring is insensitive with the small variation of K induced by optical path loss change or source power change in equation (1). This is why high precision can be achieved in above experimental setup, where the MIOC tested was set only in temperature change environment.

(2)As discussion above, when the period T of saw-tooth wave is 4τ, the measuring will achieve high accuracy. But this is not necessary for all measure, the period can be set as other value too. Assuming T>τ, we have:

$$K_S = V_s / T$$

$$\phi_1 = -K_P \cdot K_S \cdot (T-\tau) = -\frac{\pi V_S}{V_\pi T}(T-\tau) \tag{12}$$

$$\phi_2 = K_P \cdot K_S \cdot \tau = \frac{\pi V_S}{V_\pi T}\tau \tag{13}$$

$$\phi_2 - \phi_1 = \frac{\pi V_S}{V_\pi} \tag{14}$$

when $V_S = 2V_\pi$, the phase difference of ϕ_1 and ϕ_2 is still 2π and $V_1 = V_2$. To the contrary, we can still use the voltage difference ΔV to determine whether the equation (9) is satisfied and calculate V_π with equation (9). But at $T = \tau / n$ (n=1,2,3...), this scheme can not operate because the phase modulation wave is superposition and the phase difference of ϕ_1 and ϕ_2 is always zero. When T<τ and $T \neq \tau / n$, the same resulr as equation (14) can be obtained with the same derivation procedure too.

(3)In this scheme, the voltage difference between V_1 and V_2 is used to determine whether equation (9) is satisfied. Further, it can be used as feedback error signal to control the amplitude of saw-tooth wave through a set of special electric circuit completely, and then the auto measuring can be realized. The circuit is close-loop and can be illustrate as figure 5, ΔV is feedback signal used to control the amplitude of saw-tooth modulation signal and make $\phi_2 - \phi_1 = 2\pi$ and equation (9) satisfied. For comparing, draw the signal circuit loop of close-loop FOG in figure 6, where another square wave signal is added to the modulator, V^+ and V^- is the output correspond to the positive and negative half of a period. Their difference $\Delta V'$ is feedback too to control the frequency of saw-tooth modulation signal to make the total phase $\phi_T = \phi_S - \phi_{fb}$ keep zero, where ϕ_s is the sagnac phase duo to rotation, ϕ_{fb} is the feedback phase. Comparing the two loops, we can find they are very similar and the hardware will be same when this loop is realized with digital circuit. This demonstrate that the scheme can be taken to realize the self-checking and calibrating on V_π in close-loop FOG easily.

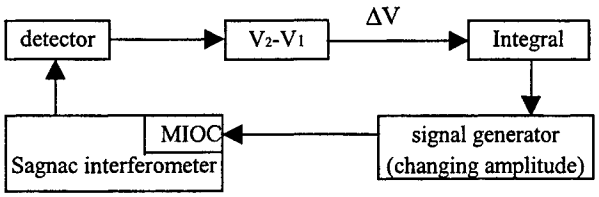


Fig 5 measuring loop for V_π

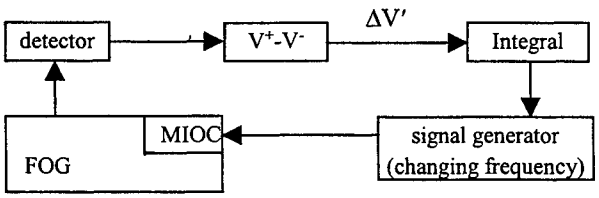


Fig 6 signal proceeding loop of FOG

(4)In figure 2, the MIOC tested take the role of splitting and combining the light of optical path too. If use another directional coupler to take place the function, we can measure V_π of other sort optical phase modulator (PM), such as single channel integrated electro-optical phase modulator and bulk phase modulator etc. The setup is shown in figure 7, where PM is phase modulator.

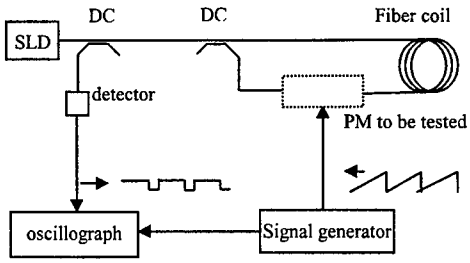


Fig 7 Measurement Scheme diagram for V_π of other type of phase modulator

5 CONCLUSION

Experimental study and detail discussion demonstrate that the scheme proposed can measure V_π of Y-tap MIOC and other sort of phase modulator conveniently and accurately. The highest accuracy can be achieved with period 4τ and other period value except for τ / n can be used to do the measurement too. This scheme is very suitable to be used in studying and calibrating V_π of MIOC used in Close-loop FOG in-site. Temperature experiment on V_π of MIOC tested demonstrates V_π is temperature-dependent and the variation is linear and repetitive.

REFERENCES

[1] Geng Fan, Shong Qiang, Li Caifen. Half-wave voltage measure with polarization interference method. Proceeding on FOG , (In Chinese) 1990.
[2] Yang Yuanhong, Ma Jing, Zhang Wei-xu. The PZT optical fiber phase modulation testing with varying amplitude modulation. CHINESE JOURNAL OF LASER, Vol.B3, No.6, Dec.1994.
[3] Herve Lefevre. The Fiber-Optic Gyroscope. Boston·London, Artech House, 1993, 33;111-126.
[4] George A. Pavlath. Close-loop fiber optic gyros. Proc. Fiber Optic Gyroscope: 20th Anniversary Conference. SPIE Vol. 2837, P46-60.

SESSION 12

Speckle Sensors

Complex monitoring of monumental structures located in seismic and mining-influenced areas with the use of laser tilt sensors and acoustical holography of the ground

Adam Lipowczan, Zbigniew Motyka, Henryk Passia, Adam Szade

Central Mining Institute, Katowice, Poland

ABSTRACT

The necessity to protect cultural heritage is indisputable. The paper presents two important metrological problems connected with the assessment of the condition and behaviour of two monumental structures located in mining-influenced area in Upper Silesia region (south of Poland). The first problem has been effectively solved through using laser tilt and vibration sensors designed at the Central Mining Institute's Laser laboratory, while the other is to be investigated by using acoustical holography to detect the structure of the castle's foundation and its surroundings in the ground. The paper presents the experimental results obtained with the use of both methods. An analysis of the tilt sensor at the castle show special effects of its behaviour which are to be related to other factors outside mining. These factors need also further detailed investigation. The method of acoustical holography has been experimentally verified in detecting the model structure of underground workings.

Keywords: Laser tilt sensor, Monitoring, Monumental structures, Seismic holography

1. LASER TILT SENSOR – DESCRIPTION OF THE OPERATION PRINCIPLE AND METHOD OF MEASUREMENT

The operational principle of the CMI-developed laser tilt sensor and measurement versions have been sufficiently widely described in the authors' publications and patents [1]-[4].

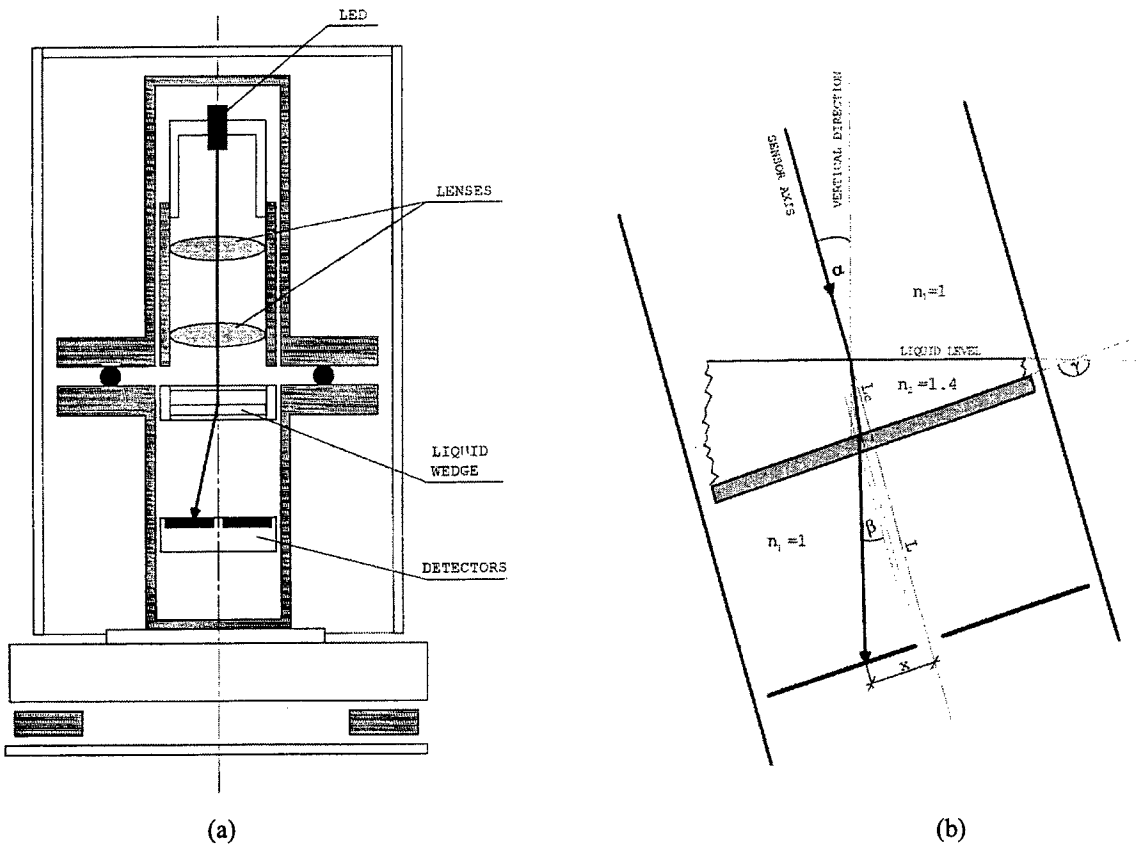
1.1 Construction and technical characteristics of the sensor

The general structure of the sensor with a liquid wedge is illustrated in Fig.1a, while the general principle in Fig.1b.

The laser tilt sensor for continuous recording of the behaviour of structures, as shown in Figure 3, is composed of the light transmitter section, glass cell containing the liquid and signal detection circuit cooperating with the amplifier, signal transmission, power supply and data acquisition system. The transmitter assembly, cell and detection circuit are firmly fastened to the analysed object through levelling head and base. The transmitter assembly is made with the use either of a diode laser or LED provided with the light collimating system. All are contained in one casing.
The beam detection contains a quadrant detector attached to a separate holder and is positioned at a specified distance under the cell. The cell contains the liquid with precisely known both the index of refraction and decrement of vibration damping.

The measurement parameters and technical characteristics of the sensor can be collected as follows:
- measurement range $\pm$ 5 mm (with the possibility to be adjusted from 2 to 50 mm/m),
- resolution $\pm$ 0.02 mm/m,
- vibration frequency to be measured < 10 Hz,
- power consumption <2W,
- galvanic separation between the transmitter and receiver circuits (only in the intrinsically-safe version of the device) – 4.5 kV,
- autonomous transmitting-supplying line (intrinsically-safe version) – L <1500m,
- weight of the measurement-transmission assembly – 2 kg (intrinsically-safe version – 10 kg).

Part of the SPIE Conference on Advanced Photonic Sensors and Applications
Singapore • November–December 1999
SPIE Vol. 3897 • 0277-786X/99/$10.00



(a) (b)

Figure 1. Schematic of laser sensor with liquid wedge (a) and operation principle (b).

(a) (b)

Figure 2. View of the Będzin castle (a) and the laser sensor at work (b).

As it has been already mentioned above, the sensor is being made in two versions: normal and intrinsically-safe.
In the normal version, the system is composed of the laser tilt sensor, measuring amplifier and analog or digital recorder.
In the methane explosion-hazard zones, the intrinsically-safe version of the instrument is used (approved by Polish High Mines Inspectorate, Nº GX-179/97).
The sensor elements are contained in dust-and-waterproof casing (type IP54).

1.2 Measurement method

The structures (objects) located within mining-influenced zones are liable to various deformations differing in their kind and size. The deflection from the vertical direction, when kept within the limits defined by the standards, is a condition of their safe operation. This deflection can proceed gradually or to be of periodical and dynamic character. Uneven subsidence, shock and vibration can result other causes, apart from mining operations. These are, as an example: hydro-geological, climate and road traffic.

For many endangered objects and engineering structures, continuous measurements are required with the use of continuous recording of amplitudes and directions of the changes.

The measurement result, in a form of a graphic printout provides the full pattern of the changes both in terms of the deflection values and their directions, and the time of their occurrence. As demonstrated in the enclosed Figures, the operation principle is based on determining the deflection of the laser beam from its initial vertical position after if has passed through the liquid wedge. There are no movable parts in the instrument.

The layer of the liquid is an absolute reference system in the measurements. The displacement of the beam spot on the photodetector is a function of the inclination angle assumed by the liquid level when the instrument (i.e. the controlled structure) is tilted, and index of refraction of the liquid.

The electrical signal is, in turn, a function of the illuminated quadrant areas of the detector. After being amplified, it is fed into a PC through A/D card for processing.

An analysis of the measurement data obtained by using the sensor system will allow the following:
- identification of factors influencing the stability of the object,
- decision taking about applying building protection measures or others.

1.3 Examples of recording of the effects at the monumental structures close to Upper Silesia Coal Basin (Poland)

The examples below demonstrate applications of the laser tilt and vibration sensor in two measuring situations relating to the monumental structures located in the mining-influenced areas, namely the medieval castle of Będzin (14-th century) and church (19-th century) in the town of Rydułtowy, both Upper Silesia region, south Poland. The structures themselves are shown in Figs.2a and 4a, respectively. The first one is influenced mainly by post-mining occurrences, the other by currently conducted mining right under the structure. The subsequent figures demonstrate various effects observed at the castle. Figures 5 presents a sudden change of the tilt resulting from ground movement caused by water from heavy rains flowing into the voids under the castle, created by previous mining operations (dewatering). As mentioned above, the second measuring situation refers to monitoring of the church (late 19-th century) strongly influenced by currently conducted mining operations right under the structure. Figure 5 shows a very strong momentary change of the tilt caused by an underground rockburst (break of the rigid sandstone plate above the working). Worth mentioning is the occurrence of the pair of foreshocks recorded several hours earlier.

When monitoring this church, the effect of the strong earthquake in Turkey was recorded as a distinct reaction of the structure to seismic waves coming from there.

Weather conditions, especially temperature differences (e.g. by illumination of one part of the structure) can cause remarkable changes of the tilt. Such an example is also presented in Fig.5, presenting the effect of the sun beginning 11 AM.

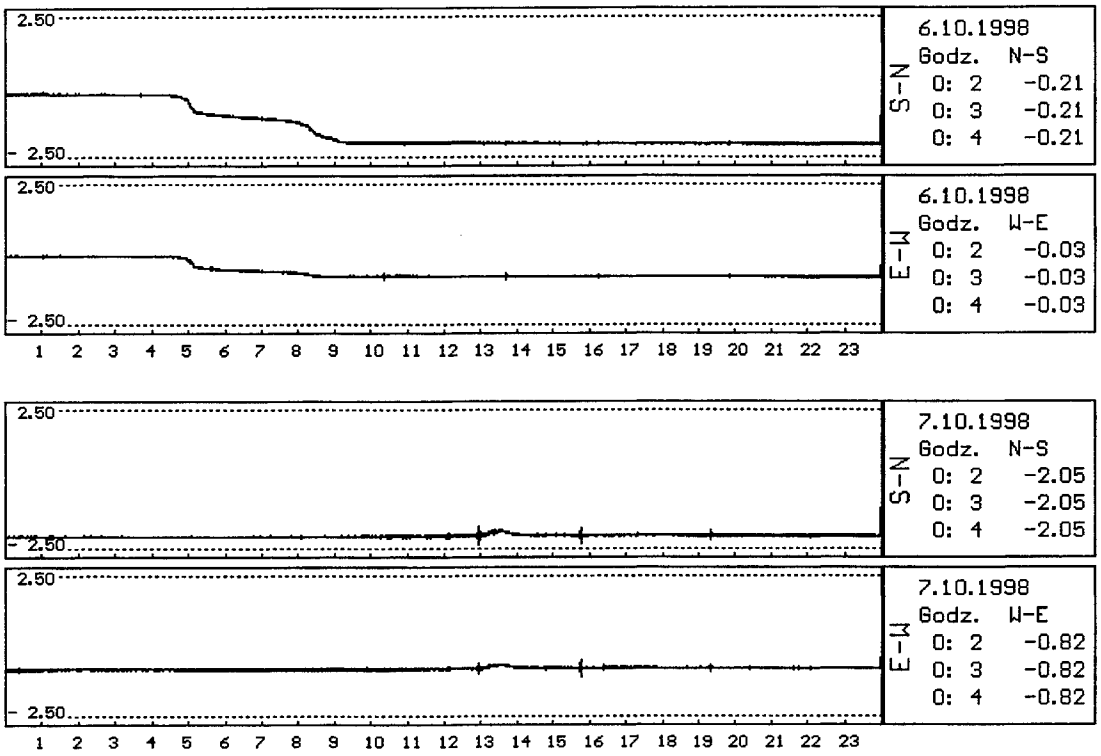
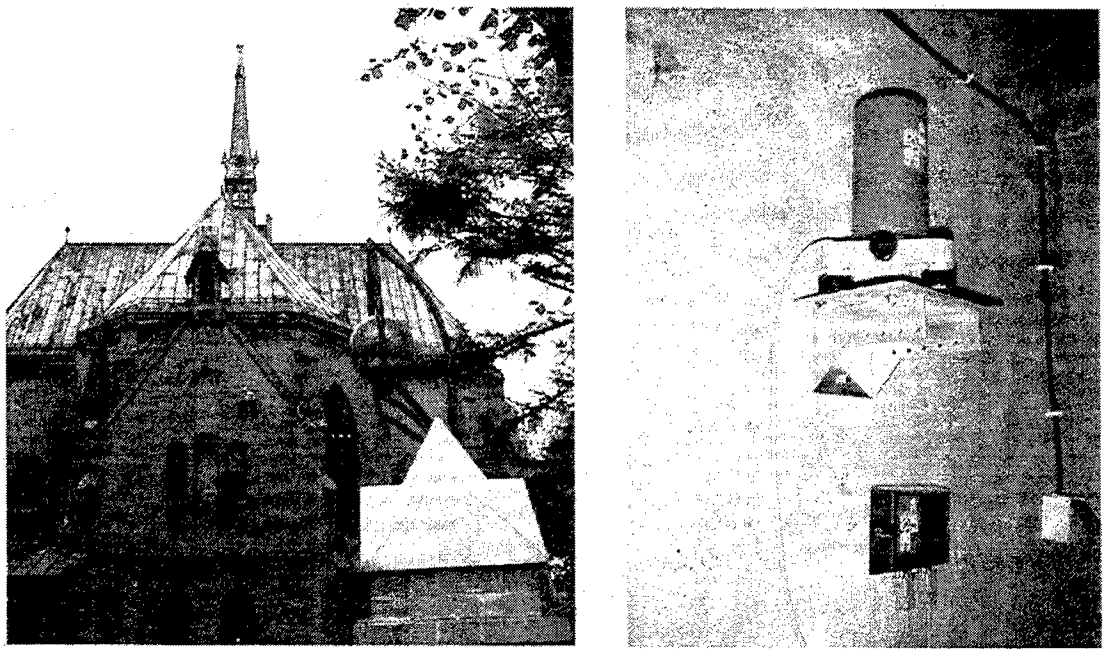


Figure 3. Sudden change of tilt due to ground movement caused by water.

One of the conclusions that can be drawn on the basis of the above discussed examples is the necessity to have reliable information about the quality of the ground under the monitored object. The effects observed particularly at the castle strongly support this idea. Finally, this will allow to undertake suitable protective actions and measures. Therefore, the next step is to apply a method being able to give effective assessments of the conditions in the ground under the structure under consideration.

The authors' proposal is to use the seismic holography verified experimentally in detecting underground structures (mine workings)



(a) (b)

Figure 4. View of the church of Rydułtowy (a) and the laser sensor at work (b).

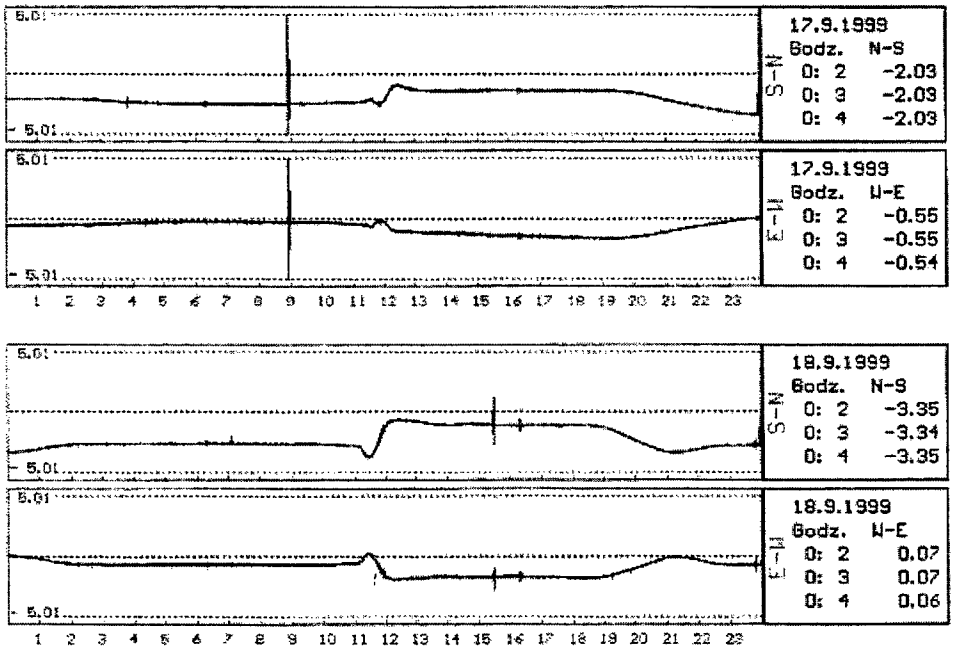


Figure 5. Strong change of the tilt caused by underground rockburst.

2. ACOUSTICAL HOLOGRAPHY OF THE GROUND

2.1 Acoustical holography in 3d seismic visualisation

The 3D seismic prospecting methods, and in particular seismic holography are found to provide the most reliable complex 3D data related to the structure of the in methods basically provide the most reliable complex 3D data related to the structure of the investigated underground space [5]-[9].

The method of acoustical holography transferred to the range of seismic elastic waves, can be seen as potentially very promising technique of seismic prospecting and visualisation of underground objects. On the basis of simple recording of seismic wavefront originated at one source-point only, it enables to obtain a hologram of distinct underground structures, if such are present between the source and the set of the geophones (accelerometers) placed at the lattice points on the ground surface, and - in the next step - to reconstruct the picture of spatial distribution of those structures on the basis of such a hologram. In principle, it is possible to obtain such an image in the case of surface location of the source (the reflective holography) as well as in the presented here case of transmission (refractive) holography. An example of this is given in Fig.6, with the task being to detect the underground structure of mine headings.

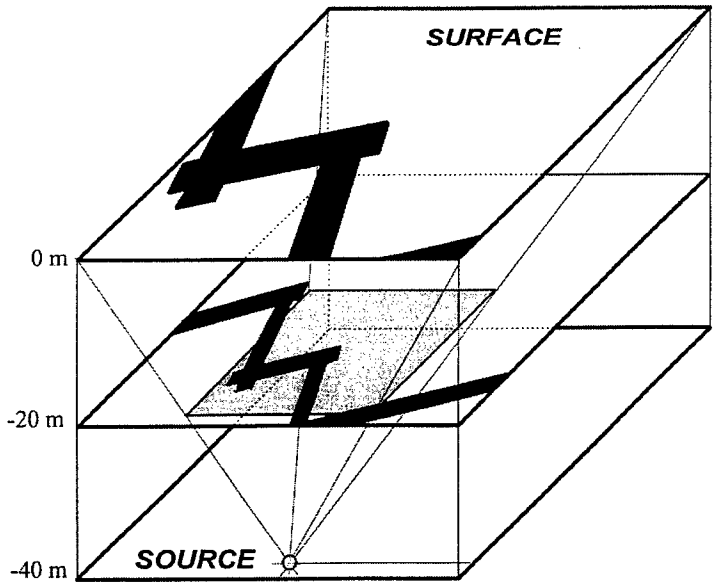


Figure 6.
An example of seismic prospecting for the needs of the transmission seismic holography method. Small charges were detonated in the roof of a mine heading, ca. 40 m below the surface at the central point of the edge of the prospected area. The seismic data were taken on the surface by 21 x 21 lattice points of accelerometer posts, forming the 50m x 50m square.

The wave field was recorded at the hologram aperture represented by N x M lattice points on the surface of the earth. (Satisfactory results may be obtained for N=M=20 or more). It is reasonable to restrict considerations to phase-only holograms which usually gives better results then the amplitude ones. Phase holograms utilises only the phase information, which can be easily obtained from the time-of-flight value, for any wave frequency from the range of recorded broad-band signal under the assumption that the signal has not been dispersed significantly.

2.2 The set up for field measurements

The authors' seismic measuring setup for field measurements consists of :

- 18 seismic transducers PS-1PI with built-in preamplifiers and I/U transducers (Fig.7)
- the packet of 18 modules of receivers of seismic signals transmission
- 24-channel recorder V-Store 24 Racal Recorders (Fig.8)

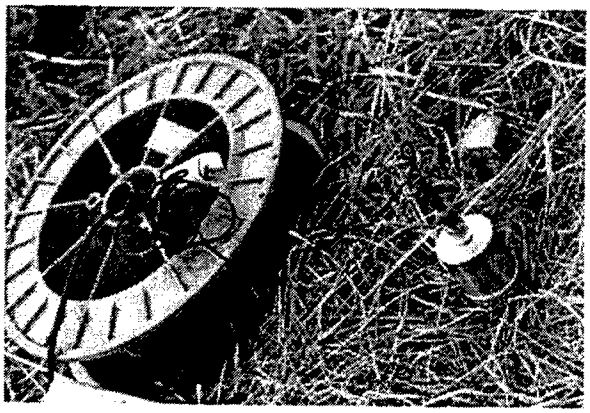


Figure 7. One of the 18 CMI-developed PS-1PI accelerometers during surface signal recording for the needs of seismic holography.

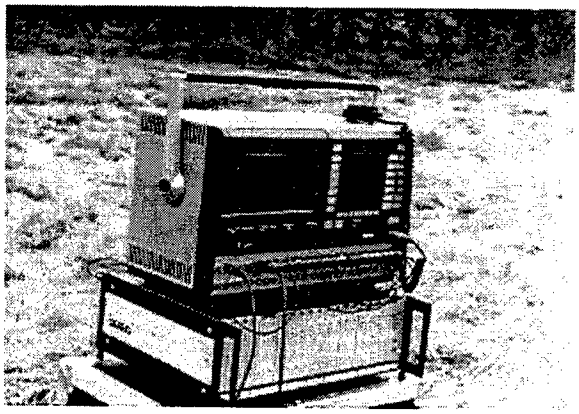


Figure 8. The recording system arrangement during seismic measurements for the needs of seismic holography.

To obtain the full set of data needed when using seismic prospecting data, a series of blasts with adequate explosive charges must be performed.

The measurement area should be limited by the square of side length of 50 -200 m over conveniently chosen blasting-point below the object. The surface area over the blasting-point should be, most conveniently, flat, and clean.

2.3 Numerical image reconstruction of seismic holograms

For reconstruction of the picture of a seismic object the numerical reconstruction on the base of Fraunhofer approximation of the Kirchhoff integral is used. Basically, it allows to obtain and reconstruct monofrequency holograms. On the other hand their sum allows to obtain multifrequency picture bringing up more details of an object then any of its components. Fig.9 presents the result of such a multifrequency reconstruction.

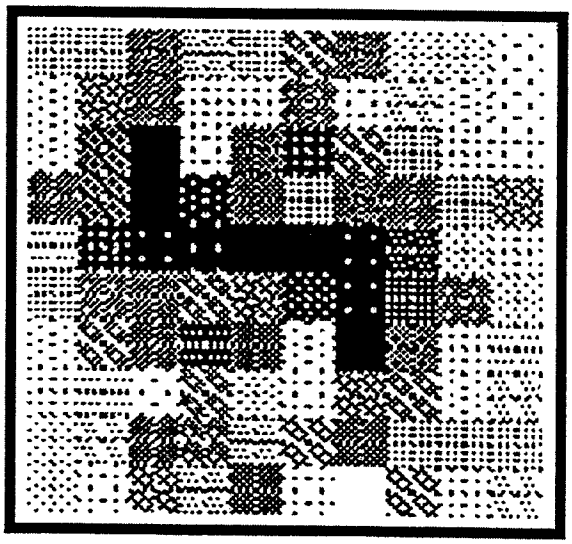


Figure 9.
This image is centered on the picture of the underground area of "Barbara" Experimental Mine in Upper Silesia region of Poland . It represents some part of underground headings at the depth of ca.20 m below the surface (schematically depicted on Fig.11). They were processed to the form of seismic phase-only hologram and reconstructed in simplified manner with the use of 2D Fourier transform. The obtained clear Z-shaped structure follows the true structure of the underground mine headings.

To summarise briefly this subject, the example presented above is a sufficient recommendation of the method of seismic holography to be used in detecting voids and disturbances of the ground under the discussed castle, and these measurements are expected to start in the nearest future.
Application of the method to other similar situations seems straightforward.

2. Conclusions

On the basis of the above presented material, the following conclucions can be drawn:

1. Underground mining and post-mining activities manifest themselves as various effects on the structures located on the surface.
2. The authors routinely conduct measurements of low-frequency vibration and tilt of the structures located on mining-influenced areas such as dwelling houses, industrial objects and others including in particular monumental structures.
3. Both short-duration ("shocks") and long-term phenomena has been observed both at the medieval castle and 19-th century church located in the Upper Silesia Coal Basin (south Poland). The measuring information obtained there with the use of the CMI-developed laser tilt sensor forces the necessity to supplement these data with a method enabling to examine the ground under the castle. Seismic holography has been choosen as this supplementary method. The extended information provide the basis to apply effective protection measures.
4. The effectiveness of the assumed protection measures should be again analysed by long-term measurements using laser tilt sensors.
5. The presented methodology and instrumentation can be directly transferred to measurements of the structures subjected to effects generated by other similar causes, like seismic ones.

REFERENCES

1. Szade A., Passia H., Lipowczan A., "Laser sensors for continuous control of tilting of buildings on mining-influenced and seismic areas; construction and field experience", *Proc. SPIE*, **2868**, pp.500-509, 1996
2. Polish patent Nº 151105 : Instrument for continuous measurement of tilt of buildings.
3. Szade A., Passia H., Lipowczan A., Bochenek W., "Intrinsically-safe laser-based system for continuous measurement of low-frequency vibration of mine shaft installations", *Proc. SPIE*, **3411**, pp.275-281, 1998
4. Szade A., Passia H., Bochenek W., Pytlarz T., "Continuous, automatic measurement of influence of mining - exploitation on objects on surface area with the use of laser sensors of vibrations and tilting" (Only in Polish) Polish Academy Of Sciences, *Conference Preceedings*, *Workshops' 98*,Ustroń,Polska pp.121-130, 1998
5. Szade A., Passia H., Bochenek W., "Laser sensor for continuous three-axial measurements of low frequency vibrations and tilt of buildings" (Only in Polish), Polish Academy Of Sciences, *Conference Preceedings*, *Workshops' 99*, Jaworze, Polska, pp.193-201,1999.
6. Aloszko S., Motyka Z.., Passia H., "Interferometric Holography of Thin Rods and Coal Beds", Proc. SPIE **859**, pp.242-244, 1987.
7. Aloszko S., Motyka Z.,"Computer simulation of seismic-acoustical hologram of the simulated point inhomogeneities of the rock mass" (Only in Polish), *II Symposium on Laser Techniques, of 7-10 September 1987,* Szczecin, Poland, pp.311-313,1987.
8. Motyka Z., "Visualisation of near-surface underground structures on mining-influenced areas – the method of three-dimensional seismic investigations" (Only in Polish), *Conference Proceedings, Acoustics in Technics, Medicine and Culture, 7-9 of November 1995,* Cracow, Poland, pp.143-148, 1995.
9. Motyka Z., „The Mehod of Seismic Holograhy in 3D Visualisation of Underground Structures" (Only in Polish) Polish Academy of Science, *Conference Preceedings*, *Workshops' 99*, Jaworze, Polska, pp.85-99, 1999.

Shearography as Thermo-Mechanical Sensor for Thermal Stress Determination of Coatings at High Temperatures

K.Habib

Materials Application Department
KISR,P.O.BOX 24885 SAFAT, 13109 Kuwait

ABSTRACT

In the present work, the thermo-mechanical behavior, temperature versus thermal deformation with respect to time, of different coating films were stydied by a non-destructive technique (NDT) known as shearography. Ceramic, i.e., Al_2O_3, as well as polymer,i.e., Epoxy, coatings on metallic alloys,i.e., carbon steels, were investigated at a temperature range simulating the severe weather temperatures in Kuwait especially between the daylight and the night time temperatutres. The investigation focussed on determinig the in-plane displacement of the coating which amounts to the thermal deformation and stress with respect to temperature and time. Along with the experimental data, a mathematical relationship was derived describing the thermal deformation and stress of a coating film as a function of temperature. Furthermore, results of shearography indicate that the technique is very useful NDT method not only for determining the thermo-mechnical behavior of different coatings, but also the technique can be used as a 2-D microscope for monitoring the deformation of the coatings in real-time at a microscopic scale.

KEY WORDS: Shearography, Optical Sensor, Thermal Stress, In-Plane Displacement, High Temperature.

1.INTRODUCTION

It is well established that metrological methods of coherent lights such as speckle interferometry are found very useful in measuring deformations at surfaces of diffrent objects in various conditions[1]. However , none of the coherent light technique has been utilized for determining the thermo-mechanical behavior of diffrent coatings at metallic samples thus far. In fact , Zhang[2] was first to apply the technique of speckle shearing interferometry ,shearography, for measuring of residual stress in solid metals. The technique of Zhang was based on drilling a hole into a stressed solid metal in order to measure the localized surface relief in the metal . As a consequence, one may speculate the possibility of applying the technique of shearography of evaluating the thermo-mechanical behavior, determination of thermal stress and strain, as a function of temperature and time, of different coatings as a result of surface relief by a localized heat impingement. In this investigation, procetures of determination of the thermo-mechanical behavior of various coatings are described.

Part of the SPIE Conference on Advanced Photonic Sensors and Applications
Singapore • November–December 1999
SPIE Vol. 3897 • 0277-786X/99/$10.00

2.THEORETICAL ANALYSIS

In speckle metrology, it is well known that an interferogram of an object normally results from two speckle wavefronts nearly superimposed but having slightly phase shift because of the deformation which usually takes place at the object surface between the first and second exposure. In the case of measuring surface relief of a coating,which associated with localized heat impingement, of an area A_h, at the coating surface, consider a metal substrate has surface area, A, covered by either an organic or a ceramic coating. The ratio of the area of heat impingement to the total area of the coating under the study is

$$R = A_h/A \qquad (1)$$

Where

A_h is the area of the localized heat impingement.

A is the total area of the coating under the study.

A_h and A can be obtained by any quantitative methods of metallugraphy.In order to calculate the surface relief of a coating caused by a localized heat impingement ,the following procedures are proposed:

1-Introducing localized heat impingement , of a heat source with known diameter d, to the coating surface of of metallic samples.

2- The area of the localized heat inpingement , A_h , can be determined from the following equation:

$$A_h = \pi/4\, d^2 \qquad (2)$$

3-The ratio of the area of heat impingement to the total area of the coating under the study can be determined by calculating the actual surface area of the coating .

4-Subsequently, a surface relief, Δd, is expected to occur around the area of the localized heat impingement because of the heat induction to the coating, see Figure 1

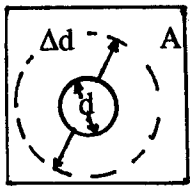


Fig.1. Localized heat impingement, d, and the surface relief, Δd, which associated with the heat induction to the surface area of the coating, A.

5- The strain relief, $\varepsilon = \Delta d/d$, is measured by the technique of shearography according to the following relationship[2]:

$$\varepsilon = \frac{\lambda}{2S} \cdot \frac{N1\,(1+\cos\theta 2) - N2\,(1+\cos\theta 1)}{\sin\theta 1\,(1+\cos\theta 2) - \sin\theta 2\,(1+\cos\theta 1)} \qquad (3)$$

Where

N1, N2 are the fringe orders, corresponding to illumination angles $\theta 1$, $\theta 2$, respectively in which $\theta 1$, $\theta 2$ are obtained from the experimental set up.

λ is the wavelength of the laser light used.

S is the constant related to the shearing camera.

For more details on the derivation of equation 3 , the reader is encouraged to refer to the literature elsewhere.

5- A relationship between Strain relief, ε , and the temperature range, ΔT, induced by the localized heat impingement is defined as follows[3]:

$$\varepsilon = \alpha\,(\Delta T) \qquad (4)$$

Where

α is the thermal expansion of the coating

ΔT is the temperature range around the localized heat impingement.

6- Eventually, the thermal stress can be measured by hook's law[3] namely, $\sigma = \varepsilon E$, Where

E is the Modolus of elasticity of the coating.

As a result, one can determine from equations 1-4, the thermo-mechanical behavior of coatings by shearography. Data on the relationship between σ , ε , ΔT, and R as a function of time for different coating will be presented in the conference.

3.REFERENCES

1- K.Robert,ERF,"Speckle Metrology",Academic Press,New York,P.2 (178).

2-H.Zhang and J.Ke,"Determination of Residual Stress by Laser Speckle Shearing Interferometry and Hole Drilling Method"Journal Of Experimental Mechanics,Vol.No.1,PP.181-188 (1986).

3-J.Shigley,Mechanical Engineering Design,McGraw-Hill,3rd Edition, New York.PP.67-69, (1977).

Mechanical Behavior Measurement of Wood by ESPI Method

Jiang Li Jun, Anand Asundi
School of Mechanical and Production Engineering, Nanyang Technological University,
Nanyang Avenue, Singapore, 639798.
Tel: (65)7905936 , 7904953 Fax: 7911859
E-mail: mljjiang@ntu.edu.sg; masundi@ntu.edu.sg

Kathrin Winkelmann
University of Canterbury, Christchurch, Newzealand
E-mail:winkelk@mech.canterbury.ac.nz;

ABSTRACT

In this paper, a non-destructive testing technique, based on electronic speckle pattern interferometry (ESPI), is presented to measure 4-point bending of wood. The mechanical behavior of the wood, such as Poisson Ratio can be obtained from the test result.
To measure the out of plane displacement produced by 4-point bending, the wood specimen which is 33x3x3mm, is secured in a special loading device. The loading device can produce 4- point bending symmetrically. The device with the specimen is put in a Michelson electronic speckle pattern interferometer assembly. Speckle pattern is produced on the natural "rough" surface of the wood when illuminated by laser. A series of laser speckle interferograms are captured by a CCD camera and then sequentially subtracted to reveal deformation of the surface. Correlation fringes can be obtained while the load is applied.
Both experimental results and theoretical analysis are presented.

Keywords: Electronic speckle pattern interferometry (ESPI), wood, Poisson's ratio

1. INTRODUCTION

Electronic speckle pattern interferometry, is a hybrid technique suitable for measuring both in-plane and out-of-plane vibrations and displacements of objects [1,2,3]. This technique has been successfully applied in non-destructive measurement of metal items, tires, etc. [4,5,6]. Such measurement usually is devised for hard or rigid objects as a real-time, non-destructive means to reveal a number of types of information about the object under test, such as temperature, stress, and certain internal defects. But no applications of ESPI for wood test are reported to the author's knowledge.
Wood is a kind of important material. Its mechanical properties, such as Possion ratio, Yang's modular , are essential elements for industry application. The mechanical behavior of the wood is different from bark to core of trunk. To investigate the property in different part of the trunk, small specimens in different sections from bark to trunk are made. The specimen is 33x3x3mm. Because it is too light in quality and weight, any attachment on the specimen will influence the properties greatly. So non-contact test method is required.
The texture of the wood specimen is rough enough to produce speckles when illuminated by laser light, and this enables ESPI to become the solution.

2. METHOD AND PRINCIPLE

The proposed system is depicted by Fig. 1. The wood specimen under test is secured on a 4 -point bending cradle form one of the two arms of the laser speckle interferometer. The other arm, is formed by a wood specimen with a similar surface to the one under test is fixed on another holder. The two surfaces are aligned to be optically superimposed and are imaged by the camera lens to the CCD to form a speckle interferogram which is captured by a PC computer. This optical system provides two ways to maximize the contrast of interferograms, the first being optically equal-path between the reference and object

arms, the second being adjustable light ratio of the two arms. These two features help improve sensitivity of the system. The test is loaded by a four point loading device(Fig.1).

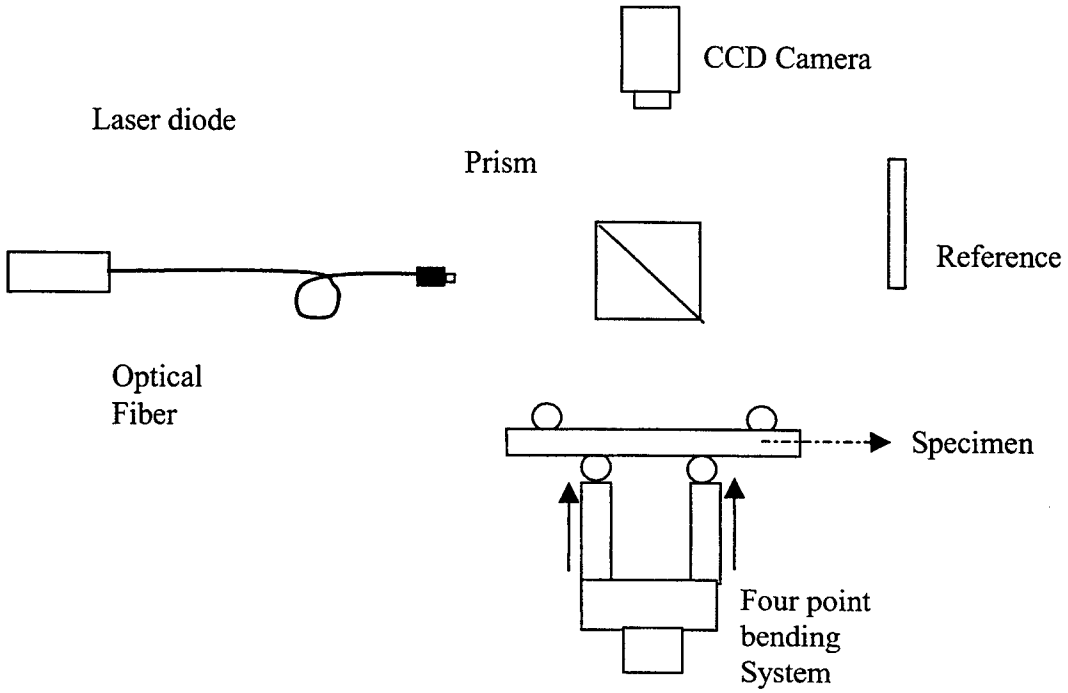


Fig1. Schematic measuring system setup

Referring to Fig. 1, the object and reference surfaces each form an image in the image plane and produce a speckle interferogram. The intensity distribution can be indicated as

$$I_{i,j}(\vec{r}) = \left| A_{1,j}(\vec{r}) + A_{2,j}(\vec{r}) \right|^2 . \tag{1}$$

where $A_{i,j}(\vec{r})$ represents the light amplitude in the image plane at point $\vec{r}$, subscript i (= 1, 2) denotes the object and reference beams respectively, and j (=1, 2) denotes the two deformation states of the object.

Now we use $\vec{r}_o$ and $\vec{r}_r$ to represent the object and reference surfaces. If the displacement corresponding to the change from state 1 to state 2 is $d(\vec{r}_o)$, then the relation between the two speckle interferograms is

$$A_{i,2}(\vec{r}_o) = A_{i,1}(\vec{r}_o - d(\vec{r}_o))\, e^{j\varphi(\vec{r}_o)}, \tag{2}$$

The phase difference $\varphi(\vec{r}_o)$ is caused by displacement and can be written as

$$\varphi(\vec{r}_o) = \frac{4\pi}{\lambda}\omega(\vec{r}_o), \tag{3}$$

where $\omega(\vec{r}_o)$ represents the corresponding out of plane deformation.

In the image plane, intensity distribution before and after deformation can be written as:

$$I_1(\vec{r}) = \left|A_{11}(\vec{r})\right|^2 + \left|A_{2,1}(\vec{r})\right|^2 + 2A_{11}(\vec{r})A_{21}(\vec{r})\cos\phi_1, \tag{4}$$

and $$I_2(\vec{r}) = \left|A_{1,2}(\vec{r})\right|^2 + \left|A_{2,2}(\vec{r})\right|^2 + 2\,A_{12}(\vec{r})\,A_{22}(\vec{r})\cos\phi_2 \quad (5)$$

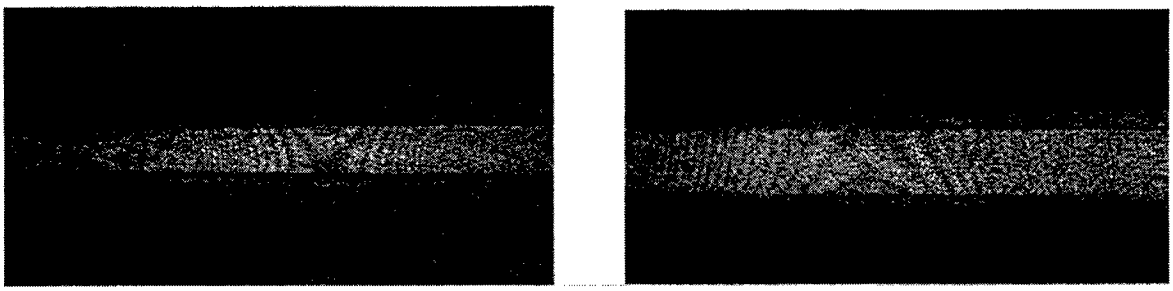
Generally, under certain statistical conditions, $|A_{i,j}|$, can be thought as the same, so

$$|I_1 - I_2| = 2|A_r||A_o|\left|\sin\frac{\phi_2+\phi_1}{2}\right|\left|\sin\frac{\Delta\phi}{2}\right| \quad (6)$$

Here $\left|\sin\frac{\phi_2+\phi_1}{2}\right|$ represents a high-frequency term which usually has less influence on the correlation fringes, $\left|\sin\frac{\Delta\phi}{2}\right|$ is a low- frequency term, and $\Delta\phi$ is the phase difference, equivalent to $\varphi(\vec{r}_o)$ in Equation (5). The fringes resulted from subtraction can be considered as an intensity modulation by $\left|\sin\frac{\Delta\phi}{2}\right|$ Thus, intensity distribution yields phase distribution, which can be used to calculate deformation distribution.

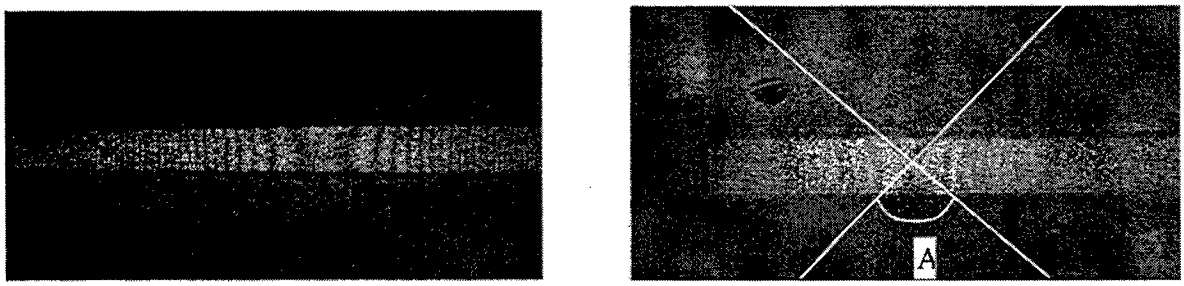
3. EXPERIMENTS

Referring to the schematic diagram of the measurement system shown in Fig. 1, a 670-nm diode laser beam is incident at normal angle on a 33X3X3mm wood specimen securely held in a 4 point bending device. The device allows longitudinal deformation along the optical axis. The speckle interferograms are acquired by a CCD camera, and transferred to a computer. Subsequent interferograms are each compared with the initial interferogram by means of image subtraction. Thus, out-of-plane deformation can be indicated by resultant fringes. The computer is programmed to process information in real-time so that the procedure of fringe changing is visible. Fig. 2 are correlation images acquired. The fringes consist of two sets of hyperbolae. The symmetry of the loading will significantly influence the shape of the fringes. The symmetry of the pattern is very important for calculating the Poisson's ratio accurately[7]. Fig.3 shows two different kinds of asymmetric image patterns caused by unbalanced loading. For the case in fig.2 (a), the loading on the upper part of the specimen is bigger than that of lower part. Fig 2(b) is just the contrary case.



(a) (b)

Fig. 2 Asymmetric pattern



(a) (b)

Fig. 3 Symmetric pattern

Fig.3 are the symmetric fringe patterns after the 4-point bending device is improved to load symmetrically. The fringe of Fig.3 (a) is less dense than that of Fig.3(b) because of the less loading than that of fig.3 (b). In the case of Fig. 3(b), the angle, A, between the hyperbolas satisfies the relationship

$$V = \tan^2 (A/2) \qquad (7)$$

V is the Poisson's ratio of the material.
The interferometric observation of four-point bending deformation therefore provides a useful technique for the measurement of Poisson's ratio.
As so far, measuring accuracy is still low because of the lower fringe contrast, lower fringe resolution. We are trying to improve the image quality to elevate accuracy. Data analysis and comparison will be presented by a coming paper.

4. CONCLUSION

The measuring systems based on electronic speckle pattern interferometry for testing four point bending has been described with experimental results. The proposed technique is non-destructive, non-contact, whole field with high sensitivity. Nevertheless, there are still many improvements to do, such as enhancing fringe quality, increasing contrast to get "finer fringe" so that to improve measuring accuracy, etc.

ACKNOWLEDGMENTS

We are grateful for my colleagues' help in the lab of strength of material.

REFERENCES

[1] J. C. Dainty, *Laser Speckle and Related Phenomena*, Springer-Verlag, London, 1975, pp. 2-253.
[2] J. N. Butters and J. A. Lenders, "Holographic and Video Techniques Applied to Engineering Measurements," J. Meas. Control, vol. 4, 1971, pp. 349-354.
[3] A. Macovski, D. Ramsey, and L. F. Schaefer, "Time Lapse Interferometry and Contouring using Television Systems," Appl. Opt., vol. 10, 1971, pp. 2722-2727.
[4] R. Jones, "The Design And Application Of A Speckle Pattern Interferometer for Total Strain Field Measurement," Opt. Laser. Tech., vol. 8, 1976, pp. 215-219.
[5] R. P. Khetan, F. P. Chiang, "Strain Analysis by One Beam Laser Speckle Interferometry: 1. Single Aperture Method," Appl. Opt., vol. 15, 1976, pp. 2205-2215.
[6] Y. Y. Hung, C. Y. Liang, "Image Shearing Camera for Direct Measurement of Surface Strains," Appl. Opt., vol. 18, 1979, pp. 1046-1051.
[7] R. Jones & C. Wykes, "Holographic and speckle interferometry" , 2nd Edition, Cambridge Univ press,1989, 104-106.

Pattern Measurement of Speckle Noise in Optical Fiber and Its Application to Sensor

Masanari TANIGUCHI[*a], Masato OKI[b] and Tasuku TAKAGI[c]

[a] Faculty of Science and Technology, Meijo University, Nagoya, Japan
[b] Faculty of Engineer, Ashikaga Institute of Technology, Ashikaga, Japan
[c] Faculty of Science and Technology, Tohoku Bunka Gakuen University, Sendai, Japan

ABSTRACT

An optical fiber has been used in a wide area such as in communication system, a view guide for diagnosis, applied optical sensor, etc.. In such applications of the optical fiber especially in the sensor, some new sensor may be expected. In this study, possibility of a speckle noise used sensor was discussed. For this, the speckle noise patterns were observed to make clear their characteristics with regard on the miss-match of optic axis and an experimental study was carried out on the influence of edge surface condition, core diameter and distortion. In addition, the speckle noise pattern was measured by using the digital image processing technique, and its characterization was partly made. From these investigations, the possibility of speckle noise used sensor was discussed.

Keywords : Optical fiber, Speckle noise, Sensor, Pattern measurement, Digital logical operation,

1. INTRODUCTION

As is well-know, an optical fiber has many excellent characteristics such as low loss, and wide-band transmission, non-leak, thin diameter and light-weight. Therefore, the optical fiber is applied not only in the electric communication,[1-3] but also, the mesurement[4,5], the sensor.[6,7] etc. However, the speckle noise in optical fiber is one of the most annoying problems in treating a coherent light.[8,9]

In this study, both the optical transmitted loss and the speckle noise pattern were investigated. That is, the speckle noise pattern was observed with regard to the axial mis-match, and the experimental study was carried out on the influence of a fiber loop shape. In addition, the experimental study was also carried out on the influence of edge surface condition.

In order to apply the speckle noise to a sensor, the relationship between the speckle noise pattern and the of mis-match in axes between the fiber and the light source center, and the deformation of an optical fiber loop were investigated. The speckle noise patterns were measured by using the digital signal processing system (DSPS) along with the digital logic operating technique (DLOT), then the characteristics of speckle noise were obtained for the present purpose.

2. TEST SAMPLE OF OPTICAL FIBER AND EXPERIMENTAL SET-UP

2.1. Test Sample of Optical Fiber

In this study, as the test sample of optical fiber, two kind of quartz fibers were chosen, which are the small size core diameter (Test sample A) and the large size core diameter (Test sample B) as

* Correspondence ; E-mail:taniguchi@meijo-u.ac.jp, Tel:+81-52-832-1151, Fax:+81-52-832-1244,

shown in Table 1.

For the sample A, both step-index (SI) and graded-index (GI) type were used, and for B, 3 different diameter (step-index) fibers were used. The core diameter and length are shown in Table 1.

2.2. Experimental Set-Up

The schematic diagram of experimental set-up is shown in Fig. 1 for the measurement of transmitted loss and the speckle noise pattern. The transmitted image was irradiated on a screen for the speckle noise pattern observation.

The transmitted loss was measured with respect to the optic axis mis-mach. That is, when the receiving connector was moved, the light intensity level from the optical fiber was measured by using an optical power meter (Ando Elec. Co. AQ-1111).

In this set-up, the light source was a He-Ne gas laser (wave length : 6328Å, optical power : 2mW).

Table 1. Test samples of optical fiber

Test sample	A		B		
	A1	A2	B1	B2	B3
Type	SI	GI	SI	SI	SI
Core diameter [μ m]	50	50	200	400	600
Length [mm]	500	500	1000	1000	1000

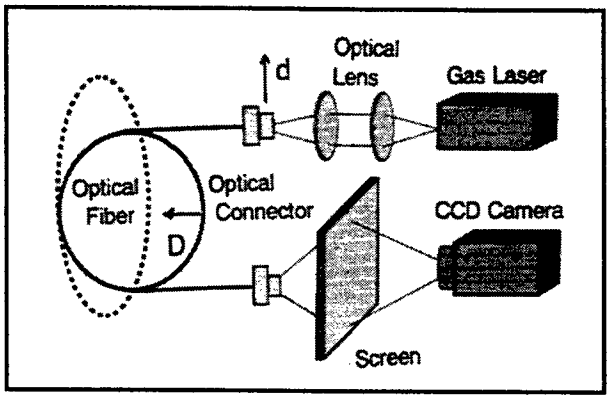


Fig. 1. Schematic diagram of experimental set-up.

3. SPECKLE NOISE PATTERN OF SMALL CORE DIAMETER (Sample A)

3.1. Optical Transmitted Loss

Fig. 2 shows the experimental results of small core diameter in Test sample A, in which, the truth line shows the case of SI mode fiber and a dotted line shows the case of GI mode fiber. From this measurement, the same shapes of curve were obtained, however the magnitude of loss was different, of which reason has not been clear.

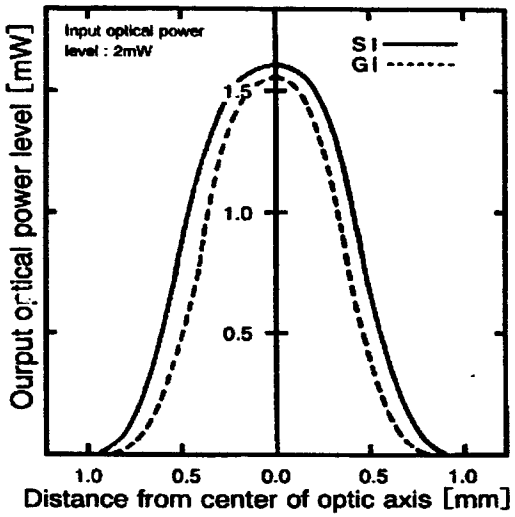


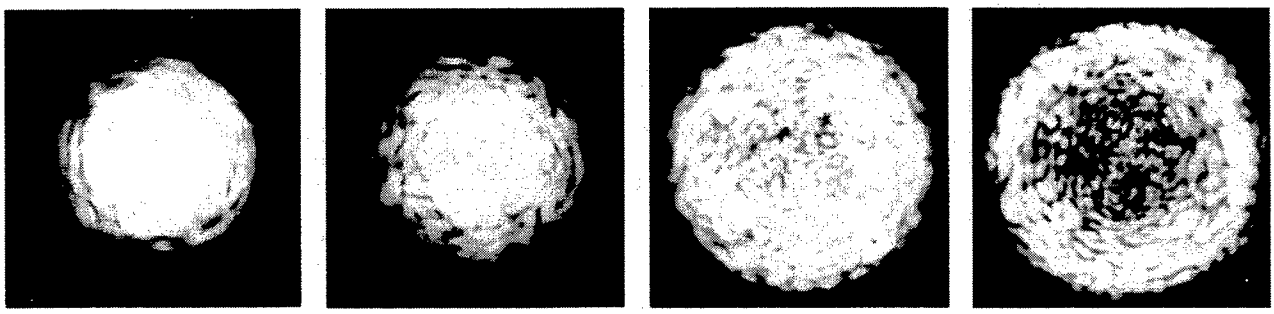
Fig. 2 The experimental results of optical transmitted loss. [Test sample A]

3.2. Observation of Speckle Noise Pattern

Figs. 3 and 4 show the optical image through the test sample A in case where the distance from the center axis of light source (d in Fig. 1) was moved. Fig. 3 shows the speckle noise pattern in SI mode fiber, and Fig. 4 shows the image in GI mode. In these figures. (a) is the distance level d=0.0mm, (b) d=0.2mm, (c) d=0.4mm and (d) d=0.6mm, respectively.

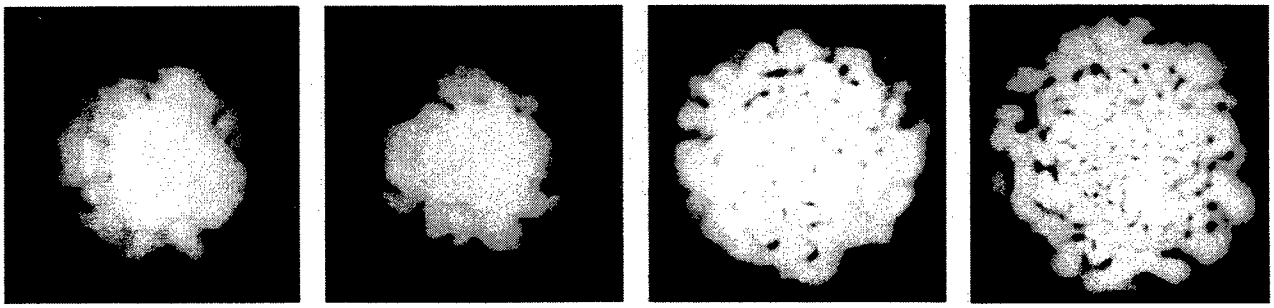
Figs.5 and 6 shows the speckle noise pattern through the test sample A in case where the optical fiber loop (Fig.1) was deformed (D changed). Fig.5 shows the image in SI mode fiber, and Fig.6 the image in GI mode. In these figures. (a) is the displacement level D=0mm, (b) D=2mm, (c) D=4mm and (d) D=6mm, respectively.

As shown in Figs.3 and 4, the size of optical image spreads more, and the intensity level decreased as d increased, and, at the same time, the speckle noise itself increased. And, as shown in Figs. 5 and 6, when the optical fiber loop was deformed, both the size of optic image and the intensity level did not changed, however, the speckle noise pattern was changed although its quantitative changes was not clear at present. From the experimental results, the speckle noise pattern was observed clearly in SI mode type fiber



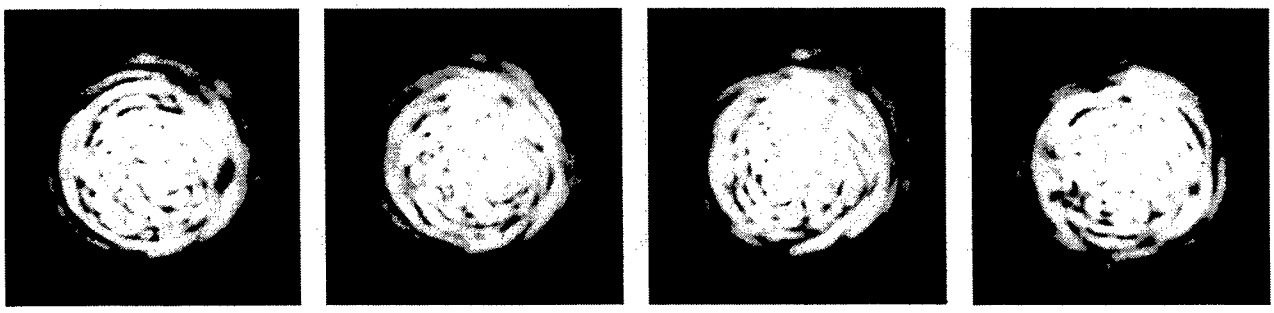
(a) Distance d=0.0mm (b) Distance d=0.2mm (c) Distance d=0.4mm (d) Distance d=0.6mm

Fig.3 The optical image through the SI optical fiber in case optic axis was moved. [Test sample A1]



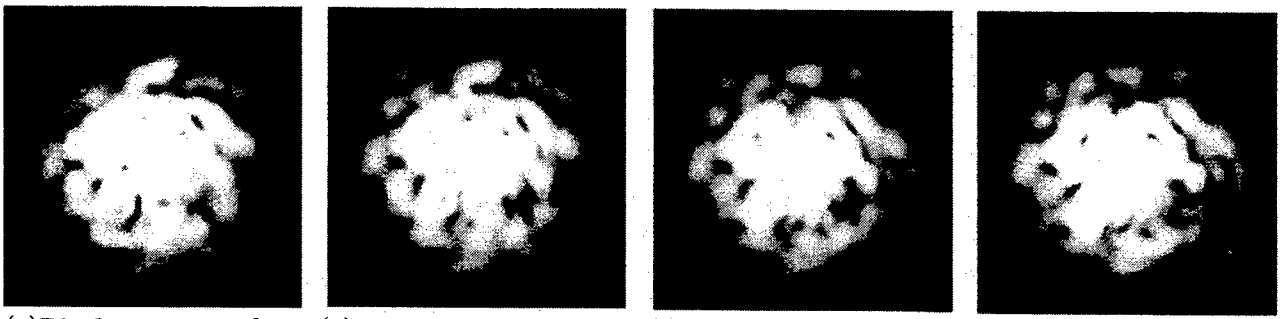
(a) Distance d=0.0mm (b) Distance d=0.2mm (c) Distance d=0.4mm (d) Distance d=0.6mm

Fig.4 The optical image through the GI optical fiber in case optic axis was moved. [Test sample A2]



(a)Displacement D=0mm (b)Displacement D=2mm (c)Displacement D=4mm (d)Displacement D=6mm

Fig.5 The optical image through the SI optical fiber in case optic fiber loop was distorted. [Test sample A1]



(a)Displacement D=0mm (b)Displacement D=2mm (c)Displacement D=4mm (d)Displacement D=6mm

Fig.6 The optical image through the GI optical fiber in case optic fiber loop was distorted. [Test sample A2]

4. SPECKLE NOISE PATTERN OF LARGE SIZE CORE DIAMETER

4.1. Optical Transmitted Loss

Fig.7 shows the experimental results of large size core diameter (Test sample B). The optical transmitted loss in case where the optical axis of receiving connector was moved in the same way with sample A.

The effect of the core size has appeared, however those differences with respect to core diameter are not so much important at the present purpose.

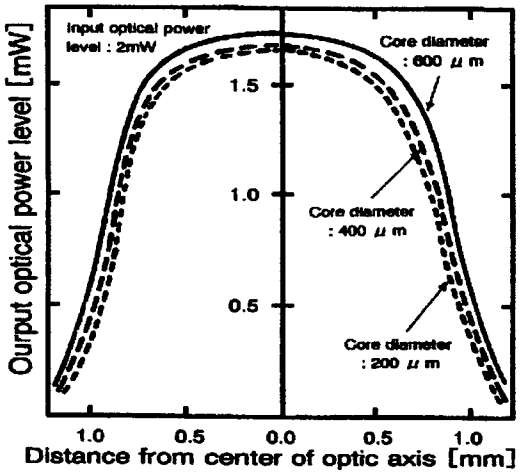


Fig.7 The experimental results of optical transmitted loss. [Test sample B]

4.2. Observation of Speckle Noise Pattern

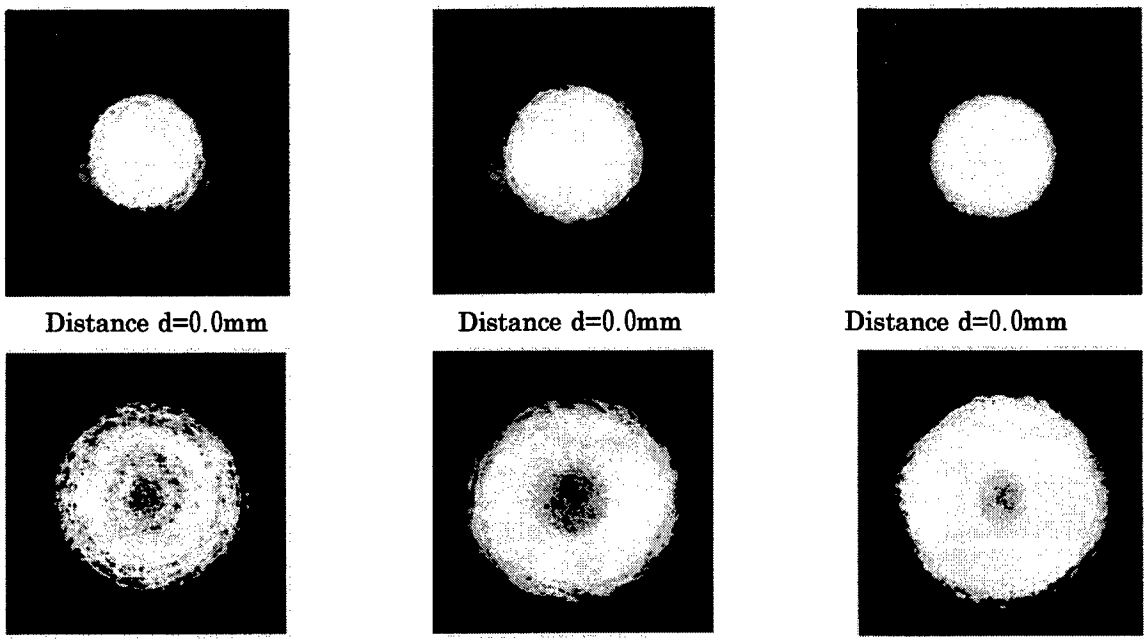
Fig.8 shows the optical image through the test sample B in case where the distance d from the center axis of light source was moved. The distance were 0.0 mm and 0.6mm, respectively. Fig.8(a) shows the case of diameter 200 μ m, (b) the case 400 μ m and (c) the case 600 μ m.

As shown in Fig.8, in the small size core diameter, the size of optical image spreads more and the intensity level decreased as d increased. At the same time, the speckle noise itself increased. Especially, the speckle noise pattern was observed clearly in case of small size core. In addition, when the optical fiber loop was deformed, both the size of optic image and the intensity level did not changed, and the speckle noise pattern was changed although its quantitative changes was not clear now.

5. EDGE CONDITION ON CORE SURFACE AND SPECKLE NOISE PATTERN

5.1. Edge Condition on Core Surface and Transmission Loss

In this study, the relationship between the speckle noise pattern and the edge condition on core surface were investigated. In the experiment, the test sample of optical fiber is a test sample B1. (the core diameter : 200 μ m, the length of fiber : 1000mm, step-index multi mode). The four kinds of edge condition on core surface were settled as shown in Table 2. Microscopic surface image of each condition are shown in Fig.9.



Distance d=0.0mm　　Distance d=0.0mm　　Distance d=0.0mm

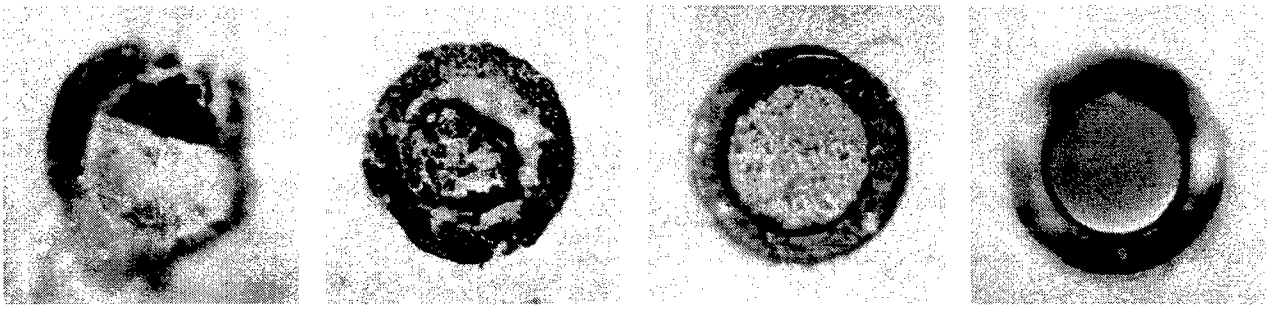
Distance d=0.6mm　　Distance d=0.6mm　　Distance d=0.6mm

(a) Test sample B1　　(b) Test sample B2　　(c) Test sample B3

Fig.8 The optical image through the test sample B optical fiber in case optic axis was moved.

Table. 2 The edge condition on core surface of optical fiber.

Condition	Process on core surface of optical fiber
Edge condition A	In case of cut by using shap knife only.
Edge condition B	In case of polished simply by using emery paper after edge condition A.
Edge condition C	In case of polished enough by using emery paper after edge condition A.
Edge condition D	In case of polished enough by using powder after edge condition C.



(a) Edge condition A　(b) Edge condition B　(c) Edge condition C　(d) Edge condition D

Fig.9 The picture image of the core surface for each condition.

As shown in Fig.9, in case of condition A, the whole core surface is damaged heavily. In case of B, the damage of core surface was improved, and, in case of C, the core surface was more improved. Finally, in case of polished enough by using powder, the core surface was made as mirror surface.

Fig.10 shows the experimental results of optical transmitted loss in case where the optical axis of receiving connector was moved (d in Fog.1 was changed). From this, condition D is enough in the present purpose in carrying out of the basic stage of investigation.

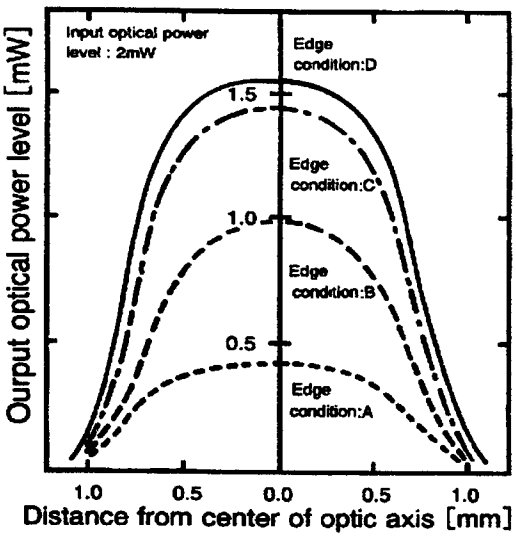
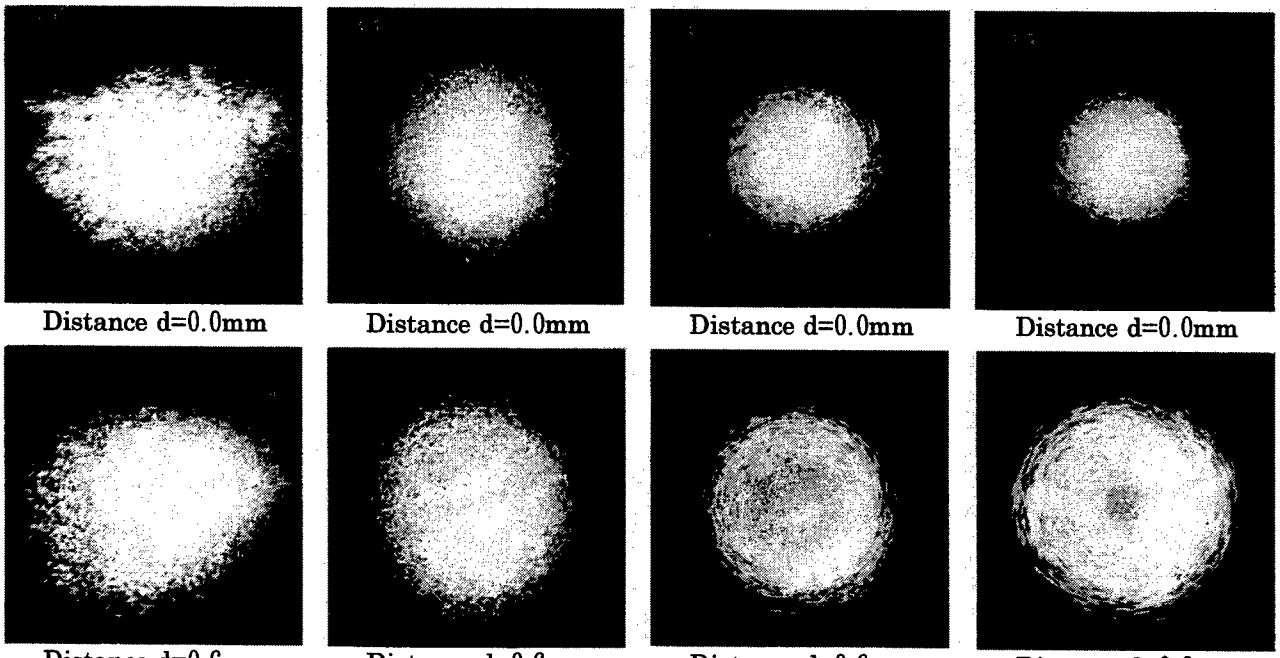


Fig.10 The experimental results of optical transmitted lose for each surface condition.

5.2. Observation of Speckle Noise Pattern

Fig.11 shows the optical image through the test sample of optical fiber in case where the distance d from the center axis of light source was moved . Fig. 11(a) is the edge condition A, (b) condition B, (c) condition C and (d) condition D.

As shown in Fig.11, the speckle noise pattern spreads widely when the optic axis of receiving connector goes away from the center axis of the light source, at the same time, the intensity level decreased. From the speckle noise pattern, surface condition of D may be used as a new sensor.



Distance d=0.0mm | Distance d=0.0mm | Distance d=0.0mm | Distance d=0.0mm

Distance d=0.6mm | Distance d=0.6mm | Distance d=0.6mm | Distance d=0.6mm

(a) Edge condition A (b) Edge condition B (c) Edge condition C (d) Edge condition D

Fig.11 The optical image through the edge surface processed optical fiber in case optic axis was moved.

6. DIGITAL IMAGE PROCESSING OF SPECKLE NOISE PATTERN

6.2. Digital Image Processing System (DIMS)

In order to extraction of speckle noise pattern from the optical image through the optical fiber, the digital image measuring system (DIMS) was developed. The schematic diagram of the DIMS is shown in Fig.12. The DIMS consists mainly of the combination of an optical camera, a digital signal processing unit (include an A/D, D/A converter and a frame memory), a multi bus, a personal computer, etc.. One flame image is 256 × 256 pixels.

In this study, for quantitative extraction and characterization of the speckle noise pattern variations, the digital logic operating technique (DL OT) was applied.

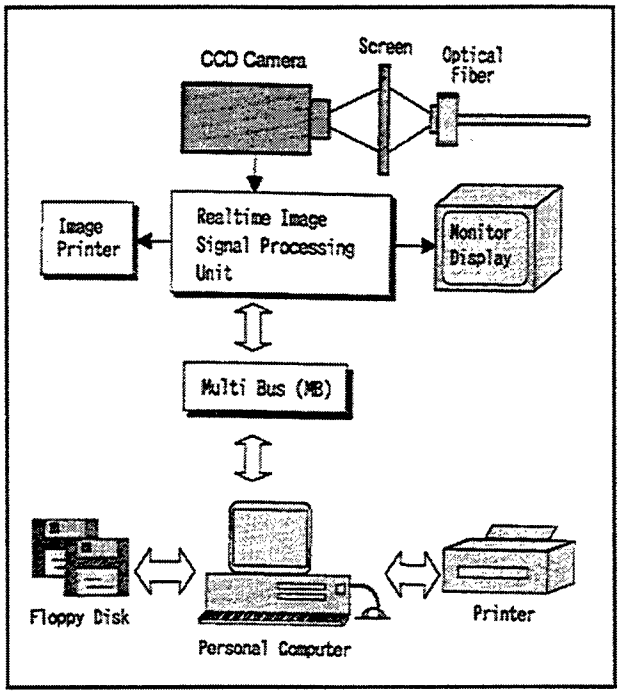
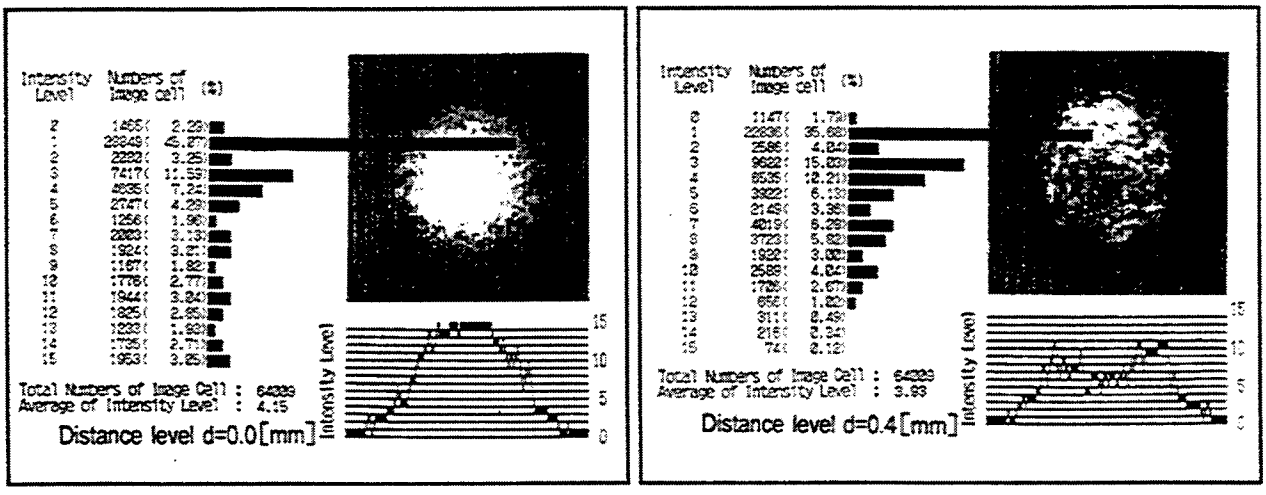


Fig.12 Schematic diagram the digital image measuring system (DIMS)

6.2. Measurement of Intensity Level

Fig.13 shows the digitally measured results of intensity level by using the DIMS. The distribution of the intensity level of optical image were shown along with the histogram of intensity levels which are divided into 16 levels.

As shown in Fig.13, the distribution of light intensity level shows gloomily on the whole, and the level of center decreased as d increased. However, the intensity level did not changed in case where the optical fiber loop was deformed due to press.



(a) Distance level d=0.0mm (b) Distance level d=0.4mm

Fig.13 Digital image measurement results of speckle noise pattern in case where optic axis was moved.

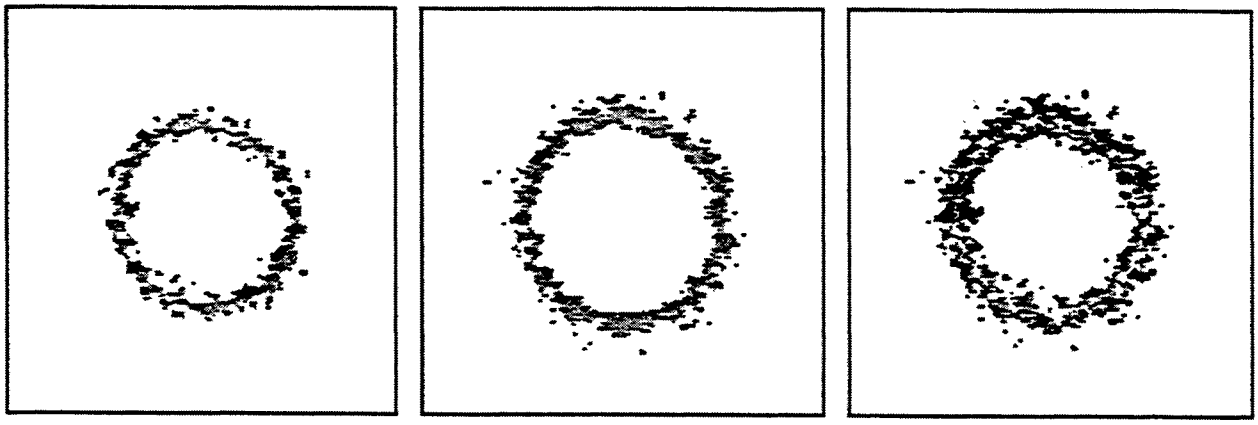
6.3. Extraction of Speckle Noise Pattern

In this study, for quantitative extraction and characterization of the change of speckle noise pattern, the digital logic operating technique (DLOT) was applied. Figs.14 and 15 show the experimental results of speckle noise pattern computed by using a exclusive logical OR method (XOR)in case where the displacement level (D in Fig.1) were 0mm and 4mm.

Fig.14(a) and (b) show the speck noise patterns of the intensity level 6. Figs.15(a) and (b) show the speckle noise patterns of the intensity level 10, respectively.

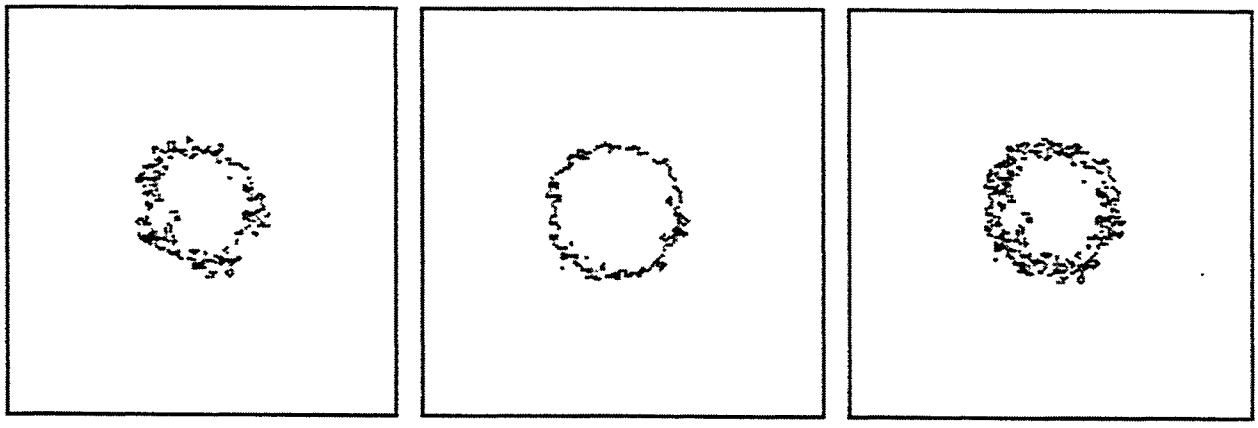
Fig.14(c) and 15(c) show the speckle noise patterns computed by using a exclusive logical OR method (XOR), in which each pixel (256×256)was computed from Figs.14(a), (b) and Figs.15(a), (b), respectively. Fig.16 shows the XOR operated and extracted pixel number from the speckle noise pattern by using a exclusive logical OR method from the above data.

As shown in Fig16, by applying the digital logical operation technique (DLOT), the level change of the speckle noise pattern could be extracted. However, it is difficult to mention for making this method could be possible to apply to a sensor, However, by using more sophisticated way for analyzing the speckle noise patterns, we are convinced of making the speckle noise based optical fiber sensor.



(a) Displacement D=0mm (b) Displacement D=2mm (c) DLOT computed result

Fig.14 Extracted results of speckle noise pattern by using DLOT in case where the intensity level was 6.



(a) Displacement D=0mm (b) Displacement D=2mm (c) DLOT computed result

Fig.15 Extracted results of speckle noise pattern by using DLOT in case where the intensity level was 10.

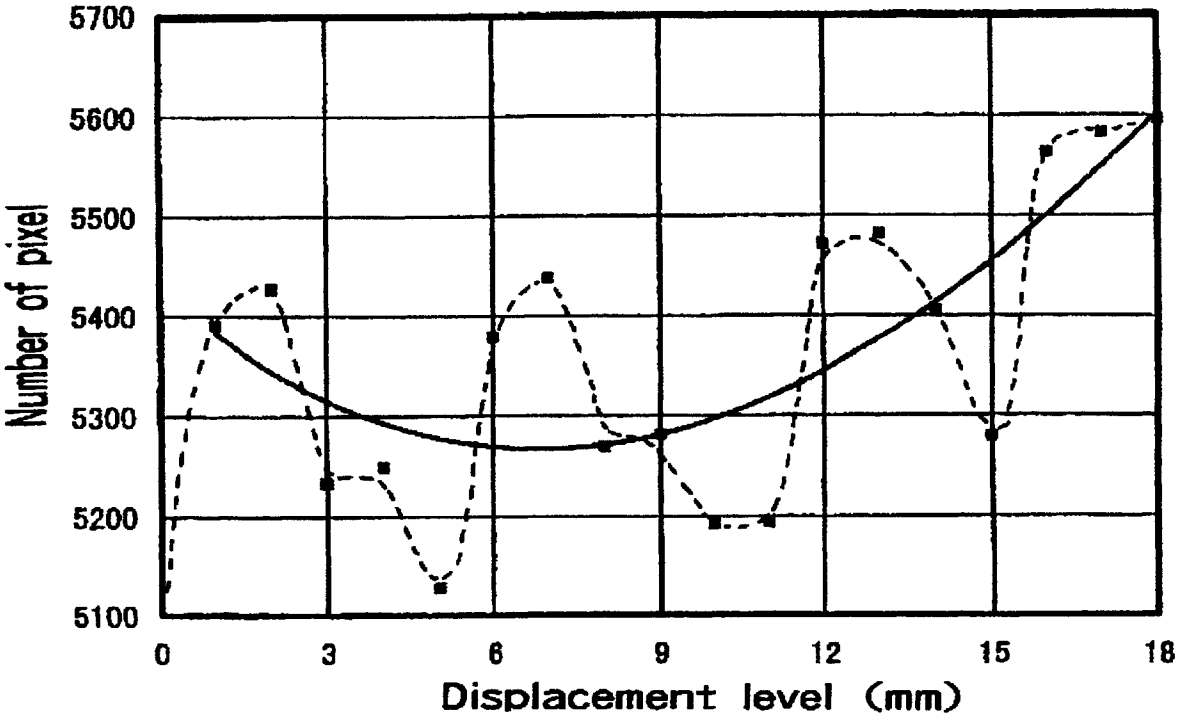


Fig.16 XOR operated and extracted pixcel number from the speckle noise patter.

7. CONCLUSIONS

The authors present purpose was to find some possibility to develop a new optical fiber sensor by means of positive use of the speckle noise. The present investigations described in this paper have shown only its possibility, however we have had the possible result on this matter. We can summarize the present investigations as follows;

1) Overall speckle noise pattern variation with respect to the axis mis-match between fiber and light source was obtained.
2) Overall speckle noise pattern variation with respect to fiber loop variation was obtained.
3) By using newly developed Digital Image Measuring System (DIMS) was well applied to analyze the speckle noise pattern.
4) Two possibilities in applying the speckle noise to sensor for displacement were obtained : one is to use the axis mis-match for small displacement measurement less than 1 mm, another is to use the fiber loop deformation for larger displacement, the range can depend upon the size of loop.

Finally, we conclude here that some new analytical and quantitative means in treating the speckle noise should be developed in future.

REFERENCES

1. T.Kaneda:"Fiber Optic Video Transmission Technology", J. Institute of Television Eng. Japan,45(5), pp.618-623,(May.1991)
2. Y.Unami:"Optical Transmission for Fiber Channel", Optronics,15(9),pp.144-149,(Sep.1996)
3. M.A.Perez,J.C.Campo,C.J.Blanco and J.C.Alvarez:"A New Sensors Network Based upon Fiber Optic Link to Work under Intrinsic Safety Conditions", Proc.of IEEE IMTC,pp.1531-1536,(May.1999)
4. S.Banerjee,C.Zan,R.Bhat and C.Caneau:"Fiber Otptic Photoluminescence Setup for Spatially Resolved Measurement", IEEE Trans. on IM,48(1),pp.45-49,(Feb.1999)
5. W.J.Bock and W.Urbanczyk:"Fiber-optic Sensor System for Absolute Measurement", Proc. of IEEE IMTC,pp.942-947,(May.1998)
6. A.Wang,G.Z.Wang,K.A.Murphy and R.O.Claus:"Fiber Optic Temperature Sensors Based on Differential Spectral Transmittance / Reflectivity mand Multiplexed Sensing System", Applied Optics,34, pp.229-300,(1995)
7. T.Takemoto:"Observation of Crustal Deformation Using Optical Devices", Optonics,15(6),pp.140-144, (Jul.1996)
8. K. Sato and K. Asatani :"Speckle Noise Reduction in Fiber Optic Analog Video Transmission Using Semiconductor Laser Diode", IEEE Trans. on COM,29(7),pp.1017-1024,(Jul. 1981)
9. R.E.Epworth:"Model Noise-Couses and Cures", Laser Focus,pp.109-115,(Sep. 1981)

SESSION 13

Structural Health Monitoring

Invited Paper

Multiplexed FBG sensors and their applications

Wei Jin
Department of Electrical Engineering, The Hong Kong Polytechnic University,
Hung Hom, Kowloon, Hong Kong

ABSTRACT

Fibre Bragg gratings (FBGs) are considered excellent sensing elements for a number of applications such as load monitoring in civil structures. Over the past decade, significant progress has been made in FBG fabrication, interrogation and multiplexing. This paper briefly reviews recent progress on FBG multiplexing and applications with emphasis on the work conducted at the Hong Kong Polytechnic University.

Keywords: Fibre Bragg gratings, optical fibre sensors, Bragg grating sensors, smart structures, multiplexing

1. INTRODUCTION

FBG sensors have been an area of intense research and development since the discovery of photosensitivity in optical fibre by Hill et al [1]. Numerous papers have been published in this area and the state-of-the-art may be found in recent review articles by Morey et al [2], by Kersey et al [3], by Rao [4] and by Othonos and Kalli [5]. The Bragg wavelength of an FBG shifts with applied strain, temperature and pressure and can therefore be used to sense these parameters. Multiple fibre gratings can now be written at arbitrary locations along a single optical fibre, with reflectivity ranging from 0.1% to approaching 100% and spectral width tailored from 0.1nm to a few tens of nm. This makes them ideal for multi-point or quasi-distributed sensing.

Practical applications are based on the detection of the relatively small shift in the Bragg wavelength caused by the measurands (e.g., strain). This area has received significant attention recently, with a variety of wavelength detection schemes demonstrated [2-5]. These include the use of CCD spectrometers [6,7], wavelength to intensity conversion using passive optical filters [8-10], tracking using tunable filters [11-13], wavelength tunable sources [14,15], Fourier transform spectroscopy [16,17], interferometric detection [18,19], mode-locked fibre lasers [20,21], wavelength-swept fibre laser [22], etc. The detection resolution is scheme-dependent and for most of the aforementioned schemes resolution on the order of 1-με has been achieved that is sufficient for many applications.

FBGs have been identified as an enabling technology for a variety of applications from aerospace, marine, to civil engineering structural monitoring, from medical, nuclear power industry, to electrical power industry. In many applications such as condition monitoring of civil structures such as highways, bridges and buildings, FBG sensors need to be multiplexed in order to achieve quasi-distributed measurement and to reduce cost per sensing point to compete with conventional electrical or optical sensors [2-5].

Two popular techniques for FBG multiplexing are wavelength division multiplexing (WDM) [13,23-25] and time division multiplexing (TDM) [26]. The maximum sensor number that can be multiplexed using WDM is determined by the ratio of the source spectral width over the width of the wavelength slot allocated to individual FBGs. The wavelength slot is related to the dynamic range of the sensors. The sensor number in a TDM system is also limited [26]. In a TDM system, the pulse duty cycle is low when the number of sensors in the array is large. This requires the use of high peak power optical pulses in order to obtain a reasonable average power level for each sensor at the photo-detector and thus a good signal to noise ratio for each individual sensor[26, 27]. WDM, TDM and other multiplexing techniques such as spatial division multiplexing (SDM) can be combined to multiplex a larger number of FBG sensors [2-5, 28-30].

The state-of-the-art is a 60-sensor array using WDM combined with the sequential addressing of separate fibre arrays via a single-mode fibre switch [28]. A recent paper [29] reported a FBG sensor array capable of supporting 128 FBGs or 64 temperature-referenced strain sensors (with 8 sensors tested experimentally) using a combination of WDM and SDM.

Part of the SPIE Conference on Advanced Photonic Sensors and Applications
Singapore • November–December 1999
SPIE Vol. 3897 • 0277-786X/99/$10.00

Dense WDM of FBG sensors using code division multiple access (CDMA) was reported recently [31]. The CDMA technique has the advantage of larger average sensor output power over the TDM and consequently better signal to noise ratio. Although a two-sensor system was experimentally demonstrated, it is expected that the technique be capable of potentially multiplexing over 100 FBGs in a single serial sensor array using a source of 40nm bandwidth.

In addition to improving wavelength detection resolution and multiplexing capacity, an important issue is to enhance the overall measurement accuracy of each sensor within the array over a dynamic range required by a specific application. Consider a multi-sensor system where an optical Fabry-Perot tunable filter is used to detect wavelength shift and to realize WDM [23,28-31], although high wavelength resolution of better than 1 pm has been achieved with this approach, the non-linear wavelength response due to Fabry-Perot tunable optics and the mechanical actuator (PZT), drift and instability from thermal coefficients of Fabry-Perot filters and other components could cause an overall wavelength measurement inaccuracy of several nanometers (several thousand times the resolution) [32]. By taking differential measurement between the sensing FBG and a reference FBG with a similar center wavelength, some of the above problems (e.g., drift) can be minimized [4,23]. The sensing-reference FBG pair approach has been used in a multiplexed FBG array to provide temperature-referenced quasi-static measurements of strain [29]. The referencing technique also minimizes the effect of non-linearity of the tunable optical filter by use of multiple pairs of sensing-referencing FBGs. However, by using only the sensing-reference pair approach, temperature-reference and drift-compensation can not be performed at the same time [29]. A multi-wavelength referencing system using a fixed Fabry-Perot interferometer in combination with a reference FBG has been used to solve these linearity, stability and accuracy problems [32]. The system was demonstrated with overall system accuracy of ± 5 $\mu\varepsilon$. The combination of the Fabry-Perot referencing and the sensing-referencing FBG pairs approach would provide a temperature-referenced strain measurement with good overall system accuracy.

In this paper, we report our recent work on FBG multiplexing, interrogation and applications. Two multiplexing schemes will be reported in sections 2 and 3 respectively. A number of experimental studies in applying FBGs for structural sensing are presented in section 4. A summary is given in section 5.

2. RF-BAND FMCW MULTIPLEXING OF FBGS

The basic principle of the FMCW technique has been described in [33]. The FMCW technique has been used for conventional fibre-optic sensor multiplexing [34,35]. The FMCW principle can be applied by either sweeping the frequency of light in the optical-frequency domain or by sweeping a subcarrier in the RF range (termed as RF-Band FMCW).

We have recently applied the RF-band FMCW principle for multiplexing FBG sensors [36]. Fig.1 shows an example of FMCW multiplexed serial-connected single fibre FBG array. Light from a broadband source is modulated in intensity by a saw-tooth (or triangular) swept frequency carrier generated from a voltage-controlled oscillator (VCO) and launched into the FBG sensor array. The intensity modulation can be performed by using an external modulator or by direct modulation of the driving current if a semiconductor source is used. The reflected signals from FBGs are guided back to a tunable optical filter (TOF), then a photo-detector, and mixed with a reference signal from the VCO subsequently. The mixing between the reference signal and the reflected signals produces beat notes with their beat frequencies determined by the time delay differences between the reflected FBG signals and the reference signal. The magnitudes of the beat signals are proportional to the convolution of the individual grating reflection spectrum and the transmission spectrum of the TOF. The Bragg wavelength of a particular FBG can be interrogated by scanning the TOF and record the control voltage of the TOF that corresponds to the peak (or zero crossing point if derivative response is measured) of that particular beat note signal.

In general, each beat note has a multiple line spectrum at intervals of the saw-tooth repetition rate. The position of the peak in the envelope of the line spectrum gives the beat frequency. Because of the multiple line characteristics of the beat-note spectra, the crosstalk performance of the FMCW technique was in general regarded as "not good" when it is applied to multiplex a number of sensors. However, we found that, for small time delay difference between the signal and the reference wave, the beat note spectrum can be essentially a single line when the beat frequency coincides with one of the harmonics of the sweeping signal [36]. This property of the RF-band FMCW was also reported in an early paper on multiplexing of vibrating fibre optic sensors [35]. This suggests that a low crosstalk multiple FBG sensors array may be realized if the network is designed in such a way the beat frequency corresponding to a particular FBG sensor is an integer time of the saw-tooth repetition frequency. This requires that the adjacent FBGs are spaced in equal distance with their

respective delays (relative to the reference) satisfying: $2\Delta f\tau_i=mf_s$, where m is an integer, Δf is the excursion of the sawtooth frequency sweeping, τ_i is the delay of the i-th sensor relative to the reference and f_s is the rate of sawtooth modulation. Our simulation results show that, for the particular set of parameters given in [36], 32 FBGs of the same nominal wavelength can be multiplexed in a reflective tree type network with a crosstalk level between any two of the sensors below -48dB. A sideline suppression ratio of –30dB was obtained in a preliminary experiment [36] and has been improved to about –40dB recently [37]. Based on these results, we can predict that a sensor number of up to 32 with the same nominal Bragg wavelength may be multiplexed with a crosstalk level of better than -30dB between any two of the sensors. This level of crosstalk is equivalent to a strain measurement accuracy of a few microstrains if FBGs with spectral width of ~0.1nm are used [38]. The FMCW technique provides a similar average power level at the photo-detector but possible lower cost compared with the CDMA approach because of the simple electronic processing circuitry required.

When combined with WDM and/or SDM techniques, the FMCW technique should be capable of potentially multiplexing a few hundreds of FBGs. A number of multiplexing topologies may be used. Fig.2 and Fig.3 show two examples. In Fig.2a, a parallel topology is shown where M FBGs of the same nominal Bragg wavelength are FMCW multiplexed along a fibre string, while N different strings are WDM multiplexed. In the system output, there would be M frequency components corresponding to M sensors along a fibre string, and for a particular frequency component, there would be N maximums when the TOF is scanned to cover the reflection spectra of the whole grating array, corresponding to N strings. Fig.2b shows an example of system output in the frequency-wavelength domain, where M=N=3 has been assumed. A total number of MxN sensors can be multiplexed using this topology. Fig.3 shows a serial topology where gratings within each set are FMCW/WDM multiplexed, while different sets are FMCW multiplexed. Different delay lines may need to be used for different set of sensors in order to approximately "match" the delays of a particular set of sensors to that of the reference signal. A single delay line for all sets of sensors would result in the delay between some sets of the gratings and the reference be too large thus producing appreciable sidelines and affecting the system crosstalk performance. Other topology [5] may also be used.

3. INTERROGATION OF FBGS BY USE OF A TUNABLE LASER SOURCE

So far, most of the reported approaches for wavelength detection and sensor multiplexing are based on the use of broadband sources coupled with various wavelength-selective devices. These systems, while having achieved reasonable multiplexing gain by adopting a combination of multiplexing techniques such as WDM, TDM and SDM, the number of sensors that can be multiplexed is still limited simply because of the limited spectral density of the source at the Bragg wavelengths. The use of TDM, SDM or FMCW technique reduces the average power per sensor at the photo-detector and therefore the signal-to-noise ratio of individual sensors. The use of tunable laser sources, which provide high power within a narrow spectral width (i.e., high spectral density) and a reasonable tuning range, would increase dramatically the signal-to-noise ratio and allow a larger number of sensors to be multiplexed by combining WDM with TDM, SDM or FMCW techniques. Recent development in instrumentation using laser sources includes addressing a number of FBG sensors in several branches using TDM [39] and addressing several sensors along the same fibre using WDM [22,40]. Activities on the development of practical tunable sources (fibre tunable laser sources) have also been reported [14,41] and a tunable range of 11nm has been achieved [41].

However, the use of highly coherent laser sources does introduce additional problems such as coherent multi-path effects that limit sensor resolution [38,42]. The multi-path effects can cause noise and crosstalk between sensors and should be reduced to an acceptable level before a TDM (or FMCW) system using a laser source can be practically implemented.

In this section, I report on the results of our recent investigation on the noise and crosstalk performance of FBG sensors or sensor arrays that use tunable laser sources.

3.1 Coherent multi-path effect in a single FBG sensor

Fig.4 shows a reflective type FBG sensor using a tunable laser source. Because of the long coherence length of the source, residual reflections (examples of reflections are shown in (a) to (c) in Fig.4)) would interfere coherently with the signal and cause fluctuation in the light intensity at the photo-detector that results in errors in the measurement of the Bragg wavelength. The noise characteristics of both direct detection (peak wavelength detection) technique and the first derivative technique have been studied [38]. For the first derivative technique, the laser wavelength is dithered sinusoidally while the average wavelength is scanned across the grating spectrum, the zero-crossing point of the first harmonic signal,

under ideal condition, gives the Bragg wavelength of the grating. For both techniques, the coherent multiple path effect would be largely reduced by use of proper low-pass filters, either analog or digital [43]. This is because the interferometric signal that has a $\sin(2\pi c\tau/\lambda_0+\phi)$ dependence varies periodically and would usually oscillate a number of periods when the laser wavelength (λ_0) is scanned across the FBG spectrum. However, there is a possibility that time delay difference (τ) between the reflected wave and the signal wave varies (due to environmental effect) in such a way that it partly cancels the periodic variation of $\sin(2\pi c\tau/\lambda_0+\phi)$ induced by the wavelength scanning. The interferometric signal may then varies at a similar frequency as the Bragg signal, and in the case, the low pass filtering technique would not be very effective.

A different situation is when the system is working in a "lock-in" mode where the laser wavelength is locked to a particular point (e.g., the peak) of the reflection spectrum by a closed feedback loop [44,45]. In this case, the effect of the interferometric signal is to affect the lock-in accuracy and the use of low-pass filter may not help. We have studied the modulation characteristics of the interferometric effects [38] and found that they can be reduced by proper wavelength modulation. Fig.5 shows the normalized wavelength detection error $|\Delta\lambda|/B_G$ (B_G is the spectral width of the grating) for the first derivative technique as functions of grating reflectivity (R_G) and delay time difference (τ) for a particular example where one and only one reflection point before the grating is assumed. The amplitude of wavelength dithering was assumed to be λ_m=0.5pm. For the case of high modulation frequency [f=100MHz Fig.5a], the error is generally large except for some isolated points corresponding to zeros of $J_1(\eta)=0$, where $\eta=(c\lambda_m\tau)/(f(\lambda_0)^2)\sin(\pi f\tau)$. The error signal is periodic and has a period of 10ns or 1m in terms of fibre length. This implies that reflections occurring at most of the positions along the fibre will make a significant contribution to the wavelength measurement error. For the case of low modulation frequency, [f=10kHz, Fig.5b], only reflections occurring very close to the FBG (τ=0, within a few meters) would contribute significantly to the measurement error (contribution from around $\tau=10^5$ns would not appear because it exceeds the coherence time of most practical semiconductor laser sources). Similar results were obtained for a reflection point after the grating. For multiple coherent reflections, the resulted wavelength detection error would be the summation of contributions from all the reflections. Experimental investigations were also conducted and the results obtained agree with that from the theoretical analysis [46].

3.2 Crosstalk performance in a TDM sensor array due to limited extinction ratio of the optical switch

For a multiplexed system using time division multiplexing, apart from the multi-path interference effect mentioned in section 3.1. An important performance-limiting factor is the inter-channel (sensor) crosstalk limited by the extinction ratio of the optical switch. The extinction ratio $(\alpha_e)^2$ of an optical switch is defined as $(\alpha_e)^2=P_{OFF}/P_{ON}$ or $10\log P_{OFF}/P_{ON}$, where P_{ON} is output power when the switch is ON and P_{OFF} is the power level when the switch is OFF. Consider the TDM system shown in Fig.6 and assume that the signal from channel one (sensor 1) is selected by use of an electric switch after photo-detection. The limited extinction ratio means that light signals from other channels would "leak" into channel 1 and thus result in crosstalk. The crosstalk may be divided into three categories: Incoherent power leakage from other channels to channel 1 (the incoherent crosstalk), the first order crosstalk caused by coherent mixing between signal from channel 1 and leakage waves from other channels, and the coherent mixing between the leakage waves from channels other than channel 1. We have simulated the crosstalk level as functions of both modulation and array parameters. Fig.7 shows the results of the incoherent crosstalk level in terms of normalized wavelength detection error $(|\Delta\lambda|/(B_G\ (\alpha_e)^2)$ as a function of sensor numbers for λ_m=1pm. For a commercial single Mach-Zehnder type integrated optic switch with an extinction ratio of -30dB, the crosstalk level is around 0.2pm for FBGs of spectral width of 0.1nm and for a sensor number of up to 100. Figs.8a and 8b show, respectively, the first order coherent crosstalk (in terms of normalized wavelength detection error $(|\Delta\lambda|/(B_G\ \alpha_e))$ as functions of the amplitude of wavelength modulation (λ_m) and time delay difference (τ) between adjacent channels for a sensor number of N=20. A crosstalk level of ~1pm may be achieved with a -30dB switch if the amplitude of the wavelength dithering is bigger than λ_m=1pm and time delay is greater than 100ns. Experimental studies were also carried out and the results will be reported at the conference.

4. APPLICATIONS OF FIBRE BRAGG GRATING SENSORS

The FBG sensors have been studied for a number of applications including the measurement of flow-induced structural vibrations [47-49] and the condition monitoring of concrete and composite structures [50,51]. Fig.9 shows an experimental setup for the measurement of lateral vibration of a cylinder within a flow field [47]. An FBG sensor was bonded at the mid-span along the cylinder at 90° from the leading stagnation line to measure strain (ε) due to the lateral displacement. The lateral fluctuating displacement (Y) at the same point was monitored by use of a commercial laser vibrometer so that

the relationship between the two quantities can be quantified. The streamwise fluctuating velocity (u) was measured by a single hot-wire.

Experiments were conducted for a range of the free stream velocity U_∞. Fig.10 shows the simultaneous time histories for a particular stream velocity (corresponding to Reynolds number Re=6100) obtained from the FBG sensor (ε signal, upper trace), the laser vibrometer (Y signal, middle trace) and the hot wire (u signal, lower trace).

The measured strain variation is consistent, in terms of the vortex shedding frequency and the natural frequency of the fluid-structure system, with the bending displacement measurement by the laser vibrometer. The fibre-grating sensor provides an alternative technique for the measurement of the flow-induced structural vibration. The fibre sensor, because of its very small size and lightweight, is relatively non-intrusive. Its attachment to a structure appears to have a negligible effect on the integrity and the vibrational characteristics of the structure. As light is guided by optical fibres which are flexible (can be bent), the FBG sensors may be used to measure fluctuating strain at any point on a structure even when it is positioned within an array, overcoming the drawback of the laser vibrometer. Using wavelength division multiplexing and/or other multiplexing techniques, the FBG sensors have potential to be used for simultaneous multi-point measurement on a single structure, or simultaneous measurement of many structures. These measurements are of significance in determining the instantaneous operational mode shape of a structure in a crossflow and in studying flow-structure interactions.

A second example is the use of FBG sensors for monitoring the internal strain of composite, concrete, and composite strengthened concrete structures [50,51]. Fig.11 shows an example where FBG sensors were used to monitor strains at three particular positions at the composite/concrete interfaces of a composite strengthened concrete beam specimen [51]. The three sensors were of the same nominal Bragg wavelength and shared the same source and processing optics/electronics by use of FMCW multiplexing. Fig.12 shows the strain measured from grating G3 (position 3) when the beam was subjected to a three point bending test as shown in Fig.11. For the purpose of comparison, the strain measured from a strain gauge attached on the surface of the reinforcements at the same location of the FBG sensor embedded is also showed.

There are also active research on applying fibre Bragg gratings for textile composite sensing in the Institute of Textile and Clothing (ITC) [52-55]. The work at ITC includes reliability study of FBG sensors embedded in the textile composites, discrimination between strain and temperature effect and etc.

5. SUMMARY

Some of the research work on FBG multiplexing and applications at the Hong Kong Polytechnic University have been reviewed. Our aim is to develop advanced multiplexed FBG sensor systems and apply them to mechanical sensing applications. Current work on FMCW is focused on examining the performance limiting factors and identifying anti-measures. For the system that uses narrow band laser sources, although the effect of coherent multiple reflections and crosstalk may be reduced by a combination of low pass filtering and proper wavelength dithering, significant experimental work needs to be done to implement a practical system. Both FMCW and TDM systems using laser sources have the potential of multiplexing a very large number of sensors. Our objective is to develop prototype instrumentation based on these principles. Other work on FBG sensors includes the development of multi-point dynamic strain sensor systems for flow-induced vibration measurement and for monitoring concrete/composite structures.

ACKNOWLEDGEMENTS

The author would like to thank research student Julian C. C. Chan and Peter C. K. Chan for their work on the TDM of FBG sensors by use of narrow band laser sources and on the FMCW multiplexing, respectively. The projects on applying FBGs for flow-induced vibration and in monitoring of composite/concrete are collaborated with Y. Zhou and L. M. Zhou and K.T. Lau of the Mechanical Engineering Department. The information on applying FBG sensors for textile structures is provided by X. M. Tao of the Institute of Textile and Clothing. The work described in this paper was partly supported by research grants from the Hong Kong Polytechnic University (Project Nos. V415, V390, and YB79) and a grant from the Research Grant Council of the Hong Kong Special Administrative Region (Project No. PolyU 5215/98E)

REFERENCES

1. K.O. Hill, Y.Fujii, D.C.Johnson, and B.S. Kawasaki, "Photosensitivity in optical fibre waveguides: application to reflection filter fabrication," Appl. Phys. Lett. Vol. 32, 647-649, 1978
2. W.W. Morey, G.A. Ball, and H. Singh, "Applications of fibre grating sensors," in Fibre Optics and Laser Sensors XIV, R.P. De Paula and J. W. Berthod III, (eds.), Proc. SPIE 2839, 2-7, 1996.
3. A.D. Kersey, M.A. Davis, H.J. Patric, M. LeBlanc, K.P. Koo, C.G. Askins, M.A. Putnam and E.J. Friebele, "Fiber Grating Sensors," Journal of Lightwave Technology, Vol.15, 1442-1462, 1997.
4. Y.J.Rao, "In-fibre Bragg grating sensors," Meas. Sci.. Technol. Vol.8, 355-375, 1997.
5. A. Othonos and K. Kalli, Chapter 7 in book "Fiber Bragg Gratings: Fundamentals and Applications in Telecommunications and Sensing," 301-396, Artech House, Norwood, 1999.
6. S. Chen., et al, " Digital spatial and wavelength domain multiplexing of fibre Bragg grating based sensors," Proc. of OFS-12, Williamsburgh, USA, 1997, 448-451.
7. C.G. Askins, M.A. Putnam, and E.J. Friebele, "Instrumentation for interrogating many-element fiber Bragg grating arrays," SPIE Vol.2444, 257-266.
8. S. M.Melle, K.Liu, R.M. Measures, "Practical fiber-optic Bragg grating strain gauge system," Applied Optics, Vol.32, 3601-3609, 1993.
9. M.A.Davis and A.D.Kersey, "All fiber Bragg grating strain sensor demodulation technique using a wavelength division coupler, Electronics Letters, Vol.30, 75-77, 1994.
10. Q. Zhang., et al. "Use of highly overcoupled coupler to detect shift in Bragg wavelength," Electronics Letters, Vol.31, 1510-1511, 1995.
11. A.D.Kersey, T.A. Berkoff, and W.W. Morey, "Fiber Fabry-Perot demodulation for Bragg grating strain-sensors," Proc. of OFS'9, Florence, Italy, 39-42,1993.
12. M.G.Xu, H. Geiger, J-L. Archambault, L.Reekie, and J.P. Dakin, "Novel interrogation system for fiber Bragg grating sensors using an acousto-optic tunable filter," Electron. Lett., Vol.29, 1510-1512, 1993.
13. D.A. Jackson, A.B. Lobo Ribeiro, L. Reekie and J.L. Archambault, "Simple multiplexing scheme for a fiber-optic grating sensor network," Optics Letters, Vol.15, 1192-1194, 1993.
14. G. A. Ball, W.W. Morey and P.K. Cheo, "Fiber laser source/analyzer for Bragg grating sensor array interrogation," IEEE Journal of Lightwave Technology, Vol.12, 1994, 700-703.
15. T. Coroy, et al., "Peak detection demodulation of a Bragg fiber optic sensor using a gain-coupled distributed feedback tunable laser," Proc. of OFS-12, Williamburg, USA, 1997, 210-212.
16. M.A. Davis, A.D. Kersey, "Application of a fiber Fourier transform spectrometer to the detection of wavelength encoded signals from Bragg grating sensors," Journal of Lightwave Technology, Vol.13, 1289-1295, 1995.
17. D.A. Flavin, R.McBride and J.D.C. Jones, " Short optical path scan interferometric interrogation of a fibre Bragg grating embedded in a composite," Electronics Letters, Vol.33, 319-321, 1997.
18. A.D.Kersey, T.A. Berkoff, and W.W. Morey, "High-resolution fiber-grating based strain sensor with interferometric wavelength shift detection," Electronics Letters, Vol.28, 236-238, 1992
19. A.D.Kersey, T.A. Berkoff, and W.W. Morey, "Fiber-optic Bragg grating strain sensor with drift compensated high-resolution interferometric wavelength-shift detection," Optics Letters, Vol.18, 72-74, 1993.
20. M.L.Dennis, et al., "Grating sensor array demodulation by use of a passively mode-locked fiber laser," Optics Letters, Vol.22, 1362-1364, 1997.
21. M.A. Putnam, M.L. Dennis, I.N. Duling III, C.G. Askins and E.J. Friebele, "Broadband square-pulse operation of a passively mode-locked fiber laser for fiber Bragg grating interrogation," Optics Letters, Vol.23, 138-140, 1998.
22. S.H. Yun, D.J. Richardson, and B.Y. Kim, Interrogation of fiber grating sensor arrays with a wavelength-swept fiber laser," Optics Letters, Vol.23, 843-845, 1998.
23. A.D.Kersey, T.A. Berkoff, and W.W. Morey, "Multiplexed fiber optic Bragg grating strain-sensor system with a fiber Fabry-Perot wavelength filter," Optics Letters, Vol.18, 1370-1372, 1993.
24. M. Volanthen, J.P. Dakin, et al., "Simultaneously monitoring of multiple fiber gratings with a single acousto-optic tunable filter," Electronics Letters, Vol.32, 1228-1229, 1996.
25. K.P. Koo and A.D. Kersey, "Bragg grating-based laser sensors systems with interferometric interrogation and wavelength division multiplexing," Journal of Lightwave Technology, Vol.13, July 1995.
26. R.S. Weis, A.D.Kersey and T.A. Berkoff, "A four-element fiber grating sensor array with phase-sensitive detection," Photonics Technology Letters, Vol.6, 1469-1472, 1994.
27. J.L. Brooks, B. Moslehi, B.Y.Kim, and H.J. Shaw, "Time-domain addressing of remote fiber optic interferometric sensor arrays, " Journal of Lightwave Technology, Vol.LT-5, 1014-1023, 1987.

28. M.A. Davis, D.G. Bellemore, M.A. Putnam, and A.D. Kersey, "Interrogation of 60 fiber Bragg grating sensors with microstrain resolution capacity," Electronics Letters, Vol.32, 1393-1394, 1999.
29. P.J. Henderson, D.J. Webb, D.A. Jackson, L. Zhang and I. Bennion, "Highly-multiplexed grating-sensors for temperature-referenced quasi-static measurements of strain in concrete bridges," Proc. of OFS'13, 320-323, Korea, April 1999.
30. T.A. Berkoff, et al., "Hybrid time and wavelength division multiplexed fiber Bragg grating sensor array," Proc. SPIE 2444, 288-294,1995.
31. K.P.Koo, A.B.Tveten, S.T. Vohra, "Dense wavelength division multiplexing of fiber Bragg grating sensors using CDMA," Electron. Lett., Vol. 35, 165-167, 1999.
32. C. Miller, T. Li, J. Miller, F. Bao and K. Hsu, "Multiplexed fiber gratings enhance mechanical sensing," Laser Focus World,119-123,March 1998.
33. A.J. Hymans and J. Lait, "Analysis of a Frequency-Modulated Continuous-Wave Ranging System," Proc. of IEE, Vol.107B, 365-372, July 1960.
34. R. Kist, "Point sensor multiplexing principles," Chapter 14, in Optical Fibre Sensors, Vol.2, B. Culshaw and J. Dakin, (eds), Artech House, 1989.
35. K.I.Mallalieu, R.Youngouist and D.E.N. Davies," FMCW of optical source envelope for passive multiplexing of frequency-based fibre-optic sensors," Electronics Letters, Vol.22, 809-810, 1986.
36. K.C.Chan, W.Jin, J.M.Gong and M.S.Demokan, "Multiplexing of fiber Bragg grating sensors using FMCW technique," IEEE Photonics Technology Letters, in press
37. P. K.C. Chan, W.Jin and M.S.Demokan, "Frequency domain technique for multiplexing of fiber Bragg grating sensors," SPIE Conference on Advanced Photonic Sensors and Applications ISPA'99, Paper no: 3897-62, 29 Nov.-3 Dec. 1999, Singapore.
38. W.Jin, "Investigation of interferometric noise in fiber optic Bragg grating sensor by use of tunable laser sources," Appl. Opt.,37, 2517-25, 1998.
39. R. Hjelme, L. Bjerkan, S. Needgard, J. S. Rambech and J. V. Aarsnes, "Application of Bragg grating sensors in the characterization of scaled marine vehicle models", *Appl. Opt.* 36, pp. 328-336, 1997
40. J. Dunphy, G. Ball, F. D. Amato, P. Ferraro, S. Inserra, A. Vannucci and M. Varasi, "Instrumentation development in support of fiber grating sensor arrays", *Proc. SPIE* 2071, pp. 2-11, 1993.
41. Hassaun Jones-Bey, "Tunable fiber laser yields 11-nm range," 20-21 Laser Focus world, July 1999.
42. W.Jin, "Performance analysis of a time-division-multiplexed fiber optic gas sensor array using wavelength modulation of the DFB laser," Applied Optics, 1 September 1999, in press.
43. C.C.Chan, W.Jin, and M.S.Demokan, "Enhancement of measurement accuracy in fiber Bragg grating sensors by using digital signal processing," OSA Technical Digest Series, Vol.6, 311-312, CLEO'98, San Francisco, May 3-8, 1998.
44. M. G. Xu, H. Geiger, J. P. Dakin, "Modeling and performance analysis of a fibre Bragg grating interrogation system using an acousto-optic tunable filter", *J. Lightwave Technol.* 14, n11, pp. 391-396, 1996.
45. B. Lissak, A.Arie and M. Tur, "High resolution strain sensing by locking lasers to fiber-Bragg gratings", in *European Workshop on Optical fiber Sensors*, Brian Culshaw, Julian D.C. Jones, Editors, Proceedings of SPIE Vol.3483, 250-254, 1998.
46. C. C. Chan , J. M. Gong , W. Jin and M. S. Demokan, "Minimization of interferometric noise in a fiber Bragg grating sensor using a tunable laser and a first derivative interrogation technique", Proc. of SPIE Vol. 3746, OFS-13, pp.534-538, Kyongju, KOREA, April 12-16, 1999.
47. W. Jin, Y. Zhou, P.K.C. Chan and H.G. Xu, "A fibre optic grating sensor for the study of flow-induced vibrations," Sensors and Actuators, in press.
48. Y. Zhou, R.M.C. So, W.Jin, H.G. Xu and PETER K.C. Chan, " Measurement of the fluctuating strain on a cylinder in a crossflow using a fibre grating sensor," Experiments in Fluids, in press.
49. Y. Zhou, R.M. C. So, H.G. Xu, and W. Jin, "Interference between two cylinders in a uniform stream," Proc. of 13th Australian Fluid Mechanics Conference, Monash University, Melbourne, 13-18 December 1998. 799-802.
50. K.T. Lau, L.M.Zhou, C.H.Zhou, C.H.Woo, K. C. Chan and W.Jin, "Strengthening and monitoring concrete structures using glassfibre composites and fibre-optic Bragg grating sensor," Materials Science Research International, in press.
51. P. K. C. Chan, W. Jin, K. T. Lau, and L. M. Zhou, "Multi-point strain measurement of composite-bonded concrete materials with a FWCM multiplexed FBG sensor array," in preparation for publication.
52. W. Du, X.M. Tao, H.Y. Tam and C.L. Choy, Fundamentals and applications of optical fiber Bragg grating sensors to textile structural composites," Composite Structures, Vol.42, 217-229, 1998.

53. L. Tang, X. Tao and C-L. Choy, "Effectiveness and optimization of fiber Bragg grating sensor as embedded strain sensor," Smart Mater. Struct. Vol.8, 154-160, 1999.
54. L. Tang, X. Tao W.-C. Du and C.-L. Choy, " Reliability of fiber Bragg grating sensors embedded in textile composites," Composite Interfaces, Vol.5, 421-435, 1998.
55. W. Du, X. Tao and H.-Y. Tam, Temperature independent strain measurement with a fiber grating tapered cavity sensor," IEEE Photonics Technology Letters, Vol. 11, 596-598.

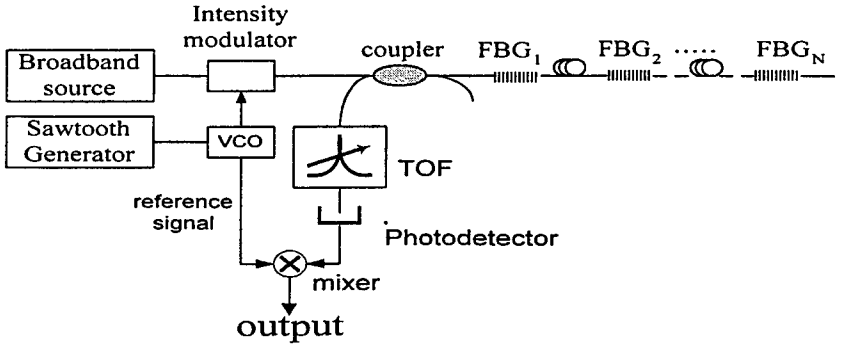
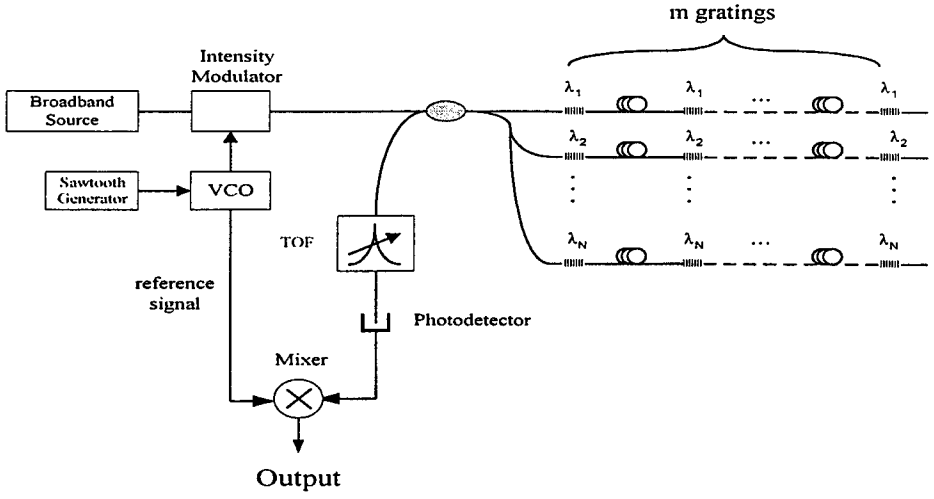
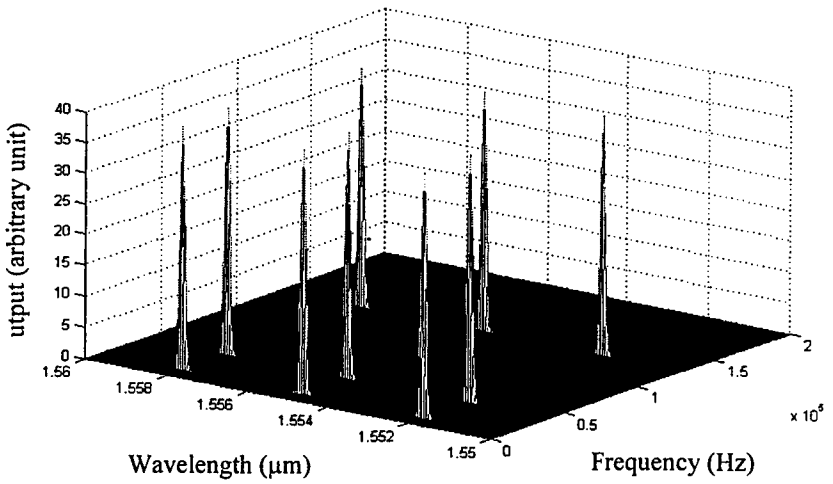


Fig.1 FMCW multiplexed serial-connected single fibre array



(a)



(b)

Fig.2 (a) Hybrid FMCW/WDM parallel FBG sensor array (b) System output in the frequency-wavelength domain

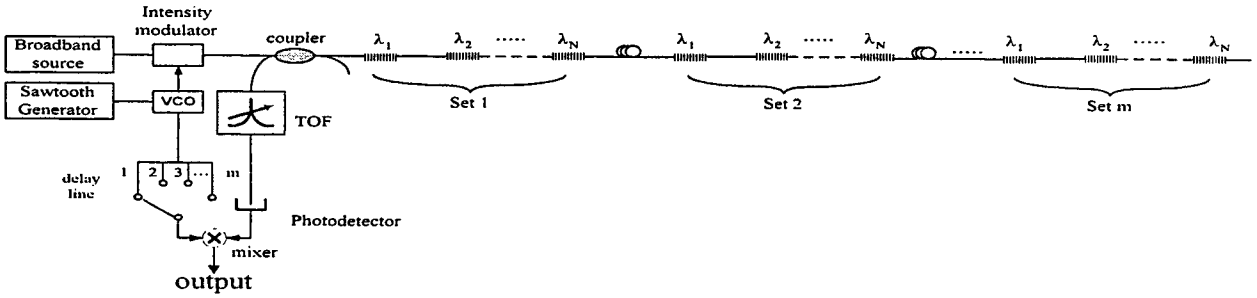


Fig.3 Hybrid FMCW/WDM serial FBG sensor array

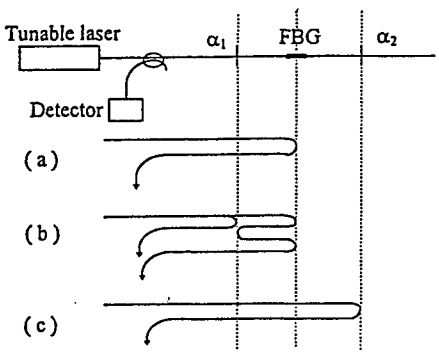


Fig.4 Interrogation of an FBG by a tunable laser source. (a) Reflection from the FBG. Reflection from a point before (b) and after (c) the FBG

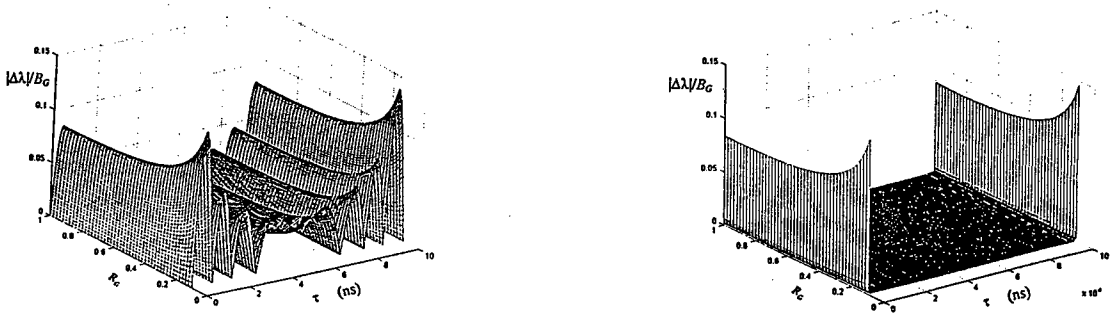


Fig.5 Normalized wavelength detection error as a function of R_G and τ. λ_m=0.5pm. Modulation frequency (*left*) 100MHz (*right*) 10kHz.

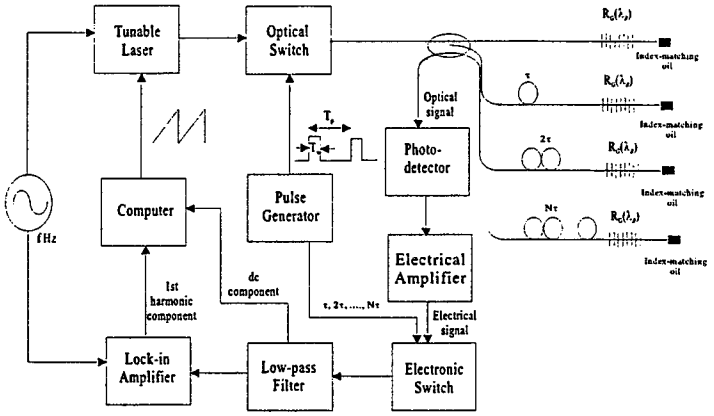


Fig.6 TDM of FBG sensors by use of a tunable laser source--Parallel topology

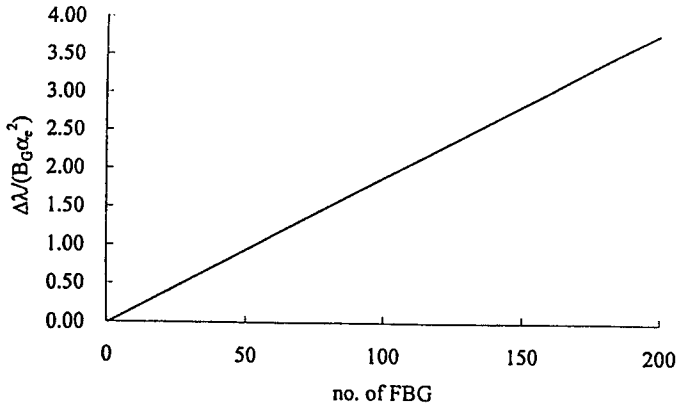


Fig.7 Incoherent crosstalk. Normalized wavelength detection error as a function of sensor number (λ_m=1pm)

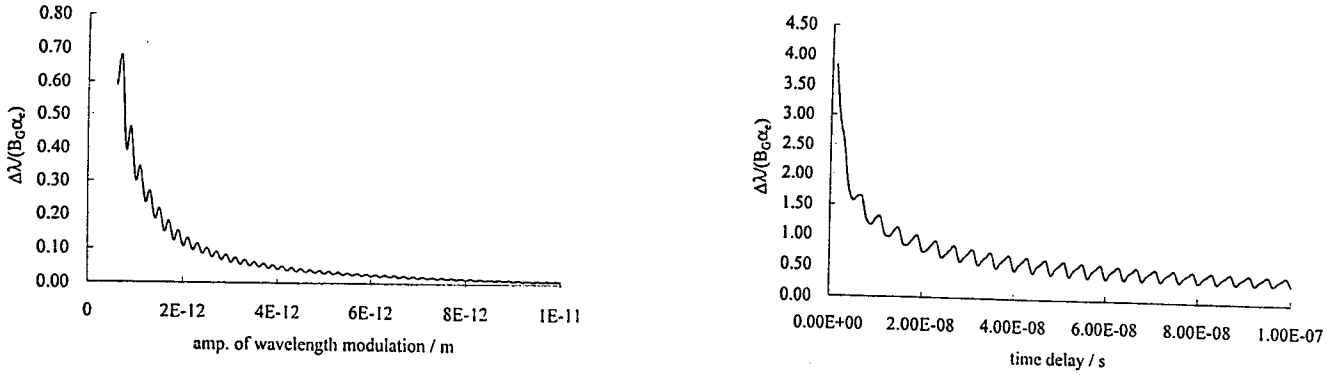


Fig.8 First order crosstalk for N=20. Normalized wavelength detection error as function of (*left*) amplitude of wavelength modulation with τ=100ns; (*right*) time delay difference between adjacent sensors with λ_m=1pm.

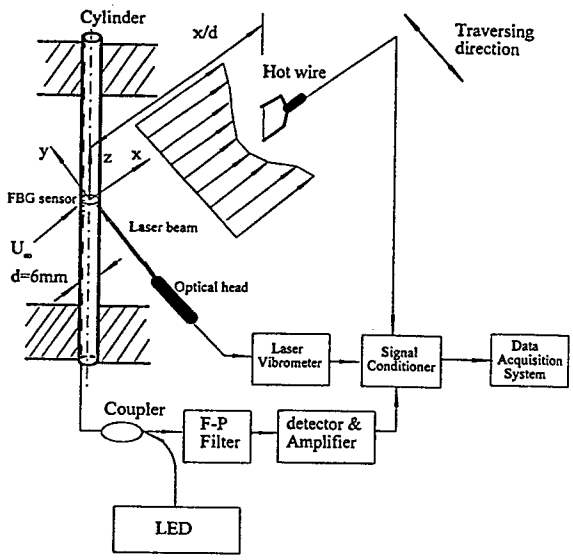


Fig.9 Experimental setup for flow induced vibration measurement

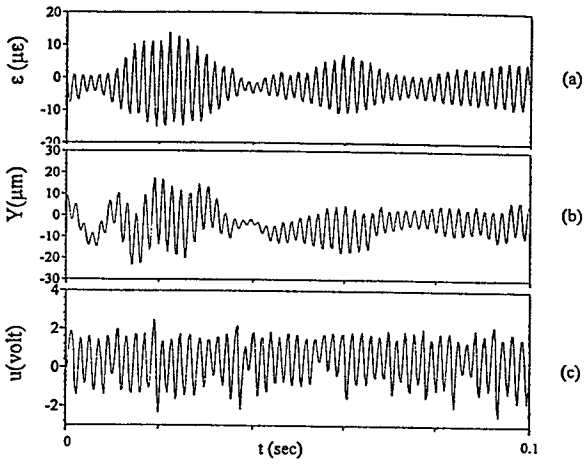


Fig.10 Time traces of (*a*) the fluctuating strain ε measured from the FBG sensor, (*b*) the bending displacement Y from laser vibrometer, (*c*) the streamwise velocity u from hot-wire (*Re*=6100).

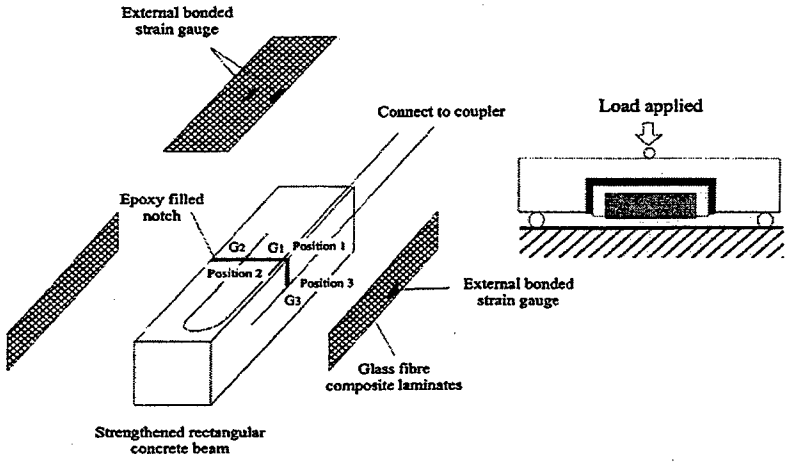


Fig.11 Schematic of the test specimen showing the location of the FBG sensors and the loading setup

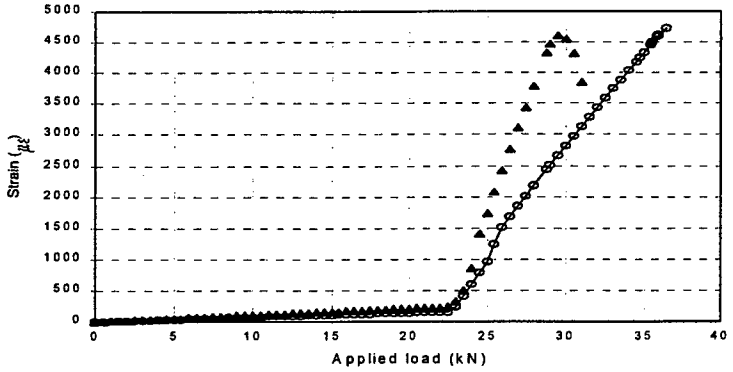


Fig.12 Results of strain measurement (▲) strain measured by grating 3 (○) strain measured from a strain gauge attached at the surface

An Approach to Analysis and Measurement of Large Array Bragg Grating Sensors

T.Srinivas, Indu Shekhar Das, A.Selvarajan

Department of Electrical Communication Engineering
Indian Institute of Science, Bangalore 560 012 INDIA
Tel: +91 80 3092852, Fax: +9180 3340563
Email: srinu@ece.iisc.ernet.in

ABSTRACT

Arrays of Fiber Bragg Gratings (FBG) can be embedded in engineering structures to study the effect of various deformations and perturbations and thus enabling its health monitoring. In this paper, two multiplexing and detection schemes useful in the case of high resolution Bragg Gratings are presented. Instead of the conventional detection schemes, like an array of non- uniform grating and wavelength based addressing or an array of uniform grating with time-delay based addressing, we suggest a non-uniform cluster of uniform gratings. A continuous wave or pulse train can be input to address and distinguish the gratings from each other. The first technique uses continuous wave input and measures the reflected amplitude response, while in the second technique we input pulsed light and measure the height as well as width of the reflected pulse. In both the cases, a minimum response time is suggested for the measurement to be made. This is the time when response from the last grating is also obtained and a high degree of saturation leading to stable condition is obtained. We have combined the results for a single grating using the coupled mode theory with an equivalent matrix operator for the array to obtain an effective and simple mathematical tool to handle the large array problem.

1. INTRODUCTION

Optical techniques have multiple roles to play antages in smart structure applications[1]. The Bragg grating seems to be the answer to many requirements of smart structures, in particular health monitoring[2,3]. It has the structural integrity of an optical fiber, it does not cause higher stress concentrations, and reflectivity and bandwidth of FBG can be tailored to a large extent. An array of Fiber Bragg Gratings, with each FBG located at certain identified points, supporting either same or different Bragg wavelength (uniform or non-uniform grating, respectively) forms the typical sensing network used in smart structures. For both the cases, the response of each grating plays a significant role in designing the network and deciphering the output signal.

One of the difficulties in the use of array of gratings is due to the bandwidth and instrumentation limitations[4,5]. Bandwidth does not allow us to put a large number of gratings, which is often required for a distributed measurement. And the common technique to measure the output and distinguish each grating from the others is either time delay or wavelength shift. If we use the grating of uniform pitch, the only information, which differentiates the grating elements, is the time delay. For a typical separation like 10 cm, the time delays between pulses getting reflected by two consecutive gratings is about 0.5 nsec. The distortion in grating structure due to a perturbation is stored in the pulse height (or continuous wave amplitude) and measuring these two simultaneously rather puts a stringent requirement from the instrumentation point of view.

For the case of non uniform grating array, each grating element can pick one wavelength and so one need to measure only the amplitude, or say, monitor the envelope function of reflected light intensity. A distortion in one or some of the gratings will off-set this function and can be detected very easily. But then each grating has its own response profile i.e. spread, side-lobes, and tail shape, which together decides the maximum number of gratings that can be accommodated in a given source bandwidth.

Part of the SPIE Conference on Advanced Photonic Sensors and Applications
Singapore • November–December 1999
SPIE Vol. 3897 • 0277-786X/99/$10.00

2. ANALYSIS OF FBG CLUSTERS

A combination of the above two is expected to overcome the difficulties and will give a better trade-off between the number of gratings, required bandwidth and measurement options. Here we explore the possibility of putting a series of uniform gratings in one array, each series characterized by its own particular wavelength or pitch. So, the array may have, say 'n' gratings of uniform pitch, followed by next 'n' gratings of different pitch and so on. Thus the whole array of gratings will look like the one shown in the Fig. 1. The direct advantage gained by this configuration is that for a given bandwidth the number of gratings that can be accommodated is directly increased by n times.

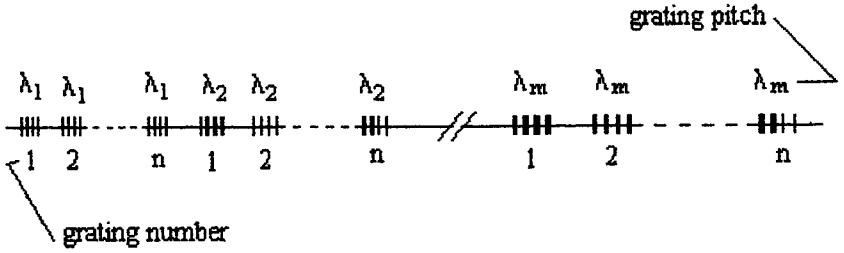


Fig.1 An array of non-uniform clusters of uniform gratings.

Out of a total 'n' uniform gratings, perturbation because of various reasons may occur to any extent to any of the gratings. Therefore measurement should be such that it should reveal the amount and location of perturbation. So, it necessarily involves finding out two variables, say i, denoting the site and Δr, denoting the change of reflectivity or extent of distortion. But then, it is evident that a single measurement can not reveal both these things. We suggest two ways to resolve this.

We have a series of gratings, whose individual elements act as a reflector and transmitter. So every ray which comes into this array after traversing one integrating separation, say, L, breaks into two beams forward propagating and backward propagating. For two or three grating case, and for the initial two or three round trip times, it appears trivial to manually calculate and find out the beam propagation. But once the number of gratings is large or if we are interested in some five or more round trips, handling each and every ray is very difficult. Denoting the entire array by one matrix, such problems can be drastically simplified and can be handled safely.

Each ray in the grating has got four distinguishable properties, namely, amplitude, phase, the propagation direction and the site of origin. Based upon these four properties, we can build a basis of rays which will be able to span the entire input-ray-space. This basis consists of 2 (n + 1) rays, (n + 1) rays forward propagating, each originating from a different site and another (n + 1) rays backward propagating, again each originating from a different site. The grating matrix, which has to work as output operator on this input ray vector, should have a 2(n + 1) by 2(n + 1) matrix, where each element, say, (i, j) gives the i^{th} output component obtained due to j^{th} input component. These elements store the reflectivity, transmittivity and the phase information of that ray.

One typical array of 'n' gratings having $r_1, r_2, \ldots r_n$ reflectivity, and $L_1, L_2, \ldots L_n$ inter-grating separation will look like the following. A=

$$\begin{bmatrix} 0 & \cdots & 0 & 0 & \cdots & 0 \\ T_1 e^{i\beta L_1} & & & R_1 e^{i\pi/2} & & \\ & \ddots & & & R_2 e^{i\pi/2} & \\ & T_{n-1} e^{i\beta L_{n-1}} & & & \ddots & \\ & & T_n & & & R_n e^{i\pi/2} \\ R_1 e^{i\pi/2} & & & T_1 e^{i\beta L_1} & & \\ & \ddots & & & \ddots & \\ & & & & T_{n-1} e^{i\beta L_{n-1}} & \\ & & R_n e^{i\pi/2} & & & T_n \\ 0 & \cdots & 0 & 0 & \cdots & 0 \end{bmatrix} \quad (1)$$

where the input ray vector has first (n + 1) forward propagating paths and (n + 1) backward propagating paths being counted from the left side.

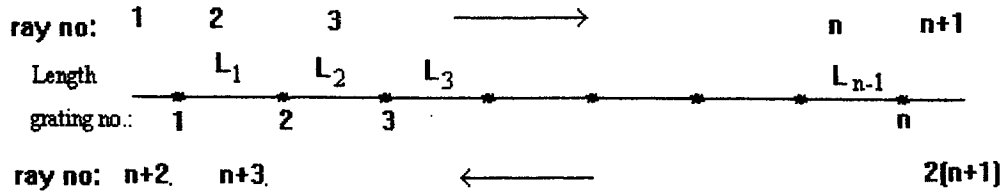


Fig. 2 Localized rays forming the basis of grating-array output

One point to note here is that the first and last rows of the Matrix A are zero. This is because we are assuming the output to be a result of only reflection and transmission, i.e., we are not assuming any source or sink inside the grating array. This implies that the ray number 1 and 2(n+1) can not be generated by the successive transmission and reflection and so the corresponding row entries are found to be zero.

At time t = 0, ray 1 strikes the grating array, so input vector I = {1, 0, 0, ... 0}. The output after one time slot will be O = A.I. The output after second time slot will be A^2. I and so on.

Having got this matrix, we move on to see the two approaches for analyzing the perturbation.

3. TWO WAVE TECHNIQUE

In this case, two wavelengths, say λ_1 and λ_2 close enough to give a suitable reflectivity are chosen. Light of wavelength λ_1 is input and the output amplitude is monitored for saturation. This value is noted down. A change in this value means the onset of abnormality. The new value of R_{max} is noted down and then light of another wavelength, say λ_2 is passed. Again the saturated value for λ_2 is noted down. These two together are expected to yield the required information.

Selection of λ_1 and λ_2 is a trade off between theoretical and experimental ease. We know that r vs λ curve is a sharply peaked curve at $\lambda = \lambda_{Bragg}$. Taking two wavelengths, at either sides of λ_{Bragg} would increase $\Delta\lambda$, a favorable experimental way, but then difference between r_1 and r_2, would be minimal and so will be the two values of R_{max}, which amount to a decreased value of sensitivity. Putting $\lambda_1 = \lambda_{Bragg}$ though decreases $\Delta\lambda$, increases Δr and so this difficulty can be overcome. Another advantage in choosing $\lambda_1 = \lambda_{Bragg}$ is the convenience of detecting the change in reflectivity explained mathematically as follows.

λ_1 and λ_2 are the two wavelengths, such that $\lambda_2 - \lambda_1 << \Delta\lambda$.

Let, $$R_{max}(\lambda) = f(r_1(\lambda), r_2(\lambda), ... r_n(\lambda)) \quad (2)$$

Now for uniform gratings and in the absence of any external perturbation,

$$R_{max}(\lambda) = f(r(\lambda)) \quad (3)$$

After perturbation, one of the grating, say i^{th}. one, changes its reflectivity to $r_1 + \Delta r_1$, then

$$R_{max}(\lambda) = f(r_1(\lambda), \Delta r(\lambda), i) \quad (4)$$

The two wavelengths, λ_{Bragg}, $\lambda_{Bragg} + \Delta\lambda$, r-values and the values of R_{max}'s are known. The amount of perturbation is constant but for different wavelengths the corresponding transformation of reflectivity is different. A general dependence between these two may be shown as follows.

$$r = r(\lambda, S) \quad (5)$$

where S is some strain, i.e. some external perturbation. Differentiation of equation (5) gives the contribution of change in λ and S in the change in r as shown below.

$$\therefore \Delta r = \partial r / \partial \lambda |_{\lambda_0, S} \Delta\lambda + \partial r / \partial S |_{\lambda_0, S} \Delta S \quad (6)$$

here $\lambda = \lambda_{Bragg}$, $\Delta\lambda = 0$ and only the strain is applied. So,

$$(\Delta r)_1 = \partial r / \partial S |_{\lambda_0, S} \Delta S. \quad (7)$$

When $\lambda = \lambda'$, $\Delta\lambda \neq 0$ and so

$$\Delta r = (\Delta r)_1 + \partial r / \partial \lambda \,|\, \lambda_{Bragg} \Delta\lambda \ . \quad (8)$$

Now we know that λ_{Bragg} is that wavelength where the reflectivity is maximum. So $\partial r / \partial \lambda \,|\, \lambda_{Bragg} = 0$. This is another advantage of choosing $\lambda_1 = \lambda_{Bragg}$. Varying the three parameters viz. the number of grating elements 'n', the amount of perturbation and location of perturbation, the output saturated value can be studied. Variation of bounded output with individual grating reflectivity for a n=4 element FBG array is shown in Fig 3. Saturated output R_{max} is defined when the output power change only beyond third decimal point upon the reception of next pulse. For number of FBG elements n = 4, and individual grating reflectivity R = 0.7 it was found that the output saturates within 24 pulses. Once this value of or Bounded output is known, instead of measuring R_{max} again and again we can directly measure R_{max} after, say, 25 pulses. The two R_{max} values for the two λ's can be determined graphically which gives the required information about the perturbation.

4. PULSE HEIGHT-WIDTH PROCEDURE

Instead of going for two wavelength measurement another way could be to excite the grating array with a light pulse train and measure the effect on pulse height and width due to external perturbation.

The validity of this technique lies in the fact that a pulsed input is equivalent to a spectrum on the frequency or wavelength domain. For a 1550 nm, 10 ps pulse, the spread is about 1nm which lies very much within the grating response curve. This property is utilized to analyze the grating behavior. Different frequency components of the pulse sees different reflectivity and so out-bounds the output at different values at different time intervals.

Out of the available frequency band, if we see that the reflectivity varies between r_{min} to r_{max}, we can find out the corresponding number of pulses required. Any observation beyond this time will yield a saturated output and so the entire output spectrum will be a fixed one. i.e. it would not change with time. On the time domain it would mean to yield a saturated height and saturated width pulse. A rectangular pulse is uniquely characterized by two parameters, pulse width and pulse height. The pulse height and width on the frequency domain, does have the information of these two and therefore can uniquely determine the amount of perturbation and the location of perturbation.

5. RESULTS AND DISCUSSIONS

In case of monochromatic CW input (two wave technique), different (wavelength) input saturates at time-intervals as they see the different reflectivity. Fig. 4 shows the value at which they saturates. Fig. 5 shows the response of a grating which is phase matched at 1550 nm. Based on these curves and the phase matching condition; two suitable wavelengths can be chosen, whose reflectivity varies by an amount suitable to give enough difference in total saturated output values. Firstly the Bragg wavelength, say 1550 nm in this case, can be used as an 'analyzing signal'. Upon the initiation of abnormality, the reflected power level changes. This may correspond to different sites of defect. Then a 'confirmation beam' of second wavelength is input and its saturated output is noted. From these two values the required site and amount of defect. This can be done graphically by pre-calibrating a set of curves of bounded reflectivity vs individual grating reflectivity for perturbation at different locations, such as Fig. 6

When a square pulse is input (pulse height-width technique), interaction with the series of grating gives a bounded output Fig 7(a) and Fig. 7(b). The change in the pulse shape amounts to spread in the frequency domain which can be established

by taking the Fourier transform of the input pulse and the perturbed pulse. The spectrum of the input pulse and that f perturbed pulse are shown in Fig. 8(a) and Fig. 8(b) We can see that the pulse broadens in height as well as in width. Changed value is a characteristics of the transfer function Now as we did for a CW, we can here also, compute the new pulse height and width, for different sites and amount of perturbation and then can calibrate the pulse height and width so as to yield the required information.

6. REFERENCES

1. Measures R M, "Advances towards fiber optic smart structures", *Optical Engg.*, **31**, 1,pp.34-47, 1996.
2. William W. Morey, Gray A Ball and Singh Harmeet, " Applications of fiber grating sensors", pp. 2-17, **2839**, Proc. SPIE on fiber optics and laser sensors, XIV,Denver, USA, 1996
3. Das I .S, Srinivas T, and Selvarajan A, " Analysis of fiber Bragg graing sensor for smart structures". Conference on Smart materials, structures and integrated systems, Dec. 10-13,1997, Adelaide, Austraila.
4. D. Jackson, "Simple multiplexing schemes for a fiber optic grating sensor network", *Optics Lett.*, **18,** pp.1193, 1993.
5. Xu M G, "Novel interrogation systems for fiber Bragg graiting sensors using an acousto optic tunable filter", *Elect. Letters*, **.29,** pp.1510, 1993.

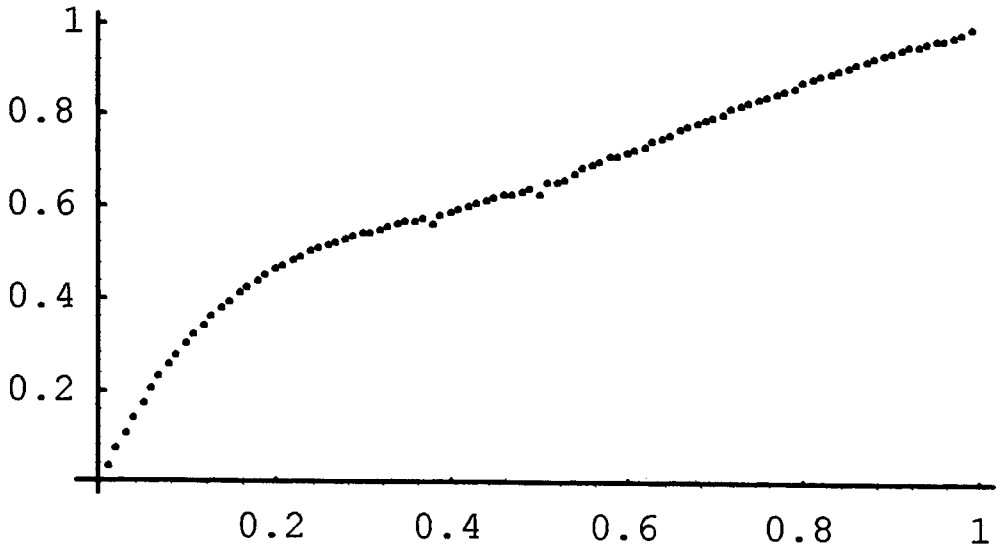


Fig. 3 Bounded Output Vs. individual grating reflectivity

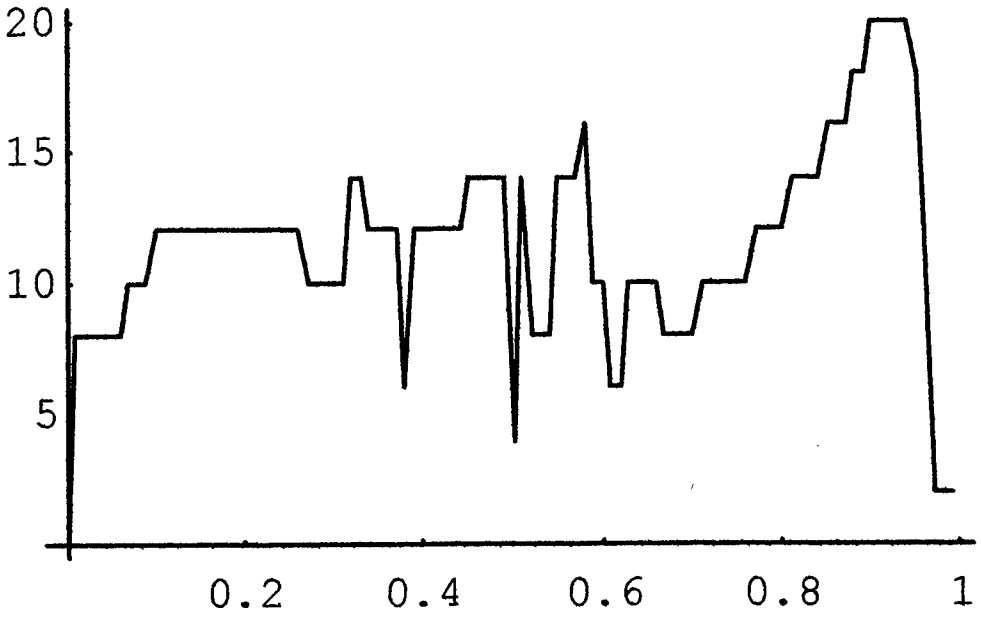


Fig. 4 Bounded Output Vs. no. of pulses

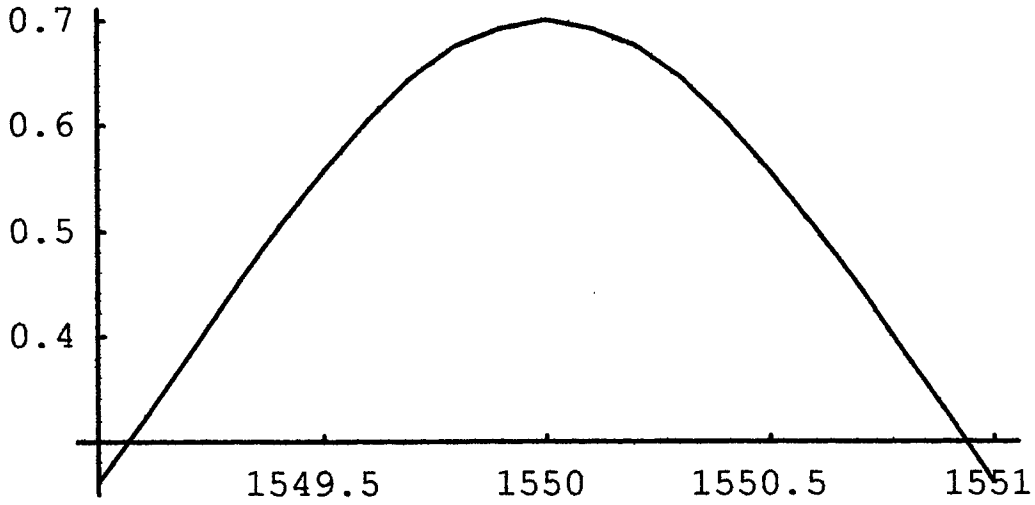
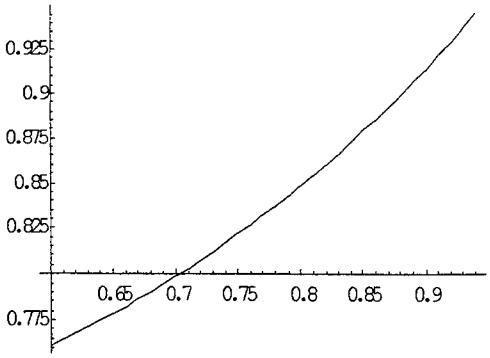
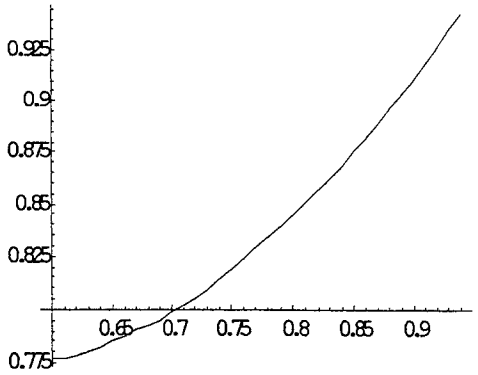


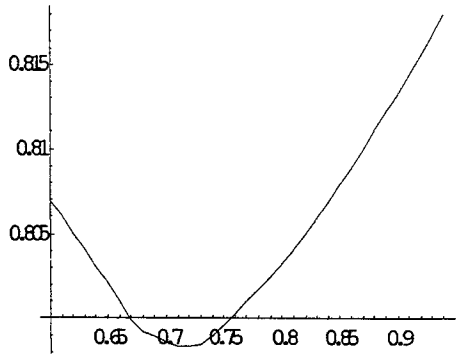
Fig. 5: Transfer Function of a single grating



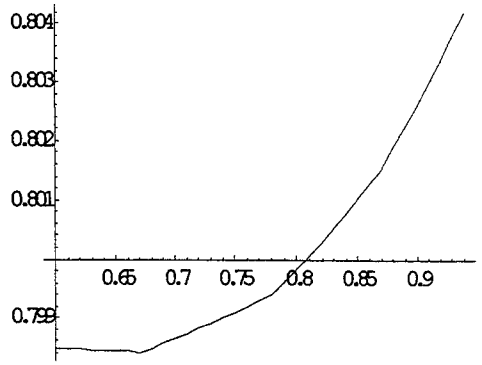
(a) First Grating Perturbed



(b)Second Grating Perturbed



(c) Third Grating Perturbed



(d) Fourth Grating Perturbed

Fig.6 Total bounded reflectivity Vs. individual grating reflectivity

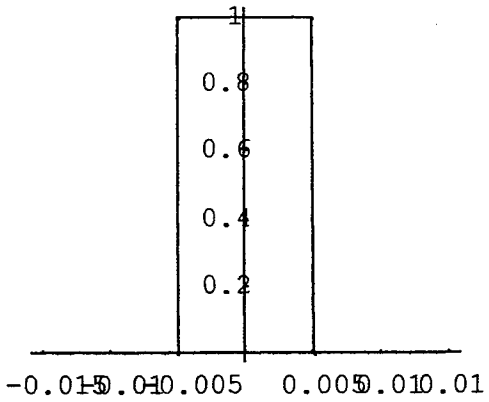


Fig. 7(a) Input Wave Pulse
width = 0.02 ns

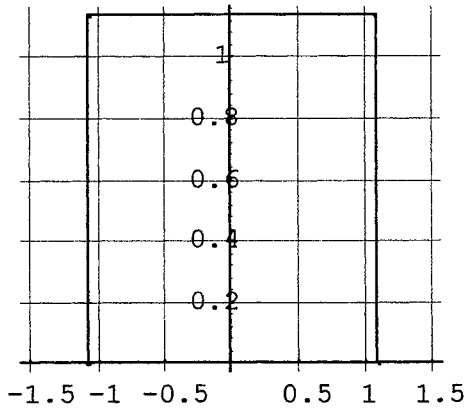


Fig. 7(b)Output Pulse Shape
(normalized to input pulse width)

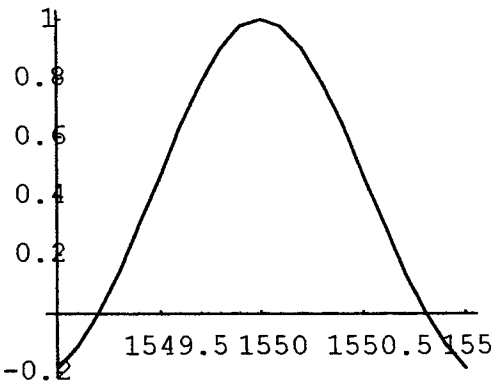


Fig. 8(a) Input wave on Frequency domain

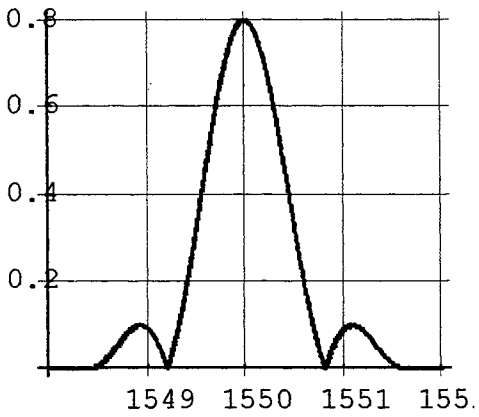


Fig. 8(b) Bounded Output of a Series of Grating on Frequency domain

Frequency domain technique for multiplexing of fiber Bragg grating sensors

Peter K. C. Chan*, W. Jin and M. S. Demokan

Department of Electrical Engineering, The Hong Kong Polytechnic University
Hung Hom, Kowloon, Hong Kong
Tel: +852 27666190, Fax: +852 23301544, e-mail: eekcchan@ee.polyu.edu.hk

ABSTRACT

We report on the use of a frequency-domain reflectometry technique for multiplexing fiber Bragg grating (FBG) sensors. This technique is based on the modulation of light intensity from a broadband source by a swept-frequency RF carrier. Signals from the FBG sensors located at different positions in an array are separated in the frequency-domain and demodulated using a tunable optical filter. A three FBG sensor system is experimentally demonstrated. The potential of the technique for multiplexing a large number of FBG sensors is discussed.

Index terms: optical fiber sensors, fiber Bragg gratings, frequency division multiplexing.

1. INTRODUCTION

Fiber Bragg Gratings (FBGs) have been identified as very important sensing elements especially for strain measurement in smart structures [1]. In many applications, arrays of FBG sensors are required for multi-point or quasi-distributed measurements; multiplexing of FBG sensors is therefore essential in order to reduce the cost per sensing point and to increase the competitiveness of FBG sensors against conventional electrical sensors.

The most popular technique for multiplexing FBG sensors is the wavelength division multiplexing (WDM) technique [2]. The maximum sensor number that can be multiplexed using the WDM technique is determined by the ratio of the source spectral width over the spacing between the Bragg wavelengths of the

* Correspondence: Email:eekcchan@ee.polyu.edu.hk; Tel.: 852+27666190; Fax:852+23301544

Part of the SPIE Conference on Advanced Photonic Sensors and Applications
Singapore • November–December 1999
SPIE Vol. 3897 • 0277-786X/99/$10.00

FBGs. For applications requiring a larger multiplexing gain, a combination of WDM and time-division multiplexing (TDM) and/or spatial division multiplexing (SDM) techniques may be used.

A code-division multiple access (CDMA) scheme has recently been demonstrated for dense wavelength division multiplexing of FBG sensors [3]. The CDMA technique has the advantage of larger average sensor output power over the TDM technique and consequently better signal to noise ratio. Although a two sensor array was demonstrated experimentally with a channel isolation (cross talk) of >20dB. A larger number (>100) of sensor and a better channel isolation is expected by increasing the code sequence length.

In this paper, we report on the multiplexing of FBG sensors using a frequency modulated continuous wave (FMCW) approach [4,5]. The FMCW technique provides a similar average power level at the photo-detector but possible lower cost compared with the CDMA approach because the FMCW technique uses simpler electronic processing circuitry.

2. PRINCIPLE OF OPERATION

Fig.1 shows the concept of FMCW technique where a triangular frequency sweeping is assumed. A time difference between a triangular frequency-swept reference waveform and a signal (a delayed version of the reference) produce a difference frequency (beat frequency f_{beat}) that is equal to the product of the rate of frequency excursion D and the time delay difference τ between two waves; i.e. $f_{beat}=D\tau=(2\Delta f f_s)\tau$ [4]. Here Δf is the frequency excursion; $f_s=1/T_s$, T_s is the period of the triangular sweeping. The resultant output beat note has a line spectrum at intervals of f_s. The position of the peak in the envelope of the line spectrum gives the beat frequency f_{beat} [5].

The detailed spectral characteristics of the beat note for a particular set of parameters (i.e., Δf=40MHz, f_s = 5kHz) have been studied by means of computer simulation. These parameters are the approximate values used in our experiments. It was found that, for small τ (e.g., τ < 400ns), the beat note spectrum is essentially a single line if the beat frequency coincides with one of the harmonics of the sweeping signal, i.e., $f_{beat}=D\tau=2f_s\Delta f\tau=mf_s$, or $2\Delta f\tau=m$, where m is an integer. The simulation results suggest that the magnitudes of the sidelines are -48dB below that of the central line. If N sensors are to be multiplexed with their respective time delay difference τ_m (m=1, 2,..., N) chosen in such a way so that the respectively beat frequencies satisfying condition $f_{beat_i} = mf_s$ (m=1, 2,..., N), the cross talk (determined by the magnitude of the spectral sidelines) between any two of the sensors should be below –48dB. The number of sensors that can be multiplexed is 400/12.5=32. When combined with the WDM technique, the FMCW technique should be capable of potentially multiplexing a few hundreds of FBGs with the exact number to be determined by the ratio of the source spectral width over the sensor dynamic range. The

aforementioned condition requires an equal time delay difference between adjacent sensors of 12.5ns or a physical separation of 1.25m between adjacent FBGs. The magnitude of the side lines and hence the cross talk level would increase if the physical separation is biased from the optimal value, i.e., 1.25m.

Fig. 2 shows an example of a three-sensor array using the FMCW technique. Light from a broadband source is modulated with a triangular swept frequency carrier generated from a voltage controlled oscillator (VCO) and launched into the FBG sensor array. The reflected signals from FBGs are guided back to a tunable optical filter (TOF), then a photo-detector, and mixed with a reference signal from the VCO subsequently. The system output will consist of a number of beat notes with their beat frequencies determined by the delay differences between the FBG sensor signals and the reference signal. The magnitudes of the beat signals are proportional to the convolution of the individual grating reflection spectrum and the transmission spectrum of the tunable filter. The Bragg wavelength of the individual FBGs can be interrogated by scanning the TOF and record the control voltage of the TOF that corresponds to the peak of the different frequency components.

3. EXPERIMENTS AND RESULTS

Experiments were conducted using the system shown in Fig.2. Light from an Er-doped fiber ASE source was modulated in intensity through a VCO driven integrated optic modulator by a triangular frequency-swept carrier from about 58 MHz to 99 MHz with 5 kHz repetition rate. The TOF used was made by JDS with linewidth 0.37 nm and a voltage-tuning coefficient of 2.95 nm/Volt. All FBGs were manufactured in our laboratory by using the same phase-mask of nominal Bragg wavelength of 1545.61 nm. The spectral widths of the FBGs are 0.17nm. The reflectivity of grating G_1, G_2 and G_3 are approximately 40%, 100% and 40% respectively. The spatial separations between G_1 and G_2 and between G_1 and G_3 were arbitrarily chosen to be around 5m and 12m, respectively. No effort was made to control the OPD between sensors to match the beat frequencies to the harmonics of the frequency-sweeping. The electrical signal produced by the photo-detector was mixed with the reference signal from the VCO. The beat signals generated from the mixer were observed from an electrical spectrum analyzer. Fig.3 shows a typical display of the spectrum analyzer. The three major peaks at 50kHz ($10f_s$), 75kHz ($15f_s$) and 105kHz ($21f_s$) correspond to sensor G_1, G_2 and G_3, respectively.

The sensor system was used for strain measurements. During the measurements, the three FBG sensors were strained independently by translation stages and the TOF was controlled through a computer by applying a slow varying voltage, via a 14-bit D/A converter. The system has a minimum resolvable wavelength shift of approximately 1.8 pm corresponding to 1.5 με applied strain. To demonstrate that this technique can be used to multiplex FBG sensors with the same Bragg wavelength, grating G_1 was shifted to 1546.29nm and used as a reference grating. The amplitude of the beat frequencies from G_2 and G_3 (at 75 kHz and 105 kHz) were monitored while gratings G_2 and G_3 were being strained from 0 to 1500 με. The

measurements were repeated 10 times for each applied strain and an average was taken as the TOF control voltage. Fig.4 shows the relations between the measured control voltages of the tunable filter and the applied strains to G_2 and G_3. The measured control voltage corresponding to the peak of the beat signal at 105 kHz has a very good linear relation with the strain applied to G_3 (lower diagram) with root-mean-square strain error of 2.1με. However, the measured voltage corresponding to the peak of the beat signal at 75 kHz shows some non-linearity when the strain applied to G_2 was around 550 με (upper diagram). This corresponds to the spectral overlapping of G_1 and G_2 and the maximum bias from the linear relation was approximately 0.04V corresponding to 100 με. This value agrees well with the crosstalk level estimated from the 'spectral-shadowing' effect [1]. This effect can be reduced using low reflectivity gratings. To estimate the contribution from the non-zero sidelines, gratings G_2 and G_3 were tuned by applying strain so that only light reflected by grating G_1 was allowed to pass through the tunable filter. The magnitudes of the sidelines at 75kHz and 105 kHz were found to be below the noise floor. Simulations were done for various values of OPD to match the simulated spectrum to that of G_1 as shown in Fig.3. The difference between the actual and the optimal separation was found to be about 30cm. The magnitudes of the sidelines at 75kHz and 105kHz were estimated to be -28 and -35dB below that of the central line at 50kHz. These correspond to a maximum crosstalk level of less than 10με [6]. For other strain values where the spectra of G_1 and G_2 are not overlapping, the strain resolution in terms of root-mean-square value of the strain is calculated to be less than 1.8με.

To demonstrate the feasibility of obtaining a single line spectrum by matching the beat frequency to a harmonic of the repetition frequency, a separate single grating system was constructed. Before fusion splicing, the fiber with a FBG to the lead of a directional coupler, the length of the fiber was cut 1 cm by 1 cm and the output beat frequency spectrum was being monitored until the maximum of the beat note envelope was matched to one of the harmonics. Fig.5 shows that an experimentally obtained single line spectrum at 15 kHz ($3f_s$). The magnitudes of sidelines were 38dB below the central line indicating that cross talk level between sensors should be below –38dB. This is 18 dB better than that demonstrated by the CDMA approach [3] and would cause a strain measurement error in the worst case of less than 4με [6]. The discrepancy between the experimental value (-38dB) and the theoretical prediction (-68dB) may be caused by the non-ideal response of the VCO used in our experiments. We expect that the sideline suppression can be improved if a VCO with a constant output voltage over the operating frequency range is used.

4. STRAIN MEASUREMENT IN GLASS-FIBER COMPOSITES

The most promising method for internal strain measurement in composite structures is to integrate optical fiber inside the structures [7-9]. In this section, we are going to demonstrate the use of our FMCW FBG sensors for internal strain measurement of a composite sample. In fabricating the composite sample, E-glass (0°/90°) woven mat and Araldite MY750 epoxy resin were used to form 20 layers of the Gr/Ep

composite. The laminated beam was cured at room temperature for 3 days and was 235.0 mm in length, 26.6 mm in width and 3.8 mm in thickness. Two FBG sensors G_3 and G_2 as shown in Fig.2 were placed at the middle section of the beam between the 1st and 2nd, 19th and 20th layers respectively. The physical length of the gratings is 10 mm. Two electrical strain gauges were bonded on the top and the bottom surfaces of the beam closest to the two FBG sensors in order to compare with the strain readings from the FBG sensors.

When a vertical load was applied, the top layers of the beam were compressed and the bottom layers were under tension. Consequently, the FBG sensor G_2 was shifted to longer wavelength and G_3 was shifted to shorter wavelength. At zero load (bending displacement = 0), the Bragg wavelength of grating G_1, G_2 and G_3 are 1555.49 nm, 1556.24nm and 1556.24nm respectively. G_1 was kept strain-free and used as a reference in recording the applied voltage of TOF. The TOF was driven by a slow varying ramped voltage which was scanned over the spectrums of G_1, G_2 and G_3. During the scanning, the applied voltages corresponding to the peak of the beat signals (50 kHz for G_1), (75 kHz for G_2) and (105 kHz for G_3) were also recorded. When making measurement, the variation in the wavelengths of G_2 and G_3 were not determined from the applied voltage of TOF, but determined from the difference between the applied voltages that correspond to the peak spectral position of G_1 and G_2, and, G_1 and G_3. This method reduces the error arisen from the hysteresis effect of the piezo-electric tunable filter. The test results are shown in Fig. 6. The results obtained from the FBG sensor agrees very well with that from the electrical strain gauge.

5. SUMMARY

In conclusion, we have demonstrated a FMCW based multiplexing technique for addressing FBG sensor arrays. This technique can be used to multiplex FBG sensors with the same Bragg wavelengths. When coupled with WDM approach, the technique has the potential to multiplex a few hundreds of FBG sensors with με resolution.

The FBG sensors were embedded inside glass-fiber composite sample and were used to measure the strain of the sample under three-point bending test. The strain readings from FBG sensors agree well with that from the electrical strain gauges. The FMCW provides alternative to the existing TDM technique that allows to a multiplex a larger array of FBG sensor arrays.

ACKNOWLEDGEMENT

This work was supported by the Hong Kong Polytechnic University Research Grants (Project No. V390).

REFERENCES

1. A. D. Kersey, M. A. Davis, H. J. Patrick, M. LeBlanc, K. P. Koo, C. G. Askins, M. A. Putnam and E. J. Friebele, "Fiber Grating Sensors", J. Lightwave Technol., vol. 15, No.8, pp1442-1463, Aug. 1997.
2. R. L. Idriss, M. B. kodindouma, A. D. Kersey and M. A. Davis, "Multiplexing Bragg grating optical fiber sensors for damage evalution in highway bridges", Smart Mater. Struct. vol. 7, pp. 209-216, 1998.
3. K.P.Koo, A.B. Tveten and S.T.Vohra, "Dense wavelength division multiplexing of fibre Bragg grating sensors using CDMA," Electron. Lett., Vol.35,165-167,1999.
4. Bassem R. Mahafza, "Introduction to Radar Analysis", CRC Press, 1998.
5. A.J. Hymans and J.Lait, "Analysis of a Frequency-Modulated Continuous-Wave Ranging System", Proceedings of IEE, vol. 107B, pp. 365-372, July 1960.
6. W.Jin, "Investigation of interferometric noise in fiber-optic Bragg grating sensors by use of tunable laser sources," App. Optics, vol.37, no.13, May 1998, pp.2517-2525.
7. Y. J. Rao, D. A. Jackson, L. Zhang and I. Bennion, "Strain sensing of modern composite materials with a spatial/wavelength-division multiplexed fiber grating network", Optics Letters, Vol.21, No.9, May 1996, p.p.683-685.
8. H. Storoy, j. Saether and K. Johannessen, "Fiber Optic Condition Monitoring During a Full Scale Destructive Bridge Test", Journal of Intelligent Material Systems and Structures, Vol.8, August 1997, p.p.633-643.
9. M. A. Davis, D. G. Bellemore, M. A. Putnam and A. D. Kersey, "High Strain Monitoring in Composite-Wrapped Concrete Cylinders using Embedded Fiber Bragg Grating Arrays", SPIE Vol.2721 1996, p.p.149-154.

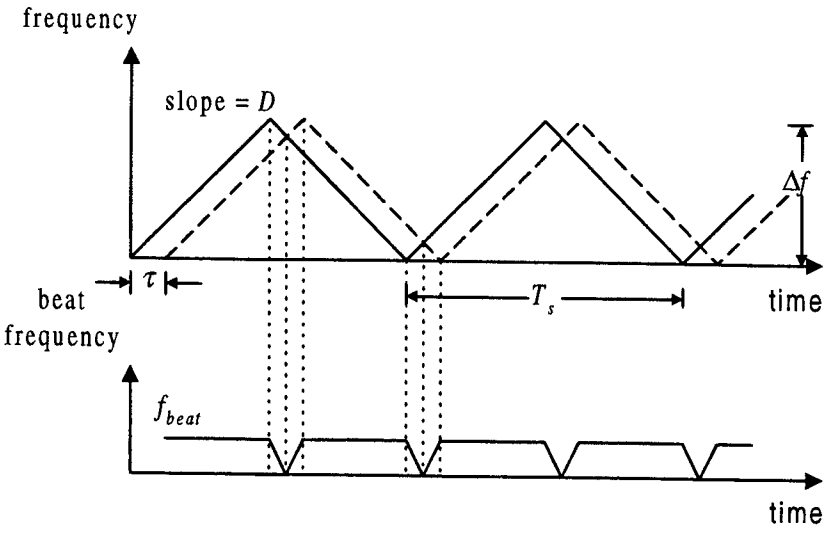


Fig. 1 Production of beat note

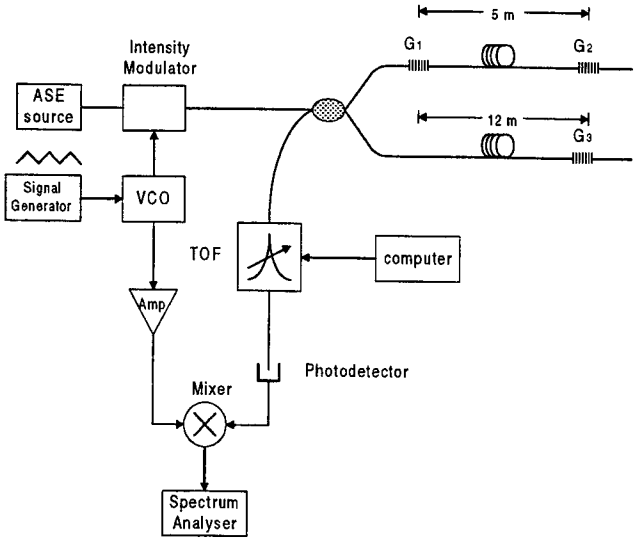


Fig. 2 Experimental setup

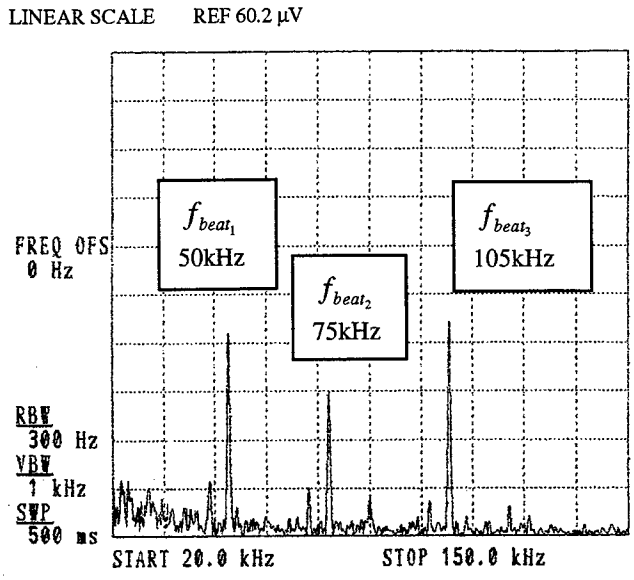


Fig. 3 Beat-note spectrum of the three sensor array

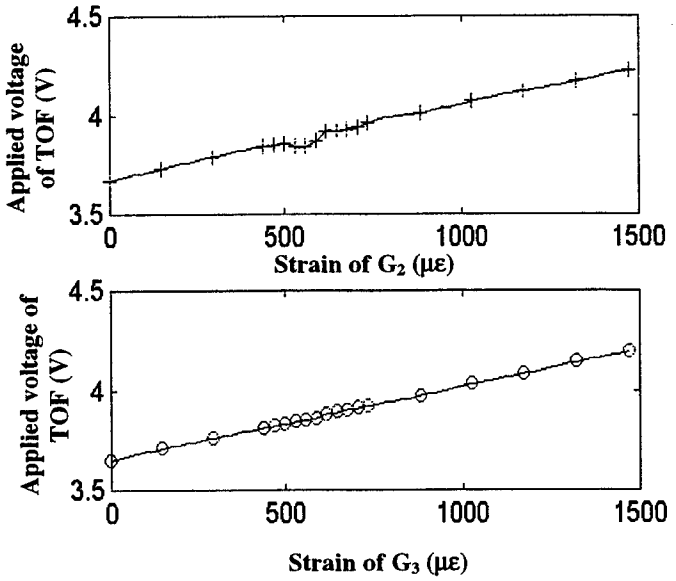


Fig. 4 Applied voltage of TOF as function of the strain applied to grating G_2 (upper diagram) and G_3 (lower diagram).

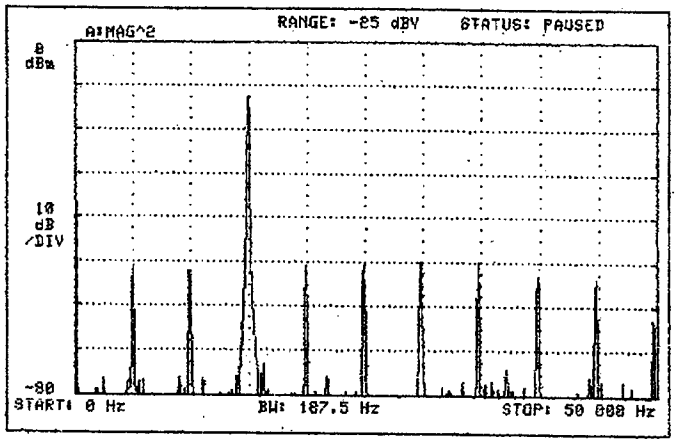


Fig. 5 Beat note spectrum of single grating when aligned to the $3f_s$ line.

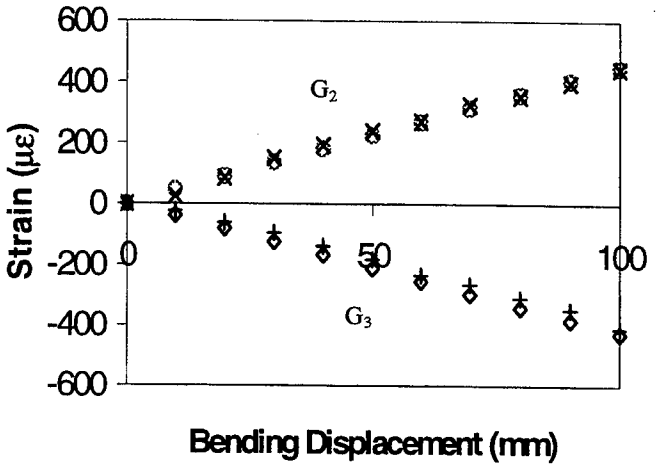


Fig. 6 Internal strain measured by FBG sensors and strain measured on the surfaces of the composite by bonded strain gauges (×: Strain from G_2; o: Strain from the strain gauge on the bottom surface of composite; +: Strain from G_3; ◊: Strain from the strain gauge on the top surface of composite).

Optical Time Domain Reflectometry for Distributed Sensing of the Structural Strain and Deformation

Xie Guangping*, Seah Leong Keey, Anand Asundi

School of Mechanical & Production Engineering, Nanyang Technology University, Singapore 639798
*mgpxie@ntu.edu.sg

ABSTRACT

A unique structure of microbend optical fiber sensor (MOFS) for measuring tensile and compressive strain is described in this paper. The average measuring sensitivity for tensile strain is 35 με using 3 MOFS arrays. The repeatability and stability of MFOS are better than 18 με . The loss sensitivity of single-mode (SM) fiber and multi-mode (MM) fiber used in MOFS, as well as the relationship between the pulse width of diode laser and loss sensitivity are also studied in this paper. From these studies, some conclusions have been obtained. There are 1) the loss sensitivity and repeatability of SM fiber are better when compared to MM fiber in MOFS, and 2) the variation of pulse width of laser would only influent the signal-to-noise ratio and dynamic range, but has no contribution to loss sensitivity. Experimental results also show that loss of SM fiber highly depends on the wavelength of laser, but MM fiber has no such property. The loss of SM fiber between the wavelength of 1550nm and 1310nm is about the ratio of 6.5. Therefore, the experiments reported in this paper used wavelength of 1310nm to measure tensile strain and 1550nm to measure compressive strain based on the above property of SM fiber, without changing the configuration of MOFS.

Keywords: Optical time domain reflectometry, microbend sensor, distributed sensing, optical fiber, strain, loss, optical fiber switch, spatial resolution, dynamic range

1.INTRODUCTION

Optical time domain reflectometry (OTDR) has become a valuable diagnostic tool which is used during the manufacture, installation, and maintenance stages of optical fiber system for the past several years. However, recently there are widespread interests in the development of distributed optical fiber sensors based on OTDR technique for the real-time monitoring of the deformation of large and important engineering structures such as bridges, space vehicles, pressure vessels, submarines, dams etc. The main problems in using OTDR technique for these applications are the sensitivity of loss and spatial resolution. In our experiment, we found that single-mode (SM) and multi-mode (MM) fibers are only sensitive to transverse compressive strain but not to axial tensile and bending strain, even though the fiber is subjected to tensile strain of greater than 5%.

In order to resolve these problems and assure the usefulness of OTDR instrument in deformation detection of large and critical engineering structures, microbend technique is studied to enhance the strain sensitivity of optical fiber in this project.

2. MICROBEND SENSORS FOR MECHANICAL SENSING

Although microbend-induced loss in optical fibers is a drawback in communications, it has proven itself useful as a transaction mechanism in optical fiber intensity sensors for detecting environmental changes[1]. Specific microbend sensor designs have been already utilized for detecting pressure[2], displacement[3], acoustic wave[4], temperature[5], water[6] etc. A fiber microbend sensor, however, despite its appearance, has a very complicated sensing performance depending on the deformation of fiber and fiber parameters. Previous microbend sensors mostly rely on non-uniform deformations of fiber caused by microbenders such as a pair of corrugated plates, a chain rollers or wire winding. In the case of these sensors, an initially straight sensing fiber is sandwiched between a pair of comngated plates. When they are squeezed together, this action causes it to undulate quasi-sinusoidal. Consequently the guided mode of SM fiber couples to cladding modes, or the highest guided mode of MM fiber couples to the first radiated mode due to the deformations of the refractive-index distribution or the geometry of the fiber. Coupling to cladding modes or radiated modes causes power to be lost from the

guided core mode, and the power traveling in the cladding modes will be rapidly dissipated, because of the high losses in these modes.

Fig.1 shows the first kind of idealized and uniform generic microbend sensor studied in this paper for the strain and deformation measurement of structures such as bridge, dam and other civil engineering structures. The sensing mechanism is based on the principle of additional losses produced by series of microbend on the optical fiber. The mechanical sensing element of the microbend sensors is a pair of groove plates. One of the plates is fixed by epoxy on the complex material frame that has a very perfect elasticity. Both ends of an elastic diaphragm are fixed to the frame of microbend sensor that is firmly bonded on the measured structure and suffer the same strain with measured structure.

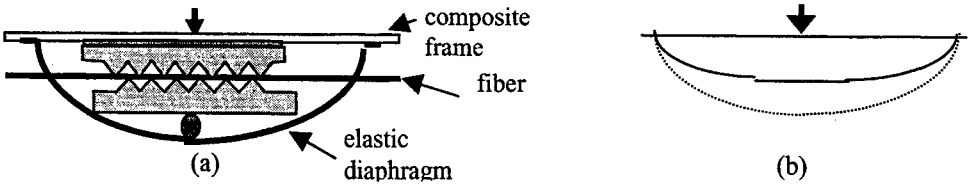


Fig.1 The microbend sensor configuration for strain and deformation measurement

However, due to the restriction of material performance and machining technique, it is difficult to make the saw tooth on an aluminum plate with a thickness of less than 1mm. In our experiment, the thickness of plate is 1mm, with the saw teeth size of 1.12mm× 0.56mm. But when this plate is fixed to the frame using epoxy, it restricts the frame to deform. As shown in Fig.1 (b), the dashed line is the normal deformation curve, and the continued line is the actual curve. This phenomenon is reduced when the plate thickness is reduced. Due to this limitation, a second microbend sensor is developed.

The schematic shown in Fig.2 illustrates the basic configuration and arrangement of the microbend sensor on the measured structure. The mechanical transducers used are actually several standard paper clips, which are fixed to the elastic frame of complex material. Under the action of structure strain, the two ends of the elastic diaphragm will displace along the strain direction, thus lower the arched height of the diaphragm. This action causes the diaphragm to press against the paper clips, causing the paper clips to move downwards and subsequently, causes the optical fiber to suffer from the

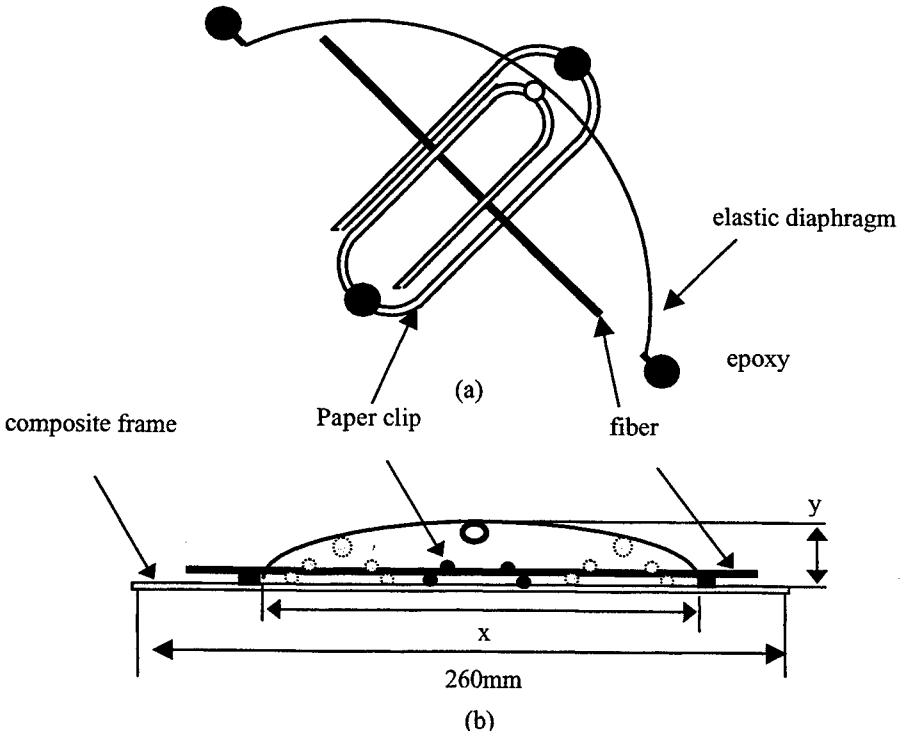


Fig.2 3D view and Schematic of microbend sensor and its equation diagram

modulation of microbend. The extent of microbend loss is depending on the gap between the paper clips and the extent of the paper clips motion.

Supposing the variation of the elastic diaphragm is geometrical only, in Fig.2, the variation δy of the gap between the paper clips in the y-direction under the action of the strain ε in x-direction has the following relationship:

$$\delta y=-m\delta x=-mx\varepsilon \tag{1}$$

where[7]:

$$m=\frac{4n^2-2n\arcsin[4n/(1+4n^2)]}{(1-4n^2)\arcsin[4n/(1+4n^2)]-4n}$$

m represents the sensitivity factor of the microbend strain sensor and n=y/x. The calculated result of the relationship between m and n is shown in Fig.3. It can be seen from this figure, the sensitivity of this microbend sensor can be higher due to the magnification effect of the sensitivity factor m. The strain sensitivity of the sensor can be set up in an appropriate value by choosing the parameter n.

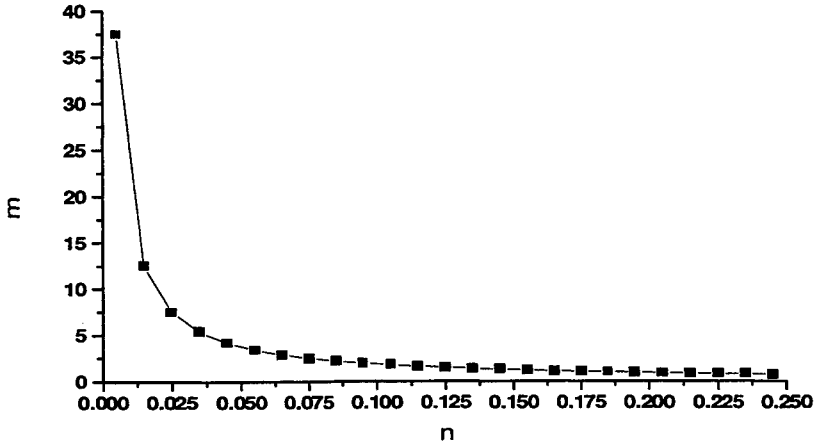


Fig.3 The relationship of m depending on the parameter n

3.EXPERIMENTS

In the experiment, a fully modular portable platform designed OTDR system (WAVETEK Inc., M5100) is used, in which pulse diode laser with center wavelength 1310nm and 1550nm are used as its sources. Resolution of loss is 0.01dB, pulse widths can be selected from 5ns to 10μs. In the preliminary tests, a 9/125μm step-index single-mode fiber and a

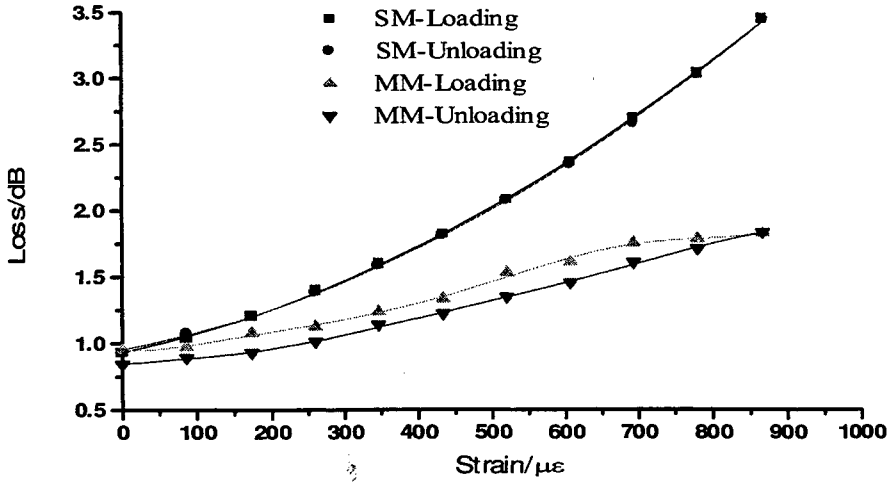


Fig.4 The measured relationship curves between microbend loss and tensile strain for SM and MM fibers

62.5/125μm multi-mode fiber are used as the sensing fiber. Fig.4 shows the measured relationship curves between microbend loss and tensile strain for SM and MM fibers. Experimental results show that the sensitivity, stability, as well as the repeatability of the microbend sensor of single-mode fiber are all better than multi-mode fiber. Therefore, in the subsequent tests, the 9/125μm step-index single-mode fiber (Cutoff wavelength $1150nm \leq \lambda_c \leq 1310nm$, mode-field diameter $9.3 \pm 0.5\mu m@1310nm$, b=62.5μm, n1=1.468) was selected as the sensing fiber of microbend sensor. The microbend sensor is made by three paper clips as the mechanical transducer with the designed parameters of x= 120mm, y= 8mm, and the gap among the three clips is 40mm, as shown in Fig.2. The thickness of the frame of the complex material is 1.2mm.

3.1. Tensile strain measurement and results

In order to verify the tensile strain sensitivity of the microbend sensor shown in Fig.2, an loading machine (INSTRON) is used to apply the tensile with a load step of 0.015mm. The corresponding tensile strain was also measured by the loading machine for comparison with the OTDR results. The two ends of the specimen are held on the Instron Machine with two fixtures. For the measurement of tensile and bending strain, 1310nm wavelength is selected as the light source of OTDR. Fig.5 shows the measured results of tensile strain during loading and unloading. 20ns and 100ns pulse widths of diode laser are selected which correspond to the spatial resolution of 2m and 10m, respectively. Measured results show that the values measured by 20ns and 100ns are in full agreement, with a maximum error less than 0.1dB, which means that the increment of pulse width of laser has no significant contribution to loss sensitivity. But the difference of pulse width will

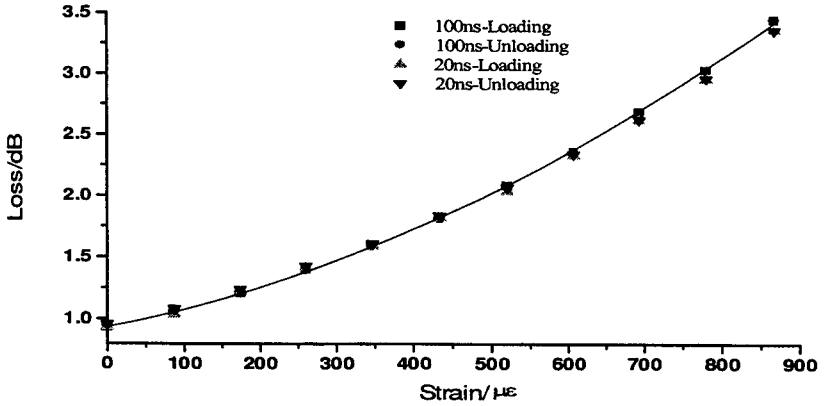


Fig5. Measured result of tensile strain at pulse widths of 20ns and 100ns

influent the stability of microbend sensor system and the dynamic range of OTDR. Fig.6 is the measured results of stability for a duration of 45 minutes of four different pulse widths of 5ns, 20ns, 100ns, and 300ns. It is obviously that the stability

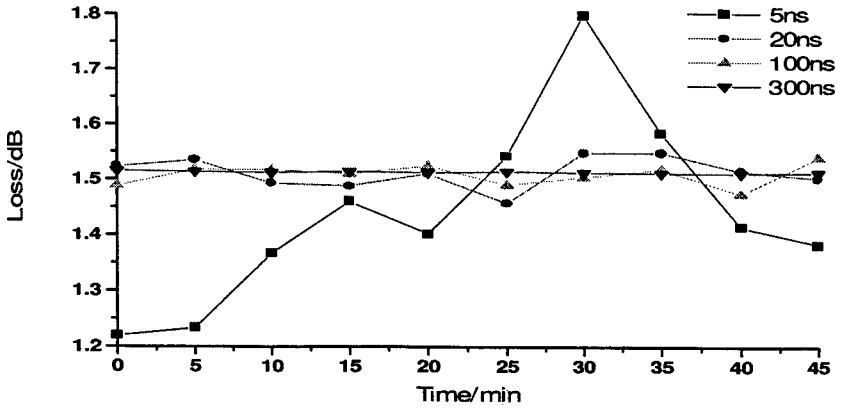


Fig.6 Stability of system for different pulse widths

will be improved notable when the pulse width increases. The stability tolerance are 0.5dB, 0.1dB, 0.05dB and 0.001dB for 5ns, 20ns, 100ns and 300ns pulse width, respectively. It is about 300 times for the stability tolerance error between 5ns and 300ns, though the pulse width is only 60 times for 5ns and 300ns. So, for high accuracy measurement of strain using OTDR, it is necessary to select a wider pulse. It was observed that a wide pulse would reduce the spatial resolution, but this disadvantage can be compensated by other method such as fiber delay ring or 1×N fiber switch technique, which will be introduced in section 4. Fig.5 also shows that the loading and unloading paths are very much the same, this means that the repeatability of the microbend sensor system is very good, even though the paths are not straight line. The sensitivity for strain measurement is better than 35 με for the microbend sensor.

3.2. Bending strain measurement and results

The same specimen was used with the microbend sensor at the bottom for bending strain measurement. The specimen was loaded using a 3-points bending frame with loading applied at the mid-length of the specimen. The selected light source is also the 1310nm diode laser with 100ns and 300ns pulse widths. Fig.7 shows the measured results of loading and unloading for the two different pulse widths of 100ns and 300ns, respectively. The maximum error between the 100ns and

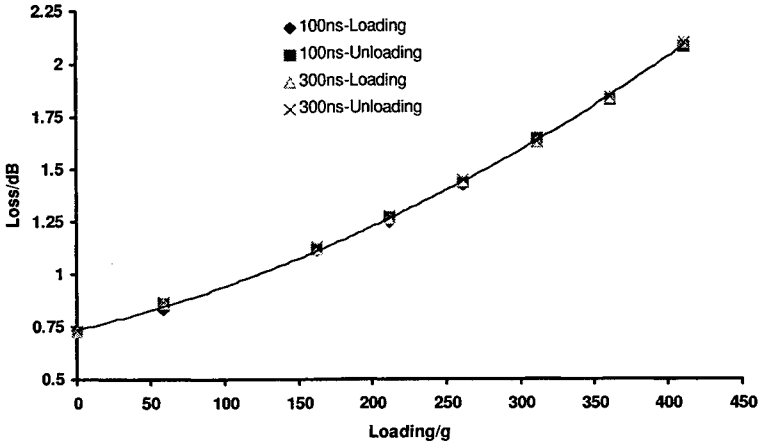


Fig.7 Measured result of bending at pulse widths of 100ns and 300ns

300ns is 0.05dB. The results also demonstrated that the repeatability of the sensor system is very high. Therefore, the sensor is practical because it can be calibrated due to its high repeatability and stability, as well as the good elasticity of the microbend sensor system.

3.3 Compressive strain measurement and results

Usually, the structure of the past microbend sensor for tensile and bending strain is different from that for compressive strain[7]. The structural difference of the sensor for different type of strain leads to the difficulty in using the microbend sensor in the past. In this paper, the same microbend sensor was used to measure the tensile, bending and compressive strain based on the dual-wavelength measuring technique.

According to Petermann [8], the microbend loss coefficient takes the following form:

$$<2\alpha>=\frac{1}{2}(kn_2w_{02})^2 \bullet \phi_p(1/(w_{01}^2kn_2)) \qquad (2)$$

where $\phi_p(\Omega)=\Omega^4\sqrt{\pi}\sigma^2L_c\cdot\exp\{-[(\Omega L_c)/2]^2\}$ denotes the curvature power spectrum, n_2 is the refractive index of the cladding, and $w_{01}=w_0+a\cdot\exp[3.34-3.28V]$, $w_{02}=w_0-a\cdot\exp[2.45-3.31V]$, w_0 is fiber mode field diameter, V is spectral frequency. L_c is the coherent length.

When L_c=500μm, the microbend loss coefficient <2α> calculated from equation (2) will increase when the wavelength drift to a longer wavelength. The calculated results show that for the same strain or loading, different wavelength will create

different loss, the longer the wavelength, the larger the loss. The ratio of the microbend loss between 1550nm and 1310nm is about 6.0 calculated by equation (2), as shown in Fig.8. Fig.9 shows the measured results of loss for the wavelength of 1310nm and 1565nm. As shown in this Fig.9, the loss between the wavelength of 1550nm and 1310nm is in the ratio of

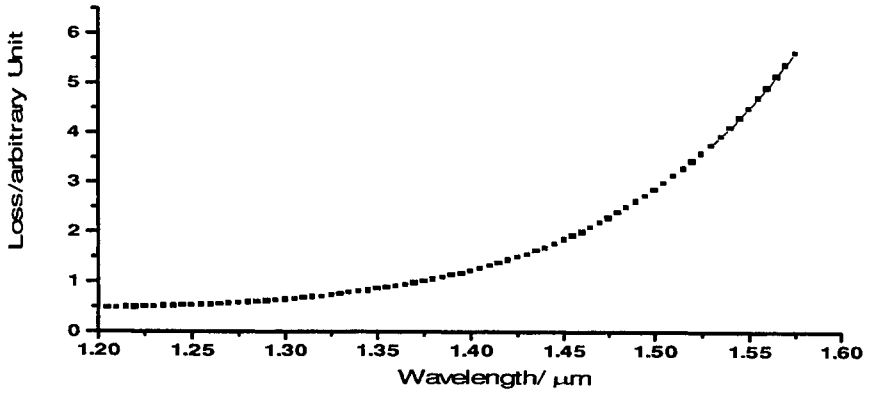


Fig.8 Relationship between the wavelength and microbend loss

about 6.5, which is almost the same as the calculated value. Therefore, in this project, we preset the initial loss of 1310nm to a lower level of 0.7dB, then the loss of 1565nm will get to a higher level near to 5.0dB. Based on this performance of microbend sensor, we use the 1310nm wavelength to measure the tensile and bending strains, and use 1565nm wavelength to measure the compressive strain, with the same structure of the microbend sensor. When the tensile or bending strain increases, the loss gradient of microbend sensor for 1310nm will passivity, in contrast, the loss gradient for 1565nm will

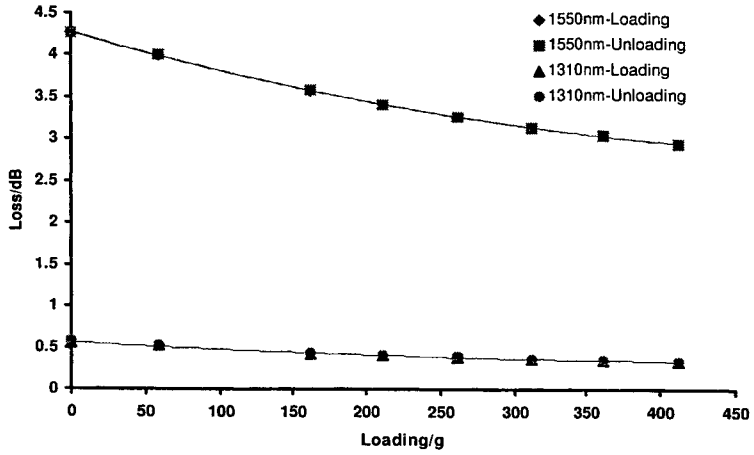


Fig.9 Measured results of compressive strain at pulse width of 300ns and wavelength of 1550nm

decrease when the compressive strain increases. According to the loss value for 1565nm, we are not only about to measure the compressive strain, but also to know whether the structure is under the tensile, bending, or compressive strain.

For compressive strain measurement, the same specimen with the microbend sensor was used upside down manner, where the sensor now was on the top of the specimen. The specimen was also loaded using a 3-points bending frame with loading applied at the mid-length of the specimen. Fig.9 shows the measured results of compressive strain for loading and unloading. The results also show a very high repeatability for loading and unloading.

4.TECHNIQUES FOR IMPROVING THE SPATIAL RESOLUTION AND DYNAMIC RANGE

Spatial resolution and dynamic range are two contrary parameters decided by pulse width of laser in OTDR. Decreasing the pulse width can improve the spatial resolution, but will decrease the dynamic range and signal-to-noise. In this paper,

two methods are introduced to improve the spatial resolution and dynamic range without decreasing the signal-to-noise, which are based on fiber delay ring and 1×N fiber switch, respectively.

Fig. 10 shows the method based on fiber delay ring. The length of fiber delay ring is decided by the spatial resolution of OTDR. Usually, commercial OTDR has in-line filter inside to improve signal-to-noise, this technique is necessary for a commercial OTDR instrument because of its very weak received optical intensity, but will seriously decrease the spatial resolution almost down to 2 times. Fiber delay ring shown in Fig.10 is a simple and effective method to improve the spatial resolution of OTDR without decreasing its dynamic range and signal-to-noise.

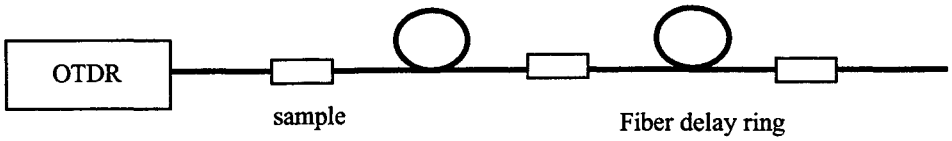


Fig.10 Using fiber delay ring to improve the spatial resolution of OTDR

Fig.11 shows another method based on 1×N fiber switch, which can improve the spatial resolution and dynamic range of OTDR up to N times, simultaneously. The commercial 1×N fiber switches are designed for 9/125μm single-mode, as well as 62.5/125 and 50/125μm multi-mode operation. Switching is bi-directional. The 1×N switching module is built around a precision stepper motor, available with 2 to 100 output channels. The switches can be easily integrated into automated test systems using SCPI-compatible commands via GPIB or RS232 interfaces. Fig.12 shows the measured dynamic range with and without a 1×2 single-mode fiber optic switch (Newport). The difference is quite small, which means the using of fiber switch does not influent the dynamic range of OTDR. So, if we use a 1×N fiber switch, we can improve the spatial resolution and dynamic range of OTDR up to N times, simultaneously.

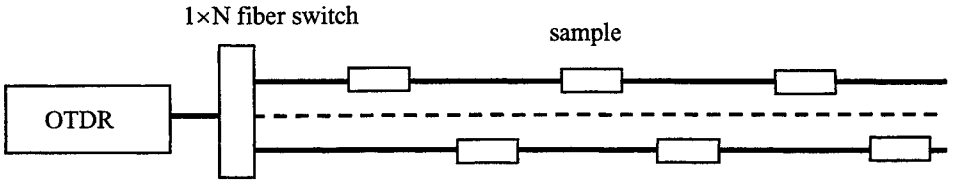


Fig.11 Using 1×N fiber switch to improve the spatial resolution and dynamic range of OTDR

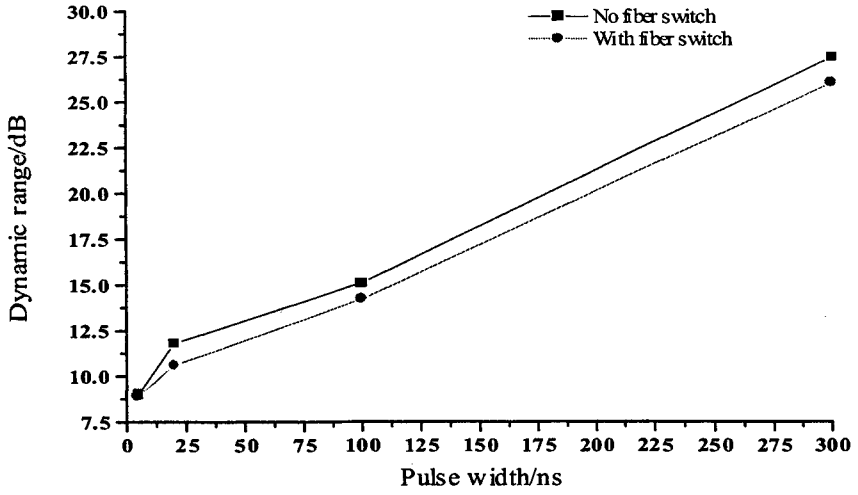


Fig.12 Dynamic range of OTDR with and without fiber switch

5. CONCLUSION

In this paper, we described a distributed strain sensor based on the OTDR technique, where the novel structures of the optical fiber microbend sensor are used for measuring of the tensile, bending, as well as the compressive strain. The use of dual-wavelength technique makes it possible to measure the tensile and compressive strain using the same microbend sensor, without changing the structure of the microbend sensor. This technique could also distinguish the compressive strain from tensile and bending strain. Experimental results also show that the variation of pulse width of laser would only influent the signal-to-noise ratio and dynamic range and has no contribution to loss sensitivity. The distributed optical fiber sensor system studied in this paper is practical in the sensing and monitoring the strain and deformation in large engineering structures.

REFERENCES

[1]. Bruce L.Danielson, "Optical time-domain reflectometry specifications and performance testing," *Applied Optics*, **Vol.24, No.15**, pp2313-2322, 1985.

[2]. Varshneya D., Ghering W.L., etc., "High-temperature fiber optic microbend sensor," *Proceedings of 3rd optical fiber sensors conference,* San Diego, pp140, 1985.

[3]. N.Lagakos, T.Litovitz, P.Macedo, etc. "Multimode optical fiber displacement sensor," *Applied Optics*, **N0.20**, pp167, 1981.

[4]. Yoshino T., Nara M., "Remote and multi-point fiber sensors using optical time domain reflectometry," *Proceedings of 13th congress of the international commission for optics*, Sapporo, pp324-325, 1984

[5]. Fields J.N., Cole J.H., "Fiber microbend acoustic sensor," *Applied Optics*, **No.19**, pp3236-3267, 1980.

[6]. Tomita A. S., Tachino H., "Water sensor with optical fiber," *Journal of Lightwave Technology*, **No.8**, pp1829-1832, 1990.

[7]. Luo Fei, Meng Aidong, Zhang Deying, "Distributed optical fiber sensor for the monitoring of the structural strain and deformation," *Proceedings of SPIE*, **Vol.2838**, pp296-300, 1996.

[8]. A.Bjarklev, "Microdeformation losses of single-mode fibers with step index profiles," *Journal of Lightwave Technological*, **LT-4, No.3**, pp342-346, 1986.

Fiber optic vibration sensor -based smart civil structures

Jinsong Leng [a,b] A.K.Asundi [a]

[a] School of Mechanical and Production Engineering, Nanyang Technological University, Nanyang Avenue, Singapore 639798

[b] Center for Smart Materials Systems and Structures, P.O.Box, #2147, Harbin Institute of Technology, Harbin 150001, P.R.China

ABSTRACT

Two kinds of fiber optic vibration sensors for civil structure vibration monitoring are proposed in this paper. The first is based on the detection of spatial mode speckle of a multi-mode optical fiber. A multi-mode optical fiber, diameter is 200/230μm, is used in the present experiment. The second method uses an e-core two-mode optical fiber. The experiment on a composite beam with an embedded statistical fiber optic sensor and surface mounted e-core two-mode fiber optic sensor are demonstrated. The experimental results show that the two types of the fiber optic vibration sensors have higher sensitivity compared to piezoelectric sensor along with other benefits of fiber optic sensor, including lower cost.

Keywords: Smart Civil Structure, Statistical-Mode Fiber Optic Sensor, E-Core Two-Mode Fiber Optic Sensor, Vibration Monitoring

1. INTRODUCTION

Smart material systems and structures (SMSS), designed according to the inspiration from nature, should have the ability to adjust their structure, morphology, shape and properties to the changing environment and aging process. SMSS should thus have functions that allow them to change their shape, monitor their own health, control vibrations, and in general display adaptive and kinetic features.[1-4] Smart materials and structures, not only have traditional structural materials' functions, but also have actuation, sensing and microprocessing capability. SMSS is hybrid composite material system that is composed of three important parts: sensor, actuator and microprocessor. Sensors sense the variations of the circumstance; the microprocessors analyze these signals and then instruct the actuators to change the global mechanical characteristics of the structure. Fiber optic sensor, shape memory alloy, electrorheological fluid, magnetorheological fluid and piezoelectric and piezoceramic materials, etc. [1-4] are the component materials which make up the SMSS.

Fiber optic sensors (FOSs) have many advantages used in smart materials and structures. They are small and therefore do not affect the properties of the concrete into which they are embedded. They are light yet rugged and low cost and are unperturbed by electro-magnetic interference (EMI). They are sensitive to any parameter which can modify the intensity, frequency, polarization, or phase of light travelling through the fiber. One of the advantages is that these sensors are very light in weight and small enough that they can be embedded in composite materials in a nonobtrusive manner that does not degrade structural integrity. So smart composite structures embedded with fiber optic sensors can monitor the health of the structures under service condition. Fiber optic sensors can be developed to selectively detect variation due to strain, corrosion, temperature, crack formation. A lot of fiber optic sensors have been developed for use in smart material and structures. [2-5] Fiber optic smart structures are an enabling technology that will allow engineer to add a nervous system to their designs, enabling damage assessment, vibration damping, and many other capabilities to structures that would be very

Futher author information

Jinsong Leng : Email: mjsleng@ntu.edu.sg

WWW: http://www.hit.edu.cn/englishver/e/school/smart/js.html Fax: (65) 7911859

A.K.Asundi : Email: masundi@ntu.edu.sg

WWW: http://www.ntu.edu.sg/home/masundi Fax : (65) 7911859

Part of the SPIE Conference on Advanced Photonic Sensors and Applications
Singapore • November–December 1999
SPIE Vol. 3897 • 0277-786X/99/$10.00

difficult to achieve by other means. The smart civil structure is one major potential market for smart structure technology. The potential market for the application of smart civil structures can be quite large. The most probable candidates will be smart civil structures such as smart building and skyscrapers, smart bridge, dams, bridge decks etc.. Fiber optic sensors offer many potential advantages for application to civil structural systems. The two kinds of fiber optic vibration sensor based-smart civil structure have been developed in this paper. One sensor is the statistical-mode fiber optic vibration sensor. The other sensor is the e-core two-mode fiber optic sensor. The experimental results show that the two fiber optic vibration sensors have higher sensitivity compared to piezoelectric sensor along with other benefits of fiber optic sensor..

2. STATISTICAL-MODE FIBER OPTIC SENSOR

A coherent light when injected into a multimode optical fiber gives a grainy interference pattern called speckles at the output end of the fiber. When a external perturbation is applied to this moltimode optical fiber, the speckle patterns change with the total intensity remaining constant. The intensity of the *i*th speckle can be written as [6-7]

$$I_i = A_i\{1 + B_i[\cos(\delta_i) - F(t)\phi_i \sin(\delta_i)]\} \qquad (1)$$

Where A_i is the amplitude of the *i*th speckle. A_i depends on the self-interaction mode. B_i is the steady state mode-mode interaction of the *i*th speckle. $F(t)$ is the external perturbation. . When a external perturbation is applied to this multimode optical fiber, the mode-mode interaction term is modified by Φ_i . Where A_i, B_i, Φ_i and δ_i are constants for any given i.

It can be seen that the speckle pattern can be considered as an array of interferometers. There are two methods for processing this individual intensity of the *i*th speckle. First method is summing the changes of a part of the speckle pattern at two different times so that statistical averaging of the speckles does not resulting in a constant value. The second method is summing the absolute of the difference of all speckles at two different times. The second method has higher sensitivity because of summing of absolute changes. But a CCD camera is needed for detection of speckles. The dynamic range of this method is limited by the speed of CCD. For first method, the basic principle of this sensor is shown in Figure 1. In this case, a simple spatial filter is used for signal processing. The sum of n components ($n<<N$) of all speckles is taken. In fact, only the time varying element are considered. So

$$\Delta I_T = \left[\sum_{i=1}^{n} C_i \sin(\delta_i)\right] F(t) \qquad (2)$$

From equation (2), we can sense external perturbation such as vibration frequency.

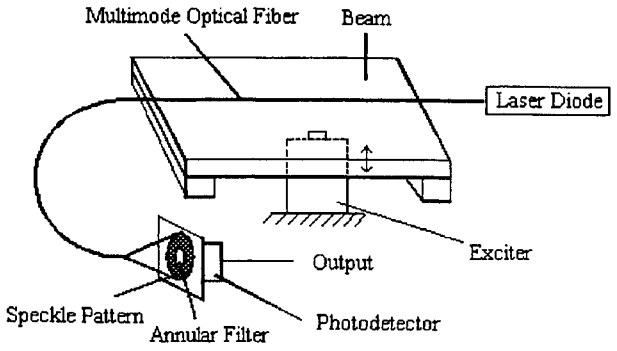


Fig.1 Experimental schematic of the fiber optic vibration sensor based on the speckle detection

3. E-CORE TWO-MODE FIBER OPTIC SENSOR

A two-mode fiber sensor includes the sensing and reference arms within the same fiber. In such a sensing technique, the phase difference between the first two modes is modulated by the parameter being measured, and the intensity monitored at the output is a measure of the perturbation. This approach eliminates the need for a reference arm since both modes propagate within the same fiber. The trade-off is that a slightly lower sensitivity (30dB less than the theoretically achievable single-mode fiber sensor sensitivity) is obtained. The major advantage of such a sensor is that it is rugged and maximum performance levels can be achieved without invoking the need for stabilization schemes that are an essential ingredient of

the single-mode Mach-Zehnder and Michelson arrangements.
Two- and three-lobe far-field intensity distributions obtained at the output by exciting appropriate modes within the fibers are used for measurement of pressure, strain, and vibrations. The standard method of analysis involves a measurement of the exchange of optical power between these lobes to determine the nature of the external perturbation. The major obstacles in practical implementation of sensing systems that use circular-core fibers have been the strict input conditions required to launch specific mode and the instability of the cross-sectional intensity distribution of the second-order modes.
In the weakly guiding approximation, single-mode circular core fibers can support the two orthogonal polarizations of LP_{01} mode as well as the four degenerate, higher-order, LP_{11} modes when operating just below the single mode cutoff wavelength. Two-lobe patterns can be obtained in the far field at the output of the fiber, and the oscillation of the power distribution between the lobes can be used to sense strain or vibration, as described below. However, environmental conditions can also introduce differential phase shifts between the almost degenerate four eigenmodes, leading to an instability of the second-order mode pattern. This limits the practical implementation of such sensors. [2]
The use of optical fibers with highly elliptical cores has been shown to remedy this situation. Since the circular symmetry of the fiber has been eliminated, only two second-order modes, the LP_{11}^{even} modes, are guided by the elliptical-core fibers just below the single-mode cutoff wavelength. The intensity distribution of the second-order modes is stable, and practical operation of a sensor system is possible since there is a considerable range of the optical spectrum over which the LP_{11}^{odd} modes are unguided. The configuration of e-core two-mode fiber optic sensor is shown in Fig.2. An e-core single-mode optical fiber is used as the lead-in fiber, a two-mode e-core fiber comprises the sensing section and a circular core single-mode or multimode fiber is used as the lead-out fiber. The resulting sensor output is sinusoidal and can be expressed as [8-9]

$$I(t) = I_0 + I_1 \cos[\Delta\beta \int_a^b p(x)\varepsilon(x,t)dx] \qquad (3)$$

Where ε is the strain experienced by the fiber, $\Delta\beta$ is the difference in the propagation constants of the LP_{01} and the LP_{11}^{even} modes, and a and b denote the two end-points of the two-mode sensing region of the fiber.

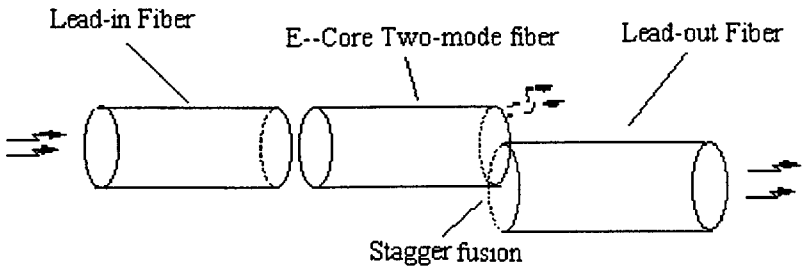


Fig.2 The configuration of e-core two-mode fiber optic sensor

4.EXPERIMENT AND RESULTS

The statistical fiber optic sensors were fabricated using conventional 200/230 μm multi-mode optical fibers .The schematic diagram of the vibration measurement of the smart beam embedded the statistical-mode fiber optic vibration sensor (FOVS) is shown in Figure 3. The fiber optic vibration sensor is embedded within a Carbon/Epoxy composite beam. In order to detect the summation of changes of a part of the speckle pattern at two different times, an annular spatial filter is used. Because the annular area of the speckle pattern is detected by detector, so that statistical averaging of the speckles changes with time. The frequency response curves of the embedded fiber optic sensor at vibration frequency 12.5 Hz and vibration frequency 400Hz are shown in Figure 4 and Figure 5. It can be seen that the statistical-mode FOVS has higher sensitivity and dynamic range compared to the traditional piezoelectric sensor. The e-core two-mode fiber optic sensor has been fabricated by using three-part fiber. The lead-in fiber is e-core single-mode fiber 1.25 μm X2.5 μm / 80 μm at 633 nm. The sensing fiber is e-core dual-mode fiber 1.25 μm X2.5 μm / 80 μm at 820 nm. The lead-out fiber is a circular core single-

mode fiber 4/125 μm at 633 nm. The schematic diagram of the vibration measurement of the smart beam surface mounted with the e-core two-mode fiber optic sensor is shown in Fig.6. The measurement results at vibration frequency 12.5 Hz and 100 Hz are shown in Fig.7 and Fig.8.

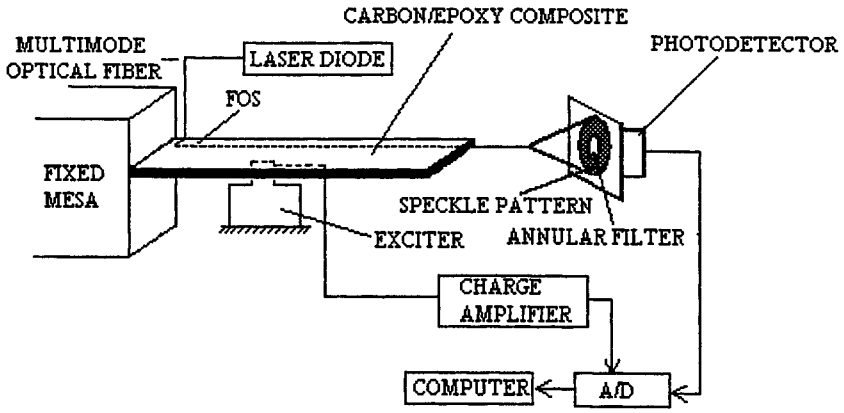


Fig.3 The schematic diagram of the vibration measurement of the smart beam embedded statistical-mode fiber optic vibration sensor

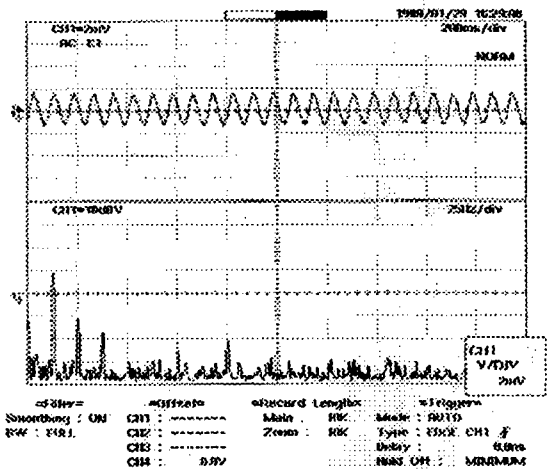


Fig.4 Frequency response curve of smart composite beam embedded the statistical-mode fiber optic vibration sensor at 12.5Hz

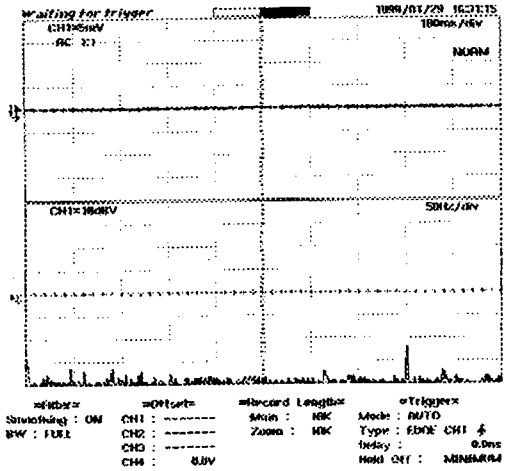


Fig.5 Frequency response curve of smart composite beam embedded the statistical-mode fiber optic vibration sensor at 400Hz

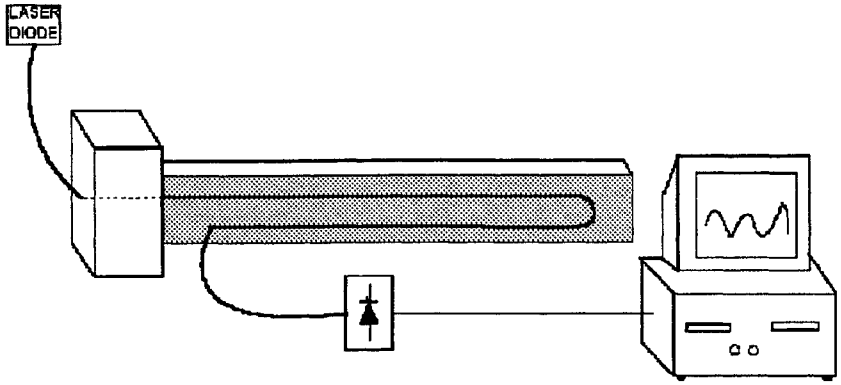


Fig.6 The schematic diagram of the vibration measurement of the smart beam embedded e-core two-mode fiber optic sensor

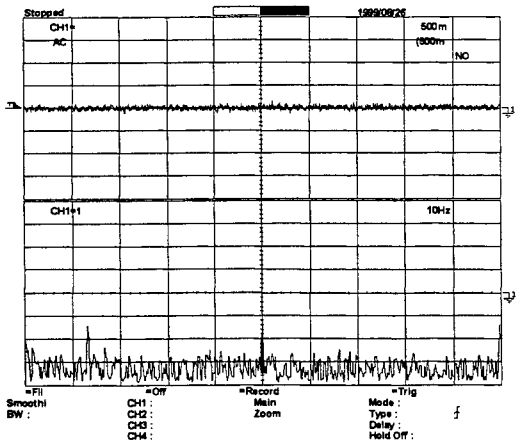


Fig.7 Frequency response curve of smart composite beam embedded the two-mode fiber optic sensor at 12.5Hz

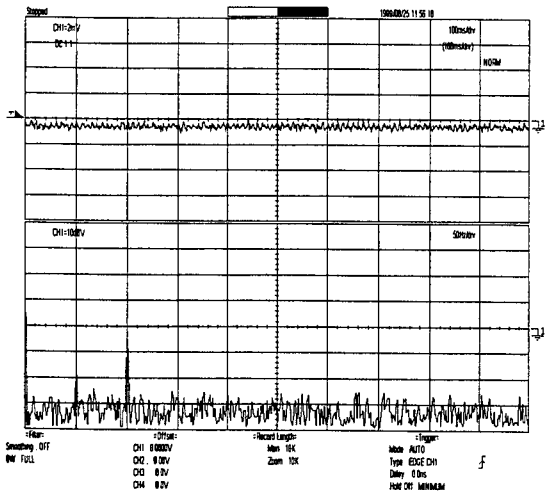


Fig.8 Frequency response curve of smart composite beam embedded the two-mode fiber optic sensor at 100 Hz

5.CONCLUSIONS

Two types of fiber optic vibration sensor based-smart civil structure are proposed and demonstrated in this paper. The first sensor is the statistical-mode fiber optic vibration sensor. The second sensor is the e-core two-mode fiber optic sensor. From above discussions, some important conclusions can be obtained. The statistical-mode fiber optic vibration sensor and the e-core two-mode fiber optic sensor can be used to sense the vibration of smart civil structure. It has comparable sensitivity and dynamic range to the traditional piezoelectric sensor. Also there are lower in cost. Hence, the two kinds of fiber optic sensor-based smart civil structures are found to be most suitable for health monitoring of buildings, bridges and highways.

6.REFERENCES

1. M.V.Gandhi and B.S.Thompson , Smart Materials and Structures, Chapman &Hall, London, 1992.
2. Eric Udd, Fiber Optic Smart Structures, John Wiley & Sons, Inc.,1995.
3. S.Y.Du and J.S.Leng , " A Study On The Optical Fiber Intelligent Composite Materials and Its Monitoring System" , Advances in Mechanics, Vol. 4 ,pp.496-503,1992 .
4. S.Y.Du and J.S.Leng, " Fiber optical NDE method for advanced composites" , Aerospace Materials and Technology, Vol.5,pp.14-18,1992.
5. Selvarajan,A., and A. Asoundi, " Photonics, Fiber Optic Sensors and Their Applications in Smart Structures " ,Journal of Non-Destructive Evaluation , Vol. 15(2),pp.41-56,1995.
6. W.B.Spillman,B.R.Kline,L.B.Maurice and P.L.Fuhr, " Statistical-mode sensor for fiber optic vibration sensing use," Applied Optics, Vol.28,No.15,pp.3166-3176,1989.
7. J.S.Leng, A. Asundi, "Vibration measurement of smart composite material based on statistical-mode fiber optic sensor", International Conference on smart materials, structures and systems, ISSS-SPIE' 99,July,1999,India.
8. A.M.Vengsarkar,K.A.Murphy, et al., "Two-mode elliptical-core weighted fiber sensors for vibration analysis",SPIE,Vol.1586,pp.53-66,1991
9. B.M.Zhang, J.S.Leng, "The Few-mode Optical Fiber Sensors in Smart Structures", Laser and Progress of Photoelectronics,(2),1997

LASER VIBROMETER APPLICATION ON BRIDGE CABLE TENSION MEASUREMENTS

Shen-En Chen*[a] and Samer Petro[b]

[a] Dept. of Civil & Environmental Engrg., Univ. of Alabama/Birmingham, AL 35294-4461

[b] Federic R. Harris, Inc., Morgantown, WV 26505

ABSTRACT

To ensure the safety and integrity of tied arch bridges, it is crucial that tension levels in cables do not exceed their design levels. Currently, visual inspection is required since there are no reliable techniques that can accurately determine the tension levels of these cables. A possible approach would be to correlate the vibration measurements with the tensions in these cables. However, due to their long length, access to these cables for mounting contact sensors is not easy. An attempt has been made to use a He/Ne Laser Vibrometer for non-contact cable vibration measurements on a tied arch bridge (I-470 Vietnam Veteran Memorial Bridge, Wheeling, West Virginia, USA). The objectives of this test are to assess the quality of vibration measurements from the He/Ne laser under regular daylight environment and to determine the vibration signatures of these bridge cables under ambient condition. The results of study indicate that using non-contact measurements, a quick and easy prediction of tension levels in bridge cables can be made. This paper presents the results of the study and presents discussion on the advantages of non-contact measurements and possible testing difficulties.

Keywords: laser vibrometer, tension levels, tied arch bridge, bridge safety

1. INTRODUCTION

Due to their load-carrying efficiency, cable suspended structures are common for applications such as radio towers, power lines, ski lifts and cable bridges. The cables used are usually manufactured from high strength steel wires. Due to its flexibility, cable supports its loads in pure tension, which provides stability to the overall structure. Application of cables may take the form of sloped shape or straight vertical shape. For sloped cables, the cables resist the load using the shape called funicular curve, which usually involve extensive geometrical analysis. This is because the tension in the cable changes continuously in both magnitude and direction along the cable's length. For straight cables, load is resisted in the transverse direction, the analysis of which is simpler.

Regardless of the type of cable design, the integrity of these cables is critical to the stability of the overall structure. Different attempts have been made on evaluating cable integrity, for example: An automated MPC (Magnetic Perturbation Cable) system has been used for fracture detection along the cables using magnetic flux leakage detection method[1], a transverse impulsive wave propagation method has also been attempted to capture defects[2]. Recently attempts have been made in correlating vibration frequencies of the cables with their tension levels [4,5]. Petro[6] reported a series of field tests that was conducted to determine the accuracy of a handheld device for tension measurement. The tests were conducted using contact transducers, and were typically attached to the near-bottom of the cables or tensioned members for measurement. The handheld device, called the Axial Load Measurement (ALM), is based on the simple physical correlation between the tension level and the cable vibration frequencies and limited assumptions. Petro[6] also reported that the tension behavior is complex near the boundary of the tensioned members; hence it is hard to distinguish the fundamental vibration mode. Preferable, the measurements, should have been made as close as possible towards the center of the cable. However, due to the size of these cable structures, which may reach several hundred feet long, it is difficult for any contact transducer setup.

* Correspondence: Email: schen@eng.uab.edu, spetro@frharris.com, Telephone: (205)934-8433, (304)292-8145

For large cable structures, it is prudent to utilize remote sensing techniques, in particular laser optic vibration measurements. Recent advancements in Laser Interferometry have provided a useful tool for such purpose. This paper describes a proof-of-theory test that has been done with the laser technology.

In August 1998, a test was setup to inspect the hanger cables of the I-470 Vietnam Veteran Memorial Bridge (a tied-arch bridge) in Wheeling, West Viginia, USA. This bridge has a steel box arch for carrying the dead and live loads of the bridge. Vertical cables do not have the complex problem of geometric nonlinearity that inclined cables in cable-stayed bridges have. The tension correlation for such system (Section 2) should be straightforward. The purpose of this test is to assess the feasibility of using a Laser Vibrometer (LV) for tension measurements on large structures. Instead of contact transducers, a commercial LV was used for the vibration measurements. Using the LV, the measurements were made to a height of at least 10 ft (3 m) high above bridge deck. The results of this study clearly show the potential of laser technology in civil engineering application.

2. THEORY

Based on the theories of vibration of elastic solids. A member under pure tension or compression vibrates with characteristic frequencies when physically stimulated. Traditionally, cable vibration measurements of natural frequencies are used to predict cable tension using the taut string model, where the natural frequency, ω_n, for the motion out of the plane of a suspended cable is given as [7]:

$$\omega_n = n\pi\sqrt{\frac{T}{mL^2}} \tag{1}$$

In Equation (1), $n = n^{th}$ mode of vibration, m = distributed mass per unit length, T = cable tension, and L = total or effective length of the cable. Bridge cable natural frequencies are usually low, however, loads due to wind pressures consisting of varying wind gusts, traffic, and coupling of modes may induce aerodynamic flutter, which may excite higher order modes. It is shown that if the vibration frequencies can be clearly identified, then tension level can be predicted by back calculation:

$$T = \left(\frac{w_n}{n\pi}\right)^2 mL^2 \tag{2}$$

As shown in equation (2), to calculate the tension value, exact measurements of four unknowns must be made. The mass value usually does not alter significantly, unless the metal has been heavily corroded. The effective length is a function of the boundary condition, which typically was assumed to be pined-pined. For long cables, the boundary effects reduced and the pin-pin assumption is accurate. The determination or selection of a boundary condition is, however, dependent on the actual fixation of the cable to the bridge. The mode number, **n**, is an integer number and is not a measurable number. However, the mode number must be correlated to the measured frequency.

3. BRIDGE TESTING

3.1 The I-470 Vietnam Veteran Memorial Bridge

The I-470 Bridge, also called the Vietnam Veteran Memorial Bridge, is located in Wheeling, West Virginia, USA, and has a span of 779 ft (237.4 m) long. It is a four-lane highway bridge, spanning across the Ohio River with eastbound traffic towards the state of West Virginia, and westbound traffic towards Ohio. The bridge deck is about 69 feet (21.1 m) above the water level. The arch or rib and the tie of the bridge are both box sections. Figure 1 shows the bridge over the Ohio River. The load on the bridge deck is transferred to the rib cage through 16 hangers with intervals 45 feet (13.7 m) between the two piers on each side of the bridge.

Each hanger is made up of four strands of cable (Figure 2). To prevent the cables from impacting each other and damping out higher frequency vibrations, elastic spacers have been placed between the cables (Figure 2). The cables are fabricated from class A zinc-coated steel structural strand. The zinc-coated strand consists of layers of wire about a center wire. 8 inch PVC

pipe has been used as protectors on the lower part of the cables. Hence, all measurements are made at least above the protectors. The strand has a minimum breaking strength of 310 tons. The nominal diameter of the cable is about 2 ¼ inches (5.72 cm). Each strand is pre-stretched under tension to about 155 tons. The modulus of elasticity of the cable is about 24,000,000 psi (165.47 x 10^9 N/m^2). The hanger acts as stress transmitter, for the box arch to carry the dead load of the bridge deck. The design tension levels in each cable were not available, however, it was assumed that the tension level should not exceed 10% of the cable's breaking strength (31 tons). Hence, this limit was used as a check on cable safety.

3.2 Laser Doppler Vibrometry (Laser Interferometry)

The Ometron VPI 4000 Laser Vibrometer (LV) is a laser interferometry system that measures vibration based on the detection of the Doppler shift in the frequency of a laser beam reflected from a moving target. Figure 3 shows the basic configuration of a LV. A single source laser beam is split into a reference beam and a measurement beam. The measurement beam is shot on the target along the line of sight, and upon reflection is captured by a FM detector. The He/Ne beam is observed as a red spot on the target with a nominal size of ~ 2 mm in diameter. The mixing of the two signals at the detector gives the output voltage that is proportional to the speed of target. VPI 4000 is a He/Ne laser with a 1 mw output, hence a class II laser safe product. It is capable of operation up to 5 MHz and has a shooting range of up to 200 m under idealized condition.

The LV has been used extensively in automobile industry for vibration testing[9]. Comparing with conventional contact transducers, the LV offers several advantages:

1. Capability to access difficult to reach locations
2. No added mass to structure
3. Detection from a distance away
4. Relatively quick setup

However, LV application to large civilian structures, such as bridges and buildings has been not as extensive, due to the question of the shooting range of commercially available systems. Only limited literatures have reported using laser for vibration measurements on large structures: Cunha et al. [10] has attempted using LV to detect cable vibrations on a suspension bridge from 5 m away; Bourgard et al. [11] has done extensive measurements of building vibrations, including an eight story high steel frame building, a six story timber framed building and a concrete building.

3.3 Vibration Testing of Cables

Current testing was conducted with the LV located on the shoulder of the bridge deck (Figure 4). The laser was fixed on a tripod without a damper. Since we are interested in the transverse vibration (not in the direction along the cable length), it is assumed that the bridge deck vibration will not affect the horizontal measurements (in reality, large deflection due to heavy truck loads constantly kick the laser out of target). The vibration measurements were made under ambient conditions. The direct velocity data were collected by a Techtronic VXI system at a sampling rate of 1 kHz with 50,000 data points. Spectrum analysis is then conducted on the time domain signals. The environmental condition on the test date is sunny with mild breezes.

The excitations are typically due to traffic load and wind load. All hangers along the westbound traffic (towards Columbus, Ohio) have been tested; except for the first and last ones (R1-T1 and R1'-T1'). For each hanger, at least two strands were measured; this is to ensure that the vibration measurements are true representation of the actual cable behavior. Except for one incidence, where a strand has known loose problem, all cables of the same hanger have reasonably close vibration frequencies.

Figure 5 shows the basic structure of the I-470 Bridge with labels for the hangers. Table 1 lists the cables, their effective length and the fundamental vibration frequencies. The nominal lengths (design lengths) of the cables have been used as the effective length. The actual unconstrained lengths of the cables, are may be 0.5 m to 1 m shorter or longer than the nominal length, due to the gripping fixtures. Also, it is recommended to measure the cable lengths in-situ wherever possible, rather than using the as-built plans.

Also shown in Table 1 are the determined tension levels using equation (2). Assuming that each strand carries the same load, the total tension in each hanger would be four times the tension level in the strands. Several tests were conducted on strands from the same hanger to confirm similar frequencies exist in each strand. It is shown that all the tension levels in the hangers are within the required safety limits. The cables are conclusively secure for daily operation.

In most of the measured spectrums, bending harmonic frequencies are shown as peaks with the fundamental (first) mode as the most significant peak. For all collected data, a high Signal to Noise (S/N) ratio is observed, which shows that the LV is capable of collecting vibration signal even under sunny condition. Figure 6 shows the measured spectrum for cable 2 (the shortest cable measured), where the fundamental mode 7.8 Hz is identified along with second vibration mode of 15.7 Hz. Other unidentified non-harmonic peaks also exist (such as 1.5 Hz and 6.4 Hz), due to other induced vibration signals from the boundary, which can be reduced by measurements closer to the center of the cable. However, for short cables, such as cable 2, the boundary effects are unavoidable.

In general, the laser results collected are very clean with very low noise level. However, repetitions have to be made, when the cable swings out of sight of laser, which results in dropouts in the captured data. To capture the fundamental mode, the laser should reach the middle of the cable as much as possible. With our existing test setup, it is not possible to reach very high. In addition, due to the limitation of angle of reflection (assuming 15 degrees on either side along the line of sight), to reach a 9 m height, would require the laser to shoot from at least 110 ft (33.5 m) away.

4. DISCUSSION

Currently, there is no direct technique in determining tension levels in axially stressed members such as cables. A quick and reliable technique to determine the tension level of these members can provide significant information on the integrity of the whole structure. The test reported in this paper shows that using accurate measurements and based on simple correlation between frequency and tension level in cables, vibration measurements can provide realistic assessments. The use of laser interferometry technology becomes crucial, because it greatly simplified the testing procedure (the whole test only took 4 hours to complete). The distant sensing technique also provided safety to the bridge inspectors.

However, application of the technique may not be feasible for structures larger than the I-470 Bridge, due to the limitation in the shooting range of available lasers. With consideration of the limitation of allowable angles of reflection, to shoot on a high bridge, may require a laser to shoot from several hundreds of meters away. Existing systems may achieve 200 m with perfect reflective surface, which almost does not existing for all real life structures. A higher power laser, such as a long range CO_2 laser may overcome this problem.

Another limitation of the technique is the laser spot, which is only approximately 2 mm in diameter, is targeting at an inherently rough surface. Figure 7 shows the cables, which consist of intertwining strands, makes it difficult for the laser to dwell at the same spot. The high wind may also create large displacements, and put the laser out of target. A larger laser spot and a tracking technique should be implemented to enhance continuous target tracing.

Another problem is that laser is subjected to environmental effects. No study has been found in the literature on effects of exhaust gas from the tailpipes of automobiles, dust and air turbulence resulting from traffic motion, on laser transmission path. These effects may create losses in beam return through absorption, scattering and turbulence. Bruce [12] reported that air turbulence within 100 m distance might create phase variation in Doppler signals.

While the theory of cable vibration may be simple to adopt, the actual vibration phenomenon, however, has been complicated by the actual bridge setup: 1. The addition of spacers between cables may create constraints and damping effect on the bridge cable vibration; 2. The installment of cable collars right above the cable fixture may also create extra constraints on the cables; 3. The fixation of cables onto the bridge deck is critical in defining the boundary condition of the cables; and 4. the sagging of inclined cables due to self-weight.

Since our current study only deals with straight, vertical cables, the issue of geometric nonlinearity due to cable self weight is not related. However, to extend this technique to embrace all cable structures, the effect of geometric nonlinearity on cable tension/frequency correlation will have to be resolved. The assumption of boundary conditions made in the string equation, requires that the method apply to cables of definite lengths. In our study, it is clear that cable 1 is too short for the application, which due to boundary effect has very noisy signature.

Finally, actual testing of bridge cables using the LV located in shoulders exposes the bridge inspector to potential danger, a better idea would be to mount the laser permanently, at a location to provide continuous vibration monitoring.

5. CONCLUSION

In this paper, vibration measurement of bridge cables on a tied-arch bridge using laser vibrometer is reported. Through simple correlation, the vibration frequencies of the bridge cables have been translated to the tension levels within the cables, which indicate that the existing cables are in safe working condition. This test shows that the technique is much more efficient than contact measurements and provides accurate results for bridge assessment. The laser data are shown to have low S/N ratio. The test setup is also much easier and safer than contact measurements or other bridge assessment techniques.

Several limitations of existing laser technology have been identified. In particular, while this technique, with further modification of the testing approach, is ideal for application on tied-arch bridges such as the I-470 Bridge, a more powerful laser system may be required for larger structures. The basic theoretical basis of this technique, such as the tension equation, needs to be modified to embrace a wider variety of cable problems.

Nonetheless, the test results indicate great potential of LV technology in civil structure applications.

ACKNOWLEDGEMENTS

The authors wish to thank West Virginia State Division of Highway District 6 for their assistance in the bridge testing. The authors also wish to thank the Constructed Facilities Center at West Virginia University for supplying the laser system.

REFERENCES

1. C.M. Teller, G.A. Matzkanin and S.A. Suhler, *Nondestructive inspection of fatigue damage in a stay-cable specimen using the MPC system*, final report n. DTFH61-89-P-00874, FHWA, Nov 1989.
2. H. Kwun, G.L. Burkhardt and C.M. Teller, "Inspection of ropes and cables using the transverse-impulse vibration technique," *Review of progress in quantitative nondestructive evaluation*, Vol. 12, pp.2215-2222, 1993.
3. A.B. Mehrabi and H. Tabatabai, "Unified finite difference formulation for free vibration of cables," *Journal of Structural Engineering*, ASCE, v. 124(11), pp.1313-1322, Nov 1998.
4. H. Zui, T. Shinke and Y. Namita, "Practical formulas for estimation of cable tension by vibration method," *Journal of Structural Engineering*, ASCE, v. 122(6), pp.651-656, Jun 1996.
5. S. Petro, D. Reynolds, S.E. Chen and H.V. GangaRao, "Measurement of axial forces via natural frequency," *Structural Materials Technology III*, SPIE v3400, pp.82-91, 1998.
6. S. Petro, *Preliminary Marketing Plan: A non-destructive evaluation instrument for axial load measurements*, Final Report, West Virginia DOT, 1998.
7. J.W. Leonard, *Tension structures – behavior and analysis*, McGraw-Hill, 1988.
8. A. Wicks and A. Barker, *Notes to Laser Vibration Measurement Seminar*, Livonia MI, 1996.
9. D. Oliver, "Scanning Laser Vibrometers" *Automotive Engineering*, pp.71-75, March 1996.
10. A. Cunha, and E. Caetano, "Dynamic measurements on stay cables of cable-stayed bridges using an interferometry laser system," *Experimental Technique*, 38-43, June 1999.
11. A.J. Bougard and B.R. Ellis, "Laser measurement of building vibration and displacement," 3rd International Conference on Vibration Measurements by Laser Techniques: Advances and Applications, SPIE, v.3411, pp.3411-36, 1998.
12. R.A. Bruce, *Remote transducer based on coherent mixing of backscattered laser light*, Report of investigation n.8220, U.S. Bureau of Mines, 1977.

Table 1 Cable vibration frequencies and corresponding tension levels

CABLE No.	MASS/unit length (gm/cm)	AREA (cm^2)	EFFECTIVE LENGTH (m)	FREQUENCY (Hz)	TENSION (Tons)	TOTAL TENSION (Tons)	SAFETY LIMIT (Tons)
R1-T1	410.209	102.608	4.57	-	-	-	31-139.5
R2-T2	410.209	102.608	8.38	8	29.07	116.29	31-139.5
R3-T3	410.209	102.608	11.43	4.8	19.46	77.85	31-139.5
R4-T4	410.209	102.608	14.48	3.7	18.55	74.21	31-139.5
R5-T5	410.209	102.608	16.38	3.3	18.90	75.60	31-139.5
R6-T6	410.209	102.608	17.91	2.9	17.44	69.75	31-139.5
R7-T7	410.209	102.608	19.05	2.7	17.11	68.42	31-139.5
R8-T8	410.209	102.608	19.43	2.5	15.26	61.03	31-139.5
R8'-T8'	410.209	102.608	19.43	2.4	14.06	56.25	31-139.5
R7'-T7'	410.209	102.608	19.05	2.8	18.40	73.56	31-139.5
R6'-T6'	410.209	102.608	17.91	2.7	15.11	60.46	31-139.5
R5'-T5'	410.209	102.608	16.38	3.3	18.90	75.60	31-139.5
R4'-T4'	410.209	102.608	14.48	3.5	16.60	66.41	31-139.5
R3'-T3'	410.209	102.608	11.43	5	21.12	84.47	31-139.5
R2'-T2'	410.209	102.608	8.38	7.8	27.64	110.55	31-139.5
R1'-T1'	410.209	102.608	4.57	-	-	-	31-139.5

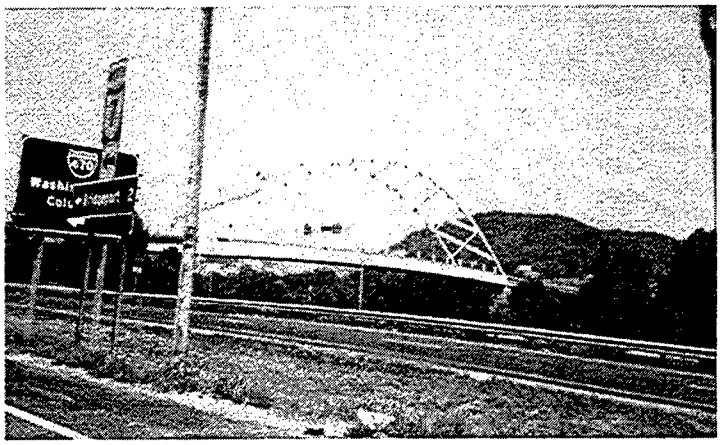


Figure 1 The I-470 bridge

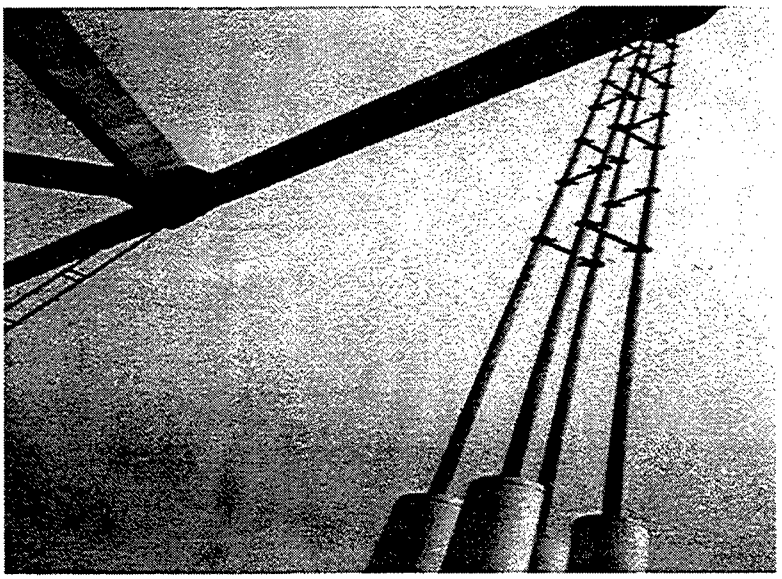


Figure 2 Hanger with four cables and interlace separators

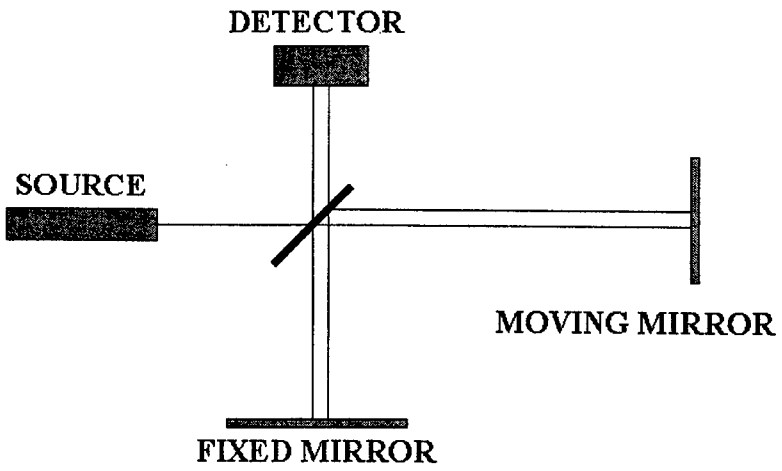


Figure 3 Basic Laser Interferometry[8]

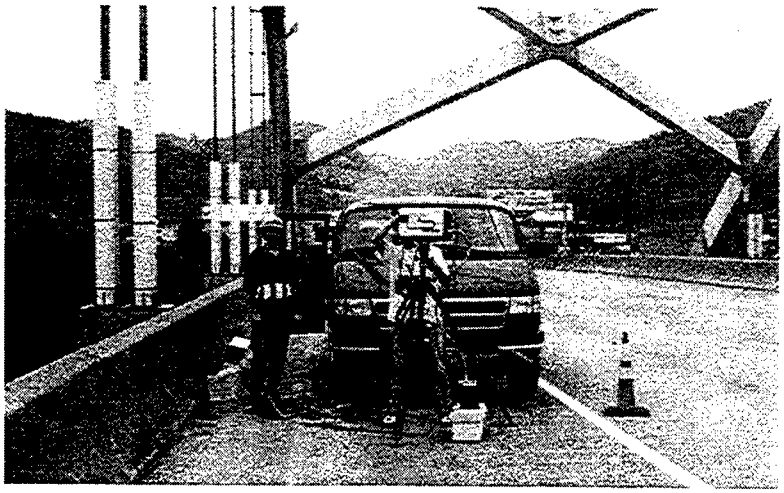


Figure 4 Bridge test setup

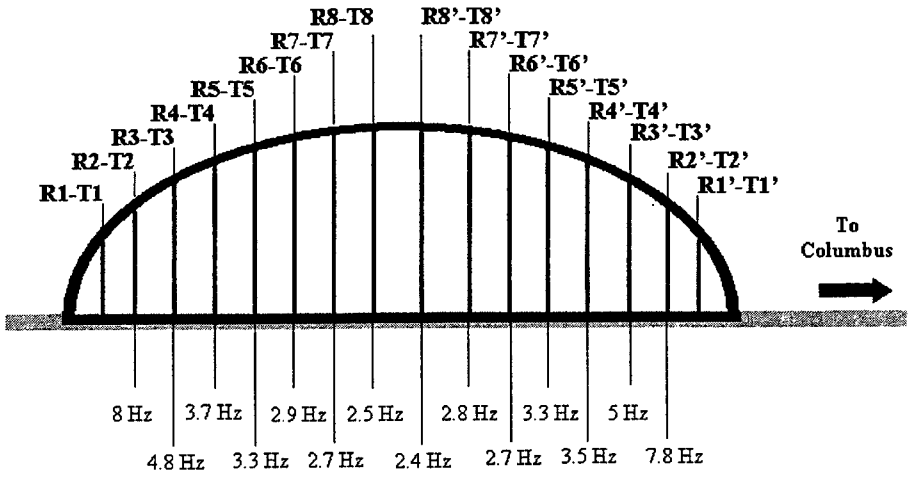


Figure 5 Fundamental vibration frequencies of the bridge cables

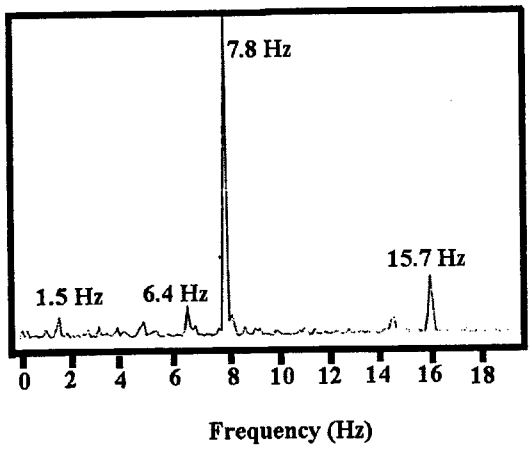


Figure 6 Spectrum for Cable R2'-T2'

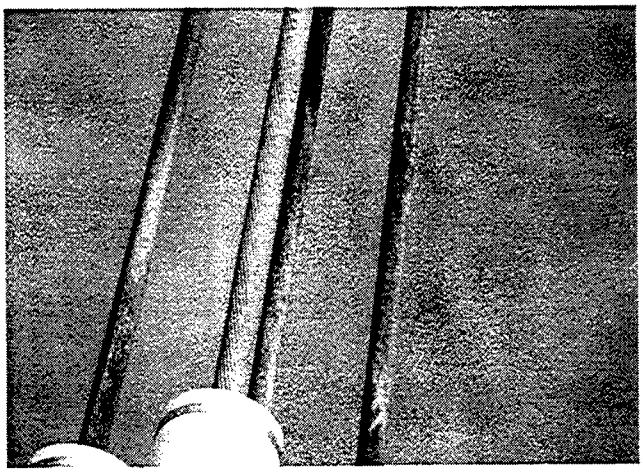


Figure 7 Texture on the cable due to intertwining wire strands

SESSION 14

Novel Sensor Designs I

LOW COST OPTICAL LIQUID LEVEL SENSORS

Prof. Anjan K. Ghosh, Navjot Singh Bedi and Prashanta Paul

Dept. of Electrical Engineering
Indian Institute of Technology Kanpur
Kanpur – 208016, India.
e-mail: anjn@iitk.ac.in
Fax: +91-512-590063

ABSTRACT

Design of simple, low cost optical sensors for continuous liquid level measurements is discussed. The sensor consists of only a source and a photo-receptor placed in a cylindrical tube containing the liquid. Various parameters of this simple package are selected through diffraction calculations.

Keywords: Optical Sensors, Packaging of Optical Sensors, Optical Liquid Level Sensing.

1. Introduction

Condition monitoring and measurement systems are important for automation, critical for modern technology. Sensors are key components in these systems. As the push towards automation continues, the design and development of sensors are growing. Optical technologies have long played an important role in instrumentation and sensors, especially for non-contact measurements. In an optical sensor, the information is passed from the physical/ chemical variable being measured to an optical beam. The changes in the characteristics of the optical beam are then determined. Owing to this structure, optical sensors are immune to EMI/ RFI, have a high bandwidth, easy for remote control or operation and have a small,compact size.

In the developed nations of the world, optical sensor industry is a multimillion-dollar business and is projected to grow into a 560 million-dollar industry by the turn of the century [1, 2]. Yet in India, optical sensors represent an emerging field. Most of the optical components are generally imported which makes them costly. Thus, there is a need for optical sensor design and development. The optical sensor design should be cost effective for it to be acceptable to the Indian market.

In general optical sensors consist of: (a) Optical Sources, (b) Sensing elements through which or in which the physical or chemical variables and optical fields interact, (c) Photo-receptors, and (d) Electronic circuitry for signal processing and display [1, 2]. The components (a) - (d) are combined in a package. The performance of an optical sensor depends on the relative placement of all mechanical and electrical components in a package. For an optical sensor desired to be cost effective and simple, it should comprise of the fewest possible number of optical components, inexpensive transducers with which the light beam interacts, and a simple low cost package. Apart from cost, a package has many roles in determining the overall performance of a sensor, such as: (a) alignment of optical, opto-mechanical and opto-electronic components [3], (b) thermal stability, (c) shielding from stray radiation, (d) providing proper spacing between the components, and (e) overall reliability.

A low cost optical liquid level sensor that can be easily fabricated and the possible packages to implement such a sensor are analyzed in this paper. The proposed sensor consists of only a source, a receptor and a tube to keep the liquid in it. The source can be a light emitting diode (LED), a laser diode (LD), or the tip of an optical fiber bringing light from a distance source. The receptor can be a PIN photodiode, an avalanche photodiode or tip of an optical fiber carrying the light

Part of the SPIE Conference on Advanced Photonic Sensors and Applications
Singapore • November–December 1999
SPIE Vol. 3897 • 0277-786X/99/$10.00

to a distant photodiode. In this sensor the intensity of light beam detected by the photodetector changes as a function of the liquid level because of the reflection/ refraction of light from the liquid surface and/or absorption in the liquid. No other transducer is needed in this type of sensor keeping the overall cost low. If a lens were used it will improve the light collection efficiency but it will make the package design more complex. Thus no lens is included in our design. This type of liquid level sensors can be used in the automobile industry, in the aeronautical industry and in measurement of the level of oil in electrical transformers.

While considering packaging, we take into account the relative importance of the size of the source and detector in the package and the overall arrangement of these items. The effects of variation of the dimensions of the container tube are considered. We observe how the output of the sensor varies with the following parameters related to spacing and sizes and the arrangements: (a) Beam spot size, (b) Beam divergence angle, (c) Height & radius of base of cylindrical package, (d) Detector size, and (e) Surface tension, density, contact angle and refractive index.

2. Simple Optical Liquid Level Sensors

There are four different ways in which a source and a detector can be arranged, to design a simple low cost optical liquid level sensor. In the first arrangement the source is kept at the top of the tube containing the liquid and the photo-receptor at the bottom of the tube. This topology will hereafter be referred to as Sensor Configuration No. 1 and is shown in Fig. 1. As depicted in Fig. 2, the receptor is kept at the top of the tube containing the liquid and the source is at the bottom of the tube in the second arrangement. In the third arrangement, both the source and receptor are kept at the top of the tube containing the liquid. This topology will hereafter be referred to as Sensor Configuration No. 3 and is shown in Fig 3. The bottom of the tube should be painted / coated so that no reflection can take place from the bottom. The light reflected by the liquid surface only should reach the photo-receptor. In the last arrangement, Configuration No. 4, both the source and detector are kept at the bottom of the tube as shown in Fig. 4. Of these four topologies listed above, the first two are examples of transmissive type. The last two, that is Configuration nos. 3 and 4, are reflective type of sensors.

3. Analysis and Design

For the package design of our optical level sensors, the possible architectures in Figs. 1-4 have been analyzed using diffraction theory.

The shape of the liquid meniscus in any tube is curved owing to surface tension [4]. The meniscus is concave or convex upwards depending on whether the liquid wets or does not wet the solid container [4]. The meniscus is assumed to be spherical in shape with a radius of curvature R which can be expressed in terms of the tube radius r_t as

$$R = (r_t^2 + s^2)/(2s). \tag{1}$$

where s is the height to which the liquid rises while wetting the tube walls. The height s is given by [4]:

$$s^2 = 2F_t(1 - \sin\theta)/\rho g \tag{2}$$

where θ is the angle of contact between the liquid and the tube wall, ρ is the density of the liquid in kg/m^3, g is the acceleration due to gravity and F_t is the surface tension of the liquid.

We assume that the source emits a linearly polarized circularly symmetric gaussian beam. This is almost true for laser light coming out of a monomode optical fiber. The radiation from a good quality LED or the LED light radiated by a multimode fiber can also be approximated by a gaussian beam. A laser diode would emit an elliptical gaussian beam. With simple optics attached with the laser it can be converted to a circularly symmetric gaussian beam. The intensity of the circularly symmetric Gaussian beam at a distance z from the output end of the source is given by [5]

$$I(r,z) = [2P_s/\{\pi w^2(z)\}] \exp[-2r^2/w^2(z)]. \tag{3}$$

where the spot radius at the distance z is obtained from

$$w(z) = w_0\sqrt{1+(z\lambda/\pi n w_0^2)^2} \tag{4}$$

w_0 is the minimum spot radius of the beam at z = 0. $w_0 << r_t$ so that optical field reflected from the side walls and interfering with the main beam at the detector plane is negligible.

If the beam falls on a circular photodiode of radius d and responsivity γ, the power coupled P_d is given by

$$P_d = \gamma P_s[1-\exp(-2d^2/w^2(z))]. \tag{5}$$

In eqn. (5) we have neglected the absorption loss inside the medium and any loss due to reflection at the medium-detector interface. Thus for a fixed value of d and P_s the power received at detector P_d is a function of w(z) only. The expression for P_d given by eqn. (5) is based on the assumption that there is no offset between the source and detector . While it may be possible to achieve this condition in case of the first two configurations (i.e. transmissive type of sensors), but in configuration no's 3 and 4 (i.e. reflective type of sensors) where both the source and detector are on the same side of the tube, it may not be possible. Even if the source and detector are kept adjacent to each other, still the center of the detector will be offset from the center of the source by a distance equal to the sum of the radii of the source and the detector. The power coupled from a circularly symmetric gaussian beam to a circular photodiode with an offset q along the x-axis is given by

$$P_{dq} = [2\gamma P_s/\{\pi w^2(z)\}]\exp[-2q^2/w^2(z)]\int_0^{2\pi}\int_0^{d}\exp[-2(r^2-rq\cos\theta)/w^2(z)]rdrd\theta. \tag{6}$$

This integral is usually evaluated numerically though an approximate close form can be derived. Because of circular symmetry the x-axis can be always chosen to be the axis joining the centers of the beam and photodetector.

As shown in Fig 1, if some liquid of refractive index n_2 partially fills up the gap of length z between the optical source and photodiode, then the beam propagates partially through air and partly through a liquid column of height h. The light travelling from the top of the tube encounters a concave dielectric interface. Thus for configuration no. 1 we find that the spot size at the photodetector is given by

$$w_{31} = w_0\sqrt{[1+(n_2-n)h/(n_2R)]^2+[\{[(n_2-n)h/(n_2R)+1](z-h)+(n/n_2)h\}\lambda/(\pi n w_0^2)]^2} \tag{7}$$

where R is the radius of curvature of the liquid surface.

In configuration no 2, the detector is at the top of the tube and the source is at the bottom as shown in Figure 2. It can be shown that for a laser diode with a small divergence angle there will be almost no total internal reflection and we obtain the expression for the spot size at the detector

$$w_{32} = w_0\sqrt{[1+(n_2-n)(z-h)/(nR)]^2+[\{[(n_2-n)h/(nR)+n_2/n](z-h)+h\}\lambda/(\pi n_2 w_0^2)]^2} \tag{8}$$

In the sensor configuration no 3 with both the source and detector at the top of the tube the expression for the spot size at the detector is given by

$$w_{33} = w_0\sqrt{[1-2(z-h)/R]^2+\{2(z-h)\lambda/[\pi n w_0^2]\}^2[1-(z-h)/R]^2} \tag{9}$$

where h is the height of liquid column from the bottom of the tube. In the sensor configuration no. 4 with both the source and detector at the bottom of the tube the spot size at the detector is given by

$$w_{34} = w_0\sqrt{(1+2h/R)^2+(2h\lambda/[\pi n_2 w_0^2])^2(1+h/R)^2}\ . \tag{10}$$

We assume that the glass plate used at the top or bottom of the tube is not optically flat and so multiple reflections in these plates are not considered in this paper. Also, for reducing multiple reflections, light absorbing paint or coating is

assumed to be applied on interior surface of the tube wall or the glass plates at the top or the bottom. Unless otherwise mentioned the lateral, longitudinal or angular offsets between the source and the detector are assumed to negligibly small. However there are reflection losses at each (uncoated) dielectric interface in the path of the light. Thus from eqn. (5) the power coupling efficiency, in the configuration no. 1 is given by

$$\eta_1 = \gamma T1[1 - exp(-2d^2/w_{31}^2)] \exp(-ah) \tag{11}$$

where a is the absorption coefficient of the liquid at the wavelength λ and T_1 is the total transmission loss consisting of products of transmission losses at the air - glass, glass- air, air- liquid, liquid -glass, glass-air, air- photodetector interfaces as the light propagates from the top to bottom in Fig 1. If the photodetector has a glass/quartz window, the transmission loss at that window should also be included in T_1. The power coupling efficiency of configuration 2 is same as that given in eqn. (11) with w_{31} replaced by w_{32} in eqn. (18). The power coupling efficiency of the reflective configuration no 3 is given by

$$\eta_3 = \gamma T_3 \rho_3 (P_{dq3}/P_s) \tag{12}$$

where ρ_3 is the reflection coefficient of the air liquid interface, T_3 is the total transmission loss from travelling through air - glass, glass - air, air - glass, glass - air and air - photodetector interface, P_{dq3} is obtained from eqn. (6) by putting $w(z) = w_{33}$ and offset = q in it. Similarly the power coupling efficiency of the reflective configuration no. 4 is given by

$$\eta_4 = \gamma T_4 \rho_4 \exp(-2ah) P_{dq4}/P_s \tag{13}$$

where ρ_4 is the reflection coefficient of liquid-air interface, T_4 is the total transmission loss from travelling through air - glass, glass - liquid, liquid - glass, glass - air and air - photodetector interfaces. P_{dq4} is obtained from (6) by putting $w(z) = w_{34}$. Because of the absorption factor exp(-2ah) the power coupling efficiency of configuration (4) would be the least of that amongst all the configurations and so this configuration not considered any more in this paper

4. Numerical Results

The tube length is selected to be 115 mm with dead zones of 10 and 5 mm at the top and bottom of the tube, respectively. The diameter of the tube containing the liquid is chosen to be 23 mm. Numerical calculations are performed using water as the liquid though testing was done with data on other liquids such as gasoline, kerosene, alcohol, etc. We chose to work with visible sources, LED, LD or fiber bundles transmitting laser beams. The beam is polarized in a transverse direction. The wavelength selected for the numerical calculations shown in this paper is 670 nm. The transmission/ reflection losses at the various interfaces and the absorption coefficient a of the liquid are dispersive in nature. Calculation of power transfer is an involved calculation involving spectroscopic data on the source, the liquid and the photodetector. So we assume that the light beam used is monochromatic. The wavelength selected for the numerical calculations shown in this paper is 670 nm. The radius of the source is about 1 to 2 mm. The power output of the source can be in the range of 1- 2 mW, adequate for the short length of our optical sensor packages. The minimum spot size of the gaussian beam at the operating wavelength is assumed to be 100 microns. However, in our numerical experiments most of the parameters are kept variable, especially the radius of the PIN photodetector and the liquid level.

The variation of the power coupling efficiency in configuration nos. 1 and 2 with liquid level are plotted in figures 5 and 6, respectively. It can be observed from these plots that in case of the transmissive sensors the power decreases continuously as the liquid level h is increased. With an increase in h, the cone of the light is spread over a larger area on the detector plane and thus the intensity of the light being received at the detector decreases. If a large area photodetector is used we can have an almost linear variation of coupling efficiency with liquid level. The coupling efficiency increases with detector radius. But in the configuration no. 1 and 2 we need to place an optical window at the bottom of the tube and it may not be possible to dip the sensor in the liquid tank. These two configurations are also highly sensitive to offsets in the positions of the source and the detector. Any tilt of the tube causes an angular offset. Most serious is that if there is any particulate impurity in the liquid these particulates will block the light path or scatter light away decreasing the coupling efficiency. So we think configuration numbers 1 and 2 are not really viable for low cost optical liquid level sensor designs.

The variation in the coupling efficiency vs. liquid level for configuration no 3, is plotted in Fig. 7 keeping the radius of detector and source at 2 mm and 1 mm respectively. The offset is varied from the minimum value of 3 mm to 6

mm. From Fig. 7, it is found that the response can be divided into two regions - the front slope and the back slope. The front slope of the plots pertains to the region of partial overlap between the detector and the cone of light reflected from the liquid surface and the back slope of the plot pertains to the region of complete overlap. The coupling efficiency increases with liquid level h in the region of front slope. This is due to the fact that initially, as the liquid level increases from the bottom of the tube the reflected beam gets more focused owing to the reflection from a concave surface. Also more and more area of the photodetector gets covered by the reflected beam. At a certain liquid level, the light reflected and focused from the liquid surface falls on the complete area of the detector and maximum power gets coupled to the detector. After this stage, though there is no further increase in the area of the detector which receives light, the intensity of light falling per unit area decreases and thus the efficiency decreases in the region of back slope of the plot.

In sensor configuration no. 3 the power received at the detector is less as compared to that in configuration nos. 1 and 2 and the coupling efficiency varies from 1% to 14%. Another disadvantage of this configuration is that the range of the measurement is limited with the two slopes and there is a dead zone. Despite this obvious disadvantage, this type of sensor has advantages such as: (i) The reflection from the top surface is not seriously affected by impurity particles in the liquid. only floating impurities can affect, and (ii) Packaging is easy with the source and detector both at the top along with all the electronics for driving the laser and amplifying the photodetector output. This sensor can look in to a tank and read the liquid level.

To appreciate the effects of the various design parameters on the coupling efficiency in this configuration we performed numerical experiments. Keeping all other design parameters the same, one particular parameter was varied and its effect was analyzed. In Fig. 8 coupling efficiency with respect to liquid level h has been plotted for values of offset varying from 3 mm, 4 mm, 5 mm, 6 mm for a given radius of the detector (d = 2 mm). It was observed that as the offset q is made to vary from 3 mm to 6 mm efficiency decreases sharply. With an increase in offset q, the amount of light being coupled to detector decreases. It is also seen that as the offset is increased the dead zone of the sensor decreases and the effective range over which sensing is possible decreases. We also notice that the maximum efficiency achieved decreases and the maxima of the efficiency curve occurs for smaller value of h. Thus configuration no. 3 is a good possible configuration for measuring lateral offsets as well.

The minimum spot size of the gaussian beam is varied from 0.1 mm to 1 mm and the variation in coupling efficiency with respect to the liquid level is plotted in Fig 9. It is seen that the coupling efficiency increases with an increase in the spot size as more power is being coupled to the detector. It is also seen that as the spot size increases, the range of measurement of the sensor also increases and there is a smaller dead zone. The maximum efficiency occurs for a larger value of h. We found that the coupling efficiency increases with the detector diameter without any change in the position of the maximum point with respect to h.

The values of the refractive index of three fluids were selected to study the variation in coupling efficiency with respect to liquid level. As shown in Fig 10, there is only a slight amount of variation in the coupling efficiency on varying the value of refractive index of the liquid. It is seen that efficiency decreases with an increase in the refractive index of the liquid. This is because with an increase in the value of n, the reflection coefficient decreases slightly and the spot size is not changed. Thus the operation of this reflection based sensor is almost independent of the type of liquid used. We need not recalibrate the sensor with a change in the liquid temperature or pH or concentration of a solute dissolved in it.

On varying the diameter of the tube, there is a change in the radius of the curvature of the liquid surface which in turn affects the coupling efficiency. From Fig. 11 it is seen that as the tube diameter is increased, which implies that as the liquid surface tends to become flat, the maximum amount of power gets coupled to the detector at a lesser value of liquid level. It is also seen that as the value of tube diameter increases the gradient of the back slope of the plot between coupling efficiency and liquid level decreases and the back slope region is suitable for the purpose of sensing.

5. Conclusions

In this paper we described how a simple optical liquid level sensor can be fabricated by using a light source emitting a narrow beam and a photodetector along with a tube containing some liquid. The light beam from the source either

travels through the tube or is reflected from the liquid surface. We derived basic formulas that show how the power received by the photodetector is related to the design parameters such as the height of the liquid column, the radii of the tube and the photodetector, refractive index, density and surface tension of the liquid, the spot size and divergence angle of the light beam, and the offset between the source and the photodetector. These formulas are useful for designing the package of the simple and low-cost optical sensor. We recommend a reflection based optical configuration. We find that in this configuration the power coupled to the photodetector is not easily affected by the temperature of the liquid or by the presence of particulates or impurities in the liquid. If multiple detectors are used for obtaining a better coupling efficiency or misalignment tolerance or for sensor fusion we can use the above formulas and results to obtain the design data for the package or the overall sensor. Thus such optical liquid level sensors can be easily deployed in aircraft or vehicles.

Acknowledgments

This project is supported in parts by a grant from the Dept. of Electronics, Govt. of India.

References

[1] J. Dakin and B. Culshaw, "Optical Fibre Sensors, Applications, Analysis and Future Trends", Volume 4, Artech House, Boston, 1997.
[2] B. Pal (Ed.), "Fundamentals of Fiber Optics in Telecommunication and Sensor Systems", Wiley Eastern, New Delhi, 1992.
[3] A.K. Ghosh and P.K. Paul, "Alignment Considerations in Extrinsic Fiber Optic Sensors", Applied Optics, Vol. 36, pp. 6256-6263, Sept. 1997.
[4] F.C. Champion and N. Davy, "Properties of Matter", Blackie, London, 1960.
[5] A. Yariv, "Optical Electronics", Holt, Rinehart and Winston, New York, 1985.

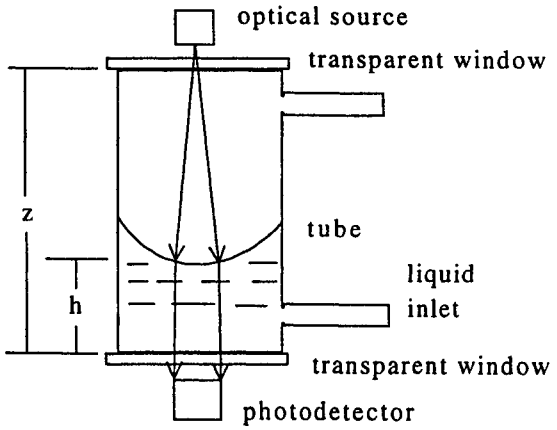


Fig. 1: Schematic Diagram of the Basic Optic Liquid Level Sensor in Configuration No. 1 with the Source at the Top and the Photodetector at the Bottom of the Tube.

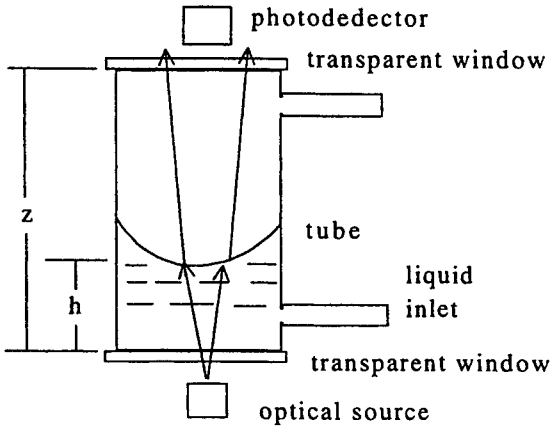


Fig. 2: Schematic Diagram of the Basic Optical Liquid Level Sensor in Configuration No. 2 with the Source at the Bottom and the Photodetector at the Top of the Tube

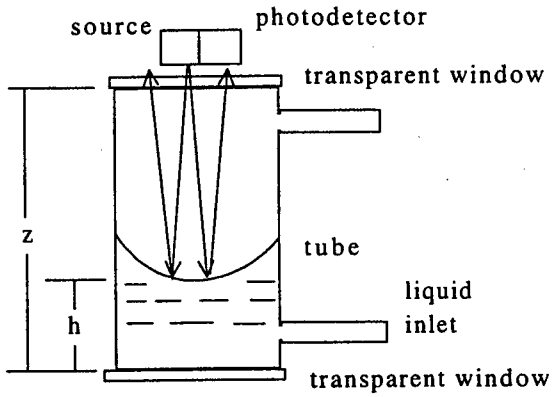


Fig. 3: Schematic Diagram of Basic Optical Liquid Level Sensor of Configuration No. 3 with both the Source and the Photodetector on the Top of the Tube.

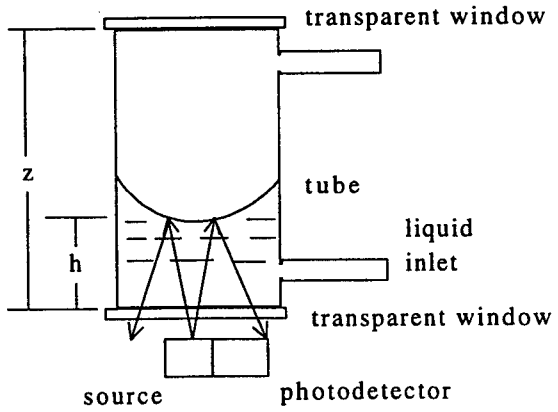


Fig. 4: Schematic Diagram of Basic Optical Liquid Level Sensor of Configuration No. 4 with both the Source and the Photodetector at the Bottom of the Tube.

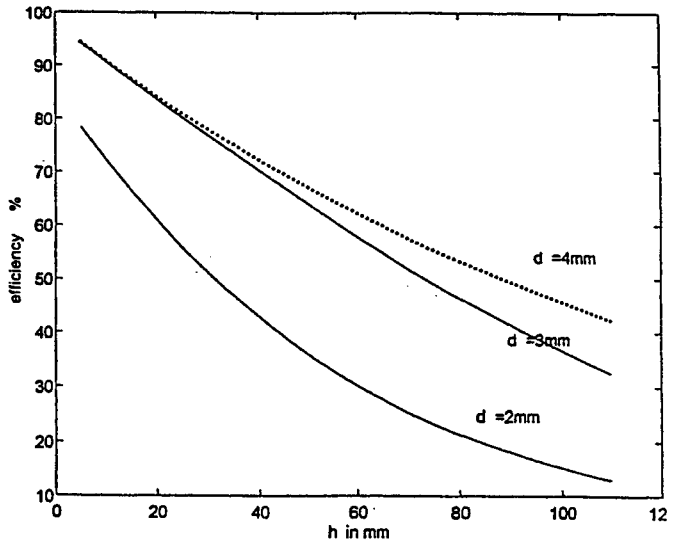


Figure 5: Variation of the coupling efficiency in the configuration no. 1 with the liquid column height for several values of the photodetector radius.

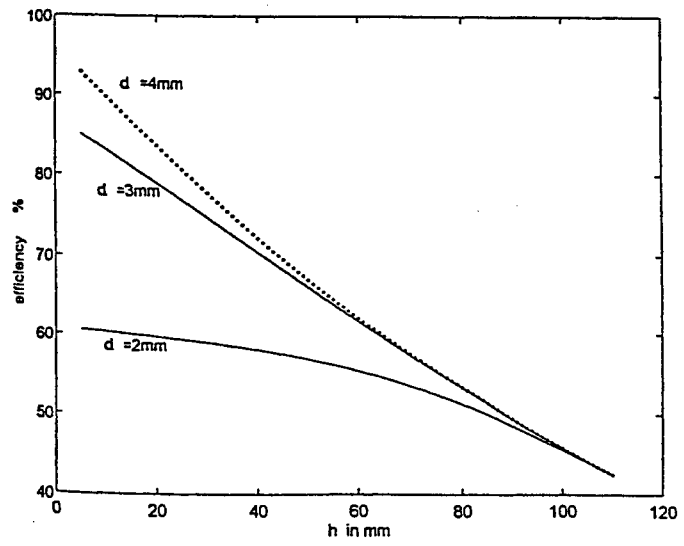


Figure 6: Variation of the coupling efficiency in the configuration no. 2 with the liquid column height for several values of the photodetector radius.

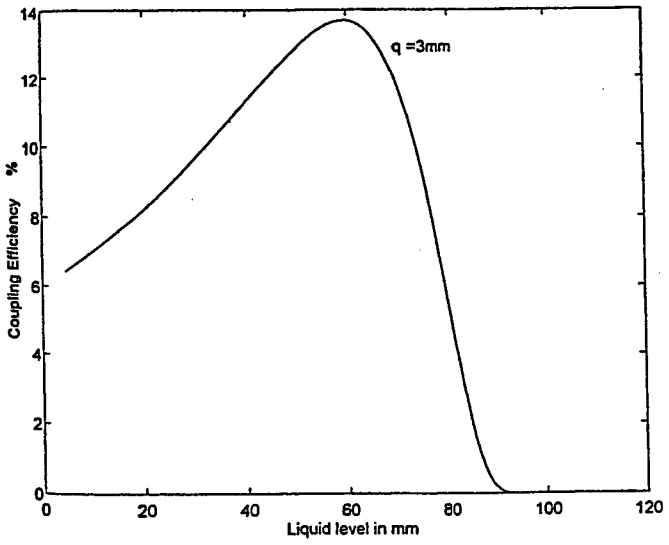


Figure 7: Variation of the coupling efficiency in the configuration no. 3 with the liquid column height. There is an offset of 3 mm between the centers of the source and the photodetector.

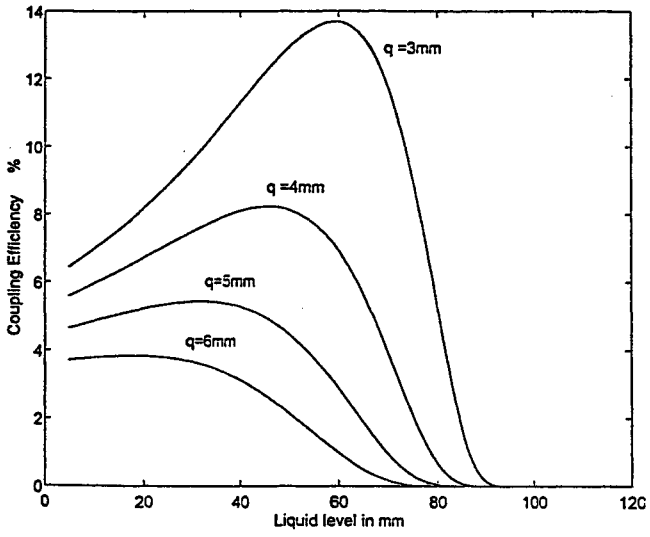


Figure 8: Variation of the coupling efficiency in the configuration no. 3 with the liquid column height for several values of offsets between the centers of the source and the photodetector.

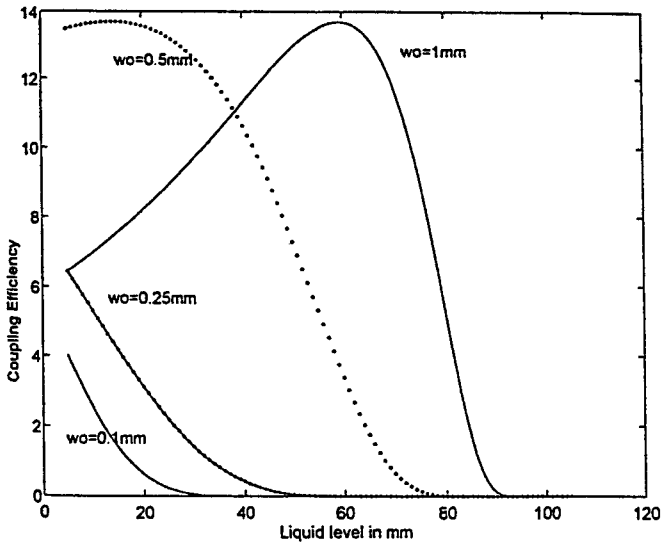


Figure 9: Variation of the coupling efficiency in the configuration no. 3 with the liquid column height for several values of the minimum spot size w_0 of the source.

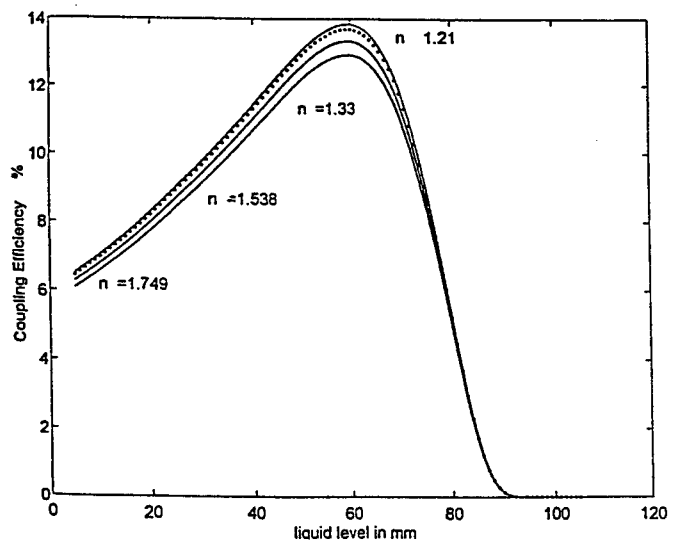


Figure 10: Variation of the coupling efficiency in the configuration no. 3 with the liquid column height for several values of the refractive index of the liquid.

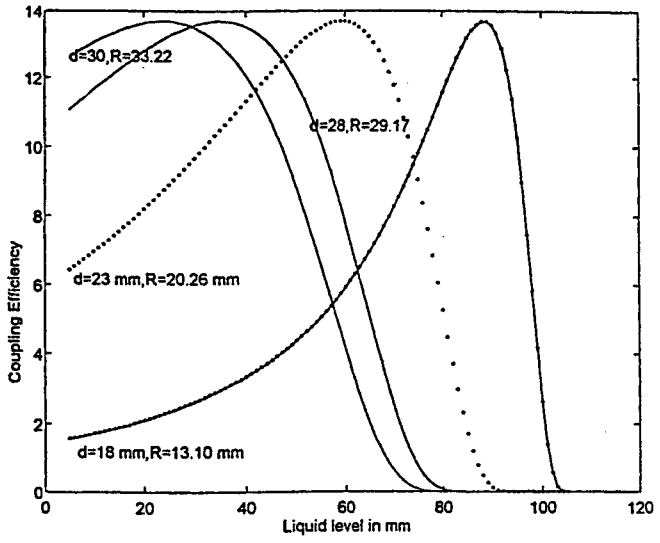


Figure 11: · Variation of the coupling efficiency in the configuration no. 3 with the liquid column height for several values of the radius of the tube and corresponding values of the radius of curvature of the liquid surface.

Optical Surface Plasmon Resonance Sensor Designs

Simanchalo Panigrahi*[a], N. B. Das[a], Aseel K. Hassan[b] and Asim K. Ray[b]

[a]Department of Physics, R.E. College, Rourkela, Orissa, India –769008

[b]School of Engineering, Sheffield Hallam University, England, (UK)

ABSTRACT

Beginning from low level concepts, the basic Physics behind the optical excitations of surface plasmon Polaritons (Surface plasmons) by volume electromagnetic wave (photon) under resonance prism coupling condition using both Otto and Kretchmann configurations is explained. Various changes to surface plasmon resonance (SPR) profile due to any modification occurring at metal/ dielectric interface are discussed. Organic thin film of Calix (4) resorcinarene are deposited on gold coated glass slides by spin coating technique The surface plasmon resonance angle shift due this organic layer as compared with bare gold film predicts the sensitivity of surface plasmons to any modification occurring at the metal/dielectric interface. Investigations are made to study the change of ambient close to the film. By the changing the layer system from glass/gold/calix/air to glass/gold/calix/water, the SPR curve is greatly changed. Exploitation of these changes on designing different optical sensors (gas, chemical, and temperature) are reviewed. Finally a short summary of potential device application is given.

Keywords: Surface Polaritons, Surface Plasmon Resonance, Otto configuration, Kretchmann configuration, Optical Sensors, dielectric constant, Spin coating, Bio gas chemical sensors, Optical excitations, Attenuated Total reflection.

1. INTRODUCTION

Surface Physics has opened up exciting opportunities to study the surfaces and interfaces of the materials, which are technologically important. Condensed matters interacts with environment consisting of vacuum/gas/vapor or with other condensed system through surfaces and interfaces respectively. Exposure of surfaces to different ambient conditions generates interesting physical and chemical interactions for basic research. Single lays or multilayered thin films exhibit many exotic properties, which are not found in bulk materials.

The interaction of volume electromagnetic wave (photon) with surface polarizing medium or dipole oscillations propagating along the interface generates polaritons or excitons like surface plasmons, surface phonons, surface excitons or surface magnones. If the interaction of incident photon is with free electrons like plasma of metal surface, the polaritons are called surface plasmons (SPs). Hence, surface plasmon is a surface electromagnetic (EM) field charge density oscillation that may exist at the interface between a metal and dielectric. The EM field is maximum along the interface and decay exponentially on both sides of the interface. The wave vector k_{SP} associated with surface plasmon is very sensitive to any modification such as surface roughness, thin dielectric coating, sample film, exposure of dielectric film to different gases and ambient temperature changes etc., occurring at the interface. It should be noted that SPs has only electric field component normal to the surface. They are TM waves and can therefore be excited by P-polarized TM light.

Organic thin films especially ordered organic films in the thickness range from monolayer to micron provide specialized surfaces, membrane interfaces and optical response function for which no inorganic analog exists. So the need for understanding thin film surface / interface phenomena involving organic material has grown unlimitedly with their increased use in polymer, synthetic material, micro electronics, corrosion, adhesive, inorganic applications and many others.

The drastic change in SPR parameters –SPR resonance angle, Resonance depth and Resonance half –width due to any modification occurring at the interface (ambient change) can be exploited to design various sensors –gas, bio, chemical, temperature and immuno etc. The merit of these sensor over other sensors are due to their non destructive nature and ability to sense low concentration of pollutants. The various sensor application includes (a) detection of small amount of poisonous gas in chemical plants, organic pollutants and organic pollutants in samples (b) monitoring small of Nox in ambient air, chemical contamination of metal surface.

*Correspondence: Fax – 0091 – 661 – 571169 or 572926

Part of the SPIE Conference on Advanced Photonic Sensors and Applications
Singapore • November–December 1999
SPIE Vol. 3897 • 0277-786X/99/$10.00

2. PHYSICS OF SURFACE PLASMON POLARITONS

We consider the interface between two semi-infinite, lenear, isotropic and non-magnetic media with dielectric constants ε_1 and ε_2. We may study two orthogonal polarizations of the electromagnetic wave (other polarizations can be resolved into these two) i.e.

1) Transverse magnetic (p-polarized) with **E** lying in the x-z plane.
2) Transverse electric (s-polarized) with **E** lying in the y-direction.

We start by considering the case of a p-polarized wave propagating in the x-direction, along the direction, along the interface. We can describe the electric fields associated with this mode by

$$E = (E_{1x}, 0, E_{1z}) \exp[iKx + \alpha_1 z - i\omega t] \text{ for } z < 0, \quad (1)$$

$$E = (E_{2x}, 0, E_{2z}) \exp[iKx - \alpha_2 z - i\omega t] \text{ for } z > 0, \quad (2)$$

Where α_1 and α_2 are the decay constants in the two media 1 and 2 respectively, K is the wavevector associated with the surface polariton of angular frequency ω. These equations must both satisfy the wave equation so that,

$$\nabla^2 E + \varepsilon_1 (\omega^2/c^2) E = 0 \text{ for } z < 0, \quad (3)$$

$$\nabla^2 E + \varepsilon_2 (\omega^2/c^2) E = 0 \text{ for } z > 0, \quad (4)$$

In this situation there are no free charges so that,

$$\nabla . E = 0, \quad (5)$$

and by applying the wave equation we obtain

$$K^2 - \alpha_1^2 = (\omega^2/c^2) \varepsilon_1 \text{ for } z < 0, \quad (6)$$

$$K^2 - \alpha_2^2 = (\omega^2/c^2) \varepsilon_2 \text{ for } z > 0, \quad (7)$$

Further we have two boundary conditions.

a) at $z = 0$, E_x is continuous so that

$$E_{1x} = E_{2x}, \quad (8)$$

b) at $z = 0$, D_z is continuous and so

$$\varepsilon_1 E_{1z} = \varepsilon E_{2z}. \quad (9)$$

Applying equation (5) to (1) and (2) we find

$$iKE_{1x} = -\alpha_1 E_1 \text{ for } z < 0, \quad (10)$$

$$iKE_{2x} = +\alpha_2 E_{2z} \text{ for } z > 0. \quad (11)$$

Thus equation (9) becomes

$$\varepsilon_1 \left(\frac{-iK}{\alpha_1} \right) E_{1x} = \varepsilon_2 \left(\frac{iK}{\alpha_2} \right) E_{2x},$$

So that,

$$\varepsilon_1 \alpha_2 E_{1x} + \varepsilon_2 \alpha_1 E_{2x} = 0. \quad (12)$$

For a non -trivial solution of this equation and equation 3.8 we require

$$\varepsilon_1 \alpha_2 + \varepsilon_2 \alpha_1 = 0. \quad (13)$$

Combining equations 3.6, 3.7 and 3.13 we obtain the dispersion relation,

$$K^2 = \frac{\omega^2}{c^2} \left(\frac{\varepsilon_1 \varepsilon_2}{\varepsilon_1 + \varepsilon_2} \right). \quad (14)$$

A similar calculation for s-polarized waves results in no dispersion relation and we conclude that surface-polaritons are p-polarized.

2.1 Existence Conditions For Surface - Polartons

For the surface mode to exist we require α_1 and α_2 to be positive and real, assuming for the moment that the media are lossless. From equation 3.6 and 3.7 we have.

$$K^2 > \frac{\omega^2}{c^2} \varepsilon_1, \frac{\omega^2}{c^2} \varepsilon_2 . \quad (15)$$

With the aid of equation (11) these conditions can be written as.

$$\frac{\varepsilon_1}{\varepsilon_2} . \frac{\varepsilon_2}{\varepsilon_1} < 0 . \quad (16)$$

Thus one existence condition of the surface modes is that dielectric constants of the two media should be of opposite sign. Having established this, equation (14) tells us that, if $\varepsilon_1 < 0$ then $|\varepsilon_1| > \varepsilon_2$ and if $\varepsilon_2 < 0$ then $|\varepsilon_2| > \varepsilon_1$.

The medium with the -ve dielectric constant is usually called the active medium. The other being the inactive one. We will take medium 1 as our active medium. Generally, ε_1 is complex that is.

$$\varepsilon_1 = \varepsilon_1' + i\varepsilon_1'' \qquad (17)$$

With $\varepsilon_1' < 0$

And $\varepsilon_1'' << |\varepsilon_1'|$

The damping introduced by the e'' term means that the mode has a complex propagation vector - and is thus attenuated. "Bulk" electromagnetic radiation which has a real propagation vector may therefore only partially couple to the surface mode.

2.2 Free Electron metals

Free electron metals are able to support surface- polaritons, naturally of the surface-plasmon-polariton type in the visible and infrared regions of the spectrum. The dielectric function over most of this range may be approximated to that of a free electron gas (plasma) and is given by.

$$\varepsilon(\omega) = 1 - \left(\frac{\omega_P}{\omega}\right)^2 \qquad (18)$$

Where ω_p is the plasma frequency. Thus $\varepsilon(\omega)$ is negative for $\omega < \omega_p$. The dispersion relation for a plasma-dielectric interface is shown in fig 1 the wave vector of the surface mode is always greater than that of bulk electromagnetic radiation in the dielectric indicated by the 'light-time'.

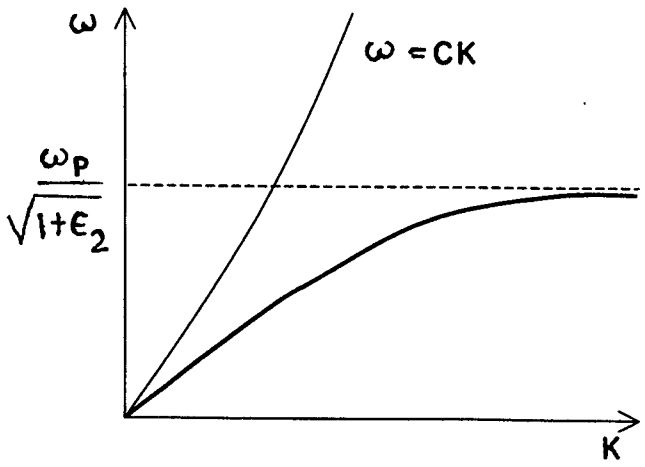


Fig 1 – Dispersion curve for a Plasma dielectric interface ω_p is the Plasma frequency and θ_2 the dielectric constant of the inactive medium. The light line is given by $\omega = cK$

2.3 Optical coupling methods

We found in section 2. from the derivation of the dispersion relation, that surface-polaritons are non-radiative. For a given frequency there is a momentum mismatch between bulk electromagnetic modes in the inactive medium, and the surface modes at the interface. Consider the situation shown in fig 2 where light is incident in the inactive medium, at an angle θ on the interface of interest. Its component of momentum in the x-direction that might couple to a surface mode is given by.

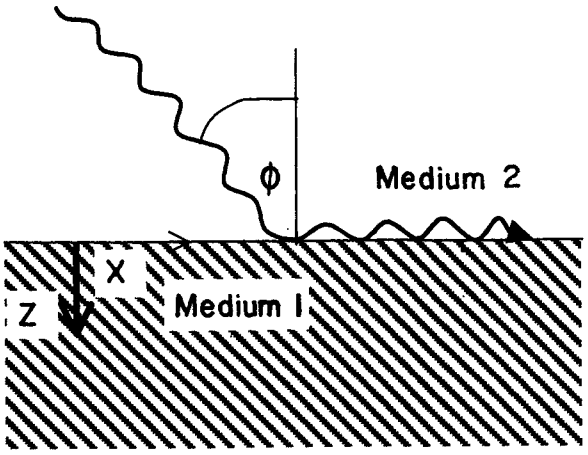


Fig 2 – Geometry of incident light and surface mode. Medium 1 is the active medium

$$K_{2\varepsilon} = \frac{\omega}{c} \sqrt{\varepsilon_2} \sin\phi \qquad (19)$$

Even for $\phi = 90^0$ this is still less than the momentum of the surface mode. i.e.

$$K = \frac{\omega}{c} \sqrt{\frac{\varepsilon_1 \varepsilon_2}{\varepsilon_1 + \varepsilon_2}} \text{ where } \frac{\varepsilon_1}{\varepsilon_1 + \varepsilon_2} > 1 \qquad (20)$$

A more subtle approach is required whereby the incident radiation can gain some momentum. We now consider in detail one of our method for achieving coupling i.e. Prism coupling.

3. OPTICAL EXCITATION OF SURFACE PLASMON POLARITIONS & PRISM COUPLING

Optical excitation of SPP's requires matching of both energy (i.e. optical frequency or wavelength) and momentum (wavevector). The x-component of the SPP wavevector at a given wavelength (i.e. free space wavevector k_0) is:

$$k_x^2 = k_0^2 \sqrt{\frac{\varepsilon_1 \varepsilon_2}{\varepsilon_1 + \varepsilon_2}} \qquad (21) .$$

Whilst light in medium 1 can provide x-components between zero and $k_0\varepsilon_i$ depending on angle of propagation.

For a metallic medium 2 it is easily proved that $K_x^2 > K_0^2 \varepsilon_1$, so the momentum is too great for excitation from medium 1.

It is also immediately clear that excitation from medium 2 cannot even provide the correct sign of K_x^2. The conclusion is that the modes of the single interface system cannot be directly excited optically, this is unsurprising as our criterion for a bound surface mode was its lack of optical radiation or excitation. There is an obvious analogy with optical waveguides where the light is confined in a film or channel to prevent radiation loss. The method of excitation for SPP's is analogous to the prism coupling technique for optical wage guides and is termed the attenuated total reflection or ATR method.

This is the most commonly used arrangement for coupling bulk and surface modes. Consider the situation shown in fig 3 Light is internally incident on the base of a prism as shown. If the angle of incidence. ϕ is greater that the critical angle ϕ_c. Then total internal reflection occurs. In medium 1 no net propagation take place in the z-direction. But an evanescent wave is set up in he region of the 0:1 interface. Of importance to us is the x-component of the wavevector of this evanescent wave: it is the same as that of the light in the prism, ie k $\sqrt{\in_0}$sin ϕ where k is the free-space wavevector. We have thus achieved the desired increase in momentum. If we now bring up a metal (medium 2) close to the 0:1 interface, the evanescent field extending from the prism may couple to the surface mode of the 1:2 interface. Momentum matching is achieved, by varying the angle of incidence, ϕ. We view this method as the tunneling of photons across a low index gap

(usually air). This method of coupling photons to SPP modes was fiorst suggested by Otto (1968)[1]. Achieving efficient coupling is a compromise. Reducing the gap increases the coupling of the fields, whilst increasing the gap avoids perturbation of the surface modes by the presence of the prism. A compromise separation of prism and metal surface is of order the wavelength involved.

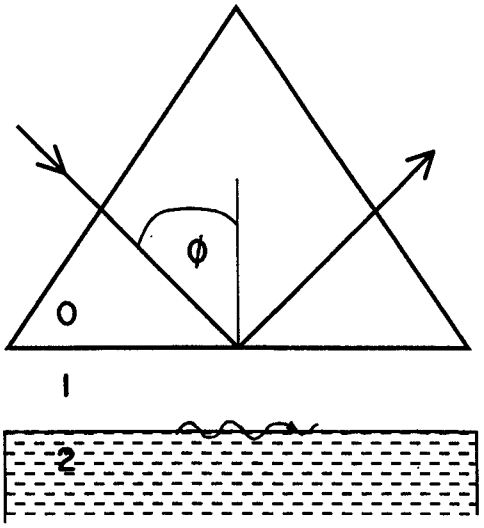


Fig 3 – Otto Configuration

A modification of this technique, devised by Kretchmann and Raether (1968)[2], is to change the roles of media 1 and 2, ie medium 1 becomes the active medium (the dispersion relation is symmetric with respect to ε_1 and ε_2), This is shown in fig 4. The active medium is now the tunneling barrier and must be thin enough to allow some of the incident energy to reach the 1:2 interface, this depending on the loss in the active medium. The Kretschman configuration is particularly useful as it allows for the direct formation of the metal film on the base of the prism by vacuum deposition. The Otto configuration on the other hand requires the setting up and maintenance of an air gap of order 1μm. Here we have used the Kretschman configuration.

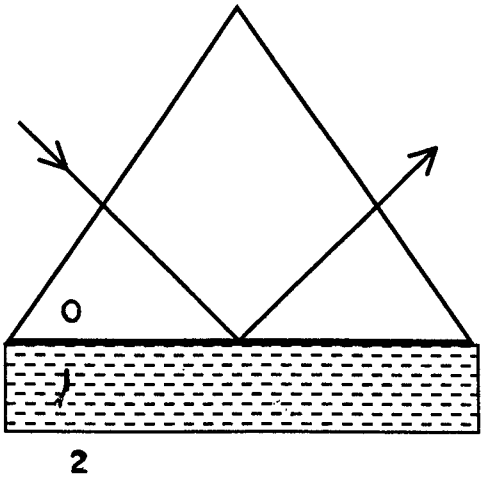


Fig –4 – Kretchmann Configuration

4. Principles of Various Optical Sensors:

The principle for gas detection involves placing a thin layer of chemically active film on the metal surface along, which SPs propagate. Changes in this layer brought about by the presence of gas will result in changes in SPR. The physical parameters that changes due to gas absorption or adsorption are refractive index, mass and thickness of the film. Nylander et

al[3] could able to detect halothane gas of different concentration by SPR method. They obtained shift in reflectivity upon gas exposure to different concentration of halothane gas and obtained good linearity between gas concentration and shift in resonance angle ($\Delta\theta_{sp}$).

Sufficient literatures are available for the design of SPR optical sensor for Nitrogen dioxide (NO_2) gas D. G. Zhu et al[4] have studied the effect of NO_2 on Langmuir Blodgett (LB) copper phthalocyanine (CuPc) film deposited on silver/nickel layer and consequent changes in SPR curve. They have shown the changes of SPR curves for cup_c LB layers when exposed to several hundred-vpm concentration of NO_x .

Vukusic at al[5] utilized S P R technique by using cobalt phthalocyanine as the basis for an optical sensors for NO_2. SPR profile changes upon exposure toNO_2 adsorption tended to deepen and narrow the SPR profile.

John. D Wright at el[6] have shown the changes of SPR measurement due to exposure of NO_2 on spun film phtholocyanine over gold layer. They have claimed that the SPR detection methods have a major advantages over semi conducting methods, because in the former method the detection for low concentration of NO_2 is possible without any signal noise problem.

In a chemical sensors design using SPR, the presence of chemical species produces a change in the refractive index at the metal/dielectric interface, which inturn results in a change in reflected light intensity or a change in the coupling angle (θ_{SP}) of SP excitation. By monitoring the resonance condition of the SP, the dielectric constant of the sample is obtained. If the sample gas or solution is contaminated by some chemical species, dielectric constant measurement gives the concentration of the chemical species in the sample.
Koji Matsubara et al[7] first obtained reflectance curve of air with SPR sensor system. Then they obtained reflectance curves if pure water and 5% ethanol in water. They found the difference in resonance angle, which is due to the difference in dielectric constant between two samples. They have also calibrated the angular shift with the concentration of ethanol.

Chadwick et al[8] designed temperature sensor based in SPR measurement. They obtained SPR curve in silver/air interface. Further SPR profile with a coating of hydrogenated amorphous silicon (S_i:H) over silver layer results in broadening and shifting the resonance angle. For temperature sensor, they set angle of prism assembly at 61.5^0 and monitored the reflected light intensity as a function of temperature. Reflected light intensity as the function of temperature shows a near linear relation in the range 20^0C to 60^0C. A remarkable factor of three changes in reflected light intensity is seen over this temperature range indicating that the sensor is sensitive to small changes in temperature.

5. THEORY

In this paper, we report the modification of SPR condition brought about due to deposition of calix (4) resocinarences thin film by spin coating[9] on gold layer and also the effect of interaction with change of ambient. This being the basic principle for designing optical sensors.
The most commonly used technique for the excitation of SPR is the attenuated total-reflection (ATR) technique[10] first demonstrated by Rather[11] illustrated in fig.5.

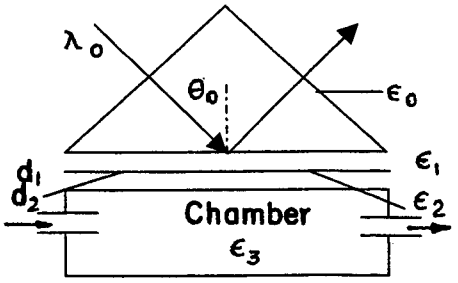


Fig –5 – Systematic diagram showing the Kretchmann ATR configuration for four-layer system.

In this configuration a metal film (thickness d_1, complex relative permitivity $\varepsilon_1 = \varepsilon'_1+i\varepsilon''_1$) which is coated with a thin layer on one surface (d_2, $\varepsilon_2 = \varepsilon'_2+i\varepsilon''_2$) is embedded in two lossless semilinfinite media characterized by dialectic function ε_0 and ε_3 respectively ($\varepsilon_0 > \varepsilon_1$). Application of Maxwell's equations and suitable continuity condition yields the reflection coefficient (r_{0123}) for the ATR system;

$$r_{0123} = r_{01}\,(Z_{0123}/D_{0123}), \qquad (22)$$

where D_{0123} is the dispersion relation of the eigenmodes of the four-layer system. The dispersion relation of the surface plasmon depends on both the dielectric constant of the metal and that of the sample medium faced on the metal. That relation is given by

$$D_{0123} = (k_{3x}/\varepsilon_3 + k_{1x}/\varepsilon_1) - i\,\tan(k_{2z}d_2)\,[\,(k_{3z}k_{1x}\varepsilon_2/k_{2z}\,\varepsilon_1\,\varepsilon_3) + (k_{2z}/\varepsilon_2)] + r_{01}\exp(2ik_{1z}d_1)[(k_{1z}/\varepsilon_1 - k_{3z}/\varepsilon_3) - i\tan(k_{2z}d_2)\,(k_{3z}k_{1z}\varepsilon_2/k_{2z}\varepsilon_1\varepsilon_3) - k_{2z}/\varepsilon_2)] \qquad (23)$$

and which defines also Z_{0123} if r_{01} is replaced by $r_{01}^{-1}.r_{01}$ $\{= [(k_{1z}/\varepsilon_1) - (k_{0z}/\varepsilon_0)]/[(k_{1z}\varepsilon_1) + (k_{0z}/\varepsilon_0)]\}$ is the Fresnel reflection coefficient for the interface between media 0 and 1, $k_{12} = [\varepsilon_1(\omega/c)^2 - k_x^2]^{1/2}$ and $k_x = \varepsilon_0^{1/2}(\omega/c)\sin\theta_0$. The absolute reflection R ($=|r_{0123}|^2$) takes on a function dependence on the angle of incidence θ_0, and it is clear that R will reach a minimum value at certain value of θ_0. This minimum in reflectance directly corresponds to the excitation of SPR. The shape of this reflectance minimum and the angle at which it occurs will clearly depend on the dielectric constants ε_0, ε_1, ε_2 and ε_3, the thickness of the metal film d_1 and the wavelength of electromagnetic radiation λ_0. Thus, by successfully characterizing the SPR the dielectric overlayer may be quantified.

6. EXPERIMENTAL PROCEDURE

Metallic gold thin film (45 mm thick) was deposited on glass slides by vaccume evaporation method. These slides were brought in optical contact with a prism by means of index matching fluid. SPR study on each gold-coated slides was taken. The experimental set up for SPR study is shown in fig. 6.

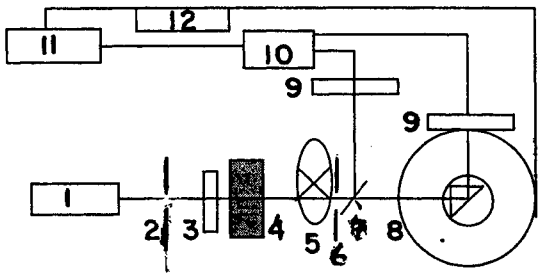


Fig-6 – Systematic representation of SPR technique [1-He-Ne Laser, 2-Aperture, 3-Neutral density filter, 4-Polariser, 5-Beam chopper, 6-Aperature, 7-Beam Splitter, 8-(θ-2θ) prism table, 9-photo detector, 10-Lock-in-amplifier, 11-microprocessor, 12- (θ-2θ drive)]

A monochromatic, highly collimated, plane polarised He-Ne (=632.8 nm) laser tube is used as the light source to illuminate the prism.

This beam is passed aperture followed by a polariser which is used in order to ensure that only p-polarised light will be incident on the prism. This beam is then modulated using a mechanical chopper at a frequency 1.6 kHz. Such a high frequency will help to eliminate noise from sources with lower frequency, such as light source inside the laboratory. The modulated beam is then passed through another aperture, which is used mainly for alignment purposes. The optical beam is then made incident on a beam splitter where a small fraction of its intensity is removed and sent to a reference photo detector. This is used for reference the main beam incident on the sample in order to eliminate any drift in the intensity of the laser source. The two beams (the reference and the reflected) now are detected and passed to a lock - in amplifier, which will remove any dc background levels from the modulating signal, resulting in a highly reduced noise in the data. To calibrate the angle of table, the normal incident must be known. This can be obtained by rotating the prism table until the first face reflection returns back down the optical path. This position is the Zero position. The detector is also calibrated from this zero angle, by positioning the detector in the second face (inside back of prism) reflection path. The angle range to be covered should start before and after the Total Internal Reflection (TIR) angle.

The prism was placed on the rotating table in such way so as to keep the point of incident on the prism's base stationary when the angle of incidence was varied. The photo detector position was adjusted so as to catch the reflected beam as closely as possible over the angle range to be used. First set of measurement was done to find the angle of total internal reflection (TIR) of the prism and also to check the refractive index of the prism used. The prism and glass slide are

clamped together by means of index matching oil. For a certain value of external angle of incidence (θ_{ext}) on any face of prism or certain value of internal angle of incidence (θ_{int}) on the base of the prism, there is some reflected intensity and transmitted intensity. Under this condition the angle of incidence is less than the angle of refraction since light comes from denser medium prism to rarer medium (air). As we go on increasing the angle of incidence, the angle of refraction also increases. AT a particular value of angle of incidence say critical angle, the value of angle of refraction is 90^0. On further increasing the angle of incidence, the light will be totally reflected and the reflected intensity is maximum and there will be no transmitted intensity. This is called total internal reflection and angle corresponding to the onset of TIR is called critical angle or angle of total internal reflection. From this angle the measured reflected intensity will be constant with respect to the increase in incidence angle.

The prism/sample assembly (prism and gold-coated glass slide) was mounted on a motorized rotation table allowing angle of incidence to be varied. The reflected light is detected by a photo diode, which was also mounted on a rotation stage. The rotation of the two stages is effected by two stepper motors driven independently by a personal computer, such that when the prism is turned by an angle θ the photo diode stage will be rotated by 2θ. This goniometre effect was used to ensure that the reflected beam always stricks the photo diode.

The whole set up interfaced to a computer. The photo diode voltage was detected by a lock-in amplifier. The data set was obtained by a computer via an IEEE488 interface, permitting software calculation of reflectivity as a function of the angle of incidence. The data set was plotted using the easy plot graphic software. Curve fitting was preformed using software based, least squarefit of the experimental data to the Fresnel reflection equation.

Reflectivity from the base of the prism was measured as a function of the internal angle of incidence. Using easy plot graphic software the reflectivity versus internal angle of incidence were plotted. At a particular value of angle of incidence, here will be a dip (Zero of minimum reflectivity) in the curve. This corresponds to optimum coupling. Few gold-coated glass slides where chosen for depositing Calix (4) resorcinarence by the process of spin coated at different spin speed ranging from 1000 to 6000 rpm. The experiment setup for photo resist spinner model-4000 includes-bench top unit and remote control unit. The gold film slide was vacuum held onto a rotating chuck whose speed of rotation can be selected. One drop of solution was allowed to fall with the help of a micro sryinge. As the slide spins the drops spreads out under the influence of a centrifugal force followed by solvent evaporation. Finally a smooth continuous film of Calix (4) resorcinarene is left over gold coated slide. SPR measurement of each spun film was made as detailed above. SPR curve of each spun film was compared with corresponding bare gold film. The curve shows SPR angular shift due to deposition of Calix (4) resorcinarene. Further investigation was made to study the change of ambient close to the film. As usual SPR measurement was made with air ambient. Then air is replaced by water and SPR measurement was made. The layer system changes from glass/gold/calix/air to glass/gold/calix/water. The SPR curve is greatly changed.

Three primary changes are seen they are (1) resonance half width increases (2) SPR angle shifted to higher angle. (3) The depth of the minimum also changes fig 7.

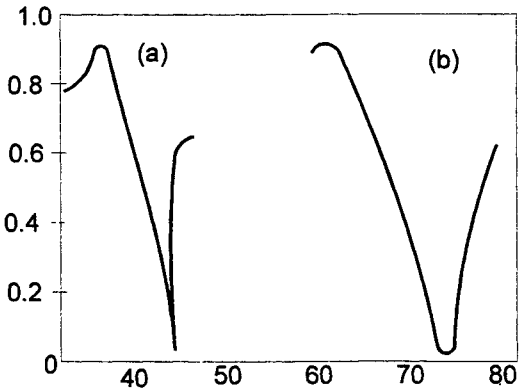


Fig7-(a) SPR for glass/gold/calix/air interface (b) SPR for glass/gold/calix/water interface.

7. RESULT AND DISCUSSION

As appeared from the plot, the SPR curve is greatly changed due to change in ambient of layer system i.e. glass/gold/calix/air to glass/gold/calix/water.

The angular reflectance spectrum shows three primary changes. The resonance curve half width increase implying the increased damping. The position of θ is changed, showing the change in effective dielectric constant of the medium. The depth of the minimum is changed which implies the change in absorption in the film. The changes which are brought about due to change in ambient can be exploited to design chemical sensors. Sufficient literatures are also available which predicts different sensing properties of organic film using SPR technique.

8. CONCLUSIONS

The wave vector sensitivity of SPs causes three primary changes due to any modification occurring at the interface. They are (a) Resonance half width change (b) Angular shift of resonance angle & (c) Depth of minimum change. These changes are properly monitored in all the optical sensor designs. Optical sensors based on SPR are promising because of their non-destructive nature and ability to sense low concentration of poisonous gases in air, chemical pollutants in a sample and organic pollutants in water. The use of photodiode array eliminates mechanically moving parts, achieving a simple, compact apparatus for in-field sensing. The SPR technique not only helps in designing hand held prototype sensors for gas chemical and temperature sensing but also finds applications in surface physics, thin film characterisation and determination of optical constants of materials. Looking from various angles, much more work on SPR both from experimental and theoretical sides seems to be very promising for further development of SPR technique.

ACKNOWLEDGEMENT

The first author (Dr. S. Panigrahi) expresses his heart felt thanks to the coordinator Indo-U.K. REC. project, Principal, HOD Physics of REC., Rourkela for deputing him to U.K. under Indo-U.K. REC. Project.

REFERENCES

1. Otto. A, Z. Phys., **216**, 398 (1968).
2. E. Kretchmann and H. Reather, Z. Naturf A, **23**, 2135 (1968).
3. C. Nylandar, B. Liedberg and T. Lind, Sensors and Actuators, 3, **79** (1982/83).
4. D. G. Zhu & M. C. Petty, Sensor and Actuators B, **2**, 265 (1990)
5. P. S. Vukusic & J. R. Sambles, Thin Solid Films, **221**, 311 (1992)
6. John. D. Wright et al Sensor and Acutators B **29**, 108 (1995)
7. K. Matsubara, S. Kawata, S. Minami, "Optical chemical sensor based on surface plasmon measurement, App.Opt. **27(6)** (1988) 1160 –1163.
8. B. Chadwick and M. Gal Jpn. J Appl. Phys, **32**, 2716 (1993).
9. Peter C. Sukanek, "Spin Coating", Journal of Imaging Technology **11(4)** (1985) 184 –190.
10. P. S. Vukusic, J. R. Sambles, J. D. Wright, "Surface Plasmon Resonance Characterization of spin deposited Phthalocyanine films", J. Mater Chem, **2(10)** (1992) 1105 –1106.
11. H. Reather, "Surface Plasmon Springer Verlag Tracts in modern Physics II, Berlin, (1988).

Separative structure ISFETs on a glass substrate

Li-Te Yin, Jung-Chuan Chou[a*], Wen-Yaw Chung[b], Tai-Ping Sun[c] and, Shen-Kan Hsiung[b]

Department of Bio-Medical Engineering, Chung Yuan Christian University,

Chung-Li, Taiwan 320, R.O.C.

[a*]Institute of Electronic and Information Engineering, National Yunlin University of Science and Technology, Toulin, Taiwan 640, R.O.C.

[b]Department of Electronic Engineering, Chung Yuan Christian University,

Chung-Li, Taiwan 320, R.O.C.

[c]Department of Electrical Engineering, National Chi Nan University, Nantou, Taiwan 545, R.O.C.

ABSTRACT

In our research, the glass was used as a substrate for H^+ ion sensitive field effect transistor (ISFET). The sensitive characteristics of five structures for separative extend gate ion sensitive field effect transistors (EGFET) were studied, which include tin oxide (SnO_2) / aluminum / micro slide glass, tin oxide / aluminum / corning glass, indium tin oxide (ITO) glass, tin oxide / indium tin oxide glass and tin oxide / micro slide glass. Indium tin oxide (ITO) thin film is the first time used as a H^+ ion sensitive film which has a linear pH sensitivity of Nerstern response, about 58 mV/pH, between pH2 and pH12. In addition, the sensing area effect of the tin oxide / glass, tin oxide / ITO glass and ITO glass structure which without Al conductive layer will be discussed.

Keywords: ISFET, EGFET, tin oxide, indium tin oxide (ITO) glass, pH response

1. INTRODUCTION

Ion sensitive field-effect transistors (ISFETs) have been developed on the basis of the MOSFET (metal oxide field effect transistor). Since Bergveld[1] first employed the field-effect transistor in neurophysiological measurement in 1970, ISFETs have developed into a new type of chemical-sensing electrode. Many theoretical and experimental studies have been

*Correspondence: Dr. Jung Chuan Chou, Address: 123 Sec. 3, University Rd. Toulin, Taiwan 640, R.O.C., Institute of Electronic and Information Engineering, National Yunlin University of Science and Technology , Tel: 886-5-534-2601 Ext.2500, Fax: 886-5-531-2029, E-mail: choujc@pine.yuntech.edu.tw

Part of the SPIE Conference on Advanced Photonic Sensors and Applications
Singapore • November–December 1999
SPIE Vol. 3897 • 0277-786X/99/$10.00

published for describing the behavior of this chemical-sensing electronic device.[2] Silicon dioxide (SiO_2) was first used as a pH-sensitive dielectric for the ISFET[1]. Subsequently, Al_2O_3, Si_3N_4, Ta_2O_5, and SnO_2 have been used as pH-sensitivity dielectric because of the higher pH response.[3-8] However, an ISFET is a kind of transistor, which works in saline water. Thus, the problem arises from the poor isolation between the device and solution. It is thus very important to develop an ISFET encapsulation process that is compatible with integrated circuit technology. Until now, several fabrication methods of ISFET-based biosensors have been reported. Esashi and Matsuo applied the anisotropic etching technique to make a needle-like ISFET device, which is completely isolated from the water.[9] However, the anisotropic etching is too complex to use in process. Another most frequently used method is based on a silicon-on-sapphire (SOS) structure.[10-13] However, SOS ISFETs have disadvantages, such as an instability of characteristics because of penetration of impurities (for example, Al) from the sapphire substrate, low current sensitivity and high cost of a sapphire wafer. In addition, Poghossian solved the ISFET encapsulation problem by Si-SiO_2-Si (SIS) structure.[14-15] But, the silicon bonding process of SIS structure ISFETs is still too complex. An extended gate field effect transistor (EGFET) is another structure to get isolation of FET from the chemical environment, in which a chemically sensitive membrane is deposited on the end of signal line extended from the FET gate electrode.[16-17] This structure has a lot of advantages, such as light insensitivity, simple to passivate and package, the flexibility of shape of the extended gate area, etc. Recently, Chi et al. introduced an improved structure of EGFET which is separated into two parts. One is sensing part with the structure of SnO_2/Al/Si and another is commercial MOSFET, CD4007UB.[18] This structure is suitable for application for a disposable biosensor, because the separate MOSFET is reusable when the sensing part has to interchange. However, the charge could be leakaged from the silicon layer to the conductor layer, Al, in the sensing part, so the critical encapsulation still need for this structure. In addition the SnO_2 sensing film, we have applied Si_3N_4, Al_2O_3 and SiO_2 to the separate structure, respectively, but all this kind of insulating materials can't show a pH response. By the way, the insulating material may not a proper sensing material for separative EGFET. In this paper, we improved the separative EGFET structure which the substrate of the sensing part is glass. Indium tin oxide (ITO) is popular for their unique electrical and optical properties, that is, high electrical conductivity and high optical transmittance in the visible region.[19] These qualities have caused them to be widely used as transparent conducting elements for addressing various kinds of alphanumeric displays.[20] In this study, the indium tin oxide glass electrode was first used as a pH-sensitivity material for separate EGFET.

2. EXPERIMENTAL

2.1. Materials

The ITO glasses (50 - 100 Ω/□, ITO coating thickness : 230 Å) were granted from Wintek corporation. The micro slide glass and corning 7059 glass were purchased from Kimax Glass Instrument and Tekstarter corporation..

2.2. Fabrication of sensing layer structure

The sensing part of separative EGFET is divided into five kind of structures which are SnO_2/Al/micro slide glass, SnO_2/Al/corning 7059 glass, ITO glass, SnO_2/ITO glass and SnO_2/glass. The Al and SnO_2 were deposited by thermal evaporation and sputtering method, 2000 and 3000 Å, respectively. Before the glasses were deposited the Al or SnO_2, they were washed by methyl alcohol and D.I. water for 20 and 10 minutes, respectively. The complete encapsulation section diagram of SnO_2/Al/micro slide glass sensing structure was shown as Fig.1.

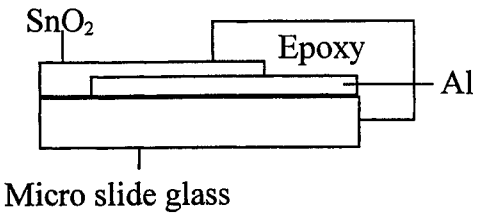


Fig.1. Cross-section of glass substrate sening structure

2.3. Measurement processes

The HP4145B Semiconductor Parameter Analyzer was used to measure the threshold voltage (V_T) of the EGFET in the pH buffer solutions. The measurement system in this study is shown as Fig. 2. The sensing structure and calomel reference electrode were dipped into buffer solution and connected with the gate of commercial MOSFET device, CD4007UB. The V_{DS} is fixed at 0.2V, where V_{DS} is the voltage between drain and source of MOSFET. The distance between the reference electrode and sensing electrode is about 0.5 ~ 2 cm which will not affect the results during the measurement.

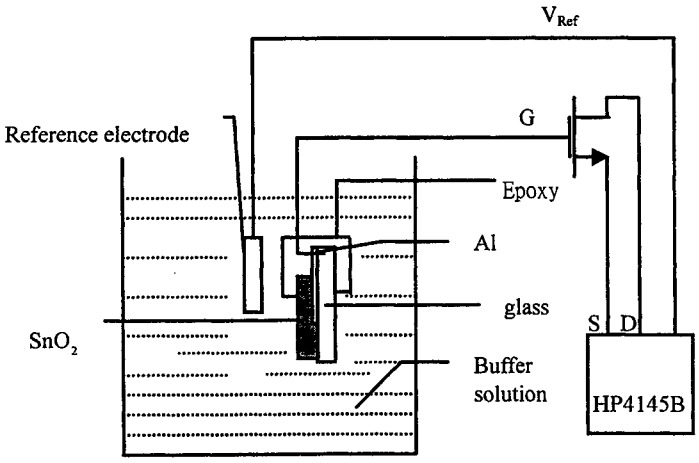


Fig 2. EGFET measurement system

A readout circuit based on an instrumentation amplifier AD620, which is shown in Fig. 3, was used to study the drift and hysteresis characteristics of the separative EGFET.

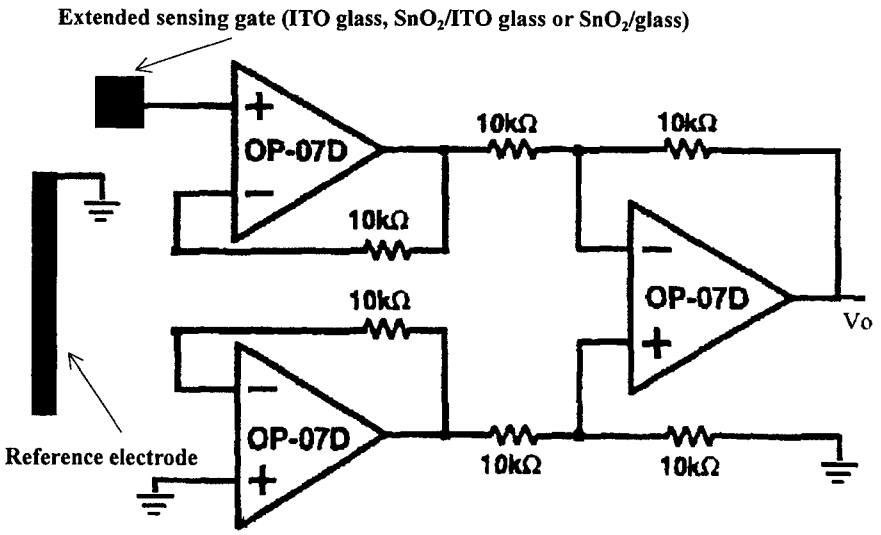


Fig. 3. Measurement circuit of instrumentation amplifier AD620

3. RESULTS AND DISCUSSION

3.1. The sensing structures with Al

In the structures with Al, the conductive layer Al became an unstable element in pH solutions. As the sensing gate put into the buffer solution, SnO_2 and Al layer peel very soon. In our experiment, we have shown that baking process can improve this peel effect. The table.1 shows that the extended sensing gates with micro slide glass and corning glass substrates have to bake for 3 and 18 hours at 150°C, respectively, to avoid that Al and SnO_2 peel into the buffer solution. The current-voltage (I-V) curves of micro slide glass and corning glass with baking process are shown in Fig. 4 and Fig. 5. In the Table.1 "Failed" means that the pH sensitivity can't be extracted as the situations of Fig.4(a) and Fig.5(a). However, the pH sensitivity of SnO_2 film ISFET decreases with the baking time.[7] So the micro slide glass substrate is more suitable for applying in SnO_2 sensitive separative EGFET.

Table. 1. Baking (150°C) effect of structures that are SnO_2/Al/micro slide glass and SnO_2/Al/corning glass.

Baking time / Structures	Without baking	3 hours	10 hours	18 hours
SnO_2/Al/micro slide glass	Failed	54 mV/pH	50 mV/pH	49mV/pH
SnO_2/Al/corning glass	Failed	Failed	Failed	46mV/pH

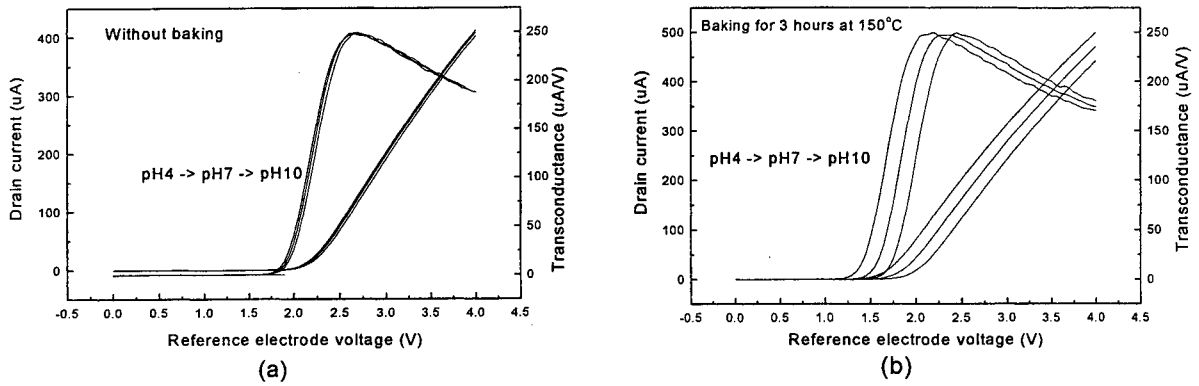


Fig. 4. I-V curves of SnO_2/Al/micro slide glass separative gate EGFET (a) without baking process (b) baking at 150°C for 3 hours

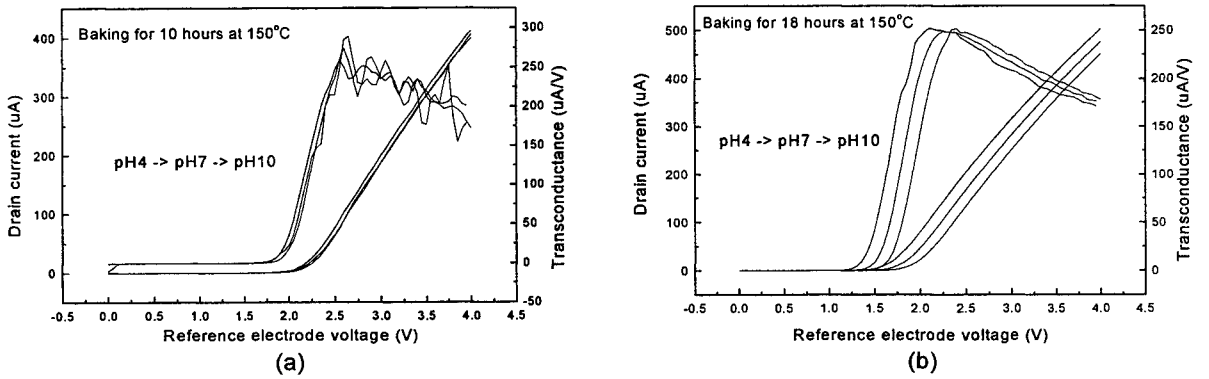


Fig. 5. I-V curves of SnO_2/Al/corning glass separative gate EGFET (a) baking at 150°C for 10 hours (b) baking at 150°C for 18 hours

3.2. The characteristic of ITO sensing gate

The indium tin oxide (ITO) thin film is a well-known material used as an electric conductor of high optical transmittance but a novel material used as a pH-ISFET sensing film. This section will show the fundamental characteristic of ITO glass electrode that applied as a separative EGFET sensing gate.

3.2.1. pH sensitivity of ITO gate ISFET

The transconductance gives the same peak value in a concentration range between pH2 and pH12 shown in Fig. 6 under the same temperature. The slop of the same I_D versus V_{ref} can be obtained around the maximum transconductance. This slop has a concept with V_{ref} that is the threshold voltage of ISFET. The pH sensitivity of the ISFET was investigated through a shift in the threshold voltage of an ISFET sensor. The result shows that the ITO gate separative EGFET sensor has a linear pH sensitivity of approximately 58 mV/pH in a concentration range between pH2 and pH12.

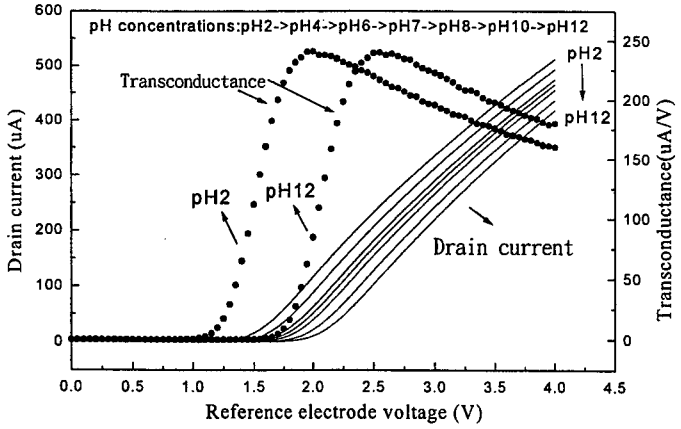


Fig.6 I-V characteristics of ITO sensing film EGFET which has a contact window of 10*10mm^2

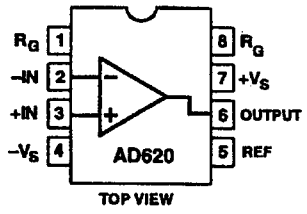


Fig. 7 Connection diagram of AD620

In this study, the instrumentation amplifier AD620 was used as a readout circuit which is shown as Fig.3. The connection diagram of AD620 is shown as Fig. 7. ITO was connected to one of the input terminal that is terminal 2 or 3 and the other was connected to ground which the output is shown as Fig. 7. The Fig. 8(a) and Fig 8(b) are the results of the ITO which were connected to terminal 2 and terminal 3, respectively, and both the results show the pH-sensitivity that is about 58 mV/pH. It shows that the AD620 can be used as a readout circuit for separative EGFET.

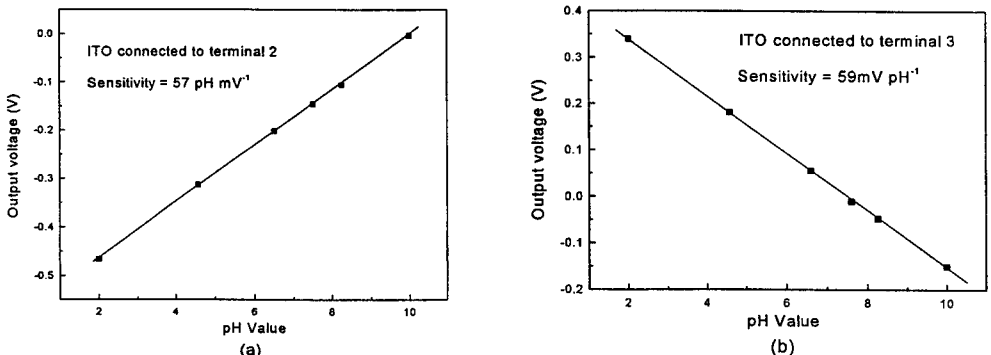


Fig.8 Output voltage versus pH value of the ITO sensing gate connected with instrumentation amplifier AD620
(a) ITO connected to terminal 2 (b) ITO connected to terminal 3

3.3.2 Drift and hysteresis of ITO sensing gate EGFET

In the study of drift and hysteresis of the ITO sensing gate EGFET sensor, the ITO was connected to terminal 2 of instrumentation amplifier AD620. The drift experiment described in Fig. 9 was tested for 18 hours. The result indicated a drift of 110 mV in this time. It shows a serious drift of ITO sensing gate EGFET sensor especially almost 1mV per minute in the first hour.

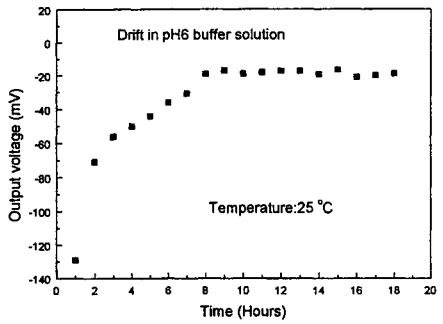


Fig. 9 Drift characteristics of ITO sensing gate EGFET

To evaluate the hysteresis of the EGFET, we measured the output offset voltage after solution change such as pH 7 → pH 4 → pH 7 and pH 7 → pH 10 → pH 7. The result shows that the hysteresis of the ITO sensing gate EGFET is about 9.8mV as shown in Fig 10.

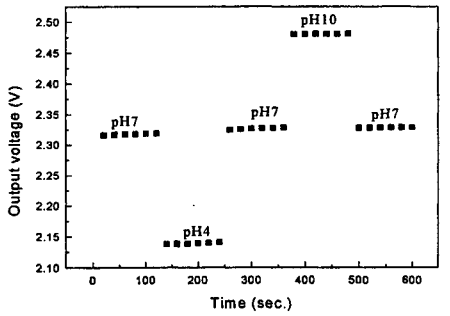


Fig 10 Hysteresis characteristic of ITO sensing gate EGFET

3.3. The structure of SnO_2/glass and SnO_2/ITO glass

In this study, we find out that the structure of SnO_2/glass sensing gate EGFET had a pH sensitivity about 60 mV/pH as the contact window between the SnO_2 layer and buffer solution is up to 2× 2 mm^2. However the pH sensitivity is decreased as the contact window is smaller than 2× 2mm^2. The structure of SnO_2/ITO glass sensing gate EGFET had a pH sensitivity about 57mV/pH, which independent of the contact window between the SnO_2 and pH buffer solution. That is caused by the higher resistance of SnO_2/glass structure with the smaller contact window, which without a Al or ITO conductive layer, sharing the voltage signal passed from the interface of SnO_2 sensing layer and pH buffer solution which the surface potential is changed with different pH values to MOSFET metal gate. The voltage sharing effect sets off a reducing of MOSFET output signal. So the pH sensitivity of small contact window of SnO_2/glass gate EGFET becomes lower. Fig.11 shows the drift of ITO glass, SnO_2/ITO glass and SnO_2/glass sensing gate structure EGFET. The results indicate the drifts of 33.9 mV and 9.1 mV in 18 hours for SnO_2/glass and SnO_2/ITO glass sensing gate EGFET, respectively. The sensing structures with SnO_2 (SnO_2/glass, SnO_2/ITO glass) have an obviously smaller drift than sensing structure with ITO.

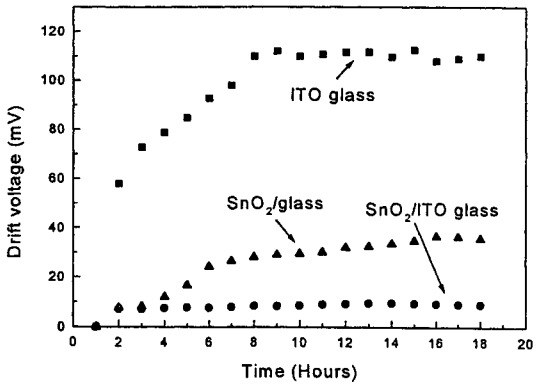


Fig. 11 Drift characteristics of ITO glass, SnO_2/ITO glass and SnO_2/glass sensing gate EGFET

4. CONCULSIONS

In summary, from the above results and discussion, the important conclusions can be obtained as follows:

1. The glass substrate sensing gate EGFETs have advantages of easier fabrication processes than traditional ISFET and lower cost than SOS structure ISFETs or silicon based EGFET.
2. The structures of SnO_2/Al/micro slide glass and SnO_2/Al/corning glass have a problem of peeling the Al and SnO_2 into the buffer. This problem can be solved by baking the SnO_2/Al/micro slide glass and SnO_2/Al/corning glass at 150°C for 3 hours and 18 hours, respectively.
3. The indium tin oxide can be used as a pH-sensitive film of EGFET, and it has a linear sensitivity about 58 mV/pH in a pH concentration in ranging from pH 2 to pH 12.
4. The instrumentation amplifier AD620 can be used as a readout circuit for separative gate EGFET.
5. The ITO gate EGFET has a drift characteristic of 110 mV for 18 hours and a hysteresis characteristic of 9.8 mV after solution change such as pH 7 → pH 4 → pH 7 and pH 7 → pH 10 → pH 7.
6. The pH sensitivity of SnO_2/glass sensing gate EGFET is decreased as the contact window smaller than $2\times\ 2mm^2$.
7. The drift characteristics of SnO_2/glass and SnO_2/ITO glass sensing gate EGFET are respectively 33.9 mV and 9.1 mV for 18 hours. The two structures that with SnO_2 sensing film have an obvious smaller drift than ITO one.

ACKNOWLEDGMENTS

This work was supported by National Science Council, the Republic of China under the contracts NSC88-2215-E033-006.

REFERENCES

1. P.Bergveld, "Development of an ion sensitive solid-state device for neurophysiological measurements, *IEEE Trans. Biomed. Eng.*, **17**, pp.70-71, 1970.
2. Clifford D. Fung, Peter W. Cheung, Wen H. Ko, "A generalized theory of an electrolyte-insulator-semiconductor field effect transistor", *IEEE Trans. Electron Devices*, **33**, pp.8-18, 1986.
3. Stanley D. Moss, Curtis C. Johnson, Jirijanata, "Hydrogen, calcium and potassium ion sensitive FET transistors a preliminary report", *IEEE Trans. Biomed. Eng.*, **25**, pp.49-54, 1978.
4. Hung-Kwei Liao, Jung-Chuan Chou, Wen-Yaw Chung, Tai-Ping, Shen-Ken Hsiung, "Study on the interface trap density of the Si_3N_4/SiO_2 gate ISFET", *Proceedings of the Third East Asian Conference on Chemical Sensors*, pp.394-400, Seoul, South Korea , November 1997.
5. Li-Te Yin, Jung-Chuan Chou, Wen-Yaw Chung, Tai-Ping and Shen-Ken Hsiung, "Characteristics of Silicon Nitride after O_2 Plasma Treatment for pH ISFET Applications", *Proceeding of 1998 International Electron Devices and Materials Symposia*, B, C, pp.267-270, National Cheng Kung University, Tainan, Taiwan, R. O. C., December 1998.
6. P. Gimmel, B. Gompf, D. Schmeiosser, H. D. Weimhofer, W. Gopel, M. Klein, "Ta_2O_5 gates of pH sensitive device comparative sepectroscopic and electrical studies", *Sensors and Actuators B*,**17**, pp.195-202, 1989.
7. Hung-Kwei Liao, Jung-Chuan Chou, Wen-Yaw Chung, Tai-Ping Sun and Shen-Kan Hsiung, " Study of Amorphous Tin Oxide Thin Films for ISFET Applications ", *Sensors and Actuators B*, **50**, pp.104-109, 1998.
8. Hung-Kwei Liao, Jung-Chuan Chou, Wen-Yaw Chung, Tai-Ping Sun and Shen-Kan Hsiung, "Study of pH_{pzc} and Surface Potential of Tin Oxide Gate ISFET", *Materials Chemistry and Physics*, **2428**,pp.1-6, 1999.
9. M. Esashi and T. Matsuo, "Integrated micro multi ion sensor using field effect of semiconductor", *IEEE Trans. Biomed. Eng.*, **BME-25**, pp.182-192, 1978.
10. B.H. van der Schoot and P. Bergveld, "ISFET-based enzyme sensors", *Biosensors*, **3**, pp.161-186, 1987.
11. S. Nakamoto, N. Ito, T. Kuriyama and J. Kimura, "A lift-off method for patterning enzyme-immobilized membranes in multi-biosensors", *Sensors and Actuators*, **13**, pp.165-172, 1988.
12. J. Kimara, T. Munakami and T. Kuriyama, "An integrated multi-biosensor for simultaneous amperometric and potentiometric measurement", *Sensors and Actuators*, **15**, pp.435-443.,1989
13. Y. Hanazato, M. Nakako, S. Shinon and M. Maeda, "Integrated multi-biosensors based on an ion-sensitive field-effect transistor using photolithographic technique", *IEEE Trans. Electron Devices,* **36**, pp.1303-1310, 1989.
14. A. S. Poghossian, "Method of fabrication of ISFETs and CHEMFETs on an Si-SiO_2-Si structure", *Sensors and Actuators B*, **13-14**, pp.653-654, 1993.
15. A. S. Poghossian, "Method of fabrication of ISFET-based biosensors on an Si-SiO_2-Si structure", *Sensors and Actuators B*, **44**, pp.361-364, 1997.
16. J. Van der Spiegel, I. Lauks, P. Chan and D. Babic, "The extended gate chemical sensitive field effect transistor as multi-species microprobe", *Sensors and Actuators B*, **4**, pp.291-298, 1983.
17. T. Katsube, T. Araki, M. Hara, T. Yaji, Si Kobayashi and K. Suzuki, "A multi-species biosensor with extended-gate field effect transistors", *Proceeding of 6th Sensor Symposium*, pp.211-214, Tsukuba, Japan, 1986.
18. Li-Lun Chi, Jung-Chuan Chou, Wen-Yaw Chung, Tai-Ping Sun and Shen-Kan Hsiung, "New structure of ion sensitivite field effect transistor", *Proceeding of the Biomedical Engineering Society 1998 Annual Symposium*, pp.332-334, Taiwan, 1998.
19. M. A. Martinez, J. Herrero and M. T. Gutierrez, "Electrochemical stability of indium tin oxide thin films", *Electrochimica Acta*, **37**, pp.2565-2571, 1992.
20. K. L. Chopra, S. Major and D. K. Pandya, "Transparent conductors – a status review", *Thin Solid Thin Films*, **102**, pp.1-46, 1983.

Optically Differential Phase Detection Using Cross-Talk Mode of Series Fabry-Perot Sensors in PMDI

Yu-Lung Lo and Chun-Han Wang
Department of Mechanical Engineering
National Cheng Kung University, Tainan, Taiwan 70101, ROC
loyl@mail.ncku.edu.tw

ABSTRACT

The signals between cascaded Fabry-Perot sensors have serious cross-talk problems due to complex phase interactions between sensors. However, cross-talk mode in cascaded Fabry-Perot sensors can be employed for demodulating optically differential-phase in a path-matching differential interferometry (PMDI). This paper presents a novel design in two series Fabry-Perot sensors for a differential-phase measurement analyzed by an optical spectrum analysis. The spectrum transfer function in a light source and Fabry-Perot cavities are simulated by a Gaussian distribution spectrum and a low-finesse reflectivity, respectively. It shows that two series Fabry-Perot sensors with the differential-phase measurement can be designed accurately. Experimental data shows that the novel design in measuring a differential-phase between two series Fabry-Perot sensors is achieved. This design in two series Fabry-Perot sensors for a differential-phase measurement can be applied to unbalanced system detection or noise compensations.

Keywords: Differential-Phase, Cross-Talk, Fabry-Perot Sensors, PMDI

1. Introduction

The demodulation schemes for common-path interferometry akin to that used in Fabry-Perot sensors are relatively limited, and perhaps the simplest and the most robust demodulator is based on the concept of white light interferometry and PMDI [1]. The primary advantage of using PMDI is that it uses a broad band source which enhances the system robustness and obviates the use of long gage length required for most demodulation schemes and, in addition allows coherence based multiplexing [2]. Several Fabry-Perot Sensors in multiplexing have a serious problem in cross-talks. To solve this problem in wavelength demodulation, Humblet and Hamdy [3] analyzed crosstalk in the optimization of single- and double-cavity Fabry-Perot filters for wavelength-division multiple-access (WDMA) networks. Different from the design of Fabry-Perot filters in networks, Chang and Sirkis [4] present several multiplexing systems to decouple Fabry-Perot sensors in parallel for phase demodulation. Later, Singh and Sirkis [5] designed two Fabry-Perot sensors: an In-Line Fiber Etalon (ILFE) and an Intrinsic Fabry-Perot (IFP) fiber in series for a two-parameter sensor to measure the axial strain and temperature simultaneously.

In this paper, quantitative design in lengths of cavities for series Fabry-Perot sensors in PMDI for a differential-phase measurement is presented. Experimental data also demonstrates that it is feasible to design a differential-phase detection between series Fabry-Perots in PMDI.

2. Analysis of Cross-Talk in Series of Fabry-Perot Sensors

The optical system illuminated with a white-light source in Fig.1 can be analyzed by using an optical spectrum analysis. The analysis of the sensing system is accomplished by treating the sensor and read-out interferometers as spectral transfer functions, and the intensity at the detection, I, can be expressed as

$$I = \eta \int (H_t^{S_1} \cdot H_r^{S_2} \cdot H_t^{S_1} + H_r^{S_1}) \cdot H_r^{R} \cdot i(k)\, dk \,; \qquad (1)$$

Part of the SPIE Conference on Advanced Photonic Sensors and Applications
Singapore • November–December 1999
SPIE Vol. 3897 • 0277-786X/99/$10.00

where $H_r^{S_1}$ and $H_t^{S_1}$ are the reflective and transmitted spectrum transfer functions in the Fabry-Peot sensor S_1, $H_r^{S_2}$ is the reflective spectrum transfer functions in the Fabry-Peot sensor S_2, H_r^R is the reflective spectrum transfer function in the read-out system, and i(k) is the power spectrum of the source. Assuming a Gaussian distribution spectrum (a reasonably good approximation for many broad-band sources), the spectral power density of the light source is given by [6]

$$i(k) = \frac{1}{\sqrt{\pi}\sigma} \exp[-(\frac{k-k_0}{\sigma})^2]; \tag{2}$$

where σ is the spectral width that is the halfwidth between points at which the spectral density is equal to the peak value divided by e. The source is assumed to have unit intensity and a Gaussian spectral distribution symmetric about wave number k_0. For the sake of simplicity, the two series Fabry-Perot sensors and the read-out system can be assumed as low-finesse Fabry-Perots, and the transfer functions in reflections can be expressed respectively as

$$H_r^{S1} = \frac{F_S}{2}(1-\cos(2kL_{S1})); \tag{3}$$

$$H_r^{S2} = \frac{F_S}{2}(1-\cos(2kL_{S2}));$$

and $$H_r^R = \frac{F_R}{2}(1-\cos(2kL_R)); \tag{4}$$

where L_{S1} and L_{S2} are the cavity lengths in sensors, and L_R is the cavity in the read-out system. Assuming that the energy is conserved and no loss occurs in Fabry-Perots, the condition $H_r + H_t = 1$ is required. Substituting Eqs.(2) to (4) into Eq.(1), the integral is given by

$$\begin{aligned} I \cong \frac{\eta}{64} \int \frac{F_R}{2} \{ & F_S[64 + 64\cos(R) + 16\cos(R-S_1) + 16\cos(R+S_2)] \\ & + F_S^2[-32 + 32\cos(R) - 16\cos(R-S_1) - 16\cos(S_1-S_2) \\ & + 8\cos(R+S_1-S_2) + 8\cos(R-S_1+S_2)] \\ & + F_S^3[12 - 12\cos(R) - 2\cos(R-2S_1) + 8\cos(R-S_1) + 8\cos(S_1-S_2) \\ & - 4\cos(R+S_1-S_2) - 2\cos(2S_1-S_2) + 6\cos(R+S_2) \\ & + \cos(R-2S_1+S_2) - 4\cos(R-S_1+S_2)]\} i(k)dk \end{aligned} \tag{5}$$

where $S_1 = 2k_0L_{S1}$; $S_2 = 2k_0L_{S2}$; and $R = 2k_0L_R$. For a white-light interferometry, the coherence length of a short coherence light source, Lc, is much smaller than the length of cavities L_{S1}, L_{S2}, and L_R in the above equation. Different from the methodology in dealing with the cross-talk effects by diminishing proposed by [7], the terms in cos(R – (S_2 - S_1)) integrated with i(k) in Eq.(5) are the real dealing signal because the differential-phase, S_2 - S_1, is included. Therefore, the term, R – (S_2 - S_1), needs to approach to zero. It should be noted that R included in cos(R – (S_2 - S_1)) is used to extract the differential-phase for the phase demodulation. In Eq.(5), the other terms including cos(R), cos(R - S_1), cos(R + S_2), cos(S_1 - S_2), cos(2S_1 - S_2), cos(R - S_1 + S_2), cos(R - 2S_1), and cos(R - 2S_1 + S_2) are the noise effects in this study; hence, to diminish the noise effects, the following conditions are assumed as

(a) $\left|L_{S_1} - L_{S_2}\right| \gg L_C$

(b) $\left|2L_{S_1} - L_{S_2}\right| \gg L_C$

(c) $\left|3L_{S_1} - L_{S_2}\right| \gg L_C$

(d) $\left|L_{S_1} - 2L_{S_2}\right| \gg L_C$

(e) $\left|3L_{S_1} - 2L_{S_2}\right| \gg L_C$ (6)

Following the above assumptions, Eq.(5) can be reduced to

$$I = \frac{I_0 F_S F_R}{4}\left(2 - F_S + \frac{3F_S^{\,2}}{8}\right) + \frac{I_0 F_S F_R}{16}\left(F_S - \frac{F_S^{\,2}}{2}\right) \cdot Cos2k_0\left(L_R + L_{S_1} - L_{S_2}\right) \tag{7}$$

It can be seen that under the proper designs in eliminating the cross-talk effects, the real signal in a differential-phase between two series Fabry-Perots can be obtained.

3. Experimental Setup and Results

To verify the differential phase measurement between two series Fabry-Perot sensors, the experimental setup is illustrated in Fig.2. The broad-band light (FWHM ~ 50 nm; coherence length is around 30 μm) from a pigtailed super luminescent diode source with a nominal wavelength 1.3 μm, 150 μW optical power was passed through a fiber-optic coupler to two series Fabry-Perot sensors. The reflected light from the Fabry-Perots is again passed through a fiber-optic coupler that formed a low finesse Fabry-Perot cavity with mirror for the read-out system. The mirror in the Fabry-Perot cavity is bonded to a PZT stack that is placed on a linear translation stage. This stack is used to provide a phase generated carrier and the translation stage is used to adjust the cavity of the read out system. It should be noticed that the Fabry-Perot cavities are all beyond the coherence length, 30 μm, for a white-light interferometry.

It should be noticed that the lead-in/out fibers connecting two Fabry-Perot sensors are much longer than the sensors so that the effect of the spectrum transfer in the spectrum analysis can be neglected. The cavity lengths of two series Fabry-Perots are designed L_{s1} as 110 μm and L_{s2} as 270 μm, and the cavity length of the read out system is 160 μm. These chooses are complying with above requirements. As a result, the fringe patterns caused by driving the PZT stack is related to the differential-phase between the two series Fabry-Perots as shown in Fig.3. The data just shows a preliminary result, and in the future, the applications in differential parameter measurements or noise compensations will be exploited.

4. Conclusions and Discussions

Differential-phase between two series Fabry-Perot sensors for demodulation in path-matching differential interferometry (PMDI) can be designed by an optical spectrum analysis. Experimental data shows that the design complying with the special condition can achieve a differential-phase measurement between two series Fabry-Perots. Accordingly, this design in two series Fabry-Perot sensors in PMDI can be applied to a differential-parameter optical fiber sensor or noise compensations.

Acknowledgments

This work has been partially supported by National Science Council Grant to the National Cheng Kung University, Tainan, Taiwan, ROC.

REFERENCES

1. Culshaw, B. and Dakin, J., Optical Fiber Sensors: Systems and Applications, Vol.2, Norwood, MA: Artech House, 1989.
2. Farahi, F., Newton, T.P., Jones, J., and Jackson, D.A., "Coherence Multiplexing of Remote Fiber-Optic Fabry-Perot Sensing System," Optics Communications, Vol. 65, pp. 319-321, 1988.
3. Humblet, P.A. and Hamdy, W.M., "Crosstalk Analysis and Filter Optimization of Single- and Double-Cavity Fabry-Perot Filters," IEEE Journal on Selected Areas in Communications, Vol. 8, pp. 1095-1107, 1990.
4. Chang, C.C. and Sirkis, J.S., "Multiplexed Optical Fiber Sensors Using a Single Fabry-Perot Resonator for Phase Modulation," Journal of Lightwave Technology, Vol.14, pp. 1653-1663, 1996.
5. Singh, H. and Sirkis, J.S., "Dual-Parameter Optical Fiber Sensors," SPIE, Vol. 2443, pp. 258-265, 1995.Lo, Y.L., Sirkis, J.S., and Chang, C.C., "Passive Signal Processing of In-Line Fiber Etalon Sensors for High Strain-Rate Loading," Journal of Lightwave Technology, Vol. 15, pp. 1578-1586, 1997.
6. Lo, Y.L., Sirkis, J.S., and Chang, C.C., "Passive Signal Processing of In-Line Fiber Etalon Sensors for High Strain-Rate Loading," IEEE, Journal of Lightwave Technology, Vol.15, pp. 1578-1586, 1997.
7. Lo, Y.L., Tsai, M.H., and Tsao, C.C., "Spectrum Analysis in Cross-Talk of Series Fabry-Perot Sensors in Path-Matching Differential Interferometry (PMDI)," Optics & Laser Technology, Vol. 30: 6-7, pp. 395-401, 1998.

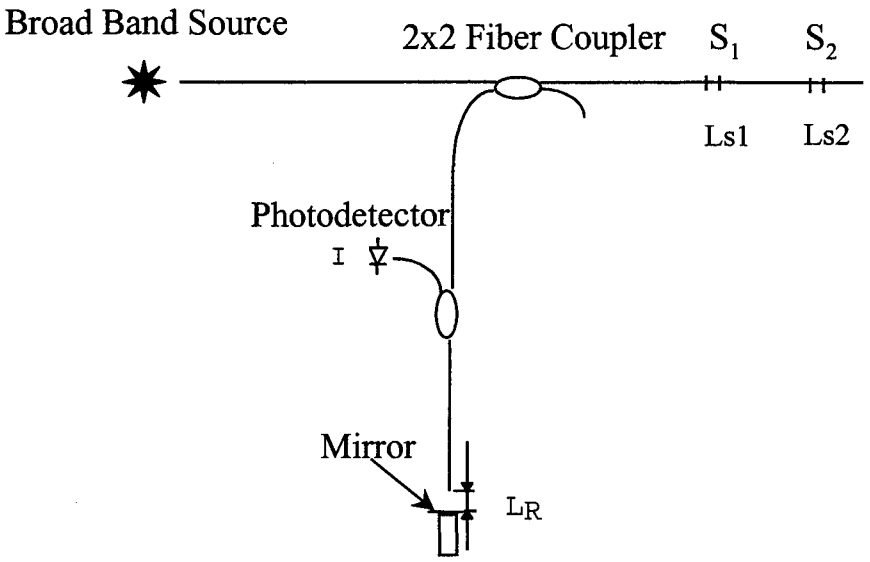


Figure 1 Configuration of PMDI for Differential-Phase Measurement

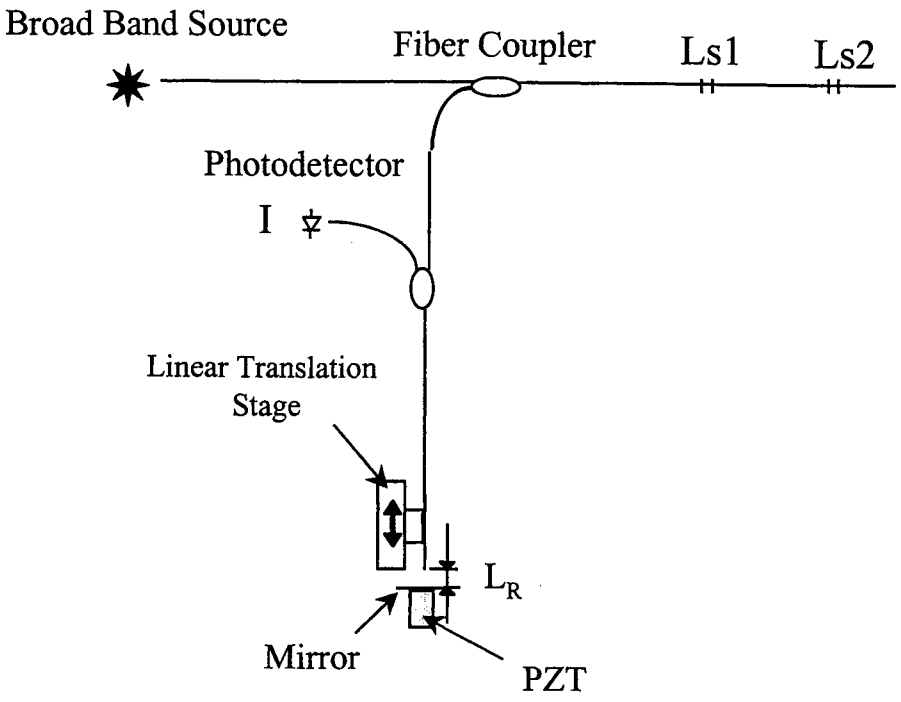


Figure 2 Experimental Setup

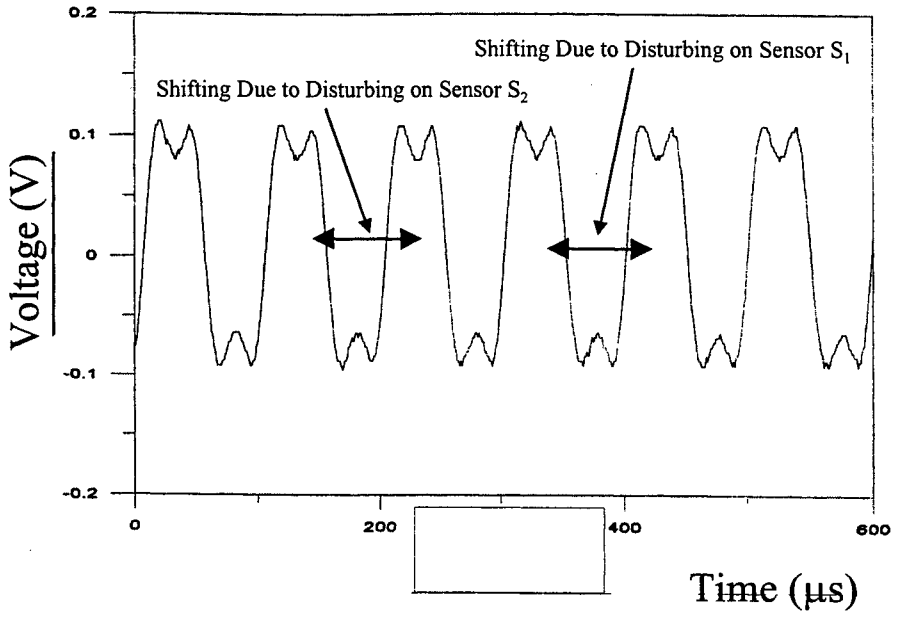


Figure 3 Experimental Data

Theory of an Amplified Closed-Sagnac Loop Interferometric Fiber Optic Gyroscope

Chin-Yi. Liaw, Yan Zhou, and Yee-Loy Lam[a]
School of Electrical and Electronic Engineering
Nanyang Technological University
Block S1, Nanyang Avenue, Singapore 639798, SINGAPORE

ABSTRACT

An Amplified Closed-Sagnac Loop Interferometric Fiber Optic Gyroscope (AC-IFOG) has been proposed. It is constructed from an open-loop interferometric fiber-optic gyroscope with an erbium-doped fiber amplifier (EDFA) that closes the sensing coil into a reentrant configuration. Earlier experiments have shown a 6.56 fold increase in rotation response factor at a rate of 5°/s over a conventional gyroscope using the same fiber length and light source. This paper reports the theoretical analysis of its operation. The analysis has been focused on the effect of the EDFA inserted within the reentrant path. The analysis shows that the rotation detection factor can be increased by at least two orders as compared to a conventional IFOG when the EDFA is an ideal device. A mathematical expression for the open-loop response has been derived and it agrees well with the experiment result.

Keywords: Active reentrant fiber gyroscope, Sagnac effect, Fiber optic sensors, Erbium-doped fiber amplifier.

1. INTRODUCTION

Conventional interferometric fiber optic rotation sensors propagate light once around a multi-turn optical fiber coil[1]. As the single trip circulation of light in the coil does not provide any amplification, the device operates passively. System noise and the sensing coil length limit the rotation detection sensitivity[2]. In the case where high precision detection capability is required, a long fiber coil with a length of a few kilometers is essential for boosting the Sagnac effect. To circumvent the long fiber requirement, active IFOGs have been experimented upon [3-4]. Generally, they make use of ARDL (Active Recirculating Delay Line) to recirculate the light to accumulate the Sagnac induced phase shift more than once within the sensing loop. An optical amplifier is generally incorporated within the loop to compensate for the loss of optical power. Two typical configurations of IFOGs using optical amplifiers are the amplified fiber ring resonator gyroscope (AFRRG) [5], and the amplified reentrant fiber gyroscope (ARFG) [3].

2. SYSTEM CONFIGURATION

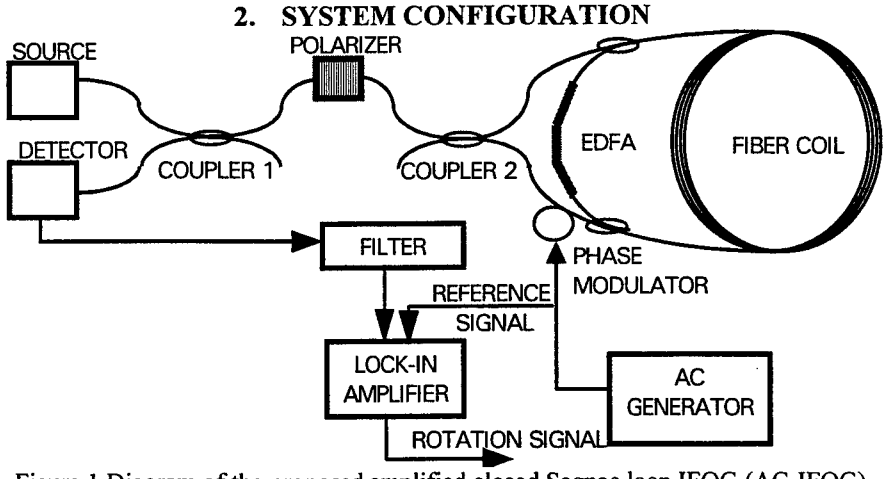


Figure 1 Diagram of the proposed amplified closed Sagnac loop IFOG (AC-IFOG).

Although the ARFG has a built in linear scale factor with frequency readout, it requires a geometrically long fiber (>10 Km) to detect low rotation rate (0.046 rad/s or 9488 °/Hr). The storage time is limited by the mechanical chopper and the rapid

[a] Correspondance: Yee-Loy Lam. Coordinator of Photonics Research Group; Email: eyllam@ntu.edu.sg; Telephone: (65)7904503; Fax: (65)7933318;

Amplified Spontaneous Emission (ASE) accumulated during each recirculation. In order to overcome these problems, we have designed and investigated a novel reentrant interferometric fiber optic gyroscope[6]. Figure 1 shows its configuration. Although it is based on the similar working principle as that of the reentrant fiber optic gyroscopes, it is however implemented in a completely different manner. With two additional couplers, a bidirectional erbium-doped fiber amplifier (EDFA) is inserted in the fiber loop to bypass the two arms of the splitter/mixer coupler (coupler 2 of Figure 1) and it links the two ends of the fiber coil together. As a result, the multi-turn optical fiber coil is closed by the insertion of this EDFA. This configuration is totally different from the previously reported Active Reentrant Fiber Gyroscope (ARFG) design[7-8], which uses only one coupler and two instead of one Sagnac loops. Also as a departure from previous ARFG operations, which are based on pulse propagation, both pulse mode and CW (Continuous Wave) operations are possible for the new design. In this paper, our new configuration is referred to as Amplified Closed-Sagnac Loop Interferometric Fiber Optic Gyroscope or AC-IFOG.

PRINCIPLE OF OPERATION OF THE AC-IFOG

For the AC-IFOG, we assume a simple case where the two EDFA couplers have exactly the same coupling ratio. Redirection of light occurs when part of the counter-propagating waves are coupled into the erbium-doped fiber before they return to the splitter/mixer coupler. These two redirected light waves will be amplified and subsequently return to the sensing coil via the *other* EDFA coupler. The recirculation of light through the multi-turn optical fiber coil causes the Sagnac phase shift induced in each and every circulation to be accumulated. The total Sagnac phase shift between a pair of counter propagating waves is therefore proportionally to the number of recirculation, *m*. If a low coherence light source is used, then only the wave pairs that have recirculated the same number of times in the loop will interfere with each other. In this case, the rotation rate detection scheme has to be interference based in order to give an output attributable to the non-reciprocal phase shift of the counter-propagating waves. The amplification provided by the EDFA should ideally overcome the loss in the light intensity during each recirculation.

Considering the Sagnac effect alone, we expect that increasing the number of signal circulation can indefinitely increase the AC-IFOG sensitivity. But this cannot be achieved in a real device because of ASE (Amplified Spontaneous Emission) noise that is built up during recirculation, and the pump intensity noise. Therefore, a possible operating technique is to have the Erbium-Doped Fiber Amplifier turned on for a specific duration to restrict the number of recirculation. Assuming that this can be carried out, the analytical expression for the output of the AC-IFOG can be derived. While detailed derivation can be referred to our previous paper[6], a simplified equation, representing the output of the AC-IFOG in CW mode right after *n* recirculation, can be given by

$$T(\Delta\phi_R) = \frac{1}{2} t_o^2 \sum_{m=1}^{n} A_{RC}^{2m} [1 + \cos(m\Delta\phi_R)] \tag{1}$$

where $T(\Delta\phi_R)$ represents the output/input transmission ratio of light exiting/entering the Sagnac loop. The net amount of amplification (or loss) realized for each round-trip (A_{RC}) is equal to the gain of the EDFA less its insertion loss (due to the WDMs and the couplers) and losses in the sensing coil. The variable *m* represents the recirculation number and *n* represents the ON-duration of the Erbium Doped Fiber Amplifier measured in the number of trips around the loop. The parameter t_o^2 is a transmission coefficient (or constant) and $\Delta\phi_R$ is the single round trip Sagnac phase shift.

SENSITIVITY OF CW EDFA OPERATION

The sensitivity in this case refers to the minimum detectable relative change in the transmission ratio of the gyro with respect to a corresponding change in the rotation-induced phase shift. This is usually the highest at the $\pi/2$ bias point on the response curve for a conventional IFOG. Using this terminology, the sensitivity of a gyroscope is defined as $S(\Delta\phi_R) = dT(\Delta\phi_R)/d\Delta\phi_R$ [9]. To determine the sensitivity of the proposed AC-IFOG, the gyro output $T(\Delta\phi_R)$ in Equation 1 can be expressed in the form below (using the relation $\sum_{m=1}^{n} A^m = \frac{A - A^{n+1}}{1 - A}$)

$$T(\Delta\phi_R)=\frac{1}{2}t_o^2\left\{\frac{A_{RC}^2-A_{RC}^{2(n+1)}}{1-A_{RC}^2}\right.$$
$$+\frac{A_{RC}^2\cos(\Delta\phi_R)-A_{RC}^4}{1-2A_{RC}^2\cos(\Delta\phi_R)+A_{RC}^4} \qquad (2)$$
$$\left.+\frac{A_{RC}^{2(n+2)}\cos(n\Delta\phi_R)-A_{RC}^{2(n+1)}\cos[(n+1)\Delta\phi_R]}{1-2A_{RC}^2\cos(\Delta\phi_R)+A_{RC}^4}\right\}$$

by differentiation we obtain the equation for sensitivity as

$$S(\Delta\phi_R)=\frac{1}{2}t_o^2\left\{\frac{(A_{RC}^4-1)}{(1+A_{RC}^4)}\frac{a\sin(\Delta\phi_R)+}{[1-a\cos(\Delta\phi_R)]^2}\right.$$
$$+\frac{an}{2}A_{RC}^{2(n+1)}\frac{[a\sin((n-1)\Delta\phi_R)-\sin(n\Delta\phi_R)]}{[1-a\cos(\Delta\phi_R)]^2} \qquad (3)$$
$$\left.-\frac{a}{2}A_{RC}^{2n}\frac{[a\sin(n\Delta\phi_R)+(an\cos(n\Delta\phi_R)-(n+1))\sin((n+1)\Delta\phi_R)]}{[1-a\cos(\Delta\phi_R)]^2}\right\}$$

where $a=\dfrac{2A_{RC}^2}{1+A_{RC}^4}$

From the above equation, the optimum $\Delta\phi_R$ that makes $S(\Delta\phi_R)$ maximum is found by differentiating and setting $dS(\Delta\phi_{opt})/d\Delta\phi_{opt}=0$. The resultant equation is lengthy but it can be expressed in a form as shown below and solved numerically by a computer.

$$\frac{x}{1-x}\left\{-n^2x^{2n}+\frac{2nx^{2n}-2(x^2-x^{2(n+1)})/(1-x^2)}{(x^2-1)}+\frac{1-x^{2n}}{1-x^2}\right\}$$
$$+\frac{y}{1-y}\left\{-n^2y^{2n}+\frac{2ny^{2n}-2(y^2-y^{2(n+1)})/(1-y^2)}{(y^2-1)}+\frac{1-y^{2n}}{1-y^2}\right\}=0 \qquad (4)$$

where $x=A_{RC}^2e^{j\Delta\phi_{opt}}$ and $y=A_{RC}^2e^{-j\Delta\phi_{opt}}$

By substituting ϕ_{opt} into Equation 3, the maximum sensitivity can be obtained. In Ref 6 it has been shown that $t_o^2=(1-\gamma)(1-\kappa)e^{-\alpha L}$, where $\sqrt{(1-\gamma)}$ is the transmission factor of the coupler, $\sqrt{(1-\kappa)}$ is the through coupling ratio of the coupler, and α is the fiber propagation loss and L is the fiber length. If we assume that there is no fiber propagation loss, i.e. α=0, the maximum sensitivity can be plotted for different values of the cross coupling ratio κ and the single round-trip gain A_{RC} as shown in Figure 2. As can be seen, the maximum sensitivity increases tremendously as A_{RC} approaches unity. However, for different coupling ratio κ, the maximum sensitivity profiles are different. Generally, it can be seen that lower coupling ratios have a much higher sensitivity as compared to higher coupling ratios. The drop in the maximum sensitivity at higher coupling ratios may be explained by the fact that less light gets coupled into the sensing coil.

The sensitivity improvement of the proposed AC-IFOG as compared to a conventional IFOG can be indicated by the ratio of the AC-IFOG maximum sensitivity against that of the fundamental IFOG. Let's define this as the sensitivity boost factor, *SBF*, then

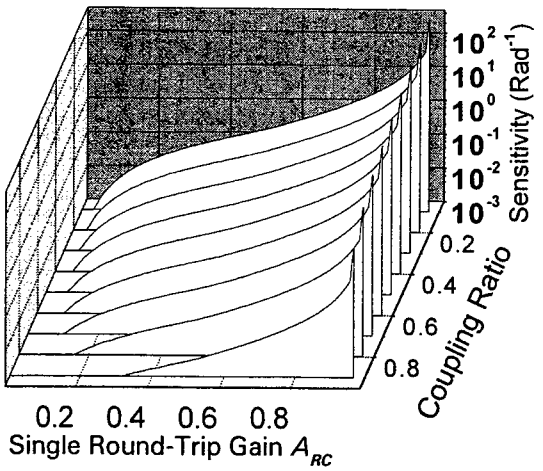


Figure 2 Maximum Sensitivity as a function of round-trip gain & Coupling Ratio.

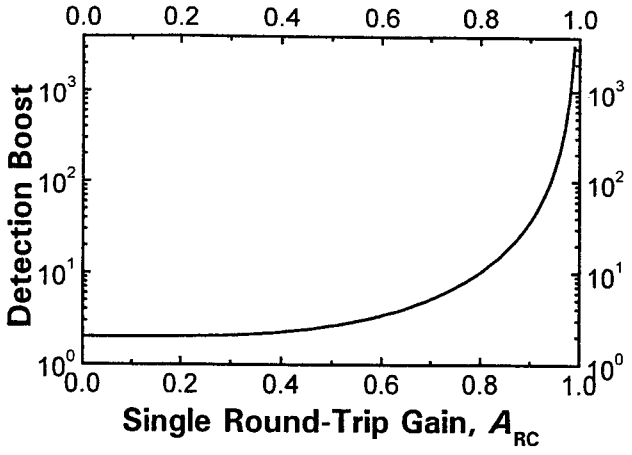


Figure 3 Detection boost versus single round-trip gain.

$$SBF = \frac{S_{\max}^{AC-IFOG}(\phi_{opt}^{AC-IFOG})}{S_{\max}^{I-FOG}(\phi_{opt}^{I-FOG})} \tag{5}$$

Recall that the AC-IFOG with $n = 1$ has a response profile equivalent to that of a fundamental IFOG, the sensitivity for this AC-IFOG can thus be found by setting n to 1 in Equation 3. By setting the bias point to $\Delta\phi_R = \pi/2$, we find that its maximum value is

$$S_{\max}^{I-FOG}(\phi_{opt}^{I-FOG}) = \frac{1}{2} t_o^2 A_{RC}^2 \tag{6}$$

Using the optimum ϕ_{opt} for Equation 3 and substituting it into Equation 5, we obtain

$$\begin{aligned} SBF = \frac{1}{A_{RC}^2}\Bigg\{ & \frac{(A_{RC}^4 - 1)}{(1 + A_{RC}^4)} \frac{a\sin(\Delta\phi_R) + }{[1 - aCos(\Delta\phi_R)]^2} \\ & + \frac{an}{2} A_{RC}^{2(n+1)} \frac{[aSin((n-1)\Delta\phi_R) - Sin(n\Delta\phi_R)]}{[1 - aCos(\Delta\phi_R)]^2} \\ & - \frac{a}{2} A_{RC}^{2n} \frac{[aSin(n\Delta\phi_R) + (anCos(n\Delta\phi_R) - (n+1))Sin((n+1)\Delta\phi_R)]}{[1 - aCos(\Delta\phi_R)]^2} \Bigg\} \end{aligned} \tag{7}$$

In Figure 3, the *SBF* is plotted against A_{RC} and it shows the substantial enhancement as A_{RC} is increased. The plot clearly indicates that if a sufficiently high single round-trip gain ($A_{RC}>0.9$) is chosen, a sensitivity boost of at least two orders of magnitude can be attained.

CW MODE OPERATION WITH OPTICAL PHASE MODULATION

In the previous section, we have discussed the operation of the AC-IFOG in CW mode and the possible enhancement in the maximum sensitivity. In order to make use of the maximum sensitivity, a gyroscope must be biased and this is generally achieved with optical phase modulation. An optical phase modulator can be placed at one arm near coupler C2 as shown in Figure 1. Once a modulation signal $\phi(t)$ is produced by the phase modulator, for the primary wave pair, the phase difference between the counter-propagating waves $\Delta\phi_{mod}(t)$ becomes,

$$\Delta\phi_{\text{mod}}(t) = \phi_{CCW}(t) - \phi_{CW}(t) = \phi(t) - \phi(t-\tau) \tag{8}$$

where ϕ_{CCW} and ϕ_{CW} represents the phases of the optical waves that propagate in the counter-clockwise and clockwise directions in the sensing loop respectively, and τ (equal to $n_f L/c$, with n_f being the refractive index of the fiber) is the transit time of light in the fiber loop or the loop transit time. The interference signal becomes

$$I(\Delta\phi_R) = I_o\{1+\cos[\Delta\phi_R + \Delta\phi_{\text{mod}}(t)]\} \tag{9}$$

where $\Delta\phi_R$ is the nonreciprocal phase difference introduced by rotation. For sinusoidal phase modulation with a modulation depth ϕ_o, the biasing phase modulation is $\phi(t) = \phi_o \sin(\omega t)$ and Equation 8 becomes

$$\Delta\phi_{\text{mod}}(t) = \phi_o\{\sin(\omega t) - \sin(\omega(t-\tau))\} \tag{10}$$

which can be simplified to

$$\Delta\phi_{\text{mod}}(t) = 2\phi_o \sin(\omega\frac{\tau}{2})\cos\left(\omega(t-\frac{\tau}{2})\right) \tag{11}$$

However, for the m^{th} wave pairs, Equation 11 becomes

$$\Delta\phi_{\text{mod}}(t) = 2\phi_o \sin(m\omega\frac{\tau}{2})\cos\left(\omega(t-m\frac{\tau}{2})\right) \tag{12}$$

Taking into consideration all the wave pairs, Equation 1 can be modified to include the optical phase modulation effect and re-written as

$$T(\Delta\phi_R) = \frac{1}{2}T_o\sum_{m=1}^{n} A_{RC}^{2m}[1+\cos(m\Delta\phi_R + 2\varphi_{\text{mod}}\cos[\omega(t-m\frac{\tau}{2})])] \tag{13}$$

where $\varphi_{\text{mod}} = \phi_o \sin(m\omega\frac{\tau}{2})$ and $T_o = t_o^2$, and n is a number that can approach infinity.

Equation 13 can be expanded into the odd and even components represented by the Bessel function of the first kind,

$$\begin{aligned} T(\Delta\phi_R) = \frac{1}{2}T_o\sum_{m=1}^{n} A_{RC}^{2m}\Big\{1+\{J_o(2\varphi_{\text{mod}})+2\sum_{k=1}^{\infty}J_{2k}(2\varphi_{\text{mod}})\cos[2k\omega(t-m\frac{\tau}{2})]\}\cos(m\Delta\phi_R) \\ -\{2\sum_{k=1}^{\infty}J_{2k-1}(2\varphi_{\text{mod}})\sin[(2k-1)\omega(t-m\frac{\tau}{2})]\}\sin(m\Delta\phi_R)\Big\} \end{aligned} \tag{14}$$

We can extract the component at ω by demodulation, i.e. multiplication by $\sin(\omega t)$ followed by low pass filtering to extract the rotation rate information and in doing so,

$$T(\Delta\phi_R) = -\frac{1}{2}T_o\sum_{m=1}^{n} A_{RC}^{2m}J_1(2\varphi_{\text{mod}})\cos[m\omega\frac{\tau}{2}]\sin(m\Delta\phi_R) \tag{15}$$

As we are concerned with the maximum sensitivity biasing point, where the differentiation of Equation 15 with respect to $\Delta\phi_R$ is of the largest absolute value, the equation can thus be differentiated to yield the sensitivity expression:

$$\frac{dT(\Delta\phi_R)}{d\Delta\phi_R} = \frac{1}{2}T_o\sum_{m=1}^{n} mA_{RC}^{2m}J_1(2\varphi_{\text{mod}})\cos[m\omega\frac{\tau}{2}]\cos(m\Delta\phi_R) \tag{16}$$

Assuming that the rotation rate is very small, we can approximate $\cos(m\Delta\phi_R) \approx 1$. The demodulated sensitivity equation can be normalized by $T_o/2$ for the ease of simulation.

1. Effect of A_{RC} on Sensitivity

Using the normalized demodulated sensitivity equation, a series of simulations are performed with A_{RC} equal to 0.1, 0.5 and 0.9. In Equation 16, for $m>10$, it can be shown that the corresponding summation terms become negligibly small, the simulation is thus limited to just n = 10. The objective of the simulation is to examine the dependence of the normalized sensitivity on A_{RC} (the net single round-trip gain), ω (the modulation frequency) and ϕ_o (the modulation depth) so as to

determine the maximum sensitivity biasing operating parameters. Figure 4(a1) shows the simulation result when A_{RC}=0.1. As can be seen, the optimum phase modulation bias point varies for each combination of modulation frequency and modulation depth. This figure also shows that as the modulation depth increases, the required optimum modulation frequency to produce the same level of sensitivity is reduced. Similar trends can be observed in Figure 4(a2) for A_{RC}=0.9. By comparing Figure 4(a1) and Figure 4(a2), it can be observed that as A_{RC} increases, the maximum sensitivity point is shifted towards the lower modulation frequency region. There is also an overall substantial increase in sensitivity (three orders of magnitude) but the optimum-biasing region becomes narrower. The sharpness of the normalized sensitivity response increases tremendously when the single round-trip gain (A_{RC}) increases. This indicates that the amount of modulation (i.e. the modulation depth) required to bias the gyroscope at its maximum sensitivity have to be reduced to within a narrower range.

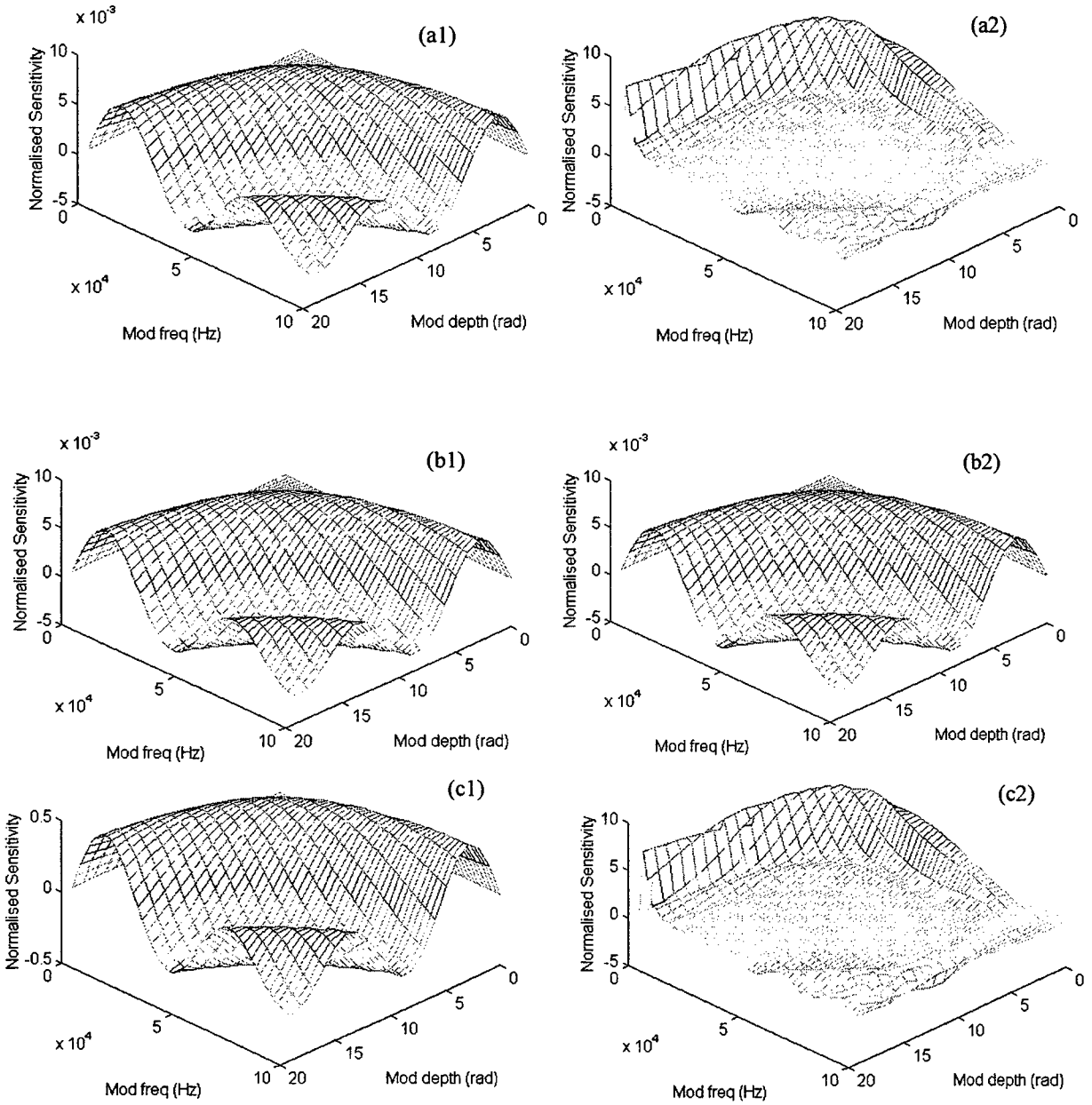


Figure 4 Normalized sensitivity plot (a1) A_{RC}=0.1,n=10, (a2) A_{RC}=0.9,n=10, (b1) A_{RC}=0.1,n=1, (b2) A_{RC}=0.1,n=100, (c1) A_{RC}=0.9,n=1, (c2) A_{RC}=0.9,n=100.

2. Effect of n on Sensitivity

The variable n indicates the total number of recirculated components that are summed at the detector to give the resultant output. In the previous paragraph, it has been pointed out that for $m>10$, the corresponding summation terms in Equation 16 are negligibly small. This claim is confirmed with the analysis of the effect of n on the normalized sensitivity here. It should be noted that although the present discussion is made on the assumption that the EDFA is operating in CW Mode without switching, the variable n could in fact be controlled by the switching-on time of the EDFA if necessary. Figure 4(b1) & Figure 4(b2) show the results of the analysis when A_{RC}= 0.1 but n=1 and n=100 respectively. It can be noted that when the single round-trip gain A_{RC} is small, there is negligible change in the 3-D profile of the normalized sensitivity, even if the total number of recirculation, n, is varied from 1 to 100. As is expected, when the single round-trip gain is large, say A_{RC} = 0.9, a change in n from 1 to 10 causes the normalized sensitivity curve to shift dramatically, as can be seen by comparing Figure 4(c1) & Figure 4(a3). However, when a comparison is made between Figure 4(a3) and Figure 4(c2), where the n value varies from 10 to 100, it can be found that the change in the profile of the normalized sensitivity is negligible. In fact, this is not surprising because in Equation 16, the dependence of each summation term on m is dominated by mA_{RC}^{2m}, which would approach zero as m increases beyond 10 simply because A_{RC} must be kept at a value of less than unity to avoid lasing of the gyro coil.

3. Open loop response

Using Equation 15, the open loop response is plotted in Figure 5 by substituting the following parameters in the equation: coil length of 200 m, diameter of 0.16 m, refractive index of 1.46, modulation frequency of 102KHz. Three curves are plotted for arbitrarily chosen A_{RC} of 0.6, 0.8 and 0.9 while the modulation depth, $\Delta\phi_{opt}$ is computed for optimum bias at these points. These are compared with the experimental result obtained[6] as shown in the inset, both the experimental results and the simulation have identical profiles at each level of A_{RC} (or EDFA pump power). When the A_{RC} or pump power decreases, the peaks begin to shift outwards. A decrease in the response slope at the origin indicates a drop in the detection sensitivity.

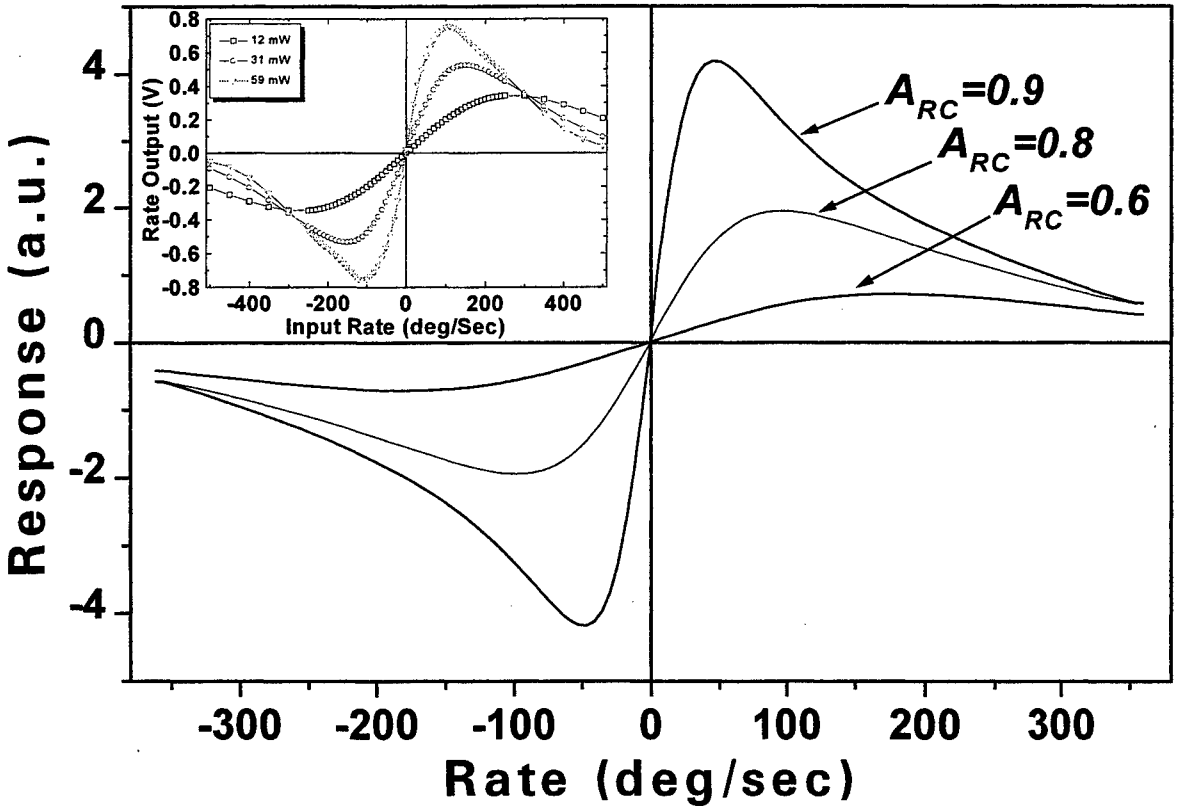


Figure 5 Open loop response for A_{RC}=0.6, 0.8, 0.9 with all other optical phase modulation parameters optimized.

3. CONCLUSION

A novel Reentrant Fiber Optic Gyroscope has been proposed and its principle of operation in CW mode has been analyzed theoretically. For CW mode operation, the analysis shows that a scheme making use of optical phase modulation and electronic demodulation is feasible. The sensitivity can be controlled by the gain of the EDFA and can be substantially improved by at least two orders of magnitude as compared to a conventional IFOG. The simulated open-loop response has profiles that match well with those obtained from the previous experiment.

4. REFERENCES

1. Vali, V., and R. W. Shorthill, "Fiber Ring Interferometer," Applied Optics, Vol.15, 1976, pp. 1099-1100.
2. Hervé Lefèvre, *The fiber-Optic Gyroscope*, Artech House, 1993, pp. 19-25.
3. E. Desurvire, B. Y. Kim, K. Fesler, and H. J. Shaw, "Reentrant Fiber Raman Gyroscope", IEEE Journal of Lightwave Technology, Vol. 6, No. 4, April 1988, pp. 481-491.
4. C.Y.Liaw, Y.Zhou, and Y.L.Lam, "An Erbium fiber based sensitivity booster used by a fiber optic gyroscope", February 22-27, 1998 Optical Fiber Communication Conference, San Jose, California, pp. 370-371.
5. J. T. Kringlebotn, "Amplified fiber ring resonator gyro," IEEE Photonics Technology Letter, Vol. 4, No. 10, 1992, pp. 1180.
6. C.-Y. Liaw, Y. Zhou, and Y.-L. Lam, "Characterization of an Open-Loop Interferometric Fiber-Optic Gyroscope with the Sagnac Coil Closed by an Erbium-Doped Fiber Amplifier", IEEE/OSA Journal of Lightwave Technology, Vol.16, No.12, December 1998, pp.2385-2392.
7. G.A. Pavlath & H. J. Shaw, "Reentrant Fiber Optic Rotation Sensors", in Fiber-Optic rotation sensors and Related Technologies, by S.Ezekiel & H.J.Arditty, Springer-Verlag, 1982, pp. 364.
8. D. N. Chen, K. Motoshima, M. M. Downs, and E. Desurvire, "Reentrant Sagnac Fiber Gyroscope with a Recirculating Delay Line Using an Erbium-Doped Fiber Amplifier", IEEE Photonics Technology Letters, Vol. 4, No. 7, July 1992, pp. 813-815.
9. Yu & A.S. Siddiqui, "Theory of a novel high sensitivity optical fibre gyroscope", IEE Proceedings-J, Optoelectronics, Vol. 140, No. 2, April 1993, pp. 150-156.
10. E. Desurvire, Moshe Tur, and H. J. Shaw, "Signal-to-Noise Ratio in Raman Active Fiber Systems: Application to Recirculating Delay Lines", IEEE Journal of Lightwave Technology, Vol. 4, No. 5, 1986, pp. 560-566.
11. N. A. Olsson, "Lightwave Systems With Optical Amplifiers", IEEE Journal of Lightwave Technology, Vol. 7, No. 7, 1989, pp. 1071-1083.
12. S. Ezekiel and H. J. Arditty, *Fiber-Optic Rotation Sensors and Related Technologies*, Proceedings of the First International Conference 1981, pp. 9.

Intensity-based fiber optic pressure sensor with an $Au/NiCr/Si_3N_4/SiO_2/Si_3N_4$ diaphragm

Jaehee Park, Myung-Gyoo Kim*, Yohee Kim**, Kwang-Soo Kim**, Ilmin Kim***

Keimyung University, Dept. of Computer and Electronics Eng., 1000, Shindang-Dong, Dalseo-Gu, Daegu, Korea, *Lightron Fiber-Optic Devices Inc., TBI Center 505, Chungnam National university, 220, Kung-Dong, Yusong-Gu, Daejon, 305-764, Korea, **Optical Technology Research Center, KERI, P.O. Box 20, Changwon, Kyungnam, 641-120, Korea, *** Hansung University, Dept. of Computer Science, 389, Samaung-Dong, Sungbuk-Gu, Seoul, Korea

Abstract

Intensity based fiber optic pressure sensor has been fabricated using micromachinning technology. The sensor consists of a multimode fiber bundle and an 100nm-Au/30nm-NiCr/150nm-Si_3N_4/300nm-SiO_2 /150nm-Si_3N_4 diaphragms supported by a micromachined frame-shape silicon substrate. The distance between the fiber bundle and the diaphragm of the pressure sensor is 900 um. Three fiber optic pressure sensors having different size diaphragms were fabricated and used in experiments. The pressure sensitivities were 20.69 nW/kPa, 26.70 nW/kPa, and 39.33 nW/kPa for the pressure sensors with diaphragms of 3x3 mm^2 area, 4x4 mm^2 area, and 5x5 mm^2 area.

Keywords: $Au/NiCr/Si_3N_4/SiO_2/Si_3N_4$ thin film, silicon micromachining, and optical fiber pressure sensor

Introduction

Recently, researchers have developed many fiber optic sensors, such as magnetic, acoustic, pressure, temperature, acceleration, gyro, displacement, current, and strain sensors . These sensors possess several advantages over conventional sensors, including immunity to electromagnetic interference, rapid response, resistance to corrosion, and capability for remote sensing [1]. Small size optical fiber pressure sensors have been produced using micromachinning technology. The pressure sensor configured as a Fabry-Perot interferometer was fabricated[2]-[3]. This sensor had good sensitivity but had the difficulty in readiing absolute applied pressure. A micromachinned intensity-based fiber optic pressure sensor for medical instruments such as catheters and endoscopes was developed using a Si diaphragm[4]-[5]. The sensor employed a simple fabricating process but had low sensitivity. In this paper we investigates high performance intensity-based fiber optic pressure sensor fabricated using micromachinning technology. The sensor consists of a multimode fiber bundle and an 100nm-Au/30nm-NiCr/150nm-Si_3N_4/300nm-SiO_2/ 150nm-Si_3N_4(Au/NiC_r/N/O/N) diaphragm supported by micromachined frame-shape silicon substrate. In this fiber optic pressure sensor, the light from source propagates into multimode fiber and reachs at the micromachined diaphragm. The diaphragm returns the light to an optical detector through other multimode fibers. When pressure is applied to the diaphragm, the diaphragm is deflected. The diaphragm deflection changes the intensity of the light reflected from the diaphragm. From the changes of reflected light intensity, pressure is measured.

Part of the SPIE Conference on Advanced Photonic Sensors and Applications
Singapore • November–December 1999
SPIE Vol. 3897 • 0277-786X/99/$10.00

Fabrication of the sensor

The fabrication procedures of a Au/NiC$_r$/N/O/N diaphragm is illustrated in figure 1. First, 150 nm-Si_3N_4/300 nm-SiO_2/150 nm-Si_3N_4 (N/O/N) dielectric films are deposited on a 610 mm thick, 5", p-type, <100> oriented silicon wafer. Top and bottom layers are grown by low pressure chemical vapor deposition (LPCVD) at 780 and a mid-SiO_2 layer is grown by atmosphere pressure CVD (APCVD) at 380 under the SiH_4 and N_2O mixture gas ambient. This triple dielectric thin film has relatively stable mechanical stress characters and very low thermal conductivity [6,7]. For protecting the N/O/N thin film and acting as the masking material during the reactive ion etching (RIE), the top and bottom surfaces of a silicon wafer deposited by a N/O/N thin film are coated with a positive photoresist. After the photolithography process, the wet etching using HF and dry etching using only CF_4 gas are carried out for patterning the one surface of the N/O/N thin film. Finally, the wafer is anisotropically etched downward from the backside up to the desired depth in the anisotropic etchant 44 wt.% KOH solution. In the anisotropic etching process, the N/O/N thin film acts as the masking material. Finally, 30 nm-NiCr/100 nm-Au films is deposited on the backside of the diaphragm by thermal evaporation, respectively.

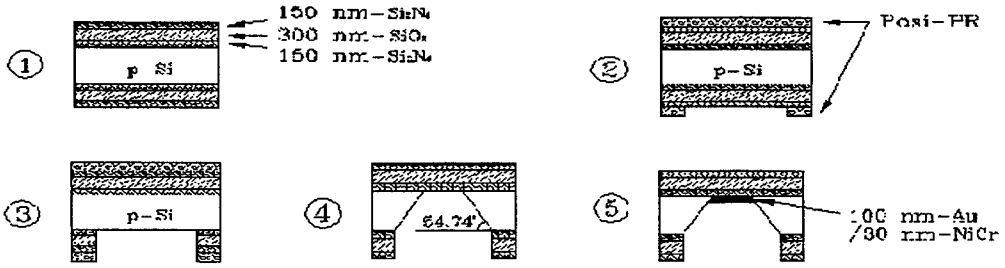


Fig. 1.Fabrication procedures of Au/NiCr/N/O/N optical reflection layer.

The sensing element is shown in figure 2. It consists of a fiber bundle having seven pieces of multimode fiber and an Au/NiCr/N/O/N diaphragm. The piece of fiber centered ar a fiber bundle is used for transmitting the light emitted from light source to the diaphragm. On the other hand other pieces of fiber receive the light reflected from the diaphragm and transmit it to an optical power meter. The Au/NiCr/N/O/N diaphragm is attached at the jig using a small amount of epoxy. And a fiber bundle is inserted into the jig and was aligned at the center of the diaphragm. Finally, the optical fiber bundle is bonded to the jig with UV curing adhesive. Three pressure sensors were fabricated using diaphragms and fiber bundles with the diameter of 780 um . The diaphragm areas of pressure sensors were 3x3 mm^2, 4x4 mm^2, and 5x5 mm^2. And the reflectances of three diaphragms were all about 95%.

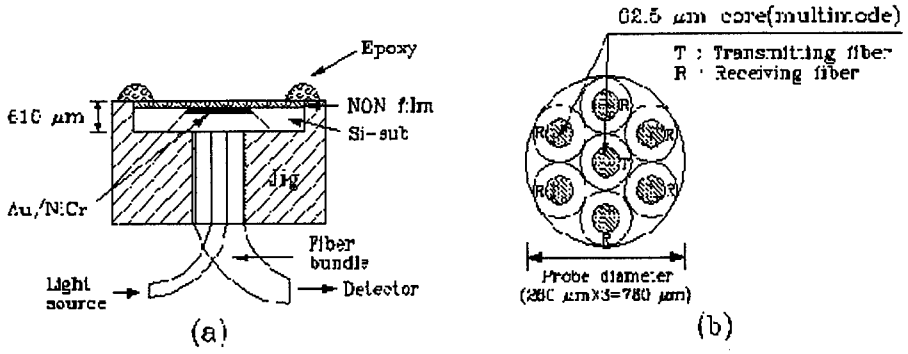


Fig. 2. An optical fiber sensing element.

Experimental results

The experiment arrangement is shown in figure 3. It composes of a light source, an optical power meter(model 731, Graseby Optronics, USA), an air pump, and a digital manometer (OMEGA, PCL-200 kit D). The light from the source is transmitted through the multimod fiber centered at the fiber bundle to the sensing element to the diaphragm on the top of sensing element. At the diaphragm the light is reflected and is coupled into other multimode fibers. The coupled light returns to the optical power meter through other multimode fibers. When the pressure is applied to the diaphragm by an air pump, the diaphragm is deflected. The diaphragm deflection induces optical intensity variations of the light reflected from the diaphragm. From the intensity variations of reflected light, the applied pressure is measured.

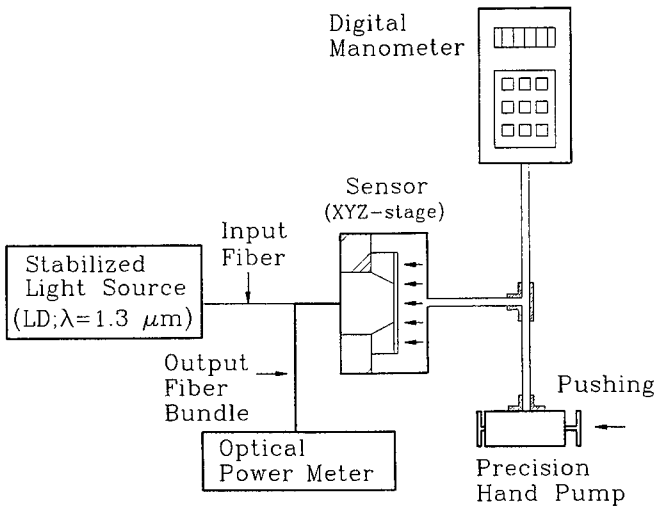


Fig. 3. Experimental setup.

Figure 4 shows the optical power of reflected light without applying pressure to the diaphragm with the area of 4x4 mm^2 as a function of the distance between the diaphragm and the fiber bundle end. From 500um to 1,150um the power increases as the distance increases and over 1,1500um the power decreases as the distance increases. In the distance range from 500um to 900um, the reflected power changes linearly along the distance and the slope is maximum. For getting maximum pressure sensitivities the distance has to be 900um.

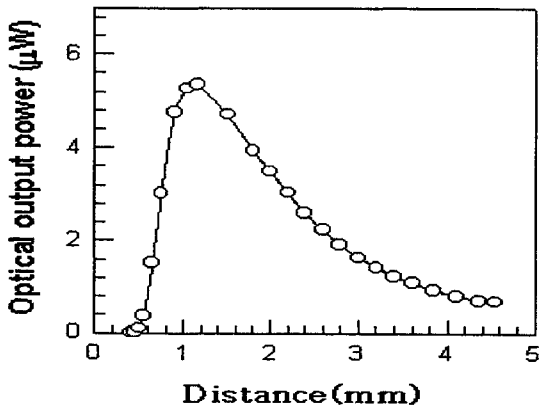


Fig. 4. The optical power along the distance between the diaphragm and the fiber bundle end.

Three fiber optic pressure sensors were used in experiments. The distances between the fiber bundle ends and the diaphragms were all about 900 um. Figure 5 shows the optical power of light reflected from the diaphragm alone the applied pressure. The pressure sensitivities are 20.69 nW/kPa, 26.70 nW/kPa, and 39.33 nW/kPa for the fiber optic pressure sensors with the diaphragms of 3x3 mm^2 area, 4x4 mm^2 area, and 5x5 mm^2 area. These results show that the sensitivity increases as the diaphragm area increases.

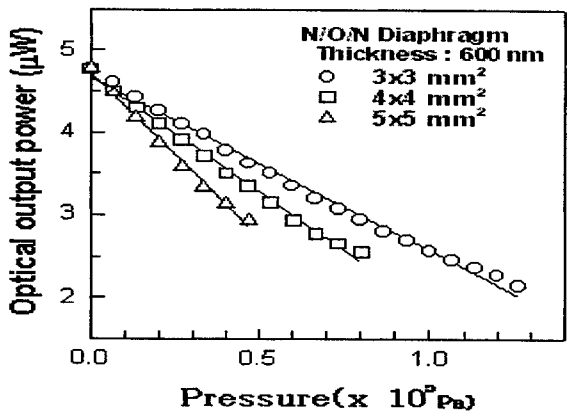


Fig. 5. Optical output power as a function of the applied pressure.

Conclusions

we demonstrates high performance intensity-based fiber optic pressure sensors consisted of a multimode fiber bundle and an 100nm-Au/30nm-NiCr/150nm-Si_3N_4/300nm-SiO_2/150nm-Si_3N_4(Au/NiC_r/N/O/N) diaphragms supported by micromachined frame-shape silicon substrates. Three fiber optic pressure sensors with the diaphragms of 3x3 mm^2 area, 4x4 mm^2 area, and 5x5 mm^2 area were used in experiments. The pressure sensitivities are 20.69 nW/kPa, 26.70 nW/kPa, and 39.33 nW/kPa for the fiber optic pressure sensors with the diaphragms of 3x3 mm^2 area, 4x4 mm^2 area, and 5x5 mm^2 area. These results show that the sensitivity increases as the diaphragm area increases. We expect that this intensity-based sensor can be used for medical applications due to small size and high sensitivity. If a diaphragm fabricated using different material and different thickness is used, the different pressure sensitivity is anticipated.

Acknowledgements: This work was supported by a grant(# HMP-98-E-1-0005) of Good Health R&D.

References

[1] T. G. Giallorenzi, J. A. Bucaro, A. Dandridge, G. H. Sigel, JR., James H. Cole, S. C. Rashleigh and R. G. Priest, "Optical fiber sensor technology," *IEEE J. Quantum Electron,* QE-18, pp. 626-665, 1982.

[2] M. Chan, "A micromachinned pressursee nsor with fiber-optic interferometric readout," *Sensors and Actuators*, vol 43, pp196-201, 1994.

[3] M. Kim, J. Park, S. Kang, and B. Shon, "Fiber optic Fabr y-Perot pressure sensor with $Si_3N_4/SiO_2/Si_3N_4$ diaphragm fabricated using Micromachinning technology," *Smart Electronics and MEMS*, SPIE vol 3242, pp.347-353, 1997.

[4] O. Tohyama, M. Kohashi, K. Yamamoto, and H. Itoh, " A fiber-optic silicon pressure sensor for ultra-thin catheters, *The proceedings of Transducers'95 and Eurosensors IX*, pp. 596-599, 1995.

[5] Y. Haga, K. Minami, S. Shoji, S. Nitta, M. Tanaka and M. Esashi, "A fiber-optic ultra miniature pressure sensor," *Technical Digest of The 12th Sensor Symposium,* pp. 135-138, 1994.

[6] F. Volklein, "Thermal conductivity and diffusivity of a thin film SiO_2-Si_3N_4 sandwich system," *Thin Solid Films*, vol. 188, pp. 27-33, 1990.

[7] Myung-Gyoo Kim, Dong-Su Park, Chang-Won Kim, Jin-Sup Kim, Jung-Hee Lee, Jong-Hyun Lee, Byung-Ki Sohn, "Fabrication of stress-balanced Si3N4/SiO2/Si3N4 dielectric membrane," J. Korea Sensors Society, vol. 4, no. 3, pp. 51-59, 1995.

Surface roughness measurement with optical scatterometry

Jyrki Saarinen*[a,b], Ilkka Kallioniemi[b], Ari Niinistö[b], and Ari T. Friberg[c]

[a]Heptagon Oy, Tekniikantie 12 (Innopoli), FIN-02150 Espoo, Finland
[b]Department of Engineering Physics and Mathematics, Helsinki University of Technology, FIN-02015 HUT, Espoo, Finland
[c]Department of Physics II - Optics, Royal Institute of Technology, S-100 44 Stockholm, Sweden

ABSTRACT

Scattering of light by random rough surfaces can be numerically simulated by using an exact electromagnetic scattering theory. Unfortunately, the characterization of surfaces is almost impossible owing to the nonuniqueness of the inverse scattering problem and highly nonlinear relationship between the surface parameters and the scattering. Thus, existing practical methods for qualitative or quantitative characterization are almost entirely experimental. Here we apply neural networks for estimating statistically the surface parameters. Previously, we have successfully demonstrated that neural network as a statistical estimator for optical scatterometry is an efficient tool for characterizing periodic microstructures [I. Kallioniemi *et al.* in *Appl. Opt.* **37**, pp. 5830–5835, (1998)]. We generate numerically random surfaces, which are characterized with the degree of roughness, i.e., root-mean-square (rms) amplitude of the roughness and correlation length. Here we are mainly interested in the most demanding region of the rms amplitude in the so-called resonance domain, corresponding to height fluctuations and correlations up to 5 times the wavelength of light. The neural network model, which is here a self-organizing map (SOM), is first trained and calibrated with the known surface parameter and scattering data pairs. At characterization stage, using only measured intensity distributions, the neural network theory classifies surface parameters into discrete classes of the rms amplitude and the correlation length. For most cases the classification result deviates at most one class, corresponding to 0.5 wavelengths, from the correct values.

Keywords: neural networks, optical scatterometry, rough surface, scattering

1. INTRODUCTION

In industry there is a need for characterizing random rough surfaces, such as ceramics, metal, paper, and wood, and for finding defects or determing the degree of roughness. Due to the lack of sophisticated and quantitative methods, surfaces are mainly examined by using visual inspection, by comparing them manually to exemplars with known roughness, or with mechanical profilometry. None of these is capable of both reliable, fast, and noninvasive quantitative measurement, which does not cause any plastic deformations on a sample.

Light interacting with wavelength-scale roughness encounters scattering phenomenon: besides specular reflection, scattering distribution has a diffusive component propagating light into all directions. The spreading of the intensity of light is dependent on the degree of the roughness of the surface in both vertical and lateral directions. In optical scatterometry, the surface properties are deduced from the scattered far-field light distribution. No measurement of the phase information is required. Inverse solution from the near-field distribution is also feasible, but recuires sophisticated and expensive equipment.[1]

Quantitative analysis of random rough surfaces by optical scatterometry requires mathematical models. In general, the inverse scattering problem has no analytical solution. Forward scattering may be simulated, in some cases, by approximate methods, but usually numerical methods based on the exact electromagnetic scattering theory are used. Solving the inverse problem is a demanding task since there exists a highly nonlinear relationship between the surface parameters and the scattering data. Here we apply the neural network theory for estimating this relationship. This approach has been suggested for surfaces with slight degree of rms-roughness, up to one wavelength.[2] We have

Further author information: (Send correspondence to Jyrki Saarinen)
Jyrki Saarinen: E-mail: jyrki.saarinen@heptagon.fi
Ilkka Kallioniemi: Email: Ilkka.Kallioniemi@hut.fi

Part of the SPIE Conference on Advanced Photonic Sensors and Applications
Singapore • November–December 1999
SPIE Vol. 3897 • 0277-786X/99/$10.00

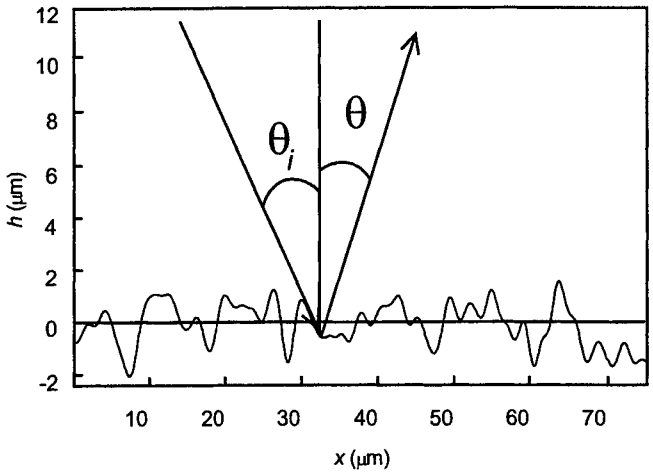


Figure 1. One-dimensional random rough surface with the height of the profile $h(x)$. The surface is characterized with two statistical parameters: the root-means-square (rms) height σ and the correlation length Λ. An incident plane wave impinges on the surface at the angle θ_i, and the angle of scattering is denoted $\theta \in [-\pi/2, \pi/2]$. The angles are measured with respect to the normal of the mean plane of the surface $h = 0$, which is the horizontal line.

already successfully demonstrated that neural network as a statistical estimator for optical scatterometry is an efficient tool for characterizing periodic microstructures.[3,4]

In this paper, we describe a method for generating numerically random surfaces characterized with their degree of roughness: root-mean-square (rms) amplitude of roughness and lateral correlation length. Here we are mainly interested in the most demanding region of the rms amplitude in the so-called resonance domain, which corresponds to height fluctuations and correlations up to 5 wavelengths of light. We simulate scattering from rough surfaces with an exact electromagnetic theory under the condition that the surfaces are one-dimensional perfect conductors. Randomness is achieved with statistical analysis of several realizations of theoretical rough surfaces from the same rms amplitude and correlation length values. The neural network model, which is here a self-organizing map (SOM), is first trained and calibrated with representative surface parameter and scattering data pairs. In the characterization stage, the neural networks classifies unknown surfaces according their surface parameters, by using measured intensity distributions as input to the network.

2. GENERATION OF THEORETICAL RANDOM ROUGH SURFACES

We generate theoretical random rough surfaces as realizations of spatially stochastic processes. For simplicity, we restrict ourselves to one-dimensional surfaces (Fig. 1). The values of height of the surface, $h(x)$, follow a Gaussian probability density function $p(h)$. The degree of roughness is characterized by the root-mean-square (rms) value σ of $h(x)$, which is averaged over the density $p(h)$. In mathematical formulation,

$$\sigma = \sqrt{\langle h^2(x) \rangle - \langle h(x) \rangle^2} = \sqrt{\langle h^2(x) \rangle}. \tag{1}$$

Here $\langle h(x) \rangle$ is chosen to a value zero. The two-point autocorrelation function between points x_1 and x_2 is defined as

$$C(x_1, x_2) = \frac{\langle h(x_1) h(x_2) \rangle}{\sigma^2}. \tag{2}$$

The random rough surface is supposed to be stationary, that is

$$C(x_1, x_2) = C(x_1 - x_2). \tag{3}$$

Furthermore, the autocorrelation function is assumed to be Gaussian, that is

$$C(x_1, x_2) = \exp\left(\frac{|x_1 - x_2|^2}{\Lambda^2}\right) \tag{4}$$

where Λ is the correlation length.

For generating theoretical surfaces, random numbers are firstly generated from a Gaussian distribution with zero mean and standard deviation σ. Secondly, the resulting random vector is convoluted with the Gaussian correlation function (4). This results in a surface sample which possesses the required statistical properties Λ and σ to a high degree of accuracy. If the surface consisted of infinitely many sample points, it would meet these statistical properties exactly. For each value of Λ and σ, 200 realizations of the surface are used with each random vector consisting of approximately $500 - 700$ surface points. The optimal sampling interval of the surface depends slightly on the value of Λ and σ since the surface profile should include at least ten structural asperities, but is typically of the order of 1/10th of the wavelength.[5]

3. NUMERICAL ANALYSIS OF RANDOM ROUGH SURFACES

Several models have been developed for evaluating the scattering from random rough surfaces, i.e., solving the forward scattering problem. The solution of the forward problem is analytically demanding, and thus various numerical techniques are required for different values of surface parameters σ and Λ.

The most commonly used methods are the Kirchhoff method and the perturbational methods.[6] The latter is restricted to 'slight' roughness, that is almost smooth surfaces. It is therefore of limited use, though it is extensively used in practice beyond its theoretical limit of validity.

In the Kirchhoff method, the scattering from each point $[x, h(x)]$ at the surface is assumed to take place from a tangent plane drawn to that point. The Kirchhoff approximation, also called as geometrical optic or 'high-frequency' approximation, is valid when the ratio $\Lambda/\lambda \gg 1$. Here λ is the wavelength of incident light in vacuum.

We utilize here a method that is based on the electromagnetic extinction theorem boundary condition.[7] The electromagnetic field inside the scattering media can be obtained from a known incident field. As a limiting value, the field at the surface is deduced. The scattered field is subsequently solved with help of the surface field and a suitable Green's function. The resulting integral equations can be solved numerically by sampling the surface at a finite number of discrete points. No assumptions have to be made concerning either the degree of roughness σ or the correlation length Λ. The only assumption is that the surface is perfectly conducting although extensions to dielectrics and finitely conducting surfaces are possible. The exact numerical implementation used here follows the one introduced by Nieto-Vesperinas and Soto-Crespo.[5,8]

For an arbitrarily accurate solution, an infinitely long surface with an infinitesimal sampling interval should be used. This requirement is unrealistic, and thus the accuracy of the calculations is restricted by available computing capacity.

The validity of the algorithm for specific values of σ and Λ is determined by the requirement of the conservation of energy, that is

$$\int_{-\pi/2}^{\pi/2} \frac{I(\theta)}{P_{\text{in}}} \, \mathrm{d}\theta \approx 1, \tag{5}$$

where $I(\theta)$ is the mean scattered intensity to angle θ (cf. Fig. 1) averaged over 200 surface realizations and P_{in} is the power of incident TE-polarized radiation. The acceptance criterion is that the left hand side of Eq. (5) should be

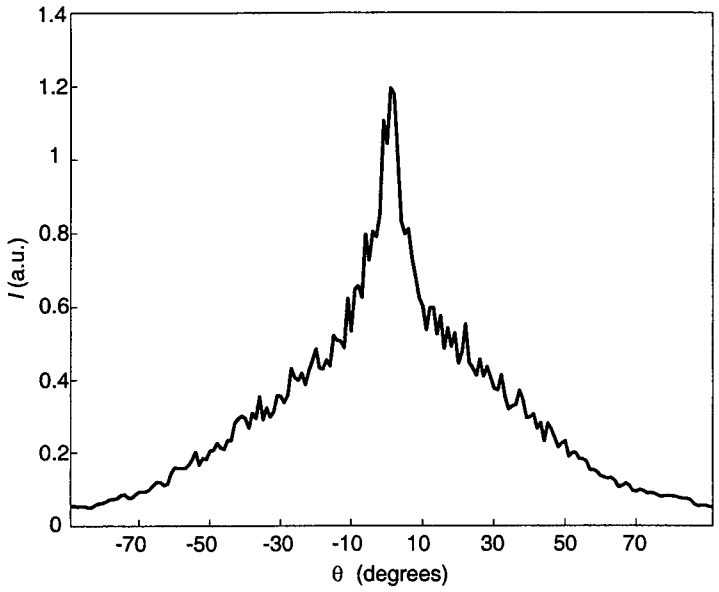


Figure 2. Example plot of the mean scattered intensity distribution $I(\theta)$ with statistical surface parameters $\Lambda = 3\lambda$ and $\sigma = 3.75\lambda$, where λ is the wavelength of light in vacuum. The incidence angle is 0°.

within ±2 % of the value one, except for the computationally most difficult values of σ and Λ, for which deviation of ±4 % is allowed.

For finding the region of parameter space, where the algorithm is valid, test intensity patterns are evaluated from a wide range of surface parameters Λ and σ. The correlation length is chosen to vary as $\Lambda = 0\lambda, \ldots, 5\lambda$, and the obtained maximum feasibility limits for σ can be read from Table 1. The calculations are done with a COMPAQ AlphaServer SG140 computer. Extending the analysis for parameters outside the achieved region would require intolerably large computation times. An example plot of $I(\theta)$ is shown in Fig. 2 with values $\Lambda = 3\lambda$ and $\sigma = 3.75\lambda$.

Table 1. Maximum feasibility limits for the root-means-square (rms) amplitude σ for the chosen values of the correlation length Λ.

$\Lambda(\lambda)$	0.5	1.0	1.5	2.0	2.5	3.0	3.5	4.0	4.5	5.0
$\sigma(\lambda)$	0.75	1.5	2.0	3.0	3.5	4.0	4.5	5.0	5.5	6.0

4. NEURAL NETWORKS

For characterizing periodic microstructures, the nonlinear relationship between the scattering data and the surface parameters can be determined by using, e.g., partial-least squares method[9] or principal component analysis.[10] Neural networks has been successfully applied to qualitative[11–13] and quantitative[3,4] analysis of the surface parameters.

For characterizing random rough surfaces, we use use a special kind of neural network model, the self-organizing-map (SOM).[14] The SOM is effective for the organization and classification of complex, multi-dimensional input data vectors. Here, the input data $\vec{I}$ consists of features extracted from the mean scattered far-field intensity pattern $I(\theta)$, which is calculated for a discrete set of the angles θ. After the SOM has organized the data patterns, the map

is subsequently used to classify unknown samples with respect to the statistical parameters Λ and σ. This specific utilization of the SOM is called associative estimation.[14]

We here consider an SOM structure that consists of nodes in two dimensions. Each node i is attached with a weight vector $\vec{W}_i$ such that $\dim \vec{W}_i = \dim \vec{I}$ where dim denotes dimensionality. The weights are adjusted iteratively by introducing calculated input vectors $\vec{I}^{(m)}$ into the network. For each intensity pattern, the best matching node b is chosen such that

$$b = \arg_{\min}[||\vec{I}^{(m)} - \vec{W}_i(m)||] \tag{6}$$

where the minimization is done over all indeces i and $||\dots||$ represents the Euclidian norm. The index m has been added to denote the iteration steps. The values of all weight vectors are further adjusted according to

$$\vec{W}_i(m+1) = \vec{W}_i(m) + N_{bi}(m)[\vec{I}^{(m)} - \vec{W}_i(m)]. \tag{7}$$

Here $N_{bi}(m)$ is the neighborhood function that may change during the iteration. This function defines the neighborhood of the best matching node b such that $N_{bi} \to 0$ as the distance $||\vec{W}_b(m) - \vec{W}_i(m)||$ increases. The exact form of $N_{bi}(m)$ varies but is often chosen Gaussian whose variance determines the 'elasticity' of the SOM network. Since the number of required iterations is typically considerably larger than the amount of available data vectors $\vec{I}^{(m)}$, each vector is introduced several times in random order. This process is terminated after the values of all the W_i stabilize. Some kind of ordering of the nodes is formed inside the two-dimensional SOM structure that best reflects the relations between the multi-dimensional input data patterns in $\vec{I}^{(m)}$.

After the organization process is completed, the SOM is calibrated by providing it with some input vectors calculated for suitable values of the statistical parameters σ and Λ. Here, suitable values are chosen to represent the classes into which the surfaces should be classified. At this stage, the weight vectors are not further adjusted but the representative values of σ and Λ are attached to the node that gives the closest matching response to the corresponding input pattern, as defined in Eq. (6). Thus some of the nodes in the SOM become labelled.

Once the calibration stage is finished, the SOM is used to classify new, unknown input patterns $\vec{I}$. The network assigns the pattern to the best matching node and the corresponding statistical parameters σ and Λ may be estimated visually from the calibrated map. However, it is possible to assign class labels also numerically. If the best matching node is not labelled, it is possible to find the second best matching node and so on, until a labelled node is found.

5. RESULTS

Two SOM networks of 10×10 and 17×24 nodes are trained to work in the accepted region of the parameter space (cf. Section 3). The training takes approximately 15000 iteration steps. Parameter combinations for the training are chosen randomly such that the entire accepted parameter region (cf. Table 1), is covered as uniformly as possible. The total number of combinations is about 450 and therefore each data vector has to be reused several times in random order for training. The relatively small number of samples is mainly due to long computation times of up to 45 CPU min required for some parts of the accepted parameter space, where both a small sampling interval and large number of sampling points are required.

The input data vectors $\vec{I}$ are obtained from the mean intensity values $I(\theta)$. The angle θ ranges from $-90°$ to $+90°$ with respect to the normal of the mean surface $h = 0$. The scattered intensities are first calculated with a step size of $\Delta\theta = 1°$, i.e., the raw data consists of 181 input parameters per sample. This number is reduced by linear principal component analysis. From this analysis, 6 most significant features are chosen to compose the final data vectors $\vec{I}$. The incident TE-polarized light impinges upon the surface at the angle of incidence $\theta_i = 0°$.

The SOMs are used to classify surfaces with respect to both Λ and σ. The classes are chosen from the accepted parameter range (cf. Table 1) with step sizes $\Delta\Lambda = 0.5\lambda$ and $\Delta\sigma = 0.5\lambda$ for the correlation length and roughness, respectively. The SOMs are calibrated with representative input data patterns which correspond to parameter values from the center of each class, thus resulting in structures where such nodes in the SOMs are labelled that best match the desired classes.

After calibration, 28 input vectors corresponding to randomly chosen sample values of σ and Λ are presented to the SOMs. Each sample m ($m = 1, \dots, 28$) is classified according to that labelled node which best matches the input data vector $\vec{I}^{(m)}$. This is not necessarily the best matching node in the sense of Eq. (6) since that may not

be labelled. Preliminary results from classification are summarized in Table 2. In this case, visual inspection of the SOM structure is not feasible option, since differences between input data are too subtle to be resolved this way. The first two columns give the actual values of Λ and σ for the chosen test data, and the last four columns give the label of the estimated class by using the SOM with 10×10 and 17×24 nodes. From Table 2 it is seen that for most cases the classification result deviates at most one class, corresponding to 0.5λ, from the correct values. Drastical misclassifications in some cases are probably due to too few training samples from the corresponding region in (Λ, σ) -parameter space. More optimized SOM classification will be published later elsewhere.

6. CONCLUSIONS

We have successfully demonstrated the use of neural networks for quantitative classification of random rough surfaces in the case of Gaussian surface statistics. A large range of statistical parameters, the rms-roughness σ and the correlation length Λ, was chosen and two neural network models were trained over the range. We emphasize that the method depends neither on the statistics used for describing the surface nor on the degree of roughness. Classification with better resolution for the parameters, $\Delta\sigma$ and $\Delta\Lambda$, could be achieved by calculating more data vectors and/or by using more angles of incidence. The latter would improve the uniqueness of the obtained solution, referring to the nonuniqueness of the inverse scattering problem.

A well-trained SOM should also respond correctly to noisy scattering data. To validate this property, noisy training samples can be formed by adding Gaussian noise with zero mean and variance of, e.g., approximately 5% relative to the intensity values $I(\theta)$ to the input data vectors. The noisy input data vectors are subsequently presented to the SOM. The results will be presented later elsewhere.

Besides the classification of random rough surfaces, more accurate predictions for σ and Λ may be required and there is no point in decreasing $\Delta\sigma$ and $\Delta\Lambda$ too much. Instead, a non-linear regression model based on the use of the multi-layer perceptron (MLP) neural network may be applied (cf. Refs. [3,4] for periodic microstructures). The results will be presented later elsewhere.

ACKNOWLEDGMENTS

We thank Professor Erkki Oja for fruitful discussion on neural networks. Financial support from the Technology Development Centre TEKES and the Academy of Finland are acknowledged. Ilkka Kallioniemi also acknowledges two grants from the Vilho, Yrjö, and Kalle Väisälä Foundation and a grant from the Jenny and Antti Wihuri Foundation.

REFERENCES

1. N. García and M. Nieto-Vesperinas, "Direct solution to the inverse scattering problem for surfaces from near-field intensities without phase retrieval," *Opt. Lett.* **9**, pp. 949–951, 1995.
2. K. Yoshitomi, A. Ishimaru, J.-N. Hwang, and J. S. Chen, "Surface roughness determination using spectral correlations of scattered intensities and an artificial neural network technique," *IEEE Trans. Ant. Prop.* **41**, pp. 499–502, 1993.
3. I. Kallioniemi, J. Saarinen, and E. Oja, "Optical scatterometry of subwavelength diffraction gratings: neural network approach," *Appl. Opt.* **37**, pp. 5830–5835, 1998.
4. I. Kallioniemi, J. Saarinen, and E. Oja, "Characterization of diffraction gratings in rigorous domain with optical scatterometry: hierarchical neural network," *Appl. Opt.* (in press).
5. J. M. Soto-Crespo and M. Nieto-Vesperinas, "Electromagnetic scattering from very rough random surfaces and deep reflection gratings," *J. Opt. Soc. Am. A* **6**, pp. 367–384, 1989.
6. J. A. Ogilvy, *Theory of Wave Scattering from Random Rough Surfaces*, Adam Hilger, Bristol, 1991.
7. D. N. Pattanayak and E. Wolf, "General form and a new interpretation of the ewald-oseen extinction theorem," *Opt. Comm.* **6**, pp. 217–220, 1972.
8. M. Nieto-Vesperinas and J. M. Soto-Crespo, "Monte carlo simulations for scattering of electromagnetic waves from perfectly conductive random rough surfaces," *Opt. Lett.* **12**, pp. 979–981, 1987.
9. S. S. H. Naqvi, R. H. Krukar, J. R. McNeil, J. E. Franke, T. M. Niemczyk, D. M. Haaland, R. A. Gottscho, and A. Kornblit, "Etch depth estimation of large-period silicon gratings with multivariate calibration of rigorously simulated diffraction profiles," *J. Opt. Soc. Am. A* **11**, pp. 2485–2493, 1994.

10. K. P. Giapis, R. A. Gottscho, L. A. Clark, J. B. Kruskal, D. Lambert, A. Kornblit, and D. Sinatore, "Use of light scattering in characterizing reactively ion etched profiles," *J. Vac. Sci. Technol. A* **9**, pp. 664–668, 1991.
11. R. Krukar, A. Kornblit, L. A. Clark, J. Kruskal, D. Lambert, E. A. Reitman, and R. A. Gottscho, "Reactive ion etching profile and depth characterization using statistical and neural network analysis of light scattering data," *J. Appl. Phys.* **74**, pp. 3698–3706, 1993.
12. N. George and S.-G. Wang, "Neural networks applied to diffraction-pattern sampling," *Appl. Opt.* **33**, pp. 3127–3134, 1994.
13. D. J. Search, C. A. Hobson, J. T. Atkinson, and J. D. Pearson, "Diffraction pattern analysis for automatic defect classification in manufactured electronic assemblies," in *Machine Vision Applications in Industrial Inspection II*, B. M. Dawson, S. S. Wilson, and F. Y. Wu, eds., *Proc. SPIE* **2183**, pp. 170–179, 1994.
14. T. Kohonen, *Self-Organizing Maps*, Springer, Berlin, 2nd ed., 1997.

Table 2. Theoretical values and predictions from 10×10 and 17×24 self-organizing map (SOM) for several combination of the rms-height σ and the correlation length Λ of random rough surfaces. The exactly correct classification is marked with bold.

Theoretical values		10×10 SOM		17×24 SOM	
Λ	σ	Λ	σ	Λ	σ
3,47	0,62	3,0	**0,5**	3,0	0,1
4,24	2,13	5,0	**2,0**	**4,0**	**2,0**
3,90	3,68	5,0	5,0	5,0	4,5
3,25	1,43	**3,0**	**1,5**	4,5	**1,5**
4,72	0,29	1,5	**0,1**	1,5	**0,1**
1,74	1,95	**2,0**	2,5	**2,0**	2,5
0,98	0,86	**1,0**	**1,0**	**1,0**	**1,0**
3,53	2,43	**3,5**	**2,5**	3,0	2,0
4,48	1,22	1,0	0,5	5,0	1,5
3,44	2,26	5,0	3,0	4,0	**2,5**
0,66	0,84	**1,0**	**1,0**	**1,0**	**1,0**
1,53	1,81	**1,5**	**2,0**	1,0	**2,0**
2,23	1,80	3,0	3,0	**2,0**	1,5
2,05	1,16	**2,0**	**1,0**	2,5	1,5
2,86	2,71	4,0	4,5	**3,0**	**2,5**
2,85	3,97	2,0	2,5	**3,0**	3.5
4,49	5,96	**4,5**	5,0	5,5	**6,0**
2,75	4,02	2,0	2,5	2,0	3,0
1,80	0,02	4,5	**0,1**	5,0	**0,1**
1,65	0,50	1,0	**0,5**	**1,5**	**0,5**
3,50	1,36	2,5	1,0	4,5	**1,5**
0,63	0,51	**0,5**	**0,5**	**0,5**	**0,5**
1,10	1,24	1,5	2,0	**1,0**	**1,0**
3,72	1,62	3,0	**1,5**	**3,5**	**1,5**
3,75	0,61	**3,5**	**0,5**	3,0	**0,5**
2,29	1,05	**2,5**	**1,0**	2,0	**1,0**
4,88	4,15	**5,0**	5,0	**5,0**	**4,0**
2,72	2,47	4,5	4,5	3,5	3,5

Bragg Optical Displacement Sensor

K. Nasrullah , AIEEE
Department of Electrical and Electronic Engineering,
University Putra Malaysia, 43400 UPM Serdang, Selangor.
(khan@eng.upm.edu.my)

ABSTRACT

A novel optical displacement sensor based on Bragg effect is described. A green light beam was used to excite a Rh6G dye solution in Shank type geometry. The beam was splitted up by a dielectric mirror and the two equal intensity vertically polarized light beams were folded by two mirrors to interfere in the dye cell. The lasing wavelength was found to be a precise function of the half angle between the beams. The vibrating body was attached to one of the folding mirror. In normal operation the lasing line wavelength was unchanged but during mirror vibration the output wavelength changed at a rate of 112.5 Å/Deg. The half of the amplitude of vibration corresponding to this 1° was 87μm. The sensor resolution was about 1.3 Å/μm. This ultrasmall displacement sensor can be used to determine the frequency of vibration or simple displacements in ultraprecise applications.

Keywords: Displacement sensors, vibration sensors, optical sensors and precision sensing.

1. INTRODUCTION

The operation of Bragg displacement sensor is based on the principle of distributed feedback dye laser (DFDL). The DFDL lasing principle was firstly reported by Derkacheva and Krymova (Russia) in 1969[1] during their four-photon resonant parametric interaction experiment. However, Kogelnik and Shank successfully demonstrated the first distributed feedback laser in February 1972[2]. One year later they reported the coupled wave theory that describes the principle of DFB laser[3].

The selective lasing wavelength is highly dependent on the stability of folding mirrors. If one of the mirrors is made part of a bulk system, under investigation for precise movements, then the net lasing wavelength can be made to vary under micro-displacements.

In distributed feedback dye lasers the positive feedback is provided by the distributed gain grating in the dye solution caused by two coherent beams interfering with each other at angle $2\theta_H$. The DFB laser action occurs if the periodicity Λ of the dynamic grating satisfies Bragg scattering condition given by

$$\Lambda = m\frac{\lambda_o}{2n} \tag{1}$$

Alternatively the periodicity Λ may also be given by

$$\Lambda = \frac{\lambda_p}{2\sin(\theta_H)} \tag{2}$$

If λ_p is the excitation laser wavelength, n refractive index of dye solution then at mth Bragg index comparison of (1) and (2) gives

Part of the SPIE Conference on Advanced Photonic Sensors and Applications
Singapore • November–December 1999
SPIE Vol. 3897 • 0277-786X/99/$10.00

$$\lambda_o = \frac{n.\lambda_p}{m\sin(\theta_H)} \tag{3}$$

The principle of distributed feedback lasers is based on coupled wave theory[4-8]. A DFB structure with or without external feedback is shown in Figure 1.

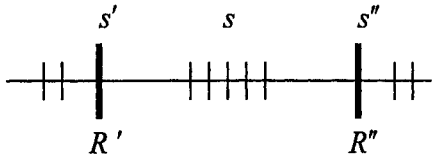


Figure 1 A generalized DFB structure

The primed s and R letters refer to the internal or external coupling strength and reflectivity. DFB gas, thin film, semiconductor, liquids and xray lasers have already been reported[5-7]. The general equation describing laser action in DFB structure shown above may be given by [9]

$$R'\,\mathrm{R}''\mathrm{e}\,xp\,(-\xi L + i\theta) = 1 \tag{4}$$

Where $R'and\,R''$ are represent end reflectivities and θ and ξ are given by

$$\theta = \Gamma + \mathrm{iK} \tag{5}$$
$$\xi = \mathrm{KL} \tag{6}$$

Here L is cavity length, Γ is effective propagation constant and parameter K is a free space wavenumber given by

$$\mathrm{K} = \frac{2\pi}{\Lambda} \tag{7}$$

The phase angle θ is related to forward and backward coupling strength s by

$$\mathbf{s_f} = \mathrm{s}\exp(\mathrm{i}\theta) \tag{8}$$
$$\mathbf{s_b} = \mathrm{s}\exp(-\mathrm{i}\theta) \tag{9}$$

In case of no external feedback the coefficients $s', s'', R'and\,R'' = 0$. However, the coupling strength s can not be zero. So the threshold oscillation condition becomes

$$-\,s^2\exp(\,2\xi L) = 1 \tag{10}$$

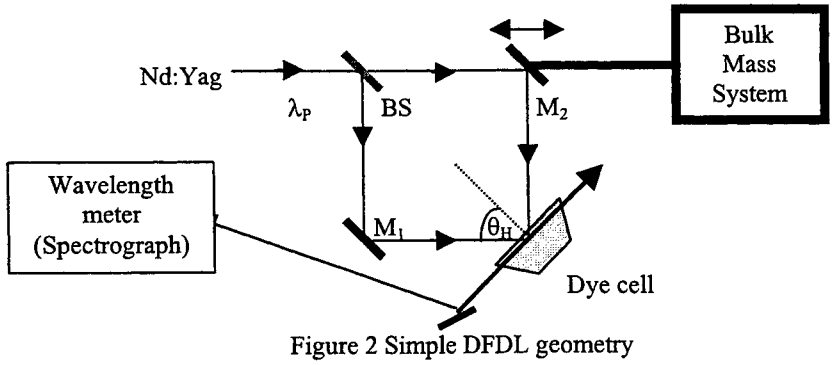
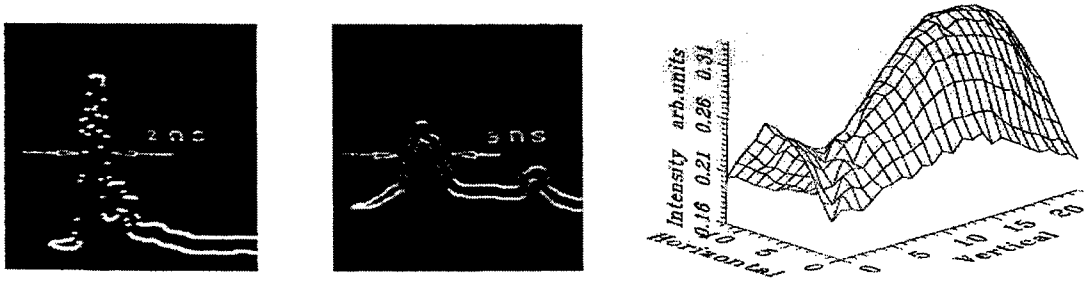


Figure 2 Simple DFDL geometry

Distributed feedback in the DFB structure may be provided by the gain grating caused by interference pattern of two beams satisfying Bragg scattering condition in the dye cell. There are a number of configurations that can be implemented to cause DFB laser action in dye solution. Typical arrangements include Shank type geometry, prism-dye cell geometry, reflection or transmission grating geometry and telescopic geometry[8-14]. A typical representative arrangement is shown in Figure 2.

The dye cell in Shank type geometry in present experiment was excited by 2nd harmonic of a passively Q.Swirched and modelocked Nd;YAG laser. The pump power was gradually increased after alignment until the DFB laser oscillator started oscillation in oval spot. The pump laser and DFB laser oscilloscope profiles recorded by a PIN diode were very close to Gaussian as shown in Figure 3 (a-b). The The pump laser burn patterns were excellently round close to the oscillator but gradually faded when measured for away at distance of some 10 meters. This fading of burn patterns was due to absorption of Nd:YAG IR (1.06 μm) lines in air. The intensity decayed exponentially away from the oscillator exit mirror. The burn patterns were scanned by a microdensitometer to find the decay rate. The burn patterns profiles obtained scanning the recorded spots on Polaroid film are shown in Figure 3(c).

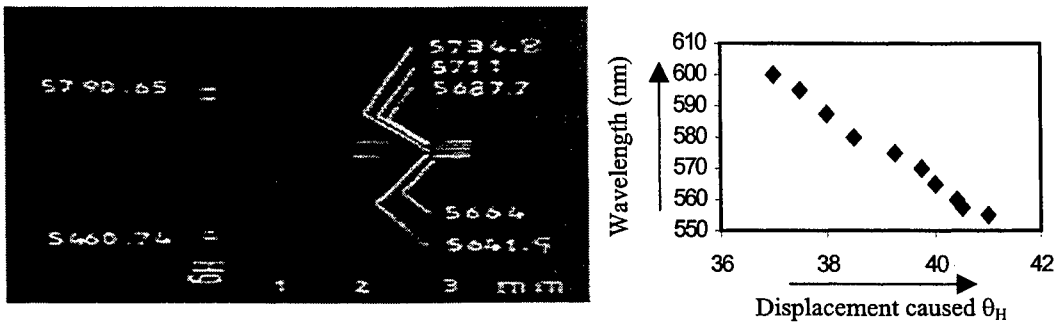


a- Nd:YAG out profile b- DFDL Output profile c- Burn pattern profile

Figure 3 Experimentally measured initial output of laser

2. DISPLACEMENT SENSING

The Rh6G laser dye based DFDL under this study was tuned by varying half angles through rotation of folding mirror attached to the bulk body. Other parameters like refractive index and Bragg index were kept constant. The half angle was increased from 37° to 41° and corresponding output lasing lines varied from 600 nm to 565 nm respectively. The results are shown in Figure 4. The results show that in Shank type geometry the average tunability index is 15 nm/Deg.



a- Spectrograph recorded variation b- Plot of wavelength variation by θ_H.

Figure 4 Variation of DFDL wavelength by micro-displacements of bulk body

The DFDL output pulse lengths were measured by operating the laser at different grating lengths controlled by micro-displacements caused by the bulk body. The pulse lengths were found temporally lengthened by increase in grating lengths. The pulse duration for similar cavity lengths were also calculated theoretically using Gales and Szatmari ideas on ultrashort pulse generation[13-14]. Calculated and measured results are shown in Figure 5.

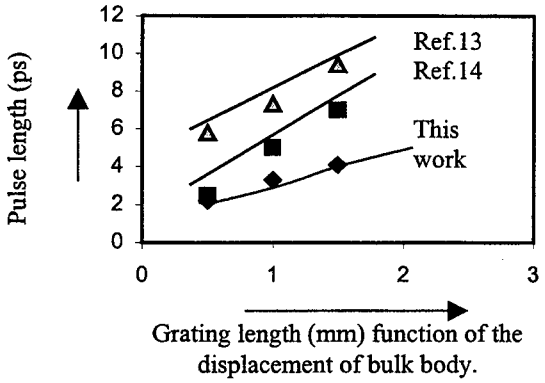


Figure 5 Variation of DFB laser output pulse as a function of grating length in dye solution.

The measured pulse lengths were too short to be measured accurately by streak camera Hadland(675). However, if the time resolution of the camera (5 ps) is deconvolved then the results fall within the theoretical boundaries. This technique can be implemented to display microdisplacements of a bulk body attached to the folding mirror. The displacements can be calibrated for display. Some more work might be needed.

3. CONCLUSIONS

Principle of microdisplacement sensor based on Bragg grating periodicity and interference overlap length modulation has been studied experimentally. The displacement length was found directly proportional to the change of wavelength as well as pulse length. In normal operation the lasing line wavelength was unchanged but during mirror vibration the output wavelength changed at a rate of 112.5 Å/Deg. The half of the amplitude of vibration corresponding to this 1° was 87μm. The sensor resolution was about 1.3 Å/μm.

ACKNOWLEDGEMENTS

Author acknowledges the assistance and efforts of technicians in UPM Laboratory and Miss Rohana for typing this manuscript at a very short notice.

REFERENCES

1- L.D.Derkacheva and A.I.Krymova, Zhet.Pis, Red, 9, No.10, PP.565-566, May 19969.
2- H.Kogelnik and C.V.Shank, App.Phys, Lett., Vol.18, No.4, PP.153-154, February 1971.
3 H.Kogelnik and C.V.Shank,J.App.Phys, Vol.43, No.5, PP.2327-2330, May 1972.
4 E. Affolter and F.K Kneubuhl Phys. Lett. Vol-74A, No.6, PP.407-409, 1979.
5 M.Ohtsu and C.H. Shin et al, Paper No. QThL, CLEO 1991.
6 I.V. Cheremiskin and T.K Chekhlova, Sov. J. Quant. Electron., Vol.4, No.3, PP.387-390, 1974.
7 G.A. Lyakhoc, Sov. J. Quant. Electron., Vol.6, No.4, PP.457-460, 1976.
8 M. Steyer and F.P Schafer et al,App. Phys B, Vol.50, PP-265-267, 1990.
9 Shyh Wang, IEEE. J. Quant. Electron., Vol.QE-10, No.4, PP.415-420, 1974.
10 S.Chandra and N. Takenchi, App. Phys. Lett., Vol. 21, No.4, PP. 144-148, 1972.
11 V.I Vashchuk and E.I. Zabello et al, Sov. J. Quant. Electron., Vol..8, No.7, PP-859-862, 1978.
12 J.Jasny,Optics. Communication., Vol.53, No.4, PP.238-241, March 1985.
13 G.M.Gales, App. Phys. B, Vol.44, PP.224-228, 1987
14 S.Szatmai and F. P. Schafer, App.Phys.B, Vol.46, PP. 305-307, 1988.

SESSION 15

Novel Sensor Designs II

Invited Paper

Laser-Diode Interferometer; Practice and Application

Yukihiro Ishii and Ribun Onodera

Department of Electronics, University of Industrial Technology, Sagamihara

4-1-1 Hashimotodai, Sagamihara, Kanagawa 229-1196, Japan

Phone: +81-42-763-9161, Fax: +81-42-763-9168, e-mail: yukidai@uitec.ac.jp

ABSTRACT

Laser diodes (LDs) have been served as light sources of a phase-measuring interferometer through the wavelength tunability of LDs by controlling their currents. Laser-diode interferometers based on a heterodyne technique are reviewed with their practice and application. A two-wavelength laser-diode interferometer is elaborately demonstrated with current control of dual LDs in opposite directions to extend the range of interferometric measurements. The wavelength is controlled by the laser injection current and is stepwise or rampwise changed to introduce a time-varying phase difference between the two beams of an interferometer with unbalanced optical path lengths. A feedback interferometer with electronics is used to lock the interferometer on a preset phase shift to ensure the measurement accuracy. The interferometric techniques with laser diodes have been applied to an optical radar such as a single detection of multiplexed phase objects and to a distance measurement.

Keywords: laser diode, wavelength tunability, phase measurement, phase-shifting method, heterodyne interferometer, feedback control, two-wavelength interferometer, optical radar, distance measurement

1. INTRODUCTION

The great progress in phase-measuring interferometry can be attributed to the development of computers and photodetectors as well as to electronic phase-extraction techniques. The sensitivity and speed of phase measurements have been increased remarkably by heterodyne and phase-shifting interferometric methods,[1] and measurement accuracy is the order of 1/100 wavelength.[2] A piezoelectric transducer (PZT) is commonly used as a phase shifter, although it has disadvantages of hysteresis and mechanical vibration.

Laser diodes (LDs) have been proved to be useful light sources in optical interferometry[3,4] because they are compact size, high efficiency, single-mode operation and frequency tunability: current changes produce frequency modulation by altering the refractive index of the active region and consequently the optical path length of the laser cavity. Thus the LD can be driven as an easy voltage-controlled oscillator in the range of a few giga-hertz frequencies.

Part of the SPIE Conference on Advanced Photonic Sensors and Applications
Singapore • November–December 1999
SPIE Vol. 3897 • 0277-786X/99/$10.00

This paper presents a laser-diode phase-measuring interferometer and its practical technique and application. The capability of direct frequency modulation in LD has been used by changing the currents. The wavelength is stepwise or rampwise changed to introduce a time-varying phase difference between the two beams of an unbalanced interferometer. In particular, a two-wavelength LD interferometer has been elaborately constructed to extend the range of interferometric phase measurement by using a long wavelength synthesized by two wavelengths together with a feedback control to lock the interferometer on a phase-shift condition by changing the bias and modulation currents of LDs. The frequency-chirped LD interferometer can be applied to an optical radar such as an optical frequency-domain multiplex imaging. Selective extraction of multiplex images can be performed by electronic tuning of multiple interference beat signals. The LD phase-shifting interferometer has been also applied to a distance measurement with a wide tunable laser diode.

2. PRACTICE IN A LASER-DIODE INTERFEROMETER

An unbalanced Twyman-Green interferometer is shown in Fig. 1 with a stepwise- or ramp-modulated LD source. The ramp waveform has a period of T and an amplitude of i_m in Fig. 2(a). A linear change i_m in the LD current generates a chirped-frequency variation $\Delta\omega$, producing the relation $\Delta\omega = 2\pi\beta i_m$ with the current-tuning ratio β in Fig. 2(b). The laser light is sent into the interferometer equipped with a polarizing beam splitter (PBS) and quarter-wave ($\lambda/4$) plates. This configuration is used to reduce feedback from the coupling optics to the LD. The lights returning from the object mirror and the reference mirror interfere with each other through an analyzer. The intensity distribution of the interference pattern in a photodetector is

$$I(x,y)=I_M(x,y)\,[\,1+\gamma(x,y)\cos\{\phi(x,y)\}], \qquad (1)$$

where I_M is the bias intensity, γ is the visibility and ϕ is the tested phase. The phase ϕ is given by

$$\phi(x,y)=2\pi\{2w(x,y)+l\}/\lambda, \qquad (2)$$

where λ is the LD wavelength, w is the profile of the test object, and l is the optical path difference (OPD) between the two arms of the interferometer. When the wavelength is changed from λ to $\lambda+\Delta\lambda$ by linearly varying the LD current by i_m such that $\Delta\lambda=\alpha i_m$ ($\alpha = \beta\lambda^2/c$) in the light speed c, the phase ϕ of the test path length w reduces mathematically to

$$\phi + \Delta\phi = \frac{2\pi}{\lambda}(2w + l) - \left(\frac{2\pi l\Delta\lambda}{\lambda^2} + \frac{2\pi 2w\Delta\lambda}{\lambda^2}\right). \qquad (3)$$

The intensity by shifting the wavelength of LD gives

$$I(x,y) = I_M(x,y)\left[1 + \gamma(x,y)\cos\left\{\phi(x,y) + \Delta\phi\right\}\right], \qquad (4)$$

by substituting Eq. (3) into Eq. (1). The modulation term $\Delta\phi$ in Eq. (4) remains, ignoring the last term in Eq. (3) since l »w. In the temporal heterodyne scheme, the modulation $\Delta\phi$ can be rewritten as $\Delta\phi=2\pi l\Delta\lambda t/(\lambda^2 T)=\omega_b t$ where the interference beat frequency ω_b is given by $\omega_b=2\pi\beta i_m l/(Tc)$ as shown in Fig. 2(b).

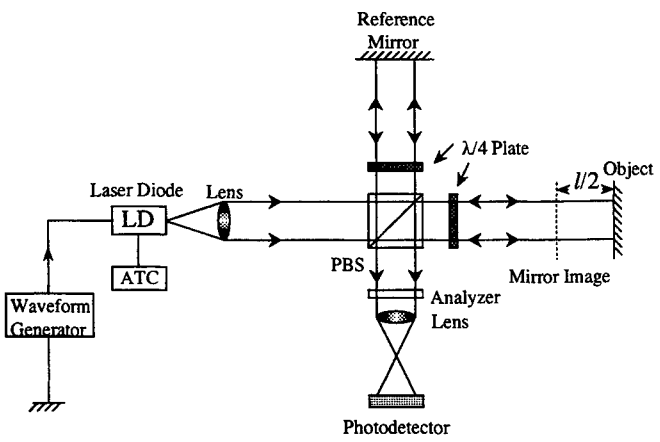


Fig. 1. A laser-diode phase-measuring interferometer with an unbalanced optical path length *l*..

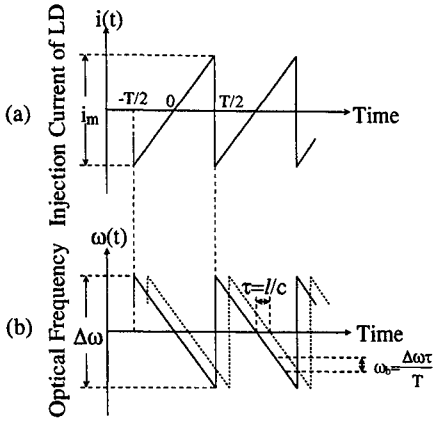


Fig. 2. (a) The variation of LD injection current and (b) the optical frequency as a function of time.

This method is based on the fact that the phase difference in an "unbalanced" interferometer is proportional to the product of the OPD and the LD temporal frequency.

If the current and its resultant wavelength change stepwise through a range of one wavelength, the *j*th phase shift δ_j corresponds to the phase-modulation term of Eq. (3), giving

$$\delta_j = j\frac{2\pi l\alpha i_m}{\lambda^2} \qquad (j=1,...,N), \qquad (5)$$

where N is the number of samples in the interferogram and in this case, i_m is the minimum current increment to give one phase step. The *j*th sampled intensity has the same form of Eq. (4) when δ_j is written in place of $\Delta\phi$. The data required for a phase-shifting interferometry (PSI)[5] are a series of interferograms recorded with different phase shifts. By using the *j*th stepped intensity $I_j(x,y,\delta_j)$ and the measurements I_j^*, the least-squares criterion[6] is minimum, and the phase ϕ including the object w(x,y) is extracted by using a conventional four-step technique[1] with four intensity readings I_j^* (j=1,...,4),

$$\phi(x,y) = \tan^{-1}\left(\frac{I_2^* - I_4^*}{I_1^* - I_3^*}\right), \qquad (6)$$

where the phase shifts δ_j are arranged as 0, $\pi/2$, π, and $3\pi/2$ for the case of N=4.[7,8] Measurement results of testing a convex lens whose arrangement is shown in Fig. 1. An LD source (HL 7801) is operated at a wavelength λ of 788 nm. By setting $\delta_4 = 2\pi$ in Eq. (5), an OPD *l* and the current increment i_m are found to be 26.6 mm and 1.25 mA for α=0.0046 nm/mA, respectively. The phase-extraction algorithm in Eq. (6) can be performed by a computer with 8-bit digital data and 128×128 sample points. Measured phase *w* is shown in Fig. 3(a), exhibiting 2π phase discontinuities. By correcting the 2π phase gaps, the spherical aberration is depicted in Fig. 3(b).

Next, we describe a phase-extraction routine that is insensitive to changes in laser power associated with current variations.[9-11] Measurement accuracy is degraded by intensity modulation of the interferogram induced by a laser-power

change. One solution is to electrically normalize the intensity of the interferogram with LD output power.[3] The normalization is followed by a numerical calculation.[7] The iterative method to find both the phase ϕ and the phase shift δ_j is also one solution.[12] An intensity stabilization of the LD source has developed that utilizes a feedback system with a superluminescent diode as an external light controller.[13,14] Inasmuch as the output power of the LD is proportional to the current increment, the bias intensity I_M associated with the phase-shift change is represented by $I_M(x,y) = I_M^0(x,y) + 2\delta_j \Delta I(x,y)/\pi$ where I_M^0 is the bias intensity under no current modulation and therefore the phase shift is zero, and ΔI is the intensity change by current change i_m to yield a phase-shift interval of $\pi/2$. As the current changes equally in six steps, the *j*th sampled intensity I_j is written [Eq. (4)] as

$$I_j(x,y) = \left(I_M^0 + 2\delta_j \Delta I/\pi\right)\left[1 + \gamma(x,y)\cos\left\{\phi(x,y) - \delta_j\right\}\right] . \qquad (7)$$

To produce an actual phase ϕ and six measurements I_j^* taken at known phase shifts $\delta_j = \pi(j-7/2)/2$ because of six unknown terms in Eq. (7), the phase ϕ gives

$$\phi(x,y) = \tan^{-1}\left(\frac{3I_1^* - 5I_2^* + 5I_5^* - 3I_6^*}{-I_1^* - 3I_2^* + 4I_3^* + 4I_4^* - 3I_5^* - I_6^*}\right), \qquad (8)$$

by estimating the least-squares solution which is a phase-extraction routine insensitive to changes in laser-diode power. A variety of methods to derive this kind of algorithm has been reported.[15-17]

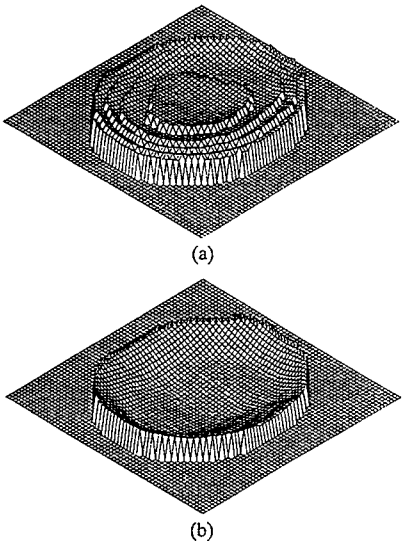


Fig. 3. (a) 3-D plot of discontinuous phase distribution. (b) 3-D plot of the spherical aberration corrected by 2π phase gaps in (a).

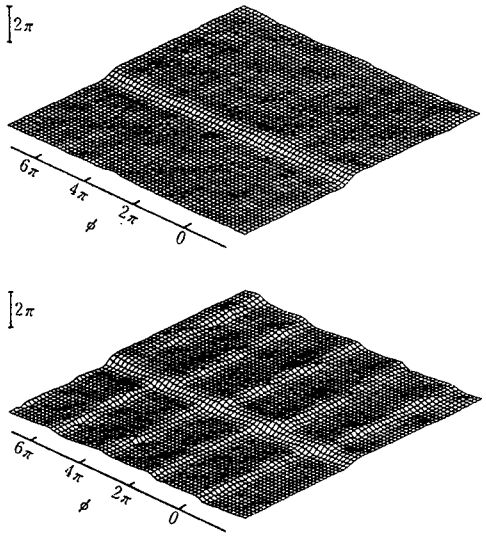


Fig. 4. 3-D plots of a step object calculated from the six-step phase-extraction algorithm (top) and the four-step one (bottom).

An experiment was performed to measure a chrome-coated step object. The single-frequency light source is an LD (TOLD9211) with an operating wavelength of λ=672 nm at a 43-mA current. Under an OPD of l=40 mm and a measured tuning rate α=0.01 nm/mA, the six phase shifts sampled symmetrically are produced by increasing the LD current whose range is associated with the power change of 0.6 mW. Figure 4 shows the measured phase calculated from the phase-extraction routine in Eq. (8) (top) and the measured one ϕ' that includes the phase error (bottom) from a four-step algorithm with intensity readings (I_1*,...,I_4*) in Eq. (6). The result at the top of the figure demonstrates freedom from the systematic error caused by LD power changes. The bottom figure shows that the periodic phase error has the same period as the interference fringe, i.e., 360° for γ=0.3 and ρ=0.1 $(\rho = \Delta I / I_M^0)$. The phase error is theoretically given by the relation of $\rho\left\{-\left(\sqrt{2}/\gamma\right)\sin\phi\right\}$.[9]

3. TWO-WAVELENGTH LASER-DIODE INTERFEROMETER

In interferometric measurements the phase distribution across the interferogram is measured modulo 2π. This problem limits the phase measurement range of PSI using a single wavelength. One solution is to use two shorter visible wavelengths to generate a longer synthetic wavelength, which makes it possible to measure distances much greater than the optical wavelength.[18] A synthetic wavelength $\Lambda=\lambda_1\lambda_2/|\lambda_1-\lambda_2|$ is obtained from two wavelengths λ_1 and λ_2. We demonstrate two-wavelength interferometry with dual wavelength-tunable LDs.[19,20]

The two-wavelength PSI where the intensity of the LD is assumed to be constant within a limited range of current alterations.[21] In an unbalanced Twyman-Green interferometer shown in Fig. 5, the intensity distributions of the interference pattern for the respective wavelength λ_k (k=1,2) are $I_k(x,y) = I_{Mk}(x,y)\left[1+\gamma_k(x,y)\cos\left\{\phi_k(x,y)\right\}\right]$, where I_{Mk} is the bias intensity, γ_k is the visibility and ϕ_k is the phase given by $\phi_k(x,y)=2\pi\{2w(x,y)+l\}/\lambda_k$. The current changes stepwise, the jth phase shift δ_{kj} is given by $2\delta_{kj} = j(2\pi l\alpha_k i_{m,k})/\lambda_k^2$ when the injection currents of the LDs change independently by $i_{m,k}$. The jth sampled intensity $a_j(x,y)$ of the two-wavelength interferogram is written by the incoherent sum of each intensity as $a_j(x,y)= I_1(x,y)+I_2(x,y)$ whose pattern is known as a moiré pattern. If $\gamma_1 I_{M1}=\gamma_2 I_{M2}$ is taken into consideration, the moiré pattern a_j becomes

$$\begin{aligned} a_j(x,y) &= I_{M1}(x,y)+I_{M2}(x,y) \\ &+2\gamma_1(x,y)\, I_{M1}(x,y)\cos\left[\left\{\phi_1(x,y)+\phi_2(x,y)\right\}/2-\delta_{1j}-\delta_{2j}\right]\cos\left[\left\{\phi_1(x,y)-\phi_2(x,y)\right\}/2-\delta_{1j}+\delta_{2j}\right]. \end{aligned} \quad (10)$$

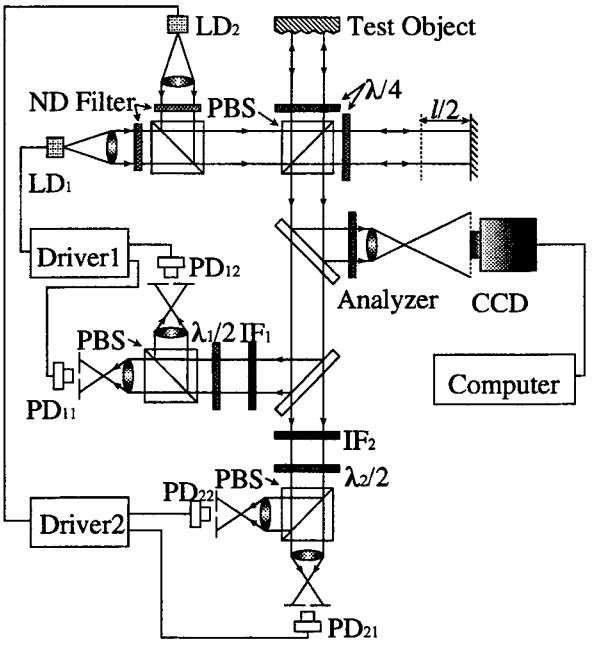


Fig. 5. A two-wavelength phase-shifting interferometer with dual LDs.

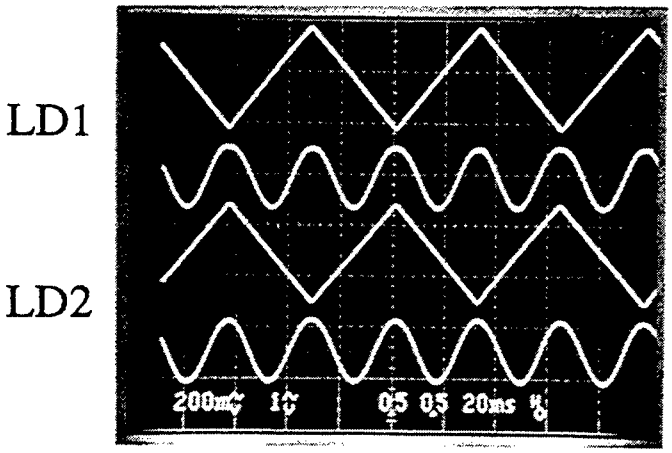


Fig. 6. Constant-intensity interference signals (the lower two traces) and mutually inverted triangular currents after calibration by feedback (the upper two traces) of dual LDs.

If the phase shifts such as $\delta_{1j} = -\delta_{2j}$ are arranged those phase-shift condition can be achieved by modulating LDs with mutually inverted ramped currents. Equation (10) becomes

$$a_j(x,y) = I_{M1}(x,y)+I_{M2}(x,y) \\ +2\gamma_1(x,y)I_{M1}(x,y)\cos\{\Psi(x,y)\}\cos\{\Phi(x,y)-2\delta_{1j}\}, \qquad (11)$$

where $\Phi=(\phi_1-\phi_2)/2=\pi(2w+l)/\Lambda$, $\Psi=(\phi_1+\phi_2)/2=\pi(2w+l)/\Gamma$ and Γ is the average wavelength given by $\Gamma=\lambda_1\lambda_2/(\lambda_1+\lambda_2)$. The first cosine term of Eq. (11) is independent of modulated phases, and the argument of the second cosine term is a function of the phase shift, which makes this moiré pattern applicable to the PSI technique, although the sensitivity depends on cosine Ψ function in Eq. (11). Using a_j of Eq. (11) and a PSI phase-extraction algorithm, one can calculate the phase Φ at a synthetic wavelength Λ with the use of the equation by replacing $\phi(x,y)$ and I_j^* in Eq. (6) with $\Phi(x,y)$ and a_j, respectively. It is interesting to note that the synthetic phase in Eq. (11) is produced by shifting equally the phases in opposite directions.

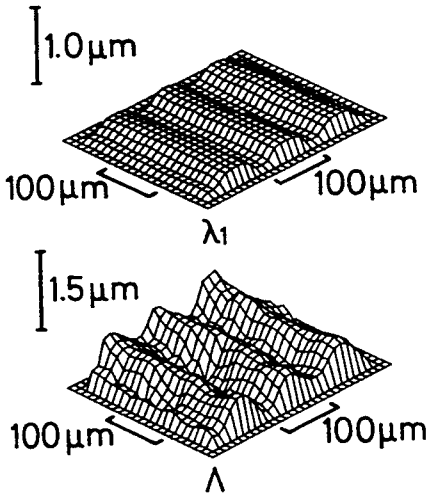


Fig. 7. Experimental results for single-wavelength (λ_1) (top) and two-wavelength (Λ) (bottom) measurements, respectively.

For measurements, two collimated beams emitted from LD_1 (λ_1=672 nm) and LD_2 (λ_2=788 nm) are incident on two neutral density filters to equalize the modulation intensities for both LDs. Two laser sources of the same linear polarization are used to illuminate simultaneously the interferometer arms with an OPD of l=40 mm. Half of the moiré pattern is measured with a CCD camera. The rest of the interference signal is transmitted through two interference filters IF_1 and IF_2 to two sets of PBSs with $\lambda/2$ plates to detect an interference signal for each wavelength. To ensure the measurement accuracy, the feedback electronics is used to lock the interferometric system on a preset phase-sfift relation $\delta_{1j} = -\delta_{2j}$ by adaptive frequency tunings of LDs. Each phase shift introduced by the bipolar triangular modulation current can be matched to the exact phase ranging from $-\pi$ to π during the half-period of a triangular current as illustrated in Fig. 6. The current variations in two LDs coincide with the corresponding interference signals. Thus, the modulated phases δ_1 and δ_2 are equally shifted in opposite directions by changing mutually inverted currents. The experimental layout of Fig. 5 has been used to test a diffraction grating with dual LDs that yield a synthetic wavelength Λ=4.6 μm from the wavelengths of λ_1=672 nm and λ_2=788 nm. The bias currents of LD_1 and LD_2 used in the experiment are 43.8 mA and 57.0 mA, respectively and the range of modulation currents are 1.26 mA and 1.80 mA, respectively. The current alterations of both LDs can satisfy the phase-shift condition. Two LD currents vary counterstepwise so that their resultant phases are shifted in opposite directions. The measured three-dimensional phase map is presented in the bottom part of Fig. 7. The measured height of grating is 0.46 μm which is in good agreement with the manufacturer's calibrated result of 0.45 μm. The top part of Fig. 7 shows the measurement made at a single wavelength (λ_1) with the PSI technique in which the measured height of 0.13 μm is subtracted half a wavelength from the actual height, causing an ambiguous measurement. To investigate the stability of phase measurement with the feedback electronics, the measurements of the same sample werерepeatedly for ten times. The good rms repeatability with an electronic calibration is obtained as ~Λ/78 compared with one of ~Λ/51 using a conventional method.[22]

4. LASER-DIODE INTERFEROMETER; APPLICATION

4.1. Selective extraction of multiple images by optical radar technique

A laser-diode interferometer with the direct frequency modulation of the LD has been applied to selective extraction of multiple phase images using an optical radar[23] technique,[24,25] and to multiplexing in a fiber interferometer[26]. Here, the laser-diode frequency is being chirping like the method as shown in Fig. 2.

In the Twyman-Green interferometer shown in Fig. 8, the frequency-modulated laser light is sent into the interferometer in which multiple phase objects O_i are located on various optical path differences (OPD's) l_i between a reference and objects arms, where the subscript i denotes an index from 1 to N and N is the total number of objects. The laser light returning from a reference mirror M_R and the objects mirrors, shown by solid line and dashed lines in Fig. 8, recombines at the beam splitter, producing time lags l_i/c. The interference beat frequencies ω_{bi} are generated as $\omega_b=\Delta\omega l_i$ /(Tc) those are proportional to the OPD's l_i. The electronic tuning makes it possible to extract selectively the interference beat signal assigned by the object on l_i. The intensity distribution of the multiple interferograms is given by the sum of the beat signals with the chirping frequency $\omega_c=\Delta\omega t/T$:

$$s(x,y,t) = A(x,y)+\sum_{i=1}^{N} B_i(x,y)\cos\left[\omega_{bi}t + \phi_i(x,y)\right], \qquad (12)$$

where ϕ_i is the phase distribution of ith object. When the interference beam on a dissector-camera plane is sampled at the frequency in synchronism with the beat frequency, the beat signals with temporal carrier frequencies[27] can be converted into interference fringes with spatial carrier frequencies. This serial-to-parallel conversion produces the spatial carrier frequency

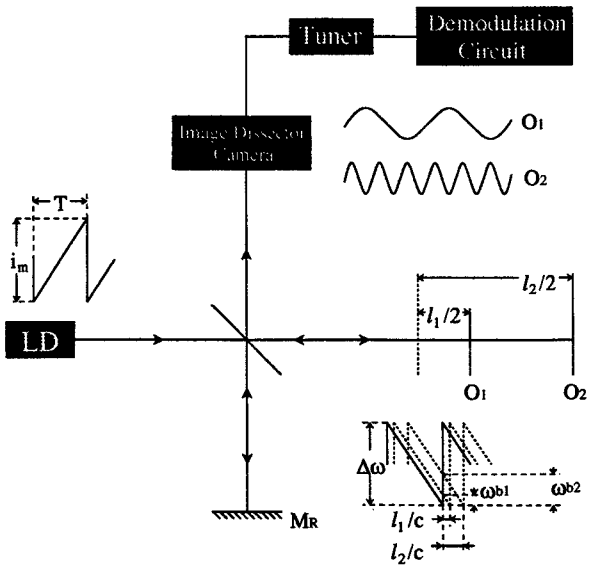


Fig. 8. Selective extraction of a beat signal by ramping LD current.

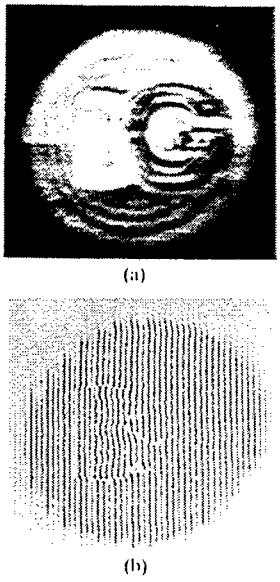


Fig. 9. Experimental results of the selective extraction of the phase object G in (b) compared with the conventional imaging in (a).

given by $\xi_i = \omega_{bi}/(f_s p)$ where f_s is the sampling frequency of the detector and p is the scan size. An image dissector camera is used that can detect the light intensity on any point of the detection plane addressed by x and y deflection signals. Selective extraction of nth interference fringe signal can be performed by tuning the center frequency of the tuner to ω_{bn} as shown in Fig. 8. Tuned interferogram I_n is spatially modulated by the phase distribution of the nth object and has no bias intensity. For the multiple images to be extracted separately, the difference $\Delta\omega_{bn}/2\pi$ between the nth and the (n-1)th beat frequencies must be greater than the bandwidth of the tuner 2B; $\Delta\omega_{bn}/2\pi \geq 2B$. This condition can be written as $\Delta l \geq l_n/Q$ where Δl is the path difference between the nth and the (n-1)th object arms and Q is a quality factor of the tuner and is given by $Q = \Delta\omega_{bn}/(2\pi \times 2B)$. The longitudinal resolution of the system is given by a minimum length of Δl; $\Delta l_{res} = l_n/Q$. The frequency-modulated laser light is coupled into an interferometer in Fig. 8 with two OPD's such as l_1=30 mm and l_2= 90 mm. Phase objects of the characters G and B are inserted into OPD's l_1 and l_2, respectively. The amplitude of the modulation current i_m is adjusted to 2.5 mA in order to make $\Delta\omega = 2\pi c/l_1$. Then the beat frequencies become $\omega_{b1} = 3\omega_{b2}$.

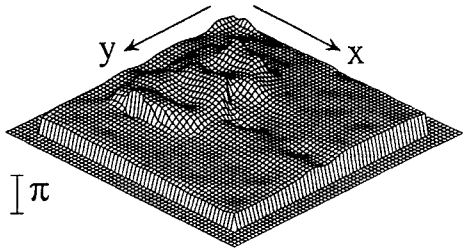


Fig. 10. Phase distribution obtained from the selective extraction by the Fourier-transform technique.

The interference pattern is focused onto the image dissector camera. The spatial carrier frequencies are $\xi_1 = 3.2\pi$ mm^{-1} and $\xi_2 = 9.6\pi$ mm^{-1}, which reduce to 1.6 and 4.8 fringes/mm, respectively. Since the detection area of the camera is 10 mm×10 mm, the measured interferograms have 16 and 48 carrier fringes located on the x axis that are phase-modulated by the characters G and B, respectively, as shown in Fig. 9(a). The longitudinal resolution is Δl_{res}=4.5 mm for l_2= 90 mm and Q=20 of a used tuner. The demodulation circuit has been used to synchronize the sampling frequency f_s of the dissector camera with the beat frequency $\omega_{bn}/2\pi$ by use of a clock of the A/D. This circuit produces the deflection signals of the camera for scanning and the sawtooth waveform for LD modulation.

Figure 9(a) shows the experimental result of the phase objects G and B in which the LD modulation and the electronic tuning are not performed. Selective extraction of the character B is illustrated in Fig. 9(b) that is modulated by 48 carrier fringes extracted successfully from the twofold interferograms. Figure 10 shows the phase distribution of the object B obtained from the extracted interferogram in Fig. 9(b) by the Fourier-transform method.[27]

This section demonstrates an optical radar technique for the selective extraction of 2-D phase objects.

4.2. Laser-diode interferometer for distance measurement

A laser-diode phase-shifting interferometer using a tunable external-cavity LD source is presented for distance measurement.[28] The phase shift produced by a wavelength tuning of the LD source[29] is measured by using Carré technique[30]

from four measurements with equal phase changes. The distances are calculated from the measured phase shifts and the emitting wavelengths of LD that are monitored by an optical spectrum analyzer.

In an unbalanced interferometer shown in Fig. 1, a phase shift in Eq. (3) can be introduced in the interferometer if the wavelength changes by $\Delta\lambda$. Thus the optical path difference l to be measured is determined by measuring the phase shift $\Delta\phi$, and the emitting wavelength λ and its wavelength change $\Delta\lambda$. Here we describe how to measure the amount of the phase shift by Carré technique. It is assumed that the phase is shifted by $\Delta\phi$ between consecutive measurements with four steps to yield four equations, $I_1 \sim I_4$ for substituting $\phi - 3\Delta\phi/2$, $\phi - \Delta\phi/2$, $\phi + \Delta\phi/2$ and $\phi + 3\Delta\phi/2$ instead of ϕ in Eq. (1). From four intensity measurements, the phase shift $\Delta\phi$ is calculated by

$$\Delta\phi = 2\tan^{-1}\left[\sqrt{\frac{3(I_2-I_3)-(I_1-I_4)}{(I_2-I_3)+(I_1-I_4)}}\right]. \qquad (13)$$

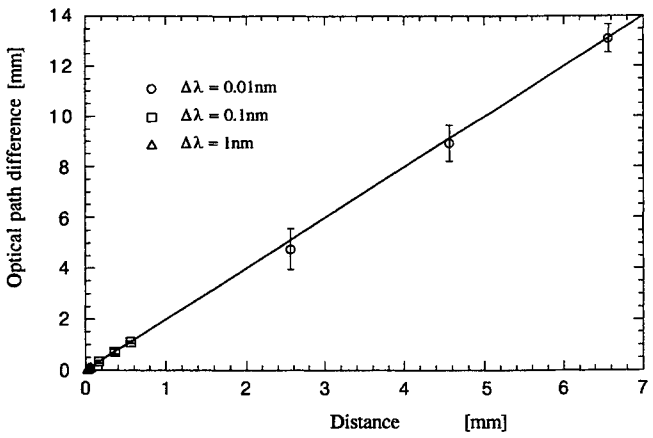


Fig. 11. Measured optical path difference vs distance for three measurernent regions $\Delta\lambda$.

Then the distance l is computed from Eq. (13). The light source is a tunable external-cavity LD. We can tune the laser wavelength by changing the angle of the tuning mirror. Its center wavelength is 780 nm with a tuning range of 20 nm. The output beam from the tunable LD is coupled into an unbalanced interferometer. The output power and the emitting wavelength are monitored during the measurement by the power monitor and the optical spectral analyzer with a resolution of 0.1 nm. Figure 11 shows the experimental result of the measured optical path difference as a function of the distance. The distance is given by the movement of the object mirror by a micrometer-driven stage in Fig. 1. As the object mirror is moved axially to nine discrete positions, the photodetector outputs are sampled at four-step wavelengths for each of the nine positions. Since the measurement accuracy is inversely proportional to the amount of the wavelength change $\Delta\lambda$, three wavelength changes, i.e., $\Delta\lambda$ = 1 nm, $\Delta\lambda$ = 0.1 nm, and $\Delta\lambda$ = 0.01 nm, are used for three measurernent regions. The experimental results denoting triangles, squares, and circles in Fig. 11 represent the average optical path lengths for ten times measurements and the error bars show the standard deviation. There is a good linear relationship between the measured optical path difference and the distance. A measurement range from 40 μm to 13 mm is experimentally demonstrated. The dynamic range in the measurement is given by $(\Delta\phi/2\pi)_{res}\lambda^2/\Delta\lambda \le l \le l_c$ where l_c is the LD coherence length used in the experiment

and $(\Delta\phi/2\pi)_{res}$ is the interferometer resolution in the phase-shifting technique those nominal resolution[1] is ~$\lambda/100$. The coherence length is ~10 m from ~5 MHz of the measurement spectral width of the used LD source.

A LD phase-shifting interferometer demonstrates the wide distance measurement. The minimum distance resolution depends on the broad wavelength tunability that can be supplied by an external-cavity laser diode.

5. CONCLUSIONS

We have introduced a practice of a laser-diode interferometer using one advantageous method to produce the phase shift by shifting the wavelength of the laser-diode source. This method is based on the phase difference in an unbalanced interferometer being proportional to the product of the optical path difference and the laser-diode wavelength shift. The laser wavelength can either be stepped to implement the phase-shifting interferometric approach or ramped to perform the heterodyne approach. A feedback electronics is used to calibrate the phase shifts to ensure measurement accuracy. Two-wavelength LD interferometry has been developed to extend the phase-measurement range. The phases have to be equally shifted in opposite directions to form the moiré pattern which can be applied to the phase-shifting interferometric technique. The frequency-chirped LD interferometer has been applied to the selective extraction of multiple phase objects and to a distance measurement using a phase-shifting technique with a wide tunable laser diode.

REFERENCES

1. D. Malacara, M. Servin and Z. Malacara, "*Interferogram Analysis for Optical Testing*," Marcel Dekker, New York, (1998).
2. K. Creath, "*Phase-Measurement Interferometry Techniques, in Progress in Optics XXVI*," ed. E. Wolf, Elsevier Science Publishers, London, (1988) Chap. 5, p.349.
3. Y. Ishii, "Recent developments in laser-diode interferometry," Opt. Lasers Eng. **14**, 293-309 (1991).
4. Y. Ishii, "Wavelength-tunable laser-diode interferometer," Opt. Rev. **6**, 273-283 (1999).
5. J.H. Bruning, D.R. Herriott, J.E. Gallagher, D.P. Rosenfeld, A.D. White and D.J. Brangaccio, "Digital wavefront measuring interferometer for testing optical surfaces and lenses," Appl. Opt. **13**, 2693-2703 (1974).
6. J.E. Greivenkamp, "Generalized data reduction for heterodyne interferometry," Opt. Eng. **23**, 350-352 (1984).
7. Y. Ishii, J. Chen and K. Murata, "Digital phase-measuring interferometry with a tunable laser diode," Opt. Lett. **12**, 233-235 (1987).
8. K. Tatsuno and Y. Tsunoda, "Diode laser direct modulation heterodyne interferometer," Appl. Opt. **26**, 37-40 (1987).
9. Y. Ishii and R. Onodera, "Phase-extraction algorithm in laser-diode phase- shifting interferometry," Opt. Lett. **20**, 1883-1885 (1995).

10. R. Onodera and Y. Ishii, "Phase-extraction analysis of laser-diode phase-shifting interferometry that is insensitive to changes in laser power," J. Opt. Soc. Am. A **13**, 139-146 (1996).

11. P. Sandoz, T. Gharbi and G. Tribillon, "Phase-shifting methods for interferometers using laser-diode frequency-modulation," Opt. Commun. **132**, 227-231 (1996).

12. K. Okada, A. Sato and J. Tsujiuchi, "Simultaneous calculation of phase distribution and scanning phase shift in phase shifting interferometry," Opt. Commun. **84**, 118-124 (1991).

13. Y. Takahashi, T. Yoshino and N. Ohde, "Amplitude-stabilized frequency-modulated laser diode and its interferometric sensing applications," Appl. Opt. **36**, 5881-5887 (1997).

14. R. Onodera, Y. Ishii, N. Ohde, Y. Takahashi and T. Yoshino, "Effect of laser-diode power change on optical heterodyne interferometry," J. Lightwave Technol. **13**, 675-681 (1995).

15. Y. Surrel, "Design of phase-detection algorithms insensitive to bias modulation," Appl. Opt. **36**, 805-807 (1997).

16. D. W. Phillion, "General methods for generating phase-shifting interferometry algorithms," Appl. Opt. **36**, 8098-8115 (1997).

17. G. Páez and M. Strojnik, "Fringe analysis and phase reconstruction from modulated intensity patterns," Opt. Lett. **22**, 1669-1671 (1997).

18. C. Polhemus, "Two-wavelength interferometry," Appl. Opt. **12**, 2071-2074 (1973).

19. R. Onodera and Y. Ishii, "Two-wavelength phase-shifting interferometry insensitive to the intensity modulation of dual laser diodes," Appl. Opt. **33**, 5052-5061 (1994).

20. R. Onodera and Y. Ishii, "Two-wavelength laser-diode heterodyne interferometry with one phase meter," Opt. Lett. **20**, 2502-2504 (1995).

21. Y. Ishii and R. Onodera, "Two-wavelength laser-diode interferometry that uses phase-shifting techniques," Opt. Lett. **16**, 1523-1525 (1991).

22. R. Onodera and Y. Ishii, "Two-wavelength phase-shifting interferometry using the frequency modulation of dual laser-diodes," Jpn. J. of Optics (Kogaku), in Japanese, **21**(1), 43-49 (1992).

23. J.C. Marron and K.S. Schroeder, "Three-dimensional lensless imaging using laser frequency diversity," Appl. Opt. **31**, 255-262 (1992).

24. R. Onodera and Y. Ishii, "Selective imaging with a frequency-modulated laser-diode interferometer," Opt. Lett. **20**, 761-763 (1995).

25. R. Onodera and Y. Ishii, "Multiplex imaging by a frequency-ramped laser-diode interferometer," Opt. Commun. **149**, 143-151 (1998).

26. I. Yamaguchi and K. Hamano, "Multiplexed heterodyne fiber interferometer that uses laser-diode modulation," Opt. Lett. **20**, 1661-1663 (1995).

27. M. Takeda and M. Kitoh, "Spatiotemporal frequency multiplex heterodyne interferometry," J. Opt. Soc. Am. A **9**, 1607-1614 (1992).

28. Y. Ishii, R. Onodera and T. Takahashi, "Phase-shifting interferometer for distance measurement using a tunable external-cavity laser diode," Proc. SPIE **3749** (ICO-18), 436-437 (1988).

29. R. Onodera and Y. Ishii, "Effect of beat frequency on the measured phase of laser-diode heterodyne interferometry," Appl. Opt. **35**, 4355-4360 (1996).

30. P. Carré , "Installation et utilisation du comparateur photoelectrique et interferential du Bureau International des Poids et Mesures," Metrologia **2**, 13-23 (1966).

A high-resolution velocimeter based on a fiber optic Sagnac interferometer

F. Ruan, C.Y. Liaw, Y. Zhou, Y.L. Lam[1] and S.H. Mei
Photonics Research Group, School of Electrical Electronic Engineering,
Nanyang Technological University, Nanyang Avenue, Singapore 639798

ABSTRACT

In this paper, we report a Sagnac interferometer based velocimeter with a resolution of 0.3μm/s. The sensor is constructed from commercially available telecommunication single mode fibers. A sensing probe is inserted on one side of the Sagnac loop to illuminate a test surface and to couple light back to the coil. Nonreciprocal phase shift appears between the counter-propagating waves when the surface is moving away or towards the sensing probe. A number of signal-to-noise-ratio enhancing techniques have been employed and a nanopositioner is also used to provide a known velocity to test the sensor response. Our experimental results have shown a linear relationship between the applied velocity and the sensor response within the nanopositioner restricted measurement range of 20μm/s.

Keywords: Fiber optic sensor, velocimeter, Sagnac interferometer.

1. INTRODUCTION

The Sagnac interferometer has primarily been developed to sense rotation rate [1] and is replacing the mechanical and ring laser gyroscope used in traditional application areas as well as opening new navigation and guidance opportunities due to its solid state and environmentally rugged nature. A great deal of investigation on the Sagnac interferometer carried out during the development of fiber gyroscopes has led to its present accuracy [2], making it an interesting candidate for sensing other nonreciprocal effects. Research work has also been done to modify the Sagnac interferometer for a wide range of sensing application such as temperature [3], current [4], nonlinear refractive index [5], magneto-optic effect [6], surface profile [7], time varying disturbance location [8] and others. In spite of the fact that velocity sensing was first proposed by Harvey [9] in 1992, the demonstrated velocity resolution was only 50 μm/s. In this paper, we report such a velocimeter with a much improved resolution. In our system, a polarizer and a Lyot depolarizer were inserted in the fiber systems. A PZT phase modulator was added to bias the sensor at its proper frequency. A Superfluorescent Er-Doped Fiber Source (SFS) was used as the light source. The recently reported zero crossing detection method [10] that is independent of the light source intensity, modulation depth and phase fluctuation was also employed to demodulate the rate output. Because of all these, a much-improved resolution of 0.32 μm/s was successfully demonstrated.

2. CONFIGURATION AND THEORETICAL STUDY

In the Sagnac interferometer the light waves counter-propagate in the fiber loop. When the loop experiences a rotation, a phase difference appears between the two waves. This is called the Sagnac effect. Since the amount of induced phase shift is proportional to the rotation rate, the Sagnac effect has been widely used to measure the rotation rate and the device is called gyroscope.

The Sagnac loop can be easily modified to sense linear velocity by adding a 2×2 directional coupler at one side of the fiber loop as shown in **Figure 2**. In order to facilitate light coupling in/out the fiber loop, a collimator can be spliced to the coupler to function as the sensing probe. During operation, the counter-propagating waves emanating from the collimator will hit the object, and part of the reflected light will be collected through the collimator. When the object is stationary, the optical path covered by the two traversing waves from the source to the detector is the same. However, when the object is in motion, the optical path covered by the wave pairs will be different. This is because the long SM fiber coil causes the waves to reach the object at different times and hence if the object is moving, the two waves will travel a different distance, resulting in a phase shift. Assume that the object is moving at a velocity v in the direction of the optical axis of the probe

[1] Correspondance: Lam Yee Loy; Email: eyllam@ntu.edu.sg; Telephone: (65) 7904503; Fax: (65) 7933318.

and during the time delay Δt, the object has moved a distance x from position M to M'. The two counter-propagating light waves will thus experience a phase shift $\Delta\Phi_D$ which can be expressed by,

$$\Delta\Phi_D = 4\pi x/\lambda = 4\pi v\Delta t/\lambda = 4vnL/c\lambda \quad (1)$$

Where L is the length of the fiber coil, n is the fiber refractive index, c is the velocity of light in vacuum, and λ is the wavelength of the light source.

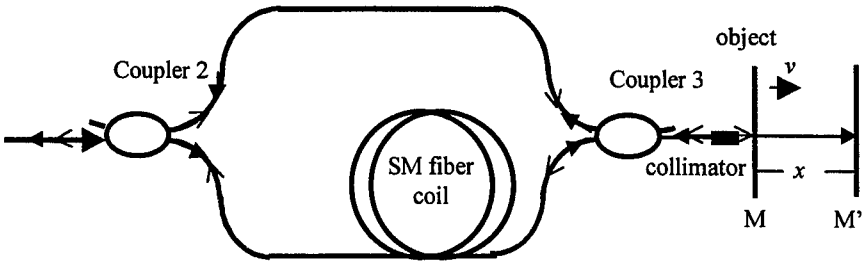


Figure 1 Phase shift induced by a moving object

Since we have constructed an open loop sinusoidally modulated gyroscope before [11], the same system has been modified to incorporate the sensing probe as shown in **Figure 1**. An overview of our fiber optic velocimeter is shown in **Figure 2**.

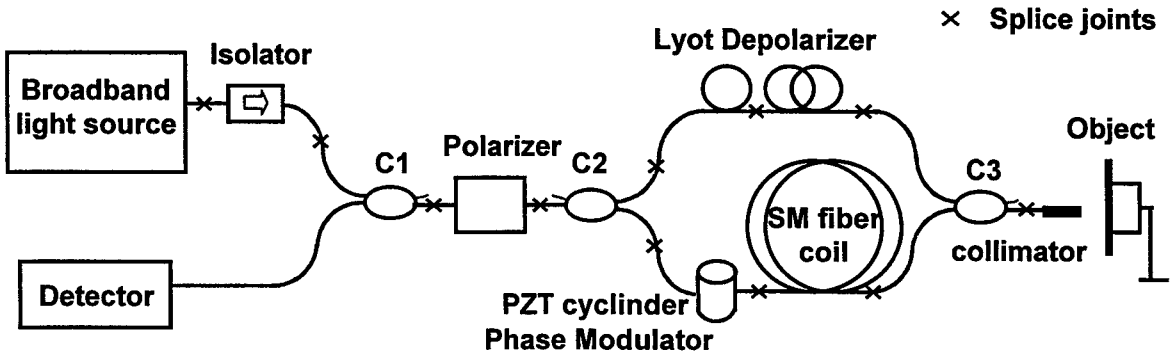


Figure 2 Configuration of the Sagnac velocimeter

In this system, a Superfluorescent Er-Doped Fiber Source (SFS) as shown in **Figure 3** was employed as the light source. Firstly, it provides an adequate power to yield a quantum-limited signal to noise ratio. Secondly, its short coherent length helps to reduce the receiver intensity noise and sensor bias instabilities; it exhibits less random wavelength fluctuation as well as predictable wavelength dependence on temperature [11]. Furthermore, it emits unpolarized light, which is very beneficial for reducing polarization-induced nonreciprocities. An isolator is placed at the output of the SFS to prevent any back reflection that could cause the SFS to oscillate and lase. The SFS has an output power of 26mW with a mean wavelength of 1550nm. The sensing loop was constructed from commercially available single mode fiber (SMF) wound on a spool with a diameter of 15 cm. The total length of fiber coil as measured by an optical time domain reflector (OTDR) is about 4.773 kilometers. A polarizer was spliced between coupler 1 and coupler 2 and was tuned for maximum transmission of the power from the source to the sensing coil so as to improve the signal to noise ratio. A Lyot depolarizer [12] was inserted in the loop to reduce polarization fading effect found in single mode fiber gyros. It was constructed from two Polarization Maintaining (PM) fibers with their fast axes spliced at an angle of 45° with respect to each other [13]. To increase the

sensitivity, a PZT phase modulator is placed at one side of the Sagnac loop to generate a time varying modulation signal. In general, the detector signal from such an open loop sinusoidally modulated Sagnac interferometer can be represented by [1],

$$I(\Delta\Phi_D) = I_0[1 + \cos(\Delta\Phi_D + \Phi_m \cos(2\pi f_m t + \Phi))] \tag{2}$$

where Φ_m is the optical phase modulation depth, f_m is the modulation frequency, and Φ is the phase uncertainty between the modulation drive signal and the output signal. Using J_n Bessel function expansion, the interference signal at the detector can be re-expressed as,

$$\begin{aligned} I(\Delta\Phi_D) = I_0 &+ I_0 \cos(\Delta\Phi_D)[J_0(\Phi_m) + 2J_2(\Phi_m)\cos(4\pi f_m t + \Phi) + ...] \\ &+ I_0 \sin(\Delta\Phi_D)[2J_1(\Phi_m)\sin(2\pi f_m t + \Phi) + 2J_3(\Phi_m)\sin(6\pi f_m t + \Phi) + ...] \end{aligned} \tag{3}$$

From this expression, we can see that the sensor output contains even and odd harmonics, the even harmonics are proportional to $\cos(\Delta\Phi_D)$ and the odd harmonics, particularly the fundamental phase modulation frequency harmonic term, are proportional to $\sin(\Delta\Phi_D)$. Thus when the object is at rest, the modulator will generate second and higher order even harmonics of f_m on the detector. In the presence of motion, there will be a relative phase shift between the counterpropagating light waves and the odd harmonic components start to appear. The amplitude of these odd harmonic components is proportional to the velocity and the sign of the phase angle shows the moving direction.

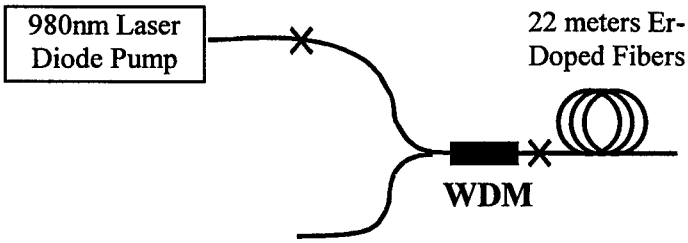


Figure 3 Superfluorescent Er-Doped Fiber Source

3. DEMODULATION SCHEME

One of the most commonly used methods of retrieving the phase information $\Delta\Phi_D$ is to synchronously detect the first harmonic of the output signal, however the signal output is also sensitive to ϕ_m, I_o and ϕ. As an improvement over this technique, a new zero cross detection approach that is insensitive to the three factors was proposed by Tselikov [10] in 1998. Tselikov verified that the phase shift to be measured is proportional to the time difference between two consecutive time intervals (ΔT) when the first harmonic output signal crosses zero. Their relationship can be expressed by,

$$\Delta\Phi_D = -\frac{\pi f_m \Delta T}{2}\left\{\Phi_m{}^2 - \arccos[J_o(\Phi_m)]^2\right\}^{1/2} \tag{4}$$

Although Φ_m still appears in this expression, it can be controlled to maintain constant by the zero cross detection scheme, thus the demodulation output is the only a function of ΔT. The block diagram of the detection circuit that we have employed to measure ΔT and meanwhile control Φ_m is shown in **Figure 4**. Detailed description of the circuit can be found in Ref 10 and is thus not repeated here.

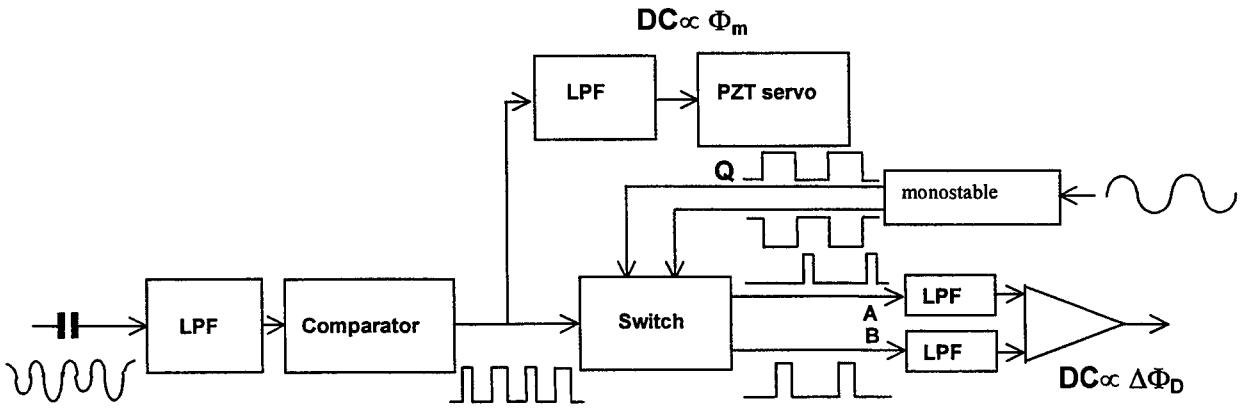


Figure 4 Block diagram of zero cross detection

4. EXPERIMENTS

A shiny silicon wafer was used as the object. It was mounted on a nanopositioner to allow precise motion control. The sensor structure was placed in an environmental chamber to achieve isolation from external disturbances. A low noise *New Focus* photo detector with a Resposivity of 0.8 A/W was employed as receiver. It is known that modulating at the proper frequency or its odd harmonics produces a phase biasing signal that does not degrade the original perfection of the system. And in general, for practical systems, fiber coil length × proper frequency ≈ 100 m•MHz [1]. For our system, the calculated proper modulation frequency is 20.9 KHz. A sinusoidal modulation of 1.5 V_{pp} at 60.6 KHz was applied to the PZT phase modulator to achieve high sensitivity and low noise. When the wafer was at rest, the detector output signal is shown in **Figure 5**, where the lower trace is the modulation signal and the upper trace is the detector output signal before the zero-cross demodulation circuit. It is noted that the frequency of the sensor output is twice that of the modulation signal, which agrees with theoretical analysis. The output after the demodulation circuit is shown in **Figure 6**, in which ch1 is the position signal for the nanopositioner, and ch2 is the demodulation output. As can be seen, demodulated sensor output contains some noise even when the nanopositioner is at rest.

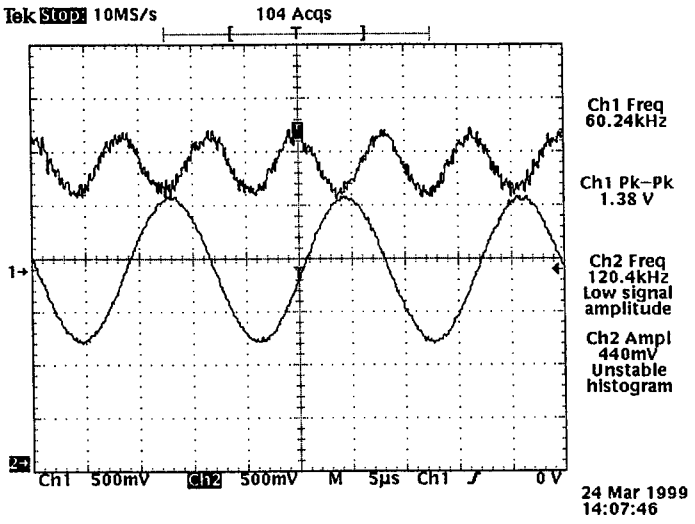


Figure 5 Detector output at rest

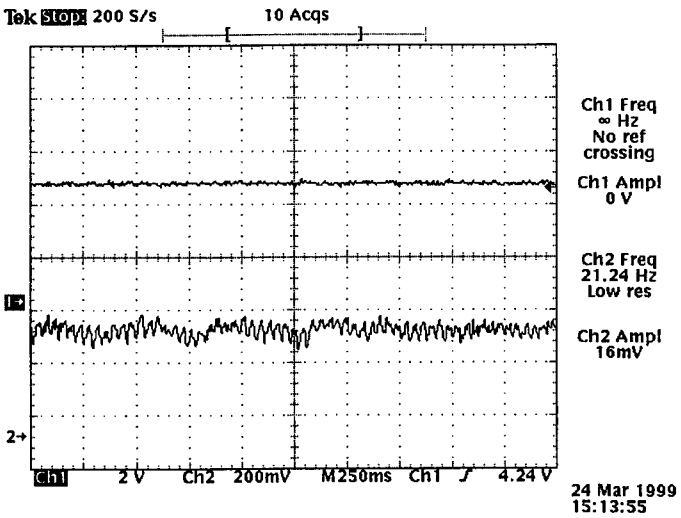
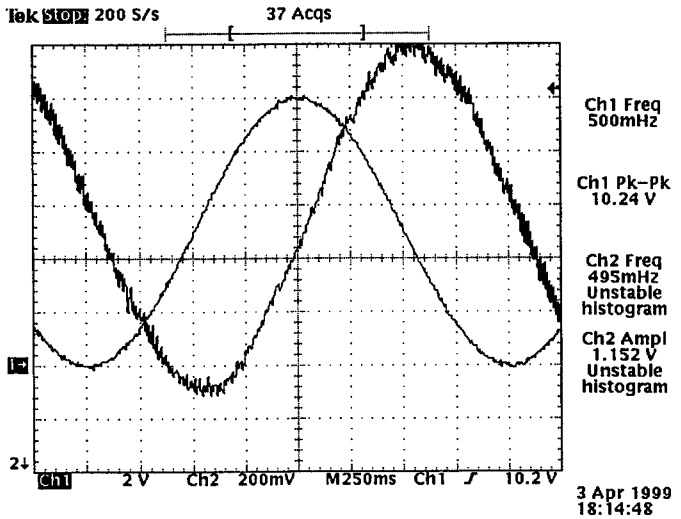


Figure 6 Demodulation output at rest

4.1 Velocity response

To verify the capability of our sensor in measuring velocity, different waveforms were applied to the nanopositioner so that the wafer moved in different ways to generate different velocity patterns. The output of the detector was fed into the zero cross detection circuit for demodulation. Both the wafer position signal and demodulated sensor output signal were monitored by an oscilloscope. **Figure 7** shows the waveforms when the wafer is driven by a sinewave, a square wave and a triangular wave respectively. In these figures, ch1 (fine curve) is the position signal and ch2 (noisy curve) is the sensor output signal. It can be seen that the sensor output signal is the differentiation of the position signal with respect to time regardless of the signal applied to nanopositioner. As is well known the position differentiation is the velocity. This verifies the feasibility of the velocity measurement of our sensor.



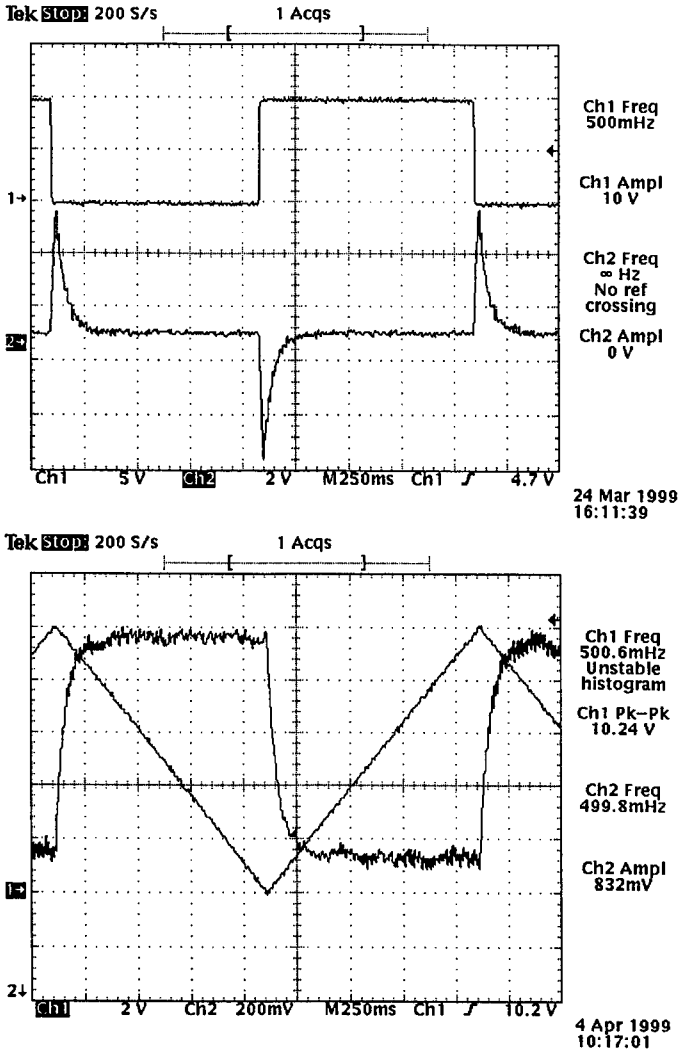


Figure 7 Sensor response when the nanopositioner moving at different traces

4.2 Linearity of the velocity response

To obtain the linearity of the sensor's velocity response, we conducted an experiment by recording the sensor output while changing the wafer velocity stepwise. The result is shown in **Figure 8**. It can be seen that the sensor output has a linear relationship with the applied velocity. The sensitivity or the slope of the response is 41mv/(μm/s).

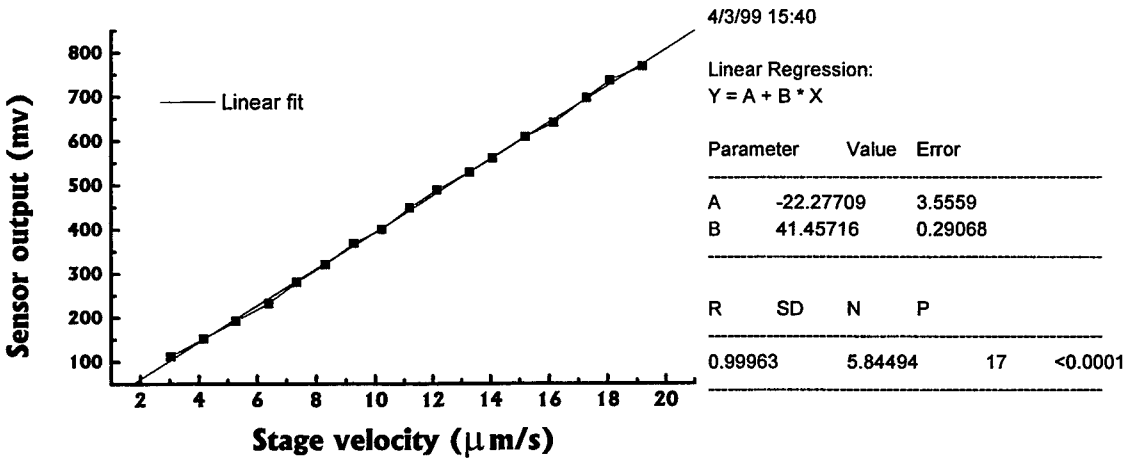


Figure 8 Velocity response of the velocimeter

4.3 Sensor velocity resolution

To determine the minimum detectable velocity of the sensor, a triangular wave was applied to the nanopositioner to move the wafer at a constant velocity. The operating frequency of the nanopositioner was chosen to be 0.5Hz. Reducing the amplitude of the triangular wave reduces the wafer velocity. **Figure 9** shows the case when amplitude of the triangular wave is decreased to such an extent that the sensor output square wave can just be distinguished. It can be seen that the amplitude of the wafer position signal is 80mV. This is equivalent to a velocity of 0.16μm/s. Consequently, from the fact that the velocity sweeps from -0.16μm/s to 0.16μm/s, the resolution of our velocity sensor is thus estimated to be about 0.32μm/s. It is worthy to highlight that the minimum detectable rotation rate for this gyroscope is 0.828°/hr [11], which is in fact equivalent to a linear velocity of 0.32 um/s ($v=R*\Omega=0.828*\pi/180*1/3600m/s=3.2*10^{-7}$m/s=0.32um/s). Hence, the results from the gyroscope and the velocimeter are in good agreement.

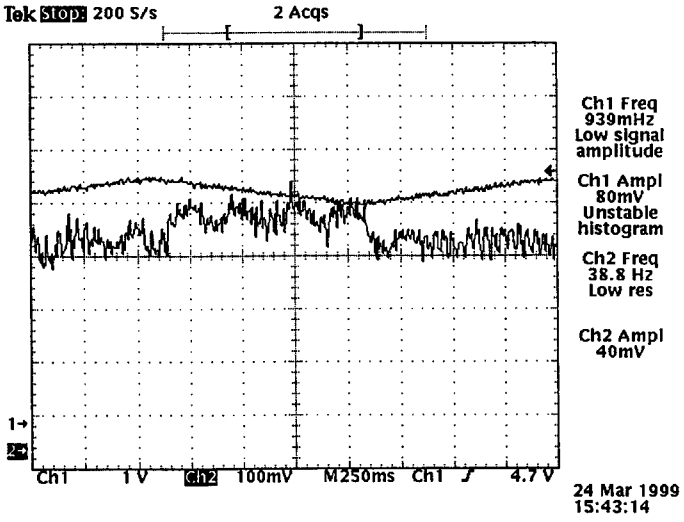


Figure 9 Sensor resolvable velocity response

5. CONCLUSION

An all fiber velocimeter, which is based on an open loop Sagnac interferometer, has been successfully demonstrated with a relatively high resolution. It is constructed from commercially available telecommunication single mode fibers. The zero cross detection method, which produces a signal output that is independent of the light source intensity, modulation depth and phase fluctuation, has been successfully adopted for demodulating the velocity information. Experiments have been performed to test the sensor's velocity response and the results have shown that the sensor output is indeed proportional to the object velocity. The minimum detectable linear velocity of 0.32 μm/s is in agreement with the minimum detectable rotation rate obtained from the same gyroscope.

6. REFERENCES

1. Herve Lefevre, *"The fiber-optic gyroscope"*, Artech House, 1992.
2. G.A. Sanders, B. Szafraniec, L. Strandjord, R. Bergh, A. Kaliszek, R. Dankwort, C. Lange, and D. Kimmel. "Progress in high performance fiber optic gyroscopes", *Technical Digest Series.* v 16, pp. 116-121, 1997.
3. Starodumov, A N. Zenteno, L A. Monzon, D. De La Rosa, E. "Fiber Sagnac interferometer temperature sensor" *Applied Physics Letters.* v 70 n 1, pp. 19-21, Jan 6 1997.
4. Yu, A. Siddiqui, A S. "Practical Sagnac interferometer based fibre optic current sensor" *IEE Proceedings Optoelectronics.* v 141 n 4, pp. 249-256, Aug 1994.
5. Li, K. Xiong, Z. Peng, G D. Chu, P L. "Direct measurement of nonlinear refractive index with an all-fibre Sagnac interferometer" *Optics Communications*". v 136 n 3-4, pp. 223-226, Mar 15 1997.
6. Kapitulnik, A. Dodge, J S. Fejer, M M. "High-resolution magneto-optic measurements with a Sagnac interferometer", *Journal of Applied Physics.* v 75 n 10 pt 2B, pp. 6872-6877, May 15 1994.
7. Somervell, A R D. Barnes, T H. "Unambiguous measurement of surface profile using a Sagnac interferometer with phase feedback" *Optics Communications.* v 150 n 1-6, pp. 61-65, May 1 1998.
8. Chtcherbakov, Anatoli A. Swart, Pieter L. Spammer, Stephanus J. "Mach-Zehnder and modified Sagnac-distributed fiber-optic impact sensor" *Applied Optics.* v 37 n 16, pp. 3432-3437, Jun 1 1998.
9. David Harvey, Roy McBride, James S Barton and Julian D C Jones, "A velocimeter based on the fibre optic Sagnac interferometer," *Meas. Sci. Technol.* n 3, pp. 1077-1083, 1992.
10. A Tselikov, J.U.de Arruda, and J. Blake, "Zero-crossing demodulation for open-loop Sagnac interferometrs," *Journal of lightwave technology IEEE,* v 16, No 9, pp. 1613-1619, September 1998.
11. Chin-Yi Liaw, Yan Zhou and Lam Yee loy, "Development of a low-cost fiber optic Gyroscope for navigation", *Defence Electronics Technology Seminar,* 13 November, pp. 126-130, 1997.
12. William K. Burns, "Optical fiber rotation sensing," *Quantum Electronics-Principles and applications, Academic Press,* pp. 69, 1994.
13. William K. Burns, "Fiber-optic gyroscope with depolarized light," *Journal of light wave technology,* v 10, n 7, pp. 992-999, July 1992.

pH Sensitivity and Hysteresis of A-WO_3 Gate ISFET Compared with Different Membranes

Jung-Lung Chiang[a], Jung-Chuan Chou*[b], Ying-Chung Chen[a]

[a]Department of Electrical Engineering, National Sun Yat-Sen University, Kaohsiung, Taiwan 804, R.O.C.
[b]Institute of Electronic and Information Engineering, National Yunlin University of Science and Technology, Touliu, Yunlin, Taiwan 640, R.O.C.

ABSTRACT

Because of the pH sensitivity is one of the important characteristic parameters of ISFET devices. The response of ISFET is mainly determined with the type of the sensing membrane, therefore the sensing material plays a significant role. In addition, the hysteresis is the non-ideal and unstable factor of ISFET devices for measuring. Hence, in this study, the pH sensitivity and hysteresis of a-WO_3 gate ISFET are investigated, and compare with different sensing membranes.
In this paper, the pH sensitivities of the different gate ISFET devices were measured in different solutions by the I-V curve of measuring system, and the hysteresis curves were measuring by exposing several cycles of pH values at different loop time. According to the experimental results, a-WO_3, Ta_2O_5 and a-Si:H have high response, the pH sensitivities are over 50mV/pH. However, it is also found that the hysteresis of a-WO_3 gate ISFET is dependent on measuring loop time and measuring path, and hysteresis width is increasing with loop time and measuring path.

Keywords: ISFET, sensitivity, hysteresis, a-WO_3, a-Si:H

1.INTRODUCTION

The application of the ion sensitive field effect transistor (ISFET) device was first explored by P. Bergveld [1] in 1970 for neurophysiological measurements. In the past few years, many researches have been done to characterize ISFET based on MOS technology and widely applied in the field of the chemical and biomedical sensors. Because of ISFET has the potential advantages over conventional ion-selective electrode in their rapid response, small size, low cost, high input impedance and low output impedance [2-3]. The ISFET is an integrated device and similar to the metal-oxide-semiconductor field effect transistor (MOSFET). In addition, the metal gate of MOSFET is replaced by a pH sensing membrane and electrolyte solution in order to expose to solution and detect ion response. When P. Bergveld first introduced ISFET, the device was operated without a reference electrode. However, later work by other investigators indicates that proper operation of ISFET requires a reference electrode to establish the electrolyte potential with respect to the semiconductor substrate [4].

The response of ISFET is mainly determined by the type of the sensing membrane, therefore the sensing material plays a significant role. Therein, the SiO_2 film was first used as the sensing membrane, but it showed an inferior pH response. Afterwards, application of sensing membranes such as Si_3N_4[5-7], Al_2O_3[8-9], Ta_2O_5[5,10-11], SnO_2[12-13], a-WO_3[14-15] and a-Si:H[16-17] deposited on the SiO_2 improved the pH response of the device. pH-ISFETs are attractive chemical sensors. However, there are several mechanisms limit the accuracy with which an ISFET can be used to measure pH sensitivity. And it has been shown that the time dependent effects are major source of measurement error. These problems limited the progress of developing a commercial device. One of the time dependent effects, hysteresis is a serious problem which results from slow response and drift[5,17-19]. In this paper, we present pH sensitivity and hysteresis of a-WO_3 gate ISFET, and compare with different materials.

Firstly, the amorphous tungsten trioxide (a-WO_3) thin films were prepared by rf sputtering system from a WO_3 target and were deposited on the double layer structure of a-WO_3/SiO_2 gate ISFET device. The a-WO_3 thin films were selected as a pH sensitive dielectric for the ISFET, because they have high sensitivity for hydrogen ion in our recent experimental results[14-15]. Another, the tantalum pentoxide (Ta_2O_5) has been found to have superior pH response[10]. In our experiments, the Ta_2O_5 thin

Part of the SPIE Conference on Advanced Photonic Sensors and Applications
Singapore • November–December 1999
SPIE Vol. 3897 • 0277-786X/99/$10.00

film was also prepared by sputtering Ta_2O_5 target. Furthermore, hydrogenated amorphous silicon (a-Si:H) thin film was also selected as a sensing material of ISFET device by plasma-enhanced chemical vapor deposition (PECVD). Because of the a-Si:H thin film can be deposited easily, at low temperature and low cost, on substrate of almost any shape and size. It is commonly applied to photoelectric devices such as thin-film transistor (TFT) and solar cell. In recent, it is also applied to the property of pH sensitive in our laboratory[16.-17.]. In order to evaluate the pH sensitivity and hysteresis of a-WO_3-gate, a-Si:H-gate and Ta_2O_5-gate ISFETs, the HP 4145B Semiconductor Parameter Analyzer was utilized to measure the I-V curves, and the devices were operated in feedback mode to obtain the hysteresis curves.

2.EXPERIMENTALS

2.1 ISFET Fabrication

In this study, the n-channel ISFETs were fabricated on p-type silicon wafers of 8 –12 Ω • cm resistivity, (100)-oriented. The channel was 50 μm long and 1000 μm wide (i.e. W/L = 20). The thickness of the thermally grown SiO_2 layer was 1000 Å. The fabrication processes were the same as *ref.* 14.

2.2. Prepared Sensing Membranes

The prepared conditions of the a-WO_3 , a-Si:H and Ta_2O_5 thin films grown on the SiO_2 layer are as follows:

2.2.1. a-WO_3 thin film

The amorphous tungsten trioxide thin film was deposited on substrate maintained at room temperature by rf sputtering from a WO_3 target, and sputtering total operating pressure of 30 mtorr in Ar gas mixed 5%O_2 gas for 2 hours and the net rf power was 90W. We finished the a-WO_3/SiO_2 gate ISFET devices.

2.2.2. Ta_2O_5 thin film

Ta_2O_5 thin film was grown by rf sputtering with a Ta_2O_5 target, and Ar ($Ar:O_2$=9:1) pressure of 10 mtorr. The mass flow rate was 20 sccm, and net rf power was 90W. The thickness of the Ta_2O_5 thin film was about 1000 Å, and the structure of Ta_2O_5/SiO_2 gate ISFETs were finished.

2.2.3. a-Si:H thin film

a-Si:H thin film was deposited by plasma-enhanced chemical vapor deposition (PECVD) at the substrate temperature of 210 ℃, and silane ($SiH_4:H_2$=1:9) pressure of 0.7 torr. The mass flow rate was 120 sccm, and net RF power was 29W. The thickness of the a-Si:H thin film was approximately 1000Å, and the structure of a-Si:H/SiO_2 gate ISFETs was finished.

According to above statement, we rewrite the prepared conditions of a-WO_3 thin film, Ta_2O_5 thin film and a-Si:H thin film are shown in Table 1. And the schematic representation of an ISFET device with reference electrode and sensing membrane is shown in Fig.1.

Table 1. Prepared conditions of a-WO_3 , Ta_2O_5 and a-Si:H thin film

	a-WO_3	Ta_2O_5	a-Si:H
System	RF Sputter	RF Sputter	PECVD
Reaction Pressure (torr)	30×10^{-3}	10×10^{-3}	0.7
Net RF Power (W)	90	90	29
Mass Flow Rate (sccm)	50	20	20
Reaction Gas	$Ar:O_2$	$Ar:O_2$	$SiH_4:H_2$
	19:1	9:1	1:9

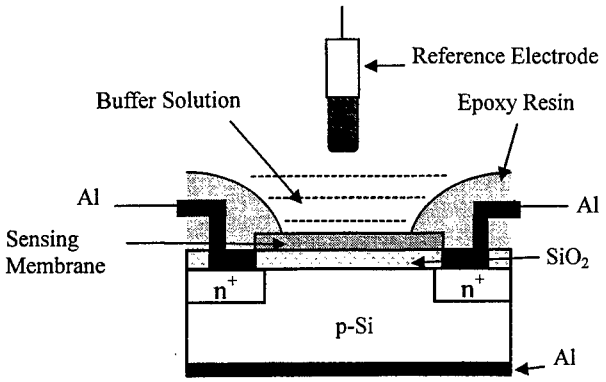
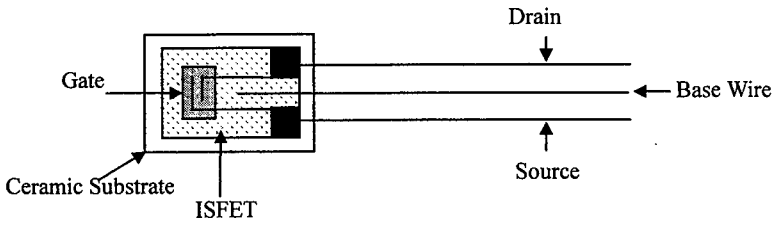
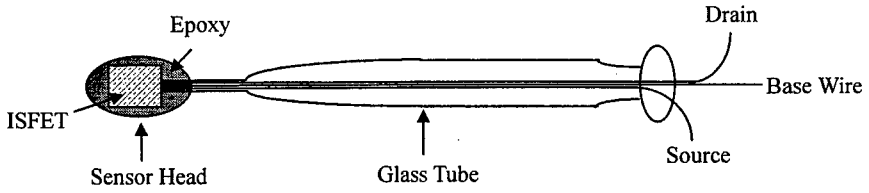


Fig.1 Schematic representation of an ISFET device

As shown in Fig.1. Since ISFET devices work in the electrolyte solutions, they must be encapsulated in such a way that only the sensing membrane is exposed to the electrolyte solutions. In this study, ISFET devices were packaged with epoxy resin for encapsulation. After bounding wire and packaging with epoxy, the ISFET devices were formed, and the section of packaged device was shown in Fig.2.



(a)



(b)

Fig.2 Diagram sketch of ISFET.
(a) Sensor head (b) ISFET device

2.3. Measurement

2.3.1. I-V measurement

Figure 3 shows the I-V measurement system. The HP4145B Semiconductor Parameter Analyzer was used to measure the threshold voltage of the ISFET devices in the different buffer solutions. In order to avoid the photoelectric and temperature

effects, all the measurements were placed in the dark box with PID temperature controller at 25℃. Afterward, we measured the I-V curves of devices to investigate the pH sensitivity. The pH sensitivities of the a-WO_3, Ta_2O_5, a-Si:H gate ISFETs were determined through the shift of the threshold voltage as shown in Fig. 4, 5, and 6, respectively.

2.3.2 Hysteresis measurement

In Fig.7, the ISFET devices were operated in the feedback mode which maintain a constant drain to source current as well as a constant drain to source voltage, then the response will reflect on the gate of ISFETs. And all test instrumentation must be placed in the dark box. Therein, the read out circuit was shown in Fig. 8, allowed the drain to source voltage and drain to source current to regulate individually by R1 and R2, respectively. The operational amplifier A2(LF356) provides a feedback so that the gate voltage varies with the change of time and pH value. We set a drain-to-source current of I_{DS}=60μA with a constant drain-to-source voltage of V_{DS}=0.2V as well as the condition of the I-V measurement. The gate voltage of device was monitored and recorded by a V-T recorder (Data chart 3000 Voltage-Time recorder).

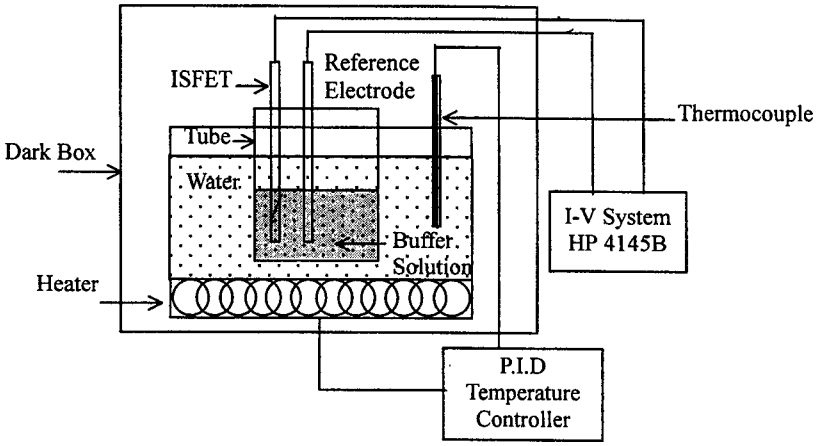


Fig.3 I-V measurement system

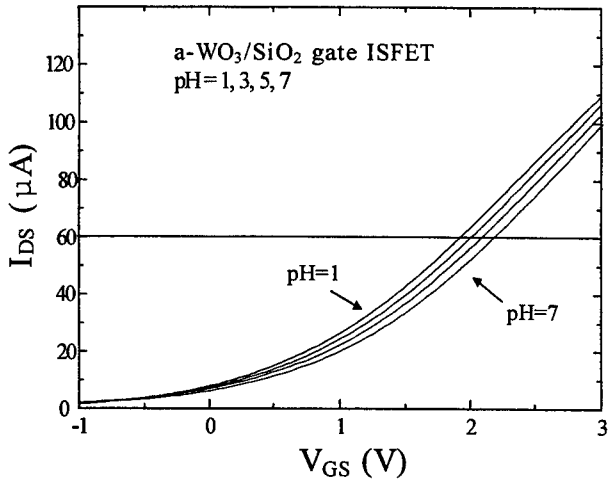


Fig.4 I_{DS}-V_{GS} curves of the a-WO_3/SiO_2 gate ISFET

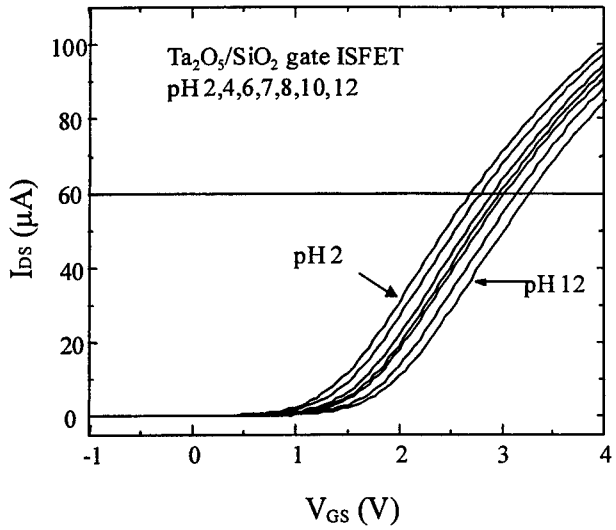


Fig.5 I_{DS}–V_{GS} curves of the Ta_2O_5/SiO_2 gate ISFET

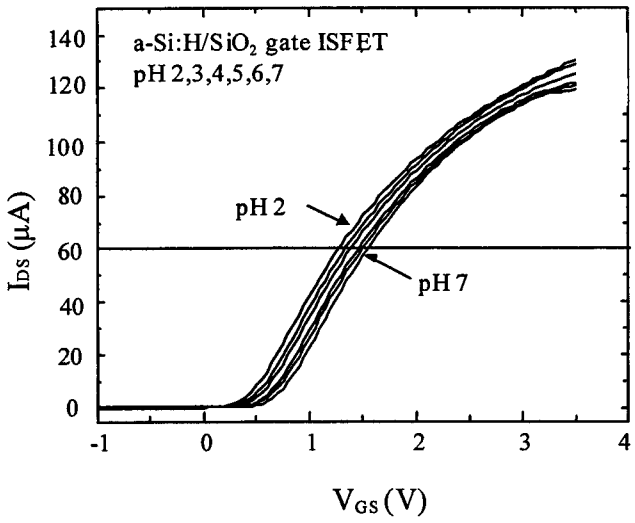


Fig.6 I_{DS}–V_{GS} curves of a-Si:H/SiO_2 gate ISFET

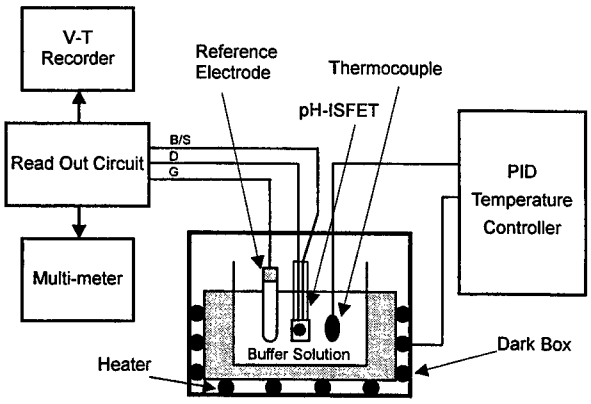


Fig.7 Diagram of the measurement setup

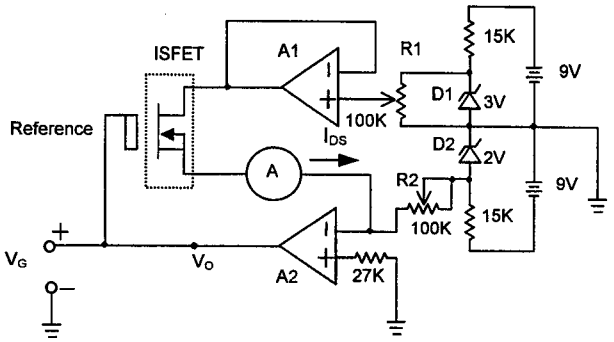


Fig.8 Schematic diagram of the read out circuit

3.RESULTS AND DISCUSSION

3.1. pH Sensitivity

In the I-V measurement, the ISFET devices and a calomel reference electrode were immersed in the different buffer solutions, a family of I-V curves can be generated by applying a series of step voltage at the reference electrode. The results shown that the I_{DS}–V_{GS} curves of a-WO_3 gate, Ta_2O_5 gate and a-Si:H gate ISFETs are shown in Fig. 4, 5, and 6, respectively. The pH sensitivities of a-WO_3 gate, Ta_2O_5 gate and a-Si:H gate ISFETs can be calculated from these figures. From the results of experimental, they are obvious that threshold voltage increases with increasing pH value.

According to J. C. Van Kerkhof *et. al.*[20] and in our previous work.[14,15] A-WO_3 thin film dissolved rapidly in the basic solutions. For this reason, in this study, the a-WO_3 gate ISFET devices were immersed in the acid solutions in the entire measured processes. As shown in Fig. 4, the pH sensitivity of a-WO_3 gate ISFET has the linear pH response, which closes to 50.2~57.6 mV/pH between pH1 and pH 7. And it is found that the pH sensitivity is independent of the thickness of thin films[14].

In the past research, the Ta_2O_5 gate ISFET has high pH sensitivity near to theoretical. Likewise, in this study, as shown in Fig. 5, the Ta_2O_5 gate ISFET has a fairly high and linear pH sensitivity about 57.1~58.3 mV/pH in the pH2 to pH12. However, the I_{DS}-V_{GS} curves of a-Si:H gate ISFET were accomplished by I-V measuring system in the acid solutions, because of the a-Si:H film would be dissolved in alkaline solutions. In this study, we also found that the a-Si:H gate ISFET

have a linear pH response. According to the measuring results, that is shown in Fig. 6, the pH sensitivity of a-Si:H gate ISFET is about 51.3~56.5 mV/pH in the concentration between pH2 and pH7.

Additionally, the pH sensitivity of property for a-WO_3, Ta_2O_5 and a-Si:H gate ISFETs will be compared with those of SiO_2[3], Si_3N_4[3], Al_2O_3[3, 5, 8, 9] and SnO_2[12]. The pH sensitivities of SiO_2, Si_3N_4, Al_2O_3 and SnO_2 used as ion-sensing membranes are listed in Table 2. The SiO_2 gate ISFET has a low pH sensitivity, and the sensitivity is unstable because it is varied by pH value of electrolyte solutions. In contrast with Si_3N_4, Al_2O_3, SnO_2, a-Si:H , Ta_2O_5, and a-WO_3 gate ISFETs indicate superior pH sensitivities.

From the results mentioned above, in our study, the a-WO_3, a-Si:H and Ta_2O_5 gate ISFETs show the near Nernstian responses, namely, the pH sensitivity is over 50mV/pH. Thus, a-WO_3, Ta_2O_5 and a-Si:H thin films are suitable for ion-sensing membranes of ISFET devices in the different testing environment.

Table 2. Comparison of sensitivity and hysteresis for different sensing membranes of ISFET

Membrane	SiO_2[3]	Si_3N_4[3,5]	Al_2O_3[3,5,8]	SnO_2[12]	a-WO_3	a-Si:H[17]	Ta_2O_5[5]
Test range	4 ~ 10	1 ~ 13	1 ~ 13	2 ~ 10	1~7	2 ~ 7	2 ~ 12
Sensitivity (mV/pH)	25 ~ 48	46 ~ 56	53 ~ 57	58	50.2~57.6	51.3~56.5	57.1~58.3
Hysteresis* (mV)	unstable	6.29~10.92	3.73~4.4	2.5	7.2~12.5	1.5~17.9	<2

* The hysteresis width was obtained in the different testing cycles.

3.2. Hysteresis of a-WO_3

Generally, a measurement starts when the sensor is stable at pH 7 until its drift rate is below 10^{-4} mV/s[5], the time needed is about 18h. Then a pH loop was started with increments of one pH unit and extreme value of pH1 and pH7. However, for the applications, the hysteresis curves of a-WO_3 gate ISFET is first investigated in our lab. The sensors is directly to placed in pH 3→1→3→5→3 and pH 3→5→3→1→3 at 600s loop time, in order to prevent the a-WO_3 thin film dissolves in the solutions during the measuring processes[14,15]. According to Fig. 9, we can obtain that the hysteresis amounts of the pH 3→1→3→5→3 and pH 3→5→3→1→3 loops are 7.2mV and 12.5mV respectively, the acid-side hysteresis is smaller than basic-side hysteresis. Otherwise, the hysteresis width in pH 3→1→3→5→3 at various loop time such as 600s, 1200s and 2400s are 7.2, 9.7 and 15.4mV, respectively. It shows that the hysteresis width increases with increasing loop time. The experimental results are the same as the *ref.* 5, 8, 17 for Si_3N_4[5], Al_2O_3[5,8], Ta_2O_5[5] and a-Si:H[17] gate ISFET.

Moreover, the hysteresis curves of a-WO_3 gate ISFET in pH 3→1→3→5→3 and pH 4→1→4→7→4 at 600s loop time. It is found that the hysteresis width increases with increasing pH loop path, and the hysteresis widths are 7.2mV and 12mV, respectively.

3.2.1 Comparison of hysteresis for a-WO_3, SiO_2, Si_3N_4, Al_2O_3, SnO_2, Ta_2O_5 and a-Si:H

According to Luc Bousses *et a.l.*[5,8] and our research, the hysteresis effect of a-WO_3 gate ISFETs will be compared with those of SiO_2, Si_3N_4, Al_2O_3, SnO_2, Ta_2O_5 and a-Si:H in the different measuring cycles. The hysteresis of widths are also listed in Table 2. Therein, the Ta_2O_5 has the best performance of hysteresis characteristic than the other sensing materials, the hysteresis width is less than 2 mV[5]. The SiO_2 shows unstable[3], the Si_3N_4 is 6.29~10.92 mV[5], the Al_2O_3 is 3.73~4.4 mV[5], the SnO_2 is 2.5 mV[12], a-Si:H is 1.5~17.9 mV[17], and a-WO_3 is about 7.2~12.5 mV.

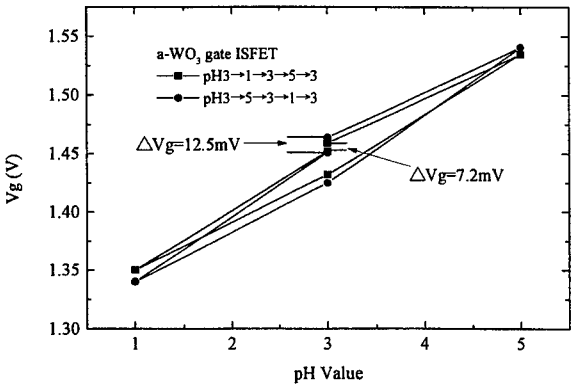


Fig.9 Hysteresis curves of a-WO_3 gate ISFET device

4.CONCLUSIONS

Owing to the pH sensitivity is a decisive factor which affects the response of a ISFET device. In this study, in order to improve the pH response, the a-WO_3 thin film was used as sensing membrane. In addition, the hysteresis is the non-ideal and unstable factor of ISFET devices for measuring. Hence, in this study, the pH sensitivity and hysteresis of a-WO_3 gate ISFET are investigated, and compare with different sensing membranes. We could summary as follows:

1. In this study, the a-WO_3 gate ISFET has a high pH sensitivity about 50.2~57.6 mV/pH. The hysteresis of 7.2 to 12.5 mV is seen in the pH 3→1→3→5→3 and pH 3→5→3→1→3 loops.
2. The hysteresis effect of a-WO_3 gate ISFET is first investigated in our lab. It shows that the acid-side hysteresis is smaller than basic-side hysteresis, the hysteresis width increases with increasing loop time in the range of 600s-2400s during the pH 3→1→3→5→3 cycle, and also finds that the hysteresis width is dependent on measuring loop time and measuring path.
3. In comparison, the a-WO_3 and Si_3N_4, Al_2O_3, SnO_2, Ta_2O_5, a-Si:H thin films are suitable for ion-sensing membranes of ISFET devices in the different testing environment, because of their pH sensing responses are fairly high.

ACKNOWLEDGMENTS

This study is supported by National Science Council, the Republic of China under the contracts NSC 88-2215-E-224-001.

REFERENCES

1. P. Bergveld, "Development of an ion-sensitive solid-state device for neurophysiological measurements", IEEE Trans. Biomed. Eng., vol. BME-**17**, pp. 70-71, 1970.
2. A. Topkar and R. Lal, "Effect of electrolyte exposure on silicon dioxide in electrolyte-oxide-semiconductor structures", Thin Solid Films, **232**, pp. 265-270, 1993.
3. T. Matsuo and M. Esashi, "Method of ISFET fabrication", Sensors and Actuators, **1**, pp. 77-96, 1981.
4. C. D. Fung, P. W. Cheung and W. H. Ko, "A generalized theory of an electrolyte-insulator-semiconductor field-effect transistor", IEEE Trans. Electron Devices, vol. ED-**33**, pp. 8-18, 1986.
5. L. Bousse, S. Mostarshed, B. van der Schoot and N. F. de Rooij, "Comparison of the hysteresis of Ta_2O_5 and Si_3N_4 pH-sensing insulators", Sensors and Actuators B, **17**, pp. 157-164, 1994.
6. D. L. Harame, L. J. Bousse, J. D. Shott and J. D. Meindl, "Ion-sensing devices with silicon nitride and borosilicate glass insulators", IEEE Trans. Electron Devices, vol. ED-**34**, pp. 1700-1706, 1987.
7. M. N. Niu, X. F. Ding and Q. Y. Tong, "Effect of two types of surface sites on the characteristics of Si_3N_4-gate

pH-ISFETs", Sensors and Actuators B, **37**, pp. 13-17, 1996.
8. L. Bousse, H. H van der Vlekkert and N. F. de Rooij, "Hysteresis in Al_2O_3-gate ISFETs", Sensors and Actuators B, **2**, pp. 103-110, 1990.
9. H. van den Vlekkert, L. Bousse and N. de Rooij, "The temperature dependence of the surface potential at the Al_2O_3/electrolyte interface", Journal of Colloid and Interface Science, **122**, pp. 336-345, 1988.
10. A. S. Poghossian, "The super-nernstian pH sensitivity of Ta_2O_5-gate ISFETs", Sensors and Actuators B, **7**, pp. 367-370, 1992.
11. D. H. Kwon, B. W. Cho, C. S. Kim and B. K. Sohn, "Effect of heat treatment on Ta_2O_5 sensing membrane for low drift and high sensitivity pH-ISFET", Sensors and Actuators B, **34**, pp. 441-445, 1996.
12. H. K. Liao, J. C. Chou, W. Y. Chung, T. P. Sun and S. K. Hsiung, "Study of amorphous tin oxide thin films for ISFET applications", Sensors and Actuators B, **50**, pp. 104-109, 1998.
13. H. K. Liao, L. L. Chi, J. C. Chou, W. Y. Chung, T. P. Sun and S. K. Hsiung, "Study on pH_{pzc} and surface potential of tin oxide gate ISFET", Materials Chemistry and Physics, **59**, pp. 6-11, 1999.
14. J.L. Chiang, J.C. Chou and Y. C. Chen, "Study on the pH sensing property of amorphous tungsten trioxide thin films by rf sputtering", Proceedings of 1998 International Electron Devices and Materials Symposia (IEDMS'98), National Cheng Kung University Tainan, Taiwan, R.O.C., Symposium B、C, **CP-4**, pp. 337-340, December 20-23, 1998.
15. J. C. Chou and J.L. Chiang, "Study on the amorphous tungsten trioxide ion sensitive field effect transistor", Proceedings of Technical Digest of the 7th International Meeting on Chemical Sensors, Beijing International Convention Center, Beijing, China, pp. 535-537, July 27-30, 1998.
16. J. C. Chou and J. S. Lin, "Effect of operating temperature on sensitivity of a-Si:H ISFET", Proceedings of The 1998 Annual Conference of The Chinese Society for Materials Science, Tatung Institute of Technology, Taipei, Taiwan, R.O.C., pp. 121-124, November 20-21, 1998.
17. J. C. Chou and C. N. Hsiao, "The hysteresis and drift effect of hydrogenated amorphous silicon for ISFET sensor", Proceedings of Technical Digest of the 7th International Meeting on Chemical Sensors, Beijing International Convention Center, Beijing, China, pp. 553-555, July 27-30, 1998.
18. Luc Bousse, Dean Hafeman and Nancy Tran, "Time-dependent of the chemical response of silicon nitride surfaces, Sensors and Actuator B, **1**, pp. 361-367, 1990.
19. P. Woias, L. Meixner, P. Frostl, "Slow pH response effects of silicon nitride ISFET sensors", Sensors and Actuator B, **48**, pp. 501-504, 1998.
20. J. C. Van Kerkhof, W. Olthuis and P. Bergveld, "Tungsten trioxide (WO_3) as an actuator electrode material for ISFET-based coulometric sensor actuator systems", Sensors and Actuators B, **3**, pp. 129-138, 1991.

* Correspondence: Email: Choujc@pine.yuntech.edu.tw ; Telephone: (8865) 5342601 Ext. 2500, 4333
Fax: (8865) 5312029, 5312063

Development in a High-Stability of an Optical Fiber Refractometer Using Path-Matching Differential Interferometry

Yu-Lung Lo *, Hsin-Yi Lai, Wern-Cheng Wang

Department of Mechanical Engineering

National Cheng Kung University, Tainan, Taiwan 70101, ROC

ABSTRACT

The paper presents a new approach to develop an optical fiber refractometer. The objective of the study is to come up with a relatively inexpensive but reliable optical refractometer that can be used to measure the change of refractive index in a resolution of 10^{-5} and to work in a dynamic range up to 6×10^{-3} at a DC frequency up to 100 Hz. It is known that the phase modulations of optical fiber sensors are very sensitive to external disturbances, especially to the effects of thermal drifts or vibrations. A cancellation technique to compensate the effect of variation on a PZT stack is proposed in this paper to stabilize the system. Two parallel Fabry-Perot sensing cavities corresponding to two path-matching cavities for read-out systems are employed to form path-matching differential interferometries (PMDI). One Fabry-Perot cavity is used as sensing head, and the other as reference sensor. As a result, the experimental data show that the changed of refractive index of a so designed sensing system can be kept in at the level of 10^{-4} without any serious variations even for a three-hour long-term monitoring. Accordingly, the proposed new system can be easily implemented and used as a long-term monitoring system in medical care environment.

Keywords: Refractometer, White-Light Interferometry, and Optical Fiber Sensors

1. INTRODUCTION

The refractive indices of liquids are conventionally measured by using an Abbe refractometer and/or some similar device[1]. However, in the past decades, several different interferometries have been proposed and implemented for use as refractometers[2, 3]. Moosmuller and Arnott[4], introduced a folded Jamin interferometer that consists of two optical elements. This system was found not quite sensitive to the variation in movement. In addition to traditional bulk optics, integrated optical sensors were also proposed[5] using the principle of evanescent wave sensing in an integrated Mach-Zehnder interferometer.

Research and development in optical fibers has generated lots of research interest in recent years. This is attributed to the fact that sensors posses several advantages over conventional electrical transmitters. These distinct advantages include inherent immunity to electromagnetic interference, safety in hazardous or explosive environments, high sensitivity, and long-distance remote measurements. The miniature size, intrinsic safety, and ease of installation of fiber-optic based sensors make the system ideal for applications in various engineering areas including numerous in-line chemical, food, beverage, or medical analysis and monitoring operations. Moreover, some researchers further employed the aforementioned optical fiber sensing techniques to develop refractometers. Takeo and Hattori[6, 7] proposed refractometers by measuring the attenuation of light guided by optical fibers within the liquid. They also applied the refractometer thus built to monitor the state of a skin hydration. Recently, Asseh[8] presented a fiber Bragg grating refractometer using an evanescent filed refractive index fiber sensor that comprise a 42 mm Bragg grating in an etched fiber. However, none of them have come up with a refractometer that can be used for a long-term monitoring. To the author's knowledge, this paper presents a newly developed fiber-optic based refractometer based on PMDI, also

* Correspondence: E-mail: loyl@mail.ncku.edu.tw; Telephone: 8866-2757575 ext. 62123; Fax: 8866-235-2973

Part of the SPIE Conference on Advanced Photonic Sensors and Applications
Singapore • November–December 1999
SPIE Vol. 3897 • 0277-786X/99/$10.00

especially on developing a highly stable measurement system. The theory and procedure of an optical fiber sensing system was developed to implement the phase-modulated stable optical fiber sensing systems are detailed. The environmental factors that affect the sensitivity of the proposed optical fiber sensing system and the associated compensation techniques are also addressed.

Rao and Jackson[9] developed a single-mode fiber-based ultrahigh pressure remote sensor based upon a Fizeau cavity using dual-wavelength coherence reading technique and a built-in temperature compensator. Lo[10] and Lo et al.,[11] presented several cancellation methods for compensating the variation of intensity in quadrature signals for Fabry-Perot sensors in PMDI[12], and for Bragg grating sensors, respectively.

This paper proposes a system that can be used to construct a high stability and high sensitivity optical fiber refractometer based on a white-light interferometry. The change of refractive index in a mixture of de-ionized water and ethanol is used to illustrate the effectiveness of the proposed system by comparing the results with the measurement from conventional Abbe refractometer. It shows a good correlation between them. Based on a proposed cancellation technique, disturbances induced by thermal drifts or vibrations in the read-out system are compensated.

2. OPTICAL FIBER REFRACTOMETERS

The optical setup arrangement and the implementation of a demodulation scheme for measuring the change of refractive index for use in long-term monitoring of an ongoing system are detailed in this section. A two-beam interferometry used in the sensors and in the read-out systems is for PMDI. The coherent length of the light source, L_C, is chosen to be less than the cavity lengths. If the lengths of sensing and read-out cavities are denoted as L_S and L_R respectively, $L_S - L_R$ is adjusted to be less than the coherent length. By doing so, the intensity, I, can be estimated as[12]

$$I \cong A + B\cos(2k_0(L_S - L_R))\exp[-\sigma^2(L_S - L_R)^2]; \tag{1}$$

where A and B are constants, σ is the spectral width of the light source that is one half the distance between two points of which the spectral density equals to the ratio of the peak value to e, and k_0 is the wavenumber. It is obvious from above equation, the intensity, I, reaches the maximum degree of visibility while $L_S \cong L_R$. This is the condition for a path matching in PMDI.

Fig.1 illustrates a schematic diagram of the proposed sensing system based on the technique of PMDI with compensations. The main source of phase noise is thermal drifts or induced vibrations on the PZT stack. The PZT stack is used for demodulating the phase signal in the single-channel phase tracker[13]. To compensate the effect of variations in read-out system, two-parallel Fabry-Perots are implemented in PMDI. One is used as the sensor for sensing the change of refractive index, and the other is used as the reference sensor formed by an in-line fiber etalon (ILFE). As illustrated in Fig.1, two read-out cavities are aligned together with the PZT stack to form the read-out systems. This arrangement is specifically designed to compensate the effect of thermal drift and vibration on the PZT stack.

Fig.2 shows a reference ILFE sensor that is composed of a hollow core fiber fused in between two standard single mode optical fibers[14]. The air gap between the interfacial surfaces of two fibers approximately in between 30 and 400 μm acts as a low finesse Fabry-Perot cavity. Since an ILFE sensor has low thermal apparent strain, thermal insensitivity makes it very suitable for compensating thermal drifts and/or vibrations on a PZT stack. Thus, the phase change of an ILFE Fabry-Perot reference sensor in PMDI shown in Fig.1 can be expressed as

$$\begin{aligned}\Delta\phi_{ILFE} &\cong \frac{4\pi}{\lambda_0}[n_{air}L_{ILFE} - n_{air}(L_{PZT_I} + \Delta L_{PZT_I})];\\ &\cong \frac{4\pi}{\lambda_0}(-\Delta L_{PZT_I});\end{aligned} \tag{2}$$

where λ_0 is the wavelength of the launching light, the refractive index of air, n_{air}, is equal to 1, L_{ILFE} and L_{PZT_I} are the cavity lengths of an ILFE sensor and the corresponding read-out cavity, and ΔL_{PZT_I} is the variation in length caused by PZT thermal drifts or vibrations. In PMDI, the cavity lengths of an ILFE sensor and the corresponding read-out cavity in L_{ILFE} and L_{PZT_I} are of path matching. It is worth noting that an ILFE sensor is thermally stable and isolated. Therefore, the variation from the PZT stack of read-out cavities is the only effect that needs to be considered.

The sensing head, another Fabry-Perot cavity, is formed by two optical fibers and placed on the V-groove substrate, as shown in Fig.3. To eliminate the thermal drifting effect in the sensing head, the V-groove substrate is made of ceramic or glass. The coefficient of thermal expansion is then limited to 5 x 10^{-6}/°C. While the refractive index of the measurand is changed in the sensing head, the corresponding phase change in Fig.1 can be estimated by

$$\begin{aligned}\Delta\phi_S &\cong \frac{4\pi}{\lambda_0}[(n_S + \Delta n_S)L_S - n_{air}(L_{PZT_S} + \Delta L_{PZT_S})];\\ &\cong \frac{4\pi}{\lambda_0}(\Delta n_S L_S - \Delta L_{PZT_S});\end{aligned} \tag{3}$$

where n_S is the original refractive index in a sensing head, Δn_S is the change of refraction index, L_S is the cavity length of the sensing head, and ΔL_{PZT_S} is the variation in length introduced by the effects of thermal drifts and/or vibrations in PZT. Here, $n_s L_S$ and L_{PZT_S} are in optical-path balance in terms of PMDI. As compared with Eq.(2), it is can be seen that ΔL_{PZT_I} and ΔL_{PZT_S} in Eqs.(2) and (3) are variations caused by the thermal drifts and/or vibrations from the PZT stack. This implies that ΔL_{PZT_I} is equal to ΔL_{PZT_S}. As the phase change in Eq.(2), it is subtracted from that of Eq.(3), and the effects of thermal and/or vibration on the PZT stack can be canceled from each other. Consequently, the result of cancellation can be obtained by estimating

$$\Delta\phi_S - \Delta\phi_{ILFE} \cong \frac{4\pi}{\lambda_0}(\Delta n_S L_S). \tag{4}$$

Thus, the phase change induced by the thermal drifts and/or vibrations in PZT is removed in essence. The stability of the sensing system in PMDI can be improved by using two parallel Fabry-Perot sensors for a subtraction operation. It is noteworthy that the two cavity lengths in two parallel Fabry-Perot cavities, Ls and L_{ILFE}, are required to be kept constant to prevent signal errors from contaminating in the refractive index measurements. If the resolution of 10^{-5} is expected in an optical fiber refractometer, the coefficient of thermal expansion for the substrate of a sensing head will have to be around 5 x 10^{-6}/°C. This is estimated at the level of temperature variation within 1 °C.

3. EXPERIMENTAL SETUP AND RESULTS

As illustrated in Fig.1, the broadband light (FWHM ~ 50 nm) of the pigtailed super luminescent diode (Model: MRV MREDSP010-1) has a nominal wavelength 1.3 μm and the optical power is 150 μW. The receiver (Model: New Focus 2011) has a noise-equivalent power (NEP) < 1 pW/$\sqrt{Hz}$. The light travels through a fiber-optic coupler to two parallel Fabry-Perot cavities. One Fabry-Perot cavity is used as the sensing head for sensing the change of refractive index, and the other, ILFE, is used as a reference sensor. The reflected lights of the two parallel Fabry-Perot cavities again travels through the fiber-optic coupler to form low finesse Fabry-Perot cavities with a mirror. The mirror is bonded to a PZT stack that is placed on a linear transition plate. The PZT stack is used to provide a phase carrier for demodulation. The translation plate is used to adjust the optical path of the read-out cavities to match with that of the sensing head and the ILFE. The cavity lengths of the read-out system corresponding to the sensing head and the ILFE are 133 μm and 100 μm, respectively. The length difference is 33 μm that is longer than the coherence length of SLD broadband light source of 20 μm[15]. This configuration is used to test the measurement capacity of the sensing head for the change of refractive index change by driving the PZT stack with a ramp function in a single-channel phase tracker[13] for demodulation. This scheme is based on an electronic feedback phase nulling technique[13] using an AD639 trigonometric

simulator chip. By driving a PZT stack with a 6 kHz sinusoidal carrier frequency and selecting 800 Hz as the cutoff frequency of the low-pass filter (usually is around 0.2 time carrying frequency), the sensitivity of a sensing system can be improved in a low frequency response. Subsequently, several tests are conducted to characterize the performance of demodulation in the single-channel phase tracker. By driving a PZT from DC to 500 Hz in various depths of modulation, the behavior of the demodulation system can be characterized. The minimum detectable change of refractive index is approximately around 10^{-5} in a dynamic range up to 6×10^{-3} at the frequency response from DC to 100 Hz.

A long-time test is conducted by inserting a sensing head into the de-ionized water in an isothermal environment. Fig.4 shows that the variation of the compensated signal in the sensing system is limited to be within the range of 10^{-4} in terms of the change of refractive index as compared to that of original signal without compensation for a test period of three hours. The fluctuation can be explained as the electric noise effects in circuit. It can be seen that the variation of the original signal without compensations is uncontrolled. In addition to the phase variation associated with thermal drifts and/or vibrations on the PZT stack, a phase noise from the broadband source can be explained as a main source that causes the variation of the sensing system. Nevertheless, as compared to the long coherence light source, the phase noise from the broadband source is much smaller since the optical path-imbalance in Eq.(1) is small while the path-matching condition is satisfied. Therefore, for long-term monitoring, the use of a white-light interferometry does not require high stability of a light source that is generally a necessity in an interferometry based on a long coherence source.

Mixing de-ionized water with ethanol solution is selected for testing the implemented optical fiber refractometer. The change of refractive indices for various mixtures of de-ionized water and ethanol are measured. The results are given in Fig.5. A linear regression, as represented by the solid line in Fig.5, returns an excellent correlation between the change of refractive index and the mixing composition. The standard deviation of the measurand, σ, is equal to 2.33×10^{-5}. As compared with experimental data conducted by using the traditional Abbe refractometer, slight differences exist. These differences are caused by different light sources used in both systems. The launching light used in an Abbe refractometer is around 638 nm in wavelength that is somewhat different from the wavelength of the light source used in our proposed system.

4. CONCLUSIONS AND DISCUSSIONS

An optical fiber based refractometer that measures the change of refractive index in a resolution up to 10^{-5} for the dynamic range of 6×10^{-3} in DC frequency to 100 Hz is developed in this paper. A cancellation method is proposed and implemented for use in compensating the effect of thermal drifts and/or vibrations on a PZT stack so that the long-term monitoring of the change of refractive index can be accomplished. The use of two parallel Fabry-Perots in PMDI to allow the cancellation of phase change is proved to be efficient in compensating phase variations on the PZT stack driven by a function generator. It should be noticed that the cancellation technique is applied to compensate thermal drifts and/or vibrations on a PZT stack in the read-out system only.

Because the distinct features such as small cavity in sensing head and high robustness are very attractive, the device is very suited for various applications where small size, high mechanical stability, and small analyte volumes are substantially required.

ACKNOWLEDGMENTS

The author would like to acknowledge Dr. J.S. Sirkis, department of mechanical engineering at the University of Maryland, College Park, for his support on hollow-core fibers.

REFERENCES

1. Longhurst, R.S., *Geometrical and Physical Optics*, 3rd ed. Longmans, London, 1962.
2. Webb, D.J., Tatam, R.P., and Jackson, D.A., "A Novel Interferometric Liquid Refractometer," Vol. 60, pp. 3347-3348, 1989.
3. Fujii, K.I., Williams, E.R., Steiner, R.L., and Newell, D.B., "A New Refractometer by Combining a Variable

Length Vacuum Cell and a Double-Pass Michelson Interferometer," IEEE Transactions on Instrumentation and Measurement, Vol. 46, pp. 191-195, 1997.
4. Moosmuller, H. and Arnott, W.P., "Folded Jamin Interferometer: a Stable Instrument for Refractive-Index Measurements," Optics Letters, Vol. 21, pp. 438-440, 1996.
5. Maisenholder, B., Zappe, H.P., Kunz, R.E., Moser, M., and Riel, P., "Optical Refractometry Using a Monolithically Integrated Mach-Zehnder Interferometer," Transducers '97, International Conference on Solid-State Sensors and Actuators, Chicago, pp. 79-80, 1997.
6. Takeo, T. and Hattori, H., "Optical Fiber Sensor for Measuring Refractive Index," Japanese Journal of Applied Physics, Vol. 21, pp. 1509-1512, 1982.
7. Takeo, T. and Hattori, H., "Skin Hydration State Estimation Using a Fiber-Optic Refractometer," Applied Optics, Vol. 33, pp. 4267-4272, 1994.
8. Asseh, A., Sandgren, S., Ahlfeldt, H., Sahlgren, B., Stubbe, R., and Edwall, G., "Fiber Optical Bragg Grating Refractometer," Fiber and Integrated Optics, Vol. 17, pp. 51-62, 1998.
9. Rao, Y.J. and D.A., Jackson, "Prototype Fiber-Optic-Based Ultrahigh Pressure Remote Sensor with Built-In Temperature Compensation," Rev. Sci. Instrum., Vol. 65, pp. 1695-1698, 1994.
10. Lo, Y.L., "Intensity Compensation for In-Fiber Fabry-Perot Sensors in High Disturbed Environments," IEE, Electronics Letters, Vol. 34, No. 4, pp. 394-395, 1998.
11. Lo, Y.L., Huang, J.F., and Yang, M.D., "Intensity Effects in Bragg Grating Sensors Scanned by a Tunable Filter," Smart Structures and Materials SPIE Conference, Sensory Phenomena and Measurement Instrumentation for Smart Structures and Material, Vol. 3670, pp. 65-73, 1999.
12. Cielo, P.G., "Fiber Optic Hydrophone: Improved Strain Configuration and Environmental Noise Protection," Applied Optics, Vol. 18, pp. 2933-2937, 1979.
13. Kersey, A.D., Moeller, R.P., Berkoff, T.A., and Burns, W.K., "Single Channel Phase-Tracker for the Open Loop Fiber Optic Gyroscope," SPIE, Vol. 1585, pp. 198-202, 1991.
14. Sirkis, J.S., Berkoff, T.A., Jones, R.T., Singh, H., Kersey, A.D., Friebele, E.J., and Putnam, M.A., "In-Line Fiber Etalon (ILFE) Fiber-Optic Strain Sensors," J. Lightwave Technol., Vol. 13, pp.1256-1263, 1995.
15. Lo, Y.L., "Study of Cross-Talk on Parallel Fabry-Perot Sensors in Path-Matching Differential Interferometry (PMDI)," Accepted by Optics and Lasers in Engineering, 1999.

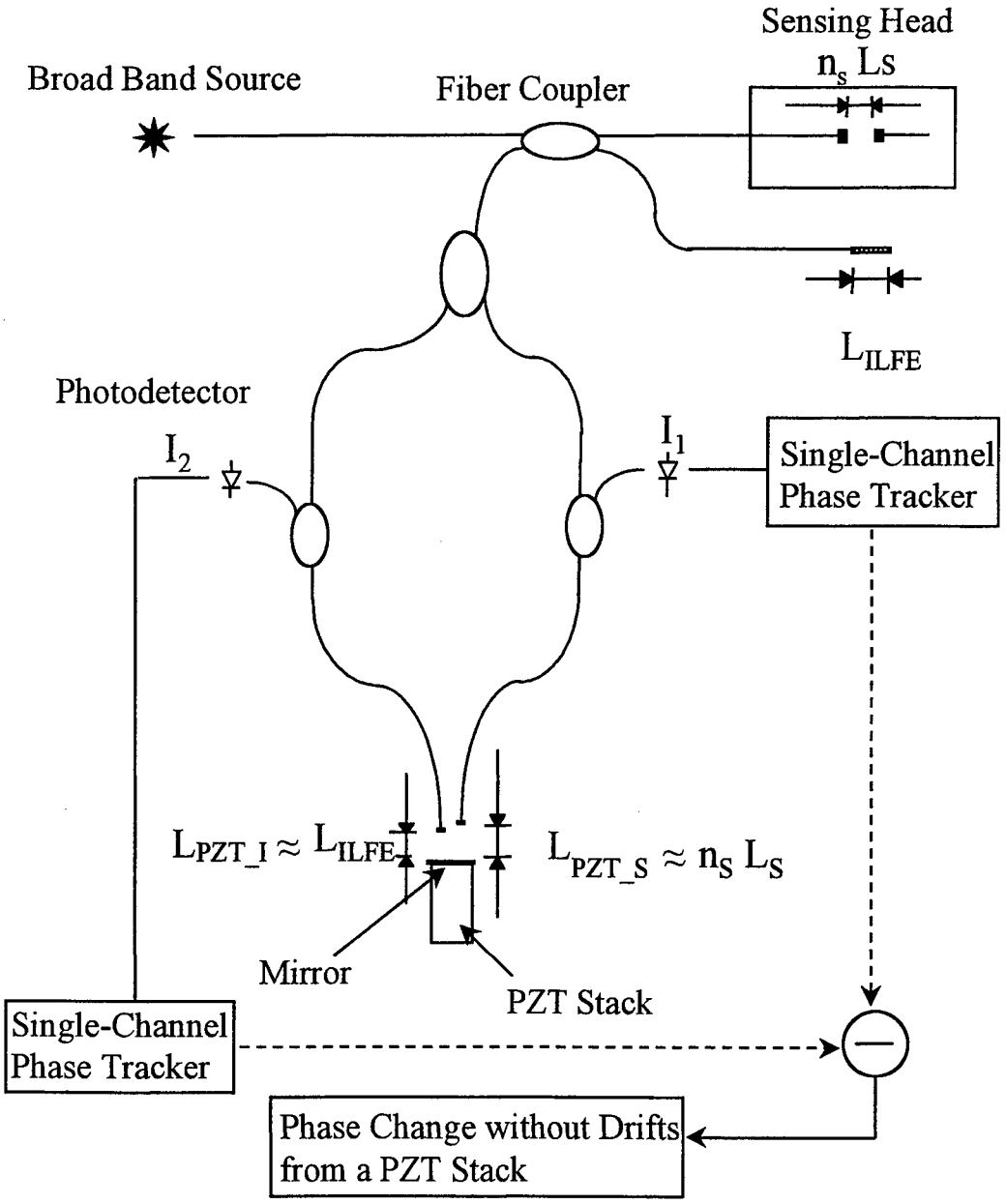


Figure 1 Schematic Diagram of Setup

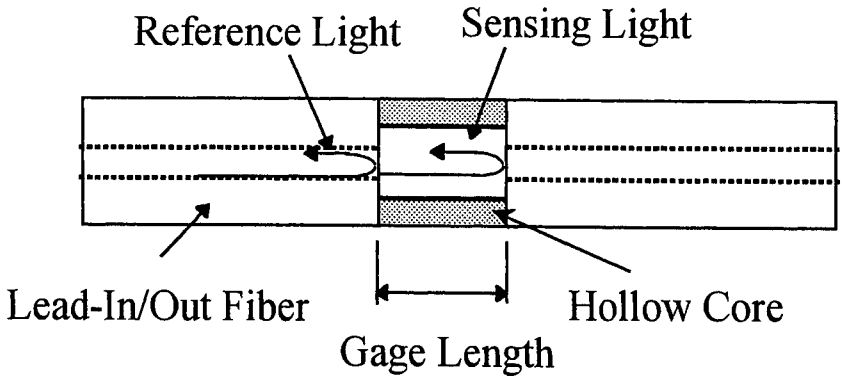


Figure 2 ILFE Sensor

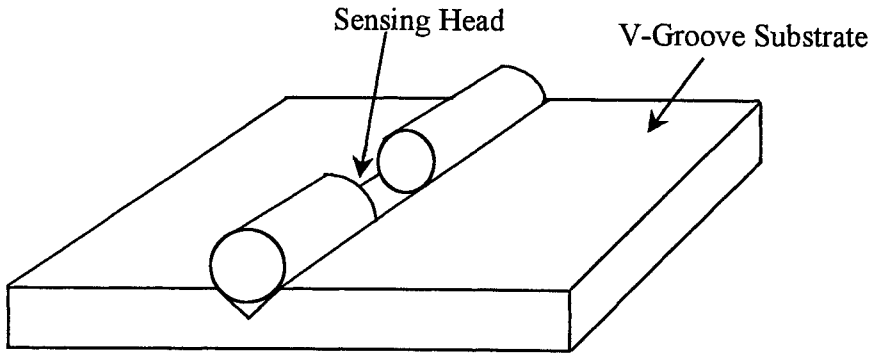


Figure 3 Sensing Head

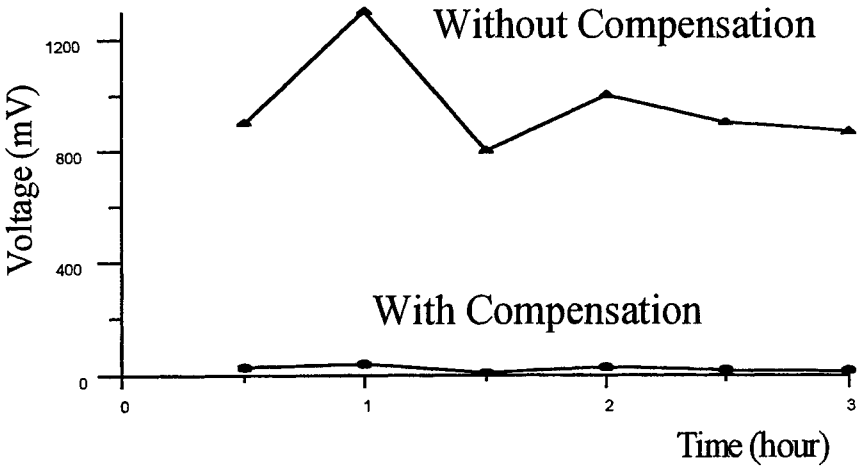


Figure 4 Long Time Monitoring

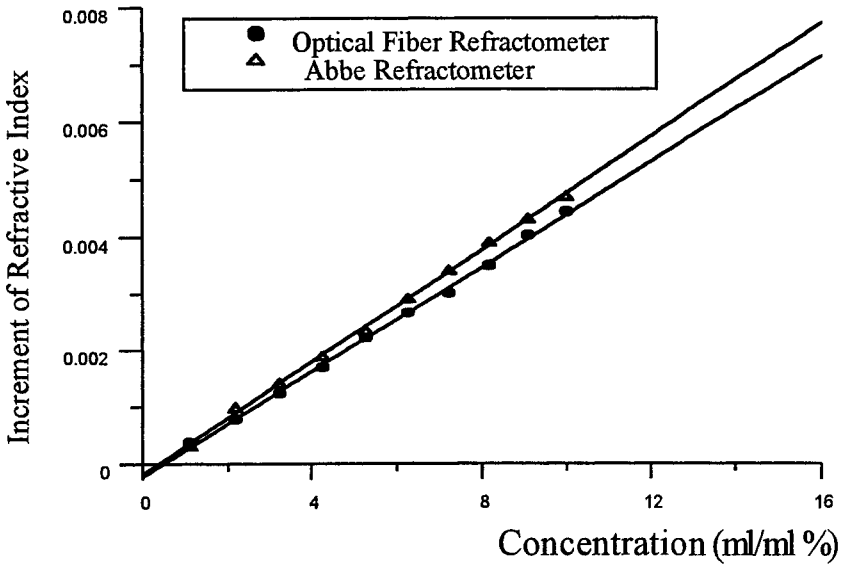


Figure 5 Comparisons between Abbe and Optical Fiber Refractometers

SESSION 16

Poster Session

Achromatic design strategies with diffractive optical elements

Zhao Liping*, Lam Yee Loy, Zhou Yan and Yun Zhisheng
Photonics Research Group, School of Electrical and Electronic Engineering
Nanyang Technological University, Nanyang Avenue, S639798, Singapore

ABSTRACT

In both design and fabrication, diffractive optical elements (DOEs) are more flexible and powerful than traditional refractive/reflective optical elements, hence an optical system with one or more diffractive elements may provide better optical performance at lower cost with small, light and compact structure. However due to its inherent large spectral dispersion, a DOE is generally designed and fabricated for one specific wavelength or a narrow spectral bandwidth. In the case of a wide band light source, chromatic aberration and loss of diffraction efficiency will occur. In this paper, the chromatic dispersion of DOE is discussed, and four achromatic strategies, namely hybrid diffractive/refractive strategy, harmonic diffraction strategy, multi-material strategy, and bi-blazed strategy are introduced and analyzed respectively. A comparison has also been made among them to guide the application.

Keywords: diffractive optical element, achromatic design, spectral dispersion

1. INTRODUCTION

Binary optics is a surface-relief optics technology based on VLSI fabrication techniques (primarily photolithography and etching) with the "binary" in the name referring to the binary coding scheme used in creating the photolithographic masks[1]. It has been widely recognized that diffractive optics, generally denoted as binary optics, allows the creation of new, unconventional optical elements as well as provides greater design freedom and new materials choices for conventional optical elements. In fact, diffractive optical elements, including mainly binary optical elements, holographic elements (especially computer-generated holograms), and gratings, have already made it feasible for optical designers to create innovative components that can solve problems in optical sensors, optical communications, and optical processors. Over the decade, binary optics has advanced sufficiently to allow the production of different kinds of diffractive elements, hybrid refractive-diffractive elements, and refractive micro-optics, which are satisfactory for use in cameras, military systems, medical applications and other demanding areas.

However, DOEs inherently suffer from large chromatic dispersion, the wavelength-dependent focal length as shown in Equ.(1) is a prime example. In the visible band, a DOE with a designed focal length f can cause an axial chromatic aberration as large as *f/3*.

$$f(\lambda) = f_0 \frac{\lambda_0}{\lambda} \tag{1}$$

DOEs themselves are unsuitable for broadband imaging, and it has been shown that an optical system consisting only of diffractive lenses cannot produce a real image. In this paper, we will first discuss the special chromatic dispersion of DOE; we will then introduce some practical design strategies, namely hybrid refractive/diffractive strategy, harmonic diffraction strategy, multi-material strategy, and bi-blazed strategy respectively; and finally we will make a comparison among these different schemes.

2. CHROMATIC DISPERSION OF DOE

The Sweat model[2], an approximation to the grating equation, models a diffractive optics element as an equivalent refractive element. It allows results derived for refractive optics to be applied to diffractive optics. In the model, a binary optics element with phase $\varphi(x,y)$ at wavelength λ_0 is replaced by a refractive equivalent with thickness and refractive index given by

*Correspondence:
Zhao Liping: Tel: (65)7905985-36, Fax: (65)7912687, e-mail: elpzhao@ntu.edu.sg

Part of the SPIE Conference on Advanced Photonic Sensors and Applications
Singapore • November–December 1999
SPIE Vol. 3897 • 0277-786X/99/$10.00

$$t(x,y) = \frac{\lambda_0}{n_0 - 1}\frac{\varphi(x,y)}{2\pi} + t_0 \tag{2}$$

$$n(\lambda) = \frac{\lambda}{\lambda_0}(n_0 - 1) + 1 \tag{3}$$

In the above equations, t_0 is a constant chosen to make $t(x,y)$ always positive and n_0 is the index at the material wavelength λ_0. As $n_0 \to \infty$, the Sweat model approaches the grating equation. DOEs' chromatic behavior can be understood by using this model, which states that a diffractive optical lens behaves like an ultrahigh index refractive lens with an index which varies linearly with wavelength (let $n_0 \to \infty$ in Equ.(3)). In practice value of $n_0 = 10,000$ is sufficiently high for accurate results[3]. For the design of achromats and apochromats, using the C,d and F lines, the effective Abbe number of DOE is

$$V_d = \frac{n_d - 1}{n_F - n_C} = \frac{\lambda_d}{\lambda_F - \lambda_C} = -3.45 \tag{4}$$

Similarly, using the g and F lines, the effective partial dispersion will be

$$P_{gF} = \frac{\lambda_g - \lambda_F}{\lambda_F - \lambda_C} = 0.296 \tag{5}$$

Fig.1 plots the partial dispersion P_{gF} versus Abbe number V_d for several commonly used glasses. From it we can see that DOE has a special chromatic dispersion with a negative Abbe number and a P_{gF} location far away from the other glasses'.

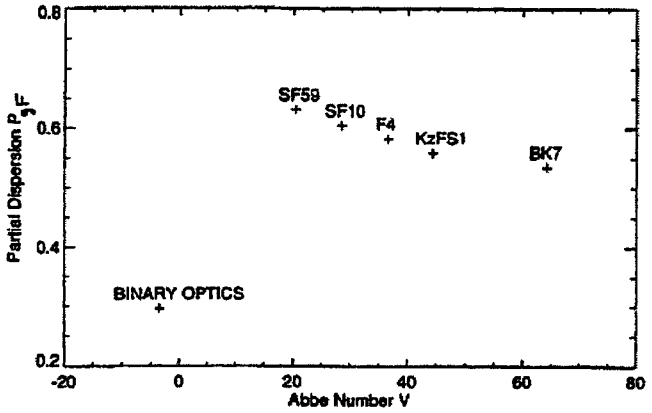


Fig.1 Partial dispersion P_{gF} vs Abbe number V

3. HYBRID REFRACTIVE/DIFFRACTIVE STRATEGY

Usually the phase function of a rotationally symmetrical DOE is expressed as follows:

$$\varphi(\rho) = 2\pi(A_1\rho^2 + A_2\rho^4 + \cdots) \tag{6}$$

Then the paraxial optical power ϕ_{dif} of the DOE depends on the coefficient A_1 through the following equation

$$\phi_{dif} = -2m\lambda A_1 \tag{7}$$

where m is the diffraction order (m is always set to 1 as the first order is used most frequently). The paraxial optical power ϕ_{dif} is proportional to the wavelength λ, and an appropriate design of A_1 can correct chromatic aberration. T. Stone and N. George[4] followed the achromatic method used in the conventional system, and introduced the well-known formula for the division of optical power in an achromatic hybrid refractive/diffractive singlet as shown in Fig.2.

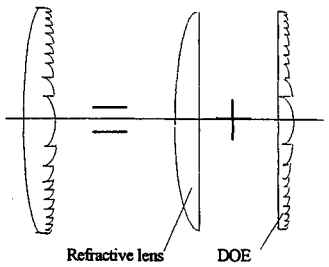


Fig.2 the schematic construction of hybrid achromatic singlet

$$\begin{cases} f'_{dif} = F'(\dfrac{V_{dif} - V_{ref}}{V_{dif}}) \\ f'_{ref} = F'(\dfrac{V_{ref} - V_{dif}}{V_{ref}}) \\ \phi = \phi_{dif} + \phi_{ref} \end{cases} \tag{8}$$

where ϕ is the total optical power of the hybrid achromatic singlet, and ϕ_{ref} is the optical power of the refractive element. For BK7 glass, V_{ref}= 64.0 then

$$\begin{cases} f'_{dif} = 19.540F' \\ f'_{ref} = 1.0540F' \end{cases} \qquad \begin{cases} \phi_{dif} = 0.051\phi \\ \phi_{ref} = 0.949\phi \end{cases} \tag{9}$$

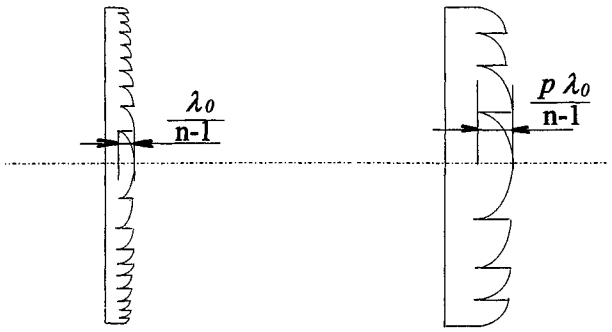
According to this division, DOE not only can correct the axial chromatic aberration, but also can share certain optical power as well. If it were not for DOE, the positive lens would have a positive optical power at least equal to or even greater than the total power. By using the DOE, the burden on the positive refractive lens is lessened, and as a result high order aberrations will be reduced greatly and the structure simplified.

If the coefficient A_2 of a DOE is used to correct the spherical aberration, a significant spherochromatism will be introduced in the mean time. In order to limit the spherochromatism, the axial chromatism must be overcorrected. So the division of the optical power should be upgraded as follows[5]

$$\begin{cases} \dfrac{\phi_{dif}}{v_{dif}} + \dfrac{\phi_{ref}}{v_{ref}} = \dfrac{1}{l'^2}\Delta l_{FC}' \\ \phi_{dif} + \phi_{ref} = \phi \end{cases} \tag{10}$$

4. HARMONIC DIFFRACTION STRATEGY

Harmonic diffractive optical element, also called high-order diffractive element, is characterized by its multiple 2π or generally stated $p*2\pi$ (p, harmonic coefficient, is a positive integer, generally $p \geq 2$) phase jumps at its zone boundaries. So as shown in Fig.3, the depth of the microstructure of harmonic DOE is p times that of the ordinary DOE. (in Fig. 3 the left is an ordinary diffractive optical lens (DOL) and the right is a harmonic diffractive lens (HDL)).



a) Ordinary module 2π DOE b) HDOE

Fig. 3 Comparison of two kinds of DOE and HDOE

The diffraction light distribution of a DOE in its focal plane actually is the interference result of the beams passing through all the phase zones. If the design wavelength is λ_0 , then the first order focal length of the ordinary DOE for the wavelength λ is:

$$f_{1,\lambda} = \frac{\lambda_0}{\lambda} f_0 \tag{11}$$

Similarly for a HDOE, since there are $p*2\pi$ phase jumps between two zone boundaries, its p-th order focal length for its design wavelength $p\lambda_0$ is f_0 . For wavelength λ , the m-th order focal length is:

$$f_{m,\lambda} = \frac{p\lambda_0}{m\lambda} f_0 \tag{12}$$

Providing that the following formula comes into existence, the design focal length f_0 will coincide with $f_{m,\lambda}$

$$\frac{p\lambda_0}{m\lambda} = 1 \tag{13}$$

Therefore the harmonic or resonant wavelength λ is obtained as

$$\lambda = \frac{p\lambda_0}{m} \tag{14}$$

This specifies that the coherent light with any harmonic wavelengths will converge to the common focus f_0. The values of these harmonic wavelengths can be determined by equation (14). The larger p is, the more harmonic wavelengths are there in the employed spectrum range. Therefore, p is a new design freedom, and can be used to make the light of several separate wavelengths converge to the same point.

According to the concept of harmonic diffractive elements and the derivation process of the monochromatic aberrations of the DOE, it is easily known that the monochromatic aberrations of HDOE are the same as those of the simple DOE.

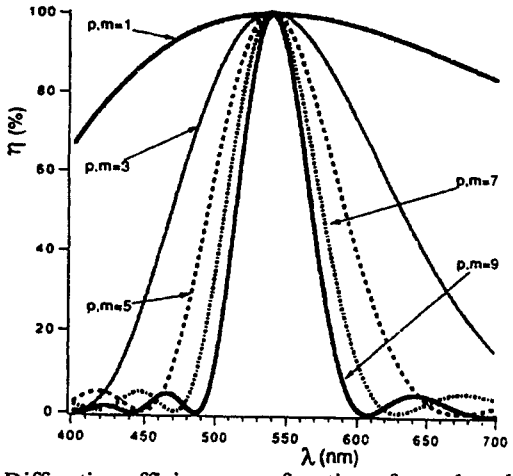


Fig.4 Diffraction efficiency as a function of wavelength for $p = m$

The m-th order diffraction efficiency of HDOE can be expressed as[6]

$$\eta_m = \sin c^2 \left\{ \frac{\lambda_0}{\lambda} \left[\frac{n(\lambda)-1}{n(\lambda_0)-1} \right] p - m \right\} \tag{15}$$

The wavelength bandwidth of the diffraction efficiency around a given diffracted order narrows with increasing values of p as shown in Fig.4. Physically, we are trading between the thickness of the diffractive lens and the achromatic performance. The important point is that at any resonant wavelength the focus is diffraction limited. At intermediate wavelengths, the focus has an axial spread that is symmetric about the nominal focal plane. Fig.5 is another visualization of the diffracted field for specific wavelengths and optical power. The lens[7] was designed with p=15 for a design wavelength of 550nm. Diffracted orders 12 to 21 (i.e. m values) contribute to the focus across the visible, and for a simple DOL the diffracted energy would be (to the first order) distributed along the dashed line labeled p=1.

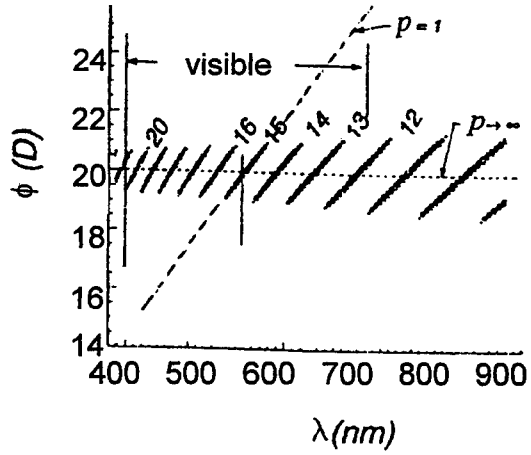


Fig.5 Distribution of energy for the p=15 harmonic diffractive lens.

5. MULTI-MATERIAL STRATEGY

The multi-material strategy sees an achromatic DOE as an array of distribution pixels[8]. The pixel is composed of two DOEs made from different dispersive materials, as depicted in Fig.6. The suitably designed thickness (d_1, d_2) at each pixel

provides the exact phase value for the design wavelength λ_0, the middle of a selected spectrum, with minimum phase aberration over the entire wavelength range. The required phase retardation φ of light propagating through the combined DOE for λ_0 is:

$$[n_1(\lambda) - n_2(\lambda)]d_1 + [n_3(\lambda) - n_2(\lambda)]d_2 = \lambda\varphi/2\pi \tag{16}$$

So the condition for which the optical phase is stationary with respect to wavelength at a designated wavelength can be described as:

$$\frac{d\varphi}{d\lambda} = 0 \tag{17}$$

then solving equ.(1) and (2) for d_1 and d_2 gives:

$$\begin{aligned} d_1 &= (\lambda\varphi/2\pi) f_1(n_1, n_2, n_3, \frac{dn_1}{d\lambda}, \frac{dn_2}{d\lambda}, \frac{dn_3}{d\lambda}) \\ &= (\lambda\varphi/2\pi)[(n_1 - n_2) + (n_3 - n_2)(n_2 - n_1 - \lambda\frac{dn_2}{d\lambda} + \lambda\frac{dn_1}{d\lambda})/(n_3 - n_2 - \lambda\frac{dn_3}{d\lambda} + \lambda\frac{dn_2}{d\lambda})]^{-1} \end{aligned} \tag{18}$$

$$\begin{aligned} d_2 &= (\lambda\varphi/2\pi) f_2(n_1, n_2, n_3, \frac{dn_1}{d\lambda}, \frac{dn_2}{d\lambda}, \frac{dn_3}{d\lambda}) \\ &= (\lambda\varphi/2\pi)[(n_3 - n_2) + (n_1 - n_2)(n_3 - n_2 - \lambda\frac{dn_3}{d\lambda} + \lambda\frac{dn_2}{d\lambda})/(n_2 - n_1 - \lambda\frac{dn_2}{d\lambda} + \lambda\frac{dn_1}{d\lambda})]^{-1} \end{aligned} \tag{19}$$

In the above two equations, if we set $n_1 = n_3 = 1$, then we can get that $d_1 = d_2 = \infty$. That means using one material one can not get this kind of achromatic DOE. If we build the pixel in the way as shown in Fig. (b), and material 2 is air, but materials 1 and 3 are different, then we can get an achromatic doublet DOE.

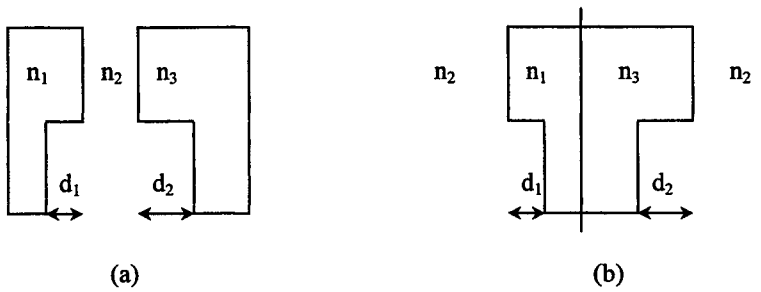


Fig.6 One pixel of the wideband DOE made of two DOE's

6. BI-BLAZED STRATEGY

We have proposed a method to employ bi-blazed diffraction[9] to partially fulfill the large spectral dispersion for wide band light whose wavelength ranges from λ_1 to λ_2 , and $\lambda_1 > \lambda_d > \lambda_2$, where λ_d is the design wavelength. The mechanism and qualitative analysis are briefly illustrated in Fig. 7 and Fig. 8 respectively. All the pictures in this paper are schematic diagrams, and are not drawn quantitatively.

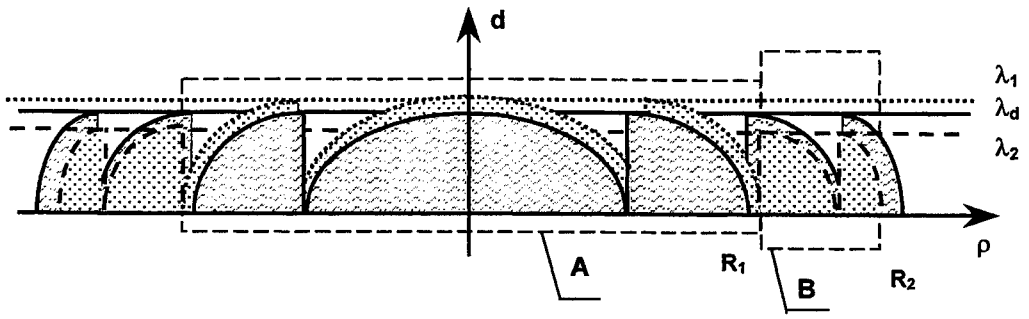


Fig.7 Profile of a bi-blazed DOE composed of two portions with different zone depth

The phase function of a bi-blazed DOL for the design wavelength is designed similarly with that of an ordinary DOL, but its parameters are proportionally changed in different aperture range according to its application bandwidth. Thus the bi-blazed DOL has different structure as shown in Fig.7, in which the profile of an ordinary DOL is illustrated by the solid lines, and that of a bi-blazed one by the broken lines. The micro-zones in the whole aperture are divided into two portions A and B in a staggered way. Because portions A and B are designed respectively for two different wavelengths λ_1 and λ_2, the bi-blazed BOL has a zone depth in A deeper than that of the ordinary DOL designed for λ_d, and a shallower zone depth in B.

When a wide band light is incident on the bi-blazed BOL, either potion A or B will still have its own spectral dispersion of a BOL. The micro profile of the two portions is specially designed to make the two focal points for wavelength λ_1 diffracted from portion A and for λ_2 from portion B coincide with each other, thus the light passing through BOL is diffracted and spectrally distributed as shown in Fig 8. Certainly, the diffraction efficiency from portion A is not constant in the whole spectral band, and its top point is at wavelength λ_1. Similarly, for the portion B the diffraction efficiency at λ_2 is at its peak. Therefore, in terms of energy distribution after diffraction, the two parts with different micro profiles seem to possess different spectral-filtering capabilities for wide band illumination. In such design, although weak dispersion still exists, we can expect to get wide band imaging with much lower chromatic aberrations at the focal plane. We note that this effect is quasi-achromatic. Moreover, the bi-blazed BOL can image with a higher contrast as compared to the ordinary BOL.

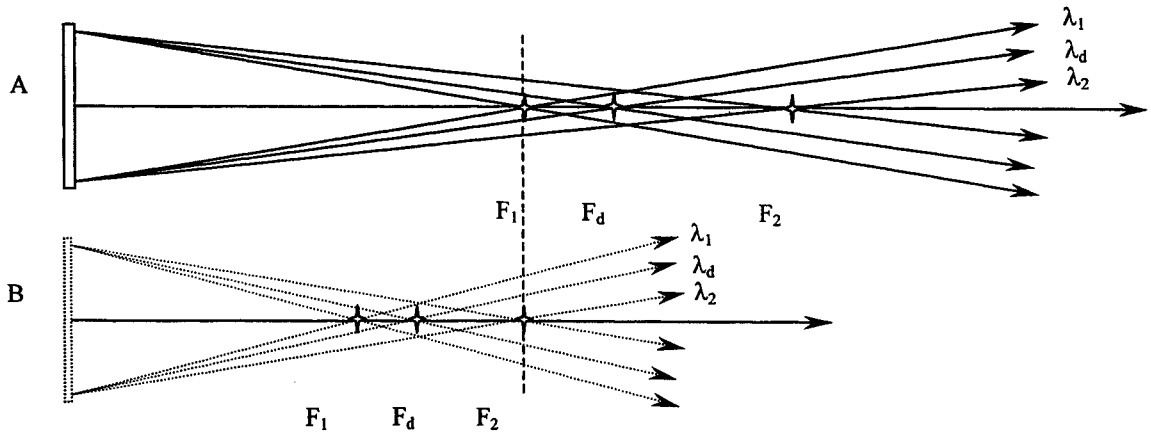


Fig. 8 Spectral dispersion in the propagating direction of a Bi-blazed DOE

7. COMPARISON OF DIFFERENT ACHROMATIC SCHEMES

In the above sections, we have discussed four strategies for the achromatic diffractive optical system design. In order to use them effectively and understand them comprehensively, we describe their differences and similarities in Table 1.

Table 1 Comparison of different achromatic strategies.

Strategy / Difference / Feature	Hybrid refractive/diffractive strategy	High-order diffraction strategy	Multi-material strategy	Bi-blazed strategy
Achromatic efficiency for continuous wavelength band	High	Mid; More suitable for several separate wavelengths	High	Mid; More suitable for several separate wavelengths
Least number of components needed	2	1	2	1
Least number of materials needed	1	1	2	1
Alignment and assemble is necessary or not	Yes	No	Yes	No
Deep-etching is necessary or not	No	Yes	Yes	Possible
Suitable for planar system or not	No	Yes	Yes	Yes

8. CONCLUSIONS

In this paper, we summarized four strategies for designing achromatic optical system with diffractive optical elements by studying their theories and structures, and analyzing their effects respectively. Finally the comparison among the different methods are given to guide their applications in this field.

9. REFERENCES

1. Michael Bass, et al, Handbook of Optics, Vol.I, Chapter 8 "Binary Optics", McGraw-Hill,Inc., 1995
2. W.C.Sweatt, "Mathematical Equivalence between a Holographic Optical Element and an Ultra-high Index Lens", J.O.S.A., 69, 1979, pp.486-487.
3. M.W.Farn, "Quantitative comparison of the general Sweat Model and the Grating Equation", Appl. Opt., 30, 1991, pp.2151-2158.
4. T.Stone and N.George, "Hybrid diffractive-refractive lenses and achromats", App. Opt., 27, 1988, pp.2960-2971.
5. Zhao, Liping. Wu, Minxian. Jin, Guofan. Yan, Yingbai. "Spherochromatism correction of hybrid refractive-diffractive singlet", Acta Optica Sinica. v 18 n 5 1998. p 621-626
6. M.Rossi, R.E.Kunz, and H.P.Herzig, "Refractive and diffractive properties of planar micro-optical elements", Appl. Opt. 34(26), 1995, pp. 6005-6006
7. Sweeney D W, Sommargren G E., " Harmonic Diffractive Lenses", *Applied Optics*, 34(14), 1995, p.2469.
8. Yoel Arieli, Shmuel Ozeri, and Naftali Eisenberg, "Design of a diffractive optical element for wide spectral bandwidth", Opt. Lett., 23(11), 1998, pp.823-824
9. Zhao Liping, Lam Yee Loy, Zhou Yan, and Yun Zhisheng," Design of planar-binary optical visor", SPIE Vol. 3778, 44th Annual Meeting, July 1999, Denver, USA.

Planar diffractive imaging element design

Zhisheng Yun, Yee Loy Lam, Yan Zhou, and Liping Zhao
Photonics Research Group, School of Electrical and Electronic Engineering
Nanyang Technological University, Nanyang Avenue, S639798, Singapore

ABSTRACT

In planar diffractive imaging systems, extra-axial imaging elements are frequently used. Since the elements have large off-axial aberrations such as coma, astigmatism, field curvature and distortion, it seems a little difficult to design a practical aberration-free element. In this paper, we reviewed the axial imaging diffractive element design procedure. Referring to designing axial elements, we present a semianalytical approach that enables one to determine the exact surface profile of an extra-axial element based on geometrical optics according to design and aberration-free requirements. The design procedure of the element can be divided into two steps: firstly to obtain the zone boundaries and then to solve the exact surface profile. Finally, a schematic is given to test and evaluate small size diffractive elements.

Keywords: Optical design, diffractive optics, imaging quality evaluation, optical testing

1. INTRODUCTION

Diffractive elements, popularly known as kinoform lenses and used as imaging lenses, are now finding applications[1]-[8] in traditional optical systems to improve the system imaging quality, they are also showing great potential in various modern photonics systems, such as planar integrated optics[9], integrated optical-disk pickup[10], planar imaging optical system[11], and planar holographic imaging visor[12]. Since most diffractive elements are employed in a system as off-axial imaging elements, how to design and obtain their exact surface profile in order to guide their fabrication is obviously very useful and important.

Generally, the exact surface profile of a diffractive element can be obtained in two ways. One is to use some commercial software design package such as CODE V, OSLO, or ZEMAX to firstly optimize and obtain its aberration-free phase function according to the optical system design requirements such as focal length, view of field, object and image distances, and then, from the obtained phase function, to find out its zone radii and the exact surface profile[13]. The other method is, by referring to the aberration-free condition, to derive and solve its zone radii and the exact surface profile directly using geometrical optics[14].

Although studies have been carried out based on the above two ways to obtain the exact phase profile, most of them were focused on axial stigmatism imaging[15]-[17]. In spite of fact that some work for designing off-axial imaging elements has been done[18][19], more research is still required in studying the off-axial planar imaging performance for different applications.

In the following, we first give a review on how to use the software design package to design and solve the radii and exact surface relief of a diffractive element. Then we calculate and derive the zone boundaries of an off-axial imaging diffractive element from geometric optics. We also report our study on how to obtain the exact surface profile of the extra-axial imaging element is given. Finally, we present a method to evaluate and test a short focal length element.

2. DIFFRACTIVE ELEMENT DESIGN

To design a diffractive element, whether starting from a commercial optical design software or directly from geometric optics, the procedure is usually almost same and can be divided into two steps. Firstly, the zone boundaries of the element (for symmetrical system it is usually a series of radii) need to be determined. Then according to the phase function obtained from the optical design software package or geometrical relation, the exact surface profile needs to be calculated.

2.1 Exact surface-relief profile of a diffractive element from its phase function

Part of the SPIE Conference on Advanced Photonic Sensors and Applications
Singapore • November–December 1999
SPIE Vol. 3897 • 0277-786X/99/$10.00

For simplifying the analysis, we start from an axial imaging diffractive element. Generally, the phase function needed for an axial stigmatism imaging diffractive element can be obtained in two ways: One is to solve the phase function to form a stigmatic image O' of an axial object point at O, i.e., $\phi(r)$ transforms the spherically diverging wave front AA' into a spherically converging wave front BB'[Fig. 1a]. The other is to obtain the phase function $\phi(r)$ for the purpose of eliminating aberrations in a multi-components optical system: the diffractive element with phase function $\phi(r)$ at PP' transforms the nonstigmatic wave front AA' into a stigmatic, spherically convergent wave front BB' [Fig. 1(b)].

In the practical design of an optical system involving surface-relief diffractive lenses using commercially available software, a diffractive element is often represented as a phase function. For a symmetrical imaging optical system, usually we can obtain a symmetrical diffractive phase profile, which can expressed as

$$\phi(r) = \frac{2\pi}{\lambda} \sum_{i=1}^{I} a_i r^{2i}, \tag{1}$$

where $\phi(r)$ is the phase at radius r, λ is the design wavelength, and the maximum value of I used in the optical design software usually do not exceed 5 for easy fabrication using the facility available at present.

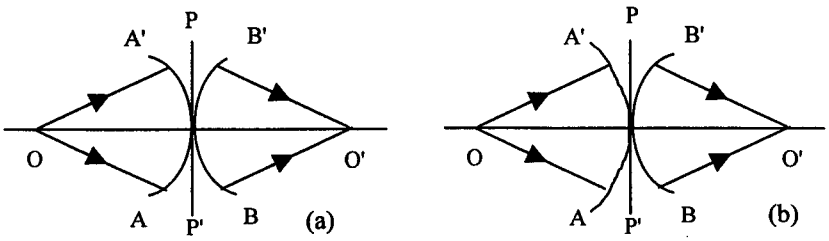


Fig. 1 Two cases for an axial diffractive lens PP' to transform (a) an incident spherical wave front AA' into another spherical wave front BB' and (b) an incident nonspherical wave front AA into a spherical wave front BB'.

The phase function can be obtained using the optical software package and this means that the coefficient a_i (I=1, 2, ···, 5) as shown in Eq. (1) can be determined, corresponding to the outer-zone radius r_m (for a symmetrical element) of the m^{th} full period zone of the diffractive lens. However, the phase function $\phi(r)$ should also be restricted by the following relation

$$\phi(r_m) = -Sm2\pi \tag{2}$$

where

$$S = \begin{cases} +1 & if \quad a_1 < 0 \quad or \quad f > 0 \\ -1 & if \quad a_1 > 0 \quad or \quad f < 0 \end{cases}. \tag{3}$$

Finding the proper r_m values using Eqs. (1) and (2) involve the solution of a 10^{th}-degree nonlinear equation and in spite of the complexities, we can determine the zone radii of the element.

The approximate surface-relief profile *t(r)* corresponding to the desired zone radius *r* can be determined from the following relationship

$$t(r) = \{[\phi(r)]_{2\pi}\} \frac{\lambda}{2\pi(n-1)} \tag{4}$$

where $\{[\phi(r)]_{2\pi}\}$ is the phase function $\phi(r)$ modulo 2π and n is the refractive index of the optical material used for the diffractive element. Here the approximate method neglects the thickness of the diffractive element and considers that the element is infinitesimally thin.

For most of the practical aberration-free imaging elements, the above approximate surface-relief profile is far from satisfying the requirements. To design a practical diffractive element, usually, we need to derive the surface-relief according to aberration-free imaging requirements. The details can be found from the references[14][16][20].

For an extra-axial imaging element design, it is almost the same as the design for an axial imaging element. However, the zone boundaries are usually not symmetrical, but are consisted of a series of curves. Since an off-axial imaging element has a large view of field, large off-axial aberrations such as coma, astigmatism, distortion and curvature field will be dominant. In order to design and make the element aberration-free, the zone boundaries have to be a series of non-regular curves according to the aberration-free condition. In this case, it is hard to fabricate the elements by any easily available facility because of the complex curves. However, we can design several symmetrical diffractive elements to replace a non-symmetrical element to simplify the fabrication. This practice is just like the use of several spherical surfaces to replace an aspherical surface in the traditional optical system design.

2.2 Exact Surface-Relief Profile of an Extra-Axial Stigmatism Imaging Element Designed from Geometrical Optics

2.2.1 Axial diffractive element design

Before we analyze the extra-axial diffractive element design using geometric optics, let us briefly review the design of an aberration-free imaging diffractive element for an axial object point.

Referring to reference[16], the design procedure starts from a commercial optical design software, firstly to determine the zone boundaries, and then to obtain the exact surface profile according to the aberration-free requirements.

In the design of a positive diffractive element, the sign convention of Kingslake[21] is generally adopted. As shown in Fig. 2, the convergence angle of the incident ray U=∠POA is positive. The axial object and image points, O and O′, are specified by l=AO and l'=AO′, respectively. T is the distance from the object point to the image point, the throw.

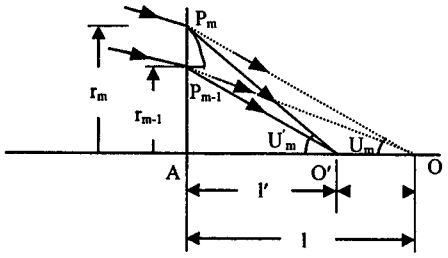


Fig. 2 Light rays traveling through mth zone of a diffractive element.

From Hazra's results[16], we know that the calculation of the zone radii r_m can be obtained from the following relationship,

$$r_m{}^2 = \left(\frac{1+ms\lambda/2T}{1+ms\lambda/T}\right)m^2\lambda^2 + \left(\frac{1+ms\lambda/2T}{(1+ms\lambda/T)^2}\right)2ms\lambda f\,, \tag{5}$$

where λ is wavelength; s is +1 if the focal length of the design element f is greater than 0 and s is −1 if f is less than 0.

Using geometrical optical analysis, the surface relief expressed in terms of the blazed thickness t_m can be described by the following relationship,

$$At^2{}_m + Bt_m + C = 0\,, \tag{6}$$

where

$$A = 1-n^2\,, \tag{7}$$

$$B = 2\left\{n\left(T+ms\lambda+\frac{r}{\sin U}\right)-\frac{1}{n}\left[(T+l)\sqrt{n^2-\sin^2 U}+r\sin U\right]\right\}, \tag{8}$$

$$C = 2T\left(l - ms\lambda - \frac{r}{\sin U}\right) - 2ms\lambda\frac{r}{\sin U} - m^2\lambda^2, \tag{9}$$

Eq. (6) shows that the zone profile of the diffractive element is parabolic. Given the external condition such as focal length, object and image distances, wavelength, Eq. (6) can be used to obtain the exact zone profile according to different zone radius r.

2.2.2 Extra-axial diffractive element design

1. Zone boundaries

For an extra-axial element design, it should be similar to the analysis of an axial element. We still have to determine the zone boundaries of the different zones and then to obtain the exact zone profile.

Referring to [10], if a diffractive element as shown in Fig. 3 is to converts a plane wave propagating with an oblique angle θ into a spherical wave converging with an optical axis vertical to the substrate. The phase-shift function of the lens is derived as

$$\Phi_T(x,y) = \frac{2\pi}{\lambda}\left[(x^2 + y^2 + f^2)^{1/2} + ny\sin\theta - f\right] - 2m\pi \tag{10}$$

where (x, y) is the coordinator of the imaging element, f is the focal length of the diffractive element, λ is the wavelength of the imaging wave in free space, n is the refractive index of the substrate, and m is the zone number, which is an integer that satisfies $0 \le \Phi_T \le 2\pi$.

If the lens can image a plane wave to a point, the phase-shift for every point on the wavefront of a plane wave should be zero, i.e. Φ_T=0. From this requirement, we can see that the imaging element pattern must have the following relationship,

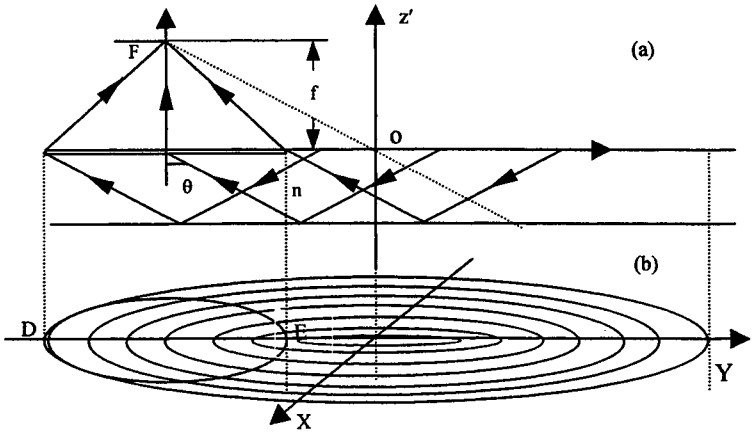


Fig. 3 (a) Design of a diffractive imaging element; (b) Bird view of the diffractive element, whose center is located at origin O. Only part of the whole lens is used for imaging.

$$\frac{x^2}{a^2} + \frac{(y - y_{cm})^2}{b^2} = 1 \tag{11}$$

with

$$a^2 = \frac{m^2\lambda^2 + 2fm\lambda + n^2f^2\sin^2\theta}{1 - n^2\sin^2\theta}, \tag{12}$$

$$b^2 = \frac{m^2\lambda^2 + 2fm\lambda + n^2f^2\sin^2\theta}{(1 - n^2\sin^2\theta)^2}, \tag{13}$$

and

$$y_{cm} = \frac{-n(m\lambda + f)\sin\theta}{1 - n^2 \sin^2\theta}. \tag{14}$$

From the above relationship, we find that the diffractive element possesses the pattern of an ellipse with its major axis lying along the Y-axis and its center position located at (0, y_{cm}). The ratio of the major axis to the minor axis of each ellipse is given by

$$\frac{b}{a} = \frac{1}{1 - n^2 \sin^2\theta}. \tag{15}$$

The diffractive element pattern is only a part of the ellipse in the negative direction of Y-axis as given by Eq. (11). In this case, the relation that shows the lens pattern can be written as

$$y(x,m) = y_{cm} + \frac{\left[-x^2\left(1 - n^2\sin^2\theta\right) + m^2\lambda^2 + 2fm\lambda + n^2 f^2 \sin^2\theta\right]^{\frac{1}{2}}}{1 - n^2\sin^2\theta}. \tag{16}$$

From the above relationship we can get the element period in the Y-axis as

$$\Lambda_m = \frac{\partial y}{\partial m}(x=0) = \frac{\lambda}{1 - n^2\sin^2\theta} \times \left[-n\sin\theta + \frac{m\lambda + f}{\left(m^2\lambda + 2m\lambda f + f^2 n^2 \sin^2\theta\right)^{1/2}} \right] \tag{17}$$

When m=0, the center period is given by

$$\Lambda_c = \lambda / (n\sin\theta). \tag{18}$$

It is noted that the diffractive element center is located at O as shown in Fig. 3(b), while the real diffractive element is only part of the whole element with its effective area restricted within the region from point D to point E.

2. Surface relief profile

With the zone boundaries determined using the above analysis, the next step is to determine the real surface relief profile using geometrical optics. In the case of a planar imaging visor using diffractive elements, the base plate of the element is generally very thick and the surface profile of the diffracitve element is thus much thinner than the base plate.

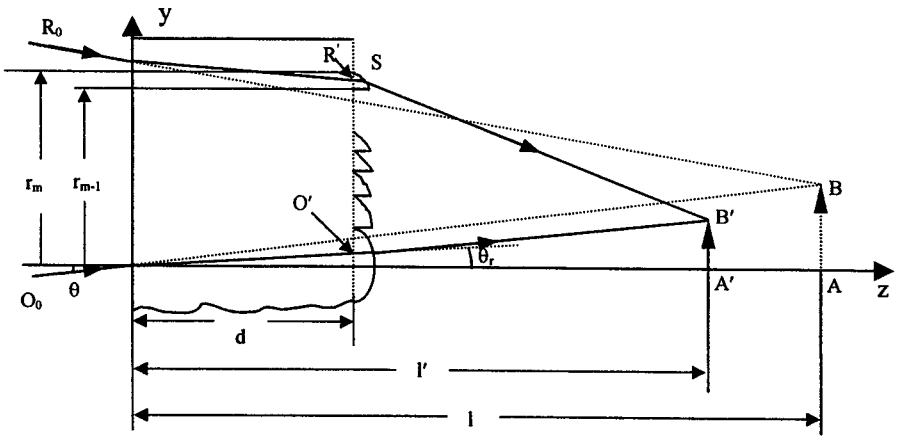


Fig. 4 An off-axial object B forms an aberration-free image at B′after imaging through a diffractive element. Two rays, one from the origin of the coordinate system and the other from the mth zone of the diffractive element, are illustrated.

Fig. 4 shows the simplified two-dimension figure of the element imaging through the mth zone of the element. Assume that the object is AB and its image is formed as A′B′. For the off axial object point B, if its aberration-free image B is formed by

the thick diffractive element, the optical path through the m[th] zone of the element and the optical path through the origin O will have the following relationship,

$$[BRR'SB'] = [BOO'EB'] + m\lambda . \tag{19}$$

Usually, when we start to design a diffractive element, we should know some external parameters such as object and image distance l, l', objective field of view θ, and diffractive element material refractive index n. From the previous zone boundaries analysis, we have determined the zone radius along y axis. It is easy to get the direction cosine (L, M, N) of the incident ray R_0R through the m^{th} zone,

$$L = -\frac{x}{\Delta},\ M = \frac{-(y + l\tan\theta)}{\Delta},\ N = \left[1 - \left(L^2 + M^2\right)^{1/2}\right], \tag{20}$$

$$\text{where } \Delta = \frac{l}{|l|}\left[x^2 + (y + l\tan\theta)^2 + l^2\right]^{1/2}. \tag{21}$$

For the incident ray OB, its refracted ray OO′ makes an angle θ_r with the z axis. So, we have, by the refractive law,

$$\theta_r = \arcsin\left(\frac{\sin\theta}{n}\right) \tag{22}$$

$$\text{and } OO' = \frac{d}{\cos\theta_r}. \tag{23}$$

The refractive ray RR_0 through the m^{th} zone can be expressed by the vectorial refractive law as

$$\mathbf{U_i} \times \mathbf{N} = \mathbf{U_r} \times \mathbf{N}, \tag{24}$$

where $\mathbf{U_i}$ and $\mathbf{U_r}$ are unit vectors along the incident ray R_0R and the refractive ray RR′, respectively, while $\mathbf{N}$ defines the normal at point R and the multiplication symbol indicates the cross vectorial product.

From Eq. (24), we have

$$L' = \frac{L}{n},\ M' = \frac{M}{n},\ N' = +\left[1 - \left(L'^2 + M'^2\right)\right]^{1/2}. \tag{25}$$

where L', M' and N' are the direction cosine of the refractive ray of R_0R.

The optical path $[BOO'EB']$ is given by

$$[BOO'EB'] = \frac{-l}{\cos\theta} + \frac{nd}{\cos\theta_r} + \frac{\lambda n}{n-1} + \frac{l'}{|l'|}\left[(-l'\tan\theta + d\tan\theta_r)^2 + (l' - d)^2\right]^{1/2}. \tag{26}$$

The optical path $[BRR'SB']$ can be written as

$$[BRR'SB'] = [BR] + [RR'] + [R'S] + [SB']. \tag{27}$$

$$\text{where } [BR] = \frac{-l}{|l|}\left[x^2 + (y + l\tan\theta)^2 + l^2\right]^{1/2}, \tag{28}$$

$$[RR'] = n\left[(x + L'd)^2 + (y + M'd)^2 + (N'd)^2\right]^{1/2}, \tag{29}$$

$$[R'S] = nt_m, \tag{30}$$

$$[SB'] = \left\{[x + L'(d + t_m)]^2 + [y + l'\tan\theta + M'(d + t_m)]^2 + [l' - N'(d + t_m)]^2\right\}^{1/2}, \tag{31}$$

Combing Eqs. (27), (26) and (19), we have the following relationship,

$$A_0 t^2{}_m + B_0 t_m + C_0 = 0, \tag{32}$$

$$\text{where } A_0 = 1 - n^2, \tag{33}$$

$$B_0 = 2(xL' + yM' - l'N' + D_0 n + d - l' \tan\theta M'), \tag{34}$$

$$C_0 = x^2 + y^2 + d^2 + l^2 + 2(xL'd - yl' \tan\theta + yM'd - l'M'd \tan\theta - l'N'd) + l'^2 \tan^2\theta - D^2{}_0, \tag{35}$$

$$D_0 = \frac{-l}{\cos\theta} + \frac{nd}{\cos\theta_r} + \frac{n\lambda}{n-1} + \frac{l'}{|l'|}[(-l' \tan\theta + d \tan\theta_r)^2 + (l' - d)^2]^{1/2} + n[(x + L'd)^2 + (y + M'd)^2 + (N'd)^2]^{\frac{1}{2}} + E_0, \tag{36}$$

$$E_0 = \frac{-l}{|l|}\left[x^2 + (y + l \tan\theta)^2 + l^2\right]^{1/2}. \tag{37}$$

From Eq.(32), we know that the profile of an off-axial imaging diffractive element is still a parabolic. The profile of the element is given by the thickness $t_m N'$ at point $S(x+t_m L', y+t_m M', t_m N')$. It should be noted that the refracted ray path RS inside the surface relief profile is specified in terms of the direction cosines (L', M', N') and the coordinates of point $R(x, y, 0)$ in the x-y plane. The parameters, t_m, L', M', and N' have unique values corresponding to a given point R at the x-y plane. A three-dimensional graph of the blaze profile of the diffractive element can be obtained from a plot of $t_m N'$ versus $x+t_m L'$ and $y+t_m M'$.

However, it is obvious that the blaze profiles of the diffractive element are symmetric with respect to the y-z plane, since both B and B′ are assumed to lie in this plane. Thus, a two-dimemsional section of the blaze profile in the y-z plane is adequate to describe the main characteristics of the diffractive element surface profile.

3. SHORT FOCAL LENGTH TESTING OF A DIFFRACTIVE ELEMENT AND ITS IMAGING EVALUATION

Although the testing and evaluation of optical lenses are very common in optical workshop testing, there is little research in the testing and evaluation of microlenses. This may be attributed to the somewhat recent emergence of microlenses, and also the specific experimental configuration that may, by size considerations, differ in multiple ways from traditional optical lenses testing. In the following, we will present an experimental configuration to test and evaluate these microlenses.

The difference between a common optical lens and a microlens is the size and the focal length, some of them even are too small to be seen by naked eyes. For testing such elements, the key problem is to enlarge the focal length. Fig. 5 shows a configuration in which a microscope objective is used to enlarge the focal length of the test element.

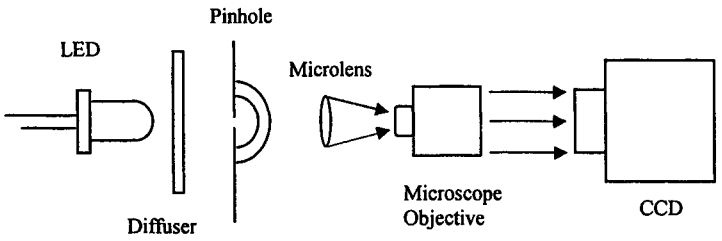


Fig. 5 Experimental configuration for testing microlens.

The schematic of the experimental setup, shown in Fig. 5, illustrates a method to test and evaluate a microlens. Narrow-band light from a high-luminance LED passes through a diffuser and condensing lens, and is spatially limited and filtered with a microscope objective and a pinhole aperture. As the microscope objective is used, the image of the point object on the pinhole formed by the microlens can also be used as the object of the microscope objective. By either having the microlens inserted in the system or not to form a sharp image on the CCD camera, two CCD camera distances can be obtained and from these image distances, the microlens focal length can be calculated. After analyzing the point-spread function, i.e., the image of the pinhole, we can obtain the MTF of the test element and hence complete the testing and evaluation of the microlens.

4. CONCLUSIONS

We have presented a geometrical method to determine the exact surface profile of an off-axial diffractive element. Our approach can be classified into two steps: firstly to find the zone boundaries according to the design requirement and then to obtain the exact surface profile for making the element aberration-free. We also presented an experimental schematic to test and evaluate micorlenses.

5. REFERENCES

1. Dale A. Buralli and G. Michael Morris, "Design of diffractive singlets for monochromatic imaging," Appl. Opt. 30(16), 2151-2158(1991).
2. Dale A. Buralli and G. Michael Morris, "Design of a wide field diffractive landscape lens," Appl. Opt. 28(18), 3950-3959(1989).
3. Donald W. Sweeney and Gary E. Sommargren, "Harmonic diffractive lenses," Appl. Opt. 34(14), 2469-2475(1995).
4. M. D. Missig, G. M. Morris, "Diffractive optics applied to eyepiece design," Appl. Opt. 34(14), 2452-2461(1995).
5. G. J. Swanson and W. B. Veldkamp, "Diffractive optical elements for use in infrared systems," Opt. Eng. 28(6), 605-608(1989).
6. D. A. Buralli and G. Michael Morris, "Design of two- and three-element diffractive Keplerian telescopes," Appl. Opt. 31(1), 38-43(1992).
7. Gary J. Swanson and W. B. Veldkamp, "Binary lenses for use at 10.6 micrometer," Opt. Eng. 24(5), 793-799(1985).
8. A. P. Wood, " Design of infrared hybrid refractive-diffractive lenses," Appl. Opt. 31(13), 2254-2258(1992).
9. Seok Ho Song, Suntak Park, El-Hang Lee, Pill Soo Kim, and Cha Hwan Oh, "Planar optical implementation of multichannel fractional Fourier transforms," Opt. Comm. 137(1997), 219-222.
10. Teruhiro Shiono and Hisahito Ogawa, " Planar-optic-disk pickup with diffractive micro-optics," Appl. Opt. 33(31), 7350-7355(1994).
11. Jurgen Jahns and Susan J. Walker, "Imaging with planar optical systems," Opt. Comm. 76(5,6), 313-317(1990).
12. Friesem, S. Reinborn, Y. Amitai, and R. Shechter, "Planar optical configurations for imaging and data processing," SPIE 2404, 201-210(1995).
13. Y. Han, L. N. Hazra and C. Delisle, "Exact surface-relief profile of a kinoform lens from its phase function," J. Opt. Soc. Am. A. 12(3), 524-529(1995).
14. L. N. Hazra, Y. Han and C. Delisle, "Plane kinoform lenses for axial stigmatism in finite conjugate imaging," Opt. Comm. 91(1,2), 1-4(1992).
15. L. N. Hazra, Y. Han and C. Delisle, "Curved kinoform lenses for stigmatic imaging of an axial object at infinity," Opt. Comm. 90(4,5,6), 201-206(1992).
16. L. N. Hazra, Y. Han and C. Delisle, "Design of plane kinoform lenses for stigmatic imaging between two prespecified points on the axis," Opt. Comm. 94(4), 203-209(1992).
17. L. N. Hazra, Y. Han and C. Delisle, "Imaging by zone plates: axial stigmatism at a particular order," J. Opt. Soc. Am. A. 11(10), 2750-2754(1994).
18. D. L. Dickensheets, "Imaging performance of off-axis planar diffractive lenses," J. Opt. Soc. Am. A Vol. 13(9), 1849-1858(1996).
19. O. Filali Meknassi, C. A. Delisle, and L. N. Hazra, "Extra-axial stigmatic imagery with a thick planar inoform," Appl. Opt. 36(26), 6577-6582(1997).
20. L. N. Hazra, Y. Han and C. Delisle, "Plane kinoform lenses for axial stigmatism in finite conjugate imaging," Opt. Comm. 91(1,2), 1-4(1992).
21. Rudolf Kingslake, *Optical System Design,* Chapter 3, pp.26-39, Academic Press, Inc. (1983).

*Correspondence:
Yun Zhisheng: Tel: (65)7905985 ext 22, Fax: (65)7933318, e-mail: ezsyun@ntu.edu.sg

Highly efficient diffraction gratings and resonant plasma layers at IR-wavelengths

J. Stiens[a*], G. Nemova[b], P. Muys[c], R. Vounckx[a]

[a]Vrije Universiteit Brussel, Lab for Micro and Optoelectronics, Electronics Department, Pleinlaan 2, B-1050 Brussels, Belgium.

[b]Institute of Radio Engineering and Electronics of RAS, Vvedensky Square 1, R-141120 Fryazino (Moscow Region), Russia

[c]Laser Power Europe, Meersstraat 138E, 9000 Gent, Belgium

ABSTRACT

In this work we investigate the interaction between rectangular-grooved transmission gratings and a resonant plasma layer at infrared (IR) wavelengths. The gratings are designed to convert quasi-vertically incident light into quasi-horizontally propagating light in a high refractive index GaAs substrate with efficiencies of almost 90% in the first diffraction order. This can only be achieved by etching highly asymmetric "λ/4" or "λ" gratings in a high refractive index material (e.g. Germanium) evaporated on the GaAs substrate. The resonant plasma layer (RPL) consists of a very thin (10-100nm) highly doped (5-$8x10^{18}cm^{-3}$) n-GaAs layer whose plasma frequency is almost equal to the frequency of the incident light. Under these conditions Drude's formula shows that the refractive index almost vanishes due to plasma oscillations. The interaction between the diffraction modes and the RPL are investigated and optimised on the base of a rigorous coupled-wave analysis. This analysis reveals to which extent the RPL can influence the distribution of light between the zero and first diffraction orders. In the optimum position of the RPL, the first order diffraction efficiency changes from 90% downto 25% for a RPL thickness change of 40nm.

Keywords: diffraction grating, free electron, plasma resonance, GaAs, optical incoupling mechanism

1.INTRODUCTION

A variety of opto-electronic devices and systems can be made more efficient by causing light to propagate horizontally inside a substrate containing light-sensitive layers. In many applications, incident light is available from a source normal to the substrate. Light-coupling mechanisms, therefore, are required substantially to change a propagation direction of the incident light as it enters the substrate. In general, coupling efficiency of the light-coupling mechanism decreases as a change in the propagation direction increases. In addition, the use of substrates with high refractive indices also lowers coupling efficiency. Many devices, such as photodetectors [1,2] and other sensors [3], would benefit from an efficient conversion of normally incident light into horizontally propagating light inside the substrate. In photodetectors, for example, an interaction length of normally incident photons within an absorbing or sensing layer of the substrate can be drastically increased when the photons propagate obliquely inside the substrate. The resulting increase in an absorption coefficient allows the absorbing layer to be made thinner. Miniaturization of the detector area[4], coupled with thinner absorbing layers, thus allows for the fabrication of high speed photodetectors.

Some types of photodetectors, such as infrared photodetectors, rely on intersubband transitions in quantum wells[5]. The intersubband transitions, however, can only be induced by a photon electric field directed along a growth direction of the substrate. Since the growth direction is vertical, normally incident light (having an electric field perpendicular to its direction of propagation) produces no interaction between the photon electric field and the quantum wells. The normally incident light, therefore, must be converted into horizontally propagating light in the substrate to induce the intersubband transitions.

- Correspondence: email: jstiens@vub.ac.be; http://etro.vub.ac.be/lami/jstiens/welcome.html
- phone: (+32)2.629.23.97 ; fax: (+32)2.629.28.83

Part of the SPIE Conference on Advanced Photonic Sensors and Applications
Singapore • November–December 1999
SPIE Vol. 3897 • 0277-786X/99/$10.00

Modulators that exploit the characteristics of bulk polariton effects are another class of devices that would benefit from the horizontal propagation of light inside the substrate. A modulation depth of the modulators increases as the angle of propagation in the substrate approaches horizontal. A number of different light-coupling techniques have been proposed. One technique attempts to couple as much light as possible into the substrate by positioning the substrate under a Brewster angle. While transverse magnetic (TM) polarized light enters the substrate with an efficiency of almost 100%, refraction of the light towards the normal reduces the efficiency of this technique, making it undesirable for detection and sensing applications. Another technique involves polishing the substrate to an angle of less than 45° to convert incident light to an almost horizontally propagating beam. The incident light may thus penetrate the substrate at the proper angle. An efficiency of the polishing technique, however, is low. While the efficiency may be increased by the use of anti-reflective coatings, the polishing technique is difficult to integrate into large arrays. Diffraction gratings have also been proposed. Diffraction gratings operating in transmission mode are used to diffract light into many orders. While diffraction gratings exhibit a high total efficiency, diffraction efficiencies related to horizontal propagation are low. Single order diffraction gratings, produced from the same material as the substrate and operating in transmission mode, have also been proposed. The gratings, however, are inherently limited in efficiency due to a mismatch between grating modes and substrate modes. The efficiency of the gratings may be increased by covering the gratings with a layer of reflective metal. The reflective layer, however, restrict the grating, allowing it to operate only in reflection mode.

A particular application of highly efficient diffraction gratings consists in the excitation of bulk plasmons in IMOS modulators[6]. It is well known that free carriers can have a great impact on the refractive index of a doped semiconductor at mid and far infrared wavelengths. One can calculate by means of Drude's formula [6] that the refractive index of this layer almost vanishes when the frequency of the incident light matches the plasma frequency of the highly doped GaAs material due to plasma oscillations. This frequency matching can be achieved for the spectrum of a CO_2 laser (9-11μm). The effective reflection on this layer of an incident electromagnetic wave is a sensitive function of wavelength, polarization, angle of incidence, thickness and doping concentration [7]. By changing the thickness of the *resonant plasma layer*, large changes in reflectivity are realized if the proper combination of wavelength, polarization and angle of incidence are picked. The strongest reflectivity changes are observed when p-polarized light is near grazing incidence. The depletion of the electron concentration in this very thin layer (typical tens on nm) can be achieved by applying an electric field resulting in a tremendous refractive index changes. A p-I-n diode structure was successfully applied before to modulate the depletion width of a RPL

It is the purpose of this paper to combine these two interesting physical objects: diffraction gratings and RPL and to investigate the process of the light diffraction by dielectric grating when interacting with the RPL

2. RECTANGULAR-GROOVED DIELECTRIC GRATING ON A GaAs-SUBSTRATE

In the first step we investigate and optimize the grating conditions for which quasi-vertically incident light can be diffracted to almost horizontally propagating light inside the substrate. The diffraction grating when the RPL is absent, is depicted in figure 1. Light is incident from the cover region (usually air) under an angle θ_c. The rectangular grooved grating consists of two materials: a high refractive index material with refractive index $n_{g,H}$ and a low refractive index material with refractive index $n_{g,L}$. The substrate has a refractive index n_s. The grating period Λ is set by the condition that the first diffraction order in transmission T[+1] or T[-1] is diffracted under a very large angle, typically more than 70°. The grating materials, height H and the duty cycle δ, defined as the ratio of the width of high index material and the period, have to be optimized for maximum diffraction in the T[+1] and/or T[-1] order. For the investigation of this structure the rigorous coupled-wave analysis (RCWA) was chosen as one from the most well known methods [8,9] which does not suffer from the various approximations, which are inherent to a number of other methods, e.g. the coupled-wave approach or the modal approach [10,11].

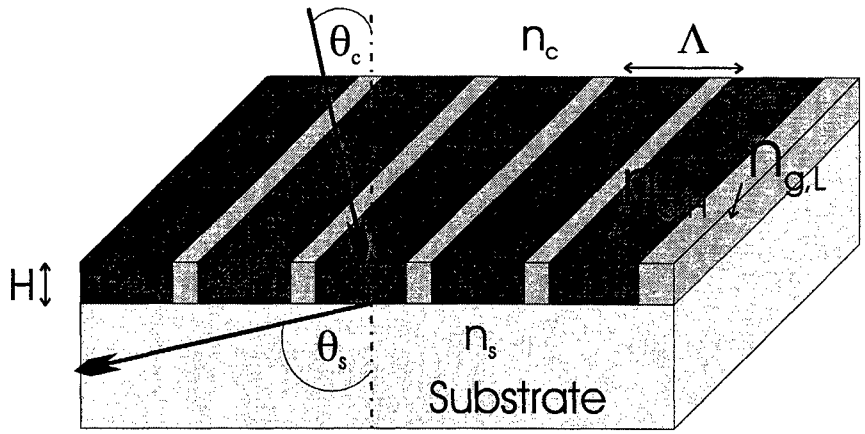


Fig.1 Schematic view on the rectangular grooved diffraction grating. Symbols are indicated in the text.

Following the RCWA we obtain the period of the grating. According to RCWA the total field in the regions 1, 2 and 3 may be written as a sum of plane waves (i-modes), each of which individually satisfies Maxwell's equation [3]. The normalized total fields in cover region 1 and substrate region 3 are respectively:

$$E_1 = \exp(-j\ \vec{k}_1\vec{r}) + \sum_{i=-\infty}^{\infty} R_i \exp(-j\vec{k}_1^{\,i}\ \vec{r}), \tag{1}$$

$$E_3 = \sum_{i=-\infty}^{+\infty} T_i \exp[-j\vec{k}_3^{\,i}(\vec{r} - h\vec{r})], \tag{2}$$

where $j=\sqrt{-1}$ and $\vec{r} = x\vec{x} + y\vec{y}$, $\vec{x}$ and $\vec{y}$ are unit vectors of x-axis and y-axis, respectively. R_i and T_i are the amplitudes of the reflected and transmitted modes. $\vec{k}_1^{\,i}$ is the wave vector of the i-th reflected wave, $\vec{k}_3^{\,i}$ is the wave vector of the i-th transmitted wave. Each i-th wave in region 1 and 3 must be phase matched to the i-th space harmonic field inside the grating. Thus the x components of the wave vectors of the i-th wave (region 1 and 3) and the x component of the wave vector of the i-th space harmonic field (region 2) must be the same. That is $\vec{k}_1^{\,i}\vec{x} = \vec{\sigma}^{\,i}\vec{x} = \vec{k}_3^{\,i}\vec{x}$, where $\vec{\sigma}^{\,i}$ is wave vector of i-th mode in the grating region. Hence we have the "grating equation" for the backward-diffracted waves:

$$n_c Sin\theta_{R[i]} = n_c Sin\theta_c - i\lambda/\Lambda \tag{3}$$

and for the forward-diffraction waves:

$$n_s Sin\theta_{T[i]} = n_c Sin\theta_c - i\lambda/\Lambda \tag{4}$$

The permittivity of the grating region is described by the function

$$\varepsilon_g(x) = \begin{cases} \varepsilon_{g,H}; x \in (0, \delta\Lambda) & \text{(5a)} \\ \varepsilon_{g,L}; x \in (\delta\Lambda, \Lambda) & \text{(5b)} \end{cases}$$

Expanding $\varepsilon_g(x)$ into a Fourier gives

$$\varepsilon_g(x) = \varepsilon_g^{(0)} + \sum_{m=-\infty, m\neq 0}^{+\infty} C_m e^{jmKx} \quad \text{, where } K = \left|\vec{K}\right| = 2\pi/\Lambda \text{'} \tag{6}$$

$\vec{K}$ is the wave vector of the grating

$$1, x \in (0, \delta\Lambda) \tag{7a}$$

$$C_m = \frac{1}{\Delta}\int_0^{\Delta} f(x)e^{-jmKx}dx;\ f(x) = \{$$

$$0;\ x \in (\delta\Lambda, \Lambda) \tag{7b}$$

Grating equation for the grating region reduces to

$$n_C Sin\theta_i^{'} = \sqrt{\varepsilon_g^{(0)}}\, Sin\Theta - i\frac{\lambda}{\Delta} \tag{8}$$

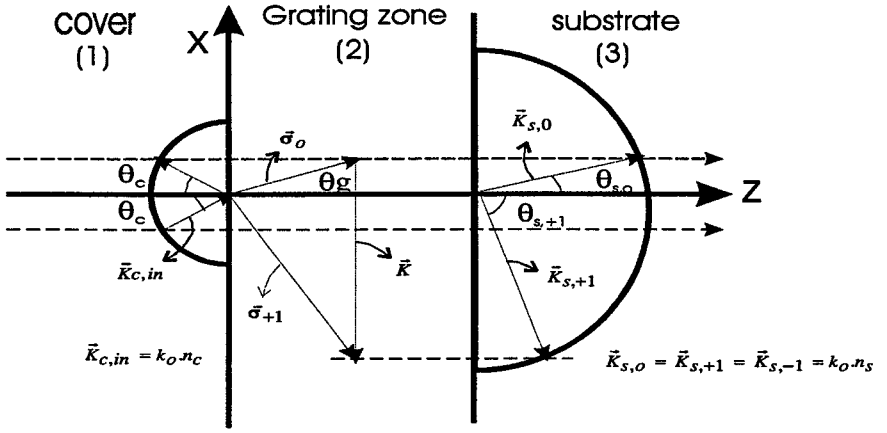


Fig.2 Schematic illustration of the phase matching conditions

In the region 1 and 3 $\left|\vec{k}_1^{\,i}\right| = \left|\vec{k}_1\right|$ and $\left|\vec{k}_3^{\,i}\right| = \left|\vec{k}_3\right| = k_0 n_s$, where n_s is the refraction index of substrate. Two semicircles demonstrate these conditions. Knowing the total amplitude and the x components of the diffracted wave vector from phase-matching requirements, the z components are then determined to be

$$\vec{k}_{1z}\vec{z} = \left[\left|\vec{k}_1\right|^2 - \left(\vec{k}_1^{\,i}\vec{x}\right)^2\right]^{1/2} = \left[k_1^2 - (\sigma_0 Sin\Theta - jK)^2\right]^{1/2} \tag{9}$$

and

$$\vec{k}_3^{\,i}\vec{z} = \left[\left|\vec{k}_3\right|^2 - (\vec{k}_3^{\,i}\vec{x})^2\right]^{1/2} = \left[k_3^2 - (\sigma_0 Sin\Theta - jK)^2\right]^{1/2} \tag{10}$$

where $k_1 = \left|\vec{k}_1\right|; k_3 = \left|\vec{k}_3\right|$ and $\sigma_0 = \left|\vec{\sigma}_0\right|$; $\vec{z}$ is unit vector of z axis.

Quantities (9) and (10) are either positive real (propagating wave) or negative imaginary (evanescent wave) in region 3 or negative real (propagating wave) or positive imaginary (evanescent wave) in region 1. Whether the i-th field in region 1 or 3 is propagating or cut off may be visualized as shown on Fig.2. The horizontal dashed lines in the figure demonstrate the phase matching conditions. In order to convert almost vertically incident light into quasi-horizontally propagating the period is set as follows:

$$\Lambda = \lambda\left[n_C Sin\theta_C - n_S Sin\Theta_{T[+1]}\right]^{-1} \tag{11}$$

In the next part we investigate the influence of the duty cycle δ, grating height H and grating material on the diffraction order T[+1] and T[-1]. We assume that light is perpendicular incident. Under this condition the diffraction efficiency T[+1] is equal to T[-1] due to symmetry. It is easy to demonstrate that by slightly tilting the angle of incidence and slightly adapting the grating period, optical power from the T[-1] order can be completely transferred to the T[+1] order upto a completely vanishing T[-1] order. The simulations are performed for λ=10.6μm. The substrate is undoped GaAs with ε_s=10.9.

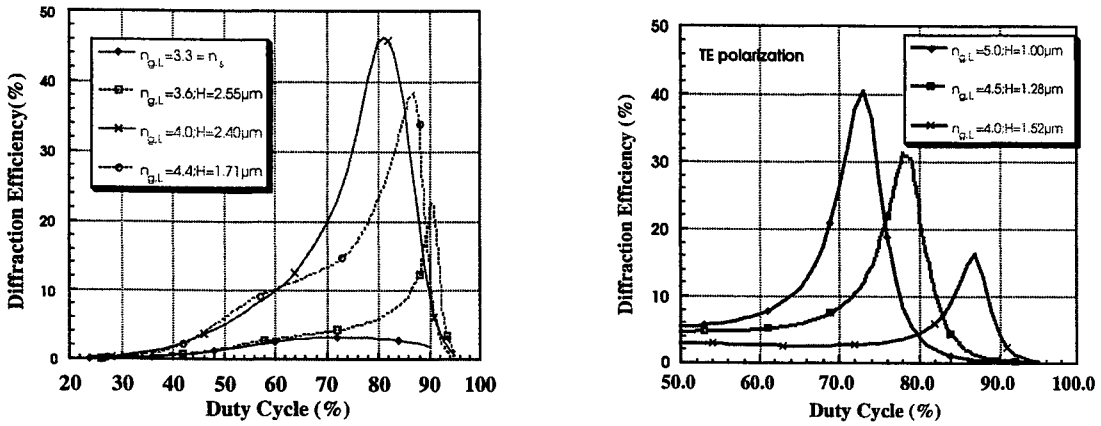


Fig.3 The diffraction efficiency T[+1] versus the duty cycle of the grating etched in different materials. The wavelength of the incident light is equal to 10.6 μm, the grating period is 3.22 μm and the grating height ≈ λ/4. In case (a) the polarization is equal to TM and in case (b) the polarization is equal to TE. The diffraction order T[+1] diffracts under an angle of 85° inside the substrate.

Turning now to fig.3, illustrated is an exemplary graph of a diffraction efficiency of the diffraction grating for (a) TM and (b) TE polarized light. The graph illustrates a relationship between the diffraction efficiency and an asymmetric duty cycle of the diffraction grating for various refractive indices. In the illustrated graph, the first refractive index $n_{g,H}$ is varied between 3.3, i.e. the refractive index of the substrate itself, and 5.0. Additionally, the second refractive index $n_{g,L}$ is assumed to be equal to 1 for simplicity in calculating the diffraction efficiency and the duty cycle of the first and second materials. As illustrated, optimal diffraction efficiencies may be obtained at highly asymmetric duty cycles (*e.g.*, for a first refractive index $n_{g,H}$ of 4.0, an optimal duty cycle is about 80% for TM polarized light). When the grating is directly etched into the GaAs substrate the diffraction efficiency T[+1] or T[-1] remains very small, less than 5%. This can be explained as follows. For small refractive index material, the average refractive index or dielectric permittivity ε_g° of the rectangular grooved grating region is too small to accommodate a high transfer of the incident light toward the required diffraction order. Only when the refractive index and the duty cycle are sufficiently large, the total structure can behave as a 3-layered waveguide such as in multilayer grating filters. This exceptional optical power transfer towards the T[+1] order is connected with the excitation of leaky guiding waves in the grating zone. Such a similar anomaly was discussed in detail by Hessel [12]. We observe that for TE polarized light the refractive index required for the maximum power transfer is larger. This is due to the fact that the waveguide index for TE modes is always larger than for TM waveguide modes. These two graphs illustrate that direct etching into a high refractive index substrate is useless when one aims at converting the vertically incident light towards horizontally propagating light. Optimal designs for TM and TE polarized light do not coincide. One dimensional diffraction gratings, therefore, should be designed for either TM or TE polarized light. The operation of the diffraction grating is reciprocal: in the given example, light is incident from air, but similar results are obtained when light is propagating quasi-horizontally inside

the substrate and coupled out almost vertically. This means that high refractive index gratings can also be used as efficient output couplers.

The height H of the grating is about "λ/4". The height of the grating is directly related to the wavelength. The results obtained here can be scaled without any problem to other wavelengths. The observed anomaly, however exists also for another grating height H as illustrated in figure 4a,b. In figure 4.a the grating height corresponds to an optimized λ/4 value, 4.b the grating height is essentially equal to the wavelength. In both figures the curves are calculated for different grating periods, corresponding to different angles of propagation inside the substrate. In fig. 4a we see that the larger the angle of propagation inside the substrate the smaller the total diffraction efficiencies in the T[+1] and the T[-1] order. Upto about 80° one can say that the total diffraction efficiency remains quasi-constant, but for larger angles it is clear that it is more difficult for the equivalent waveguide to leak to the substrate. When the grating height H is about four times larger, i.e. equal to λ, we see that the grating structure can supply three diffraction anomalies. These anomalies are again due to the excitation of leaky waveguide modes inside the grating region. From an experimental point of view we see that the "λ/4" gratings are more interesting than the "λ"gratings as less material is required. The maximum diffraction efficiencies of both types of gratings are very similar. The three anomalies for the "λ" gratings have a different behavior for different gratings periods or different angles of propagation inside the substrate. The diffraction efficiency of the first anomaly increases with larger angles of propagation. The second anomaly behaves like the anomaly of the "λ/4" gratings and the third anomaly is quite independent of the angle of propagation inside the substrate. With these graphs we demonstrate that high refractive index gratings can be used as efficient optical coupling mechanisms to provide light inside the substrate under quasi-horizontal propagation. In the remainder of this paper we will focus on a optimized "λ/4"Germanium grating ($\varepsilon_{g,H}$=16, H=2.43μm, δ=84%) for TM polarized light.

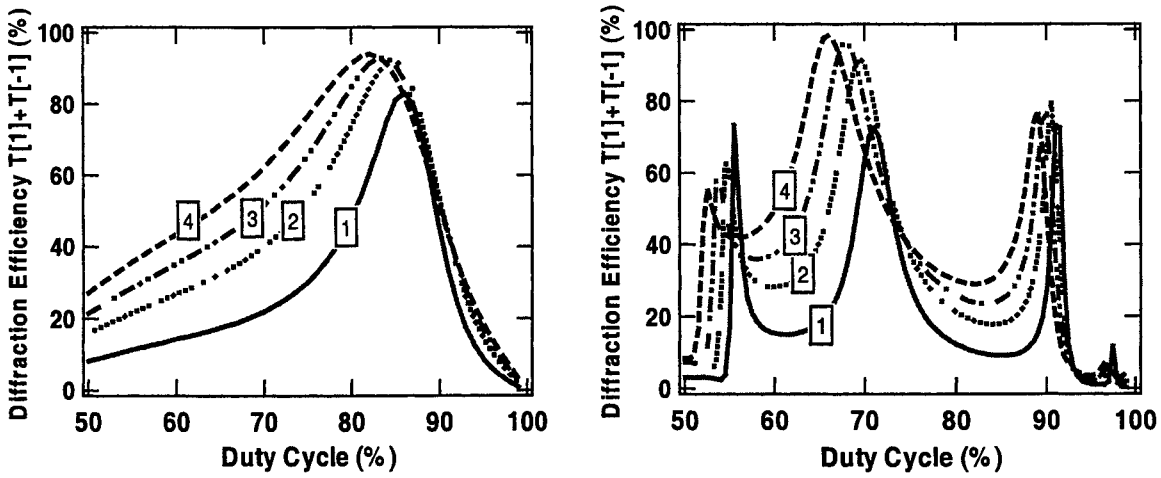


Fig.4 The diffraction efficiency T[+1] + T[-1] versus the duty cycle of the grating for different periods. The etch depth ≈ λ/4. The wavelength of the incident light is equal to 10.6 μm and the polarization is equal to TM. The etch depth ≈ λ/4 in case (a) and ≈ λ.in case (b). (1) $\theta_{T[\pm1]}$=85° (2) $\theta_{T[\pm1]}$=80°; (3) $\theta_{T[\pm1]}$=76.6°; (4) $\theta_{T[\pm1]}$=73.4°.

3. RESONANT PLASMA LAYER and Ge-GRATING

As an example of application we will study the influence of a resonant plasma layer (RPL) on the distribution of the optical power amongst the different diffraction orders. When calculating the refractive index of the highly doped n-GaAs layer, the RPL, it is important to take into account the full nonparabolic behaviour of the central conduction band as well as the contribution of carriers in the satellite valleys[13] at these high doping concentration. When the resonant plasma layer is integrated into a diode [6] it is possible to change the electronic thickness of the RPL by means of the depletion effect. By applying a reverse bias to a diode the electric field will remove the free carriers and we obtain insulating GaAs instead. For the considered high doping concentration the depletion width is almost 20nm. Since the LRP thickness is sufficiently easy changeable parameter it will be interesting to investigate the influence of the RPL thickness on the diffraction process.

3.1 A flat RPL inside or on top of the substrate

In a first step the RPL is on top of the substrate just beneath the diffraction grating. Light is again impinging perpendicularly on the grating zone. The period is again chosen such that only the following diffraction orders T[0], T[+/-1] exist inside the substrate with wave vectors $\vec{\sigma}_0$ and $\vec{\sigma}_{+,-1}$ respectively. The T[0] order propagates within the RPL and penetrates into the substrate. The T[+1] and T[-1] impinge on the RPL under an angle, which is much more than the angle of total reflection. This wave exponentially decays in the RPL but when the RPL is sufficiently thin the T[+1], T[-1] orders can tunnel into the substrate as a propagating wave.

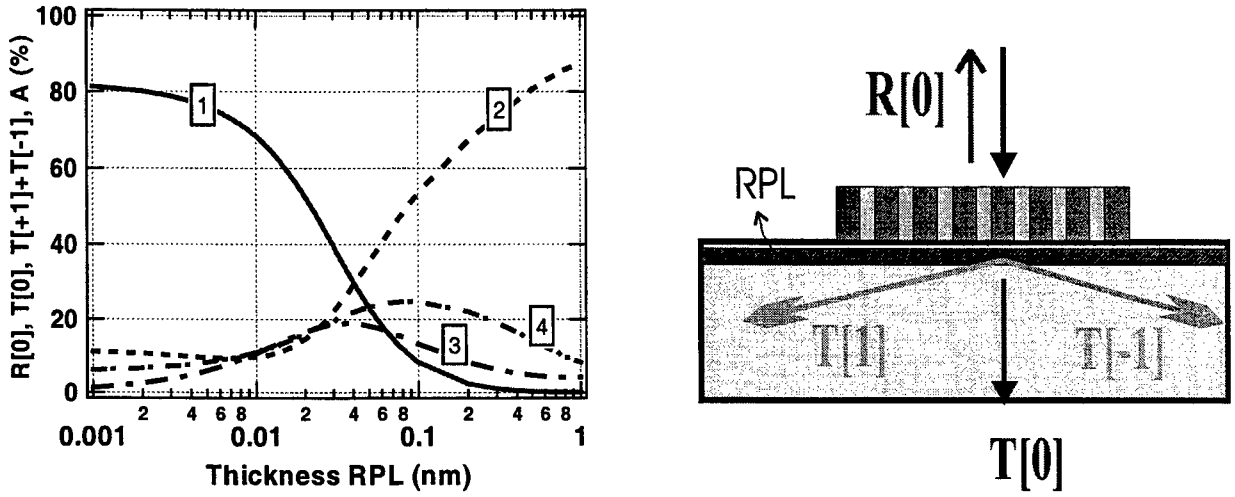


Fig.5 The diffraction efficiencies T[+1] + T[-1], T[0], R[0] and the absorption versus the thickness of the resonant layer (RL). The duty cycle is 85%, the etch depth $\approx \lambda/4$, the angle of diffraction $\theta_{T[\pm1]}=85°$. The wavelength of the incident light is equal to 10.6 μm and the polarization is TM. (1) T[+1]+T[−1]; (2) R[0]; (3) T[0]; (4) Absorption.

When the RPL is sufficiently thick (that means comparable with or more than the penetration depth) the T[+1], T[-1] orders do not penetrate into the substrate and all incident light remains only in the R[0] order. The dependence of the light intensities diffracted into the different orders is shown in Fig. 5. When the thickness of the RPL increases the T[-1] and T[+1] diffraction mode evanesces. This light is predominantly transferred into the R[0] order. The absorption inside the RPL first increases and then decreases as less light penetrates the RPL because it is reflected on it. Once the RPL is sufficiently thick the operation of the diffraction grating is completely neutralized by the RPL. We also investigated the possibility of putting the RPL deeper into the substrate such that a GaAs spacer layer is between the grating and the RPL. In this case the transmission orders propagate in the spacer layer between the grating and the RPL unaffected. As before the T[+1], T[-1] orders decay exponentially in the RPL due to total reflection. The T[+1] order can tunnel through the RPL only if its thickness is sufficiently small. The influence of the thickness of the spacer layer is minimal on the diffraction characteristics and hence can be neglected. In this case we have again reciprocal characteristics. Incoupling and outcoupling are again interchangeable. Light can be propagating inside the substrate and the vertical outcoupling will be determined by the layer thickness of the RPL.

3.2 The RPL inside the Germanium grating region

In this case the resonant plasma layer is a part of the grating itself. The resonant plasma region is positioned inside the teeth of the Germanium grating. We will first discuss the two extreme cases when the plasma layer is at the top or the bottom of the grating zone. We will start with the situation where the RPL is at the top of the grating as illustrated in figure 6a,b.

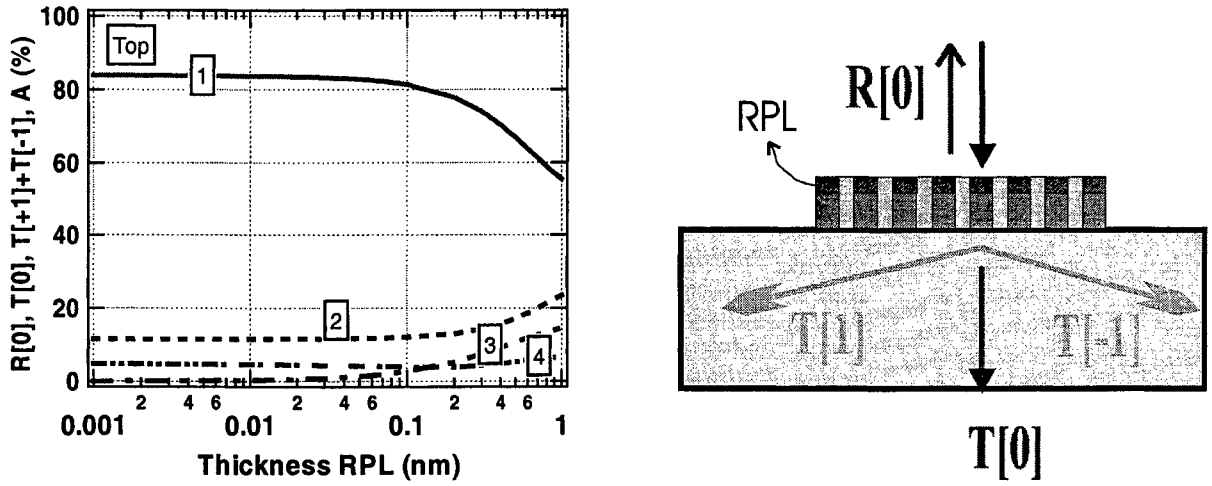


Fig.6 The diffraction efficiencies T[+1] + T[-1], T[0], R[0] and the absorption versus the thickness of the resonant layer (RPL) inside the grating region. In this case the RPL is positioned at the top of the grating teeth. The duty cycle is 85%, the etch depth ≈ λ/4, the angle of diffraction $\theta_{T[\pm1]}$=85°. The wavelength of the incident light is equal to 10.6 μm and the polarization is TM. (1) T[+1]+T[−1]; (2) R[0]; (3) T[0]; (4) Absorption.

Light is again vertically incident on the resonant plasma grating region and diffracts into two propagating orders: T[0] and T[+/-1]. All orders reach the Ge part of the grating and here they again diffract into two propagating waves. These diffraction orders propagate in the substrate. As we can see from fig. 6a when the RPL is on the top of Ge-grating the diffraction process, which takes place in the Ge-grating is only slightly disturbed. Only when the RPL is at least 100 nm thick and above the 2.4μm thick Ge-grating region the diffraction processes are disturbed. Hence in this configuration the interaction between the diffraction grating and the resonant plasma layer is very weak, such that it is not interesting for e.g. detection or sensor applications.

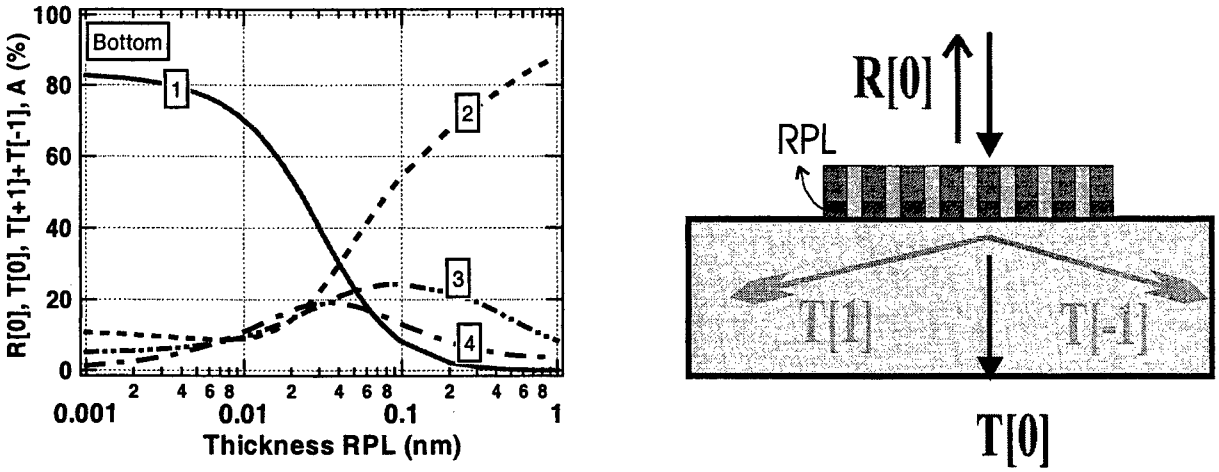


Fig.7 The diffraction efficiencies T[+1] + T[-1], T[0], R[0] and the absorption versus the thickness of the resonant layer (RPL) inside the grating region. In this case the RPL is positioned at the top of the grating teeth. The duty cycle is 85%, the etch depth ≈ λ/4, the angle of diffraction $\theta_{T[\pm1]}$=85°. The wavelength of the incident light is equal to 10.6 μm and the polarization is TM. (1) T[+1]+T[−1]; (2) R[0]; (3) T[0]; (4) Absorption.

In the second case the plasma layer is positioned at the bottom of the diffraction region as illustrated in fig.7a,b. When light is incident on the Germanium teeth, the zero order diffraction propagates through the germanium region and when it reaches the RPL it continues to diffract in the RPL into two propagating orders. The first order from the germanium region exponentially decays in the RPL due to phase matching conditions and can only tunnel through the substrate when the RPL is sufficiently thin. So in this case the RPL strongly affects the diffraction pattern. The diffraction processes in such structures conceptually take place inside the RPL.
The interaction between the diffraction orders and the resonant plasma can still be further optimized when the RPL is put inside the grating zone. We have simulated the case where a 40nm thin RPL is vertically moving inside a 2.35μm thick diffraction grating region. The results are shown in figure 8a,b. One observes that the optimum interaction between a 40nm thin RPL and the diffraction pattern takes place when this layer is inside the Ge region, 350 nm above the GaAs substrate. This maximum interaction is spread over a region of about 500 nm. In this case the first order diffraction can be modulated from 80% downto 24.5% when the RPL would be completely depleted.

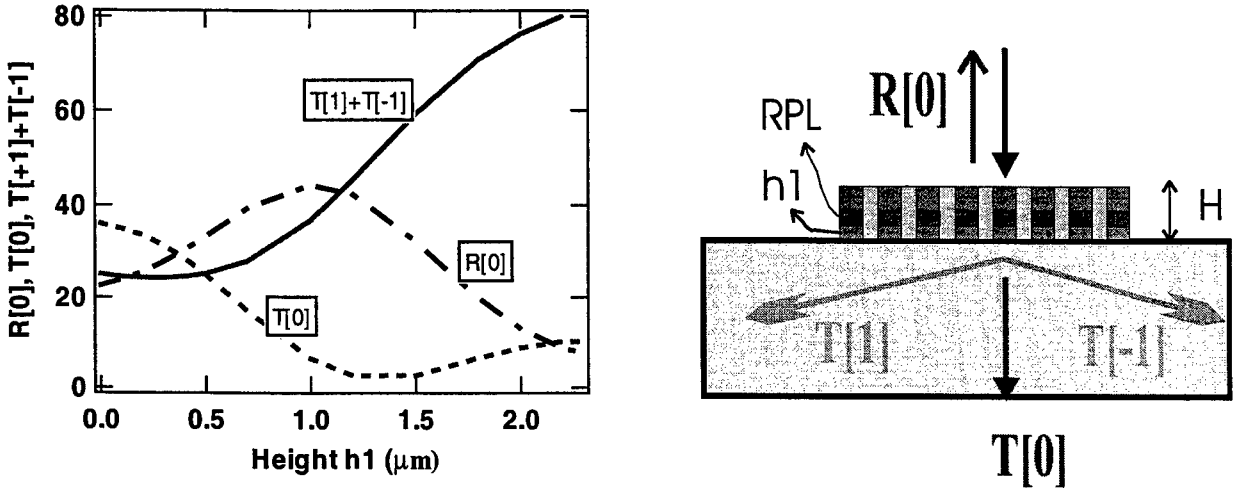


Fig.8 The diffraction efficiencies T[+1] + T[-1], T[0], R[0] versus the height h1 of the grating zone between the substrate and the RPL. The RPL is 40nm thick and the total height of the grating is 2.35 μm and the duty cycle is 85%. The angle of diffraction $\theta_{T[\pm1]}=85°$. The wavelength of the incident light is equal to 10.6 μm and the polarization is TM.

From the simulations it is also clear that the diffraction pattern difference between a spatially modulated RPL near the bottom of the grating zone and a fully flat RPL inside the substrate is quite small. This leads also to the conclusion that the modulation of the RPL from a flat layer towards a spatially modulated layer is not interesting for modulation applications.

4. CONCLUSIONS

The first goal of the investigations reported in this paper was the optimization of a diffraction grating operating in transmission for converting vertically incident light in quasi-horizontally propagating light inside a high refractive index substrate. We showed that it is necessary to cover the substrate with a very asymmetrically etched grating consisting of material with a refractive index, which is substantially larger than the index of the substrate. This high asymmetry and refractive index material are needed to virtually construct a waveguide where the diffraction grating acts as the leaky guiding film. Diffraction efficiencies of almost 90% were obtained for two different grating types. Two optimal values for the grating heights are "λ/4" and "λ" gratings. This also means that the results shown here can be easily rescaled for other wavelengths. In these two cases anomalies appear due to the resonance conditions of the leaky waveguide construction. Many (electro) optical devices such as detectors, sensors where horizontal propagation inside the substrate is vital, can benefit from these findings. A particular application of these highly efficient diffraction gratings is their interaction with a resonant plasma layer. We have shown that the position of the RPL determines to which extent the RPL affects this interaction. When the RPL is inside the subtrate or at the bottom of the grating region the original RPL-free diffraction pattern is seriously disturbed. When the RPL is positioned at the top of the grating structure, the operation of the diffraction grating is almost unaffected. A

modulator exploiting the excitation of bulk plasmon polaritions can be effectively modulated between 80% and 25% for a 40nm thickness variation.

ACKNOWLEDGMENTS

This work was funded by the VUB-GOA (Concerted Research Actions) and the IWT (Flemish Institute for the encouragement of Scientific-Technological Research in the Industry).

REFERENCES

1. S.D. Gunapala, J.K. Liu, J.S. Park, M. Sundaram, C.A. Shott, T. Hoelter, T.L. Lin, S.T. Massie, P.D. Maker, R.E. Muller, G. Sarusi, "9 μm Cutoff 256x256 GaAs- $Al_xGa_{1-x}As$ Quantum Well Infrared Photodetector Hand-Held Camera", IEEE Trans. Elec. Dev. 44(1), pp. 51 (1997S. Y. Auyang, P.A. Wolff, J. Opt. Soc. Am. B. **6**(4), 595 (1989).
2. S. Bandara, S. Gunapala, J. Liu, W. Hong, J. Park, Optical Coupling Mechanisms in Quantum Well Infrared Photodetectors" SPIE vol.2999, pp.103 (1997) and references therein.
3. K. Welford,"Surface plasmon-polaritons and their uses", Optical and Quantum Electronics 23, pp.1-27 (1991).
4. W.A. Wohlmuth, P. Fay, K. Vaccaro, E.A. Martin, I. Adesida, High-speed InGaAs Metal-Semiconductor-Metal Photodetectors with thin absorption layers, IEEE Photonics Technology Letters, vol.9 (5), pp. 654 (1997).
5. L. T. Claiborne, "Improved Performance III-V Quantum well IR photodetectors : Review of current and potential focal plane arrays" SPIE vol. 2999, pp. 94 (1997) and references therein.
6. J. Stiens, C. De Tandt, W. Ranson, R. Vounckx, and G. Borghs, Investigation of the depletion-modulation characteristics of low power 10.6 μm modulators based on the resonant plasma effect, Phot. Tech. Lett., **9** (6), pp. 770-772 (1997).
7. J. Stiens, W. Ranson, C. De Tandt, V. Kotov, G. Borghs, P. Muys, R. Vounckx, Resonant intracavity modulator for Q-switching a CO_2 laser in a new parameter regime, EOS-SPIE Conference on Industrial Lasers & Inspection, (14-18 June, 1999, Munich, Germany.
8. M.G. Moharam and T.K.Gaylord, «Rigorous coupled wave analysis of planar-grating diffraction», J. Opt.Soc. Am., 61,811-818 (1981).
9. T.K.Gaylord, M.G.Moharam, «Analysis and applications of optical diffraction by gratings», Proceedings of the IEEE,73,894-973 (1985).
10. J.A.Kong, «Second-order coupled-mode equations for spatially periodic media», J.Opt.Soc.Am.,67,825-829 (1977).
11. T.Tamir,H.C.Wang, and A.A. Oliner , «Wave propagation in sinusoidally stratified fielectric mefia», IEEE Trans. Microwave Theory Tech., MTT-12,323-335 (1964).
12. A. Hessel, A.A. Onliner, A new theory of Wood's Anomalies on Optical Gratings, Appl. Opt. 4(10), 1275-1297 (1965).
13. G. Shkerdin, J. Stiens, R. Vounckx, "Comparative study of the intra- and intervalley contributions to the free-carrier induced optical nonlinearity in n-GaAs", Journ. Appl. Phys. **85** (7), pp. 3807-3818 (1999).

Development and Application of Micro-Moiré Interferometer

Oh Kim Eng, Chai Gin Boay and Anand Asundi
Nanyang Technological University
School of Mechanical and Production Engineering
Nanyang Avenue, Singapore 639798

Chan Kai Chong and Chai Tai Chong
Institute of Microelectronics
11 Science Park Road, Science Park II
Singapore 117685

ABSTRACT

Consumers' demand for small, portable, increased functionality, more powerful and lower cost products has been the driving force in the technological advances of electronics packaging. This pushes the trend towards the development of smaller and higher power electronic packages. Due to the coexistence of variety materials with different Coefficient of Thermal Expansion (CTE) in the electronic packages, reliability becomes a concern, especially with smaller package sizes. The thermal mismatches often result in the delamination of interfaces between two materials, which eventually leads to ultimate mechanical and/or electrical failure. With its unique combinations of high sensitivity, optical contrast, range and good spatial resolution, Moiré Interferometry technique has increasingly been employed in the analysis of thermal deformation in electronic packages. However, with the commercial available moiré interferometer, spatial resolution of the system leaves much to be desired. The inability of commercial moiré interferometer to monitor local deformation such as at the die attachment level or the die corner thus limits deformation studies at those critical areas.

This paper presents the development of a fiber optic micro-moiré interferometer, which is capable of monitoring local deformation at micron level. This enables the study of micro deformation, which is the root of failure. Quad Flat Pack (QFP) and Flip Chip (FC) packages are used as specimens in this experiment. Only the QFP package is used for analysis in this paper. This system can also be used in applications like composite materials, residual stress analysis and micromechanical study of materials.

Keywords: Moiré Interferometry, micro-moiré, electronic packaging, thermal deformation, die attach level.

1. INTRODUCTION

In the area of stress analysis, there has been a persisting need for reliable tools to better understand and characterize material behaviour, reveal structural response to thermal loads, and determine distributions of stresses and strains in different geometries under various boundary conditions. Many approaches such as analytical, numerical or experimental have been applied to handle these problems. Numerical methodologies such as the finite element method or the boundary element can handle a wider class of problems that analytical models cannot solve. However, extreme care is needed as the choice of element size and shape, mesh design and boundary conditions cannot be reliably prescribed. Furthermore, both the analytical and numerical models also need validations, which are usually achieved by the experimental measurements. Experimental evaluations of stress and strain are not affected by assumptions made to facilitate the analytical and numerical solution. However, direct measurements of stresses are not feasible as they are calculated on the basis of the material properties and the knowledge of strain field. The experimental evaluation of strain is traditionally based on direct measurement of the associated displacement field. In order to measure the strain in specimen like the electronics packages or composites material, a highly sensitive experimental technique is required. Moiré interferometry which offer a unique combinations of high sensitivity, optical contrast and range, as well as good spatial resolution is a good choice.

2. BACKGROUND

Current applications of moiré interferometry are in the microelectronics industry but scope also exists for deformation analysis in micromechanics and biomechanics engineering. Also since many components are getting smaller in

Part of the SPIE Conference on Advanced Photonic Sensors and Applications
Singapore • November–December 1999
SPIE Vol. 3897 • 0277-786X/99/$10.00

size, micro-measurement techniques are needed for the deformation analysis. This paper proposes a micro-moiré interferometric system, which provides greater spatial resolution such that strains within the die attachment level can be monitored. In this paper, moiré interferometer is restricted to study the in-plane deformation of electronics packages. The electronics packages are replicated with high diffractive efficiency grating at 100°C and then cooled slowly to room temperature. The commercial moiré interferometer is used to obtain the whole field moiré fringe patterns of specimen due to the isothermal loading. However, with the trend of advancement in developing smaller and more functional packages, the reliability of the packages becomes very critical. The limitation of the commercial system to monitor the deformation at the micron level is highlighted in Figure 1 and 2. The figures show the moiré fringe patterns of a QFP package. One of the critical regions is at the die corner and copper leadframe interface of 25-50μm. The fringes appear to concentrate at the die corner with a large change in fringe gradient. Deformation analysis is thus hindered due to the shortcomings of the system to capture the local deformation at the critical region where the moiré fringes concentrate.

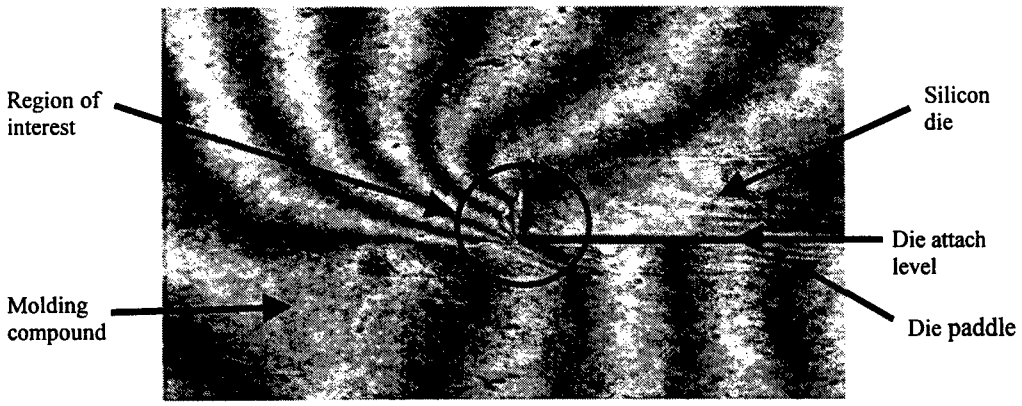


Figure 1 U-field moiré fringe pattern at the die/die paddle interface obtained by the commercial moiré interferometer.

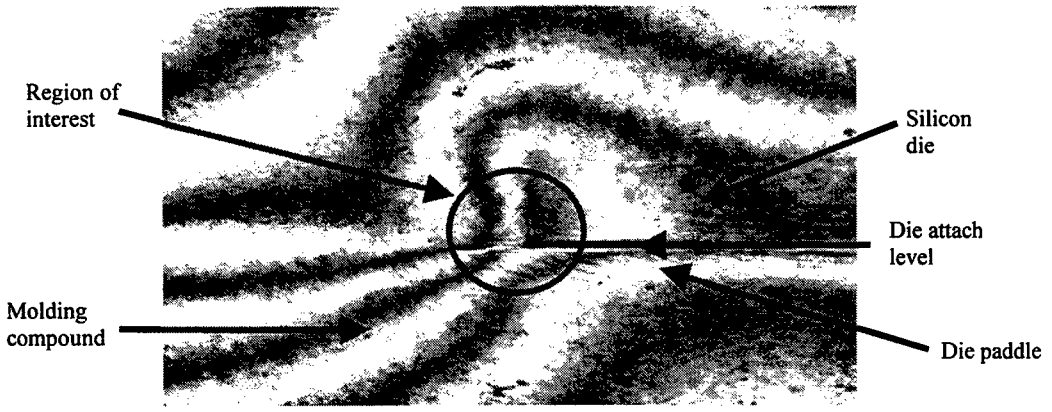


Figure 2 V-field moiré fringe pattern at the die/die paddle interface obtained by the commercial moiré interferometer.

3. MOIRÉ INTERFEROMETRY

Moiré interferometry is an optical experimental technique that provides high sensitivity on the in-plane deformation measurement. Moiré interferometry involves optical diffraction and interference principles to measure deformation in the nanometer range. The basic principle of moiré interferometry is illustrated schematically in Figure 3. In this method, the high frequency specimen grating is adhered or printed on the specimen surface. A virtual reference grating created by two coherent beams are used to interact with the deformed specimen grating to form the moiré fringe patterns. In practice, the virtual reference frequency of 2400 lines/mm is used, which provides a basic sensitivity of 0.417μm. The moiré

patterns depict contours of displacement component in a direction perpendicular to the grating lines. From the fringe patterns, displacement is obtained and the derivative of the displacement is the strain component. The displacements can be determined from the fringe orders by

$$U = (1/f)\, N_x\,, \qquad V = (1/f)\, N_y\,, \tag{1}$$

where N_x and N_y are the fringe orders in the U and V field patterns and f is the frequency virtual reference grating. When strains are to be obtained, it can be calculated from the measured displacement fields

$$\varepsilon_x = \partial U/\partial x = (1/f)\,[\partial N_x / \partial x],$$

$$\varepsilon_y = \partial V/\partial y = (1/f)\,[\partial N_y / \partial y], \tag{2}$$

$$\gamma_{xy} = \partial U/\partial y + \partial V/\partial x = (1/f)\,[\partial N_x / \partial y + \partial N_y/ \partial x].$$

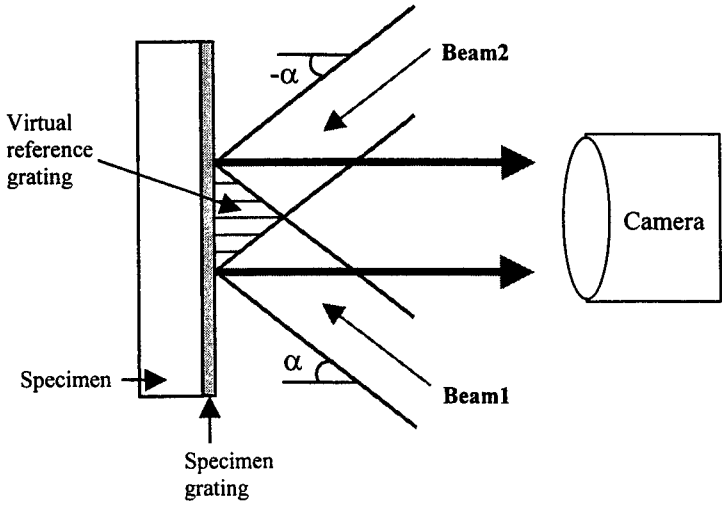


Figure 3 Schematic diagram of Moiré Interferometry.

4. EXPERIMENTAL RESULTS AND DISCUSSIONS

4.1 Specimen Preparation

The specimen used for this experiment is Quad Flat Pack package (QFP). The specimen was initially halved and then ground to expose the internal components of the package. Figure 4 shows the schematic diagram of the QFP package. A uniform cross-line grating on ultra low expansion (ULE) substrate with grating frequency of 1200 lines/mm was for the grating replication process. This is to ensure that the frequency of the specimen grating did not change at the elevated temperature. Both the specimen and the grating mold were heated to the elevated temperature and stayed at that temperature for at least two hours. This was to ensure same temperature throughout the specimen and the mold, and also to make the sample relax sufficiently. Replication steps include spreading of adhesive, curing and prying of specimen from the mold. The grating was transferred onto the QFP package at 100°C. Immediately after the prying of specimen from the grating mold, the specimen was cooled to room temperature (25°C). The specimen was thus subjected to an isothermal loading of ΔT = -75°C. The grating that was adhered onto the specimen would deform with the specimen during the cooling down process. The replication process is completed when the specimen cooled down to room temperature. The deformation of the specimen was then recorded by the moiré interferometer at room temperature. The recorded moiré fringes showed deformation that was induced by the change of temperature. The specimen was placed in the micro-moiré interferometer and aligned. Figure 5 and 6 show the moiré fringe patterns obtained from the micro-moiré interferometer. The vital

shortcoming of the commercial moiré interferometer is improved with the micro-moiré interferometer set-up. This can be compared from Figure 1,2,5 and 6. The limited spatial resolution of the commercial moiré interferometer limits the resolving of deformation at the critical corner of the electronic packages where cracks tend to initiate. The micro-moiré interferometer, however, can resolve to the critical corners. High strain concentration can be noticed from the bottom corner of the silicon die and die paddle interface.

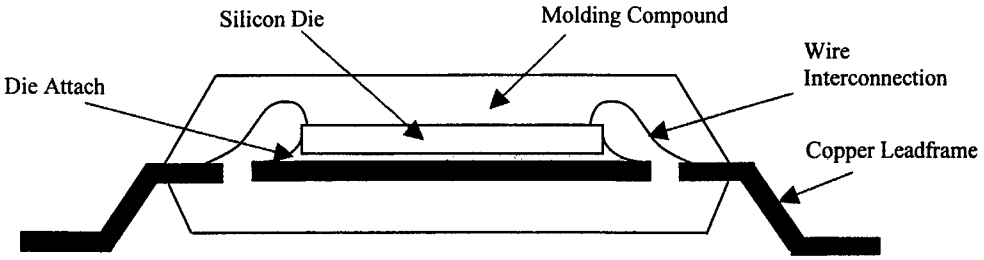


Figure 4 A schematic diagram of QFP.

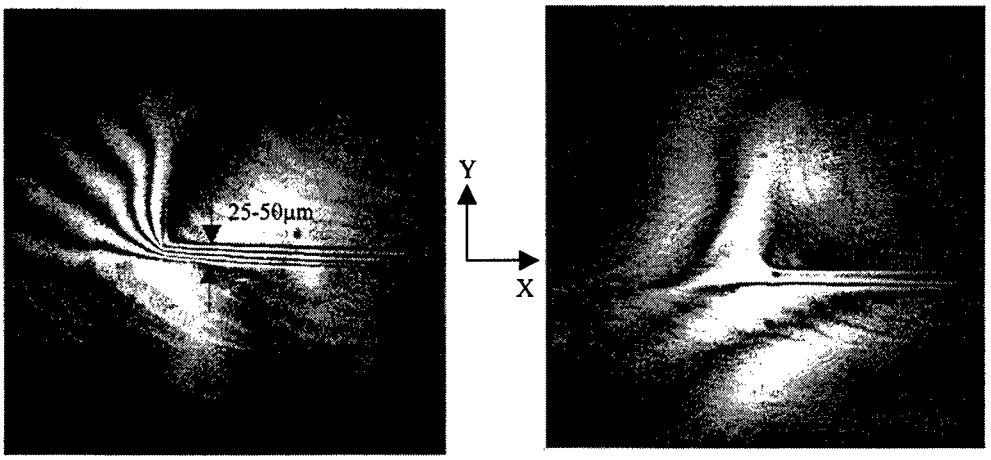


Figure 5 U-field moiré fringe patterns at the corner of the die/die paddle interface obtained by the micro-moiré interferometer.

Figure 6 V-field moiré fringe patterns at the corner of the die/die paddle interface obtained by the micro-moiré interferometer.

4.2 Strain Analysis

The recorded moiré fringe patterns, which represent the isothermal loading of Δ=-75°C, were used to obtain the strain values from the critical corner to the center of the package. Due to symmetry, analysis was performed from the die corner to the center of the package.

Figure 7 shows the distribution of normal strain dU/dx value at the lower corner of the silicon die. It can be seen that the strain value increases gradually to a peak value of 0.007mm/mm at the bottom corner of the silicon die and die attachment interface. After which, it decreases drastically towards zero. This can be noticed from the fringe pattern, as the fringes are highly concentrated at the region outside the die corner. The fringes then flow in a horizontal gradual manner into the interface of the die and the die paddle where dU/dx is zero.

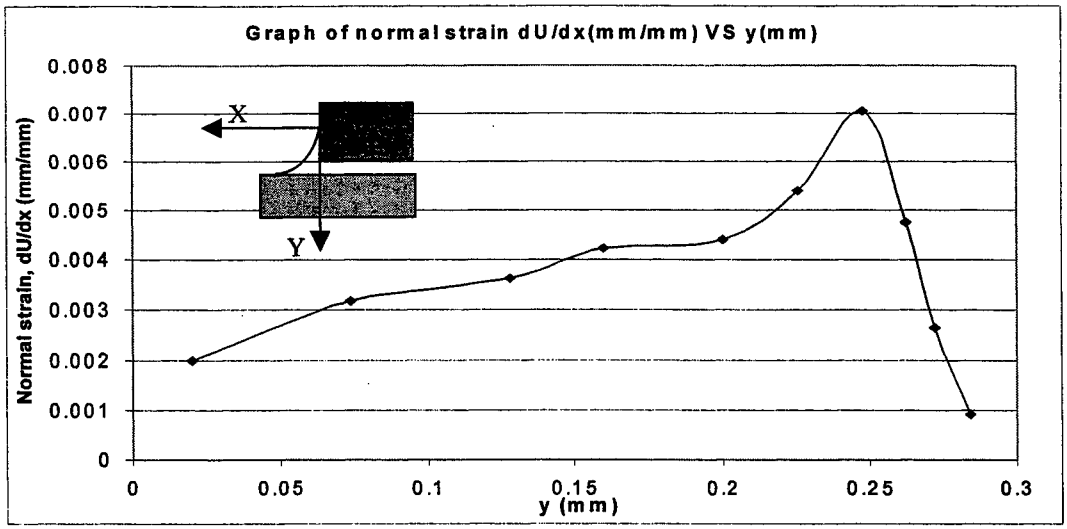


Figure 7 U-field normal strain at the corner of the of the silicon die.

Figure 8 shows the distribution of normal strain dV/dy at the lower corner to the interface along the center of the silicon die. Strain value increases steeply from the outside region of the die corner and then gradually to almost of a constant value at the center of the interface between the die and the die paddle.

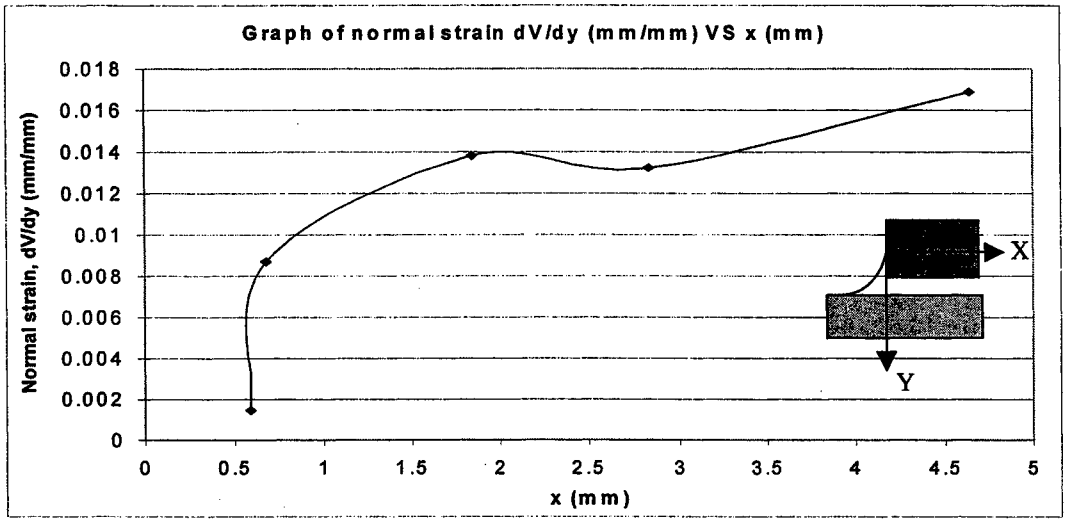


Figure 8 V-field normal strain at the corner to the center of silicon die attachment interface.

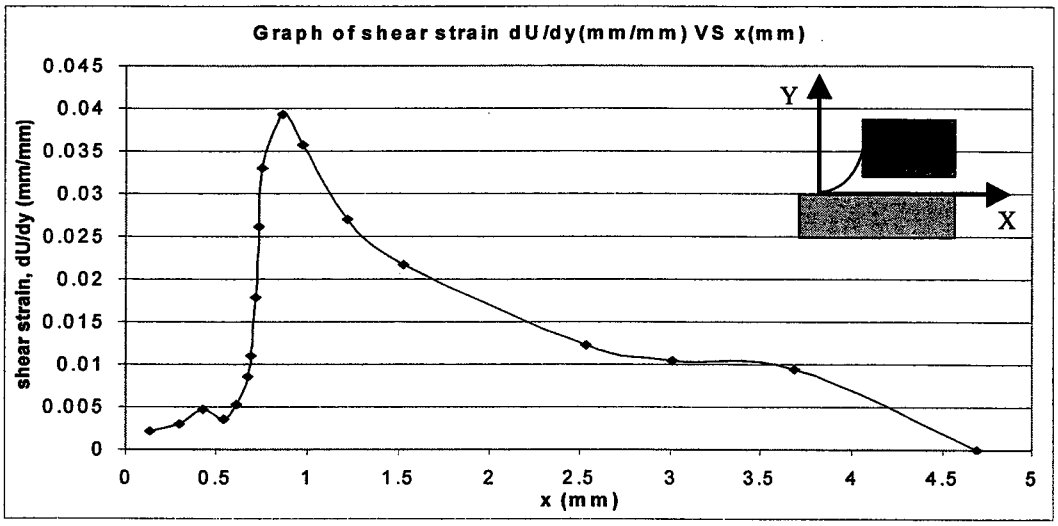


Figure 9 U-field shear strain from the corner to the center of silicon die attachment interface.

The graph in Figure 9 displays the shear strain component dU/dy, from the corner of the silicon die to the center of the package along the interface. Very high dU/dy value was obtained at the region just outside the die corner interface. However, the shear strain component diminishes gradually within the interface to the middle of the package.

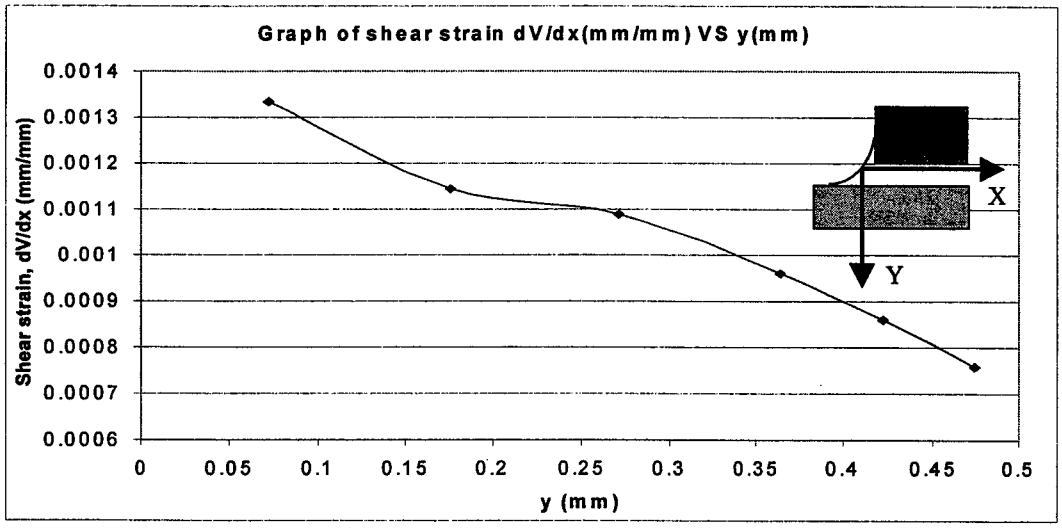


Figure 10 V-field shear strain at the corner of the die.

The shear strain component, dV/dx is relatively high only at the outside region of the die corner is highlighted in Figure 10. The dV/dx strain value decreases relatively gradually as it approaches the die corner. From Figure 6, the component approaches to zero at the die attach level interface towards the center of the specimen.

4.3 Discussion

Residual stress/strain is a result of the different of material properties in electronic packaging. When the packages are assembled at high temperature, as the temperature cools down to room temperature, residual stress/strain will develop especially in the bonding interface. In the manufacturing process of QFP packages, die attach bonding is used to mechanically attach the silicon die to the leadframe. Other functions of die attach materials include heat transfer from the dies to the leadframe and if needed to provide electrical contact to the backside of the die. After the attachment process, the specimens are send to oven for curing the die attach material after which, will develop the required bonding strength. Wire bonding process is next to provide interconnection of the active circuits to the leadframe. The subsequent process is the transfer molding process in which the specimens are encapsulated by thermoset compound. All the manufacturing processes are performed in high temperature environment, as high as 230°C in the soldering stage. The CTE of the silicon die and the copper leadframe is approximately 3ppm/°C and 18ppm/°C respectively. With the difference in material properties, thermomechanical stresses tend to develop at the bonding interface. Thus the selections of die attach materials and the leadframe materials play an important part in the production of reliable and durable electronic components. Thus, the leadframe chosen should have a CTE close to that of the silicon die to reduce the thermomechanical stress at the die leadframe interface, and the die attach used should provide very good adhesion properties to the die and the leadframe. This will largely reduced the thermal expansion mismatch between the die and the leadframe.

From the moiré fringe patterns in Figure 5 and 6, the fringe patterns develop by the isothermal deformation show that the die attach material is providing the cushioning effect to reduce the thermomechanical stress between the two different materials. Although the isothermal thermal loading is only performed from 100°C to room temperature, it is actually simulating the residual stress/strain effect of Δ=-75°C during the cooling down effect after the electronic packages are remove from the postmold curing process. The residual strain measured here is in-plane and is determined by using moiré interferometry. Deformations measured on the specimen represent the thermal deformations of the specimen induced by the temperature change. To fully determine the residual strain, the specimens need to be heated to the assembly temperature and then measured at room temperature when the residual strain are fully developed. In practice, the sample can be replicated at an elevated temperature very near the assembly temperature.

Besides the application in QFP, the micro-moiré interferometer can be used to study the solder bump of FC packages. The bumps have diameter of approximately 75-100μm. Figure 11 and 12 show the moiré fringe pattern of a solder bump of a FC package without underfill. Figure 13 and 14 show the micro-moiré interferometer being used to study the residual stress of a composite material by hole drilling method.

Figure 11 U-field moiré fringe pattern of a solder bump of a FC package without underfill of isothermal loading Δ=-75°C.

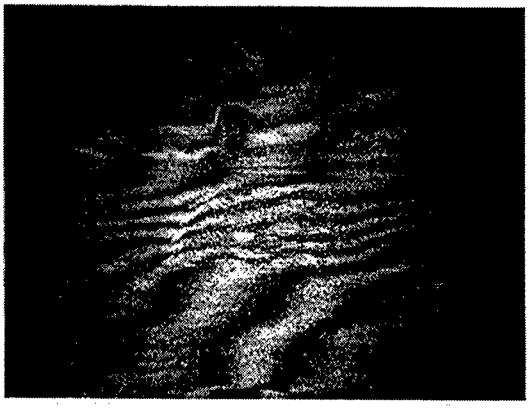


Figure 12 V-field moiré fringe pattern of a solder bump of a FC package without underfill of isothermal loading Δ=-75°C.

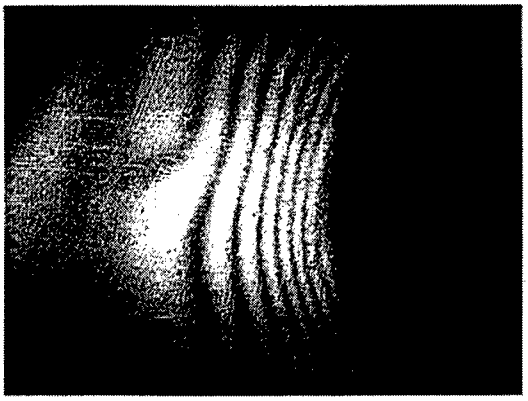


Figure 13 U-field moiré fringe pattern of a composite by hole drilling method.

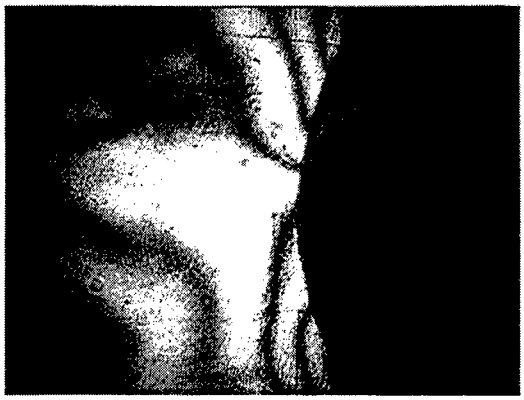


Figure 14 U-field moiré fringe pattern of a composite by hole drilling method.

5. CONCLUSION

A micro-moiré interferometer system is developed to overcome the shortcomings of the commercial moiré interferometer such as the limited spatial resolution. This enable the analysis on the deformation of QFP package due to the thermomechanical stress induced by the temperature change at the die corner and bonding interface of the silicon die and the leadframe. Besides the QFP, moiré fringe patterns of solder bump of FC packages is also obtained. The micro-moiré interferometer can also be applied to other applications such as the residual stress analysis of the composite material and biomechanics.

In summary, the micro-moiré interferometer provides several advantages that overcome the shortcomings of the commercial moiré interferometer. The preliminary specifications are listed below:

1. Displacement sensitivity of 0.417μm per fringe order with no fringe processing.
2. Fringe displacement resolution of approximately 10μm.
3. Imaging system resolution of 2μm/pixels.
4. Strain sensitivity range of 2.71E-04 to 4.17E-02 μm/μm.
5. Fiber optics beam delivery system with no mirrors or optics.
6. Digital imaging system for on-line acquisition.
7. Compact size, low weight and flexible.

ACKNOWLEDGEMENTS

The authors would like to express their appreciation to Research Fellow, Dr. Zhao Bing and Dr. Xie Guang Ping who have given some valuable advice to this work. The support of NTU through grant MLC1/97 is acknowledged. Finally, the support of IME for instrument development is also acknowledged.

REFERENCES

1. Post, D., Han, B., and Ifju, P. (1994). High Sensitivity Moiré: Experimental Analysis for Mechanics and Materials. New York. Springer-Verlag.
2. Han, B., Guo, Y., and Lim, C. K. (1995). Moiré interferometry as a tool for engineering education. Manuscript for ASME. Intersociety Electronic Packaging Conference & Exhibition. Lahaina. Hawaii.
3. Guo, Y., and Sarihan, V. (1997). Residual strain/stress in electronic packages due to assembly process conditions. ASME International Mechanical Engineering Conference and Exposition. Dallas, Texas, 1997.
4. Han, B. (1998). Recent advancements of moiré and microscopic moiré interferometry for thermal deformation analysis of microelectronics devices. Experimental Mechanics. Vol. 38(4). Dec. pp. 278-288.

Correspondence: Email: p7220775j@ntu.edu.sg;
Telephone: (65) 790 5564
MPE Strength of Materials Laboratory;
Fax: (65) 792 4062.

Photoexcitation induced current sensing on semi-insulating GaAs using a tunneling microscope tip

Kenji Kawashima*, Shinji Takai, Gou Kudou, Hideo Adachi, Misaichi Takeuchi, Kenzo Fujiwara

Department of Electrical Engineering, Kyushu Institute of Technology
Tobataku, Kitakyushu 804-8550, Japan

ABSTRACT

Detailed tunneling current measurements using a tunneling microscope tip have been performed on semi-insulating GaAs surfaces as a function of illumination power, Pt/Ir tip-surface distance, and separation between the tip and In/Ga electrode on the sample surface to elucidate previously unsolved problems of illumination-induced thermal expansion effects on the probe and of the surface depletion effects. We show that the tip-sample distance to detect a constant tunneling current is extended with increasing the optical excitation power. It is also found that the photo-induced tunneling current as high as 8 nA driven by a 746 nm laser diode is linearly proportional to the optical excitation power (2-350 μW/cm^2). This photo-induced carriers conduction is also confirmed by studying the transient photocurrent response, which is slowed down by increasing the tip-to-electrode distance. These results reveal that, in our case, the thermal effects are negligible and photogenerated electron tunneling is a dominant mechanism for the increased tunneling current from the sample surface biased at negative voltages relative to the tip.

Keywords: tunneling current, photocurrent, GaAs, STM

1. INTRODUCTION

Scanning tunneling microscope (STM)[1] has been a powerful technique to investigate various nano-scale informations such as topography, chemical composition and electronic property of surfaces for conductive materials (metals and semiconductors) but not for insulating materials. To find a way overcoming the difficulty for the insulating materials, the use of optical excitation to generate carriers has been received a great deal of attention.[2,3] Previously, it is known that, photocurrent spectroscopy (PCS)[4,5] is a convenient and highly sensitive technique to study the electrical and optical properties of semiconductor materials. Therefore, STM combined with PCS has attracted interest as a new approach to study the electro-optical properties of the semiconductor materials in a nano-meter scale.

The photo-assisted STM has been used to study the electrical, optical and geometrical properties of semiconductor nano-structures including heterojunctions and quantum well structures.[6-8] However, mechanisms for the enhancement of tunneling current by optical excitation are still a matter of controversy. One of the possible mechanisms is thermal expansion of the probe tip and the sample, which reduces the tip-sample distance.[9,10] Another possible mechanisms are originated from photocarriers, which are generated by irradiated photons whose energies are higher than the band gap of semiconductor samples. These photocarriers lead to shrinkage of a surface depletion layer,[11] increase a surface charge density, and/or reduce the sample resistance.

In this paper, we have carried out tunneling current measurements on semi-insulating GaAs surfaces as a function of illumination power, Pt/Ir probe tip-surface distance, and separation between the tip and the In/Ga electrode formed on the sample surface to clarify leading mechanisms for the photo-induced tunneling current enhancement. We show that the tip-sample distance producing the same current intensity is extended with increasing the laser irradiation intensity. In general, the extension of the tip-sample distance was estimated by the variation of the electrical feedback signal under the constant current mode in conventional STM measurements. However, in our experiments, we find the extension of the tip-sample distance by measuring the increase of movable length (namly, tip-sample separation where the tunneling current signal shows a logarithmic behavior) with increasing the illumination intensity. We also show a linear dependence of the photo-induced tunneling current on the optical illumination intensity. These results indicate

*Correspondence: E-mail: kawa@ele.kyutech.ac.jp; Fax: +81-93-884-3217

Part of the SPIE Conference on Advanced Photonic Sensors and Applications
Singapore • November–December 1999
SPIE Vol. 3897 • 0277-786X/99/$10.00

that the increase of the photogenerated charges on the sample surface leads to a direct enhancement of the tunneling current.

2. EXPERIMENTAL DETAILS

Our experimental setup operative in air is schematically shown in Fig. 1(a). We have used a block-type piezo for coarse positioning of a sample holder sliding on a smooth surface of a glass plate. A tube-type piezo was used for fine positioning of the tip relative to the sample surface. The positioning apparatus was set on a metal-rubber stack to reduce vibrations. Figure 1(b) shows a schematic diagram of electrical and optical arrangements to measure photo-assisted tunneling currents. For the preparation of the probe tip, a 0.25 mm diameter Pt/Ir wire was cut mechanically. As a sample, a semi-insulating GaAs wafer was cleaved parallel to a (110) surface. The cleaved facet piece whose size is about 0.35 mm × 6 mm was irradiated with a diode-laser light. On the both ends of the sample surface, upper and lower electrodes were formed by In/Ga solder. These two electrodes were used to measure current-voltage (I -V) characteristics to confirm the ohmic contacts and to estimate the sample resistance. Using one of the electrodes, we applied a positive voltage of 2.0 V to the tip relative to the sample and detected the current using a pico-ampere meter with a voltage application function.

We have basically carried out two kinds of measurements (experiment A and B). In the experiment A, we have investigated the dependence of tunneling currents on the Pt/Ir tip to sample surface distance at several values of the illumination power using a diode laser at 785 nm. At first, an electrical feedback circuit was set to control the tip-sample distance by detecting the tunneling current, and the tip was moved close to the sample to get a target value of the tunneling current (I_0) under a high-power illumination intensity (P_1) as illustrated by Fig. 1(c). After disabling the feedback system, the tip was carefully approached to the sample surface within the distance where the detected current (I) shows typical semi-logarithmic tunneling-current characteristics. Here, we define the distance (scaled by the voltage applied to the piezo stage) as a movable length ($z_m(P_1)$) where the current always shows a semi-logarithmic behavior. When a low power light (P_2) irradiates the sample, the probe will be closer to the sample to obtain a enough target current value. Consequently, a short movable length ($z_m(P_2)$) will be

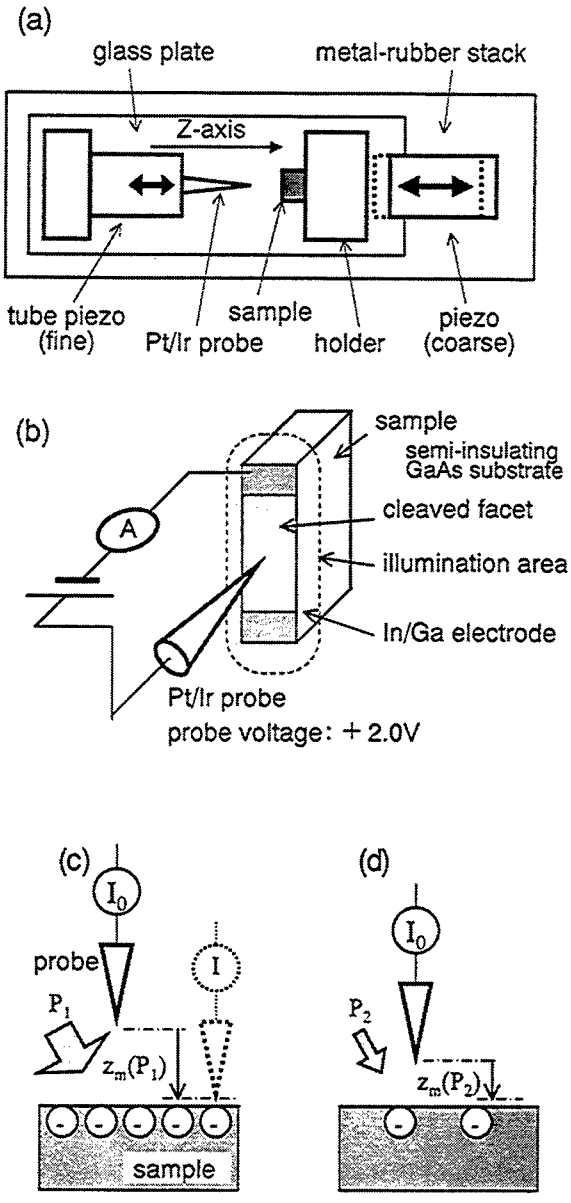


Fig. 1: (a) Schematic diagrams of the experimental setup for detection of photo-assisted tunneling current. (b) Electrical and optical arrangements of the tip and the sample. (c) Experimental procedure to measure the movable length ($z_m(P_i)$) under high-power illumination (P_1) and (d) low-power illumination (P_2).

resulted, as shown by Fig. 1(d). By checking the movable length as a function of illumination intensity, we can clarify the photo-assisted effect on the tunneling current to eliminate the possible photo-thermal expansion. In these measurements, the tip was located almost at the center of the sample surface, namely, the tip-to-electrode distance was set to be ~ 3 mm.

In the experiment B, the tip was approached to the sample to get the target value without the laser illumination by the electrical feedback system. After disabling the feedback system, we measured the dependence of the tunneling current on the illumination power using a 746 nm pulsed laser with a duration of about 5 sec. In this experiment B, we also measured transient photocurrent responses at different tip-to-back-electrode distances by keeping a constant optical power and changing the relative position between the tip and the electrode of the sample surface.

3. RESULTS AND DISCUSSION

3.1. Basic characteristics of experimental system

Before performing the main experiments, we have checked the $I-V$ characteristics of the semi-insulating sample between the upper and lower electrode (see Fig. 1(b)) under different illumination intensities. Inset of Fig. 2 indicates the $I-V$ characteristics without illumination. Although the sample is a semi-insulator, we achieved the nearly symmetric ohmic contacts for both electrodes. Using the measured $I-V$ characteristics under the laser irradiation, we calculated the sample resistance, and plot it as a function of optical power density in Fig. 2.

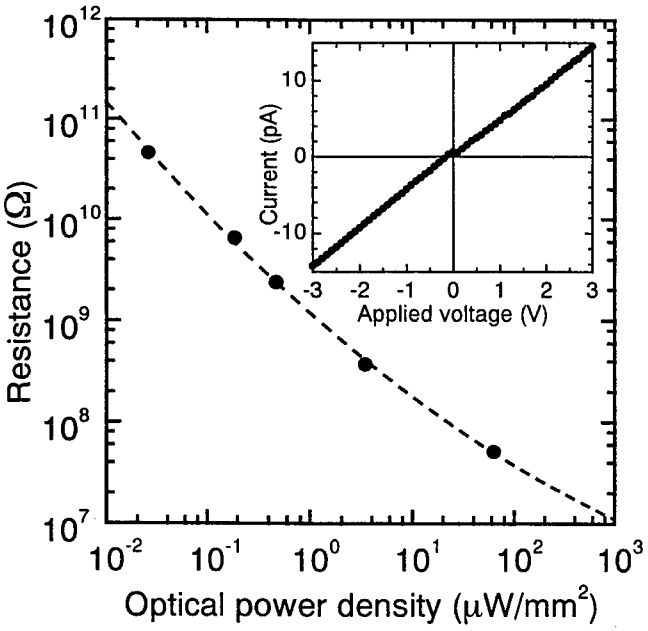


Fig. 2: Sample resistance as a function of illumination power density, calculated by the $I-V$ characteristics under several illumination levels. Inset shows the $I-V$ characteristics without illumination.

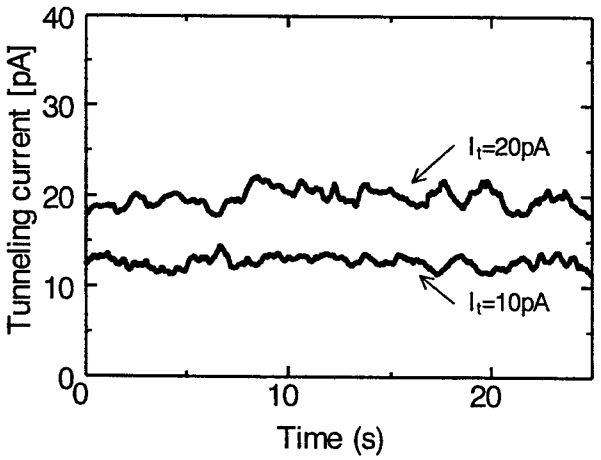


Fig. 3: Stability of the tunneling current intensity without using the feedback system. The bottom and top curves show the tunneling currents when the target values were set to be 10 pA and 20 pA, respectively.

To confirm the stability of our experimental system, we have also measured fluctuations of the tunneling current after the feedback system was turned off. Figure 3 shows the tunneling-current intensities measured when the target current intensities were set to be 10 pA (bottom curve) and 20 pA (top curve). Observed fluctuations of a few pico-amperes of the tunneling current are able to ignore in our experiments. These fluctuations probably arise from residual mechanical vibrations, thermal drifts, and/or electrical noises in the apparatus.

3.2. Extension of movable length by laser irradiation

In order to check the movable length as a function of illumination intensity, we have measured dependence of the tunneling current intensity (I_t) on the tip-sample distance by changing the voltage applied to the piezo scanner under different illumination intensities. Figure 4 shows several modulation patterns of the applied voltage to the scanner (dV_{PZ}) and the corresponding expansion length of the tube piezo (dZ) as a function of sweep time (t). Here, the expansion ratio to the voltage of the tube scanner (2.08 nm/V) is the theoretical value calculated by using parameters depending on the material, structure, and electrodes-configuration of the tube piezo.

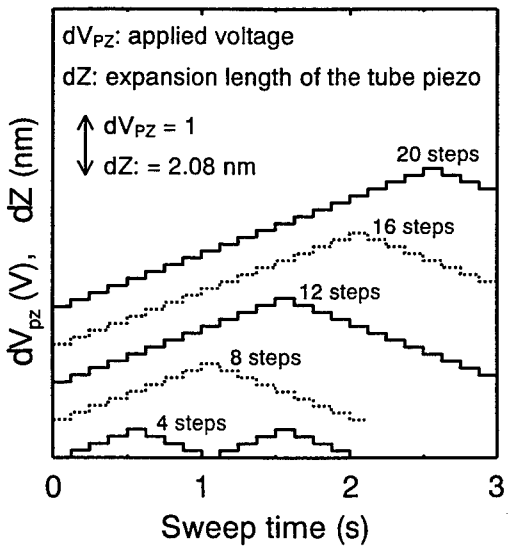


Fig. 4: Patterns of voltage modulation applied to the piezo scanner (dV_{PZ}) and the resultant expansion length of the tube piezo (dZ) as a function of sweep time.

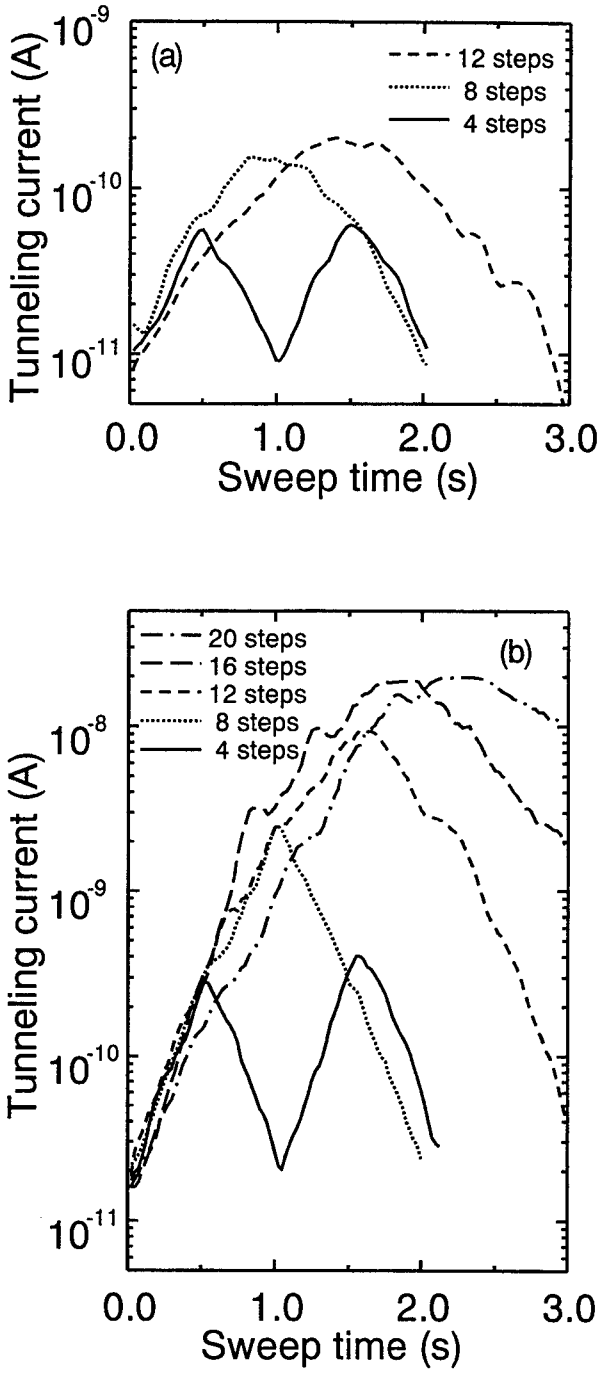


Fig. 5: Tunneling current under different modulation patterns. Illumination level at 785 nm is (a) 2.2×10^{-4} mW/mm^2 and (b) 2.1 mW/mm^2. Abscissa (sweep time) is equivalent to the piezo length (dZ), if we assume one-to-one correspondence between piezo voltage and length.

Figures 5(a) and 5(b) show intensity variations of the tunneling current under the different voltage-modulation patterns at a low illumination power density of 2.2×10^{-4} mW/mm^2 and a high illumination power density of 2.1 mW/mm^2, respectively. In these experiments, the target value of tunneling current without illumination was set to be 10 pA. In each figure, the abscissa (sweep time) is able to convert into the expansion length of the tube piezo by using Fig. 4.

As it is known the tunneling current shows the exponential dependence on the tip-sample distance (z), namely $I_t \propto \exp(-a\phi^{1/2}z)$, where a is the constant and ϕ the average value of the work functions of the tip and the sample materials.[12] In Fig. 5(a), when the tube piezo was extended or shrunk with the 4-step modulation pattern, the detected current shows an inverted V shape in response to the first cycle of the modulation pattern. This result shows that the exponential dependence is obtained in the tunneling current versus sweep time characteristics $(I_t - t)$. Because the $I_t - t$ characteristics is equivalent to $I_t - dZ$ characteristics, it is clear that the tunneling current can be detected when the tip approaches to the sample by about 4 steps. While in the large-step modulation patterns, the $I_t - t$ characteristics show saturation behaviors and do not reveal the exponential dependence. Therefore we estimate the tip-sample distance to yield the tunneling current intensity of 10 pA is about 0.8 nm under the low-power illumination intensity of 2.2×10^{-4} mW/mm^2. With increasing the illumination intensity, the exponential behaviors were observed in a larger region of the $I_t - dZ$ characteristics. In Fig. 5(b), a much finer exponential dependence is observed in the $I_t - dZ$ characteristics for the 12-step modulation. In this case which use the high-power illumination of 2.1 mW/mm^2, the tip-sample distance is expanded to about 2.3 nm. These results indicate that the tip-sample distance producing the same current intensity, namely the movable length, is extended by the laser illumination.

Figure 6 shows the measured movable length as a function of optical power density. A dashed line is drawn by applying a least-squares method. According to J. Tersoff and D. R. Hamann,[13] the tunneling current is proportional to the density of states at the tip (ρ_T) and the sample surface (ρ_S), namely, $I_t \propto \rho_T \rho_S$. We simply assume that ρ_T is a constant value and ρ_S is proportional to the illumination intensity (P). If we further assume that tunneling amplitudes of the wave function occupied by the photogenerated charge on the sample surface decays exponentially

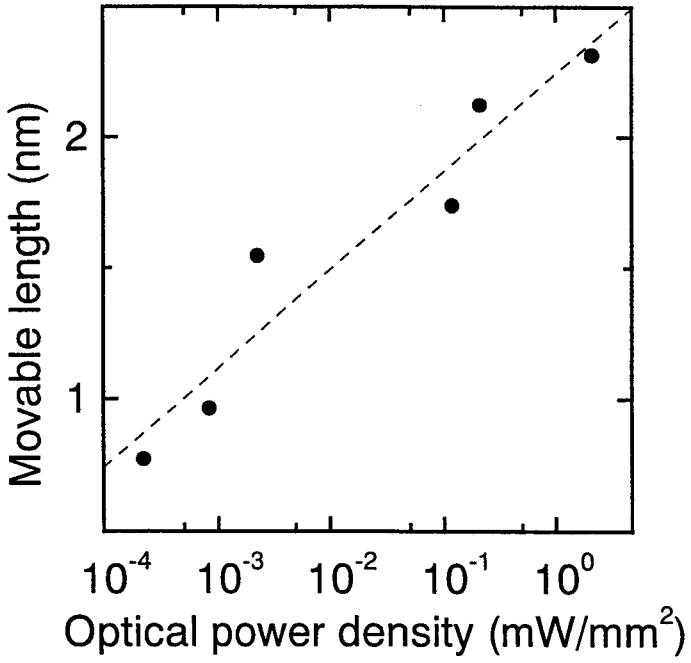


Fig. 6: Movable length as a function of illumination intensity. Dashed line is drawn by applying a least-squares method.

with the distance, the tunneling-current intensity can be expressed by $I_t = \beta P \exp(-\alpha z)$, where α and β are constants. This results in a relation of $z_m(P) \propto \ln(P)$ whose tendency agrees with the dashed line in Fig. 6. Therefore we claim that the increase of the photogenerated charges on the sample directly enhances the tunneling current. Concerning about the effect of the shrinkage of the surface depletion layer on the tunneling current, it can be omitted because the Fermi levels due to unintentionally-doped impurities are pinned at the mid gap in semi-insulating GaAs samples.[14]

With increasing the modulation steps in Figs. 5(a) and 5(b), the $I_t - dZ$ characteristics become sublinear in the logarithmic plot and show saturation behaviors. Origin of these characteristics is not clear, however, we consider that they may arise from the tip in contact with a contamination layer, a water film or an oxide layer formed on the sample surface during the measurements. In the case where the tip and the surface layers are in contact, the current intensity will increase with increasing the contact surface area and finally reach a saturation value determined by the contact resistance and the sample resistance. Thus, the $I_t - dZ$ characteristics do not show the exponential dependence in the large-step modulation of dZ.

The saturation value of the detected current under the high power illumination (estimated to be ~ 20 nA) is larger than that of the low power illumination (~ 0.2 nA). The increment of the saturation current level gives evidence for the reduction of the sample resistance due to the increase of the charge carriers by the optical excitation. However, the relative value of the current saturation level is not proportional to the illumination intensity. From Fig. 2, the sample resistance between the tip and the electrode (~3 mm) at the high and low illumination powers is estimated to be ~10 MΩ and ~2 GΩ, respectively. By assuming the voltage drop at the sample surface, the saturation current level at the high and low illumination powers is calculated to be 200 nA and 1 nA, respectively. Both calculated saturation values are found to be smaller than the measurement ones. This result suggests that the resistance is controlled by tunneling to the tip which is not in contact with the sample surface. These sublinear and saturation characteristics are not clearly explained, however, we should point out again that the fine exponential dependence below the sublinear and saturation behaviors develops with the well-defined $I_t - dZ$ characteristics by increasing the optical excitation. That is, the tunneling current is enhanced by the laser illumination.

3.3. Dependence of tunneling current on optical illumination power

Figure 7 shows the measured tunneling current intensity as a function of optical intensity (P_P). A dashed line is drawn by a least-squares method. If the photo-thermal expansion is dominant, reduction of the tip and the sample distance is proportional to the optical intensity. Consequently, the photo-thermal effect by the laser irradiation should result in the exponential dependence in the $I_t - P_P$ characteristics.[7] However, the observed $I_t - P_P$ characteristics shows the linear dependence. This result confirms that the enhancement of the tunneling current is caused by the increment of the charge carriers induced by the laser illumination.

Typical tunneling current responses to the optical pulses are shown in Fig. 8. An upper trace indicates an interruption timing of the optical pulse. The optical power density at the on-state was set to be 0.471 μW/mm^2. Two lower traces show normalized tunneling current characteristics on a semi-log scale when the distances from the tip to the In/Ga electrode are set to be dL = 1 mm and dL = 3 mm, respectively. The actual tunneling current level with the laser illumination at dL = 1 mm is about 160 pA, which is 30 times greater than that at dL = 3 mm. When the optical pulse is turned off, the current intensity in the both cases should, ideally, recover to the same value (to the target value of 10 pA). However, in this experiment the actual current intensity at dL = 1 mm was slightly greater than that at dL = 3 mm. The decay time of the tunneling current in our experiments is roughly estimated to be a several tens of secs. These slowly decay characteristics are probably caused by a large CR time constant of the electrical circuit in the experimental setup due to large values of the sample resistance and an internal resistance of the pico-ampere meter. According to the $I-V$ measurements of the sample under the illumination intensity of 0.471 μW/mm^2, the resistance values of the sample were calculated to be about 400 MΩ and 1.2 GΩ at dL = 1 mm and dL = 3 mm, respectively. The internal resistance of the pico-ampere meter was 1 GΩ. These resistance and the decay time suggest that capacitance of our experimental circuit is about a few tens of pF. Figure 8 also indicates that the decay is slower when the tip-electrode distance is larger. This dependence is probably caused by the longer transport length of photogenerated charges and the resultant larger resistance when dL = 3 mm compared to when dL = 1 mm.

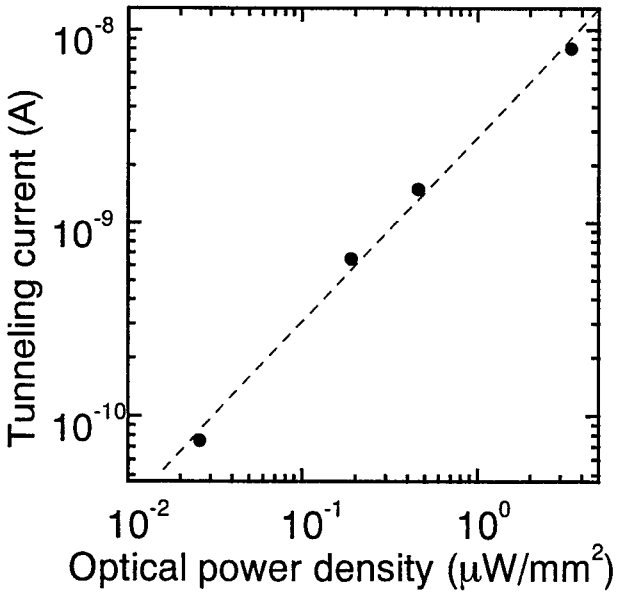


Fig. 7: Tunneling current as a function of optical power density using a pulsed laser at 746 nm with a duration of about 5 sec. Dashed line is obtained by appling by a least-squares method.

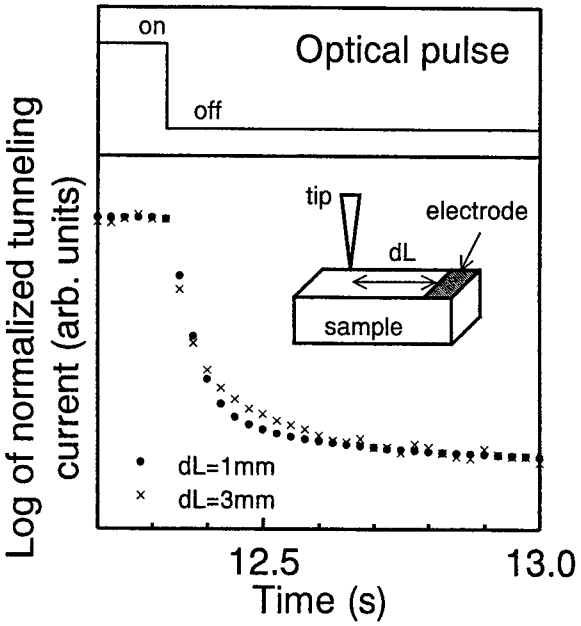


Fig. 8: Tunneling current responses to the optical pulse excitation of a power level of 0.471 μW/mm^2. Upper trace indicates interruption-timing of the optical pulse. Lower traces show normalized tunneling current characteristics on a log scale at two different tip-electrode distances, dL = 1 mm (dots) and dL = 3 mm (crosses),

4. CONCLUSIONS

Detailed tunneling current measurements have been performed on semi-insulating GaAs surfaces as a function of illumination power, Pt/Ir tip-surface distance, tip-to-back In/Ga contact separation to elucidate previously unsolved problems of illumination-induced thermal expansion effects on the probe and of the surface depletion effects. We have shown that the tip-sample distance to detect a constant tunneling current is extended with increasing the optical excitation power. It is also found that the photo-induced tunneling current as high as 8 nA driven by a 746 nm laser diode is linearly proportional to the optical excitation power (2-350 μW/cm^2) at several constant dark, purely tunneling current levels, which reveals negligible thermal effects in our case but photogenerated electron tunneling from the sample surface biased at negative voltages relative to the tip. This photo-induced charge conduction is also confirmed by studying the transient photocurrent response, which is slowed down by increasing the tip-to-electrode distance. These results indicate that absolute photo-induced current sensing is inherently useful to dynamically probe the atomic-scale photo responses of nanostructures.

5. ACKNOWLEDGMENTS

This work was supported in part by Grant-in-Aid for Scientific Research (B) (No. 10555106) from The Ministry of Education, Science, Sports and Culture of Japan.

6. REFERENCES

1. G. Binning, H. Rohrer, Ch. Gerber, and E. Weibel, "Surface studies by scanning tunneling microscopy," Phy. Rev. Lett., **49** (1), pp. 57-91, 1982.
2. G. F. A. van de Walle, H. van Kempen, and P. Wyder, "Scanning tunneling microscopy on photoconductive semi-insulating GaAs," Appl. Phys. Lett., **50** (1), pp. 22-24, 1987.
3. S. Akari, M. Ch. Lux-Steiner, M. Vogt, M. Stachel, and K. Dransfeld, "Photoassisted tunneling spectroscopy: Preliminary results on tungsten diselenide," J. Vac. Sci. Technol. B**9**, (2), pp. 561-563, 1991.
4. R. T. Collins, K. v. Klitzing, and K. Ploog, "Photocurrent spectroscopy of GaAs/$Al_xGa_{1-x}As$ quantum well in an electric field," Phys. Rev. B, **33** (6), pp. 4378-4381, 1986.
5. K. Schmidt, A. D'Intino, N. Linder, M. Tuna, G. H. Döhler, H. T. Grahn, K. Kawashima, and K. Fujiwara, "Miniband and Wannier-Stark states in GaAs/AlAs superlattices investigated by differential photocurrent spectroscopy," Superlattices and Microstructures, **12** (2), pp.181-184, 1992.
6. L. Q. Qian and B. W. Wessels, "Scanning tunneling optical spectroscopy of semiconductors," Appl. Phys. Lett., **58** (12), pp. 1295-1296, 1991.
7. L. Q. Qian and B. W. Wessels, "Scanning tunneling optical spectroscopy of semiconductor quantum well structures," Appl. Phys. Lett., **58** (22), pp. 2538-2539, 1991.
8. T. Takahashi, M. Yoshita, and H. Sakaki, "Scanning tunneling microscopy of undoped GaAs/AlGaAs heterostructures under laser irradiation," Appl. Phys. Lett., **68** (4), pp. 502-504, 1996.
9. A. V. Bragas, S. M. Landi, J. A. Coy, and O. E. Martinez, "Spectroscopic response of photoinduced currents in a laser-assisted scanning tunneling microscope," J. Appl. Phys., **82** (9), pp. 4153-4158, 1997.
10. I. Lyubinetsky, Z. Dohnalek, V. A. Ukraintsev, and J. T. Yates, Jr., "Transient tunneling current in laser-assisted scanning tunneling microscopy," J. Appl. Phys. **82** (8), pp. 4115-4417, 1997.
11. T. Takahashi and M. Yoshita, "Laser irradiation effects on tunneling properties of n-type GaAs and InAs by scanning tunneling microscopy," Appl. Phys. Lett., **68** (24), pp. 3479-3481, 1996.
12. J. G. Simons., "Generalized formula for the electric tunnel effect between similar electrodes separated by a thin insulating film," J. Appl. Phys. **34** (6), pp. 1793-1803, 1963.
13. J. Tersoff and D. R. Hamann, "Laser irradiation effects on tunneling properties of n-type GaAs and InAs by scanning tunneling microscopy," Phys. Rev. B, **31** (2), pp. 805-813, 1985.
14. S. M. Sze, Physics of Semiconductor Devices (Wiley, New York, 1981) pp. 275-278.

Large aperture continuous phase diffractive optical element for beam transform

Tan Qiaofeng*, Yan Yingbai, Jin Guofan, Wu Minxian

State Key Laboratory of Precision Measurement Technology and Instruments
Tsinghua University, Beijing, China, 100084

ABSTRACT

Beam transform, such as to obtain uniform focal spot with flat top, steep edge, low side lobes and high light efficiency, can be realized well by diffractive optical element(DOE). The DOE has many advantages, such as high light efficiency and strong phase distribution design flexibility. To increase the light efficiency and decrease large-angle scattering, continuous phase DOE should be used. The phase design is completed by a kind of multi-resolution hybrid algorithm based on hill-climbing and simulated annealing, which exploits sufficiently strong convergence ability of the hill climbing and global optimization potential of the simulated annealing. A kind of phase distribution with good geometrical structure and diameter 80mm is obtained by choosing disturbance function, receipt and refused probability and so on. The simulated results show that the light efficiency is more than 95%, and the non-uniformity is less than 5%. Because the etching depth is direct proportion to the exposure time, to obtain continuous phase DOE, a kind of hollowed-out mask, namely gray-scale mask, is used to control exposure time of each area. The mask is manufactured by linear cutting machine. The continuous phase DOE with diameter 80mm is fabricated by ion-etching with the mask. Finally, the tolerance of manufacturing error including depth error and alignment error are analyzed.

Keywords: Beam transform, diffractive optical element .

1.INTRODUCTION

Uniform illumination with flat top, sharp edge and high light efficiency is needed in many fields, such as in the Inertial Confinement Fusion(ICF) system, laser manufacturing system. Such uniform illumination can be obtained by Diffractive Optical Elements(DOE)[1,2]. Phase distribution of the DOE can be designed by algorithms of optimization, such as Gerchberg-Saxton algorithm, Yang-Gu algorithm[3], Simulated Annealing[4], Genetic algorithm[5] and so on. To take the advantages and avoid the disadvantages of each algorithm, many kinds of hybrid algorithm are proposed for designing the phase distribution of the DOE. A kind of hybrid algorithm based on hill-climbing and simulated annealing is utilized in this paper, for performing the ability of strong convergence of the hill-climbing and the global optimization potential of the simulated annealing. And To decrease time-consuming and obtain good simulated results, a multi-resolution algorithm is adopted. Phase distribution with good geometrical structure and characteristic parameters is obtained by this hybrid algorithm.

To avoid large-angle scattering, the reduce of the light efficiency and the destroy of the flat top uniformity caused by phase quantization, continuous phase DOE should be adopted. The DOE will be manufactured on the K9 glass substrate with gray scale mask by ion-etching[6]. And the gray scale mask is fabricated by the linear cutting machine. The manufacturing error is analyzed.

*Correspondence: E-mail: tanqf@post.pim.tsinghua.edu.cn; Tel: 86 10 62781187; Fax: 86 10 62784503

Part of the SPIE Conference on Advanced Photonic Sensors and Applications
Singapore • November–December 1999
SPIE Vol. 3897 • 0277-786X/99/$10.00

2.DESIGN PRINCIPLE

The transimittance function of the DOE with circular symmetry can be written as:

$$T(\rho) = \exp(i\Phi(\rho)) \tag{1}$$

where $\Phi\ (\rho)$ is the phase distribution of the DOE. Suppose plane wave is put in, and according to Kirchhoff Theory, the far field distribution $E(r, f)$ of the DOE can be written as a Hankel transform of the transmittance function $T(\rho)$ of the DOE:

$$E(r,f) = C\int_0^{\rho} J_0(\frac{2\pi r\rho}{\lambda f})\exp(i\Phi(\rho))\rho d\rho = CB(\exp(i\Phi(\rho))) \tag{2}$$

where C is a constant, J_0 is the zeroth-order Bessel function, λ is the laser wavelength, f is the focal length of the Fourier Transform Lens and B takes place Hankel transform. Because the far field intensity distribution is concerned, the additional far field phase distribution caused by relative position between the DOE and the lens is ignored in Equ.(2).

Not considering the constant C, the far-field light intensity distribution can be rewritten as:

$$I(r,f) = \left| B(\exp(i\Phi(\rho)))\right|^2 \tag{3}$$

The cost-function of the optimization can be chosen at will. In this paper, the cost-function is chosen as:

$$\cos t(\Phi(\rho)) = \left|I(r,f) - I_{ideal}(r,f)\right|^2 \tag{4}$$

where $I_{ideal(x,y,f)}$ is the ideal intensity distribution of the uniform illumination.

Two parameters describing the quality of the beam transform, including the light efficiency and non-uniformity, are defined as:

Light efficiency:

$$\eta = \sum_{r_i \ge -d/2}^{r_i \le d/2} I(r_i,f)\cdot(2i-1) \Big/ \sum_{\forall r_r} I(r_i,f)\cdot(2i-1) \tag{5}$$

Non-uniformity:

$$rms = \sqrt{\frac{1}{N-1}\sum_{i=1}^{N}\left[\frac{I(r_i,f)-\bar{I}}{\bar{I}}\right]^2} \tag{6}$$

where $\bar{I} = \frac{1}{N}\sum_{i=1}^{N} I(r_i,f)$, $N = [d/\delta d]+1$.

3.OPTIMIZATION ALGORITHM

The basic structure of the hybrid algorithm based on the hill-climbing and simulated annealing is described in the Fig. 1(a). This hybrid algorithm is exploit sufficiently the ability of strong convergence of the hill-climbing and the global optimization potential of the simulated annealing.

The phase distribution can be written as:

$$\Phi(\rho) = \sum_{i=1}^{N} A_i \sin(\varpi_i \rho + \phi_{0i}) \tag{7}$$

where A_i, ϖ_i and ϕ_{0i} are amplitudes, periods and initial phases of the sine functions.

Let $\Phi(\rho)$ be zero, and choose a sine function with random period and initial phase as choosing direction, and optimize the amplitude with hill-climbing algorithm along this direction. Whether the amplitude is accepted or not is determined by simulated annealing operator.

The continuity of the designed phase distribution is assured by the sine functions.

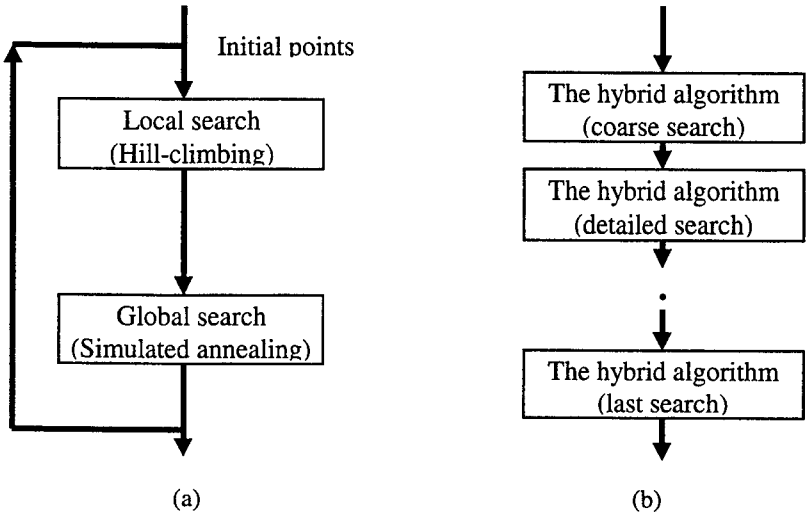


Fig.1 Schematics of the hybrid algorithm and the algorithm with multi-resolution

The hybrid algorithm is very valid to design the phase distribution of the DOE. But this optimization algorithm is a little time-consuming, and time-consuming increases sharply with the improvement of the resolution, namely, the increase of the sample number. The DOEs applied in ICF system and laser manufacturing system have large aperture, and the sample number of the space domain must not be chosen as very small. Large sample number will obtain better simulated results but need more time than small sample number. To decrease time-consuming and obtain good simulated results, a multi-resolution algorithm is adopted.

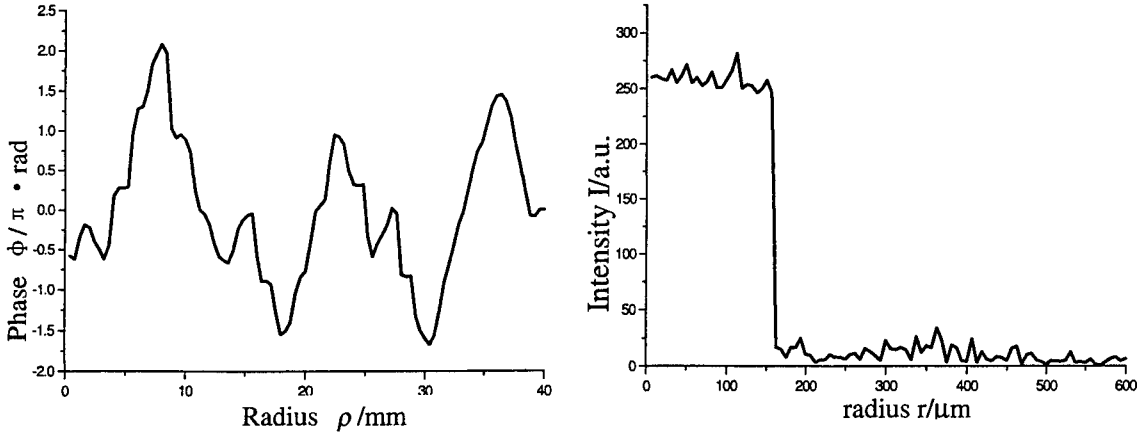
For reducing the calculation time, based on the principle of fractal, a multi-resolution design algorithm is developed as shown in Fig.1(b). Through "coarse optimization" and "detailed optimization", a distribution of phase with fine structure is obtained step by step by increasing the sampling number. With less calculation time, a DOE with high resolution, smoothed fluctuation, less period, low phase interval and an easier to be fabricated phase geometrical structure can be designed out.

4.SIMULATED RESULTS

The design parameters f, λ, the diameter D of the DOE and the diameter d of the uniform illuminated area are $400\,mm$, $1.053\,\mu m$, $80\,mm$ and $320\,\mu m$, respectively.

Using the hybrid algorithm with multi-resolution as shown in Fig. 1(b), when the sampling is 100, the phase distribution and intensity distribution are shown in Fig.2 (a) and (b), respectively. The light efficiency and the non-uniformity are 97.2% and 6.39%, respectively. At this point, the convergence rate is very slow and the calculation with the sampling 100 is stopped. Then the calculation with the sampling 200 begins. The phase distribution after optimization is shown in Fig.3. The light efficiency and the non-uniformity are 97.4% and 5.61%, respectively. The non-uniformity is improved. The phase distribution in Fig.3 is similar to that in Fig.2(a), but different in some detail. When the sampling increases to 400, and the

phase distribution after optimization is shown in Fig.4, the light efficiency and non-uniformity are 98.0% and 4.00%.



(a) Phase distribution of the DOE (b) Intensity distribution of the focal plane

Fig. 2 Simulated results with the sample number 100

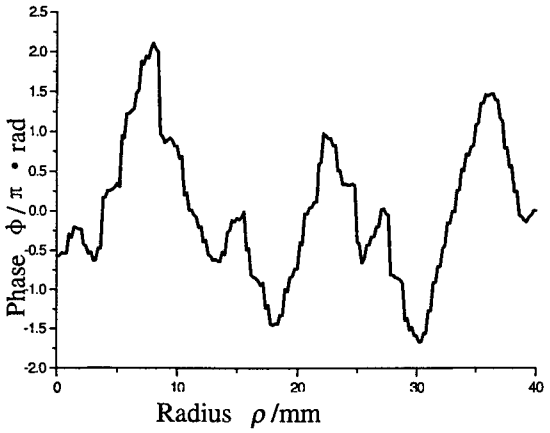


Fig. 3 Phase distribution of the DOE after optimization with the sample number 200

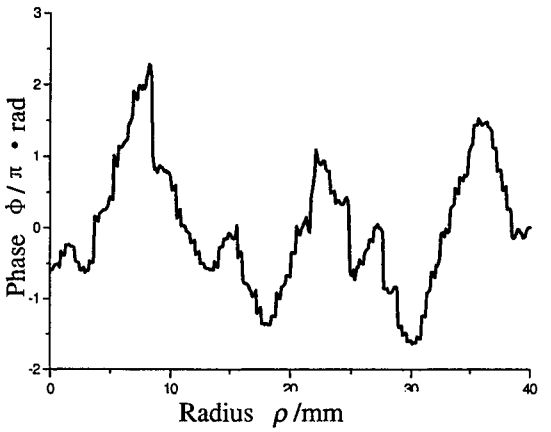


Fig. 4 Phase distribution of the DOE after optimization with the sample number 400

5.TOLERANCE ANALYSIS OF THE INPUT PHASE DISTORTION

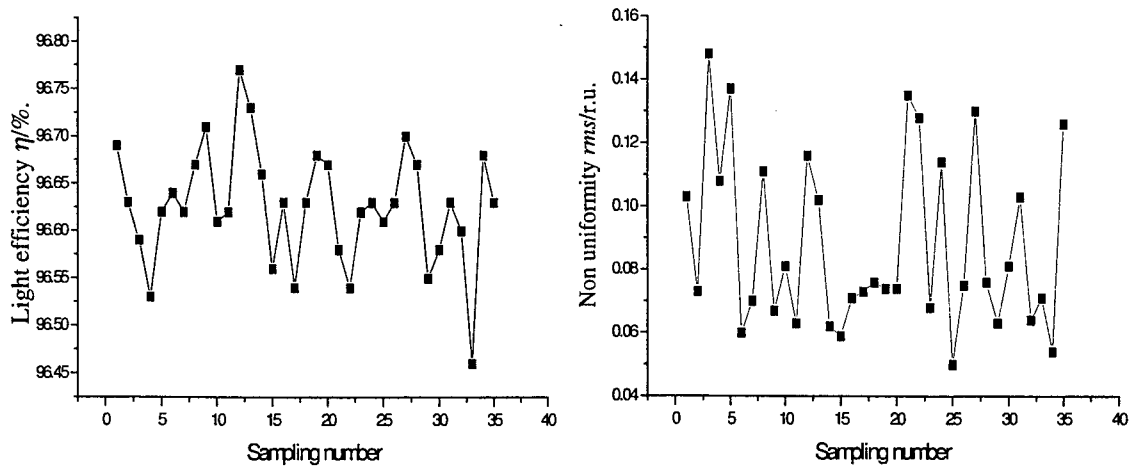
In fact, the input laser beam is not ideal plane wave. The influence to the characteristic parameters of the DOE introduced by phase distortion is much larger than that introduced by amplitude distortion[7]. Here, phase distortion of the input laser beam is only analyzed. Random noise with Gauss distribution is adopted to simulate the phase distortion. The influences of this models to the light efficiency and non-uniformity are calculated.

The random noise with Gauss distribution can be expressed as:

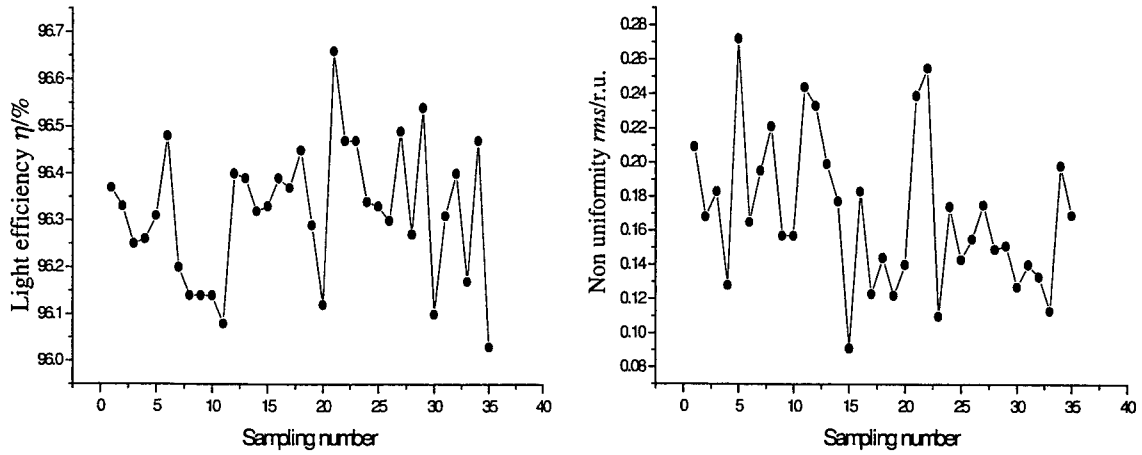
$$\Delta\phi_i = N[0, a \cdot 2\pi] \tag{8}$$

where $N[0, a \cdot 2\pi]$ takes place the Gauss distribution with zero mean and $a \cdot 2\pi$ variance, and a also reflects the magnitude of the phase distortion。

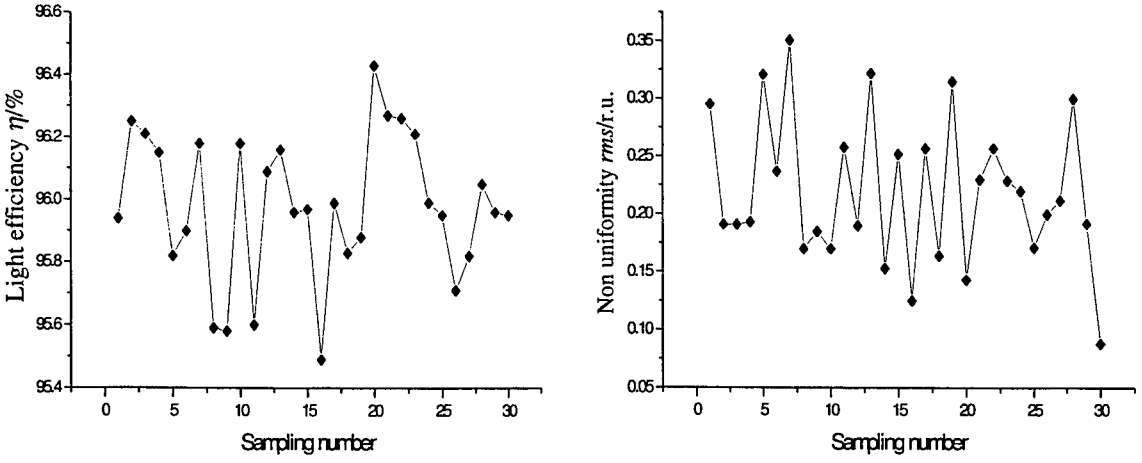
The influences introduced by the phase distortion with Gauss distribution are shown in Fig. 5(a), (b) and (c). a are 0.05, 0.10 and 0.13, respectively.



(a) a is 0.05



(b) a is 0.10



(c) a is 0.13

Fig. 5 Influence on the characteristic of phase distortion with Gauss distribution

From Fig.5, the light efficiency and non-uniformity are both destroyed with the increase of the phase distortion, and the non-uniformity extremely jumps and have no regularity. The DOE tolerance of the phase distortion is very small and can not be

increased only by design method. The DOE array will be used to increase the tolerance of the input phase distortion and improve the uniformity.

6.FABRICATION AND ERROR ANALYSIS

The gray scale mask fabricated by the linear cutting machine is shown in Fig. 6.

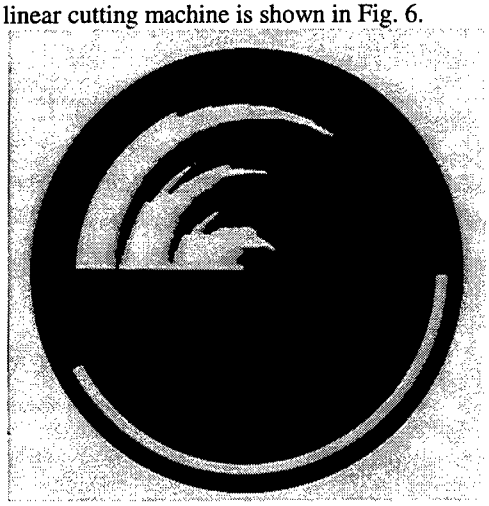


Fig.6 Gray-scale mask with $\Phi 100mm$ manufactured by linear cutting machine

The gray-scale mask is measured by a universal tool maker's microscope to determine the fabrication precision of the linear cutting machine. The excircle, one internal circle and the radius are measured, which are shown in Fig. 7.

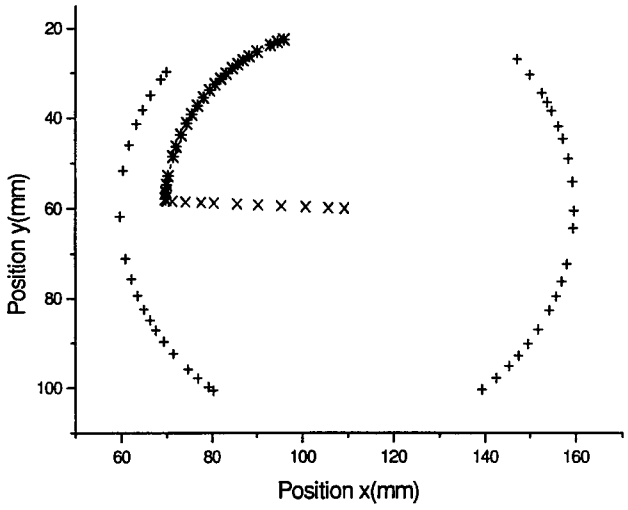


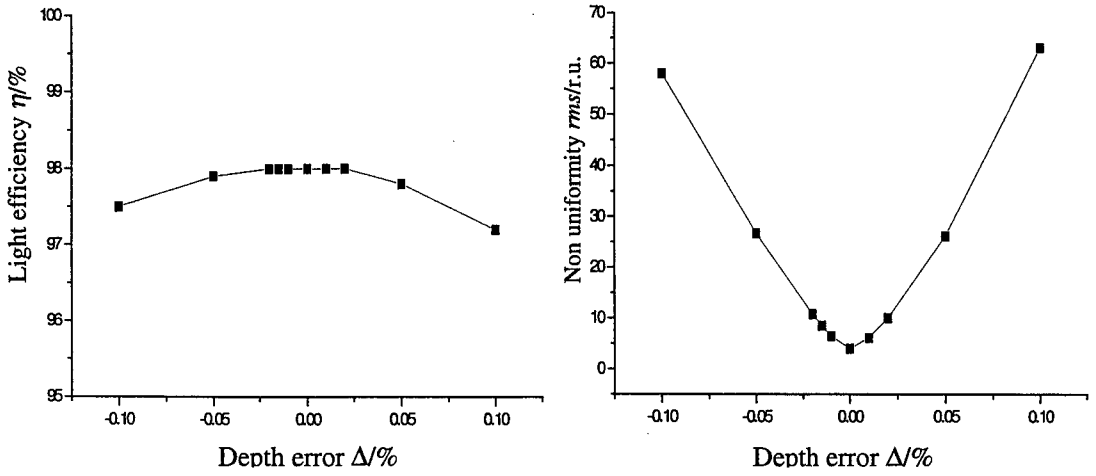
Fig.7 Data of the gray-scale mask measured by universal tool maker's microscope

After data fitting, the data of the excircle shows that the center coordinates of the mask are 109.67083mm, 60.10746mm, the outside diameter runout is about 0.0412mm. Because the fabrication of the DOE is excircle centring, the measured internal circle has variation 0.0270mm to the fitted center, the maximum error of the radius is 0.0187mm.

7.INFLUENCE OF THE MANUFACTURING ERROR

Two kinds of manufacturing error sources of continuous phase DOE are considered here, including the depth error and the

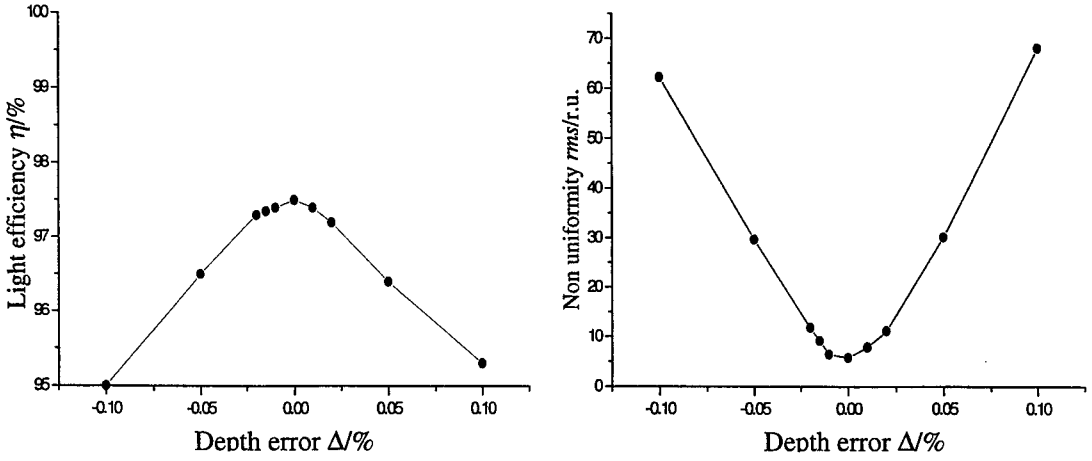
alignment error. The influence of the depth error can be calculated and the simulated results are shown in Fig. 8. The tolerance of manufacturing error is limited from -2% to +2% make the non-uniformity be less than 10%.



(a) The light efficiency (b)Non-uniformity

Fig. 8 Influence on the characteristic of the depth error

The alignment error between the mask and the substrate is determined by (r_0 , θ_0). r_0 is chosen as 0.05mm. Here, only one case is considered. When $\theta_0 = 0$, the influence of the depth error is calculated and the simulated results are shown in Fig. 9. The tolerances of manufacturing error is also limited from -2% to +2%. The 0.05mm alignment error has hardly effect on the characteristic of the DOE because the relative error is small.



(a) The light efficiency (b)Non-uniformity

Fig. 9 Influence on the characteristic of the depth error (r_0 =0.05mm, $\theta_0 = 0$)

CONCLUSION

The continuous phase DOE with good geometrical structure and good characteristic parameters is designed by the multi-resolution hybrid algorithm based on the hill-climbing and simulated annealing. The gray-scale mask for fabricating continuous phase DOE is manufactured by the linear cutting machine, and the manufacturing error is measurement by the universal tool maker's microscope. The alignment error has almost not influence on the focal spot, but the ion etching depth error is limited from -2% to +2% to assure the non-uniformity is less than 10%. The distortion tolerance of the input

laser distortion with Gauss distribution is not enough. DOE will be used to increase the tolerance of the input laser.

ACKNOWLEDGMENTS

The work is supported by the Fund of High Technology of China and by National Natural Science Foundation of China.

REFERENCES

1. S.N.Dixit, J.K.Lawson, K.R.Manes and H.T.Powell, "Kinoform phase plates for focal plane irradiance profile control", *Opt. Lett.* **19**, pp.417-419, 1994
2. Y.Lin, T.J.Kessler, G.N.Lawrence, "Distributed phase plates for super-Gaussian focal-plane irradiance profile", *Opt. Lett.* **20**, pp.764-766, 1995
3. Guozhen Yang, Bizhen Dong, Benyuan Gu, *et al*, "Gercherg-Saxton and Yang-Gu algorithms for phase retrieval in a nonunitary transform system: a comparison", *Appl. Opt.* **33**, pp.209-218, 1994
4. S.Kirkpatrick, C.D.Galatt, Jr, M.P.Vecchi, "Optimization by simulated annealing", *Science*, **220**, pp.671-680, 1983
5. Liviu I. Voicu, Wissam A. Rabadi, Harley R. Myler, "Object support reconstruction from the support of autocorrelation using multiresolution genetic algorithms", *Opt. Eng.* **36**, pp.2820-2827,1997
6. Y. Wang, *et al*, "A new approach for the fabrication of diffractive optical elements with rotationally symmetrical phase distribution", *SPIE*, **3348**, pp.94-97,1998
7. Jin Guofan, Yan Yingbai, Zhai Jinhui, *et al*, "Design and fabrication of binary optical element for ICF target uniform illumination", *1995~1997 ICF Annual Reports*, pp.61-70

Micro-optics element for color separating

Wang Jinyu*, Yan Yingbai, Tan Qiaofeng, Jin Guofan

State Key Laboratory of Precision Measurement Technology and Instruments
Tsinghua University, Beijing, China, 100084

ABSTRACT

A micro-optics element has been designed for diverting unwanted harmonic waves out of high power laser system to perform Inertial Confinement Fusion (ICF). Based on scalar diffraction theory, a basic structure of color separation gratings (CSG) has been designed, which almost meets the request of ICF target system. To increase the diffraction efficiency of third harmonic wave (referred as 3ω light) at 0.351um and decrease that of fundamental wave (called 1ω) at 1.053um and second harmonic (called 2ω) at 0.527um, the basic CSG structure has been optimized by a hybrid algorithm based on simulated annealing algorithm and hill-climbing algorithm. During the optimization procedure, the freedom variable is the step width and step height in one period and the degree of freedom is totally four. By applying several different bounds for objective function, several optimized micro-optical structures have been got, which all meet requirements of ICF target system. The influences of step width errors, step height errors and mould collimating errors of the optimized CSG have been analyzed respectively. Among these schemes, a most ideal structure has been selected which is most insensitive to fabrication errors. Under given error limitation, the diffraction efficiency's root-mean-square value at zero order for selected CSG have been investigated by statistical method.

Keywords: color separating grating, micro-optics element, ICF

1.INTRODUCTION

In Inertial Confinement Fusion (ICF) systems around the world, most of the glass laser employ non-linear frequency conversion for converting fundamental wave (1ω light) to its second harmonic 2ω and third harmonic 3ω light. Shorter wavelengths are better for driving ICF target because of their coupling efficiency can be improved [6]. However, there are still remnant energy of 1ω and 2ω frequencies remains in optical path after frequency conversion. When the unconverted lights enter fusion target, It is harmful to ICF experiments. They must be filtered away.

To divert the unconverted light away from fusion target, several strategies have been developed. On the Xingguang laser system at Southwest institute of Nuclear physics and chemistry and on the Omega laser system at the University of Rochester[1], the remnant 1ω and 2ω light is filtered out by turning mirrors with multilayer films. There are many shortcomings using the differing reflectivity of turning mirror divert remnant light: (1). Coating film's threshold of resisting light injury to ultraviolet such as 3ω light is too low; (2). There is primary theory difficulties about using multilayer film

Correspondence: Email: wangjy@post.pim.tsinghua.edu.cn ; Telephone: 86-10-62781187; Fax: 86-10-62784503s

Part of the SPIE Conference on Advanced Photonic Sensors and Applications
Singapore • November–December 1999
SPIE Vol. 3897 • 0277-786X/99/$10.00

divert unwanted light in ICF system, odd harmonic and even harmonic are incompatible. It is impossible to separate 1ω, 2ω and 3ω light by one turning mirror; (3). In order to divert the unconverted light, several turning mirrors must be used. The optical path is somewhat complex and the loss of 3ω light is serious. On the Nova laser system at Lawrence Livermore National Laboratory and on the Phebus laser system at Limeil [1], by the chromatic aberration of the focusing lens, a shadow of a diameter circular obscuration is cast to the center of the near-field profile to prevent the 1ω and 2ω light from hitting the target located at the 3ω focus. This scheme's disadvantage is: (1). Although 1ω and 2ω light are prevent to enter the holhraum of ICF target system, they have entered the target cavity. The impurity of target cavity may seriously influent target physical experiments; (2). The side lobe caused by obscure disc may decrease the coupling efficiency of 3ω light. In the baseline design for the proposed National Ignition Facility (NIF) [1], a wedged focusing lens has been employed for diverting the unconverted light. However, this will lead to higher nonlinear B-integral and therefore increase the possible degradation of beam quality of 3ω light.

Micro-optical element has advantages of high diffraction efficiency, high threshold of resisting light injury [2], realizing multi-function integration and so on. It is promising to develop the application of micro-optical element in high power laser system to perform ICF [3,4]. Currently, there are several schemes [1,5–10] has been brought forward about using binary optical element (BOE) to divert the unconverted light away from ICF target system. However, there are still some disadvantages such as low diffraction efficiency, impurity and strait fabrication error redundancy.

In this paper, a new color separating grating (CSG) structure has been designed. In section 2, possible grating structure with regular diffraction cell to separate the three harmonic waves has been discussed. Under scalar wave diffraction theory, the basic CSG structure has been designed. In section 3, Based on the basic CSG structure, the influences of step height and step width to color separating effects has been analyzed and seven kinds of CSG structure has been obtained by optimization under different bounds for objective function. In section 4, the influences of the step width errors, step height errors and collimating errors of all calculated CSG structure have been calculated. Among them, one scheme has been found to be the most insensitive to above fabrication errors.

2. THEORY AND DESIGN OF BASIC STRUCTURE

According to the requirement of ICF system, 3ω should be diffracted at 0 order, 1ω and 2ω light should be diffracted at +1 and −1 order respectively. In this paper, we only consider to separate colors by regular grating. To fulfill this aim, the relief structure in one period of regular grating is showed in fig.1, Where M (M=0,1,2, ...) denotes step number in one period; d_j (j=1,2,...,M) denotes width of step j; $d = \sum_{j=1}^{M} d_j$ denotes the period width; h_j (j=1,2,...,M) denotes height of step j.

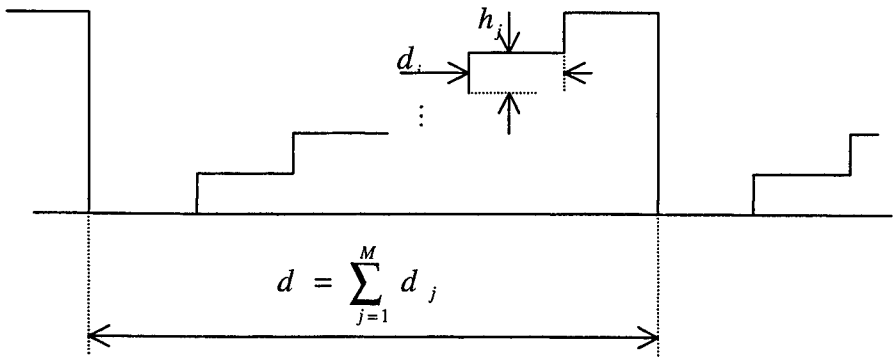


Fig 1. Possible CSG relief structure in one period

Supposed that the phase introduced for the wave which wavelength is λ by step j is $\phi(\lambda, j)$, we will have:

$$\phi(\lambda, j) = \frac{n(\lambda)-1}{\lambda} h_j, \tag{1}$$

where $n(\lambda)$ is the refractive index for the wave. Thus, the transmittance distribution of the grating is:

$$t(x) = \sum_{m=-\infty}^{\infty} \delta[x - m(d_1 + d_2 + ... + d_j + ... + d_M)] * \{rect(\frac{x - d_1/2}{d_1}) \cdot e^{i\phi(\lambda,1)} + rect(\frac{x-(d_1+d_2/2)}{d_2}) \cdot e^{i\phi(\lambda,2)}$$

$$+ ... + rect(\frac{x-(d_1+...+d_{j-1}+d_j/2)}{d_j}) \cdot e^{i\phi(\lambda,1)} + ... + rect(\frac{x-(d_1+...+d_{M-1}+d_M/2)}{d_M}) \cdot e^{i\phi(\lambda,M)}\}. \tag{2}$$

Ignoring constant phase factor, the complex amplitude distribution on the far field diffraction plane is:

$$\alpha(f) = \frac{1}{d} \sum_{m=-\infty}^{\infty} \delta(f - \frac{m}{d}) \cdot \{d_1 \operatorname{sinc}(fd_1) \cdot e^{i(\phi(\lambda,1) - 2\pi f \frac{d_1}{2})} + d_2 \operatorname{sinc}(fd_2) \cdot e^{i(\phi(\lambda,2) - 2\pi f (d_1 + \frac{d_2}{2}))} + ...$$

$$+ d_j \operatorname{sinc}(fd_j) \cdot e^{i(\phi(\lambda,j) - 2\pi f (d_1 + ... + d_{j-1} + \frac{d_{j1}}{2}))} + ... + d_M \operatorname{sinc}(fd_M) \cdot e^{i(\phi(\lambda,M) - 2\pi f (d_1 + ... + d_{M-1} + \frac{d_M}{2}))}\}, \tag{3}$$

where $f = \frac{\sin\theta}{\lambda}$, and θ is the diffraction angle. For diffraction order m, we must have $f = \frac{m}{d}$. Thus, the complex amplitude distribution at order m is:

$$\alpha_m = \frac{d_1}{d} \sin c(\frac{md_1}{d}) \cdot e^{i(\phi_1 - 2\pi \frac{md_1}{sd})} + \frac{d_2}{d} \sin c(\frac{md_2}{d}) \cdot e^{i(\phi_2 - 2\pi \frac{m(d_1 + d_2/2)}{d})} + ...$$

$$+ \frac{d_j}{d} \sin c(\frac{md_j}{d}) \cdot e^{i(\phi_1 - 2\pi \frac{m(d_1+...+d_{j-1}+d_j/2)}{sd})} + ... + \frac{d_M}{d} \sin c(\frac{md_M}{d}) \cdot e^{i(\phi_M - 2\pi \frac{m(d_1+...+d_M/2)}{d})}. \tag{4}$$

At zero diffraction order, m=0, the corresponding complex amplitude is:

$$\alpha_0 = \frac{d_1}{d}\cdot e^{i\phi_1} + \frac{d_2}{d}\cdot e^{i\phi_2} + \ldots + \frac{d_j}{d}\cdot e^{i\phi_1} + \ldots + \frac{d_M}{d}\cdot e^{i\phi_M} \tag{5}$$

and corresponding intensity is:

$$I_0 = \alpha_0 \cdot \alpha_0^* \tag{6}$$

In order to calculate the basic structure of CSG, we assume the step width in one period of CSG is equal to each other ($d_1 = d_2 = \ldots = d_M$), and there is no chromatic aberration. According to the requirements, CSG should be transparent to 3ω light at zero diffraction order, and opaque to 1ω and 2ω light. The phases introduced by step j (j=1,2,...,M) in one period should be:

$$\begin{aligned}
\phi(1\omega, j) &= 2k_1\pi + \frac{j}{M}\cdot 2\pi \\
\phi(2\omega, j) &= 2k_2\pi + \frac{j}{M}\cdot 2\pi \\
\phi(3\omega, j) &= 2k_3\pi + j\cdot 2\pi
\end{aligned} \tag{7}$$

where k_1, k_2, k_3 is integer. Ignoring chromatic aberration, the phase introduced by same step for three harmonics has following relationship:

$$\frac{1}{3}\phi(1\omega, j) = \frac{2}{3}\phi(2\omega, j) = \phi(3\omega, j) \tag{8}$$

From equation (7) and (8), we can derive that when k=1,2,3 respectively, the steps number in one period must have:

$$\begin{aligned}
MOD(\frac{2}{3}k) &= \frac{1}{M}, \frac{2}{M}, \ldots, \frac{M}{M} \\
MOD(\frac{1}{3}k) &= \frac{1}{M}, \frac{2}{M}, \ldots, \frac{M}{M}
\end{aligned} \tag{9}$$

where MOD(x) means signed remainder after divided by 1. According to equation (9), a regular grating which steps number M=3 can meet the requirements. Actually, there are different phase structures when M=3 can be used to separate harmonic waves. However, because of chromatic aberration, the phase introduced by step will changed slightly. From equation (1),

$$\frac{\partial\phi(\lambda, j)}{\partial n} = \frac{h_j}{\lambda}. \tag{10}$$

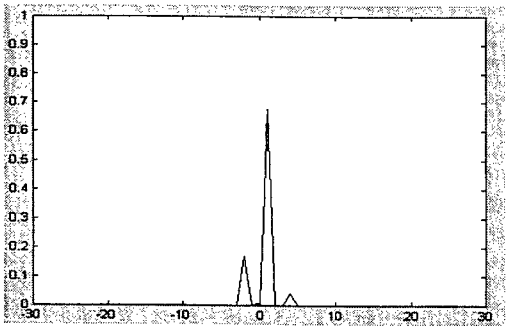
Phase changes generated by chromatic aberration can be reduced by decreasing step height. Therefore, a basic CSG structure should be applied which was the thinnest one. The refractive index to three harmonic waves of fused silica substrate on which CSG was fabricated is 1.44982,1.46094,1.47673 respectively, the refractive index to 3ω light n=1.47673 was used as designing index. Calculated result is show in table 1. With actual refractive index, the computer simulated intensity distribution in far field of the three harmonic wave is shown in table 2 and fig.2. Obviously, the intensity distribution is slightly deviated from designing result because of chromatic aberration.

Table 1. The basic CSG structure in one period

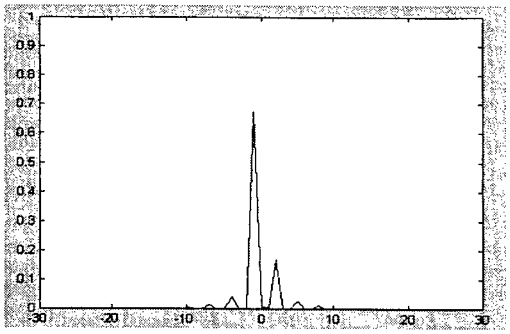
Step serial No.	Step height	1 ω phase	2 ω phase	3 ω phase
1	0. 00000	0	0	0
2	0. 73627	$\frac{2}{3}\cdot 2\pi$	$\frac{1}{3}\cdot 2\pi$	2 π
3	1. 47253	$\frac{4}{3}\cdot 2\pi$	$\frac{2}{3}\cdot 2\pi$	4 π

Table 2. CSG diffraction efficiencies in various order

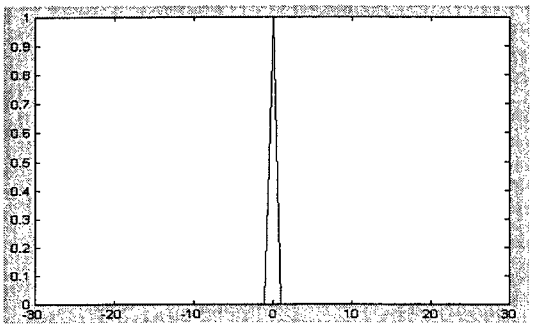
Diff. order	1 ω efficiency	2 ω efficiency	3 ω efficiency
-5	0.02710	0.00019	0.00000
-4	0.00018	0.04220	0.00000
-3	0.00000	0.00000	0.00000
-2	0.16939	0.00118	0.00000
-1	0.00296	0.67518	0.00000
0	0.00495	0.00587	1.00000
+1	0.67757	0.00472	0.00000
+2	0.00074	0.16880	0.00000
+3	0.00000	0.00000	0.00000
+4	0.04235	0.00029	0.00000
+5	0.00012	0.02701	0.00000



(a)



(b)



(c)

Fig.2 Intensity distribution in far field of three harmonic waves generated by basic CSG structure

(a). 1 ω light (b).2 ω light (c). 3 ω light

3. OPTIMIZATION OF CSG

On the assumption of the absence of chromatic aberration, we have obtained a basic CSG structure in section 2. Unfortunately, chromatic aberration really exist and the simulated calculating result therefor slightly deviate from ideal aim. In this section, we will analyze the influences to diffraction efficiency of step width and step height and then optimize the basic CSG structure.

3.1. Influences of step width

According to the basic CSG structure, every step width is same, on the absence of chromatic aberration, the sketch map of amplitude synthesizing at zero diffraction order in one diffraction period is shown in fig.3. Where a_{ij} $\mathrm{i,j}=1,2,3$ refers to complex amplitude formed by step j on fundamental wave, second harmonic wave and third harmonic wave.

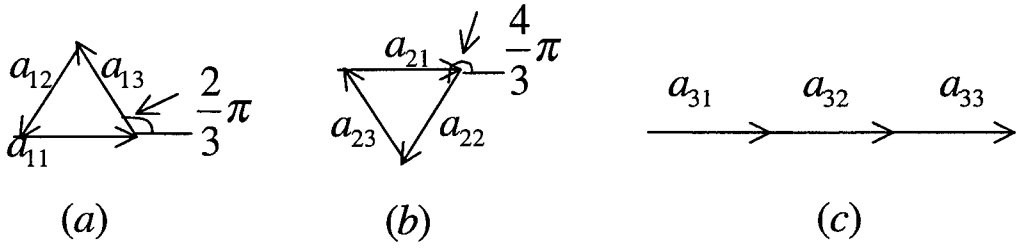


Fig.3. Sketch map of complex amplitude synthesizing at zero diffraction order on the absence of chromatic aberration

(a).1 ω light (b).2 ω light (c).3 ω light

Because of chromatic aberration, the phase introduced by steps to different color has changed. The phase shift of 1 ω and 2 ω light will decrease compared to original value, thus their complex amplitude curve will not close actually and therefor there are remnant energy of them. Considering chromatic aberration, the sketch map of amplitude synthesizing at zero

diffraction order in one diffraction period is shown in fig.4. Where a_{ij} $i,j=1,2,3$ refers to complex amplitude formed by step j on fundamental wave, second harmonic wave and third harmonic wave. Rm_k $k=1,2$ denotes the remnant amplitude of 1ω and 2ω light at zero order.

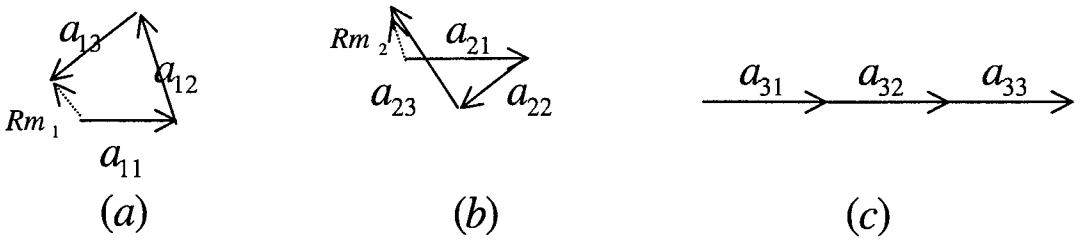


Fig.4. Sketch map of complex amplitude synthesizing at zero diffraction order taking into account of chromatic aberration
(a).1ω light (b).2ω light (c).3ω light

As we known, the absolute value $|a_{ij}|$ of corresponding complex amplitude will vary with step width. However, the phase shift remain unchanged. Thus, It is possible to decrease remnant light $|Rm_k|$ $k=1,2$ and maintain 3ω light transparent by optimize step width. By step width optimization method, a new structure has been obtained. During optimization process, since efficiency of 3ω light maintain 100% , the objective function is sum efficiency of 1ω and 2ω light at zero order. Its normalized width parameter is shown in table 3 and diffraction efficiency in various orders in table 4. Compared to table 2, the sum efficiency of 1ω and 2ω has been slightly decreased.

Table 3. CSG structure by step width optimization

Step serial No.	1	2	3
Normalized width	1.00000	1.02554	1.0000

Table 4. width optimized CSG diffraction efficiencies in various order

Diff. order	1ω efficiency	2ω efficiency	3ω efficiency
-5	0.02715	0.00003	0.00000
-4	0.00003	0.04234	0.00000
-3	0.00006	0.00008	0.00000
-2	0.16971	0.00067	0.00000
-1	0.00212	0.67619	0.00000
0	0.00613	0.00456	1.00000
+1	0.67652	0.00472	0.00000

+2	0.00127	0.00593	0.00000
+3	0.00006	0.16807	0.00000
+4	0.04198	0.00008	0.00000
+5	0.00038	0.00064	0.00000

3.2. Influences of step height

As shown in fig.4, if the phase angle of a_{ij} $i,j=1,2,3$ was changed, it may lead to decrease of 3ω light efficiency and changes of remnant light $|Rm_k|$ $k=1,2$. In fact, the phase angle of a_{ij} $i,j=1,2,3$ is varying with step height. It is possible to find a more desirable step height of CSG structure adapt for request of ICF system. By step height optimization method, a new structure has been obtained. During optimization process, the objective function is sum efficiency of 1ω and 2ω light at zero order. In order to compare with step width optimization results, bound of objective function was not set. Its height parameter is shown in table 5 and diffraction efficiency in various orders in table 6. Compared to table 2, the sum efficiency of 1ω and 2ω has been greately decreased and at the same time efficiency of 3ω has decreased to 96%.

Table 5. CSG structure by step height optimization

Step serial No.	1	2	3
Step height	0.00000	0.76515	1.53031

Table 6. Height optimized CSG diffraction efficiencies in various order

Diff. order	1ω efficiency	2ω efficiency	3ω efficiency
-5	0.02733	0.00000	0.00062
-4	0.00002	0.04273	0.00073
-3	0.00000	0.00000	0.00000
-2	0.17079	0.00002	0.00390
-1	0.00037	0.68373	0.01170
0	0.00056	0.00014	0.96009
+1	0.68316	0.00009	0.01559
+2	0.00009	0.17093	0.00293
+3	0.00000	0.00000	0.00000
+4	0.04270	0.00000	0.00097
+5	0.00001	0.02735	0.00047

3.3 Optimization with step width and height

As analyzed in above mentioned, step width can change absolute value of diffraction complex and step height can change corresponding phase shift. By adjusting step width and step height felicitously, a new CSG structure may be obtained, through which the efficiency of 3ω light is maintained as high as possible and that of 1ω and 2ω light is restrained as low as possible at zero diffraction order. In practice, it is often the situation that the simulated computing results are very ideal, however, the experimental results of the elements can not fulfill designing request because of fabrication error. Therefor, we hope to find a desirable structure satisfying requirements and at the same time it is insensible to fabrication error. By adopting different boundary condition, several optimized CSG structures have been got. According to request of ICF system, the efficiency of 3ω light should be greater than 95%., bound of objective function was set for it to be greater than 96%, 97%, 98% and 99% respectively. Its step width and height parameter is shown in table 7 and its diffraction efficiency in zero orders in table 8.

Table 7. CSG structure optimized by step width combined with height

Bounds (3ω efficiency)		Step 1	Step 2	Step 3
$\geq 96\%$	Step width	1.00000	0.94810	1.00004
	Step height	0.00000	0.76497	1.52987
$\geq 97\%$	Step width	1.00000	0.96056	1.00370
	Step height	0.00000	0.76133	1.52218
$\geq 98\%$	Step width	1.00000	0.97106	1.00007
	Step height	0.00000	0.75653	1.51308
$\geq 99\%$	Step width	1.00000	0.98727	1.00007
	Step height	0.00000	0.75058	1.50123

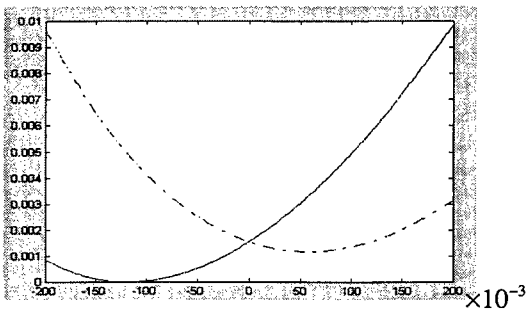
Table 8. Width and height optimized CSG diffraction efficiencies in zero order

Bounds (3ω efficiency)	1ω	2ω	3ω
$\geq 96\%$	0.00005	0.00004	0.96000
$\geq 97\%$	0.00028	0.00023	0.97000
$\geq 98\%$	0.00079	0.00065	0.98000
$\geq 99\%$	0.00185	0.00147	0.99000

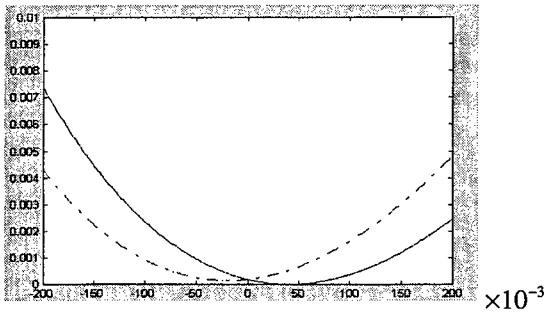
4. SIMULATED FABRICATION ERROR ANALYSIS

Actually, for a micro-optical element, the main fabrication errors comes from step width, step height and collimating errors of mould. We have investigated these kinds of error respectively. In the light of the request of ICF system: 3ω efficiency should be greater than 98%, and that of 1ω and 2ω should be less 2%. The error limit for the four CSG of step width, step height and mould collimating errors have be calculated. By comparison, the optimized CSG width bound of 3ω efficiency

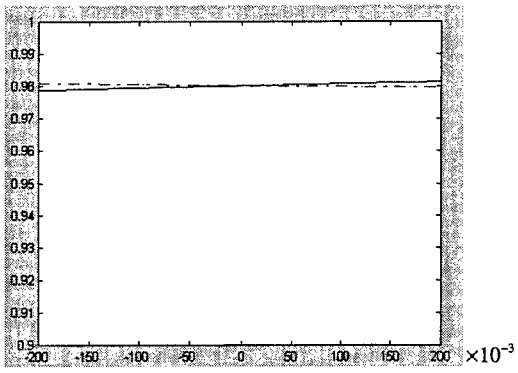
greater than 98% is found to be the best one with highest fabrication error redundancy. The curves of errors are shown in fig.5, fig.6 and fig7 respectively.



(a)



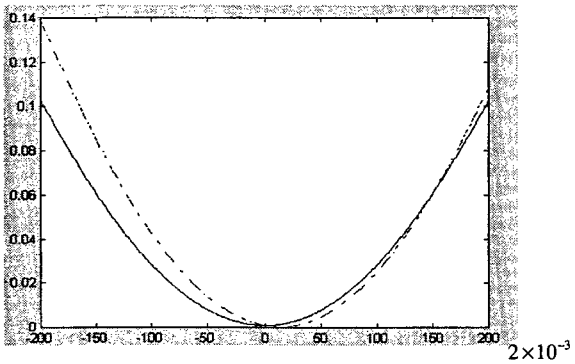
(b)



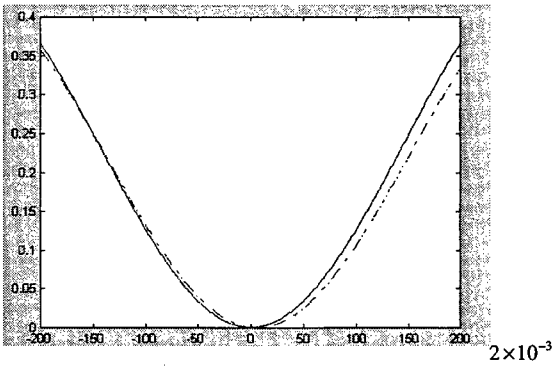
(c)

Fig. 5 Step width error curve for different harmonic waves.

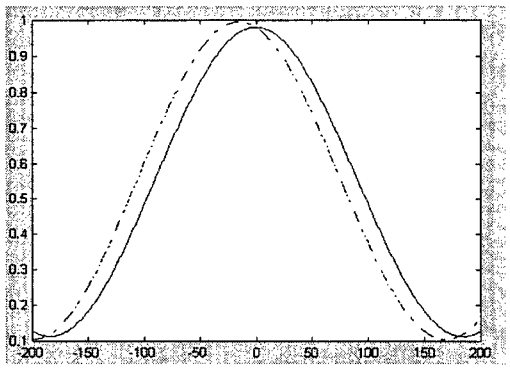
Solid curve refer to width of step 2; dash dotted curve refer to width of step 3. (a). 1 ω light (b).2 ω light (c).3 ω light



(a)



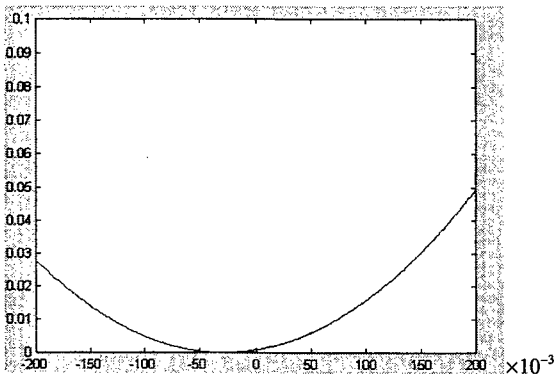
(b)



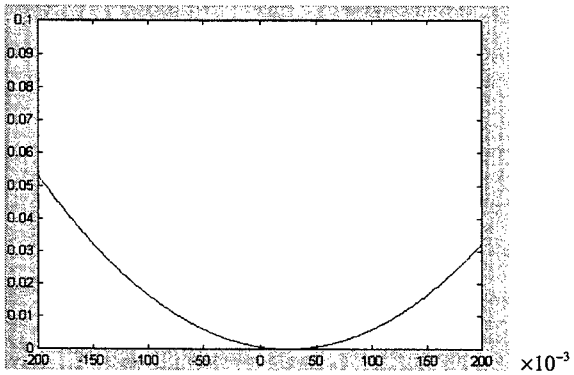
(c)

Fig. 6 Step height error curve for different harmonic waves.

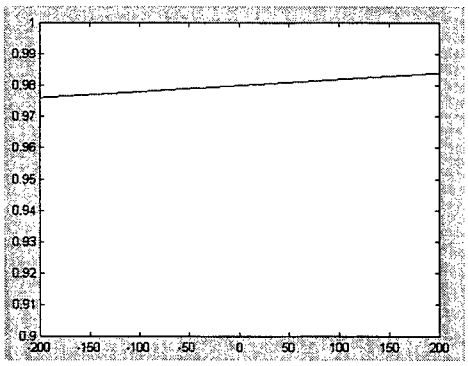
Solid curve refer to width of step 2; dash dotted curve refer to width of step 3. (a). 1 ω light (b).2 ω light (c).3 ω light



(a)



(b)



(c)

Fig. 7 Mould collimating error curve for different harmonic waves

(a). 1 ω light (b).2 ω light (c).3 ω light

In the light of the separating angle required in ICF system. According to currently fabrication precision, relative fabrication error limitation is set as 3% respective. Under the error limitation, we have calculated that the RMS(root-mean-square value) of 1ω, 2ω and 3ω light are 0.71%,1.12% and 98.69% respectively.

5. CONCLUSION

According to the request of ICF system, we have analyzed the primary theory of color separating grating. A basic CSG is derived and proved to be the most promising regular grating meets the requirement of ICF system. Based on the basic CSG model, several optimization structures have been calculated. Through fabrication error analyzing, a CSG with the highest redundancy to errors has been selected. By statistical analyzing method, the diffraction efficiency of the selected CSG to three harmonics has been given.

ACKNOWLEDGEMENT

This work is supported by the fund of High-Technology of China and National Natural Foundation of China.

References

1. S.N.Dixit, M.C.Rushford, I.M.Thomas and S.M.Herman etc., "Color separation gratings for diverting the unconverted light away from the NIF target", SPIE, **3047**, pp.463-470, 1997

2.J.R.Leger, M.R.Moharam and T.K.Gaylord, "Diffractive optics: an introduction to the feature issue", Appl.Opt, **34**, pp.2399,1995

3.Y.Kato,K.Mima,N.Miganaga etc.,Phys.Rev.Lett,**53**, pp.1057,1984.

4.R.M.Stenveson, M.J.Noemen and T.H.Bett, "Binary phase zone plate arrays for the generation of uniform focal profiles", Opt.Lett., **19**,No.6, pp.363-365,1994.

5.H.Damman, "Color separation gratings", Appl. Opt. **17**, pp.2273-2279, 1978

6.M. B. Stern and G. J. Swanson, "Color separation echelon gratings", Technical Digest Series, **5**, Optical Society of American, April 1996.

7.M. W. Farn, M. B. Stern, W. B. Veldkamp, and S. S. Medeiros, "Color separation by use of binary optics", Opt.Lett. **18**, pp.1214, 1993

8. M. W. Farn, R.E.Knowlden, M.B.Stern and W.B. Weldkamp, "Color separation gratings", NASA conf publ. **3227**, pp.409,1993

9.Guo Yongkang, SU Jingqin and DU Jinglei etc., "Binary optical element for dividing harmonic waves applied in ICF syste", SPIE **3557**, pp.161-167.

10.M.P.Cahng, O.K.Ersoy, B.Dong,G.Yong, and B.Gu,"Iterative optimization of diffractive phase elements simultaneously implementing several optical functions", Appl.Opt., **34**,No.17,pp.3069-3076.

Dynamic and Spectral Responses of a Micromachined Fabry-Perot Interferometer-Based Spectrometer

Kuo-Kang Liu*[a], Shih-Ping Lee[b], Ran-Jin Lin[b], Long Hsu[c]

[a]School of Mechanical and Production Engineering, Nanyang Technological University, Singapore 639743
[b]Industrial Technology Research Institute, Hsinchu, Taiwan 310, R.O.C.
[c]Department of Electro-Physics, National Chiao Tung University, Hsinchu, Taiwan 300, R.O.C.

ABSTRACT

The dynamic and spectral responses of a micromachined Fabry-Perot interferometer (FPI) subjected to an electrostatic force have been simulated. The FPI features a circular corrugated diaphragm supporting a flat thin film coated optical mirror that can be driven by four electrodes. These four electrodes serve not only to fine adjust the parallelism but also to tune the air gap between the mirrors and hence the characteristic wavelengths of the interferometer. Regarding to the dynamics response of the flat diaphragm, we consider a force balance among the squeezing-film damping, the spring force of the corrugated diaphragm, the inertia force of the flat diaphragm, and the electrostatic attraction force. The spectral response of a novel FPI based spectrometer that consists of two FPIs in tandem has been modeled and simulated. Contrary to the second-order oscillation normally existed in a mass-spring-damper system, for the current system, the dynamic of the optical mirror released from a pre-deflection displays an over-damping effect, which depicts the simulated mirror deflection as a function of time relaxing from a 6 μm pre-deflection. This interesting phenomenon may exist pervasively in various micro-systems due to its higher area to mass ratio compared to that of the macro-systems. Our simulated result demonstrates that the tandem FPI based spectrometer, in principle, is able to obtain a spectral response with high accuracy from a light source having continuous spectrum.

Keywords: Fabry-Perot Interferometer, Spectral Responses, Micromaching, Squeezing-film Damping, Dynamic Responses.

1.INTRODUCTION

A Fabry-Perot interferometer (FPI) is an optical device that consists of two partially reflecting parallel mirrors separated by a distance to form a resonant cavity. When the distance equals multiples of a half wavelength of the incident light, the transmission peaks are then generated by multiple reflections of the light in the cavity. In order to provide a sufficient bending rigidity and hence the higher selectivity of FPI, the optically coated diaphragm of the current FPI system is designed to be several times in thickness than the corrugated diaphragm (see Figure 1). The current firstly simulation describes that the miniature FPI subject to electrostatic forces to produce an initial deflection in the silicon membrane and then the forces are released. Subsequently, diaphragm's motion which is mainly governed by the force balance between the squeezing-film damping, the spring force of the corrugated diaphragm, the inertia force can be modeled as a second-order mass-damping-spring vibration. The second part of the current simulation is focused on the optical response of a FPI based spectrometer, which is comprised of two FPIs in series.[1] The simulation result shows that this novel spectrometer is able to achieve a high accuracy for resolving the input spectrum. In order to resolve the input spectrum, a novel algorithm for solving the corresponding compound transparency matrix of the spectrometer is developed and presented here.

2. MODELING

The spring force of the system is mainly contributed from the deflection of corrugated structure. The relationship between applied force and the deflection for a circular corrugated diaphragm without internal stress can generally be described as[2]

$$F_k = \left(\frac{Eh^3\pi}{a^2}K_P A_P\right) y + \left(\frac{Eh\pi}{a^2}L_P B_P\right) y^3 = K_{s1}\, y + K_{s2}\, y^3 \quad , \tag{1}$$

Part of the SPIE Conference on Advanced Photonic Sensors and Applications
Singapore • November–December 1999
SPIE Vol. 3897 • 0277-786X/99/$10.00

where F_k is spring force

a is radius of diaphragm

E is modulus of elasticity

y is center deflection

h is thickness of diaphragm

A_p is a dimensionless stiffness coefficient

B_p is a dimensionless coefficient of the nonlinear term

K_p is the reduce factor of stiffness coefficient for a rigid center

and L_p is the reduce factor of tension coefficient for a rigid center

For a corrugated diaphragm, the value of A_p increase with the depth of the corrugation and however B_p decreases rapidly with the depth. These two parameters may be expressed by the equations:

$$A_p = \frac{2(3+q)(1+q)}{3[1-\nu^2/q^2]} \quad , \tag{2}$$

$$B_p = \frac{32}{q^2-9}[\frac{1}{6} - \frac{3-\nu}{(q-\nu)(q+3)}] \quad , \tag{3}$$

where ν is Poisson's ratio and for a sinusoidal corrugation profile, the profile factor q is defined as

$$q = [1.5\frac{H^2}{h^2}+1]^{1/2} , \tag{4}$$

where H is the depth of corrugation.

A rigid center is necessary if the diaphragm is used to convert pressure into a force as it provides a hosting plate to which electrodes may be attached. The values K_P and L_P for various b/a (the ratio of boss radius to diaphragm radius) and q is shown in the Figure 23.6 & 23.7 of the reference.[2]

The electrostatic force is given by

$$F_e = \frac{V^2 A_e \varepsilon}{2(h_o - y)^2} \quad , \tag{5}$$

where V is the applied voltage, A_e is the total electrode area, ε is the permittivity of air, h_0 is the initial gap distance of cavity.

Squeeze-film damping is used to describe the damping experienced by the moving diaphragm. This damping force F_b for a circular diaphragm can be presented as[3]

$$F_b = -\frac{3\pi\mu a^4}{(h_0 - y)^3}\frac{\partial y}{\partial t} = -\frac{K_b}{(h_0 - y)^3}\frac{\partial y}{\partial t} , \tag{6}$$

Where μ represents the viscosity of air.

Further taking into account the inertia of the diaphragm, the governing equation for the moving diaphragm can be comprehensively expressed as[4].

$$m_d \frac{\partial^2 y}{\partial t^2} + \frac{K_b}{(h_0 - y)^3} \frac{\partial y}{\partial t} + K_{s1} y + K_{s2} y^3 - \frac{V^2 A_e \varepsilon}{2(h_o - y)^2} = 0 \quad , \tag{7}$$

It is obvious that the equation (7) is intrinsically nonlinear and can only be solved numerically.

In order to resolve a continuous spectrum, a spectrometer comprising of two FPIs in series has been modeled and simulated. Suppose the ranges of variations in the gap distances of the resonance cavities of two FPIs (g_1 and g_2) are 6.2-3.2 μm and 15-14 μm respectively and the transparency spectrum of a single interferometer is comb-like. If the gap of the first interferometer is initially set to 15 μm, when the gap of the second interferometer varies, a corresponding transparency spectrum may be read at each step of variation. When the combination transparency integration of the two resonance cavities is obtained, 16 sets of linear equations (when 16 variables are used), are obtained and can be expressed as

$$\sum_{i=1,16} a_{ij} P_j = S_j \quad (j = 1,2,...16), \tag{8}$$

where a_{ij} is the compound transparency which is the integration of the transparency of the two interferometers; P_j is the multiplication of the 16 incident spectra and the sensitivity of photosensor and the transparency of optical diaphragm; and S_j is a set of the signal currents of the photosensor at the 16 specific wavelengths. The transmission of individual FPI is described by the Airy function:[5]

$$I = \left[1 - \frac{A}{(1-R)}\right]^2 \frac{1}{1 + \dfrac{4R\sin^2(kd\cos\theta)}{(1-R)^2}} \quad , \tag{9}$$

where A is the mirror absorptance, R the mirror reflectance, d the cavity gap, θ the angle of incidence of the beam, and $k=2\pi/\lambda$ (λ is the wavelength)

In the simulation of spectral response, the gap of first interferometer is varied a step for, e.g. 0.2 μm. Then the incident is scanned again and a new set of spectrum spots are obtained. The above process is repeated and a virtually continuous spectrum is obtained. In fact during the scanning of the resonance cavity, the variation curve of the signals currents of the photosensors varies within a great range. This illustrates that much more information is contained in the incident light. Even if the gap of one interferometer is kept unchanged, more information may be obtained from the signals by adjusting the gap of the other, although the matrix may be more complicated.

3. RESULTS AND DISCUSSIONS

Simulation is conducted under Mathamatica 3.0 (a software provided by Wolfram Research). The values of parameters employed in this simulation are shown in Table 1.

The simulated results for the diaphragm in static situation (neglect the damping and inertia effect expressed in Equation (7)) are shown in Figure 2 and 3. The relationship between deflection and applied voltages shown in Figure 2 is nonlinear. The nonlinearity is mainly generated by the intrinsically nonlinear relationship between the applied voltage and the gap distance of the electrodes (see (5)). However, the relationship between the applied loads of electrodes and the deflection (see Figure 3) is shown to be very linear. This linearity is mainly contributed by the flexible corrugated structure which can significantly reduce the radial stress in the diaphragm.

Figure 4 shows the dynamics response (deflection versus time) of a diaphragm released from a 6 μm pre-deflection; the result is predicted using Eq. (7). It is worthy to be pointed out that in the current system the damping force, represented as the second term of Eq. (7), is 100 times larger than the inertia force, the first term. Hence, an over-damping effect is clearly reflected in the diaphragm's vibration as shown in Figure 4. This interesting phenomenon may exist pervasively in various micro-systems due to its higher area to mass ratio compared to that of the macro-systems.

Figure 5 shows the simulation results of the FPI based spectrometer. There are three curves: the spectrum of the hypothetical incident light (curve a), the spectrum of the multiplication of the sensitivity of photo diode and the transparency of silicon membranes (curve b) and the spectrum of the multiplication of curve a and curve b (curve c). Since a linear axis is needed to integrate the spectral transmittance of the incident light, the x-axis represents the frequency under the unit of 10^{14}Hz. The resolved spectrum (dots shown in Figure 5) which is calculated using the Eq. (8) with its corresponding algorithm shows a good agreement with the hypothetical input spectrum (curve c).

4. CONCLUSIONS

The results of simulation show a quasi-linear relationship between the applied loads versus the deflection of the diaphragm and a nonlinear relationship between applied voltage and the deflection. Regarding to the dynamics response of the optically coated diaphragm, an over-damping effect existing in the relationship of the traveling distance versus time. The results of spectral response of a FPI based spectrometer that is comprised of two FPIs in series show a good agreement with the hypothetical input spectrum. The simulated result demonstrates that the tandem FPI based spectrometer, in principle, is able to obtain a spectral response with high accuracy from a light source having continuous spectrum.

REFERENCES

1. S. P. Lee, R. J. Lin, H. H. Chen and K. K. Liu, Micromachined Non-invasive Glucose Meter, USA Patent (Pending), 1999.
2. M. D. Giovanni, "Flat & Corrugated Diaphragm Design Handbook", Marcel Dekker, New York, 1982.
3. J. B. Starr, " Squeeze-film damping in solid-state accelerometers", Proc. IEEE Solid State Sensor and Actuator Workshop, Hilton Head Island, SC, USA, 4-7 June, pp. 44-47, 1990.
4. P. B. Chu, P. R. Nelson, M. L. Tachiki and K. S. J. Pister, "Dynamics of polysilicon parallel-plate electrostatic actuators", *Sensor and Actuators* **A52**, pp. 216-220, 1996.
5. J. H. Jerman, D. J. Clift and S. R. Mallinson, "A miniature Fabry-Perot interferometer with a corrugated silicon diaphragm support", *Sensor and Actuators* **A29**, pp. 151-158, 1991.

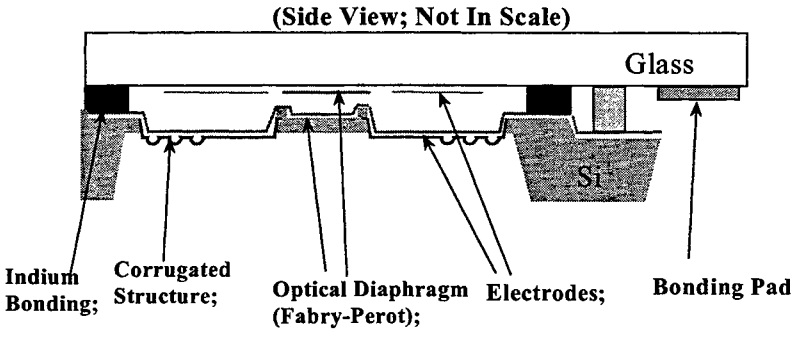


Figure 1 the major parameters of Fabry-Perot Interferometer structure

Table 1. the values of simulation parameters

E=1.657x10^5 MPa	a=2100 μm	ν=0.29	b=1850 μm
h=1.5 μm	q=4.352	ε=8.8542×10^{-12} F/m	μ=1.8×10^{-11} Kg/μm-Sec
H=5 μm	Ae=3.1 cm^2	A=0	R=0.95

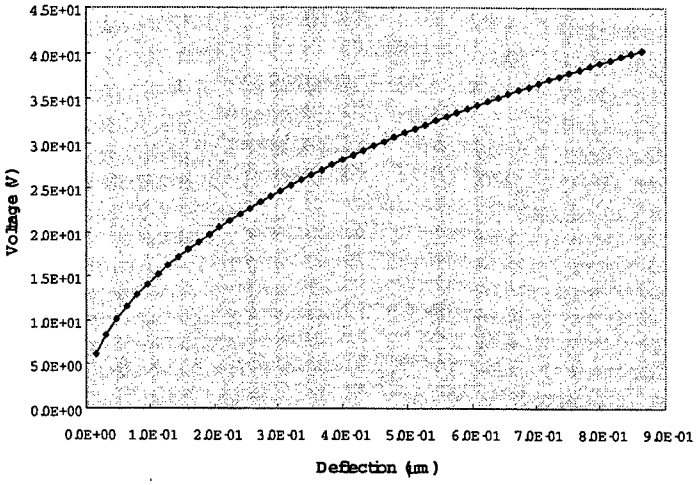


Figure 2 Deflection of the diaphragm's center versus the applied voltages of electrodes

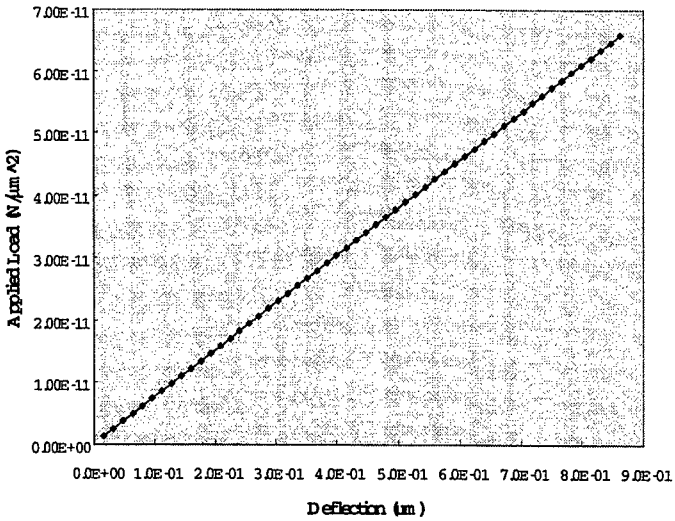


Figure 3 Deflection of the diaphragm's center versus the applied loads of electrodes

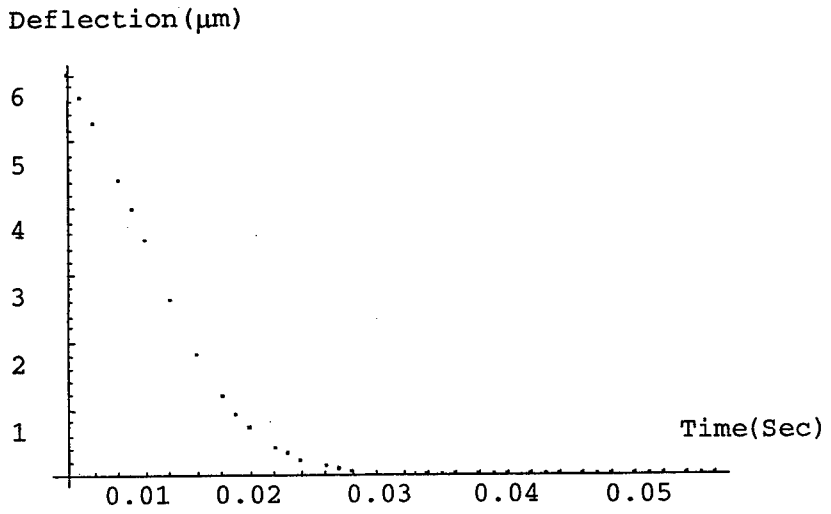


Figure 4 Deflection versus time of the mirror diaphragm released from a 6μm pre-deflection

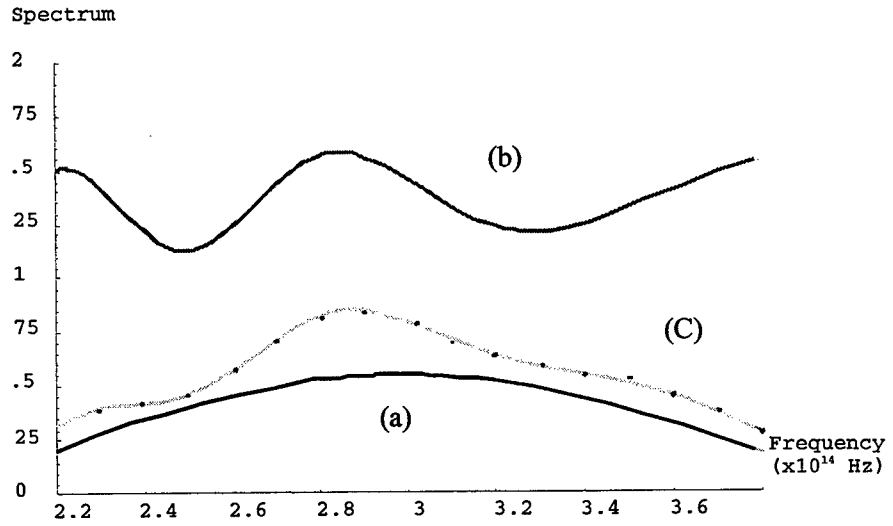


Figure 5 The simulation of FPI based spectrometer: the spectrum of hypothetical light source (curve a); the spectrum of the multiplication of the photo detector sensitivity and the transmission of mirror diaphragm (curve b); the spectrum of the product of curve a and b (curve c); and the resolved spectrum (dots) calculated by Eq. (8).

External-cavity semiconductor laser sensor

O.L.Quan[b],Y.T.Seng[b],L.P.Lan[b], J.T.K.Wah,[b]A.T.Sabaratnam[b],*G.M.Hegde[a] and A.Selvarajan[a]

[a]Department of Electrical Communication Engineering

Indian Institute of Science, Bangalore-560 012, INDIA.

[b]Department of Electronics and Computer Engineering

Ngee Ann Polytechnic, SINGAPORE-599 489.

ABSTRACT

The spectral characteristics of a diode laser are significantly affected due to interference caused between the laser diode output and the optical feedback in the external-cavity. This optical feedback effect is of practical use for linewidth reduction, tuning or for sensing applications. A sensor based on this effect is attractive due to its simplicity, low cost and compactness. This optical sensor has been used so far, in different configurations such as for sensing displacement induced by different parameters. In this paper we report a compact optical sensor consisting of a semiconductor laser coupled to an external cavity. Theoretical analysis of the self-mixing interference for optical sensing applications is given for moderate optical feedback case. A comparison is made with our experimental observations. Experimental results are in good agreement with the simulated power modulation based on self-mixing interference theory. Displacements as small as 10^{-4} nm have been measured using this sensor. The developed sensor showed a fringe sensitivity of one fringe per 400nm displacement for reflector distance of around 10cms. The sensor has also been tested for magnetic field and temperature induced displacement measurements.

Keywords: Sensor, external-cavity, self-mixing

1. INTRODUCTION

For many practical applications different types of optical sensors have been proposed, in which laser diodes are used as their light sources because of their compactness and modulation capability. It is also known that laser diode itself can be used as a sensor when operated in an external-cavity configuration[1,2,4]. Diode laser sensors based on self-mixing interference have been of interest due to their simple, compact and inexpensive nature. The major components of such systems are a laser diode, basic collimating optics and a reflector.

The spectral characteristics of a laser diode (LD) are very much affected by external optical feedback[3]. Light emitted from a laser diode can be coupled back into its cavity by an external reflection. The output intensity of such a diode laser is modulated by the incoming light from the external reflector with the change of the reflector distance. The presence of the external reflector creates a standing wave which alters the effective facet reflectivity of the LD. A slight change in the position of the external reflector alters the phase of the light reflected back into the laser cavity thus varying the effective laser facet reflectivity. The back scattered light is then self-mixed with the light inside the LD cavity causing variations of the laser output power. These variations of the optical output power due to interference are monitored by the photodiode within the laser diode package to determine displacements[12], distances[13] or velocities[14]. In these applications the laser diode acts as a sensitive optical phase detector and the sensor system has the advantages of being very simple and compact in terms of optical configuration. As in the case of fiber sensors the diode laser sensor can be configured to sense temperature, current, acoustic waves, velocity and acceleration[1,2,5]. For sensing these parameters, so far the diode laser sensor

* Correspondence: Email: sat@np.edu.sg ; hegopal@mrc.iisc.ernet.in; rajan@ece.iisc.ernet.in

Part of the SPIE Conference on Advanced Photonic Sensors and Applications
Singapore • November–December 1999
SPIE Vol. 3897 • 0277-786X/99/$10.00

was used in different configurations. In the present work we have developed a diode laser sensor which can be used to detect the displacement, temperature, magnetic field and current without changing the sensor configuration. Our experimental results for different levels of optical feedback are compared with the self-mixing interference theory[6].

2.THEORY OF SELF-MIXING INTERFERENCE

The operation of a diode laser with an external optical feedback has been studied extensively[3-10]. When the optical feedback is very weak the spectral linewidth may be narrowed or broadened, depending on the face of the feedback relative to the optical field within the active laser, and this behavior has been used for linewidth reduction and tuning. The self-mixing is explained to be the consequence of the optical spectrum and threshold modulation produced by external optical feedback. A simplified theory of self-mixing interference based on the assumptions of three mirror Fabry-Perot structure has been given by earlier authors[6,8,10], which considered that the power modulation by the self mixing was due to the reflectivity variations of the laser facet. It is known that the phase condition $\phi(\nu) = 2q\pi$ in the presence of feedback determines the lasing frequency of the diode laser in the external cavity[8] and this leads to

$$\Delta\phi(\nu) = 2\pi\,(\nu-\nu_0)\tau_d + \xi\sqrt{(1+\alpha^2)}\,\sin[2\,\pi\nu\tau_L + \arctan(\alpha)] = 0 \qquad (1)$$

where $\Delta\phi(\nu)$ represents the deviation of the oscillation phase $\phi(\nu)$ from $2q\pi$, α is the linewidth enhancement factor, $\xi = (1-R_2)\,r_3/r_2$ denotes the coupling effect from the external reflection with $R_2 = |r_2|^2$, r_2 and r_3 are the amplitude reflection coefficients of the laser facet and the external reflector respectively, ν_0 is the laser emission frequency without feedback, $\tau_d = 2n_od/c$ and $\tau_L = 2L/c$ are the round trip time periods inside the diode laser and in the external cavity, respectively, and n_o is the refractive index of the laser material. The external feedback strength is defined by the parameter[6], $C = \frac{\tau_L}{\tau_d}\,\xi\sqrt{(1+\alpha^2)}$.

In the case of weak feedback (C<1), the phase change $\Delta\phi$ and the emission frequency is monotonic function which results in only one solution for $\Delta\phi = 0$ and thus emits a single frequency. Using first order approximation eqn(1) may then be solved which gives the emission frequency ν as,

$$\nu = \nu_o - \frac{C\sin[\phi_{ext} + \arctan(\alpha)]}{2\pi\tau_L\{1 + C\cos[\phi_{ext} + \arctan(\alpha)]\}} \qquad (2)$$

This equation shows that the frequency shift is a periodic function with respect to the external reflection phase ϕ_{ext}.
In the case of weak feedback the lasing spectral linewidth is given by

$$\delta\nu = \frac{\delta\nu_0}{(1 + C\cos[2\pi\nu\tau_L + \arctan(\alpha)])^2} \qquad (3)$$

where $\delta\nu_o$ is the linewidth of the solitary diode laser in the absence of feedback.

When the length of the external cavity is varied, the output power will vary between constructive (P_{max}) and destructive (P_{min}) interference. Therefore the laser output power is expressed with respect to the laser spectral linewidth $\delta\nu$, as[6]

$$P = P_0 [\, 1 + m \, \{\exp(-\delta\nu\tau_L/2)\}\cos(\pi\nu\tau_L)] \qquad (4)$$

where P_0 is the output power in the absence of feedback and m is the modulation coefficient depending on the reflection coefficient of the reflector, which is a measure of the available signal strength in the laser optical intensity with the self-mixing interference. The above equation represents the theoretical expression describing the power modulation for self-mixing interference in a single-mode diode laser with weak optical feedback.

On similar theoretical basis the authors[8] have presented an expression for intensity modulation. Assuming that the diode laser is in a steady state and characterized by optical intensities I_0 and I in the absence and in the presence of weak optical feedback respectively, the intensity modulation is given by

$$I = I_0 [1 + m \cos(2\pi\nu\tau_L)] \qquad (5)$$

which is again a repetitive function with a period of 2π radians. The authors have also studied the effect of external reflector alignment in a diode laser sensor[11]. Analysis of optical feedback effects on the linewidth of a diode laser sensor using self-mixing interference is also reported[9].

3. SENSOR SYSTEM CONFIGURATION

The experimental set up of the sensor for displacement measurement using self-mixing effect is shown in Fig.1.

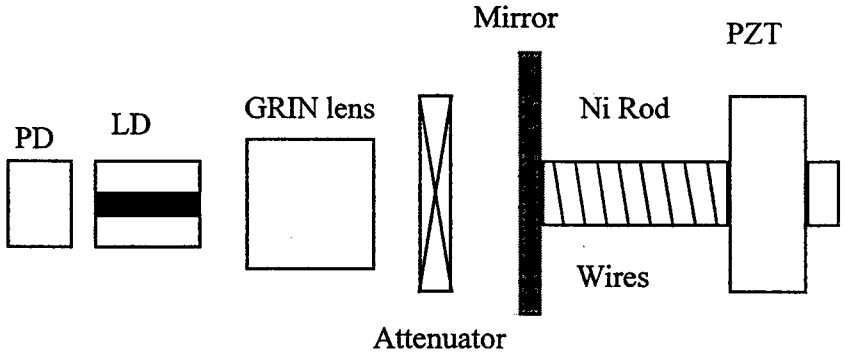


Fig.1 Experimental sensor configuration

The sensor head consists of a single mode, cw diode laser (Sharp LT022 MC/MD, λ=780nm, 5mW) package and a GRIN rod lens. A plane mirror (R≈90%) mounted on a PZT was used as an external reflector. An optical attenuator was inserted between the LD and the external reflector to supress the higher order modes whenever necessary. This also controls the feedback coefficient C. Light output was monitored by a photodiode equipped in the LD package. The LD was mounted on a thermoelectric cooler for temperature stabilization. Temperature was maintained at a constant value of 15 ^{0}C during the experiment. The entire sensor set up was mounted in an isolation chamber. The external reflector was moved by a PZT translation and the corresponding light output was monitored by the photodiode. The laser output varied sinusoidally with a periodicity of $\lambda/2$; that is approximately 390nm, corresponding to a π-rad. phase shift in the movement of the external reflector which is shown in Fig.2.

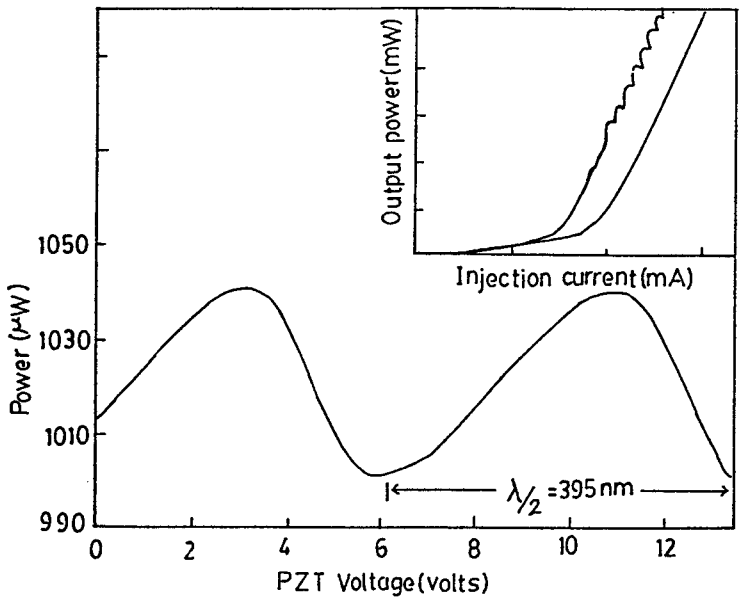


Fig.2 Output power variation of LD with external reflector displacement.

To use the system as a magnetic field sensor, the external reflector was attached to the material whose magnetic properties are such that the induced magnetic field displaces the reflector, with the displacement being proportional to the magnetic field strength. The reflector was mounted on one end of a 2cm. long, 1cm. dia. nickel (Ni) rod. The other end is attached to the PZT. The Ni rod is wrapped parallely with a thin nichrome wire and a thick(1mm) insulated copper wire which helps to induce the reflector displacement either by temperature or by magnetic field. For any sensing applications the device was stabilized at maximum sensitivity by adjusting the voltage on the PZT. The sensitivity of this device as a displacement sensor is first determined by the magnitude of the movement of the reflector to the incident field and second by the combined noise sources.

4. EXPERIMENTS , RESULTS AND DISCUSSION

The output characteristics of LD on optical feedback was measured by changing the reflector distance by a PZT. The reflector distance was about 8cm. When the drive current is fixed (dc) and the external reflector is moved along the optical axis, the light intensity is modulated periodically for every half a wavelength displacement of the reflector as shown in Fig. If the injection current is modulated with sawtooth waves the optical feedback effect observed is a periodic light modulation in the linear region, as shown in Fig2.,inset. Then the displacement is calculated by[4], $\Delta l = \Delta\phi\lambda/2T$, where T is the modulation period. In the case of temperature sensor, from the analysis of the linear expansion coefficient of the material versus the temperature it is possible to identify the regions where a small change in temperature results in maximum change in length. The operating point for maximum sensitivity is obtained by monitoring the output from the rear facet with a photodiode. Fig.3 shows the output intensity modulation of LD caused by the temperature induced displacement. It is possible to measure the displacements of about 10^{-4} nm, which corresponds to a temperature change of 0.01 ^{0}C.

When the sensor is operated for a magnetic field induced displacement, the function which relates the change in magnetostriction of the material to the applied magnetic field and the applied bias field is to be considered. By analyzing the dimensional change versus magnetic field strength curves for different materials it is possible to identify the region where maximum dimensional change occurs for a small change in the applied magnetic field.

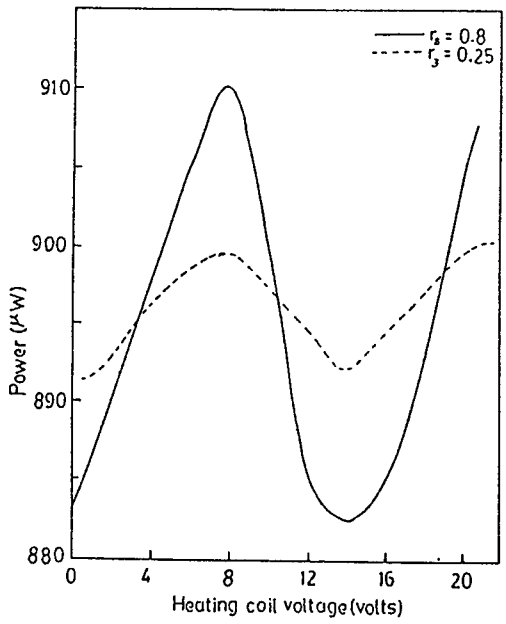


Fig.3 Experimentally measured power variation with temperature induced reflector displacement.

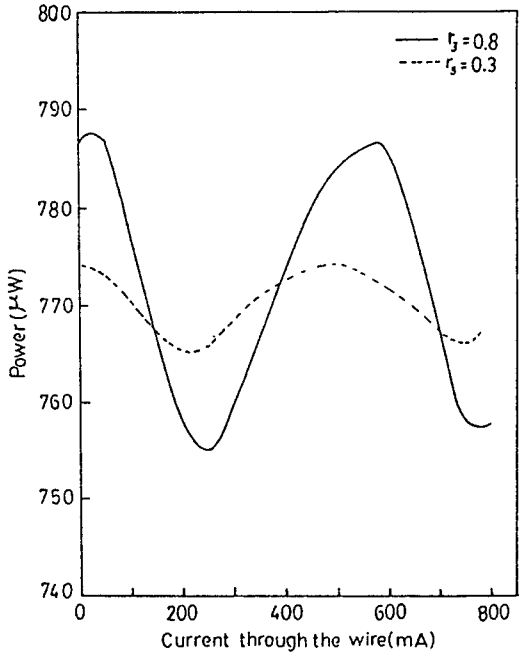


Fig.4 Experimentally measured power variation with magnetic field induced reflector displacement

For pure nickel the change in length (Δl) is maximum when the bias field is 3 Gauss (G). For this magnetic field $\Delta l / l \approx 3x10^{-6}$ H, where H is the bias filed in gauss. We have used a 2cm long nickel rod, for which Δl = 30nm for a magnetic filed of about 1.5G, this causes a phase change of about 0.2 rad. Thus if a displacement of 10^{-4} nm is detectable, then the minimum magnetic field required to produce this displacement would be about 10^{-5}G. In Fig.4, output power variation of the LD for the reflector movement induced by the magnetic field is given. If a translation element of higher magnetostrictive constant is used, the sensitivity of the sensor may be increased by an order of magnitude. Once the sensor is calibrated for magnetic field for different bias currents, the sensor can then detect the current passing through the coil associated with the magnetic field, hence making the device as current sensor.

Our experimental observations are compared with the theoretical simulation of power variation with phase change based on self-mixing interference. The emission frequency of LD in the case of feedback also depends on the external phase change and Fig.5 shows the emission frequency shifts with respect to the external reflection phase for various feedback strengths. Higher values of frequency shifts for strong feedback clearly indicates that the laser emission is multimode. Following eq(1), the power dependence on external phase change for different levels of feedback is depicted in Fig.6. The parameters used for the calculations are λ = 780nm, $\delta\nu_0$ = 80 MHz and L= 10cm. It is clear from Figs.5 and 6 that for weak feedback both frequency shift and output power are of sinusoidal form and for strong feedback the variations become saw-tooth like with the external phase ϕ_{ext} . The results obtained from experiment for power modulation for different levels of feedback shown in Figs.3 and 4, are in good agreement with the theoretical simulation. For weak feedback the self-mixing is similar to that of two beam interference and the spectrum remains single mode. In the case of stronger feedback level, because of multiple external-cavity modes the waveform becomes sawtooth like which clearly shows that the performance of the developed sensor follows self-mixing interference theory. It has been concluded in the theoretical model[8] that the inclination of the fringe pattern is related to the change of phase sign in the emission frequency shifts with the relative direction change of the external reflector movement.

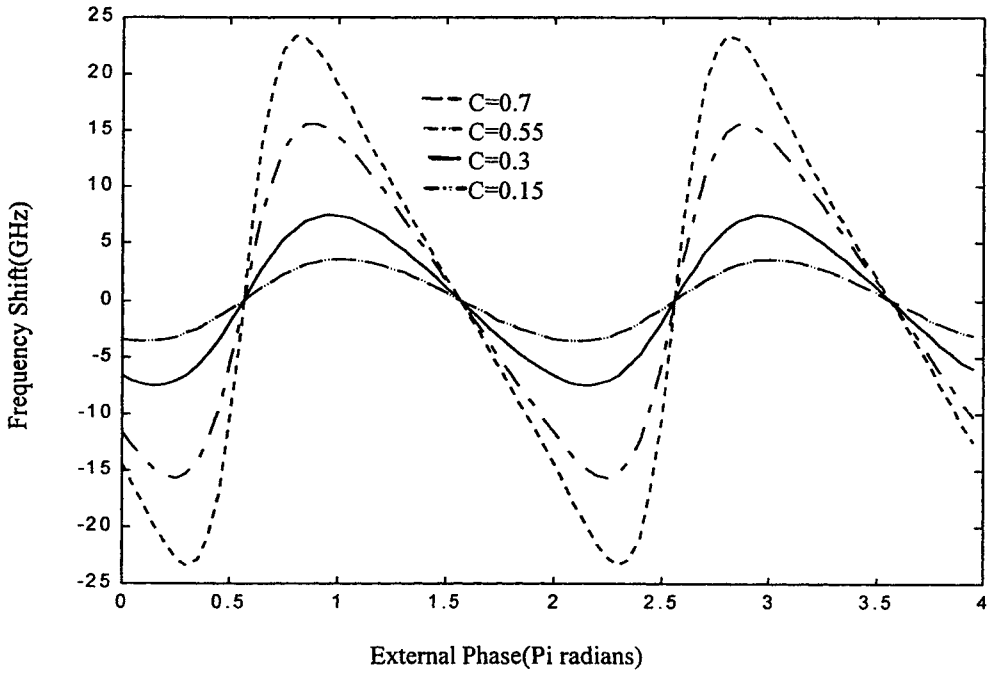


Fig.5 Theoretical simulation of laser frequency shift dependence with external phase for different values of C

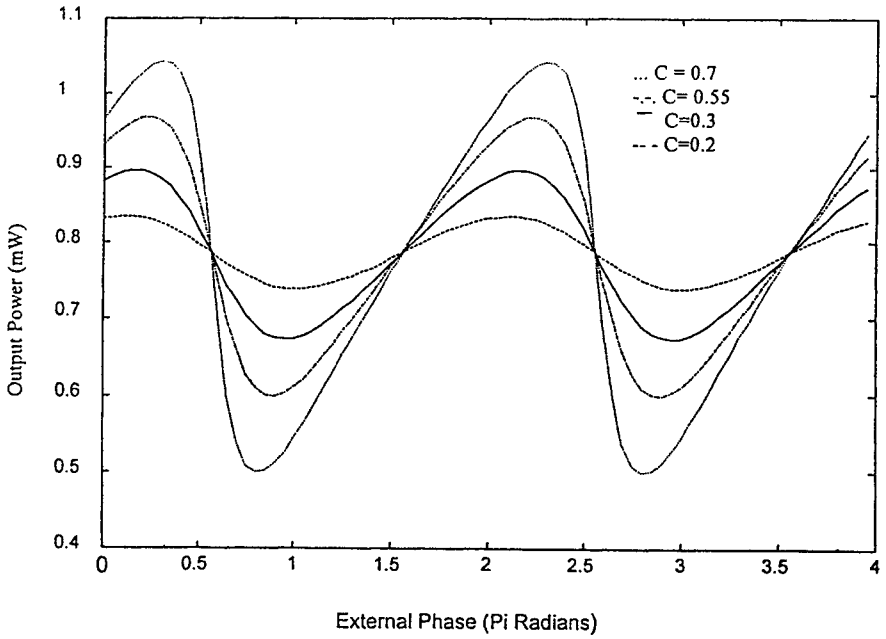


Fig.6 Theoretical simulation of power change with external phase for different C values

In conclusion, we have demonstrated a compact and low cost diode laser sensor based on self-mixing interference. The developed sensor can be used to detect very small temperature and magnetic field variations through measurement of displacements of the order of fraction of a nanometer. The sensor may find applications wherever a noncontact or nondestructive technique for detecting these parameters is required.

ACKNOWLEDGEMENT

The authors would like to thank Dr.Lim Choo Min for encouragement and Mr.Yong Saw Soon for his help in the experiments.

REFERENCES

1. A.Dandridge, R.O Milles and T.G.Giallorenzi, "Diode laser sensor" *Electron. Lett.*, **16,** pp. 948-949,1980.
2. T.Bosch and M.Lescure, Eds. *Selected Papers on Laser distance Measurements,* SPIE Milestone Series, MS **115,** 1995.
3. R.Lang and K Kobayashi, "External optical feedback effects on semiconductor injection laser properties", *IEEE J.Quantum Electron.*, **QE-16,** pp. 347-355, 1980.
4. S.Shinohara, A.Mochizuki, H.Yoshida and M.Sumi, "Laser Doppler velocimeter using the self-mixing effect of a semiconductor laser diode", *Appl.Opt.*, **26,** pp.1417-1419, 1986.
5. J.Kato, N. Kikuchi, I. Yamaguchi and S. Ozono, "Optical feedback displacement sensor using a laser diode and its performance improvement", *J.Meas.Sci.Technol.*, **6**, pp.45-52, 1995.
6. W.M.Wang,W.J.O.Boyle, K.T.V.Gratten and A.W.Palmer, "Self-mixing interference in a diode laser: Experimental observations and Theoretical Analysis", *Appl. Opt.*, **32,** pp. 1551-1558,1993.
7. M.Shunk and K Peterman, "Numerical analysis of the feedback regimes for single mode semiconductor lasers with external feedback", *IEEE J.Quantum Electron,* **QE-24,** pp.1242-1247, 1988.
8. W.M.Wang , K.T.V. Gratten, A.W.Palmer and W.J.O.Boyle, "Self-mixing interference inside a single-mode diode laser for optical sensing applications" *J.Lightwave Technol.*, **12,** pp. 1577-1587 ,1994.

9. G.Mourat., N Servagent and T.Bosch., "Optical Feedback effects on the spectral linewidth of semiconductor laser sensors using self-mixing interference", *IEEE J. Quantum Electron.*, **QE-34,** pp.1717-1721,1998.
10. N.Servagent, F.Gouaux and T.Bosch, " Measurement of displacement using the self-mixing interference in a laser diode", *J.Opt.*, **29,** pp.168-173, 1998.
11. R.C.Addy, A.W.Palmer and K.T.V. Gratten, "Effect of external reflector alignment in sensing applications of optical feedback in laser diodes" , *J .Lightwave Technol.*, **14,** pp.2672-2676,1996.
12. N.Servagent, T.Bosch, and M.Lescure, ."A laser diode displacement sensor using the self-mixing effect for modal analysis and defect detection", *IEEE Trans. Instrum. Meas.*, **46,** pp.847-850 ,1997.
13. S.Donati, G.Giuliani and S.Merlo, " Laser diode feedback interferometer for measurements of displacements without ambiguity", *IEEE J.Quantum Electron.*, **QE-31,** pp.113-119, 1995.
14. T.Shibata, S.Shinohara, H.Ikeda, H.Yoshoda, T.Sawaki and M.Sumi, "Laser speckle velocimeter using self-mixing laser diode", *IEEE Trans. Instrum.Meas.*, **45,** pp.499-503, 1996.

Nonintrusive Surface Inspection Utilizing Infrared Light Radiation

Hatim Abdul Hamid, Wojtek Wlodarski and Frank Brennan

Department Communications and Electronics Engineering,
Royal Melbourne Institute of Technology,
Sensor Technology Laboratory,
City Campus, GPO Box 2476V,
Vic. 3001, AUSTRALIA.

ABSTRACT

The need for quality control has lead to the implementation of automatic surface inspection systems, which have improved on-line monitoring of surface quality. Enamelled copper wire, cable and optical fibre are three examples where surface quality is very important. An extensive literature review conducted by the authors, has shown that with the current state of technology, there is much room for improvement in the field of non-destructive defect detection for enamelled copper wire. In this paper, the authors describe an infrared light based surface inspection system which has been developed for non-destructive defect detection on cables, optical fibre and specifically on enamelled copper wire. Finally, results from extensive trials at an enamelled copper wire manufacturing company are presented and compared to a simulatation, of the defect detection head.

Keywords - nonintrusive, defect sensor, reflection, surface inspection, surface analysis

1. INTRODUCTION

In the past and to some extent even to date, visual and touch inspection are common methods for determining the surface quality of a product. However, because of their inherently subjective and inconsistent nature, both visual and touch inspection do not provide satisfactory results. Often monotonous working conditions make the job of visual and touch inspection too difficult and demanding for the worker, and hence is not cost effective for the manufacturer. To compensate for the limitations of the human sensors automated surface inspection systems have been implemented.

Well known manufacturers have made significant technological advances in the area of surface inspection. These advances have been achieved through the application of computerised defect inspection systems. In general surface inspection systems may be categorised by their functionality. These may include: destructive of non-destructive, online or off-line, suitable for

Part of the SPIE Conference on Advanced Photonic Sensors and Applications
Singapore • November–December 1999
SPIE Vol. 3897 • 0277-786X/99/$10.00

curved or flat surfaces and do they operate for high or low production speeds. It should be noted that although there are established techniques for surface inspection, in general, there has been a very small effort put into the development of a non-destructive surface inspection technique for enamelled copper wire. Having this in mind, there has been ongoing research for the past three years to develop a non-destructive defect detection system for enamelled copper wire. The aim of this paper is to present the novel non-destructive defect sensor, and to show the practical applications and results from extensive trials of the defect sensor head.

2. SURFACE INSPECTION SYSTEMS

The need for detecting surface imperfections or defects, has existed as long as the manufacturing process itself. It is well known, that a perfect surface quality is impossible to achieve. Using modern equipment for surface examination, with resolutions down to atomic levels, it has been clearly shown that no surface is free from features which may be classified as defects[1].

Non-destructive surface inspection is an increasingly important element of the manufacturing process. Consumers consistently demand high quality products, and hence manufacturers are becoming increasingly aware of the quality of their product. In general non destructive surface inspection may be based on acoustics (sound), penetrating radiation (x-rays), light (ultraviolet infrared, visible), electric and magnetic fields.

One of the first non-destructive methods used for defect detection was based on visible light, using a simple optical comparator[3]. A cylindrical metal product was passed through an optical unit, and the light present at the receiver was then focused onto a detector. The intensity of the light on the detector was then analysed. It was almost impossible to calibrate such a system. Calibration had to be done every time a new product was to be manufactured. The calibration standard was a metal rod with a known circular diameter. During the early stages, most defect detection systems were simple analog deviation type instruments, which did not offer an absolute readout.

The need for more accurate, efficient and reliable surface inspection systems, saw the development of a wide variety of defect detection systems. These systems are based on laser radiation[4], ultrasonics[5], microwave radiation[6] and CCD imaging[7,8]. Based on an extensive literature review[2] conducted by the authors, we have discovered that a dedicated surface inspection system for enamelled copper wire, does not exist. Existing surface inspection systems offer some form of defect

detection, however these surface inspection systems, lack specific features which the authors believe to be crucial for defect detection on enamelled copper wire.

Of the above mentioned techniques for surface inspection, solid state lasers have been most effective. The principle operating mechanism for a laser based surface inspection system (LBSIS) relies on the detection of scattered laser radiation. There are essentially 2 techniques for using laser as a means for surface inspection. The first technique is a single plane surface inspection systems, which measures diameter defects in one axis of the wire only. A single plane surface inspection system, relies on focusing laser light onto a rotating mirror drum. The rotating mirror transforms the single light spot into a beam of light. A transmitter lens must see to it that parallel beams are created throughout the measuring zone, and finally the receiving lens focuses the laser beam onto a photocell.

The second technique is a triple plane surface inspection systems, which covers the total circumference of the wire, with 60 degrees between each of the 3 axis. A 3-axis LBSIS, is different in principle to a single plane surface inspection system. A 3-axis LBSIS relies on 3 sources of continuous laser energy projected onto the wire, with 3 corresponding photocells to pick up scattered laser radiation. The intensity of laser radiation on the shadow side is measured "proportionally to the diameter" and its variations are evaluated electronically.

Laser based surface inspection systems are mainly used for diameter measurement, and hence the name diameter gauge. Note however, a laser based diameter gauge has a limited capability for defect detection. A defect passing through the inspection system, is treated as a sudden change in diameter. A positive change in diameter would indicate a lump, whereas a negative change in diameter would indicate a dint or a neckdown in the surface.

LBSIS are mainly used on optical fibre[9]. A 3-axis LBSIS is capable of detecting 200 micron sized defects, on a 100 micron diameter optical fibre, at production speeds up to 1000 m/min. Although the use of a solid state lasers has proved to be very accurate, this technology is very costly. When considering that copper wire manufacturers may have in excess of 150 production lines operating simultaneously, there is a definite need in the authors opinion for a less costly alternative, however not at the cost of both quality and performance.

3. NEW DEFECT DETECTION SYSTEM

A new defect detection system has been developed for on-line, continuous, non-destructive defect detection on enamelled copper wire. The new system can monitor production on all wire diameters, operating at full production speeds. The principle of operation of the proposed novel defect sensing system utilises infrared light radiation. The configuration of the sensor head has several unique features, which are vital for its operation. The sensor head seen in Figure 1, allows for a total view of the copper wire circumference at any point in time. Another key feature of the defect sensor head is its tolerance to wire vibration. Vibration in general, is a problem common to all surface inspection systems[10]. The prototype defect sensor head has been designed, keeping in mind that the defect sensing system will operate in a harsh industrial environment. This is where interfering parameters such as variable temperature, humidity, electromagnetic interference and wire vibration pose serious problems.

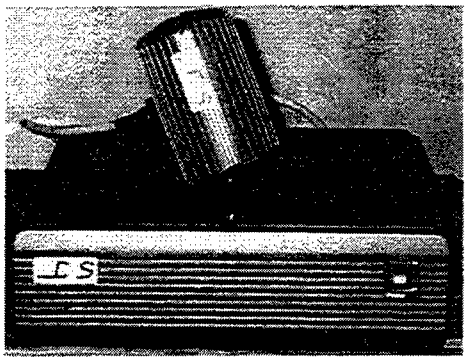


Figure 1: Prototype Defect Detection System

4. PRELIMINARY RESULTS

Over a period of 18 months, the defect detection system has been tested extensively at a factory, manufacturing enamelled copper wire. The system was installed on a 0.27mm production line, running at 4.5 meters per second. Hundreds of defect recordings were taken using a digital oscilloscope, which were then saved to disk. Following are some of the results obtained for the defects detected.

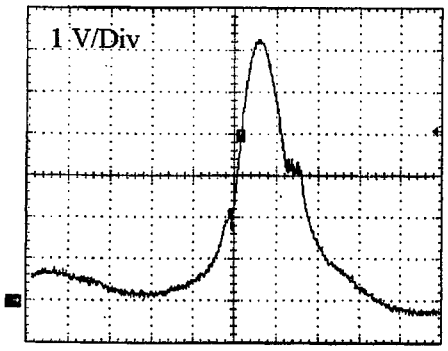


Figure 2a: Output Waveform

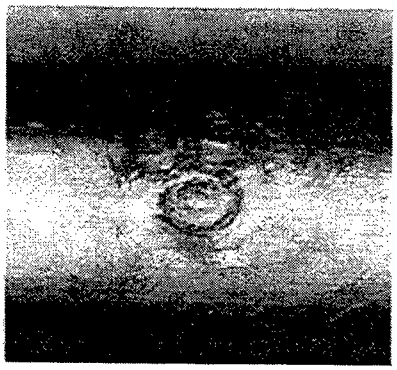


Figure 2b: Defective Wire

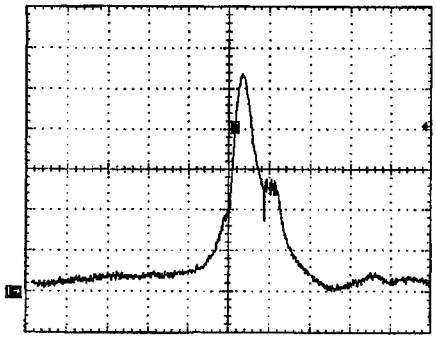


Figure 3a: Output Waveform

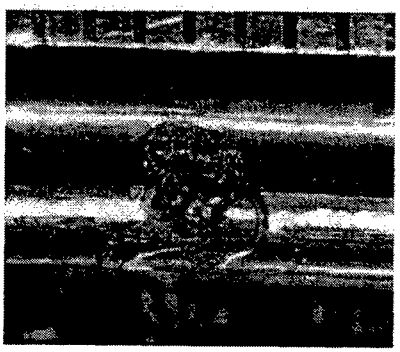


Figure 3b: Defective Wire

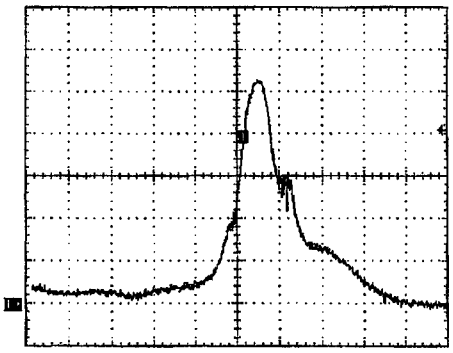


Figure 4a: Output Waveform

Figure 4b: Defective Wire

Figures 2a - 4a are the recorded output waveforms of the defects shown in Figures 4b - 4b respectively.

It can be clearly observed that for similar defects, the output waveforms are also similar. For the purpose of defect identification, this result is very promising. The consistency in the output signal suggests that the sensor head has a stable operation. The reproducibility of the wave forms suggest that the same defect or group of defects is continuously re-occurring. These results greatly enhance the possibility of using pattern recognition, which will allow for intelligent defect classification.

5. DISCUSSION

To both verify and confirm that the output of the sensor is indeed due to a defect, the sensor head was modelled mathematically, by adopting the Torrance and Sparrow reflectance model[11].

Figure 6 is the simulated response due to the defect shown in Figure 4b. It can be seen clearly that the simulated response is in very good agreement with the output recorded from the defect sensor, seen in the Figure 4b. Even though the simulated output may not at first instance resemble the actual output, it should be pointed out that on a closer examination, the particular wave form characteristics seen in Figure 5, points 1-4 closely resemble points A-D seen in Figure 6. One possible explanation for this difference is:

- The sensor response in not fast enough to resolve such a small defect, travelling at 4.5 meters a second. Hence the result is a merger of the peaks, resulting in a single waveform.

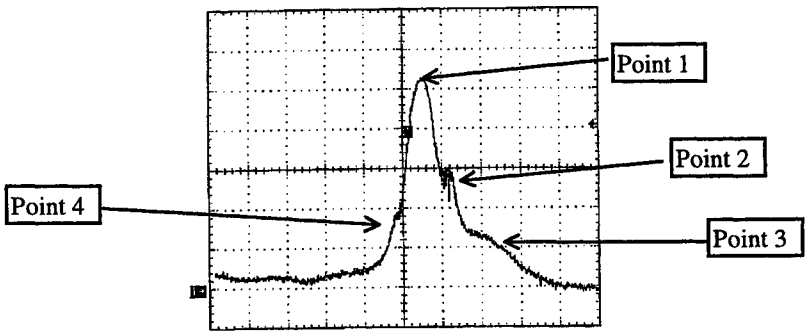


Figure 5: Output Waveform

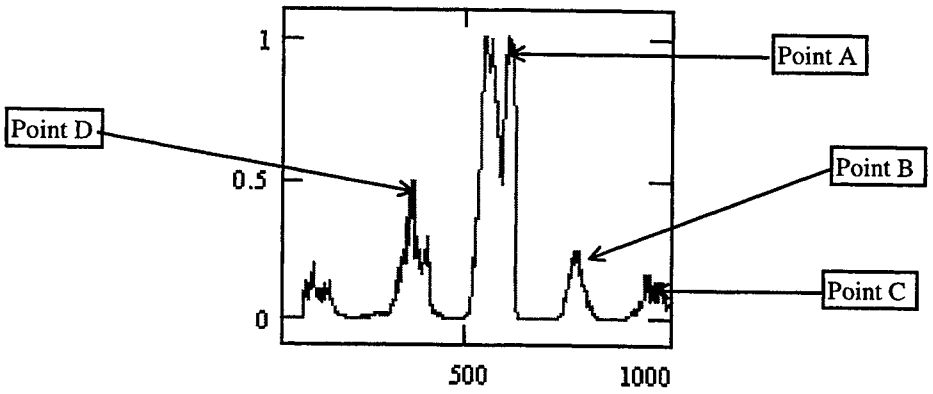


Figure 6: Simulated Sensor Output

6. CONCLUSION

Until now, there has not been a true solution solely dedicated for non-destructive defect detection for enamelled copper wire. The new non-destructive defect sensor head, has been developed for the purpose of defect detection on enamelled copper wire. The defect sensing system has been designed keeping in mind that copper wire may be manufactured as cylindrical, rectangular or triangular form. Copper wire may be drawn to very fine diameters, such as 19 micrometers, or as large as 50 mm diameter rods. Production speeds may be as high as 1000 meters per minute. Extensive trials at an enamelled copper wire manufacturing company have been very successful. The sensor head has been trialed under various conditions. Defects as small as 30 microns have been detected at operating speeds up to 10 meters a second.

In the authors opinion, a true and generic defect detection system is a system which operates on-line, covers the full circumference of the copper wire, handles various copper wire diameters, handles various production speeds, tolerant to wire vibration and resistant to external interferences. Such a defect detection system should also be intelligent and provide the capability of defect classification using a trained neural network. Such a system has been developed and was presented to the world at the 1998 Wire and Cable Fair in Duesseldorf Germany.

The reality is that the combination of scrap costs, downtime rework and inspection of one incident can be in excess of $20,000 US. There is also the cost of sending an inferior product to the customer, which could be very substantial. Surface inspection technology is fast becoming a powerful tool in management and control of continuous manufacturing processes.

There are several benefits for continuous on-line defect detection for enamelled copper wire industry. This includes real time production quality feedback, significant reduction of product wastage and hence reducing the overall production cost. And more importantly, better quality means greater customer confidence and satisfaction.

REFERENCES

1. L.R. Baker, "Inspection Of Surface Flaws By Comparator Microscopy" Applied Optics, Vol. 27, No. 22, pp. 4620-4, 15 November 1988.
2. H. Abdul Hamid, F. Brennan, W Wlodarski, "Defect Detection Systems for Enamelled Copper Wire, Cable and Optical Fibre : State of the Art and Current Trends" IEEE, 1999. (In Print)
3. E. Yamanaka, H. Kondoh, "New Defect-Detection System For A Rod And Bar Mill" Wire Journal International, pp. 54-9, August 1993.
4. G. Wei, C. Jianxing, "The Principle And Application Of A New Technique For Detecting Wire Rope Defects" Proceedings of The IEEE International Conference on Industrial Technology, 1996.
5. A. J. McGrail, D.W. Auckland, "Detecting and Classifying Flaws within Insulating Materials using Ultrasound" IEEE International Symposium on Electrical Insulation, pp. 2-4, June 5-8, 1994.
6. Z. Ding, J. Huang "The Surface Roughness Measurement of Metals Using Microwave" IEEE MTT-S Digest, pp. 253-256, January 1993.
7. M. Vascotto, "High Speed Surface Defect Identification On Steel Strip" MPT International, Vol. 19(4), pp. 70-73, August 1996.
8. C. Schenk, H. P. Müller, B. Maier, "Inspection Of Steel Strip And Advanced Defect Classification" MPT International, Vol. 4, pp. 86-92, May 1995.
9. N. Arraje, E. Evancich, "Diameter Gauges And Control Systems For Wire Extrusion Lines" Wire Journal International, pp. 90-6, August 1994.
10. K. Goszyk, "Eliminating High Speed Wire Vibration In On-Line Diameter Measurement" Wire Journal International, pp. 60-4, November 1993.
11. K.E. Torrance and E. M. Sparrow "Theory for Off-specular Reflection from Roughened Surfaces" Journal of the Optical Society of America, Vol 57, pp.1105-114, Sep. 1967.

Detection of cracks in aluminum beams using fiber optic polarimetric sensor

Jianjun Ma Anand Asundi

School of Mechanical and Production Engineering
Nanyang Technological University
50 Nanyang Avenue, Singapore 639798

Abstract

On line monitoring cracks development of the aircrafts in active service is of great significance for flight safety. Fiber optic sensor has its unique advantages over conventional sensors in this aspect. For modern aircrafts, many critical structures are made of aluminum alloys. When structural damage occurs in structures due to cracking, it reduces the stiffness. This change is reflected in the structure through changes in its dynamic characteristics such as natural frequencies, modal damping. In this paper, the ability of surface mounted fiber optic polarimetric sensors on aluminum vibrating cantilever specimen to detect cracks is studied. Different sources such as He-Ne laser, LD for fiber communication, laser pointer with longer, shorter and very short interference lengths are used in our experiments to evaluate their ability for dynamic test, respectively. The preliminary experimental results show that the measured main frequencies will be different for various number of cracks and depths. The three sources, the drift of light intensity and the change of interference contrast will not affect the measured main frequency values. Some of theoretical analysis is also given. From these results, surface mounted fiber optic polarimetric sensors is expected to find its application in monitoring structural integrity of structural members of aircraft in active service.

Key words: fiber optic sensor, vibration, cantilever, polarization

1. INTRODUCTION

The investigation of structural elements and their dynamical characteristics is of great insterest as it provide information on the damage incurred. This is vital for on line structural monitoring. Cracks in structures change the local flexibility and thus dynamical characteristics. The stiffness of a structure depends on the localization of damage and its magnitude, and is different for each natural mode of vibration[1]. Measurements of amplitudes, natural frequencies and vibration modes are used in the identification of cracks (their magnitude and location).

The most commonly used method for the measurement of vibration is through piezoelectric vibration transducers which are also called accelerometers. Fiber optic sensors are developing into an internal part of smart structures. The primary function for these integrated fibers is to sense various types of loadings continuously, and provide ability for condition monitoring during the operation. Especially, when compared with traditional sensors, fiber optic sensors can be surface mounted or embedded within structures with the unique advantages such as lightweight, flexibility and immunity from electromagnetic field interference.

Many critical structural members of aircrafts in active service are made of aluminum alloys. On line structural integrity monitoring of these members are important. Obviously fiber optic sensors are the best candidates for this issue. To reduce the complexity of the sensor system it may be necessary to limit an amount of integrated optical fiber. Fiber optic polarimetric sensor has the self-referencing capability of the optical signal so only one fiber is sufficient to determine one external parameter compared to systems composed of standard non-polarization-maintaining fibers, where both a sensor fiber and totally isolated reference fiber is required.

Part of the SPIE Conference on Advanced Photonic Sensors and Applications
Singapore • November–December 1999
SPIE Vol. 3897 • 0277-786X/99/$10.00

In this paper, experimental results for fiber optic polarimetric sensor for crack sensing through vibration test of aluminum beam is demonstrated. FFT spectrum identify the main frequency of the beam and has the ability for crack sensing.

2. THE SENSOR PRINCIPLE

High birefringent (HiBi)fibers which have a stress applying part(SAPs) such as a bow-tie fiber shown in Fig.1. An intrinsic tensile stress is applied to the core. This stress is largest in the different axis direction, which is the slow axis of the birefringence. The birefringence $\Delta\beta$ is defined as the propagation constant difference between two orthogonally polarized modes which travel along fast and slow axis, respectively, that is $\Delta\beta=\beta_x - \beta_y$. The phase change introduced by this relative birefringence over a length L is

$$\phi = \Delta\beta L \tag{1}$$

The polarimetric fiber sensor sensitivity to strain is:

$$\Lambda_\varepsilon = \frac{d\phi}{d\varepsilon} \tag{2}$$

The intensity I at the output end of the fiber is of the form

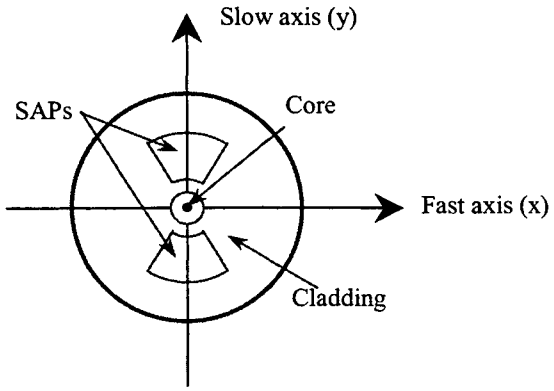


Fig.1. Cross-section of Bow-tie PM fiber

$$I = \tfrac{1}{2}(1+\cos\phi) = \tfrac{1}{2}[1+\cos(\Delta\beta L)] \tag{3}$$

So change in fiber length or birefringence causes the signal to modulate the output.

Counter a cantilever beams under viscously damped vibration. For a linear system with one degree of freedom the equation of motion is of the form[2,3]

$$m\ddot{W} + C\dot{W} + KW = F(t) \tag{4}$$

where W is the beam displacement, m is the mass of the object, C is the material damping coefficient, K is the stiffness and F(t) is the excitation force.

For the free underdamped case, under initial conditions $W=W_0$, and $\dot{W} = 0$ at t=0, it can be shown that:

$$W = \Omega_0 e^{-\zeta\omega_n t}\cos(\omega_d t - \varphi) \tag{5}$$

where

$$\omega_d = \sqrt{1-\zeta^2}\,\omega_n \tag{6}$$

$$\zeta = \frac{C}{C_c} \tag{7}$$

and

$$C_c = 2m\omega_n^2 \tag{8}$$

ζ is called the non-dimensional damping ration, C_c is the critical damping coefficient, Ω_0 and φ are constants determined from initial conditions, ω_d is the frequency of damped oscillation, and ω_n is the natural frequency which will related to the beam stiffness. Therefore, by observing the beam displacement or equivalently the strain response [see Eq. (2)]of the sensor, as a function of time, vibration monitoring is possible. Processing the time-domain data by Fast Fourier Transform (FFT) provides the frequency spectrum of the oscillation and thus the natural frequency peak can be observed. In a free damped cantilever beam usually only the first mode of vibration is observed, consequently the natural frequency peak would be predominant in the frequency spectrum.

3. EXPERIMENTS

Based on the analysis above, the experimental setup for dynamic test of cracked beams is shown in Fig. 2.

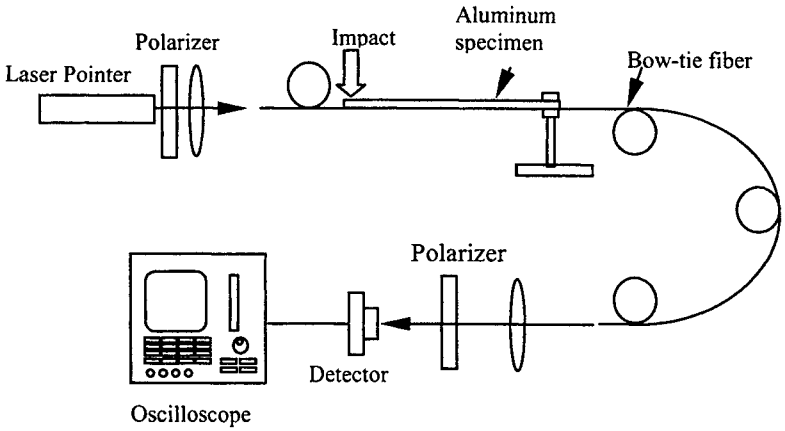


Fig.2. Experimental setup of fiber optic polarimetric sensor for crack sensing through dynamic test

Fig.3 shows the different cracked beams for the dynamic test.

For this experiment, light from a linearly polarized He-Ne laser with λ=633nm was launched into the fiber through a $\lambda/2$ waveplate and lens. The $\lambda/2$ waveplate was used to rotate the state of the input beam such that it made an angle of 45° or 135° with the fiber eigenaxes, equally populating both eigenaxes. Light output from the fiber was collimated and then passed through a analyser and send into a photodetector. A Oscilloscope was used to observe the damped vibration and give the results of FFT analysis. Fig.4 illustrates the typical underdamped oscillatory motion for uncracked beam and the its Fast Fourier Transform (FFT) analysis result.

For the different beams, Table 1 shows their main frequency for different crack numbers. As expected, because of the continuous flexibility reduction caused by cracks, the main frequencies are also decreased continuously. Table 2 shows the effects of different crack lengths. We find that the main frequency decreases rapidly with increase in crack length, though the crack number is less. This means the beam flexibility reduction strongly depends on the crack length.

4. DISCUSSION

For the interferometers, generally a monochromatic light is needed in order to produce interference pattern. The requirements for the interference length depend on the optical path difference (OPD) between the two interference lights.

For polarimetric sensor, the needed OPD for two orthogonally polarized lights is very small, so for the source with poorer interference length it is possible to observe the polarization interference. For the dynamic test of the aluminum cantilever beams, we use either a He-Ne laser, or a Laser Diode or even a Laser Pointer to do all of the experiments. All sources showed good results for both time-domain damped vibration signal and frequency-domain FFT analysis. Typically, for a laser pointer source, Fig.5. has shown a result for the beam with no cracks which is the same as that in Fig. 4. This result is useful for the commercial consideration of the sensor system.

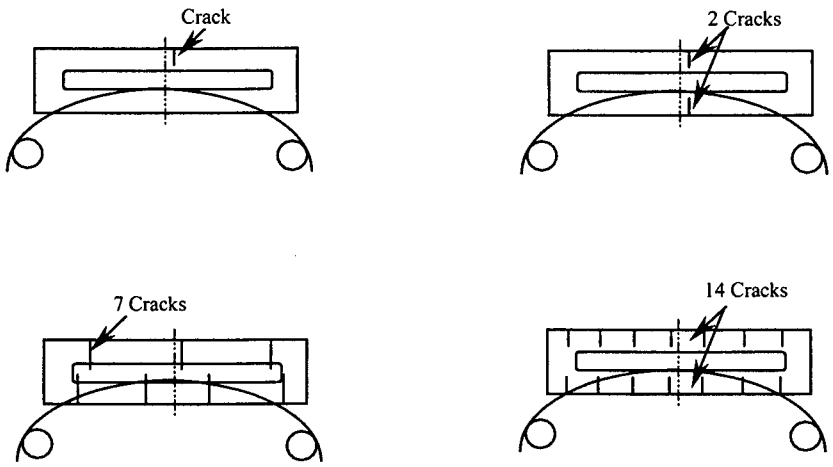


Fig.3. Different cracked beams.

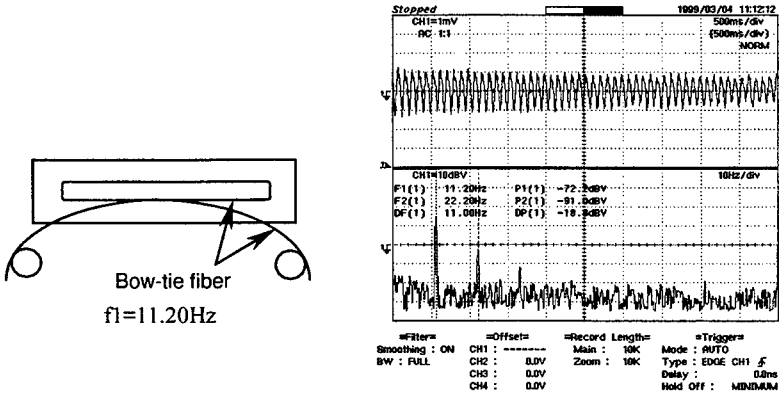


Fig. 4. The typical experimental result for the beam with no cracks

Table 1 Effect of crack number on the first frquency of beams

Crack Numbers	f1(Hz)
0	11.20
1(30mm)	11.00
2(30mm each)	10.40
7(40mm each)	8.600
14(20mm each)	10.20

Table 2 Effect of crack length on the first frquency of beams

Crack Numbers	f1(Hz)
0	11.20
7(40mm each)	8.600
14(20mm each)	10.20

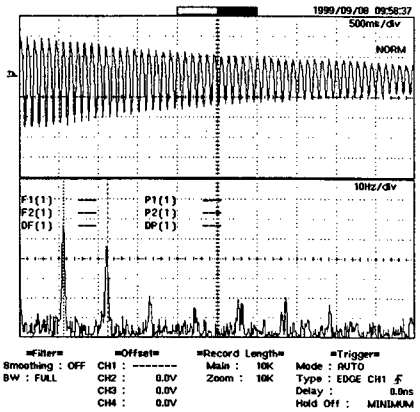


Fig.5. Experimental result for the uncracked beam using laser pointer source.

5. CONCLUSION

For aluminum beams, we have used fiber optic polarimetric sensor to detect the main frequency change caused by the cracks under damping loading. The experimental results show dynamic FOPS has its ability for on line monitoring of the crack development. A low cost light source can be used for this sensor which will be important for the sensor commercialization.

ACKNOWLEDGEMENT

The support of NTU/MOE through grant MLC1/97 is acknowledged.

References

1. W. M. Ostachowicz and M. Krawczuk, "Vibration analysis of cracked beam", Computers & Structures Vol. 36, No. 2, pp245-250, 1990.
2. Edward Tapanes, "Real-time structural integrity monitoring using a passive quadrature demodulated, localised Michelson optical fiber interferometer capable of simultaneous strain and acoustic emission sensing", SPIE Vol. 1588 Fiber Optic Smart Structures and Skins IV,1991, pp356-367.
3. W. T. Thomson, Theory of vibration with applications, second edition, Prentice-Hall, New Jersey, USA, 1981.

Spectrally resolved white light interferometry for profilometry with polarization phase shifter

S. Suja Helen, M. P. Kothiyal and R. S. Sirohi
Applied Optics Laboratory, Department of Physics
Indian Institute of Technology, Madras
Chennai – 600 036, INDIA

ABSTRACT

We describe a spectrally resolved white light interferometer with polarization phase shifter for use in surface profiling. Phase shifting is introduced by a rotating half-wave plate. The phase shifted intensity values needed for the phase calculation at each pixel are obtained from the same pixel instead of different pixels, thereby avoiding error due to variation in sensitivities of different pixels.

Keywords : Spectral interferometry, White light interferometry, Achromatic phase shifting, Interferometric metrology

1. INTRODUCTION

White Light Interferometry (WLI) has been introduced for the surface analysis of reflecting surfaces having height variation greater than $\lambda/2$, which introduces phase ambiguity problem in the conventional laser interferometry. In WLI, the visibility of the fringes is directly related to the optical path difference between the two interfering beams. The surface profile is measured by scanning the optical path difference (OPD) and locating the maximum of the fringe contrast position. In this method, a large number of white light interferograms need to be recorded and processed to obtain the surface profile. To avoid the OPD scanning, spectrally resolved WLI has been proposed for surface profiling [1-2]. In spectrally resolved white light interferometry, the height measurements are based on the phase shifts due to wavelength variation along the chromaticity axis. Phase shifting algorithms have been used for the determination of phase in spectrally resolved WLI [1]. However the intensity data used in the algorithm is obtained from different pixels. This may result in error due to variation in sensitivities of different pixels. In this paper, we propose achromatic phase shifting based on polarization components so that all the intensity data needed for the calculation of phase at any pixel are obtained from the same pixel thereby avoiding the error introduced due to variation of sensitivity mentioned above.

Part of the SPIE Conference on Advanced Photonic Sensors and Applications
Singapore • November–December 1999
SPIE Vol. 3897 • 0277-786X/99/$10.00

2. PRINCIPLE

When a two-beam interferometer is illuminated with white light, the output intensity of the interference pattern is the superposition of all monochromatic constituents of the spectrum and can be expressed as[3]

$$I(z,\bar{\sigma}) = I_{ref} + I_{surf} + 2(I_{ref} I_{surf})^{1/2} g(2z)\cos(4\pi\bar{\sigma}z) \quad (1)$$

where I_{ref} and I_{surf} are the intensities from the reference and test arms of the interferometer respectively; $\bar{\sigma}$ is the mean wave number of the light source, and 2z is the round trip OPD. The phase term $\phi(z)=4\pi\bar{\sigma}z$ gives information about OPD and is used for profilometry. In spectrally resolved white light interferometry, the spectrum of the interference pattern is analyzed to obtain the phase information. When the interference pattern is passed through a spectrometer, the exit plane of the spectrometer gives the phase variation as a function of the wave number if scanned along a direction which is perpendicular to the entrance slit. The spectrometer transforms the white light interferogram into a set of N monochromatic intensities. These intensities form a spectrum modulated by a cosine wave as a function of σ. This number N is governed by the number of CCD pixels available along the direction perpendicular to the entrance slit in the exit plane of the spectrometer. The intensity data at the pixels can be converted into phase data at each pixel.

The variation in phase $\phi(\sigma)$ with σ for a particular z-value is given by the relation

$$\phi(\sigma)=4\pi\sigma z. \quad (2)$$

Equation (2) shows that for a given point on the object $\phi(\sigma)$ varies linearly with σ. If we can determine $\phi(\sigma)$ for different σ at any object point , ie a given z value, then z can be determined from the slope calculation

$$z = \frac{\Delta\phi(\sigma)}{4\pi\Delta\sigma} \quad (3)$$

Each pixel receives light of a different wave number σ. At any pixel the intensity for a particular wave number can be written as

$$I(\sigma) = I_0\left[1+V\cos(\phi(\sigma)+\phi_0)\right] \quad (4)$$

where $I_0 = I_{ref} + I_{surf}$ and $V = \frac{\sqrt{I_{ref} I_{surf}}}{I_{ref} + I_{surf}}$. ϕ_0 is a controllable phase shift. $\phi(\sigma)$ can be calculated from this equation using the phase shifting technique. The problem is to introduce the same phase shift simultaneously at all the pixels, receiving light of different σ. We have done this by introducing nearly achromatic phase shift using a polarization phase shifter. From the four $\pi/2$ phase shifted intensity values of each spectral component the phase $\phi(\sigma)$ is calculated for each σ using the 4-point phase shifting algorithm[4]

$$\phi(\sigma) = \tan^{-1}\left(\frac{I_2 - I_4}{I_1 - I_3}\right) \quad (5)$$

where I_1, I_2, I_3 and I_4 are the four $\pi/2$ phase shifted intensity values.

3. SPECTROSCOPIC WHITE LIGHT INTERFEROMETER WITH POLARIZATION PHASE SHIFTER

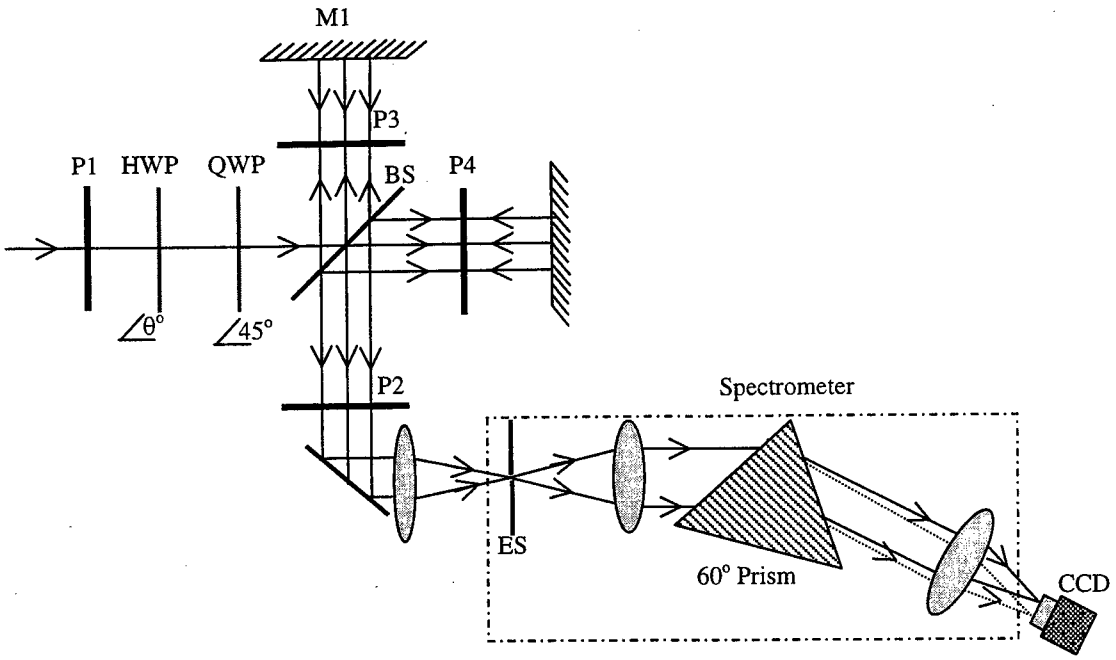


Figure 1 Schematic diagram of the spectrally resolved white light interferometer with polarization phase shifter

The schematic diagram of the white light interferometer with polarization phase shifter at the input end is shown in figure 1. The polarization phase shifter consists of a rotating half wave plate (HWP) followed by a quarter wave plate (QWP) fixed at an azimuth of 45°. This phase shifter is nearly achromatic[5]. The phase shift introduced by this phase shifter between the two interfering beams is nearly same for the wavelength range of 450-700 nm. The light transmitted by the QWP is incident on the beam splitter BS. The BS splits the incident beam and sends it to test and reference arms of the interferometer. The polarizers P3 and P4 are kept orthogonal to each other in the two arms of the interferometer. So the beams transmitted by the polarizers P3 and P4 are orthogonally polarized. These two beams are reflected back by the test and reference surfaces of the interferometer and sent towards the polarizer P2 by BS. The polarizer P2 brings them to interfere. This interference pattern is spectrally dispersed by a spectrometer and the intensity of the each spectroscopic component is recorded for phase calculation. The phase shifter introduces the necessary phase shift. The phase shifts introduced between the two interfering beams for different rotation of the HWP as a function of σ is given in Table 1.

Table 1

Phase difference between the two interfering beams as a function of σ

Wavenumber $(\mu m)^{-1}$	Phase difference ϕ_0 between the two interfering beams for different θ settings of the HWP			
	θ=0°	θ=22.5°	θ=45°	θ=67.5°
2.22	90.0	16.7	-90.0	-196.7
2.00	90.0	3.4	-90.0	-183.4
1.82	90.0	0.0	-90.0	-180.0
1.67	90.0	2.4	-90.0	-182.4
1.54	90.0	8.1	-90.0	-188.1
1.43	90.0	15.6	-90.0	-195.6

For the orientations 0°, 22.5°, 45° and 67.5° of the HWP the phase differences introduced are 90°, 0°, -90° and -180° respectively for the mean wave number $\bar{\sigma}$ (=1.82μm^{-1}). The phase difference varies with σ but can be considered nearly achromatic

In order to study the usefulness of this phase shifter in this application, we have used the values of the phase shifts from Table.1 in equation (4). From the four equations obtained at each σ, we have then calculated the phase $\phi(\sigma)$ at each σ using equation (5) in which z is assumed to be 5μm. The variation of $\phi(\sigma)$ with σ is practically linear as shown in figure (2). The figure also shows the unwrapped 2π modulo phase steps obtained from equation (5). The linearity of $\phi(\sigma)$ vs σ curve proves that the phase shifting can be assumed to be achromatic over the σ range considered.

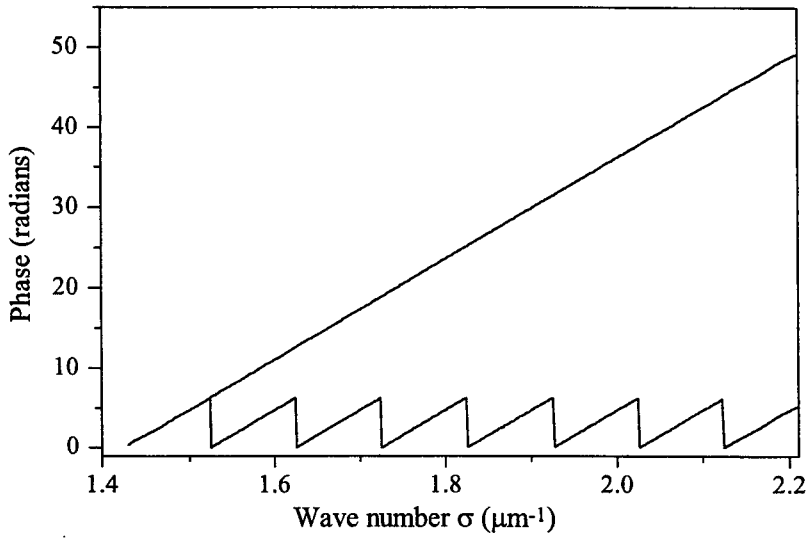


Figure 2 Variation of the calculated value of phase with wavenumber σ using phase shifting algorithm of equation (5) and phase values from Table 1

4. EXPERIMENTAL DEATAILS

The schematic diagram of the white light interferometer with polarization phase shifter at the input end of the interferometer is shown in figure. 1. The interferometer is illuminated with collimated white light. The test surface is imaged on to the entrance slit (ES) of the spectrometer which selects a line of the test surface which can be profiled. The spectrometer consists of a 60° prism in an arrangement as shown in figure.1. The entrance slit is imaged on a CCD detector. The axis perpendicular to the entrance slit represents the chromaticity axis. This means that the pixels on the chromaticity axis record the intensities of the interference pattern from a given point of the object (corresponding to a particular value of z) in different σ values as per equation (4). The σ value at each pixel is known in advance by a suitable calibration procedure. The intensity values on these pixels can be used to obtain $\phi(\sigma)$ vs σ relationship and obtain the z value at one point of the object. The pixels perpendicular to the chromaticity axis have information on the z-values of the neighbouring object points giving the line profile.

The HWP of the phase shifter is oriented at 0°, 22.5°, 45° and 67.5° to introduce the phase shift. For each phase shift the interferogram is grabbed. One line of the detected spectral intensity on the chromaticity axis for different rotation of the HWP is shown in figure (3). From these four phase shifted intensity values, the phase is calculated for each σ. The measured phase along with its linear fit shown in figure (4). The slope of this curve gives the absolute height of the object at that point which in the present case is 2.52μm.

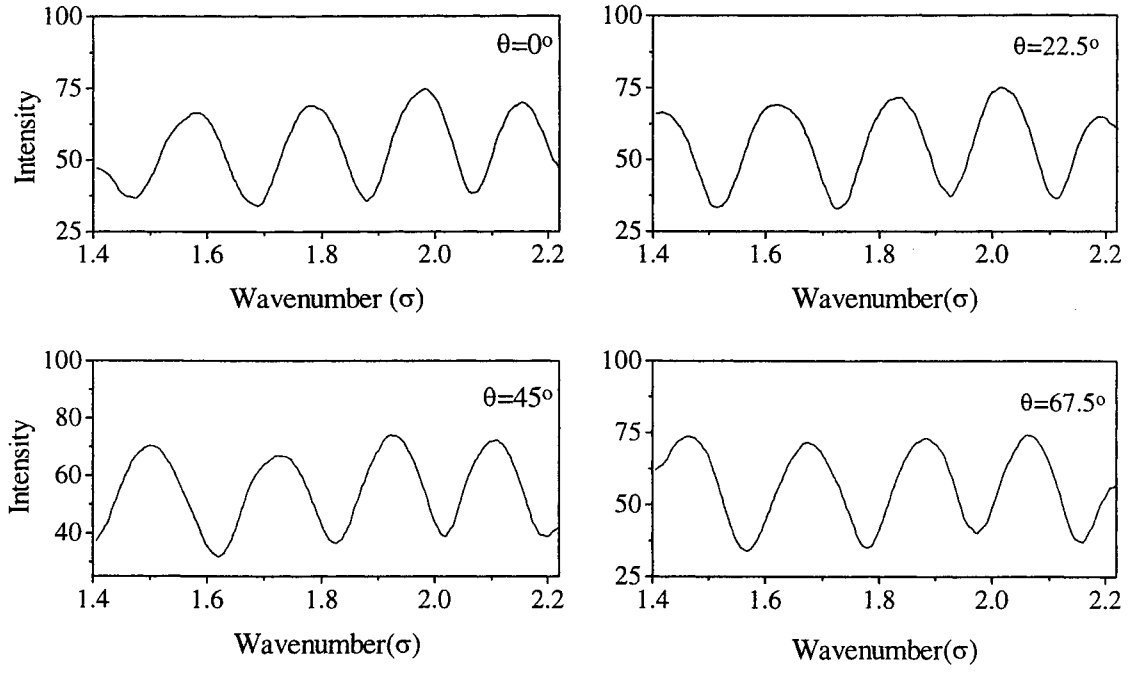


Figure 3 Intensity variation along the chromaticity axis for different rotation of the HWP θ

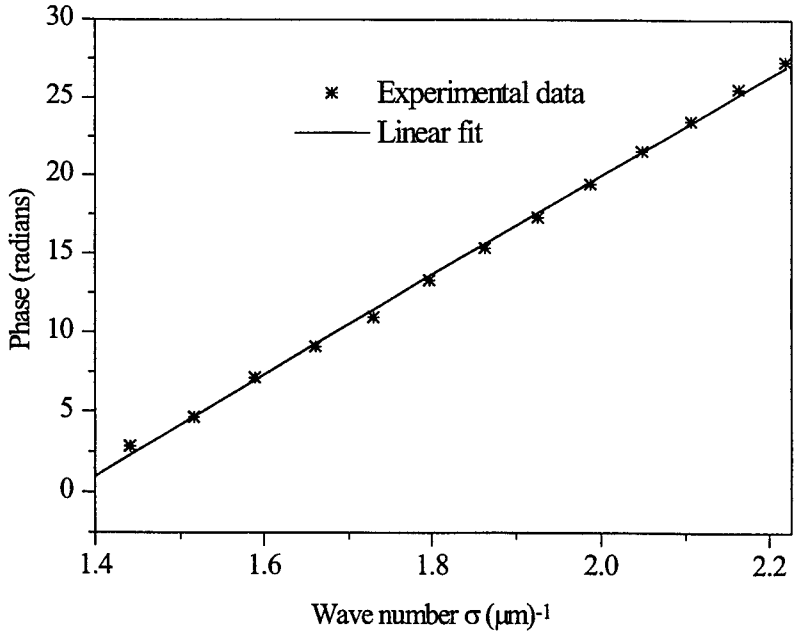


Figure 4 Experimentally determined phase variation along the chromaticity axis along with the linear fit.

5. CONCLUSION

We have demonstrated the use of polarization phase shifting method in spectrally resolved white light interferometry for surface profiling. Practically linear relationship between measured phase $\phi(\sigma)$ and σ indicates the nearly achromatic behaviour of the phase shifter. The main advantage of the proposed method is that all the phase shifted intensity values are obtained from the same pixel.

REFERENCES

1. P. Sandoz, G. Tribillon and H. Perrin, "High resolution profilometry by using phase calculation algorithms for spectroscopic analysis of white light interferograms", J. Mod. Opt.,vol.**43**, pp701-708, (1996)
2. J. Calatron, A. L. Guerrero, C. Sainz and R. Escalona, "Spectrally resolved white light interferometry as a profilometry tool", Opt. Laser Technol., vol.**28**, pp485-489, (1996)
3. S. Suja Helen, M. P. Kothiyal and R. S. Sirohi, "Phase shifting by a rotating polarizer in white light interferometry for surface profiling" J. Mod. Opt., vol.**46**, pp993-1001, (1999)
4. D. Malacara, *Optical Shop Testing,* pp501-515, John Wiley & Sons, New York, (1992)
5. S. Suja Helen, M. P. Kothiyal and R. S. Sirohi, "White light interferometry for surface profiling with phase shifter at the input end" (Communicated)

A Kind of Method of Measurement With High-Speed External Diameter Outline

National Expert of Ditinguished Contribution. Senior Engineer Professor Baoying Liu[*a,b],
Senior Engineer Professor Bingsheng Chen[b], Kun Shi[b]
[a] Tianjin Institute of Laser Technology
[b] Tianjin O.E. TEC. CO.LTD
Keyan West Road No. 6, Nankai District, Tianjin 300192, PR CHINA

ABSTRACT

Defect of external outlines of wire rods which running at speed of 100 m/s need be measured in process of production which high-speed wire rods in iron and steel industry, At present, it adopts mainly method of rotation which measures section shape. This article will introduce a kind of disposal plan which high-speed camera and picture by multi-detector. The article will analyze that number of calculation detectors and range of measurement which defect of outline, and advantages and disadvantages which in comparison rotation plan with multi-detector plan. It proved the accuracy of the analysis by simulated test.

Key Words: measurement, diameter-measurement, CCD, photonics

1. INTRODUCTION

With technological development of metallurgical industry, many companies have set up production lines of high-speed wire rod. To ensure quality of products, it is important that measure diameter which high-speed wire rod on these production lines. Now it has advantage what is can 360°whole-section measurement that most of this kind of devices which measure diameter takes single-head of measurement or double-heads of measurement spinning measurement as the dominant factor. But it only can measure spiral section with the wire rod high-speed forward. Higher-speed wire rod forward, longer the thread pitch of spiral. If shut down the thread pitch of measurement, we must speed up spinning, For instance, rotational speed of head of measurement is 200 revolutions per minute (i.e. 3.3 revolutions per second); speed of wire rod is 100 m/s which is measured, so separation distance of measurement is 15 m.

It is very sensitive and it has improved true time and reliability except a few of "fade zone" which use multi-heads of

measurement (e.g. eight heads of measurement) at the same section to measure, and its separation distance is 0.15 m.

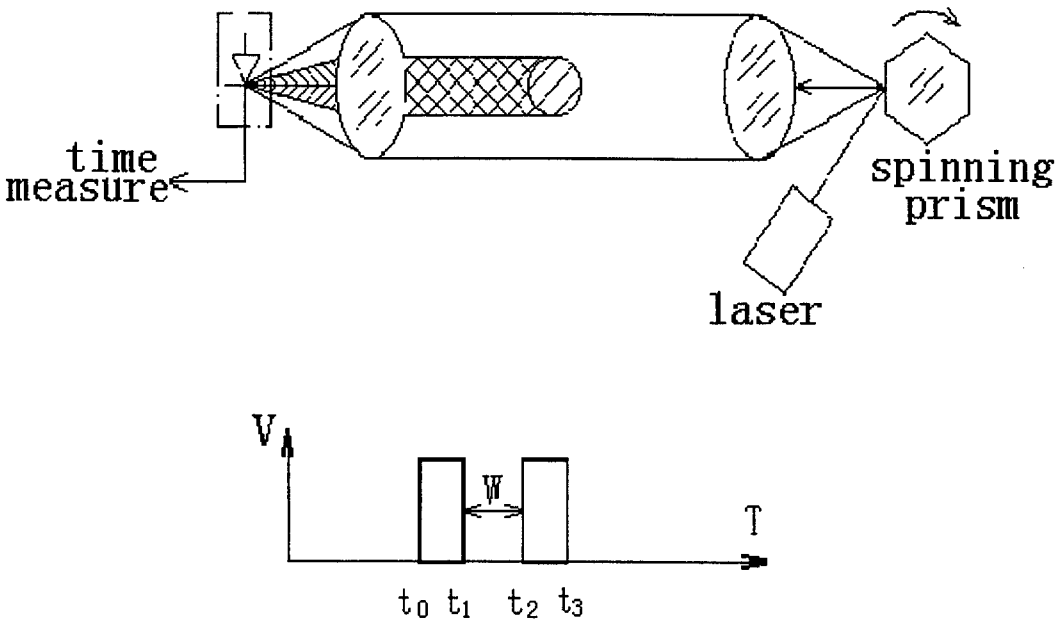
The key special subjects of this measurement program as follows: 1. Use line array CCD to detect high-speed moving object; 2. Confirm the quantity of head of measurement according to the range which defect detection of outline; 3. Use mathematical model to improve precision of measure diameter. The major points of this article as follows: analyze and calculate the quantity of head of measurement and the range which defect detection of outline; compare the advantages and disadvantages of the programs which spinning with multi-heads of measurement; and offer results of simulated test.

2. SELECTION OF PRACTICAL PROGRAM

2.1 Selection of Basic Program Which Measure Diameter

At present the technological programs which non-contact measure external diameter mostly includes two kinds: laser scan and photoelectric camera. Their characteristics as follows:

2.1.1 Laser Scan



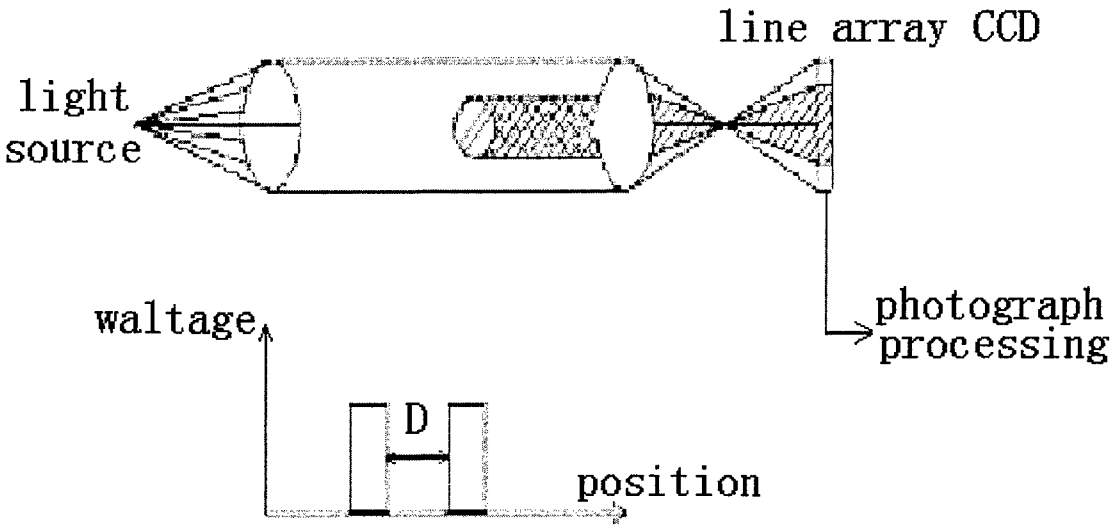
Picture 1: principle of laser scan

The spinning prism at the focus on launch lens, which is drove by motor, reflects laser beam to launch lens to become a parallel band of scan. It can calculate the diameter (or width) of external diameter that sheltered time multiply scanning speed of the parallel laser beam when the diameter (or width) of the measured object shelter the parallel scanning laser beam on the

laser beam. The advantage of this method is what static precision can reach order of magnitude of μm and bulk can make very small. But the major disadvantage is dispatch dynamic characteristic. Because the highest speed of this kind of mechanical scan, which depend on electrical machine spinning, is about 300 times per second around the world, but the average speed is only tens times per second. To production of high-speed wire rod, it difficulty meets demand of measure error what it will bring superior error when the measured object horizontally shake. Additionally, wear and tear of axis will lead to change of reflecting laser beam because spinning lens is spinning in long term. So some spin measurement diameter need to replace cores parts—scan device when it continues to work for two years.

2.1.2 Method of Photoelectric Camera

It's fundamental as picture 2.



Picture 2: principle of photoelectric camera

A band of parallel light, which is become by optics system, become electronic image on line array CCD over receive lens. It can calculate external size of corresponding object which shelter the above-mentioned parallel light beams through image processing. The advantage of this measure method is what it has no mechanical motional part and no error which wear and tear lead to. It is a shadow what is filmed by electronic camera when object shelters the lights beam at the same time. So it's error of horizontal shake is more less-than that of laser scan and the speed of film can reach one thousand sizes per second or even faster. It is advantageous to measure size of external outline that high-speed motional object which we use special film

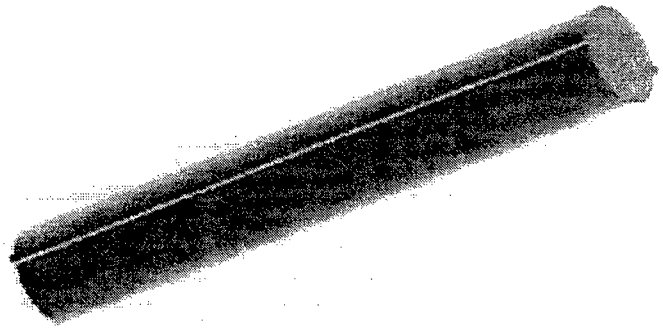
technology to lead to film time in one ten thousandth second or even faster.

2.2 Analysis of Measured Object Which High-Speed Wire Rod

The measured object is round steels of φ5-φ18, motional speed is 65-120 m/s, and no rotation.

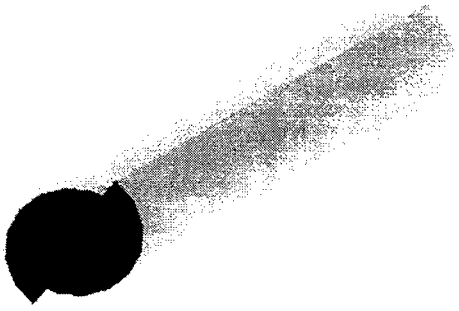
It need detection and alarm in 20 seconds after products have defects.

The first kind of defect as the picture (or called ear) is one of major defects. So it must be detected.



Picture 3

The second kind of defect as the picture (or called offset).



Picture 4

2.3 Selection of Multi-Group of Head of Measurement

Because speed of wire rod is greater than 63 m/s regardless of shake or rotation, it's equivalent that scanning rotate speed of one round per second and space is 63 m. The space of every sampling cycle should be 0.126 m if sampling cycle of CCD is 500 Hz.

If use eight groups of head of measurement at the same section, 180°/ 8 = 22.5°; if use twelve groups of head of

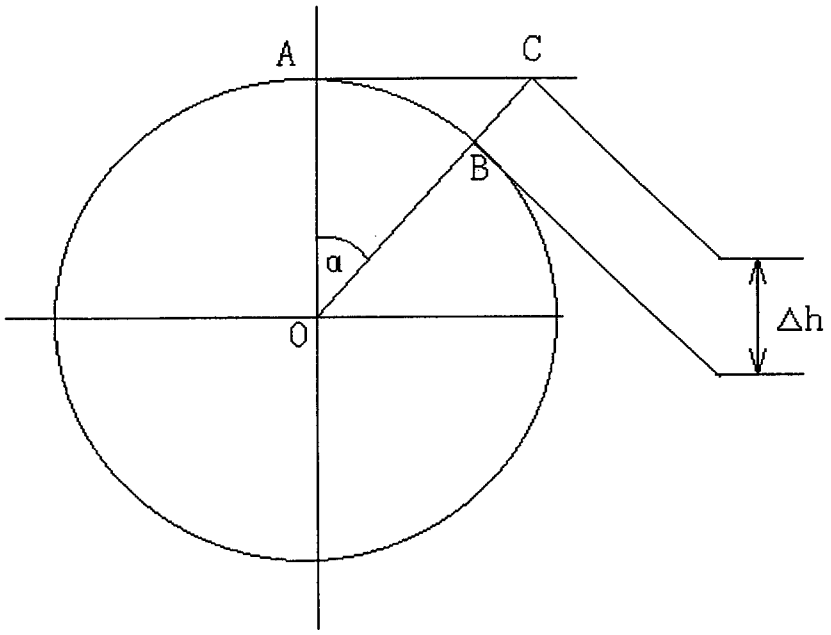
measurement, 180°/ 12 = 15°.

It is said that someone make it what nine heads of measurement, space between angles is 20°and no rotation in international.

Table 1

Quantity of head of measurement	group	2	4	8	12
Space between angles	degrees	90	45	22.5	15
Space which circumference of φ5	mm	3.927	1.964	0.982	0.655
Space which circumference of φ18	mm	14.14	7.069	3.534	2.356
measured Δh (Δh/r)	%		0.0824	0.0196	0.00863

So the effect what selection of 8-12 groups of heads of measurement can realize measurement of all diameters at the same section is better than real time of spinning. To try to explain detection effect of heave (ear) by picture as follows.



Picture 5

According to the above-mentioned picture: $\Delta h = -R * (1- 1/ \cos.\alpha) = R * (1- \cos.\alpha)$

So we can obtain data as follows:

Table 2

Quantity of head of measurement	Group	4	8	12
Space between angles	degrees	45	22.5	15
Δh/R	%	8.240	1.960	0.863
Practical measure Δh(φ5)	mm	0.206	0.049	0.022
Practical measure Δh(φ18)	mm	0.742	0.176	0.078
When h=1mm, Δh (φ5)	mm	0.007	0.932	0.969
When h=1mm,Δh(φ18)	mm	0.238	0.808	0.914

3. COMPARE DIAMETER MEASURERS WHICH SPINNING WITH MULTIi-HEADS OF MEASUREMENT

Spinning head of measurement can be applied to measure slow-speed forward measured object. It is more applicable that multi-heads of measurement be applied to measure high-speed wire rod. So we compared that as follows: to get lengthways displacement ΔL which a size of section image; space which measure every size of section image; measure track; reliability; maintenance and so forth.

To set technological parameters: speed of wire rod is 100 m/s; spinning measure use one head of measurement which is 200 revolutions per minute (i.e. 3.3 revolutions per second); space of readings is 0.3°(i.e.4000 times per second). Multi-heads of measurement select eight groups of head of measurement.

The result of comparison as follows:

Table 3

Items Kinds Of Diameter Measurers	Tine which measure one cycle (second)	Space which measure one cycle (m)	Dot track	Time displac-ement of filming (mm)	Find contin-nue heave or can't	Photo electr-ic shift	Mech-anical struct-ure	Conne-ction of electric signal	Data proc-essing	Reliability of principle
Spinning	0.15	15	Helical line of 0.3°	25	Can find	Samp-le	Comp-lex	Compl-ex	Sample	It has mechanical electric spinning machine, so reliability is low.
Eight groups Of head of measurement	0.015	0.15	16 share the same section	10	Can find wire rod which isϕ5,>0 .049mm; ϕ18,> 0.176 mm	Comp-lex	Not too comp-lex	Sample	Comp-lex	It has no mechanical electric spinning machine, so reliability is high.

According to the above-mentioned comparison, it is not only its function correspond to that of spinning program but also its real time is better, that is measure of the same section is only 0.15m. It don't like that of spinning diameter measurer which spiral forward is 15 m. Although sampling density in circumference of external diameter is less than that of spinning diameter measurer, It's ability of finding heave defect is stronger than that of spinning diameter measurer in 1.5 m space. But spinning diameter measurer can detect the heave defect in 15m space. Additionally, it's reliability and maintenance is stronger than that of spinning diameter measurer. It leads to corresponding increase which speed of microcomputer processing and capacity that photoelectric shift from one group of head of measurement to eight groups of head of measurement because mainly "difficulty shift". But solution of these problems simplified the trouble which spinning method leads to.

4. RESULT OF SIMULATED TEST

Measured data of threading steel stems from samples which Club Products Works of Tangshan Iron and Steel Company offer threading steel.

Measurement method: use diameter measurer to measure 3 times; use micrometer and calipers to measure it.

Measurement results as the sheet:

Table 4

	JDC-J diameter measurer	Micrometer	Calipers	Remarks
Internal diameter	11.208 11.221 11.230	— — —	11.18 11.20 11.22	Can't use micrometer
Endlong heave	14.316 14.317 14.318	14.316 14.314 14.318	14.32 14.32 14.32	
Sidelong heave	13.667 13.667 13.674	13.653 13.672 13.656	13.62 13.64 13.62	

5. CONCLUSION

It is a good measurement program, which it uses multi-heads of measurement to measure diameter of high-speed wire rod, ellipse oral cross section, defect of wire rod and characteristic size of threading steel, what is good real time, high reliability and good maintenance.

SUMMARY

With the speed of round steels rolling improving, product lines of high-speed wire rod rolling must have external diameter measurer on lines. At present it mostly use program of spinning head of measurement to check outline size of products in international. This method can measure outline size of spiral section. This thesis put forward the program which use 8-12 fixed heads of measurement (or even more heads of measurement) to measure external diameter of the same section at the same time. This article compares advantages and disadvantages of two programs as follows: obtain lengthways displacement of a size of section image; separation distance of every size of section image; measure track; size which can find heave; reliability; maintenance and so forth. It is more superiority that uses the program of multi-heads of measurement to the program of spinning measurement from current technological ability by comparison. Its advantage is what lengthways displacement is small and its separation distance of measure section is decuple smaller. It has no electric mechanical rotation, so its ability that real time, reliability and finding asynechia heave and defects is stronger. It may be a utility program that solves measurement of threading steels through proper sort processing the measured data. Because its measure displacement is very small (about 1 mm) and it can measure eight equidistace diameters at the same section. The simulated test proves that the above-mentioned analysis is right.

ACKNOWLEDGEMENT

Thanks our colleagues who have given us help in experiments and proofreading of the thesis. Additionally, thanks Qunli Chen and Jinwei Liu that have given us help in written course.

REFERENCES

1. Baoying Liu, "Company Standard of JDC-J (K) Series of Diameter Measurer and Width Measurer", Otc.1997.
2. Bingsheng Chen," Direction of Selection of JGC Series", No. 27 Institute of Electric Ministry.
3. Baoying Liu, "Apply Line Array CCD to Measure High-Speed Motional Object", *Optical Technology supplemtnt* pp.**87-88**, Sept. 1996.
4. Commercial of SIKORA, Zumbach, BETA Company etc.
5. Sarras G. Chamberlain, William D. WashkwaK, High-Speed, Low Noise, Fine Resolution TDI CCD Imageries , DALSA INT.
6. Baoying Liu, "Raising precision of Measuring diameter by Mathematical model" *Journal of test and measurement technigue* vol.**12** No.**3** pp.**370-375**, 1998.

1

[1] *correspondence: Email: chntjcbs@public.tpt.tj.cn; Telephons: 86 22 23686054;
Fax: 86 22 23686054[1]

220kV combined optical voltage and current transformer

Sunan Luo*, Miaoyuan Ye, Yong Zhu, Ying Cui, Yan Xu,
KaiCheng Li, Chaolong Ou, Zhiping Chen

Department of Electrical Engineering,
Huazhong University of Science and Technology
Hubei, Wuhan, 430074, China

ABSTRACT

This paper presents a new optical transformer which can measure high voltage (up to 220kV) and large current (up to 3000A) simultaneously. The measurement of voltage is based on the Pockels effect. A novel structure is proposed to realize an optical voltage sensor which has better stability. The measured high voltage is directly applied on the optical voltage sensor without using any capacitive voltage divider. The measurement of current is based on the Faraday effect. A bulk glass closed-loop structure for optical current sensor is used. The optical voltage sensor is in the middle of the HV polymeric insulator which is full of SF_6 gas, and the optical current sensor is on the top of the same insulator. This transformer offers many advantages such as immunity of electromagnetic interference, no saturation, no oil inside, wide band range, excellent transient characteristics and light weight etc. The principles, structures and testing results of the combined optical voltage and current transformer are described in this article.

Keywords: Combined Transformer, Optical Voltage Transformer, Optical Current Transformer, Voltage Sensor, Current Sensor, Pockels Effect, Faraday Effect, $Bi_4Ge_3O_{12}$ Crystal, Dense Flint Glass ZF-7, Polymeric Insulator

1. INTRODUCTION

It's very important for an electric power system to measure accurately the system parameters, in particular, current and voltage. Traditionally this has been achieved by expensive, bulky electromagnetic induced current transformers (CTs) and voltage transformers (VTs). With the development of electric power systems, traditional CTs and VTs become more and more unsuitable for modern power systems due to their cost, safety implications for personnel and surrounding installations if failure occurs, substation land requirements etc.

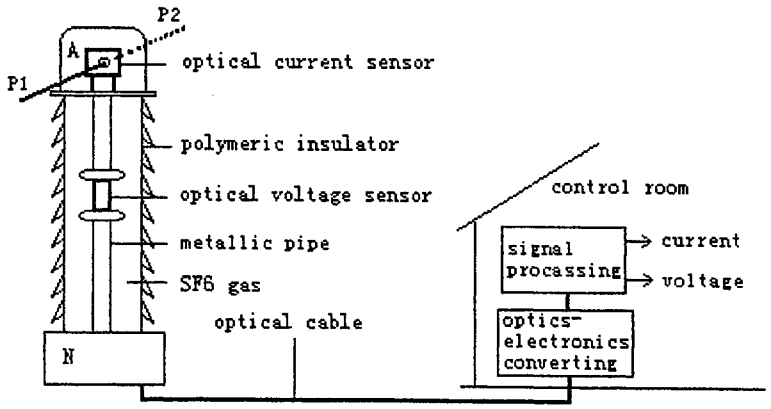


Fig. 1 Block Diagram of the COVCT

*Correspondence: Email: luosunan@263.net; Telephone: 86-27-87552532

Part of the SPIE Conference on Advanced Photonic Sensors and Applications
Singapore • November–December 1999
SPIE Vol. 3897 • 0277-786X/99/$10.00

Research effort into viable alternatives to CTs and VTs has been ongoing for many years. Optical current transformers (OCTs) and optical voltage transformers (OVTs) are viewed the most ideal CT and VT replacements[1,2]. OCTs and OVTs employ optical fibers to transit signals between the sensors near the high voltage buses and the control rooms. Using all optical devices and optical fibers, OCTs and OVTs offer many advantages such as reductions in size and weight, inherent electrical safety, immunity of electromagnetic interference, wide band range and excellent transient characteristics etc.

In this paper, a new transformer which can measure high voltage (rated value 220kV) and large current (rated value 3000A) simultaneously without interference each other is presented. And a novel structure to realize an optical voltage sensor which has better stability is proposed. Figure 1 shows the block diagram of the combined optical voltage and current transformer (COVCT).

2. PRINCIPLES OF THE COVCT

2. 1. Principle of the Optical Current Sensor

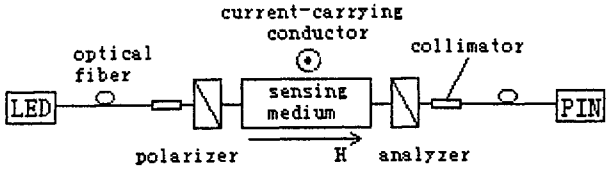


Fig. 2 Principles of the Optical Current Sensor

The optical current sensor operates based upon the Faraday effect. As shown in Fig.2, the polarizer turns a light coming from a LED into a linearly polarized light. Because of the presence of magnetic field generated by the measured current, the polarized plane of the linearly polarized light undergoes a rotation when the light which direction parallels the direction of the magnetic field passes through the sensing medium. The rotation θ_F can be expressed as,

$$\theta_F = VHl \tag{1}$$

where V: Verdet constant (rad/A)

H: magnetic field intensity (A/m)

l: length of light ray in the sensing medium (m)

If the interaction length *l* encircles the current-carrying conductor, equation (1) can be rewritten to express θ_F directly in terms of the measured conductor current *I*,

$$\theta_F = \oint \vec{H} \cdot d\vec{l} = VI \tag{2}$$

Thus the measured current can be perceived by measuring the angle of rotation θ_F.

The Faraday rotation θ_F can be measured by transforming the rotation to a modulation of the transmitted light intensity, by passing the light through an analyzer. The analyzer is orientated with its axis of transmission at 45° to the axis of transmission of the polarizer. As a consequence, the output light intensity of the analyzer can be expressed as,

$$P = P_0 \cos^2(45 + \theta_F) = \frac{1}{2} P_0 (1 + \sin 2\theta_F) \tag{3}$$

where P_0 is the output light intensity of the polarizer.

According to equation (3), the measured current can be obtained with the optic-electric converting and signal processing units in the control room.

2.2. Principle of the Optical Voltage Sensor

The optical voltage sensor operates based upon the Pockels effect which refers to the property of a material becoming birefringent when placed in an electric field. As shown in Fig.3, a linearly polarized light injecting into the sensing medium to which voltage V is applied is divided into two linearly polarized lights propagating in the same direction but at different speeds. As a result the output light of the sensing medium becomes an elliptically polarized light, and the phase difference

between the two linearly polarized lights can be expressed as,

$$\Delta\theta = \frac{2\pi}{\lambda} n_0^3 \gamma E l = \frac{2\pi}{\lambda} n_0^3 \gamma \frac{l}{d} V \tag{4}$$

where λ is the wavelength of the light used, n_0 is the refractive index of the sensing medium, γ is the linear electro-optic coefficient of the sensing medium, E is the electric field strength produced by the measured voltage V, l is length of the sensing medium and d is the height of the sensing medium.

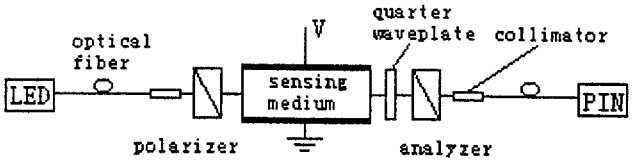


Fig. 3 Principles of the Optical Voltage Sensor

Thus the measured voltage V can be perceived by measuring the phase difference $\Delta\theta$. Similarly to the optical current sensor, the same methods are employed for measuring the phase difference: transforming the phase difference to a modulation of the transmitted light intensity. The elliptically polarized light emitted out of the sensing medium obtains a 90° optical bias through a quarter waveplate. Then the light passes through an analyzer, which axis of transmission is at 90° to the axis of transmission of the polarizer, to be converted into a intensity signal p which can be expressed as,

$$P = \frac{1}{2} P_0 [1 + \sin(\Delta\theta)] \tag{5}$$

where P_0 is the output light intensity of the polarizer.

According to equation (5), the measured voltage can be obtained with the optic-electric converting and signal processing units in the control room.

3. STRUCTURES OF THE COVCT

3.1. Structure of the Optical Current Sensor

The sensing material chosen and the structure of an optical current sensor are responsible for the performance of the sensor. Many materials have Faraday magnetic-optic effect, dense flint glass ZF-7 is chosen as the sensing material because dense flint glass ZF-7 is a diamagnetic medium which Verdet constant is approximately independent of temperature variation. As shown in Fig.4, in order to make the interaction between a linearly polarized light and a magnetic field produced by the current-carrying conductor encircles the current-carrying conductor, a bulk glass closed-loop structure is adopted. For each 90° change of the direction of light, there are twice the orthogonal reflections, thus a complementary method[3] can be used to cancel reflection-induced phase difference. Then the state of polarization of the light can be preserved within the whole circle. Because of the special structure shown in Fig.4, the output light and the input light are almost in the same plane, the no-closure of the light circle is less than 0.5%.

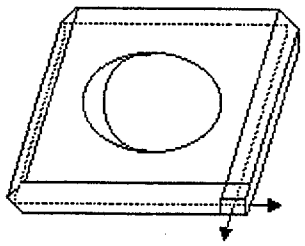


Fig. 4 Closed-Loop Structure of the Optical Current Sensor

The optical components used in the optical current sensor, including the ZF-7 glass ring, a polarizer, an analyzer and two collimators, are glued together with optical glue, and are housed in a box filled with heat-insulated and vibration-prevented

materials. As shown in Fig.1 the optical current sensor is located at the top of a polymeric insulator.

3.2. Structure of the Optical Voltage Sensor

For an optical voltage sensor, it is expected that the modulated output signal varies only with the measured voltage. This is impossible, because for an electro-optic material there are also many disturbing factors such as elasto-optic effect, thermo-optic effect, optical activity and natural birefringence etc which are superimposed on electro-optic effect and affect the accuracy of an optical voltage sensor. The structure of an optical voltage sensor and the optical components chosen are also responsible for the performance of an optical voltage sensor. $Bi_4Ge_3O_{12}$ crystal (BGO) is used as the sensing medium because it belongs to the point group $\bar{4}3m$, it has neither natural birefringence nor optical activity.

The structure of optical voltage sensor is shown in Fig.5. The sensing medium BGO and two 90° prisms are glued on the ground electrode. The two 90° prisms have both functions, one is to change the direction of light, another is to replace a quarter wave-plate to produce a 90° optical bias[4]. Other optical components including a polarizer, two analyzer, a 90° prism and three collimators are fixed inside the ground electrode. In order to avoid the BGO crystal and the two 90° prisms being affecting by SF_6 gas pressure which varies with temperature, a plexiglass cylinder with a round fillister as shown in Fig.5 is employed to protect the BGO crystal and the two 90° prisms. The plexiglass cylinder has also another function which is to make the distance between the HV electrode and the ground electrode maintain constant as much as possible, this is very important for accurate measurement of voltage. There is also a hollow plexiglass pipe between the HV electrode and the ground electrode which has a function of supporting the HV electrode. This optical voltage sensor has two channel output signals which can be used to improve the accuracy and stability of the optical voltage sensor[5].

As shown in Fig.1, the optical voltage sensor is fixed at the middle of the polymeric insulator filled with 0.35MPa SF_6 gas with two metallic pipes. The measured voltage is directly applied to the optical voltage sensor, so careful dielectric design of the sensor is required for avoiding dielectric breakdown. The COVCT is designed for a 220kV substation, the phase-to-ground voltage that is applied to the voltage sensor is 127kV rms (220kV rms/ $\sqrt{3}$). According to IEC standards[6] the sensor must be able to withstand not only the nominal operating voltage of 127kV rms, but also a short-duration ac voltage of 460kV rms for 1 min and lightning impulse voltages of 1050kV peak of positive and negative polarity. The sensor is designed such that the field strength at the 1050kV lightning impulse voltage does not exceed an empirically determined safety limit of 23kV/mm at any location in the SF_6 gas.

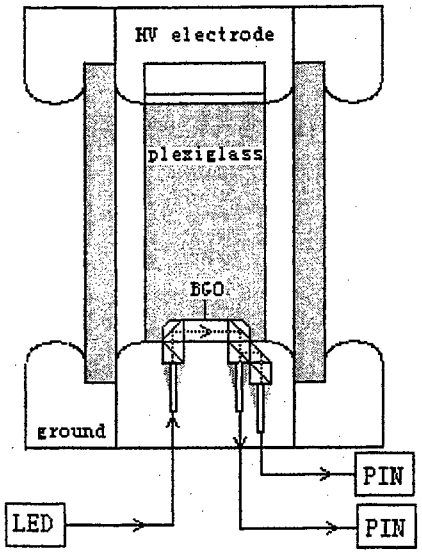


Fig.5 Structure of the Optical Voltage Sensor

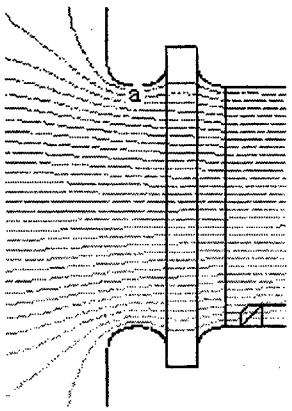


Fig.6 Lines of equipotential of the optical voltage sensor

As shown in Fig.6, the lines of equipotential obtained by numerical field computation show a fairly homogeneous field distribution. The label a indicates the location of relative field strength maximum, and the strength at the point a at 1050kV voltage is 12.6kV/mm.

4. TESTS OF THE COVCT

4.1. Dielectric Reliability Tests

The high voltage tests were performed to verify the dielectric reliability of the COVCT. These tests included 460 kV power-frequency short-duration (1 minute) withstand test and 1050 kV lightning impulse withstand tests. No dielectric breakdown occurred. These tests show that the dielectric design of the COVCT is reliable.

4.2. Linearity Tests

Fig.7 shows the linearity of the optical current sensor of the COVCT. Fig.8 shows the linearity of the optical voltage sensor of the COVCT. These tests show that the linearity of the COVCT is better.

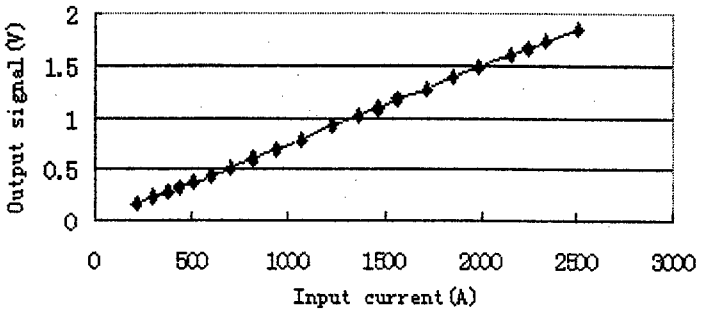


Fig.7 Linearity of the Optical Current Sensor

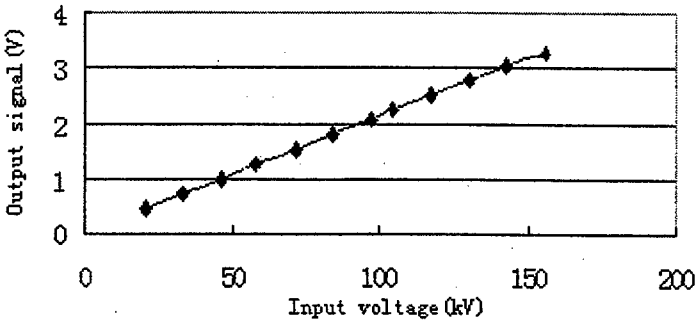


Fig.8 Linearity of the Optical Voltage Sensor

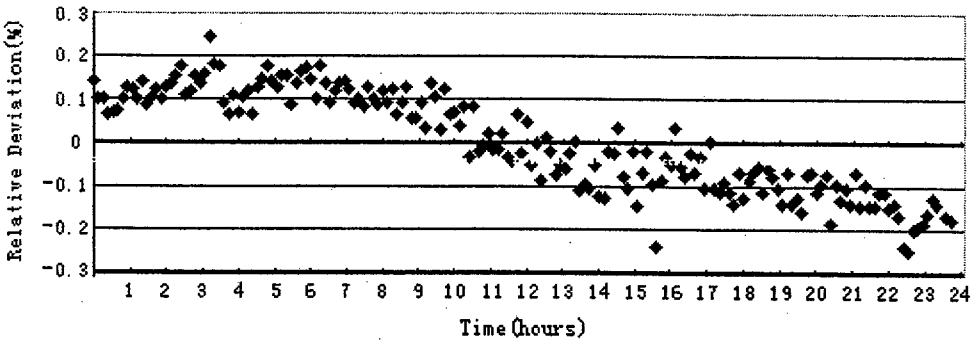


Fig.9 Stability of the Optical Voltage Sensor Over a 24-h Period

4.3. Stability Tests

Fig.9 shows the relative deviation of the ratio of the output signal to the input voltage of the optical voltage sensor over a 24-h period.

5. CONCLUSION

The combined optical voltage and current transformer can measure high voltage and large current simultaneously without interference each other. A novel structure is used to construct the optical voltage sensor. The testing results showed that the performances of the COVCT are better. And the insulation structure of the COVCT is simple and reliable. The temperature performance and the long term stability of the COVCT will be further examined.

REFERENCES

1. Emerging Technologies Working Group, "Optical Current Transducers For Power Systems: a Review," IEEE Trans. Pwr. Delivery, Vol.9, pp.1778-1788, 1994.
2. Lars Hofmann Christensen, "Design, Construction, and Test of a Passive Prototype High Voltage Instrument Transformer," IEEE Trans. Pwr. Delivery, Vol.10, pp.1332-1337, 1995.
3. Ye Miaoyuan, Hu Shichuang, "Design of a Magneto-optic Current Sensor," J. Huazhong Univ. of Sci. & Tech. Vol.15, pp.119-124, 1987.
4. Henrik Fabricius, "Achromatic Prism Retarder for Use in Polarimetric Sensors," Applied Optics, Vol.30, pp.426-429, 1991.
5. Kyung S. Lee, "New Compensation Method For Bulk Optical Sensors With Multiple Birefringence," Applied Optics, Vol.28, pp.2001-2011, 1989.
6. IEC 44-7:1993, Instrument Transformers-Part 7: Requirements For Electronic Voltage Transformers.

Research On Overcoming Temperature Offset and Noise Affection In Optical Current Transducer

GuoPing Yan, YaoJun Huang, Haiqing Chen, ZhongGe Sun
Huazhong University of Science&Technology ,Wuhan, 430074, P.R.China
Tel:+86-27-87543435 E-mail:hustradio@263.net

Abstract

Optical Current Transducer (OCT) is a new kind of transducer applying in electric power measurement system. This paper analyses the questions of temperature offset and noise affection in measuring DC current with OCT. The solutions to these questions are described in this paper. One is adopting AC modulation technology to overcome temperature offset. The other is an adaptive filter algorithm which is used in OCT to refrain noise. It is proved by experiments that the two methods are effective and can improve the measurement accuracy.

KeyWords: Optical Current Transducer, DC current measurement, AC modulation, Adaptive filter algorithm

1. INTRODUCTION

With the development to high voltage and large current of electric power system , Higher demands are put forward to electric power measurement system. Traditional electric-magnetic current transducer is not fit for modern electric power system because of low accuracy and huge volume. OCT ,which is based on Faraday electric optical effect , is a kind new current measurement transducer using optical and electric measurement technology. OCT has many advantages such as high accuracy, noncontactive measurement etc. Because of these advantages, OCT becomes more and more important nowadays in electric power measurement field.

The measurement on DC current is an important part of electric power system. So it is necessary to measure DC current with OCT. But, different with the measurement of AC current, There exist temperature offset and noise affection in the measurement of DC current because DC signal measurement is easy to be affected by temperature and noise. To improve the DC current measurement accuracy with OCT, some methods must be taken in OCT to solve these questions. In this paper, we adopt AC modulation technology to overcome the affection that is caused by temperature change. Meanwhile, a kind of adaptive filter algorithm is used to refrain noise affection in the measurement with OCT.

2.DC CURRENT MEASUREMENT THEORY WITH OCT

OCT is put forward on the basis of Faraday effect. Magnetic field can change the polarization state of light . According to this theory, a new kind of transducer called Optical Current Transducer (OCT) is proposed and applied in electric power system. Current including AC current and DC current can be measured with OCT. Many researches have been done on AC current measurement with OCT, so this paper only discusses the DC current measurement. The frame of OCT used in DC current measurement is showed in the following Fig.1

Part of the SPIE Conference on Advanced Photonic Sensors and Applications
Singapore • November–December 1999
SPIE Vol. 3897 • 0277-786X/99/$10.00

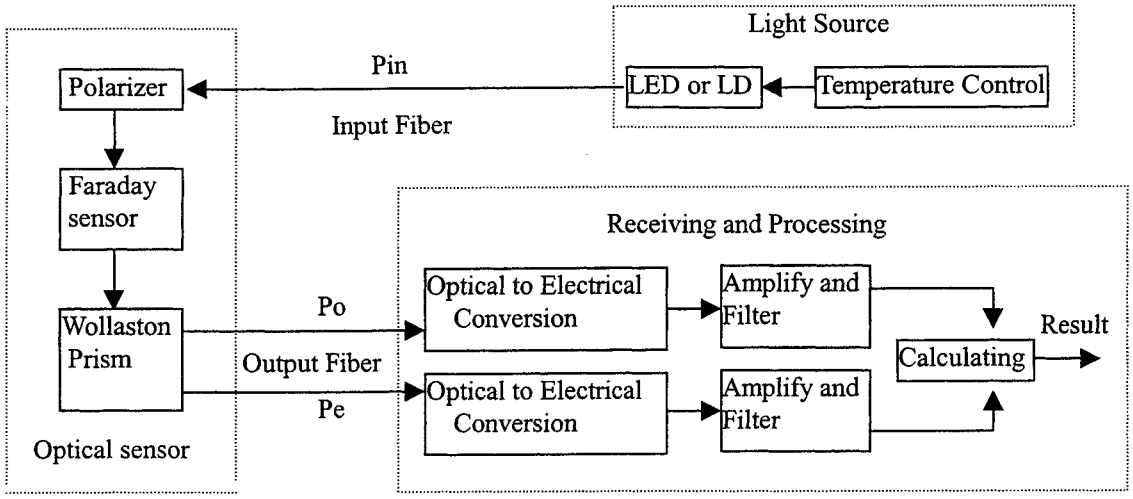


Fig.1 Frame of OCT in DC measurement

Fig.1 shows the basic components of an OCT used in DC current measurement. A beam of light is produced through LED or LD in light source and is used as light carrier wave . Traditionally, the intensity of this light is constant . Input optical fiber transmit this light to the optical sensor .In the sensor the light carrier wave is modulated by the magnetic field which is produced by the DC current to be measured.
Equation (1) describes the rotation of polarized light by the Faraday effect in a simple form:

$$\theta = \mu VHL \qquad (1)$$

Where θis the angular of the plane of polarization,μis the permeability of the material, V is the Verdet constant of the Faraday sensor, H represents the magnetic field intensity, L is the path length in the glass over which the field interacts with the light. What's more, the exact relationship between H and the DC current to be measured is proportional , so we can write like this:

$$\theta = \mu V(KI)L \qquad (2)$$

Where K is a geometry-dependent constant.
I is the DC current to be measured

From equation (2), it can be seen that the value of θis proportionate to the value of the DC current . So the value of the DC current can be obtained through the measurement onθ. Pin is the light which is produced from the light source and is used as light carrier wave in OCT. The optical sensor outputs two beams of light whose polarization states are vertical with each other. One is called ordinary light ,like Po in Fig.1. The other is called extraordinary light, like Pe in Fig.1. The optical power of Po and Pe is defined as the following equation (3).

$$P_o = \frac{1}{2}P_{in}\ (1+\sin\ 2\theta) \qquad (3)$$
$$P_e = \frac{1}{2}P_{in}\ (1-\sin\ 2\theta)$$

In fact, in equation (1) and (2) , the Verdet constant V for a given material describes the magnitude of the Faraday effect as the rotation per unit field per unit length. And this constant is so small that the value of θis comparable with zero. So we can say Sin2θ≈2θ. Then, equation (4) is brought about from equation (3)

$$\dot{E} = \frac{1}{2}\frac{P_o - P_e}{P_o + P_e} \qquad (4)$$

From equation (2) and (4) , the final measurement equation is brought about in the following form:

$$I = \frac{1}{2\,\mu\,VKL}\frac{Po - Pe}{Pe + Pe} \qquad (5)$$

As is showed in equation (5), to get the value of the DC current , the optical power of Po and Pe must be measured first, then the result can be obtained by the calculation according to equation (5).

3.TEMPERATURE OFFSET AND AC MODULATION

2.1 Temperature Offset and Its Affection

From Fig.1, it can be seen that OCT is composed of three parts: light source, optical sensor, receiving and processing part. Light source generates light carrier wave which is modulated by magnetic field in optical sensors. Two beams of light are produced from optical sensors. In receiving and processing part ,the modulated lights are measured and calculated according to equation (5) and then the value of DC current is obtained. In receiving and processing part, there are some analog electronic device uesd, for example operational amplifier. These devices exist temperature offset. In theory, the output of these devices should be zero when no signal is input, but in fact the output signal is not zero and its value is changed with the temperature. This is called temperature offset and this nonzero output is called temperature offset signal. This offset signal changes slowly and can be considered as a DC signal. Traditionally, in the measurement with OCT, the ligth wave generated by the light source is constant and this is called DC modulation technology. This kind of modulation technology is easily practiced but not fit for DC current measurement. If DC modulation technology is applied in OCT, the intensity of the output light Po and Pe is constant . So the useful signal Po and Pe are also DC signal and they have the same frequency and phrase (DC signal) with the temperature offset signal. Thus Po and Pe can not be distinguished from the temperature offset signal. Then the measurement result contains the true value and the affection which is caused by temperature offset.

2.2 AC Modulation

Temperature offset is an important factor which restricts the application of OCT in measuring DC current. To solve this question, AC modulation technology is adopted . The principle of AC modulation is that the carrier wave of OCT adopts AC light signal, not DC light signal which is often used in traditional OCT. AC light signal is produced by driving light source LED or LD with AC current signal. This AC light signal is used as the carrier wave of OCT and modulated in optical sensor. So the output light Po and Pe are also AC signal, and they can be distinguished from the temperature offset signal because of different frequency. The carrier light wave and the output light wave are defined in the following equations (6) and (7),

$$P_{in} = P_m(1 + \mathrm{Sin}\,\omega t) \qquad (6)$$

Where Pm is the DC component of Pin,
ω is the frequency of Pin.

From equation (3) and (6), Po and Pe can be defined as this,

$$P_o = \frac{1}{2} P_m (1+\sin \omega t)(1+\sin \varphi)$$
$$= \frac{1}{2} P_m (1+\sin \varphi) + \frac{1}{2} P_m (1+\sin \varphi) \sin \omega t$$

$$P_e = \frac{1}{2} P_m (1+\sin \omega t)(1-\sin \varphi) \quad (7)$$
$$= \frac{1}{2} P_m (1-\sin \varphi) + \frac{1}{2} P_m (1-\sin \varphi) \sin \omega t$$

From equation (7), it can be seen that the output light Po and Pe both contain two parts: one is DC light signal , the other is AC light signal. So, the electric signals corresponding to the light signals also contain DC electric signal and AC electric signal. And the value of electric signal is proportionate to that of the light signal because the conversion of light signal to electric signal is linear. To refrain temperature offset, A narrow band pass filter is set in the signal process part. Through this filter, the AC signal is passed while the DC signal including temperature offset signal is bypassed. And then the signal obtained is like this:

$$P_o' = \frac{1}{2} P_m (1+\sin \varphi) \sin \omega t$$
$$P_e' = \frac{1}{2} P_m (1-\sin \varphi) \sin \omega t \quad (8)$$

Then Po' and Pe' are changed to digital signal through analog-digital conversion device and are processed according to equation (5) to get the final true result. Because AC modulation technology is adopted in OCT, temperature offset can be avoided and the measurement accuracy is improved.

4.NOISE AND ADAPTIVE FILTER ALGORITHM

3.1 Noise and Its Affection

To get the final result, signal must be processed in OCT. The signal process in OCT contains optical signal process, conversion of optical signal to electric signal and electric signal process. In the period of signal process, noise is introduced and then affects the measurement accuracy. Noise introduced mainly comes from two parts: one is optical noise which is introduced in the period of light modulation and light transportation. This kind of noise is produced by optical device like optical sensor and optical fiber. The other is taken placed in light receiving and processing part where light signal is first turned into electric signal and then processed. And that is the main source introducing noise. This is because the output signal is very weak and need to be amplified large enough. In the period of amplification, noise is also amplified and mixed with the true signal . So the final signal is destroyed by noise and the measurement accuracy is decreased.

3.2 Adaptive Filter Algorithm

Noise affects the measurement accuracy. To decrease the affection of noise, some methods must be adopted to purify the signal. A kind of adaptive filter algorithm is adopted in OCT to refrain noise affection. Adaptive filter algorithm can trace the change of signal and noise. The principle of this kind of algorithm is based on the statistical characteristic of the signal and the noise. Under this adaptive filter algorithm, the useful signal is intensified and the noise is refrained. The model of this algorithm used in OCT is showed in the following Fig.2.

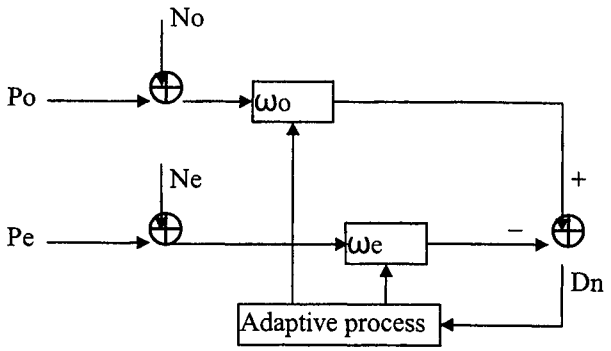


Fig.2 Model of Adaptive Filter Algorithm in OCT

As is showed in Fig.2, Po and Pe represent the true signal. No and Ne represent the noise. Here noise is the total of all noise introduced. ωo and ωe respectively represent the quotiety of the adaptive filter. From Fig.2, it can be seen that No comes from ordinary light signal channel and the corresponding electric channel while Ne comes from extraordinary light signal channel and the corresponding electric channel. Because the two channels are independent, noise No and Ne are independent with each other. Also noise No and Ne are independent with the signal Po and Pe. So we can draw the following conclusion:

$$\begin{aligned} &E(No\cdot Ne)=0\\ &E(No\cdot Po)=0\\ &E(No\cdot Pe)=0\\ &E(Ne\cdot Po)=0\\ &E(Ne\cdot Pe)=0 \end{aligned} \qquad (9)$$

Where $E(X\cdot Y)$ represents the expectation value of variable X and Y.

Meanwhile, equation (5) can be changed into the following form:

$$I = \frac{1}{2\mu VKL}\,\frac{(P_o/P_e)-1}{(P_o/P_e)+1} \qquad (10)$$

From equation(10), it can be seen that the final result is only relevant with the ratio between Po and Pe. So it is not necessary to measure the absolute value of Po and Pe. Since the result is only relevant with the ratio between Po and Pe, we can assume ωo =1 to simply the calculation. In Fig.2, Dn represents the difference between signal of ordinary light and extraordinary light. Then Dn is defined like this:

$$Dn = Po + No - (Pe + Ne)\omega e \qquad (11)$$

The control rule of adaptive process is that the value of ωe are adjusted so that the power of signal Dn can get a minimize value. The average power of signal Dn should be adjusted to a minimize value.

$$E(D_n^2) = \min \Rightarrow \frac{\partial E(D_n^2)}{\partial \omega_e} = 0 \qquad (12)$$

From equation(9) , (11) and (12), ωe can be obtained in the following form:

$$\omega_e = \frac{E(Po\,Pe)}{E(P_e^2)+E(N_e^2)} \qquad (13)$$

Because a narrow band filter is set in the receiving and processing part to restrict noise. Signal is passed while noise is bypassed mostly. Thus the value of signal Pe is much greater then that of the noise Ne. Meanwhile, this OCT is used to measure DC current, the value of the signal is constant after procession (before adaptive filter algorithm), so the statistical characteristic of signal Po and Pe have the following equation:

$$E(P_e^2) >> E(N_e^2)$$
$$E(Po) = P0$$
$$E(Pe) = Pe \qquad (14)$$
$$E(Po\ Pe) = E(Po)\ E(Pe) = Po\ Pe$$

Combining equation (13) and (14) , We can draw the following conclusion:

$$\text{¡}\emptyset_e = \frac{P_o}{Pe} \qquad (15)$$

Equation(15) shows that ωe is the value of ratio of Po and Pe, having nothing to do with the noise No and Ne. So equation (10) can be converted to the following form:

$$I = \frac{1}{2\mu VKL} \cdot \frac{\text{¡}\emptyset_e - 1}{\text{¡}\emptyset_e + 1} \qquad (16)$$

From the equation (16), it can be seen that the final result is only relevant with ωe, Noise is refrained by this adaptive filter algorithm. Thus the measurement accuracy can be improved.

5. EXPERIMENTS AND ANALYSIS

4.1 Experiments of AC Modulation

To test the effect of AC modulation, two comparative experiments are done. One experiment adopts DC modulation technology in a traditional way. The other adopts AC modulation technology. The carrier wave frequency in AC modulation experiment is set to 10KHZ. These two experiment both measure a standard 500A DC current under a series of different temperature. The result of experiment is showed in the following Fig.3

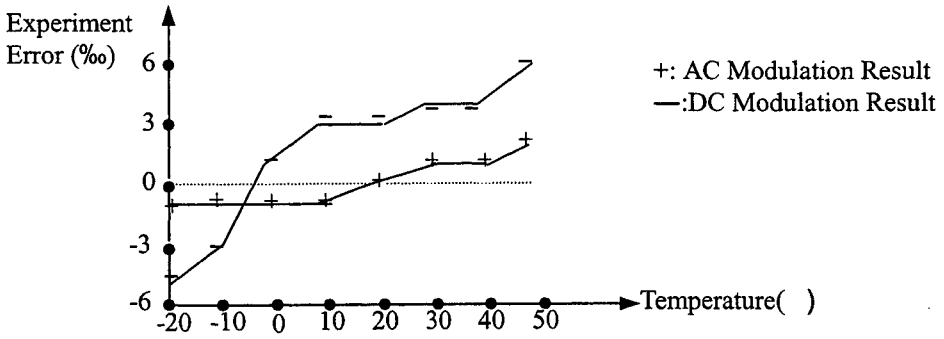


Fig.3: Experiment accuracy comparison under different modulation

From Fig.3 it can be seen that the result of AC modulation changes little while the result of AC modulation changes heavily. That is to say , AC modulation technology is stabler than DC modulation. And from Fig.3, also we can see that the accuracy of AC modulation is higher than that of DC modulation. It proves that the measurement adopting AC modulation in OCT is little affected by temperature offset and the accuracy is improved.

4.2 Experiments of Adaptive Filter Algorithm

Two comparative experiments are done to test the effect of adaptive filter algorithm. One experiment doesn't adopt adaptive filter algorithm, calculating the result according to equation (5). The other experiment adopt adaptive filter algorithm, calculating the result according to equation (16). These two experiments measure a series of different DC currents under the same conditions. The result of experiment is showed in the following Fig.4

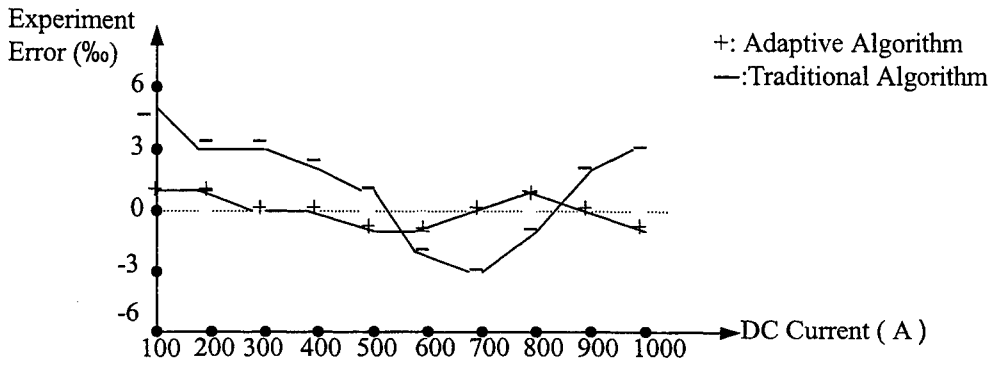


Fig .4 Experiment accuracy comparison under different filter algorithm

From Fig.4 it can be seen that the measurement error of experiment adopting adaptive filter algorithm is smaller than that of experiment which works in the traditional way. So it can be proved from the two comparative experiments that this kind of adaptive filter algorithm is effective in refraining noise affection and can improve the measurement accuracy.

6. CONCLUSION

In DC current measurement system with OCT , temperature offset and noise affection are two key factors affecting measurement accuracy. To improve the application of OCT in DC current measurement of electric power system, these questions must be solved. In this paper, two methods are put forward respectively to refrain temperature offset and noise affection. AC modulation technology, which is used to distinguish the true signal from the temperature offset signal through different frequency, can overcome the affection of temperature offset. Noise affection can be decreased by adopting adaptive filter algorithm which is proposed on the basis of different statistical characteristic for the true signal and the noise. From the analysis of theories and experiments, it can be seen that these two methods are effective and can improve the measurement accuracy.

ACKNOWLEDGMENTS

This research and paper are supported by Circuit&System Laboratory of Electronic and Information department , Optical and Electronic department of Hua Zhong University of Science and Technology. Also this work is benefited from Electric Power Research Institute of Wuhan, P.R.China.

REFERENCES

1 . Edward A.Ulmer, A high-accuracy optical current transducer for electric power systems,IEEE trans on Power Delivery,Vol 5,No 2, pp892-897,April 1990.

2 . H.Katsukawa, H.Ishikawa, H.Okajima, T.W.Cease, Development of an optical current transducer with a bulk type Faraday sensor for metering, IEEE trans on Power Delivery, Vol.11,No.2, pp702-707,April 1996.

3 . TianRen Yao, Hong Sun, Modern digtal signal process, Chapter 3,HuaZhong University of science and technology, Wuhan, 1996.

4 . Y.Yamagata, T.Oshi, H.Katsukawa, S.Kato, Y.Sakurai, Development of optical current transformers and application to fault location systems for substations, IEEE trans on Power Delivery ,Vol.8, No.3,1993

5 . Oppenheim .A .V, Schafer .R.W, Digtal signal processing, Chapter 4 , Prentice Hall, New Jersey,1988.

Effect of target position on signal amplitude in laser Doppler vibrometer

Vyacheslav Aranchuk

Scientific Center for Machine Mechanics Problems, Akademicheskaya Str. 12, Belarus, Minsk 220072

ABSTRACT

Laser Doppler vibrometers are widely used to obtain velocity information from vibrating targets. Doppler shift of the light scattered from a vibrating target is detected by using heterodyning with a reference beam. To obtain high spatial resolution and high intensity of the backscattered light and to produce speckles of large size at a photodetector plane the target is placed in the waist of a probe laser beam. At that target position the curvatures of the wavefronts of the reference beam and the scattered light can be different. The difference in the wavefront curvatures results in mismatching of wavefronts of the heterodyning beams and decreasing in the Doppler signal amplitude.

This paper investigates how the target position relative to a focusing lens can influence the matching of the heterodyning wavefronts and the Doppler signal amplitude in a two-beam interferometer. The investigation is carried out for the Gaussian beam normal to the target plane. The relationships between the coefficient showing the matching of the heterodyning wavefronts and geometrical parameters of the interferometer is obtained. Results of theoretical and experimental investigations are presented.

Keywords: laser vibrometer, Doppler signal, optical heterodyning

1. INTRODUCTION

Laser vibrometers are widely used for the measurement of out-of-plane vibration of diffuse objects[1-3]. In a laser vibrometer the laser beam is focused on the surface of the diffuse object. Vibration of the object produces the Doppler shift of the backscattered light. The Doppler signal is detected by heterodyning the backscattered light with a reference beam.

To obtain high intensity of the backscattered light the object is placed in the waist of the probe light beam. Positioning object in the waist of the probe light beam provides high spatial resolution and large size of speckles at a photodetector plane.

However, at that target position the curvatures of the wavefronts of the reference beam and the backscattered light can be different. The difference in the wavefront curvatures results in mismatching of wavefronts of the reference beam and the scattered light and decreasing in the Doppler signal amplitude. Vibration of the object results in changing the wavefront curvature of the scattered light. That also decreases the amplitude of the Doppler signal and limits the vibration range.

This paper investigates how the target position relative to the waist position of a probe beam can influence the matching of wavefronts of the reference beam and the scattered light and the Doppler signal amplitude. The investigation is carried out for the two-beam Michelson type interferometer. The investigation is carried out for the Gaussian beam normal to the target plane on the assumption that the light backscattered from the diffuse target is a homocentric beam.

2. THEORY

Two-beam Michelson or Mach-Zender type interferometer is used for optical heterodyning in laser vibrometers. The optical arrangement of the Michelson interferometer is shown schematically in Fig.1.

A Gaussian laser beam enters the interferometer. The input beam is splitted by a beam splitter in reference and probe beams. The probe beam is focussed by a lens onto the target. The backscattered light passes through the lens and is superimposed with the reference beam at a photodetector producing the interference signal. The interference signal generates the photodetector current $i_d(t)$ which can be expressed by the relationship[4].

$$i_d(t) = i_r + i_s + 2(i_r i_s)^{1/2} \operatorname{Re} \chi[- j(\omega_s + \omega_d - \omega_r)t], \quad (1)$$

Part of the SPIE Conference on Advanced Photonic Sensors and Applications
Singapore • November–December 1999
SPIE Vol. 3897 • 0277-786X/99/$10.00

where i_r, i_s is the DC component of the photodetector current caused by the reference beam and the scattered light respectively, χ is the coefficient which shows the matching of the heterodyning wavefronts, ω_s, ω_r is the frequency of the scattered light and the reference light respectively, ω_d is the Doppler shift, j is the imaginery unit. $i_s + i_r$ gives the DC component. The third term is the Doppler signal.

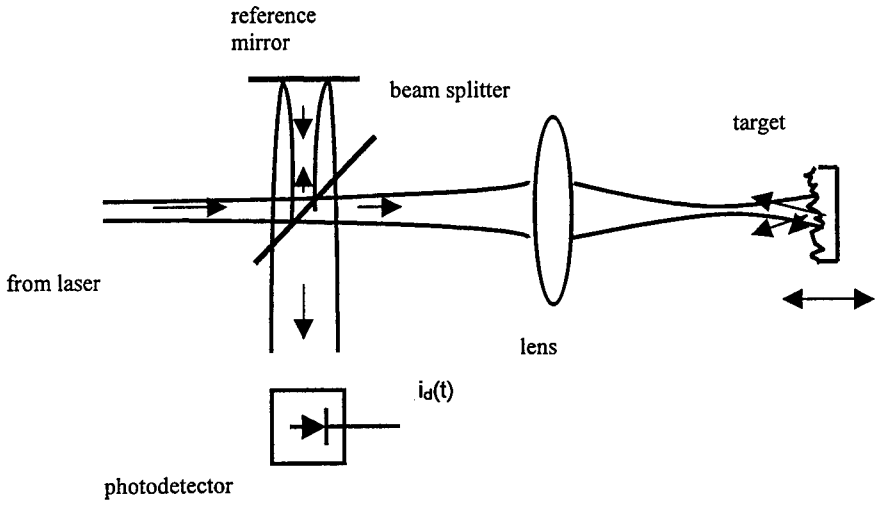


Fig. 1. Michelson type interferometer

The coefficient χ for the photodetector having circular aperture is expressed by[4]

$$\chi = \sin c\left[\left(\frac{1}{R_s} - \frac{1}{Rr}\right)\frac{r^2}{2\lambda}\right]\exp[-j(\varphi_r - \varphi_s)], \tag{2}$$

where R_s is radius of the wavefront curvature of the scattered probe light at the photodetector, R_r is the radius of the wavefront curvature of the reference beam at the photodetector, λ is the wavelength of light, r is the radius of the photodetector aperture, $\varphi_r - \varphi_s$ is the phase difference between the reference and scattered lights.

The light scattered from the target is the speckle field. The envelope of the the wavefront of the scattered light is a sphere. The lens changes the radius of curvature of the transmitted wavefront. The radius of curvature of the wavefront of scattered light at the photodetector plane is given by the well known expression

$$R_s = \frac{lf}{f - l} + l_0, \tag{3}$$

where l is the distance between the lens and the target, f is the focal length of the lens, l_0 is the distance between the lens and the photodetector plane.

$$l = l_w - s, \tag{4}$$

where l_w is the distance from the lens to the probe beam waist position, s is the displacement of the target relative to the beam waist.

Substitution of Eq. (4) into Eq. (3) yields

$$R_s = \frac{(l_w - s)f}{f - l_w + s} + l_0, \tag{5}$$

The position of the beam waist is defined by the lens focal length and the geometrical parameters of the input beam.

$$l_w = \left\{1 - \left(1 - \frac{l_1}{f}\right)\left[\left(1 - \frac{l_1}{f}\right)^2 + \frac{a^2}{f^2}\right]^{-1}\right\} f \tag{6}$$

where l_1 is the distance from the input beam waist position to the lens, a is the confocal parameter of the input beam.
The radius of curvature of the wavefront of the reference beam is given by

$$R_r = z\left[1 + \left(\frac{a}{z}\right)^2\right] \tag{7}$$

where z is the path length from the input beam waist position to the photodetector plane.
So, depending on the geometrical parameters of the interferometer the radii R_r and R_s can be equal at different positions of the target.
Figure 2 gives different examples for the absolute value of the coefficient $|\chi|$ as a function of the displacement s for different values of geometrical parameters of the interferometer. As seen from Fig.2 the $|\chi|$ can be a maximum at different target position relative to the waist of the probe beam. That is, the best matching of the heterodyning wavefronts occurs at different target position relative to the probe beam waist depending on the geometrical parameters of the optical configuration of the interferometer.
In a laser vibrometer the target is preferably be placed in the waist of the probe light beam. That provides high intensity of the backscattered light and speckles of large size at a photodetector and high spatial resolution of the vibrometer. Matching the wavefronts of the reference beam and the scattered light should be performed to obtain the maximum value of the Doppler signal amplitude. That is, the wavefronts of the reference beam and the scattered light must have the same radius of curvature at the photodetector.
The equality of the radii of curvature of the wavefronts of the reference and scattered light can be obtain by placing the positive lens in the reference arm of the interferometer (Fig.3).
The lens in the reference arm of the interferometer is placed so that the radius of curvature of the wavefront of the reference beam is equal to the radius of curvature of the wavefront of the light scattered by the target located at the waist of the probe beam.
In that case the radius of curvature of the reference wavefront is expressed by

$$R_r = \frac{l_w f}{f - l_w} + l_0 \tag{8}$$

The target vibration changes the curvature of the wavefront of the scattered light. That results in decreasing the amplitude of the Doppler signal due to the mismatching of the heterodyning wavefronts. The amplitude will decrease to the 0.2 level of the maximum value when[4]

$$|R_r - R_s| = \frac{4\lambda R_s R_r}{r^2} \tag{9}$$

When s is positive (the target moves to the lens) $R_r > R_s$. The Eq.(9) can be written as following

$$\frac{1}{R_s} - \frac{1}{R_s} = \frac{4\lambda}{r^2} \tag{10}$$

Substituting Eqs.(8) and (9) into Eq. (10) and solving the equation for s yields

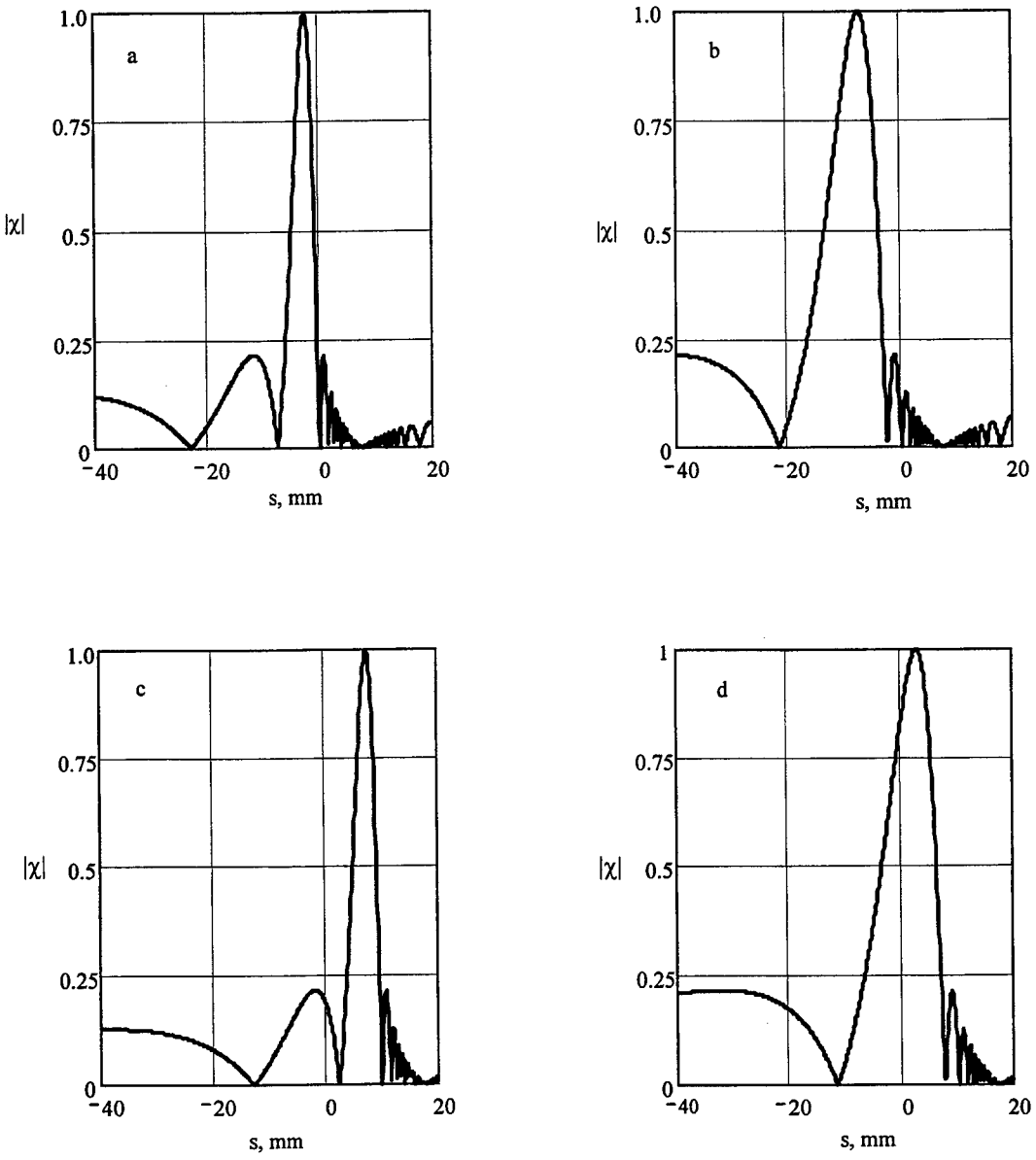


Fig.2. Absolute value of the coefficient |χ| versus the displacement s for f=75mm, r=1.5mm, λ=632 nm, l_0=400mm. a - l_w=100mm, R_r=120mm; b - l_w=100mm, R_r=150mm; c - l_w=110mm, R_r=120mm; d - l_w=110mm, R_r=150mm.

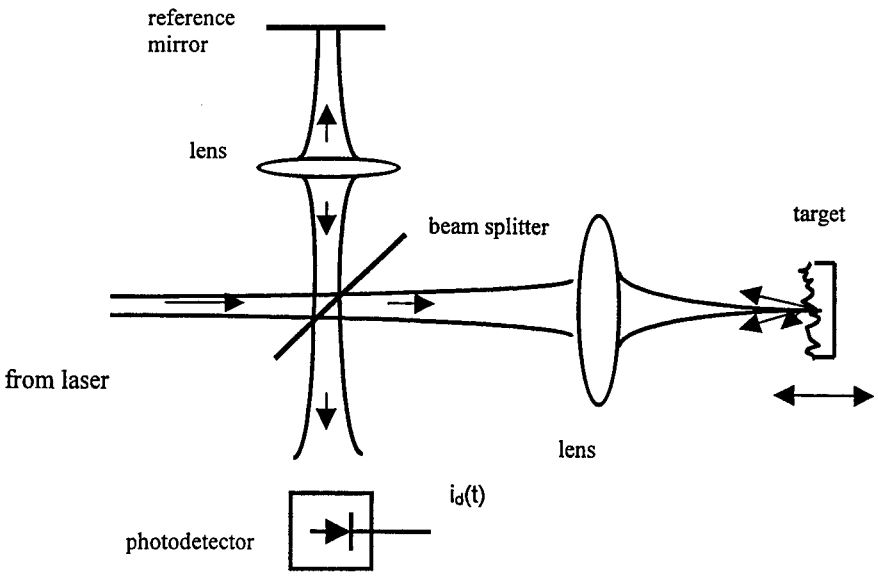


Fig. 3. Michelson type interferometer providing equal radii of curvature of the heterodyning lights

$$S_1 = \frac{\left(4\lambda/r^2\right)b^2}{b + \left(f - l_0\right)\left[\left(f - l_w\right) + 4\lambda b/r^2\right]}, \tag{11}$$

where $b = l_0(f - l_w) + l_w f$

When s is negative (the target moves from the lens) $R_r < R_s$. The Eq.(9) becomes

$$\frac{1}{R_r} - \frac{1}{R_s} = \frac{4\lambda}{r^2} \tag{12}$$

Substituting Eqs.(8) and (9) into (12) and solving the equation for s yields

$$S_2 = \frac{\left(-4\lambda/r^2\right)b^2}{b + \left(f - l_0\right)\left[\left(f - l_w\right) - 4\lambda b/r^2\right]} \tag{13}$$

S_1 and S_2 define the values of vibration at which the amplitude of the Doppler signal decreases to the 0.2 level of the maximum value due to the changing in the wavefront curvature of the scattered light. The vibration range defined by wavefronts mismatching is expressed by

$$\Delta S = S_1 - S_2 = \frac{4\lambda b^2}{r^2}\left[\frac{1}{b + (f - l_0)\left[(f - l_w) + 4\lambda b/r^2\right]} + \frac{1}{b + (f - l_0)\left[(f - l_w) - 4\lambda b/r^2\right]}\right] \qquad (14)$$

The Eq. (14) indicates that the vibration range depends on the geometrical parameters of the probe beam and the interferometer. Fig.4 shows how the vibration range ΔS depends on the distance l_0 between the focusing lens and the photodetector for different values of the waist position l_w.

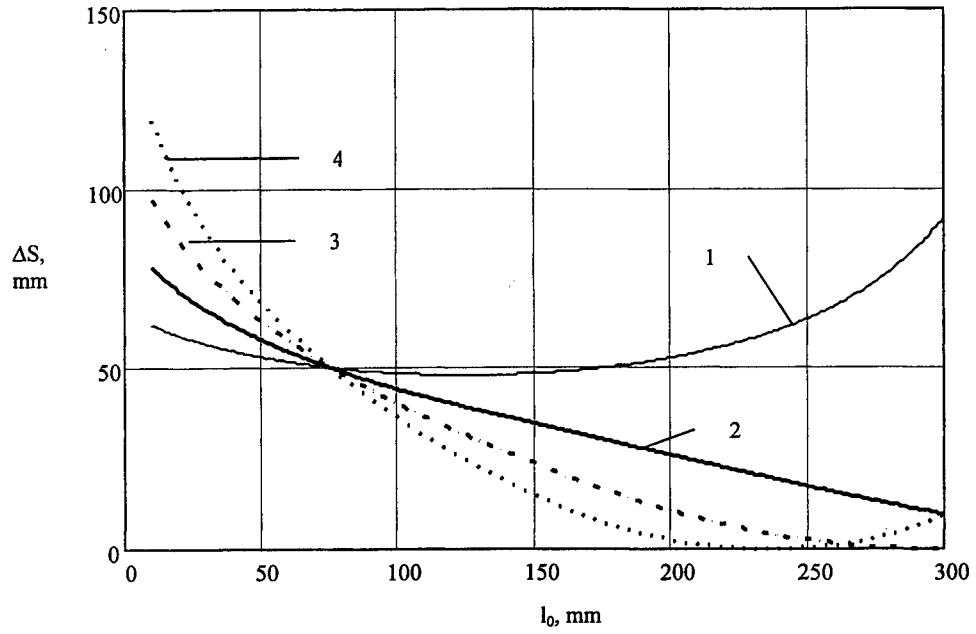


Fig.4. Vibration range ΔS versus the distance l_0 from the focusing lens to the photodetector for r=0.75mm, f=75mm; λ=632 nm. 1. l_w=80mm; 2. l_w=90mm; 3. l_w=100mm; 4. l_w=110mm.

3.EXPERIMENTAL RESULTS

Experimental investigation of the Doppler signal amplitude have been performed by using the optical configurations shown in Figs. 1, 3. A diffuse object (aluminum plate) vibrates with the amplitude of about 2 μm. The vibration was created by means of a piezoelectric vibrator. The vibrator was supplied with a 150 Hz AC signal. The vibrator was translated along the probe beam axis. The amplitude of the alternating current (the Doppler signal) of the photodetector was measured. The relationship between the Doppler signal amplitude and the target position was investigated by using the optical configuration shown in Fig. 1. The optical configuration has the following parameters: f=75mm, r=1mm, l_w=100mm, l_0=420mm, R_r=140mm. The experimental results are presented in Fig. 5. The theoretical relationship between the coefficient $|\chi|$ and the target displacement s is also presented. As seen from the Fig. 5 the Doppler signal amplitude is dropping faster then the coefficient $|\chi|$. It is likely caused by decreasing the intensity of the scattered light at the photodetector plane due to increasing the probe beam diameter on the target surface. The amplitude maximum occurs at the target placed behind the waist of the probe beam.

The relationship between the Doppler signal amplitude and the target position when the target was initially placed at the waist of the probe beam was investigated by using the optical configuration presented in Fig. 3. The lens in the reference

arm of the interferometer is placed so that the radius of curvature of the reference wavefront is equal to the radius of curvature of the wavefront of the light scattered by the target located at the waist of the probe beam. The equality of the wavefronts curvature was detected by maximum of the Doppler signal. The optical configuration has the following parameters: f=75mm, r=0.5mm, l_w=110mm, l_0=195mm. The experimental results are shown in Fig. 6. The theoretical relationship between the coefficient $|\chi|$ and the target displacement s is also presented. As seen from the Fig. 6 the peak positions of the theoretical and the experimental curves are practically coincide. The Doppler signal amplitude is dropping faster then the coefficient $|\chi|$. It is likely caused, like in the previous case, by decreasing the intensity of the scattered light at the photodetector plane due to increasing the probe beam diameter on the target surface.
The experimental results obtained basically confirm the theoretical results.

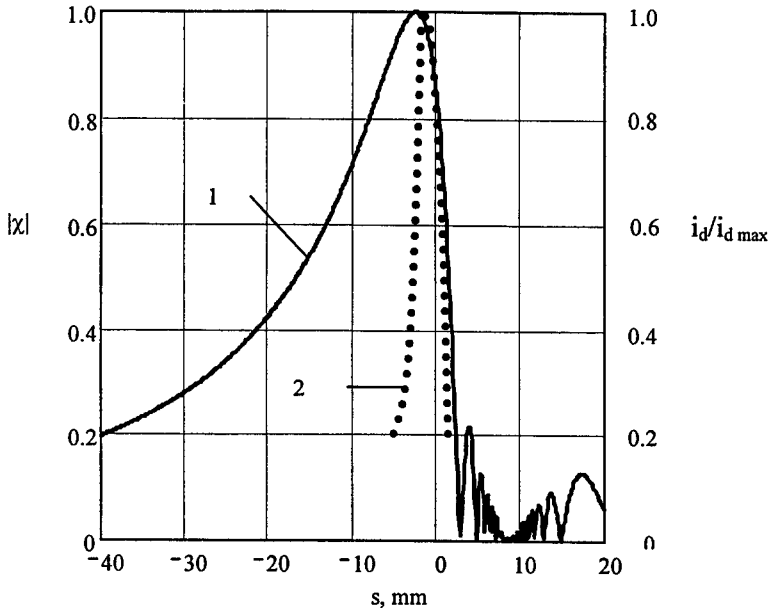


Fig.5. Coefficient $|\chi|$ and relative value of Doppler signal amplitude $i_d/i_{d\,max}$ versus the displacement s of the target for f=75mm, r=1mm, l_w=100mm, l_0=420mm, R_r=140mm λ=632 nm. 1- $|\chi|$, 2- $i_d/i_{d\,max}$

4.CONCLUSION

The investigation of the matching of the heterodyning wavefronts in the laser vibrometer based on a Michelson type interferometer is carried out.
It is shown that the target position at which the best matching of the heterodyning wavefronts and the Doppler signal amplitude maximum is obtained depends on the focal length of the focusing lens, the waist position of the probe beam, the photodetector position and the radius of curvature of the reference beam.
The maximum value of the Doppler signal amplitude occurs when the target is placed in waist position of the probe beam and the wavefronts of the reference beam and the scattered light have the same radius of curvature at the photodetector.
The expression for the vibration range defined by the wavefronts mismatching when the target is placed in the waist position of the probe beam is obtained.

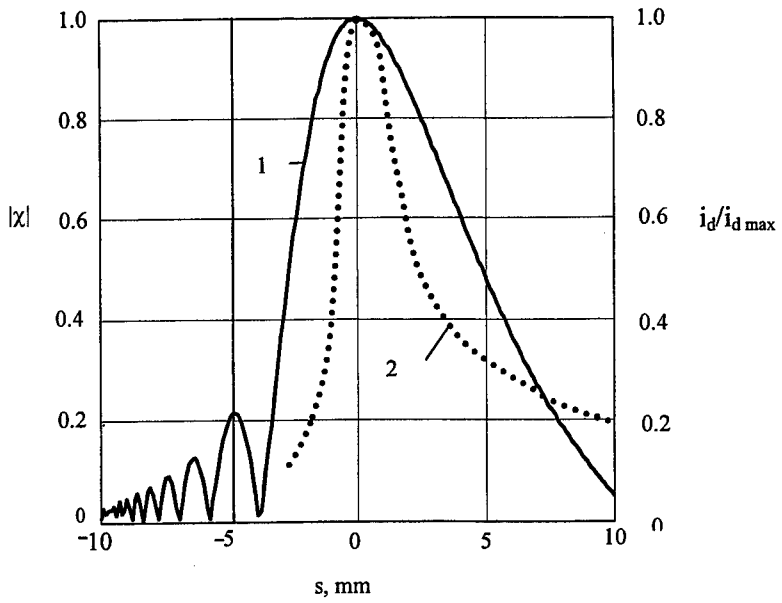


Fig.6. Coefficient $|\chi|$ and relative value of Doppler signal amplitude $i_d/i_{d\,max}$ measured according to the Fig. 3 versus the displacement s of the target for f=75mm, r=0.5mm, l_w=110mm, l_0=195mm, λ=632 nm. 1- $|\chi|$, 2- $i_d/i_{d\,max}$

REFERENCES

1. Product data: Scanning Vibrometer PSV-200. Polytec GmbH,Polytec-Platz 1-7, 76337 Waldbronn, Germany.
2. P.Castellini and C. Santolini, "Vibration measurements on blades of naval propeller rotating in water", Proceedings SPIE 2868, pp. 186–194, 1996.
3. A. B. Stanbridhe and D.J.Ewins, "Measurement of total vibration at a point using a conical-scanning LDV", Proceedings SPIE 2868, pp. 126–136, 1996.
4. Y. N. Dubnistchev and B. S. Rinkevichius, *Methods of Laser Doppler Anemometry,* Nauka, Moscow, 1982.

Simulation and Study on the Temperature Effect of the A-Ta_2O_5 ISFET

Jung Chuan Chou*, Ying Shin Li
Department of Electronic and Information Engineering, National Yunlin University of Science and Technology, Touliu, Yunlin, Taiwan 640, R.O.C.

ABSTRACT

In the research, we simulated the temperature characteristics of the a-Ta_2O_5 ISFET by the Gouy-Chapman-Stern theory. The values of the pK_a and pK_b would be induced to calculate the temperature coefficients of the a-Ta_2O_5 ISFETs for predicting the behaviours of the a-Ta_2O_5 ISFET under different temperatures. In the experiment we used the method finding the V_{GS} values of the experimental curves by fixed I_{DS} value to get the pH sensitivities at different temperatures. By using the same way we could change the temperatures to find the temperature coefficients in different pH solutions. The relationship of the pH sensitivities of a-Ta_2O_5 ISFET versus the temperatures were the linear. Oppositely the curves of the temperature coefficients is not linear obviously.

Keywords: a-Ta_2O_5 ISFET, temperature coefficient, pH sensitivity, pK_a , pK_b

1. INTRODUCTION

The ion-sensitive field effect transistor (ISFET) is an electronic device using an exposed insulator or membrane to measure a selected ion concentration in electrolytes. The response of the ISFET to pH is commonly explained by using the do-called site-binding model[1], which describes the charging mechanism of an oxide as the equilibrium between the AOH surface and H^+ ions in the bulk of the solution. These theories, together with the Gouy-Chapman-Stern model[2] of the potential profile in the electrolyte and with MOSFET physical[3], give a complete description of the ISFET behaviour. Investigations have demonstrated that ISFETs show large thermal instability, which leads to inaccuracy in measurement. The main focus of interest research is the measurement of pH, which can be realized with insulators common in semiconductor technology such as SiO_2, Si_3N_4, or Al_2O_3. Extension of sensitivity to other ions by adding certain types of membranes or surface coatings is also being extensively investigated[4-6]. The amorphous tantalum pentoxide (a-Ta_2O_5) has various applications in semiconductor technology such as its use as capacitor material in VLSI-technology, as material for electrical resistors, as gate insulator for field-effect transistor (FETs), and as pH-sensitive layer for ion-sensitive FETs (ISFETs)[7]. During the experiments we used the Ta_2O_5 for the ISFET coating film material. Among different materials (Si_3N_4, Al_2O_3, TiO_2, etc.) used as membranes in the pH-sensitive ISFET, one of the most suitable is Ta_2O_5 films. The Ta_2O_5-gate ISFETs have pH sensitivity (55-57mV/pH) near to theoretical, lower response time, higher selectivity and comparatively good long-term stability[7-9]. It is often necessary for ISFETs to be used under thermostatic conditions. Since this is not convenient for many applications, it is important to investigate the thermal behaviour of ISFETs. The important parameters pK_a and pK_b were unknown for the researches of the behaviour of ISFETs and how to find the pK_a and pK_b from known parameters should be done first, then the temperature characteristics of the a-Ta_2O_5 ISFET could be investigated and analyzed.

*Correspondence : Telephone: 886-5-5342601 Ext. 2500
Fax: 886-5-5312029
E-mail: choujc@pine.yuntech.edu.tw

Part of the SPIE Conference on Advanced Photonic Sensors and Applications
Singapore • November–December 1999
SPIE Vol. 3897 • 0277-786X/99/$10.00

2. SIMULATION

For the oxides on which we will focus our attention, namely SiO_2 and Al_2O_3 , it is usual to consider that only one type of site is present, of the type A-OH where A represents Si or Al[10-12]. To account for the fact that both signs of charge have been experimentally observed in colloid chemical studies[13], the site considered must be amphoteric, which means it can act as a proton donor or acceptor. We therefore assume that the oxide surface contains sites in three possible forms: $A\text{-}O^-$, A-OH, and $A\text{-}OH_2^+$. The acidic and basic character of the neutral site A-OH, can be characterized by two equilibrium constants K_a and K_b.

$$A-OH \Leftrightarrow A-O^- + H_s^+ \quad \text{, with} \quad K_b = \frac{[A-O^-][H_s^+]}{[A-OH]} \tag{1}$$

$$A-OH_2^+ \Leftrightarrow A-OH + H_s^+ \quad \text{, with} \quad K_a = \frac{[A-OH][H_s^+]}{[A-OH_2^+]} \tag{2}$$

$[H_s^+]$ represents the surface concentration of H^+ ions, which is related to the bulk value by Boltzmann statistics

$$[H_s^+] = [H^+]\exp(-q\phi_0 / kT) \tag{3}$$

The pH_{pzc} (the point of zero charge), which is one of the important parameters of pH sensitive membranes, determines the linear range of pH sensitivity and allows suppositions to be made about types of surfaces active centers and absorption abilities of surfaces. In order to calculate the site-binding model equations, we need to know pH_{pzc} of the material. Unfortunately, the information available from the colloid chemical literature is very sparse. Akiyama et al.[7] reported a pH_{pzc} of 2.8 and Bousse et al.[14] and Kerhof et al.[15] found 3.0; none of these authors describes the type and origin of the material tested or show their experimental data for Ta_2O_5. Poghossian et al.[16] tried to obtain a pH_{pzc} value from ISFET measurements, and they found $pH_{pzc}=2.8 \pm 0.3$. According to most of experimental values of references[14-16] about pH_{pzc}, they used the value of $pH_{pzc}=3$. We also used the $pH_{pzc}=3$ to be our simulation value. The equation as following[17]:

$$pH_{pzc} = -0.5\log(K_a K_b) = 0.5(pK_a + pK_b) \tag{4}$$

where $pK_a = -\log(K_a)$, $pK_b = -\log(K_b)$. The pK_a, pK_b are two of important parameters of influencing temperature behaviour of the ISFETs. If the pK_a and pK_b were very close, the curve of the temperature coefficient of the surface potential would approach to linear[18].

We adopt the Gouy-Chapman-Stern theory, and the literatures of Bousse et al.[19] to get relative functions of pH_{pzc} and K_a, K_b as following

$$2.303(pH_{pzc} - pH) = \frac{q\phi_0}{kT} + \sinh^{-1}\left(\frac{q\phi_0}{\beta kT}\right) \tag{5}$$

where $\beta = \dfrac{2q^2 N_s (K_a / K_b)^{1/2}}{kTC_{DL}}$, ϕ_0 is the surface potential of electrolyte/ insulator interface with respect to electrolyte bulk, and C_{DL} is electrode/oxide double layer capacitance, and N_s is surface charge density.

If differentiating above eqn.(5) with respect to pH and rearranging it yields

$$\frac{q}{kT} \times \frac{d\phi_0}{dpH} + \frac{1}{\sqrt{x^2+1}} \times \frac{q}{\beta kT} \times \frac{d\phi_0}{dpH} + 2.303 = 0 \tag{6}$$

where $x = \dfrac{q\phi_0}{\beta kT}$

The pH sensitivity, $d\phi_0/dpH$, has been carried out; thereupon all equations about inducing pK_a and pK_b were obtained. If we assumed that the pH sensitivity approach the ideal Nernstian value (59mV/pH), the pK_a and pK_b values could be gotten from eqn.(4) and eqn.(6). The program written by MATLAB was run on Pentium 166 computer to simulate the variance of pK_a and pK_b and we assumed that the dielectric constant was 20 when the Ta_2O_5 was annealed at 500~600℃ which was the optimal temperature [20-21] and the other parameter conditions were in Table I. The simulation results of the pK_a and pK_b were shown as. In the Fig. 1 we found that when the sensitivity was lower, the difference of the pK_a and pK_b was higher; otherwise if the pH sensitivity of a-Ta_2O_5 ISFET was close Nernstian response, the pK_a and pK_b had approached 2.16 and 3.84 respectively. These values was similar to the value of Bousse et al.[14].

Table I. Parameters of a-Ta_2O_5

Parameter	Value
N_s	10^{15} cm^{-3}
C_{DL}	20× 10^{-6} F
d (Ta_2O_5)	150nm
d (SiO_2)	50nm
Temperature	300K
C_o(Electrolyte concentration)	0.1 mole
ε_w (Electrolyte dielectric constant)	80
S (Sensitivity)	58~40mV/pH

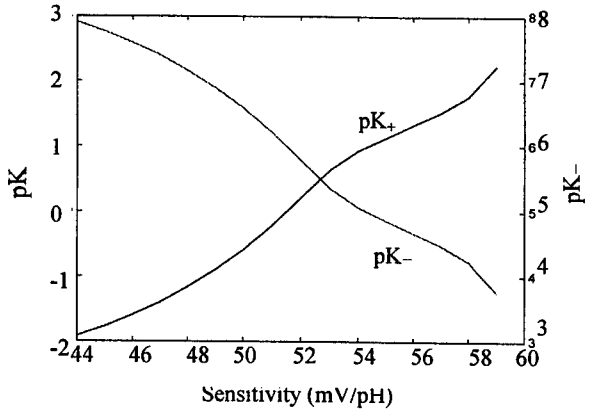


Fig. 1 Curves of the pK_a, pK_b versus sensitivity.

P. V. Bobrov et al.[22] explained that a-Ta_2O_5 had practically an ideal Nernstian pH response in a very broad pH range. In the basic region, the absence of alkali ions results in the decrease of pH sensitivity. So we chose pH=5 to be our input value of functions for simulating. The a-Ta_2O_5 ISFET had higher pH sensitivity and was more stable than other kinds of membranes, since the pK_a and pK_b were very small and close to each other.

In order to simulate behaviour of temperature variance of a-Ta_2O_5 ISFET, mathematical equation below in the reference[18] was used.

$$\frac{d\phi_0}{dT} = -\left(\frac{-2a_{H^+} q\phi \left(1-\frac{\phi_0}{\phi_{max}}\right)}{kT^2} + \frac{K_a C_1 \exp\left(\frac{q\phi_0}{kT}\right)\left(\frac{\phi_0}{\phi_{max}}\right)}{kT^2} + \left\{ \frac{\frac{4K_aK_b(C_1+C_2)}{kT^2} + \left[\frac{2K_a^2C_1}{kT^2} - \frac{4K_aK_b(C_1+C_2)}{kT^2}\right]\frac{\phi_0^2}{\phi_{max}^2}}{4a_{H^+}\left(1-\frac{\phi_0}{\phi_{max}}\right)\exp\left(-\frac{q\phi_0}{kT}\right) - 2K_+\frac{\phi_0}{\phi_{max}}} \right\} \exp\left(\frac{q\phi_0}{kT}\right) \right) \times \left(\frac{2a_{H^+}}{\phi_{max}} + \frac{K_a \exp\left(\frac{q\phi_0}{kT}\right)}{\phi_{max}} + \frac{2\left(1-\frac{\phi_0}{\phi_{max}}\right)qa_{H^+}}{kT} + \frac{2(K_a^2-4K_aK_b)\left(\frac{\phi_0}{\phi_{max}^2}\right)\exp\left(\frac{q\phi_0}{kT}\right)}{4a_{H^+}\left(1-\frac{\phi_0}{\phi_{max}}\right)\exp\left(-\frac{q\phi_0}{kT}\right) - \frac{2K_a\phi_0}{\phi_{max}}} \right)^{-1} \tag{7}$$

where $a_{H^+} = \dfrac{\left\{K_a\phi_0/\phi_{max} + [(K_a\phi_0/\phi_{max})^2 + 4(1-\phi_0^2/\phi_{max}^2)K_aK_b]^{1/2}\right\}\exp(q\phi_0/kT)}{2(1-\phi_0/\phi_{max})}$

$C_1 = -kT\ln(K_a)$, $C_2 = -kT\ln(K_b)$, $\phi_{max} = qN_s/C_{ox}$, C_{ox} is oxide capacitance

Fig. 2 showed the curve of the temperature coefficient of surface potential versus pH. The slope of the curve was negative indicted that the larger pH value caused the larger temperature coefficient value. It means that a-Ta_2O_5 membrane in the strong acid is influenced by temperature easily.

Table II. Deposition conditions of a-Ta_2O_5 thin film

RF Power(W)	90
Reactive Pressure(10^{-3} torr)	10
Gas Mixed Ratio(Ar/O_2)	9:1
Gas Flow Rate(SCCM)	20

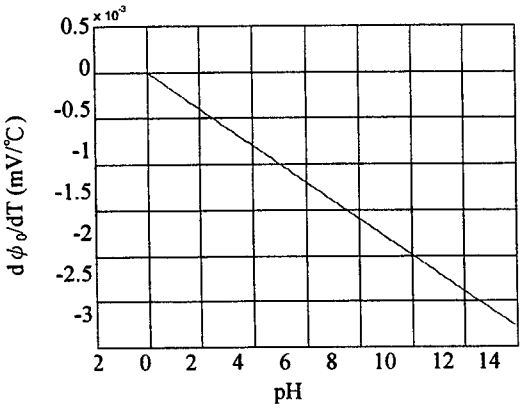


Fig. 2 Curve of $d\phi_0/dT$ versus pH

The relation of the surface potential ϕ_0 and the threshold voltage V_{th}^* of ISFET can be expressed as[23-24]:

$$V_{th}^*(ISFET) = (E_{ref} + \phi_{lj}) - (\phi_o - \chi_{eo}) - \left[\frac{Q_{ss} + Q_{sc}}{C_{ox}} - 2\phi_f + \frac{\phi_{sc}}{q}\right] \quad (8)$$

where E_{ref} is the reference electrode potential, ϕ_0 is the surface potential of electrolyte/ insulator interface with respect to electrolyte bulk. χ_{eo} is the electrolyte-insulator interface dipole potentials, ϕ_{sc} is the silicon work function, ϕ_f is the Fermi potential, Q_{SC} is the charge located in the oxide, Q_{SS} is the equivalent insulator-silicon interface charge per unit area, and C_{ox} is the insulator capacitance per unit area. ϕ_0, ϕ_f, Q_{SC} , Q_{SS} is temperature function, and if the working temperature changed, there would be different V_{th}^* values. If ISFETs operates in linear region, the output voltage of ISFET would be[25]

$$\frac{\partial V_{GS}}{\partial T} = \frac{1}{V_{DS}}\left(\frac{3}{4T}\right)\left(\frac{I_{DS}}{\mu C_{ox}\frac{W}{L}}\right) + \frac{\partial E_{ref}}{\partial T} - \frac{\partial \phi_0}{\partial T} + \frac{\partial V_{th}}{\partial T} \quad (9)$$

We know that the $\partial E_{ref}/\partial T + \partial V_{th}/\partial T$ was negative from reference[25]. If $\partial V_{GS}/\partial T$ is zero (at the point of zero charge), the increasing of the $\partial \phi_0/\partial T$ will make I_{DS} decreasing.

3. Experimental

3.1. Preparation of a- Ta_2O_5 pH-ISFET devices

The a-Ta_2O_5 pH-ISFETs were fabricated on p-type Si(100) wafer (8~12Ω · cm). The channel was 50 μm long and 1000 μm wide. The thickness of gate oxide (SiO_2) was 1000Å. The a- Ta_2O_5 thin film was deposited by sputter system. Table II. shows the deposition conditions of a-Ta_2O_5 thin film. The thickness of a-Ta_2O_5 thin film is 2000 Å, which was measured by Dektak[3] surface profile measuring system. And then, we deposition a layer of aluminum on a-Ta_2O_5 thin film through thermal evaporation system to form Al/a-Ta_2O_5/SiO_2 gate MOSFET. The structure of the a-Ta_2O_5 pH-ISFET was shown as Fig. 3.

3.2. Measurement of a-Ta_2O_5 pH-ISFET and MOSFET devices

We utilize Keithley 236 Semiconductor Parameter Analyzer to measure the I_{DS}-V_G curves of pH-ISFET for different pH values at temperature 15℃ 、25℃ 、35℃ 、45℃ 、55℃. The measurement system in this study was shown as Fig. 4. The sensing structure and reference electrode were dipped into buffer solution and connected with the gate of a-Ta_2O_5 ISFET. Measured conditions of our experiment was as follows: V_{DS}=0.2 V; starting V_G=0 V; stopping V_G= +5 V stepping V_G= 0.1 V. Under these conditions, the a-Si:H pH-ISFET was operated at linear region.

4. RESULTS AND DISCUSSION

4.1 The sensitivities of the experimental and simulation

The operating point of a-Ta_2O_5 ISFET was at I_{DS}=40 μA, and measuring the V_{GS} value on the I-V curve could get the Fig. 5. The slope of the line is the sensitivity of a-Ta_2O_5 ISFET. At different temperatures the a-Ta_2O_5 ISFET would have different

sensitivities. The curve collecting the experimental values at different the temperatures was drawn as the Fig. 6. The slope of experimental curve was about 0.134mV/pH℃ and the slope of the simulation curve was 0.196mV/pH℃. The reason, which caused the difference between the experiment and simulation, was that the a-Ta_2O_5 ISFET fabricated in the laboratory was not perfect device. And the building chemical functions on the surface of sensitivity thin film were not completion.

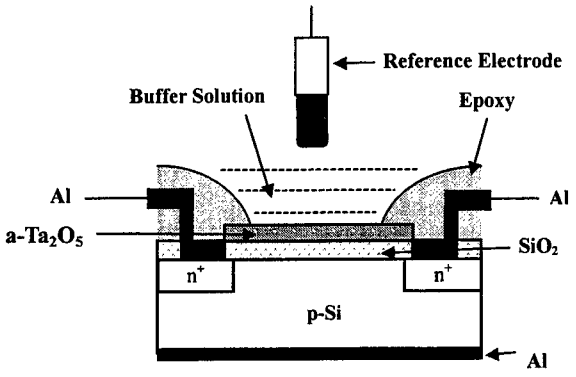


Fig. 3. Structure of the a-Ta_2O_5 pH-ISFET

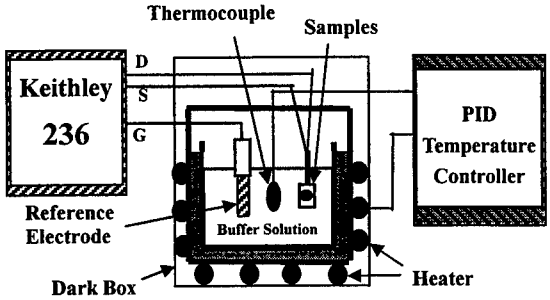


Fig. 4. I-V measuring system

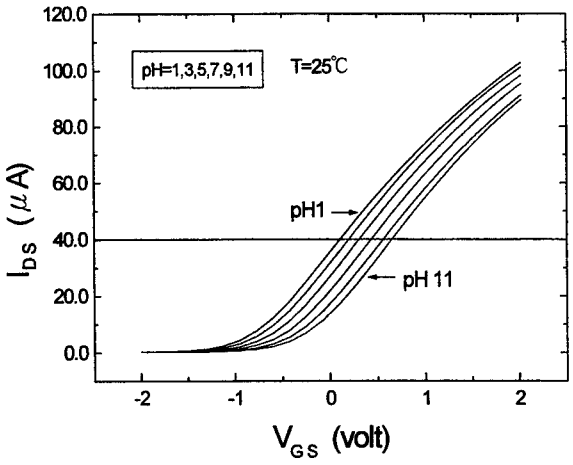


Fig. 5. I-V curves of Ta_2O_5-ISFET at 25℃

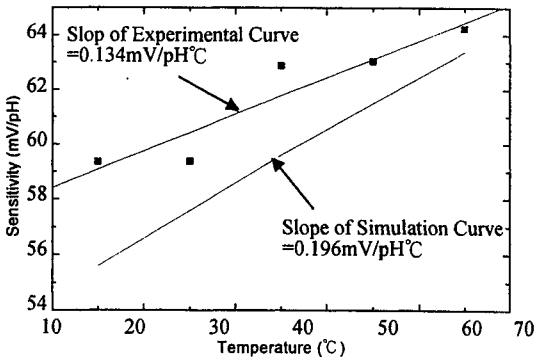


Fig. 6. Curves of sensitivities of Ta_2O_5-ISFET versus temperatures.

Otherwise we studied the temperature coefficients of the a-Ta_2O_5 ISFETs in pH1, 3, 5, 7, 9, 11 solutions. The I_{DS} was fixed at 40 μA and 20 μA respectively for measuring V_{GS} of the I-V curve at different temperatures. It could be calculated the temperature coefficients of the a-Ta_2O_5 ISFETs. The collected values were shown as Table III. The relationships of temperature coefficients versus the pH values were not linear obviously. The range of temperature coefficient was about -5.4 ~ -8.6mV/℃. The temperature coefficients were unstable because the MOSFET was effected by external environment easily.

Table III. Temperature coefficients of the a-Ta_2O_5 ISFET

PH		1	3	5	7	9	11
Temperature coefficients (mV/℃)	I_{DS}= 40 μA	-5.41	-6.46	-7.45	-6.59	-5.49	-7.52
	I_{DS}= 20 μA	-7.81	-7.53	-8.68	-7.00	-6.51	-8.45

5. CONCLUSIONS

From above simulation, we found the pK_a=2.16 and pK_b=3.84 to be the dissociative constants of a-Ta_2O_5-gate ISFET. We could analyze its pK_a and pK_b values by the same way for other ISFET materials. It provides us a convenient way to find pK_a and pK_b values from known pH_{pzc} and sensitivity given by experiments. When the values of pK_a=2.16 and pK_b=3.84 were inputted into this simulation functions, the curve of $d\phi_0/dT$ versus pH had been linear. The sensitivities of the a-Ta_2O_5-gate ISFETs will over the 59mV/pH when the temperature was 40℃. The temperature coefficient of sensitivities was about 0.196mV/pH℃ and the sensitivity increased with temperature. In this paper, we studied the sensitivities and the temperature coefficients of the a-Ta_2O_5 pH-ISFET. The curve of the sensitivity versus temperature was linear and the temperature coefficient of sensitivities was about 0.134mV/pH℃. The relationships of temperature coefficients versus the pH values were not linear obviously. The range of temperature coefficient was about -5.4 ~ -8.6mV/℃.

ACKNOWLEDGEMENTS

The authors acknowledge the financial support of the National Science Council of Republic of China under grant NSC 89-2215-E-224-003.

REFERENCES

1. W. M. Siu and R. S. C. Cobbold, "Basic Properties of the Electrolyte-SiO_2-Si System: Physical and Theoretical Aspects", IEEE Trans. Electron Devices, vol. ED-26 (1976) 1805~1808.
2. Meng-Nian Niu, Xin-Fang Ding, Qin-Yi Tong, "Effect of Two Types of Surface Sites on the Characteristics of Si_3N_4-Gate pH-ISFET", Sensors and Actuators B, vol. 37 (1996) 13-17.
3. M. Grattarola, G. Massobrio, and S. Martinoia, "Modeling H^+-Sensitive FET's with SPICE", IEEE Trans. Electron Devices, vol. ED-39 (1992) 813-819.
4. S. D. Moss, J. Janata, and C. C. Johnson, "Potassium Ion Sensitive Field Effect Transistor", Anal. Chem. , vol. 47 (1975) 2238-2243.
5. R. P. Buck and D. E. Hackleman, "Field effect potentiometic sensor", Anal. Chem. , vol. 49 (1977) 2315-2321.
6. S. D. Moss, J. Janata, and C. C. Johnson, "Hydrogen, Calcium and Potassium Ion-Sensitive FET Transducers: A Preliminary Report," IEEE Tran. Biomed. Eng., vol. BME-25 (1978) 49-54.
7. T. Akiyama, Y. Ujihira, Y. Okabe, T. Sugano, and E. Niki, "Ion-Sensitive Field-Effect Transistors with Inorganic Gate Oxide for pH Sensing", IEEE Trans. Electron Devices, vol. ED-29, pp.1277-1283, 1982.
8. T. Matsuo and M. Esashi, "Methods of ISFET Fabrication", Sensors and Actuators, vol. 1 (1981) 77-96.
9. A. S. Poghossian, V. M. Harootunian, K. S. Karapetian, A. V. Aviazian and A. A. Vartanian, "The Research of Ion-Sensitive Field-Effect Transistors", Microelectronics, vol.17 (1988), 105-112.
10. D. E. Yates, S. Levine and T. W. Healy, "Site-Binding Model of the Electrical Double Layer at the Oxide/Water Interface," J. Chem. Soc. Faraday Trans., vol. 70, pp. 1807-1818, Nov. 1974.
11. J. A. Davis, R. O. James, and J. O. Leckie, "Surface Ionization and Complexation at the Oxide/Water Interface," J. Colloid Interface Sci., vol. 63, pp. 480-499, Mar. 1978.
12. J. Westall and H. Hohl, "A Comparison of Electrostatic Models for the Oxide/Solution Interface," Adv. Colloid Interface Sci., vol. 12, pp. 265-294, 1980.
13. T. W. Healy and L. R. White, "Ionizable Surface Group Models of Aqueous Interfaces," Adv. Colloid Interface Sci., vol. 9, pp. 303-345, 1978.
14. L. Bousse, S. M., B. V. D. Shoot, "Zeta Potential Measurements of Ta_2O_5 and SiO_2 Thin Films", J. Colloid Interface Sci., vol. 147, pp.23-32, 1991.
15. J. C. van Kerkhof, J.C.T. Eijkel and P. Bergveld, "ISFET Responses on a Stepwise Change in Electrolyte Concentration at Constant pH", Sensors and Actuators B, vol. 18-19, pp.56-59, 1994.
16. A. A. Poghossian, "Determination of the pH_{pzc} of Insulators Surface from Capacitance-Voltage Characteristics of MIS and EIS Structures", Sensors and Actuators B, vol. 44, pp.551-553, 1997.
17. C. D. Fung, P. W. Cheung, and W. H. Ko., "A Generalized Theory of an Electrolyte-Insulator-Semiconductor Field-Effect Transistor", IEEE Trans. Electron Devices, vol. ED-33, pp.8-18, 1986.
18. Wang Gui-Hua, Yu Dun and Wang Yao-Lin, "ISFET Temperature Characteristics", Sensors and Actuators B, vol. 11, pp. 221-237, 1987.
19. L. Bousse, N. F. de Rooji, and P. Bergveld, "Operation of Chemically Sensitive Field-Effect Sensors as a Function of the Insulator-Electrolyte Interface", IEEE Trans. Electron Device, vol. ED-30, pp.1263-1270, 1983.
20. K. Chen, M. Nielsen, G.R. Yang, E.J. Rymaszewski, and T.-M. Lu, "Study of Amorphous Ta_2O_5 Thin Films by DC Magnetron Reactive Sputtering", J. Electronic Materials, vol. 26, pp.397-401, 1997.
21. T. Mikolajick, R. Kühnhold, H. Ryssel, "The pH-Sensing Properties of Tantalum Pentoxide Films Fabricated by Metal Organic Low Pressure Chemical Vapor Deposition", Sensors and Actuators B, vol. 44, pp.262-267, 1997.
22. P. V. Bobrov, Yu. A. Tarantov, S. Krause, and W. Moritz, "Chemical Sensitivity of an ISFET with Ta_2O_5 Membrane in Strong Acid and Alkaline Solutions", Sensors and Actuators B, vol. 3, pp.75-81, 1991.
23. Anthony S. Wong, "Theoretical and Experimental Studies of CVD Aluminum Oxide as a pH Sensitive Dielectric for the Back Contacts ISFET Sensor", Department of Biomedical Engineering Case Western Reserve University, Ph. D. Thesis, May 1985.
24. M. Grattarola, G. Massobrio, and S. Martinoia, "Modeling H^+-Sensitive FET's with SPICE", IEEE Trans. Electron Devices, vol. ED-39, pp. 813-819, 1992.
25. P. R. Barabash, R. S. C. Cobbold, and W. B. Wlodarski, "Analysis of The Threshold Voltage and Its Temperature Dependence In Electrolyte-Insulator-Semiconductor Field Effect Transistors (EISFET's)", IEEE Trans. Electron Devices, vol. ED-34, pp.1271-1282, 1987.

pH Response of A-Si:H ISFET

Jung-Chuan Chou*, Jin-Sung Lin
Institute of Electronic and Information Engineering, National Yunlin University of Science and Technology, Touliu, Yunlin, Taiwan 640, R.O.C.

ABSTRACT

The ion-sensitive field-effect transistors (ISFETs) with hydrogenated amorphous silicon (a-Si:H) were fabricated. In this paper, the hydrogenated amorphous silicon acting as sensing membrane was used to investigate the pH response of a-Si:H ISFET. The I_{DS}–V_G curves were carried out by I–V measuring system. The basic parameter of a-Si:H ISFET, namely sensitivity was obtained from I_{DS}–V_G curves. It exhibited a superior pH response of 50.6 mV/pH at temperature of 25℃. Moreover, other characteristic parameters such as hysteresis and drift were proposed. Finally, the effects of operating temperature on sensitivity and drift were presented.

Keywords: ISFET, hydrogenated amorphous silicon, sensitivity, hysteresis, drift

1. INTRODUCTION

In the past few years, many researches have been done to characterize ion-sensitive field-effect transistor (ISFET) based on MOS technology. ISFET was first reported by Bergveld in 1970[1], which is an integrated device composed of a conventional ion selective electrode and a MOSFET. This device is similar to MOSFET except that the metal or polysilicon gate is replaced by sensing membrane and electrolyte solution in order to expose the sensing membrane to solution. When Bergveld first introduced ISFET, the device was operated without a reference electrode. However, later work by other investigators indicates that proper operation of ISFET requires a reference electrode to establish the electrolyte potential with respect to the semiconductor substrate[2]. The schematic representation of an ISFET device with reference electrode and sensing membrane is shown in Fig.1. In recent years, ISFETs have been widely applied in the field of the chemical and biomedical sensors due to the potential advantages over conventional ion-selective electrode in their rapid response, small size, low cost, high input impedance and low output impedance[3-6].

Initial studies of the hydrogenated amorphous silicon were done by Chittick *et al.* in 1969[7], who demonstrated the radio frequency (RF) glow discharge deposition of a-Si:H from silane gas. The hydrogenated amorphous silicon (a-Si:H) can be deposited easily, at low temperature and low cost, on substrates of almost any shape and size. It is widely applied to thin film transistor (TFT), charge-coupled device (CCD) and solar cell, etc.[8,9]. Therefore, in this study we attempt using a-Si:H thin film as the ion-sensitive membrane to investigate the pH response of ISFET.

*Correspondence: Telephone: (8865)5342601 Ext. 2500
Fax: (8865)5312029
E-mail: choujc@pine.yuntech.edu.tw

Part of the SPIE Conference on Advanced Photonic Sensors and Applications
Singapore • November–December 1999
SPIE Vol. 3897 • 0277-786X/99/$10.00

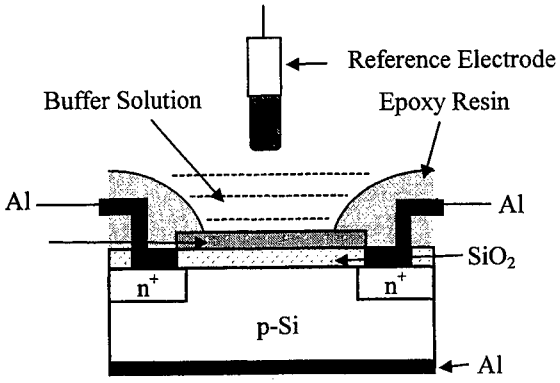


Fig.1 Schematic representation of an ISFET device

Although ISFET devices have been widely applied in recent years, they still have some non-ideal phenomena such as drift , hysteresis and temperature effect, etc.. From experimental results, the output voltage of ISFET monotonically increased with time. Such a phenomenon can be explained by drift behavior[10]. According to the literature[11], drift exists during the measurement. The hysteresis can be observed when ISFET is exposed to consecutive upward and downward pH sweeps. The hysteresis is related to the slow response of ISFET. Some investigators have mathematically derived a relationship between the hysteresis and slow response[12]. In addition, the temperature is the most important factor to ISFET device. Many investigations have demonstrated that ISFET devices showed large thermal instability, which brought about inaccuracy in measurements[13]. As a result, the effects of the temperature on pH responses of ISFET such as sensitivity and drift will be studied.

2. THEORY

In several theories that describe the mechanism of ISFET, the commonly accepted model to explain the pH response of the ISFET is the site-binding model that was first proposed by Yates *et al.* [14]. When ISFET was immersed in electrolyte solution, there was a surface potential at the interface between the sensing membrane and electrolyte solution. The surface potential was dependent on the membrane materials and pH value of the electrolyte solution. The relationship between surface potential (Ψ_0) and pH value is given by following expression[15]

$$2.303(pH_{pzc}-pH)=q\Psi_0/kT+\sinh^{-1}(q\Psi_0/\beta kT) \quad (1)$$

where k is the Boltzmann's constant, T is the temperature of the system in Kelvin, $pH_{pzc}=-\log_{10}(K_a \times K_b)^{1/2}$ is the pH value

at the point of zero charge[16], K_a, K_b are equilibrium constants, and $\beta = 2q^2N_S (K_b / K_a)^{1/2} / (kTC_{DL})$ is the sensitive parameter[17]. N_S is the total number of sites per unit area, C_{DL} is the double layer capacitance[18].

If $q\Psi_0/kT \ll \beta$, Eq. (1) can be given by[15]

$$\Psi_0 = 2.303(kT/q)[\beta/(\beta+1)](pH_{pzc} - pH) \tag{2}$$

From Eq. (2), the sensitivity $S \equiv d\Psi_0/dpH$ can be obtained

$$S = 2.303\ (kT/q)\ [\beta/(\beta+1)] \tag{3}$$

From Eq.(3), at room temperature (T=27℃) and $\beta \to \infty$, the pH sensitivity (S) is theoretically about 59.6mV/pH.

Based on electrochemical theory, the relationship between surface potential and ionic activity (a_{H^+}) of univalent ions in the solutions is given by[19]

$$\Psi_0 = E_0 + \frac{2.303RT}{F} \log a_{H^+} \tag{4}$$

where E_0 is the standard membrane potential, which can be regarded as a constant, R is the gas constant, and F is the Faraday constant.

In the structure of ISFET, the interface potential E_{ref} between reference electrode and solution, and surface potential Ψ_0 between the sensing membrane and solution are included. The work function difference between the metal and the semiconductor (ϕ_{ms}) of MOS structure is replaced by the work function difference between the chemically sensing membrane and the semiconductor (ϕ_{cs}) of EIS structure. Therefore, the drain current-voltage characteristics of an ISFET in the unsaturated region is given by[2]

$$I_{DS} = \frac{\mu_n C_O W}{L}\left\{\left[V_G - E_{ref} + \Psi_0 - \left(\phi_{cs} + \phi_s - \frac{Q_{OX}}{C_O} - \frac{Q_B}{C_O}\right)\right]V_{DS} - \frac{1}{2}V_{DS}^2\right\} \tag{5}$$

where V_G is the reference electrode voltage, Q_B is the charge in the depletion layer of the semiconductor, Q_{OX} is the charge in the oxide, C_O is the capacitance of the gate oxide. Q_B, Q_{OX} and C_O are constants. E_{ref} and E_0 of surface potential (Ψ_0) in Eq.(5) can be also regarded as constants, and $\frac{2.303RT}{F} \log a_{H^+}$ in Eq.(4) is related to ionic activity(a_{H^+}).

If these constants in Eq.(5) are combined, we can obtain

$$I_{DS}=\frac{\mu_n C_O W}{L}\left[\left(V_G+\frac{2.303RT}{F}\log a_{H^+}-V'_T\right)V_{DS}-\frac{1}{2}V_{DS}^{\,2}\right] \tag{6}$$

where $V'_T=E_{ref}-E_0+\phi_{cs}+\phi_s-\frac{Q_{OX}}{C_O}-\frac{Q_B}{C_O}$ is a constant. If $\frac{2.303\mathrm{RT}}{\mathrm{F}}\log a_{H^+}$ of the surface potential (Ψ_0) is in combination with V'_T, threshold voltage (V_T^*) is given by

$$V_T^*=V'_T-\frac{2.303RT}{F}\log a_{H^+}$$

$$=V'_T+\frac{2.303RT}{F}pH \tag{7}$$

Substituting Eq.(7) into Eq.(6), the drain current I_{DS} of the ISFET in the unsaturated region is given by

$$I_{DS}=\frac{\mu_n C_O W}{L}\left[\left(V_G-V_T^*\right)V_{DS}-\frac{1}{2}V_{DS}^{\,2}\right] \tag{8}$$

Likewise, I_{DS} in the saturation region is given by

$$I_{DS}=\frac{\mu_n C_O W}{2L}\left(V_G-V_T^*\right)^2 \tag{9}$$

From Eq.(4) and Eq.(7), it is found that $|\Delta\Psi_0|$ is equal to $|\Delta V^*_T|$. From Eq.(8) or Eq.(9), at a constant drain current I_{DS}, $|\Delta V^*_T|$ is equal to $|\Delta V_G|$, that is, $|\Delta\Psi_0|=|\Delta V_G|$. Therefore, at a constant drain current I_{DS}, the pH sensitivity $|\Delta\Psi_0/\Delta pH|=|\Delta V_G/\Delta pH|$ can be obtained from $I_{DS}-V_G$ curves at different pH values.

3. EXPERIMENTAL

3.1 Thin film deposition

The n-channel ISFET used in the experiment was fabricated on p-type silicon wafer of 8 –12 Ω • cm resistivity, (100)-oriented. The channel was 50 μm long and 1000 μm wide (i.e., W/L = 20). The thickness of the thermally grown SiO_2 layer was 1000 Å. The hydrogenated amorphous silicon thin film was deposited by plasma-enhanced low pressure chemical vapor deposition (PE-LPCVD) at the substrate temperature of 210℃, and silane (SiH_4:H_2=1:9) pressure of 0.7 torr. The

mass flow rate was 120 sccm, and net RF power was 29W. The thickness of the a-Si:H thin film was 2000Å ~3000 Å. The deposition conditions of a-Si:H thin film are shown in Table 1

Table 1 Deposition conditions for a-Si:H thin film

Deposition conditions	
reaction pressure (torr)	0.7
net RF power (W)	29
mass flow rate (sccm)	120
SiH_4:H_2	1:9

3.2 Measurement set-up

Because the a-Si:H thin film would be dissolved in alkaline solutions easily, all measurements were carried out in acidic solutions. Then the I_{DS}–V_G characteristic curves of ISFETs were measured by I–V measuring system of HP 4145B Semiconductor Parameter Analyzer shown in Fig.2. In addition, the hysteresis and drift rate were measured by the measuring system shown in Fig.3.

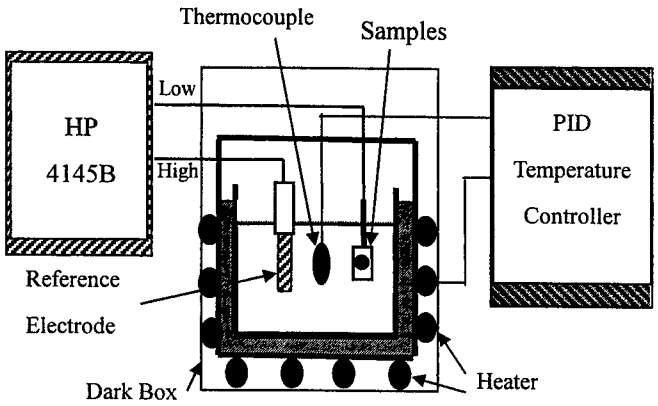


Fig. 2. I–V measuring system

At the beginning, the a-Si:H ISFET was immersed in the electrolyte solutions from pH1 to pH7 at the temperatures of 25℃ to evaluate the pH sensitivity. Moreover, in order to study the effects of operating temperature on sensitivity, a-Si:H ISFET was also operated at different temperatures from 25℃ to 65℃. To evaluate hysteresis behavior, the a-Si:H ISFET was immersed in the hysteresis loop of pH4-pH1-pH4-pH7-pH4. Finally, the measurements of drift rates in different pH solutions of pH2, pH3, pH6, pH7 at 20℃ were achieved. On the other hand, the drift rates at different temperatures of 5℃,

15℃, 25℃, 35℃, 45℃ in pH3 solution were also measured.

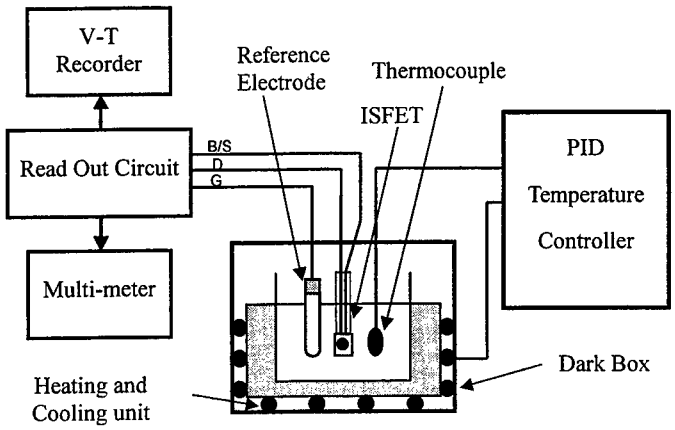


Fig.3 Hysteresis and Drift measuring system

4. RESULTS AND DISCUSSION

The I_{DS}–V_G curves of a-Si:H-gate ISFET at 25℃ are shown in Fig.4. If I_{DS} is fixed at a constant value of 80 μA, the V_G–pH characteristics is shown in Fig.5. From Fig.5, it is found that the sensitivity of a-Si:H ISFET at 25℃ is 50.6 mV/pH. Likewise, the sensitivities at temperatures of 35℃, 45℃, 55℃ and 65℃ are 52.8 mV/pH, 54.2 mV/pH, 58.4 mV/pH and 60.0 mV/pH, respectively, are shown in Fig. 6. Eq.(3) points out clearly that sensitivity (S) is proportional to temperature (T). In our experiments, we have demonstrated that sensitivity increases with increasing temperature. It agrees with the results presented by Vlekkert *et al.*[15] and Bergveld *et al.*[20].

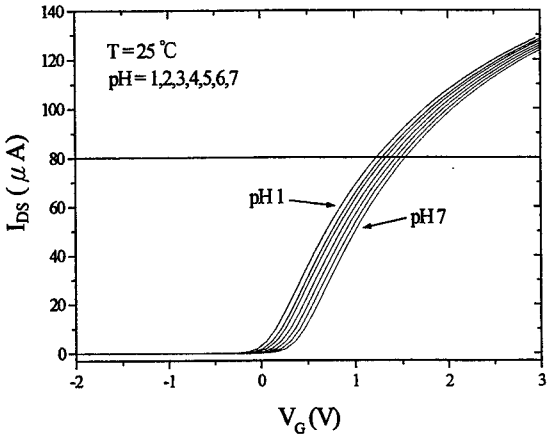


Fig.4 I_{DS}–V_G curves of a-Si:H ISFET at 25℃

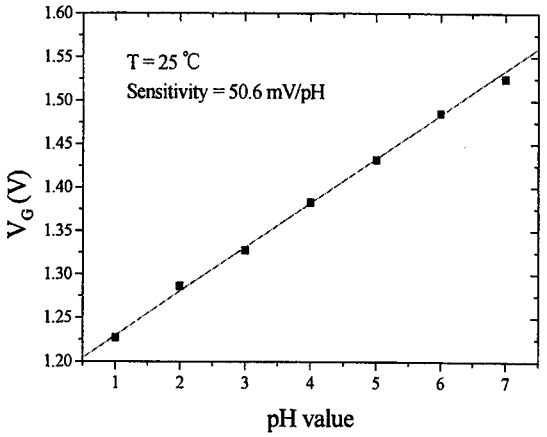


Fig. 5 V_G–pH characteristics of a-Si:H ISFET at 25℃

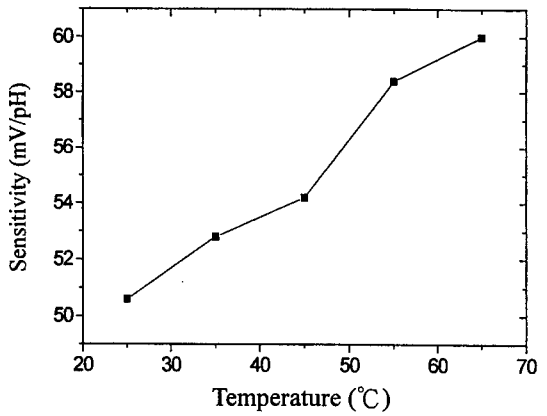


Fig.6 Relationship between temperature and sensitivity

The hysteresis of a-Si:H ISFET in the loop of pH4-pH1-pH4-pH7-pH4 is 3 mV. In addition, the drift behavior is a slow, monotonic, temporal change in the threshold voltage of the ISFET under constant conditions like pH value and temperature[21, 22]. In the experiments, the drift rates in pH2, pH4, pH6 and pH7 at 20℃ were 0.64 mV/h, 1.58 mV/h, 2.17 mV/h and 3.24 mV/h, respectively, shown in Fig.7. It is found that drift rate is increased with increasing pH value. This pH dependence implicates the hydroxyl ion (OH^-) in the drift mechanism[21]. Furthermore, the drift rates of a-Si:H ISFET at temperatures of 5℃, 15℃, 25℃, 35℃ and 45℃ in pH3 solution are 0.09 mV/h, 0.55 mV/h, 2.63 mV/h, 11.38 mV/h and 30.96 mV/h, respectively, as shown in Fig. 8. The measurement data indicate that the higher is operating temperature, the larger is drift rate. It is in agreement with the results presented by P. Hein *et al.*[22]. The reason for this drift behavior seems to be a chemical surface reaction which depends on the temperature[22].

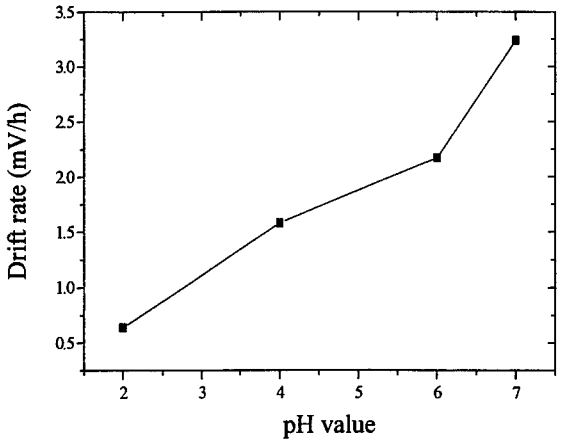


Fig.7 Relationship between drift rate and pH value.

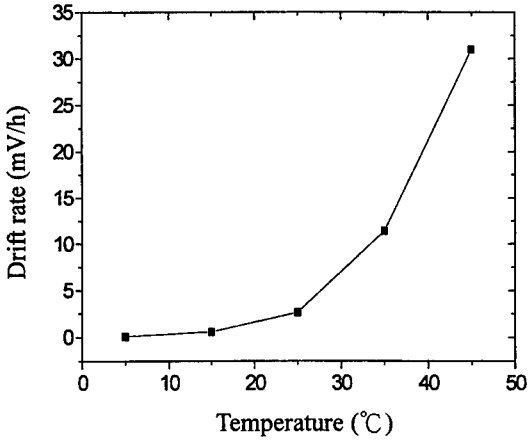


Fig. 8 Relationship between drift rate and temperature.

5. CONCLUSIONS

The pH sensitivity is a decisive factor which affects the response of ISFET device. From the experimental results, the pH sensitivity of a-Si:H ISFET is 50.6 mV/pH at 25℃, it exhibits a superior pH response in contrast with SiO_2 ISFET. From this point of view, a-S:H thin film is suitable for ion-sensitive membrane. In the hysteresis loop of pH4-pH1-pH4-pH7-pH4, the hysteresis of a-Si:H ISFET is 3 mV, which denotes a relatively low magnitude. Moreover, in the experimental drift rates in pH2, pH4, pH6 and pH7 at 20℃ are different values. In other words, drift rate is a function of pH value. Finally, in this paper the effects of operating temperature on sensitivity and drift are presented. It has been demonstrated that sensitivity and drift rate are increased with increasing temperature.

ACKNOWLEDGMENTS

The authors acknowledge the financial support of the National Science Council of Republic of China under grant NSC 89-2215-E-224-003.

REFERENCES

[1] P. Bergveld, "Development of an Ion-Sensitive Solid-State Device for Neurophysiological Measurements", *IEEE Trans. Biomed. Eng.* **17**,pp. 70-71, 1970.

[2] C. D. Fung, P. W. Cheung and W. H. Ko, "A Generalized Theory of an Electrolyte-Insulator-Semiconductor Field-Effect Transistor", *IEEE Trans. Electron Devices* **33**, pp. 8-18, 1986.

[3] A. Topkar and R. Lal, "Effect of Electrolyte Exposure on Silicon Dioxide in Electrolyte-Oxide-Semiconductor Structures", *Thin Solid Films* **232**, pp. 265-270, 1993.

[4] T. Matsuo and M. Esashi, "Method of ISFET Fabrication", *Sensors and Actuators* **1**, pp. 77-96, 1981.

[5] P. Woias, S. Koch, E. Muller, D. Barrow, J. Cefai, G. Curtis and H. Hughes, "An ISFET-FIA System for High Precision pH Recording", *Sensors and Actuators* B **15**, pp. 68-74, 1993.

[6] J.M. Chovelon, J.J. Fombon, P. Clechet, N. Jaffrezic, C. Martelet, A. Nyamsi and Y. Cros, "Sensitization of Dielectric Surfaces by Chemical Grafting: Application to pH ISFETs and REFETs", *Sensors and Actuators* B **8**, pp. 221-225, 1992.

[7] R. C. Chittick, J. H. Alexander and H. F. Sterling, The Preparation and Properties of Amorphous Silicon, J. Electrochem. *Society Solid State Science* **116**, pp. 77-81, 1969.

[8] M. Gatoh, S. Oda, I. Shimizu, A. Seki, E. Tamiya and I. Karube, Construction of Amorphous Silicon ISFET, *Sensors and Actuators* **16**, pp. 55-65, 1989.

[9] A. Pecora, G. Fortunato, L. Mariucci and A. Bearzotti, Chemically Sensitive Hydrogenated Amorphous Silicon Thin-Film Transistors, *J. Non-Crystalline Solids* **137&138**, pp. 1253-1256, 1991.

[10] Y. Dun, W, Ya-Dong and W. Gui-Hua, "Time Dependence Response Characteristics of pH-Sensitive ISFET", *Sensors and Actuators* B **3**, pp. 279-285, 1991.

[11] P. Woias, L. Meixner and P. Fröstl, "Slow pH Response Effects of Silicon Nitride ISFET Sensors", *Sensors and Actuators* B **48**, pp. 501-504, 1998.

[12] L. Bousse, D. Hafeman and N. Tran, "Time-Dependent of the Chemical Response of Silicon Nitride Surfaces", *Sensors and Actuator* B **1**, pp. 361-367, 1990.

[13] Wang Cui-Hua, Yu Dun and Wang Yao-Lin, "ISFET Temperature Characteristics", *Sensors and Actuators* **11**, pp. 221-237, 1987.

[14] D. E. Yates, S. Levine and T. W. Healy, "Site-Binding Model of the Electrical Double Layer at the Oxide/Water Interface", *J. Chem. Soc. Faraday Trans.* **70**, pp. 1807-1818, 1974.

[15] H. van den Vlekkert, L. Bousse and N. de Rooij, The Temperature Dependence of the Surface Potential at the Al_2O_3/Electrolyte Interface, *J. Colloid Interface Sci.* **122**, pp. 336-345, 1988.

[16] G. Massobrio, M. Grattarola, G. Mattioli and F. Mattioli, JR.. "ISFET-based Biosensor Modeling with SPICE", *Sensors and Actuators* B **1**, pp. 401-407, 1990.

[17] D. L. Harame, L. J. Bousse, J. D. Shott and J. D. Meindl, "Ion-Sensing Devices with Silicon Nitride and Borosilicate Glass Insulators", *IEEE Trans. Electron Devices* **34**, pp. 1700-1707, 1987

[18] A. van den Berg, P. Bergveld, D. N. Reinhoudt and E. J. R. Sudhölter, Sensitivity Control of ISFETs by Chemical Surface Modification, *Sensors and Actuators* **8**, pp. 129-148, 1985.

[19] E. J. R. Sudhölter, P. D. van der Wal, M. S. Ptasinska, A. van den Berg and D. N. Reinhoudt, "Ion-Sensing Using Chemically-modified ISFETs", *Sensors and Actuators* **17**, pp. 189-194, 1989.

[20] P. Bergveld and N. F. de Rooij, The History of Chemically Sensitive Semiconductor Devices, *Sensors and Actuators* **1**, pp. 5-15, 1981.

[21] S. Jamasb, S. Collins and R. L. Smith, "A Physical Model for Drift in pH ISFETs", *Sensors and Actuators* B **49**, pp. 146-155, 1998.

[22] P. Hein and P.Egger, "Drift Behaviour of ISFETs with Si_3N_4-SiO_2 Gate Insulator", *Sensors and Actuators* B **13-14**, pp. 655-656, 1993.

Precise measurement of the difference of the air refractive indices between visible and near-infrared wavelengths using two-color interferometer

Ichiro Fujima [a], Shigeo Iwasaki [a], Guangping Xie [b], Katuo Seta [a]
[a] National Research Laboratory of Metrology, Tsukuba 305-8563 JAPAN
[b] Nanyang Technological University, Singapore

e-mail: fujima@nrlm.go.jp

ABSTRACT

Optical measurement technique is widely used in length measurement, such as laser interferometers and electronic distance meters.[1)-4)] The correction of air refractive index is necessary because the optical path length should be converted to the geometrical length in most cases. In order to obtain an average refractive index through the optical path, Edlen's formulae[5), 6)] have been widely used with environmental measurements of air temperature, pressure, humidity and carbon dioxide concentration. In recent days, more precise formulae[7), 8)] for calculating a refractive index of air have been desired in infrared region. However, the measurement accuracy for near infrared region is not sufficient.

In this paper, we have constructed a two-color interferometer[9)] using a Nd:YAG laser to determine the difference of air refractive indices between visible wavelength (532nm) and near infrared wavelength (1064nm). Two non-linear crystals were used to generate the second harmonic wavelength of the YAG laser, while the fundamental YAG laser was also used. After the first non-linear crystal, both the fundamental and the SHG travel and are reflected by a corner cube reflector. The returned beams go through the second non-linear crystal. After the second crystal, two kinds of 532nm beams exist, one by the first crystal, the other by the second crystal. These beams interfere with each other during the movement of the reflector because the optical path lengths between the two nonlinear crystals for the two beams are different. The interference signal has an information about the difference of the air refractive indices between 532nm and 1064nm.

The difference of air refractive indices obtained by our method is compared with the values by Edlen's formulae and Ciddor's formulae.[7)]

Keywords: Air refractive index, Refractive index difference, Two-color interferometer, Nd:YAG laser, NRLM Optical Tunnel, Edlen's formulae, Ciddor's formulae

1. INTRODUCTION

Optical length measurement is widely used in many fields of science and technology. Because the geometrical length is needed instead of optical length in most cases, the measured optical length should be divided by air refractive index for a conversion to the geometrical length. The air refractive index is obtained by substituting the measured environmental values of temperature, pressure, humidity and carbon dioxide concentration through the optical path to Edlen's formulae[5), 6)]. The accuracy of Edlen's formulae is in the order of 10^{-8} to 10^{-9}. The calculated value of air refractive index slightly changes depending on the selection of the conversion formulae such as Edlen's formulae and Ciddor's formulae.[7)] In these days, the improvement of the conversion formula has been required because high accurate length measurement is needed. The international Union of Geodesy and Geophysics (IUGG) has been discussing the improvement of the formulae.[8)] The accuracy of the formula is not so good in infrared region because the number of the measured data for air refractive index is not sufficient in infrared region. Therefore, accurate refractive indices in infrared region are desired.

In this paper, two color interferometer is constructed in a common optical path using visible (532nm) and near infrared (1064nm) wavelengths and the measurement were made up to eight meters. The measurements give the difference of the optical path lengths through the common path for the two wavelengths. The difference in the path lengths corresponds to the difference of the air refractive indices for the two wavelengths. Consequently the refractive index in infrared region (1064mm) can be accurately determined by using the difference and the visible refractive index which had been accurately

obtained. It means that the infrared refractive index can be determined at one point in infrared region. That will contribute to the improvement of the air refractive index formulae.

2. PRINCIPLE

Matsumoto and Zeng proposed a two-color interferometer[9] using two non-linear crystals (KTP). They utilized both fundamental and SHG of Nd:YAG laser. They obtained the refractive index by observing the interference signal from two kinds of SHG beam. An SHG beam is generated by the first non-linear crystal arranged at the transmission part, the other is generated by the second non-linear crystal arranged at the reception part. The refractive index for the SHG beam, n_2, is obtained by the following equation.

$$n_2 - 1 = \frac{Am\lambda_{20}}{2D} \qquad \cdots (1)$$

where A represents Abbe's number for air, m the order of the interference signals by the two kinds of SHGs, λ_{20} wavelength in vacuum for the second harmonic wavelength and D geometrical length. We notice the difference of the refractive indices for the two wavelengths and the following equations has been established.

$$n_2 - n_1 = \frac{m\lambda_{20}}{2D} \qquad \cdots (2)$$

When Eq.1 and Eq.2 are compared with each other, Eq.2 does not have the multiplication of A. It means that the difference of the refractive indices can be accurately determined by Eq.2 because the A coefficient is about 66 for the wavelengths of the fundamental Nd:YAG laser and the SHG. The left side of Eq.(2) is several parts per million for the wavelengths of 1064nm and 532nm around the ordinary environment. Therefore, the relative accuracy of about 3 to 4 order is sufficient even when 10^{-9} figure of the refractive index difference is desired. On the other hand, the right side of Eq.(2) has the form of multiplication and division. The refractive indices difference can be determined with an absolute accuracy of 10^{-9} if m, λ_{20} and D have relative accuracy of 10^{-4}. It is easy to measure the geometrical length D and the wavelength in vacuum for the second harmonic wavelength λ_{20} with relative accuracy of 10^{-4}. It is not easy to get the number of the interference signals by the two kinds of SHGs m with relative accuracy of 10^{-4}, by a short change of the optical path. However, it becomes possible if the path length is several meters and the phase measurement accuracy for the interference signal is several degrees.

3. EXPERIMENTAL SETUP

Figure 1 shows the experimental setup. The light source is a Nd:YAG laser with a wavelength of 1064nm and the optical power of 40mW. The second harmonic beam is generated when the Nd:YAG laser goes through the first non-linear crystal (KTP1). The fundamental beams are expressed by solid lines while SHG beams are shown by broken lines in Fig.1. Both the fundamental and the SHG go through the air to the target of corner cube mirror. The returned fundamental beam is converted to the SHG when it passes through the second non-linear crystal. Thus the two kinds of SHG beams are generated. The two SHG interfere with each other depending on the position of the corner cube mirror when they are detected at PD3 after a pinhole. The optical intensity signal from PD3 makes a period of sinusoidal wave while the corner cube mirror on the movable table changes the position by 67mm. The signal is stored in a wave memorizer. On the other hand, the fundamental beam is introduced to a wave meter made by Anritsu Corporation after separation by a prism for measurement of the wavelength. Meanwhile, an interferometer by the fundamental beam is constructed and the interference signals with the 90 degrees different phases are the inputs to a fringe counter. The interferometer gives a rough value of the geometrical length D. Air temperature, pressure, humidity and carbon dioxide concentration were measured for calculation with the conversion formulae. A He-Ne interferometer shown in the upper side of Fig.1 is subsidiary one in order to confirm the validity of the measurement values by the interferometer of the fundamental YAG laser.

The experimental instruments are arranged in the NRLM Optical Tunnel. The tunnel is a semi-underground tunnel constructed for precise measurements with total length of 310 meters. The annual change of the tunnel temperature is only two degrees peak to peak because the tunnel is covered with soil of 3.5 meters and the tunnel has a double shell structure with several tens of centimeters thick. There are 53 platinum resistance thermometers with an interval of six meters for the accurate temperature measurements through the optical path. Also, three barometers and humidity sensors arranged at the both end and the middle. There is a fifty meters long precise moving table near the entrance of the tunnel. The moving table is utilized for high accuracy measurement linked with a commercial He-Ne laser interferometer.

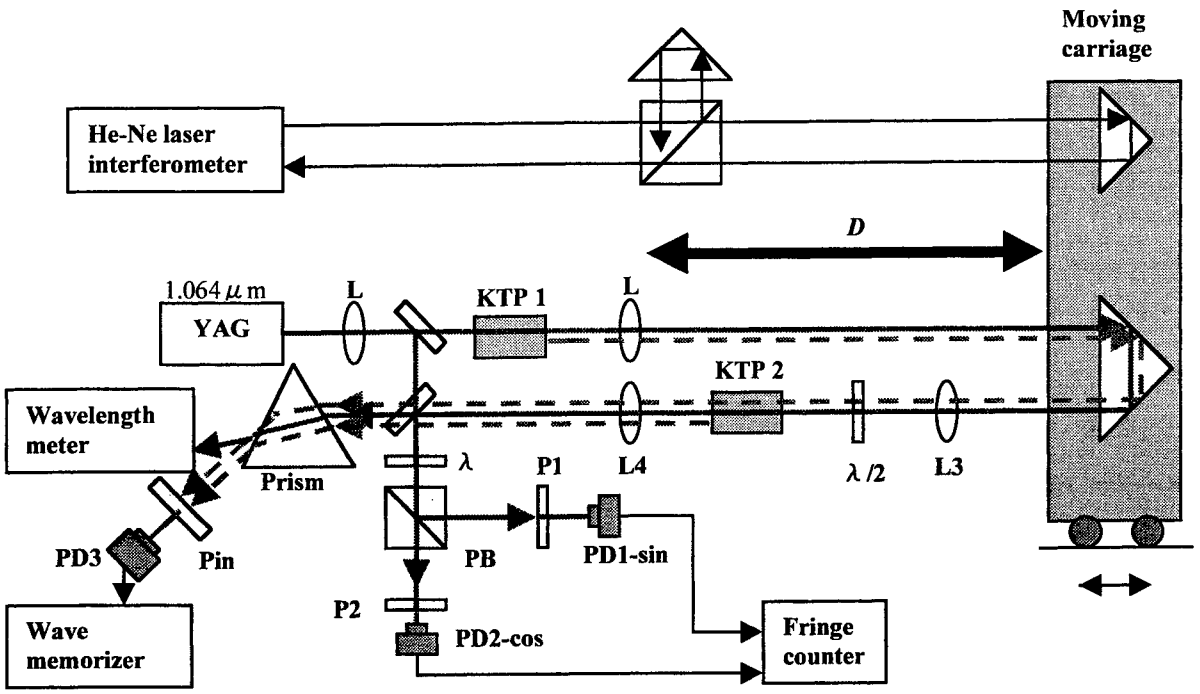


Fig.1 Experimental setup of two-color interferometer

4. DATA PROCESSING

The interference fringe signals of the two SHG beams are stored in the wave memorizer for one record while the moving carriage is moved by 800mm. The measurements were done for ten records with the total displacement of eight meters. Figure 2 shows an example of the measured record. The horizontal axis represents the point number of the data, the vertical axis the voltage of the optical intensity signal.

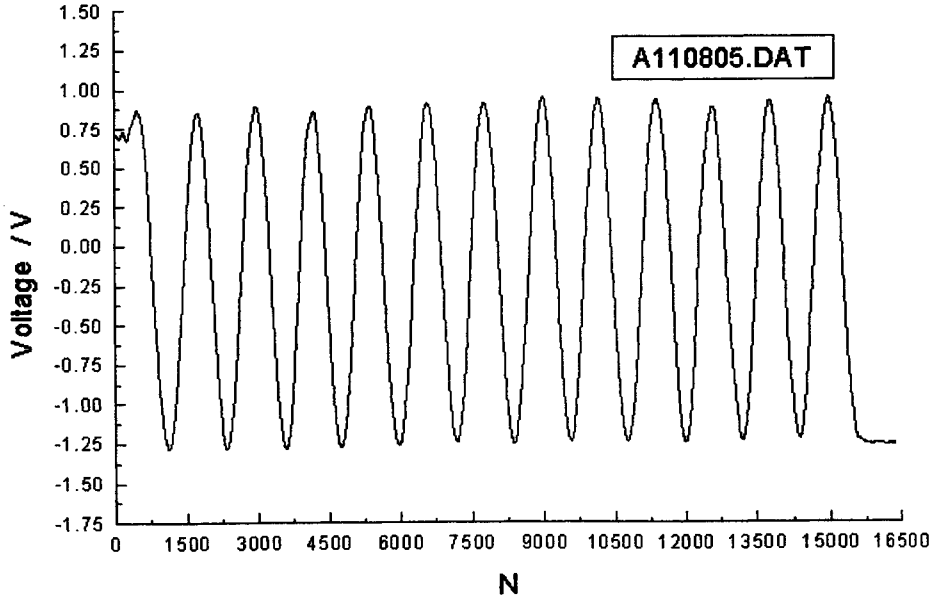


Fig.2 Example of the measured interference fringe signal of the two SHG beams

Because the interference fringe number should be determined with a high accuracy, the phases of the interference signal of 800mm are determined in the beginning area and the end area by using a curve fitting technique. Figure 3 (a) and 3 (b) show the fitting examples in the beginning area and the end area, respectively. The fittings were performed for the following four parameters of a sinusoidal curve by a data analysis software (Origin for Windows).

$$y = m_1 + m_2 \sin(m_3 x + m_4)$$

where m_4 represents the phase of the interference fringe number. It is important to determine the parameter m_4.

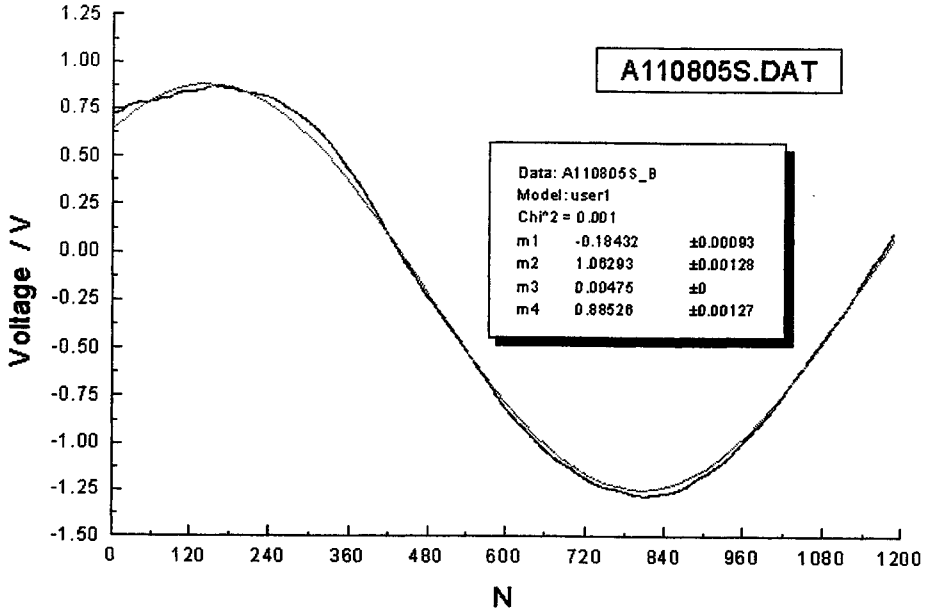


Fig. 3(a) Curve fitting example in the beginning region

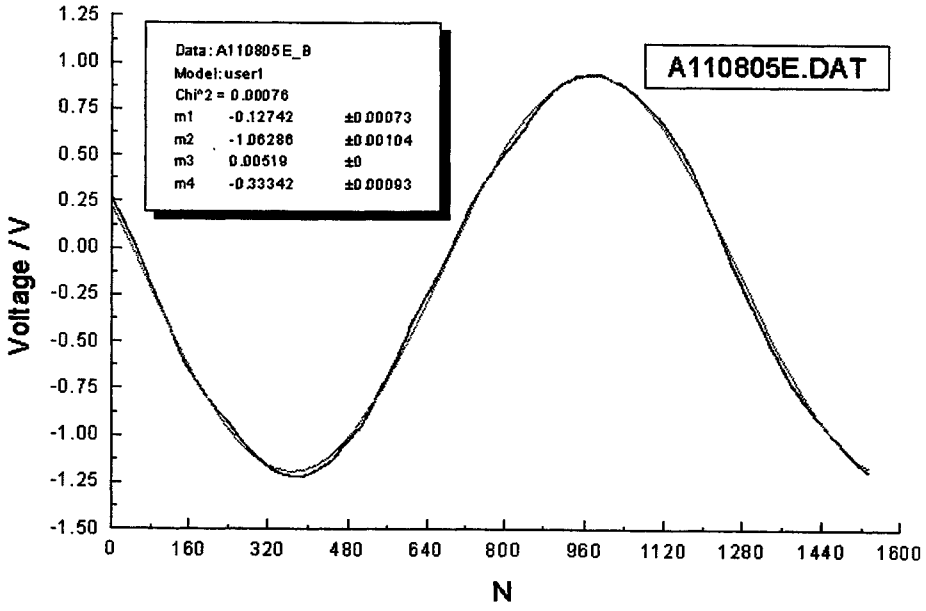


Fig.3(b) Curve fitting example in the end region

5. RESULT AND DISCUSSION

Table 1 shows the difference of the refractive indices between 1064nm and 532nm. The first (left) column represents the displacement of the corner cube mirror. The second column represents the refractive indices difference estimated by our method, the third column and the last column correspond to the refractive indices difference from the calculated values of n_1, n_2 by the Ciddor's formula and the Edlen's formula, respectively. The average values of the indices differences obtained by our method and the two kinds of conversion formulae agree with one another within an absolute value of 10-8. The difference between our method and Ciddor's formulae is 0.002ppm, the difference between our method and Edlen's formulae is 0.005ppm. This result implies that our method has an absolute accuracy of 10^{-9} order because our method agrees with the two calculated results within 0.005ppm and in the middle of the two values although it cannot be confirmed surely. It is difficult to discuss the uncertainty of the our result quantitatively at the moment. However, the uncertainty of the indices difference could become small enough if the prominent error factors listed in the following are considered. The main error factor is in the accuracy of the number of the interference fringe, which has been already mentioned in the chapter 2. In other word, the phase parameter m_4 in the chapter 4 should be determined more accurately. There are two points of improvement. First, the interference data record is limited to 800mm due to the limitation of the wave memorizer and the data acquisition is divided to ten records, each of which has a start and a stop of the moving table. Totally, twenty curve fittings are necessary. It causes an accumulation of the fitting errors. Second, the recorded data have distorted areas in the beginning and in the end region because there are transient states between stillness and constant velocity due to inertia. It makes the fitting accuracy worse. It is possible to improve these two points by using a wave memorizer with bigger storage amount and by getting the optical interference signal of the SHG's and the interference signal from the fundamental interferometer at the same time during the movement of the corner cube mirror.

Displacement /mm	(n2-n1) *10^6 Measured value	(n2-n1) *10^6 Ciddor's Eq.	(n2-n1) *10^6 Edlen's Eq.
799.6916	4.184	4.191	4.183
1601.0716	4.184	4.188	4.182
2400.7612	4.188	4.188	4.182
3201.8299	4.189	4.188	4.181
4002.8166	4.187	4.188	4.180
4803.9317	4.185	4.189	4.182
5605.3817	4.186	4.189	4.183
6406.2528	4.187	4.190	4.183
7207.3462	4.189	4.190	4.184
8009.0617	4.187	4.191	4.184
Average	**4.187**	**4.189**	**4.182**

Table 1 Comparison of the refractive indices difference between 1064nm and 532nm

6. CONCLUSION

Two-color interferometer has been made in the NRLM Optical Tunnel with the maximum arm length of eight meters. The difference of the air refractive indices at wavelengths of 532nm and 1064nm has been estimated by the two-color interferometer. The results have been compared with other calculation by Edlen's formulae and Ciddor's formulae using environmental measurement of temperatures and so on. The average result from two-color interferometer agreed with the other two methods within an absolute value of 0.005ppm. The determination of the difference in the air refractive indices between visible and near infrared region with an absolute accuracy of 10^{-9} will be possible if the phase measurement of the interference signal of SHG becomes more accurately by some modifications of the experiment system.

7. REFERENCES

1. I.Fujima, S.Iwasaki and K.Seta: "High-resolution distance meter using optical modulators at 3.5GHz", *Optical Engineering*, **36** (12) pp.3349-3352 (1997)
2. I.Fujima, S.Iwasaki and K.Seta: "High-resolution distance meter using optical intensity modulation at 28GHz", *Measurement Science and Technology*, **9** (7) pp.1049-1052 (1998)
3. I.Fujima, K.Seta, N.Brown and P.E.Ciddor: "Precision Measurement of Australian EDM Baseline using a He-Ne Laser Beat Distance Meter", *Journal of the Geodetic Society of Japan*, **39** (4), 377-386 (1993)
4. J.M.Rueger: "Electronic Distance Measurement", third totally revised edition, *Springer-Verlag* (1990)
5. B.Edlen: "The refractive index of air", *Metrologia*, **2** (2), pp.71-80 (1966)
6. R.Muijlwijk: "Update of the Edlen's Formulae for the Refractive Index of Air", *Metrologia*, **25**, 189-189 (1988)
7. P.E.Ciddor: "Refractive index of air: new equations for the visible and near infrared" *Applied Optics*, **35** (9) pp.1566-1573 (1996)
8. J.M.Rueger (Convener): "Report of the Ad-Hoc Working Party on REFRACTIVE INDICES OF LIGHT, INFRARED AND RADIO WAVES IN THE ATMOSPHERE of the IAG Special Commission SC3 – Fundamental Constants", (1999)
9. H.Matsumoto and L.Zeng: "High-accuracy measurement of air refractive index using two frequency-doubling crystals", *Optics Communications*, **104**, pp.241-244 (1994)

Author Index

Adachi, Hideo, 658
Anwander, Marcus, 404
Aranchuk, Vyacheslav, 742
Asundi, Anand Krishna, 187, 204, 253, 260, 380, 388, 414, 453, 497, 505, 650, 708
Bao, Zhengkang, 355
Bedi, Navjot Singh, 522
Bergh, Arpad A., 2
Bi, Xijun, 304
Bolte, Michèle, 46, 63
Brennan, Frank, 700
Chai, Gin Boay, 650
Chai, Tai Chong, 650
Chan, Kai Chong, 650
Chan, Peter K. C., 488
Chan, Yuen Ming, 94
Chandra, Dinesh, 139
Chen, Bingsheng, 719
Chen, C. Y., 372
Chen, Haiqing, 304, 734
Chen, Jian, 239
Chen, Jun Wei, 204
Chen, Rongsheng, 79
Chen, Shen-En, 511
Chen, Ying-Chung, 605
Chen, Zhihao, 87, 94
Chen, Zhiping, 728
Chia, Teck-Chee, 124
Chiang, Gilbert, 124
Chiang, Jung-Lung, 605
Chiang, Kin Seng, 87, 94
Chiou, Arthur E., 106
Chollet, Franck A., 380
Chou, Jung-Chuan, 543, 605, 750, 758
Chu, Beatrice Chi Bik, 87
Chung, Wen-Yaw, 543
Cui, Ying, 728
Das, Indu Shekhar, 480
Das, N. B., 534
De Wolf, Ingrid, 239
Demokan, M. Suleyman, 488
Deppe, Dennis G., 12
Ding, Zhihua, 355
Esposito, Enrico, 272
Fang, Yuen Keun, 372
Feyaerts, Filip, 193
Friberg, Ari T., 570
Fujima, Ichiro, 767
Fujiwara, Kenzo, 658
Fukano, Takashi, 342
Gan, Leong Huat, 150
Gan, Yik Yuen, 150
Georgiadis, Christos, 115
Ghailane, Fatima, 46
Ghosh, Anjan K., 522
Graham, Luke A., 12
Gröneweller, Jörg, 328
Gu, Yuzhang, 214
Günther, Bernd, 328
Habib, Khaled J., 450
Hadrboletz, Alfred, 404
Hamid, Hatim Abdul, 700
Hassan, Aseel K., 534
He, Duo Min, 204
He, Jingyi, 296
Hegde, G. M., 692
Ho, Jyh-Jier, 372
Hsieh, M. C., 372
Hsiung, Shen-Kan, 543
Hsu, Long, 686
Huang, Dexiu, 180, 296
Huang, H. H., 12
Huang, Hongxin, 335
Huang, Yajian, 180
Huang, YaoJun, 734
Huang, Zhiwei, 124
Huffaker, Diana L., 12
Hußmann, Stephan, 328
Ishii, Yukihiro, 584
Itoh, Masahide, 335
Iwasaki, Shigeo, 767
Jacobsen, Harald, 424
Jayanth, Puttappa, 56
Jiang, Dazong, 131
Jiang, Li Jun, 453
Jin, Guofan, 666, 674
Jin, Wei, 468, 488
Joblin, Anthony J., 158
John, M. Shelly, 173
Jüptner, Werner P., 314
Kallioniemi, Ilkka J., 570
Kawashima, Kenji, 658
Ke, Hong, 94
Khan, Nasrullah, 578
Kim, Ilmin, 565
Kim, Kwang-Soo, 565
Kim, Myung-Gyoo, 565
Kim, Yohee, 565
Kleuver, Wolfram, 115, 328
Koch, Stephan W., 24
Kothiyal, M. P., 713
Krishnan, S. M., 124
Kudou, Gou, 658
Kuehnert, Rolf, 224

Kumar, Virendra, 139
Lafond, Christophe, 46
Lai, Hsin-Yi, 614
Lam, Yee-Loy, 150, 557, 597, 624, 632
Lan, L. P., 692
Lee, Shih-Ping, 686
Lee, Sing, 124
Leng, Jinsong, 505
Lessard, Roger A., 46, 63
Li, Jianbai, 398
Li, KaiCheng, 728
Li, Min, 79
Li, Xiaoyun, 398
Li, Ying Shin, 750
Liao, Yanbiao, 79
Liaw, Chin-Yi, 557, 597
Liew, Oi Wah, 204
Lin, Jin-Sung, 758
Lin, Ran-Jin, 686
Lin, S. M., 372
Lipowczan, Adam, 442
Liu, Ai Qun, 380
Liu, Baoying, 719
Liu, Deming, 180, 296
Liu, Kuo-Kang, 686
Liu, Xiangdong, 214
Liu, Zhengdong, 355
Lo, Yu-Lung, 552, 614
Lu, Zukang, 355
Luen, Sia Yii, 56
Luo, Sunan, 728
Ma, Jianjun, 708
Ma, Jing, 436
Mei, S. H., 597
Meier, Torsten, 24
Miao, Jianmin, 380
Michel, Bernd, 224
Mok, M. L., 87
Motyka, Zbigniew, 442
Murukeshan, Vadakke M., 388
Muys, Peter F., 640
Nampoori, V. P. N., 173
Narayana, T. Badri, 364
Natarajan S. Ramaswamy, 388
Nemova, Galina A., 640
Ng, Mei Nar, 94
Niinistö, Ari, 570
Oh, Kim Eng, 260, 650
Oki, Masato, 457
Onodera, Ribun, 584
Osten, Wolfgang, 314
Ou, Chaolong, 728
Panigrahi, Simanchalo, 534
Park, Jaehee, 565
Passia, Henryk, 442
Pattnaik, Prasant K., 364
Paul, Prashanta K., 522
Petkov, Ivan, 46
Petro, Samer H., 511
Piazza, Francesco, 272
Pillai, A. S., 36
Pizzocaro, Christine, 63
Pollet, P., 193
Popov, Alexander, 72
Possanzini, Monica, 272
Quan, O. L., 692
Radhakrishnan, P., 173
Rasras, Mahmoud, 239
Rath, Holger, 328
Ray, Asim K., 534
Reinhand, Nadya O., 72
Revel, Gian M., 283
Ruan, Fang, 597
Saarinen, Jyrki, 570
Sabaratnam, Andrew T., 692
Seah, Leong Keey, 497
Selvarajan, Ananth, 364, 480, 692
Semenova, Irina V., 72
Seng, Y. T., 692
Seta, Katuo, 767
Shi, Kun, 719
Simons, Veerle, 239
Sirohi, Rajpal S., 713
Srinivas, Talabuttala, 364, 480
Stiens, Johan H., 640
Suja Helen, S., 713
Sun, Tai-Ping, 543
Sun, ZhongGe, 734
Szade, Adam, 442
Takagi, Tasuku, 457
Takai, Shinji, 658
Takeuchi, Misaichi, 658
Tan, Qiaofeng, 666, 674
Tan, Timothy T. M., 150
Taniguchi, Masanari, 457
Tarte, B., 158
Thomas, Jayan, 173
Tian, Ronglong, 335
Tian, Weijian, 355
Ting, S. F., 372
Tiziani, Hans J., 424
Tomasini, Enrico P., 272, 283
Tork, Amir, 46
Totzeck, Michael, 424
Tsai, Minghong, 335
Unnikrishnan, K. P., 173
Vallabhan, C. P. G., 173
van Doorn, T., 158
Van Gool, Luc J., 193
van Spengen, W. Merlijn, 239
Veeredhi, Vasudeva Rao, 187
Vogel, Dietmar, 224
Vounckx, Roger A., 640
Wah, J. T. K., 692
Wambacq, Patrick, 193
Wang, Chun-Han Wang, 552

Wang, Jinyu, 674
Wang, Wern-Cheng, 614
Wang, Yongjun, 290
Waterworth, Michael D., 158
Weiss, Brigitte, 404
Winkelmann, Kathrin, 453
Wlodarski, Wojtek, 700
Wu, Chongqing, 290
Wu, K. H., 372
Wu, Minxian, 666
Xiao, Shaorong, 398
Xie, Guangping, 497, 767
Xin, Xiangjun, 290
Xu, Yan, 728
Yamaguchi, Ichirou, 342
Yan, GuoPing, 304, 734
Yan, Yingbai, 666, 674
Yang, Kewu, 79
Yang, Xinhui, 131
Yang, Ye, 131
Yang, Yuanhong, 436
Yatagai, Toyohiko, 335
Ye, Miaoyuan, 728
Yeo, Joon Hock, 131
Yi, Benshun, 87
Yin, Li-Te, 543
Ying, Aihan, 398
Yong, T., 150
Yun, Zhisheng, 624, 632
Zagar, Bernhard G., 404
Zhang, Weixu, 436
Zhang, Xiaoling, 398
Zhang, Zhenxi, 131
Zhao, Bing, 253, 260, 414
Zhao, Liping, 624, 632
Zheng, Wei, 124
Zhou, B., 150
Zhou, Yan, 150, 557, 597, 624, 632
Zhu, Yong, 728
Zhuo, Anqing, 398