

AD-A248 517



N/S



Final Report

Development of CASI (Computer Aided Speckle Interferometry) and LSS (Laser Speckle Sensor) with Application to Material Response under High Strain Rate, High Temperature, Vibration and Fatigue

Feb. 10, 199²

by

F.P. Chiang, Leading Professor of
Mechanical Engineering and
Director of Laboratory for Experimental Mechanics Research
State University of New York at Stony Brook
Stony Brook, N.Y. 11794-2300

to

U.S. ARMY RESEARCH OFFICE
P.O. Box 12211
Research Triangle Park, NC. 27709-2211

Grant No. DAAL0388K00³83
April 1, 1988 - Sept. 30, 1991



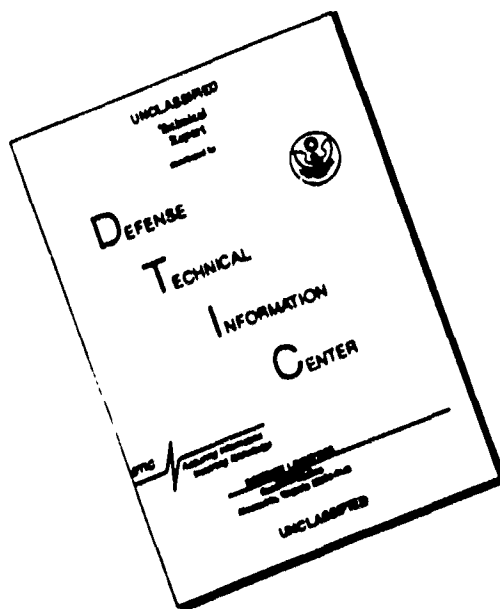
Approved for public release, distribution unlimited

92 4 10 012

92-09279



DISCLAIMER NOTICE



THIS DOCUMENT IS BEST QUALITY AVAILABLE. THE COPY FURNISHED TO DTIC CONTAINED A SIGNIFICANT NUMBER OF PAGES WHICH DO NOT REPRODUCE LEGIBLY.

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188), Washington, DC 20503

1. AGENCY USE ONLY (Leave blank)	2. REPORT DATE Feb. 10, 1991	3. REPORT TYPE AND DATES COVERED Final (1/4/88-30/9/91)
----------------------------------	---------------------------------	--

4. TITLE AND SUBTITLE Development of CASI and LSS with applications to material response under high strain rate, high temperature, vibration and fatigue	5. FUNDING NUMBERS Engineering Science Division Contract No. DAAL0388K0003 3
---	---

6. AUTHOR(S) F.P. Chiang, Leading Professor of Mechanical Engineering & Director of Lab. for Experimental Mechanics Research	
---	--

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Dept. of Mechanical Engineering State University of New York at Stony Brook Stony Brook, N.Y. 11794-2300	8. PERFORMING ORGANIZATION REPORT NUMBER
--	--

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) U. S. Army Research Office P. O. Box 12211 Research Triangle Park, NC 27709-2211	10. SPONSORING/MONITORING AGENCY REPORT NUMBER
---	--

11. SUPPLEMENTARY NOTES
The view, opinions and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy, or decision, unless so designated by other documentation.

12a. DISTRIBUTION/AVAILABILITY STATEMENT Approved for public release; distribution unlimited.	12b. DISTRIBUTION CODE
--	------------------------

13. ABSTRACT (Maximum 200 words)
Two noncontact, remote sensing experimental mechanics techniques using laser speckle are developed. CASI (Computer Aided Speckle Interferometry) is a full field method whereby the speckle formed on the entire specimen surface is captured by a CCD camera and digitally processed to yield stress/strain data. It has been applied to the study of gun oscillation during firing and high strain rate material testing under high temperature and high heating rate. LSS (Laser Speckle Sensor) is a pointwise technique particularly useful for detecting plastic strain, surface texture change due to damage or fatigue.

14. SUBJECT TERMS Laser speckle, NDT technique gun oscillation, high strain rate testing, material damage and fatigue.	15. NUMBER OF PAGES
	16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT UNCLASSIFIED	18. SECURITY CLASSIFICATION OF THIS PAGE UNCLASSIFIED	19. SECURITY CLASSIFICATION OF ABSTRACT UNCLASSIFIED	20. LIMITATION OF ABSTRACT UL
---	--	---	----------------------------------

Final Report

Development of CASI (Computer Aided Speckle Interferometry) and LSS (Laser Speckle Sensor) with Application to Material Response under High Strain Rate, High Temperature, Vibration and Fatigue

Feb. 10, 1992

by

**F.P. Chiang, Leading Professor of
Mechanical Engineering and
Director of Laboratory for Experimental Mechanics Research
State University of New York at Stony Brook
Stony Brook, N.Y. 11794-2300**

to

**U.S. ARMY RESEARCH OFFICE
P.O. Box 12211
Research Triangle Park, NC. 27709-2211**

**Grant No. DAAL0388K00³83
April 1, 1988 - Sept. 30, 1991**

Approved for public release, distribution unlimited

Table of Content

1. Forward	3
2. Statement of Problems Studied	4
3. Summary of the Most Important Results	5
3.1 Introduction	5
3.2 Fundamental Studies of Pattern Recognition and Image Processing	5
3.3 Development of CASI (Computer Aided Speckle Interferometry) and Its Applications	6
3.4 Development of LSS (Laser Speckle Sensor) and Its Applications	8
4 List of Publications	10
5 List of Participating Scientific Personnel	14
6 Appendices - Reprints	15

Accession For	
NTIS GRA&I	<input checked="" type="checkbox"/>
DTIC TAB	<input type="checkbox"/>
Unannounced	<input type="checkbox"/>
Justification _____	
By _____	
Distribution/	
Availability Codes	
Dist	Avail and/or Special
A-1	23

1. Forward

Novel experimental techniques are needed to assess material response at extreme environments such as high strain rate, high temperature, and high heating rate. In these situations the common resistance strain gage technique is not applicable. Photoelasticity (or any birefringent method) is an indirect modeling technique whereby the birefringence of a transparent material under load is related to the response of an actual material. This technique is obviously not applicable to the extreme environments cited above. Moire method (including moire interferometry and grid (grating) methods) requires that an alien piece of material (i.e. grating) be somehow attached or engraved onto the specimen surface. Apart from possible reinforcing or weakening effect to the specimen, this foreign material may itself become disintegrated in the extreme environment. Holographic interferometry has none of the above drawbacks. However, it is not a technique that is suitable for measuring in-plane displacement (and hence strain) except through rather involved optical arrangement and complicated calculation. Since it measures total optical path change between two states of an object the fringes resulting from rigid body movement tends to overwhelm that due to deformation. And its extreme sensitivity also renders it being nonapplicable to finite strain analysis. Furthermore its stringent requirement on vibration isolation essentially prevents it from being used in the field.

The laser speckle method, on the other hand, does not have the restrictions that limits holographic interferometry. Similar to holographic interferometry no foreign material needs to be attached to the specimen except the impingement of photons. But unlike holographic interferometry it requires no stringent vibration isolation and it measures in-plane displacement. It can be applied to finite as well as small deformation; it is not affected by

high temperature and its response is instantaneous (at the speed of light). This project addresses the development of two automated speckle techniques called CASI (Computer Aided Speckle Interferometry) and LSS (Laser Speckle Sensor) and their applications to the study of gun oscillation during firing (jointly with Army Ballistic Research Laboratory), the testing of materials under high strain rate, high temperature and high heating rate (jointly with Army Materials Technology Laboratory), the investigation of the mechanics of plastic deformation and surface roughness, and the evaluation of material damage and fatigue.

2. Statement of the Problems Studied

The aim of the project is to develop quantitative, non-contact, non-destructive and remote sensing techniques for the study of material response under high strain rate, high temperature and fatigue. Two laser speckle methods have been developed to meet this aim. One is CASI (Computer Aided Speckle Interferometry) and the other is LSS (Laser Speckle Sensor) and they have different realms of application. The former is a full field technique more suitable for the mapping of small strain distribution and the latter a pointwise technique more useful for the determination of finite plastic strain. The laser speckle interferometry technique was successfully demonstrated at BRL (Ballistic Research Laboratory) for monitoring the oscillation gun muzzle during firing and at MTL (Material Technology Laboratory) for studying material response under high strain rate, high temperature, and high heating rate. The new speckle sensor technique has been shown to be very effective for measuring plastic strain, surface roughness, material damage and the state of fatigue. The latter has the potential of being developed into a field tool for monitoring material aging. Some important results are summarized in the following section.

3. Summary of the Most Important Results

3.1 Introduction

A total of thirty three papers together two M.S. theses and one Ph.D. dissertation have resulted from this investigation. They may be grouped into the following categories. The first group is a series of six papers concerning the basic study of pattern recognition and image analysis. The second group of eleven papers involves the development of CASI (Computer Aided Speckle Interferometry) and its application. The third group of fourteen papers concerns with the development of LSS and its applications. And there are two miscellaneous papers. One on contouring by moire interferometry and the other on crack tip plasticity of a single crystal.

3.2 Fundamental Studies of Pattern Recognition and Image Analysis [5, 8, 11, 14, 16, 21]¹

In the development of CASI and LSS the laser speckle is digitized into different gray levels using a CCD (charged Couple Device) camera. It is from this large volume of data (e.g. 1000×1000 pixels \times 16 gray levels/pixel) that we need to extract useful information pertaining to specimen deformation. To this end we devoted part of our efforts to some fundamental issues concerning pattern recognition and digital image analysis. This has resulted in six publications. The insight gained in this endeavor has helped materially in the conceptual development of CASI and LSS. In particular the study of 3D moment method [8, 14] has had direct bearing on one of the approaches used in LSS.

¹Number in bracket indicate the papers in the publication list in section 4.

3.3 Development of CASI and Its Applications [1,3,4,7,9,10,17,26,27, 32]

Speckle is the result of multiple interference of a large number of wavelets scattering from an optically rough surface when illuminated by a coherent laser beam. It is formed on the specimen surface as well as in space. Traditionally the practice of one beam laser speckle interferometry (speckle photography) is to photograph the speckle on the specimen surface before and after the application of load to the specimen. The two speckle patterns are superimposed on film via double exposure and subsequently developed (a wet process) in a dark room. The resulting specklegram is then optical Fourier processed using a laser. In the pointwise approach a narrow laser beam probes a point giving rise to a diffraction halo containing Young's fringes directly related to the displacement vector at the probed point. In the full field approach the specklegram is inserted in a Fourier processing optical bench whereby spatial filtering results in an image of the specimen covered with isothetic fringes (contours of equal displacement component). In either case the fringe pattern needs to be recorded again on film and developed (a second wet process). The procedures are tedious, time consuming and error prone.

CASI (Computer Aided Speckle Interferometry) on the other hand, eliminates all this procedures. Instead of recording speckles on film the pattern is directly digitized into gray levels using a CCD (Charged Couple Device) camera. The digitized speckle pattern is processed in a computer using different techniques of digital image analysis and pattern recognition. The result can be either in terms of Young's fringes or isothetic fringes identical to that produced by the photographic process. Or more advantageously, the result can be directly in terms of displacement, strain or stress contours. Furthermore the traditional photographic approach, excessive rigid body displacement between exposures often results

in speckle decorrelation leading to the disappearance of fringes. In CASI, however, rigid body displacement between two speckle patterns is unlimited in practical terms, for the two digital images can always be shifted digitally towards each other for superposition.

The development of CASI evolves gradually. We started out by digital processing the photographically obtained Young's fringe patterns [1]. Then we studied the optimal resolution for correlation calculation for both laser [3] and white light [4] speckle patterns. It finally culminated via a different approach to the development of CASI [20, 27, 28] And we checked the range of applicability of CASI using the well established strain gage and moire technique [33]. CASI may be considered as a major milestone in the development of experimental techniques of stress analysis. For the first time, full field stress/strain information can be obtained without sophisticated instrumentation, without modification of, or attachment to, the surface of a specimen. And the process is fully automated. Many applications await.

While carrying out the fundamental developments of CASI, we concurrently applied the one beam laser speckle interferometry technique to two important problems of Army interest. The first is an application of the technique to the study of gun oscillation during firing. We first tested the approach to investigating the transient vibration of a cantilever beam [7]. After convinced of its accuracy and sensitivity we applied it to the monitoring of muzzle oscillation of a 20 mm Vulcan gun during actual firing. This was done at the Army Ballistic Research Laboratory. The data is still being analyzed for the purpose of developing a computer model. Some data analysis scheme has been developed [10].

The second is the application of laser speckle method to monitoring material response under testing conditions of high strain rate, high temperature and high heating rate [17].

This experiment was carried out at Dr. Tony Chou's laboratory in Army Material Technology Laboratory. Traditional strain measuring methods give an average strain over a certain gage length over which the strain distribution may or may not be uniform. Laser speckle technique is capable of giving the variation of strain over the entire gage length. Being optical the response is instantaneous. And it has been demonstrated that the speckle is not influenced by high temperature and high heating rate. More extensive studies along this line will be pursued by both Dr. Chou of MTL and us.

3.4 Development of LSS (Laser Speckle Sensor) and Its Applications [2, 6, 12, 13, 15, 18, 19, 22, 23, 24, 25, 29, 30, 31]

LSS (Laser Speckle Sensor) is a completely new technique that is developed under this project. A narrow laser beam (instead of an expanded laser beam used in CASI) impinges upon a specimen surface. The scattering laser speckle field which carries the surface texture information is directly digitized using a CCD camera. When the specimen deforms plastically the surface roughness changes resulting in a different scattering pattern. Various image processing techniques such as statistical contrast [6], fractal [15], correlation [18], spectrum analysis [24] and moment [30] have been used to analyze the pattern and to correlate with the deformation. A light scattering theory [12, 23] is developed to predict the optical field and the surface roughness and plastic strain. LSS has been successfully applied to the determination of plastic strain of metals under various loading conditions [2, 18, 23, 24, 25] and the mapping of plastic zone [13, 29]. In the process we also investigated the underlying mechanism that gives rise to surface roughness upon plastic deformation [31]. We find that while there are many factors contributing to the formation of surface roughness, the vertical RMS roughness is largely due to the low frequency signal resulting

from grain rotation. The horizontal roughness parameter in terms of correlation length is found to be linearly proportional to average grain size and becomes saturated at certain plastic deformation. We also find that for the four metals (copper, aluminum, S. steel and h.r. steel) we studied the process of surface roughening under plastic deformation is isotropic and that for a given RMS roughness the total effective strain is the same irrespective of loading path provided that there is no reverse loading [18, 25]. And we find that LSS can be successfully employed as a tool to assess material damage [19,30].

We discover that the conventional surface roughness measurement using mechanical profilometer does not give adequate information concerning surface texture. We find that surfaces produced by different machining processes give rise to distinctly different speckle patterns while their RMS roughnesses as measured by mechanical profilometer are the same. It is expected that further research along this direction may result in better surface texture characterization standards to be used in industry.

Surface roughness produced by fatigue is fundamentally different from that due to simple plastic deformation. And the resulting speckle pattern is also quite different. We have found that we can use the statistical contrast value of the scattered speckle patterns [2] and other image processing parameters to assess a material's stage of fatigue and damage [2, 6, 19, 22, 30]. In one study [2] we were able to predict the initiation of fatigue crack and its propagation path to a considerable accuracy. But in-depth study awaits to be done in order to exploit the full potential of LSS, especially its application to the study of material aging. It is worth emphasizing that LSS is a non-contact technique that can probe a surface area only a fraction of mm^2 and reveals detailed information on surface texture unobtainable by conventional means.

4. List of Publications

- [1] Chen, D.J. and Chiang, F.P., "Digital Processing of Young's Fringe in Speckle Photography", *Optical Engineering*, vol. 29(11), 1413-1420, 1990.
- [2] Chiang, F.P., "Determination of Plastic Strain and Fatigue in Metal and Metal Composite Using Laser Speckles," Invited paper, Army Sym. in Solid Mechanics: "Mechanics of Engineering Materials and Application," Newport, R.I., May 16-18, 1989
- [3] Chen, D.J. and Chiang, F.P., "Optimal Sampling Resolution and Range of Measurement in Digital Speckle Correlation I: Laser Speckle Method," *Proc. Annual Spring Meeting of the Society for Experimental Mechanics*, Cambridge, MA., pp. 133-138, May 29-June 2, 1989.
- [4] Chen, D.J. and Chiang, F.P., "Optimal Sampling Resolution and Range of Measurement in Digital Speckle Correlation II: White Speckle Method," *Proc. Annual Spring Meeting of the Society for Experimental Mechanics*, Cambridge, MA., pp. 289-292, May 29-June 2, 1989.
- [5] Lo, C.H. and Don, H.S., "Representation and Recognition of 3-D Curves," *Proc. Computer Vision and Pattern Recognition '89*, San Diego, CA., June 4-8, 1989.
- [6] Chiang, F.P., Qian, H.H. and Don, H.S., "Determination of Plastic Strain Using Laser Speckles," *Proc. 2nd Int. Sym. on Plasticity and Its Current Applications*, Mie University, Japan, July 31-Aug. 4, 1989.
- [7] Gupta, P.K. and Chiang, F.P., "Laser Speckle Interferometry Applied to Studying Transient Vibration of a Cantilever Beam," *J. of Sound and Vibration*, 133(2), pp. 251-259, 1989.
- [8] Lo, C.H. and Don, H.S., "3-D Moment Forms: Their Construction and Application to Object Identification and Positioning," *IEEE Trans, on Pattern Analysis and Machine Intelligence*, Vol. 11 No. 10, pp. 1053-1064, Oct. 1989.

- [9] Chen, D.J. and Chiang, F.P., "Investigation of Optimal Sampling Resolution in Digital Laser Speckle Correlation," *Experimental Mechanics*, (in press).
- [10] Gupta, P. and Chiang, F.P., "Resolution of Resultant Displacement into Components in Double Exposure Speckle Photography," *Applied Optics*, Vol. 29, No. 11/10, pp. 1642-1645, April, 1990.
- [11] Chen, J.C. and Don, H.S., "Segmentation of Bilevel Images Using Mathematical Morphology," *Proc. of Vision Interface '90 Conf. Halifax, Nova Scotia, Canada, May 14-18, 1990.*
- [12] Dai, Y.Z. and Chiang, F.P., "Application of Scattering Theory to Plastic Strain Estimation," *Proc. IV Int. Sym. on Non-Destructive Material Characterization, Annapolis, MD., June 11-14, 1990, (in press)*
- [13] Chiang, F.P., Tay, J. and Dai, Y.Z., "Determination of Elastic-Plastic Boundary by Speckle Pattern Correlation," *Proc. Int. Sym., on Non-Destructive Material Characterization, Annapolis, MD., June 11-14, 1990, (in press).*
- [14] Lo, C.H. and Don, H.S., "Pattern Recognition Using 3D Moments," *10th ICPR, Atlantic City, June 16-21, 1990.*
- [15] Dai, Y.Z. and Chiang, F.P., "Estimation of Plastic Strain by Fractal," *Proc. SPIE 1990 International Symposium on Optical & Optoelectronic Applied Science and Engineering, San Diego, CA., 1332-85, July 8-13, 1990.*
- [16] Lo, C.H. and Don, H.S., "Recognition of Range Data Using Distance Measure of Attributed Graphs," *Proc. Int. Conf. Automation, Robotics & Computer Vision, Singapore, Sept. 18-21, 1990.*
- [17] Emslie, J., Green, J., Chou, S.C. and Chiang, F.P., "The application of Laser Speckle Interferometry to Measuring Strain and Strain Rate under Dynamic Loading Conditions," *Proc. Soc. of Exp. Mach., 1990 Fall Meeting, Baltimore, MD., pp. 59-66, Nov. 4-7, 1990.*

- [18] Chiang, F.P., Dai, Y., Xu, B. and Kato, A., "Study of Surface Roughening under Different Stress Modes by Correlation," Proc. Soc. of Exp. Mech., Fall Meeting, Baltimore, MD., pp. 36-42, Nov. 4-7, 1990.
- [19] Dai, Y.Z. and Chiang, F.P., "Damage Monitoring of Composite Material by Image Processing," Experimental Technique, 39-441, July/Aug., 1990.
- [20] Chen, D.J. and Chiang, F.P., "Computer Speckle Interferometry," Proc. Soc. of Exp. Mech., 1990 Fall Meeting, Baltimore MD., pp. 49-58, Nov. 4-7, 1990, (Full paper to appear in Applied Optics).
- [21] Wang, C.C. and Don, H.S., "A New Approach to Evidential Reasoning," Proc. ISMM Int. Conf. Computer Applied to Design, Simulation & Analysis, Las Vegas, Mar. 19-21, 1991.
- [22] Dai, Y.Z. and Chiang, F.P., "Fatigue Monitoring by Laser Speckle," Int. J. of Fatigue, Vol. 13, No. 3, 1991.
- [23] Dai, Y.Z. and Chiang, F.P., "Scattering from Plastically Roughened Surfaces and Its Applications to Mechanics," Optical Engineering, Vol.30, No. 9, pp. 1269-1276, 1991.
- [24] Dai, Y.Z. and Chiang, F.P., "Assessment of Flow Stress, Plastic Strain by Spectrum Analysis," Experimental Mechanics, Vol. 31(3), pp. 197-200, Sept. 1991.
- [25] Dai, Y.Z. and Chiang, F.P., "Strain Path and Surface Roughness," Mechanics of Material, 1991, (in press).
- [26] Dai, Y.Z. and Chiang, F.P., "Contouring by Moire Interferometry," Experimental Mechanics, pp. 76-81, Mar. 1991.
- [27] Chen, D.J., Chiang, F.P., Tan, Y.S. and Don, H.S., "Computer Aided Speckle Interferometry (CASI) II: An Alternate Approach Using Differential Phase Information and Neighbor Prediction," Proc. SPIE 2nd Int. Conf. on Photomechanics and Speckle Metrology, San Diego, CA., July 22-26, 1991. (Full paper submitted to Applied Optics)

- [28] Chen, D.J. and Chiang, F.P., "Range of Measurement of Computer Aided Speckle Interferometry (CASI)," Proc. SPIE 2nd Int. Conf. on Photomechanics and Speckle Metrology, San Diego, CA., July 22-26, 1991.
- [29] Chiang, F.P. and Dai, Y.Z., "Determination of Plastic Zone at Crack Tip Using Laser Speckle Decorrelation," Proc. Plasticity '91, 3rd Int. Symposium on Plasticity, Genoble, France, Aug. 12-16, 1991, (in press).
- [30] Kato, A., Dai, Y.Z. and Chiang, F.P., "Damage Monitoring of Metal Materials by Laser Speckle Assisted Image Processing Technique," JSME Int. Journal, Series I Vol. 34, No. 3, pp. 374-380, 1991.
- [31] Dai, Y.Z. and Chiang, F.P., "Plastically Induced Surface Roughness: Phenomena and Mechanism," Army Symposium on Solid Mechanics, Nov. 4-7, 1991, (in press).
- [32] Chiang, F.P. and Li, X.M., "Experimental Measurement of Crack Tip Strain Field in a Single Crystal," Proc. MECAMAT '91 Int. Seminar on Large Plastic Deformations, Fontainbleau, France, Aug. 7-9, 1991, (in press).
- [33] Wang, Y.Y., Chen, D.J. and Chiang, F.P., "Material Testing by Computer Aided Speckle Interferometry," submitted to Experimental Technique.

Dissertation and Thesis Published

D.J. Chen, "Digital Processing of Young's Fringe in Speckle Photography," M.S. Thesis, Dept. of Mech. Eng., SUNY at Stony Brook, NY., May, 1990.

J. Emslie, "Application of Laser Speckle Interferometry to Measuring Dynamic Strain", M.S. Thesis, Dept. of Mech. Eng., SUNY at Stony Brook, NY., May, 1990.

Y.Z. Dai, "On Plastic Deformation Induced Surface Roughness of Metals," Ph.D Dissertation, Dept. of Mech. Eng., SUNY at Stony Brook, NY., May, 1991.

5. List of Participating Scientific Personnel

Principal Investigator: F.P. Chiang

Co-investigator: H.S. Don

Graduate Student:

Y.Z. Dai (earned Ph.D. degree)

D.J. Chen (earned M.S. degree)

J. Emslie (earned M.S. degree)

J.C. Chen, P. Gupta, X.M. Li, C.H. Lo, H.H. Qian, C.C. Wang, Y.Y. Wang

Visiting Scientists:

A. Kato, J. Tay, Y.S. Tan, B. Xu

Army Laboratory Scientists:

J. Green, S.C. (Tony) Chou

6. Appendices

Digital processing of Young's fringes in speckle photograph.

D. J. Chen
F. P. Chiang, MEMBER SPIE
State University of New York
Department of Mechanical Engineering
Laboratory for Experimental
Mechanics Research
Stony Brook, New York 11794-2300

Abstract. A new technique for fully automatic diffraction fringe measurement in pointwise analysis of speckle photographs is presented. The fringe orientation and spacing are initially estimated with the help of a 1-D fast Fourier transform (FFT). A 2-D convolution filter is then applied to enhance the estimated image. A fringe pattern with high signal-to-noise ratio is achieved, which makes feasible precise determination of the displacement components. The halo effect is also optimally eliminated in a new way by a halo-division technique. High reliability and accurate determination of displacement components are achieved over a wide range of fringe densities, with the computation time comparing favorably with those for the 2-D autocorrelation method and the iterative 2-D FFT method.

Subject terms: speckle photography; diffraction fringe pattern; halo effect; fast Fourier transform; spatial domain; spectral domain; convolution; Hanning window.

Optical Engineering 29(11), 1413-1420 (November 1990).

CONTENTS

1. Introduction
2. Description of the system
3. Diffraction fringe pattern and related spectral analysis
 - 3.1. Diffraction fringe pattern
 - 3.2. Spectral analysis
4. Procedure of the program
 - 4.1. Initial estimation of fringe spacing and orientation
 - 4.2. Enhancement: spatial convolution method
 - 4.3. Elimination of the halo effect
 - 4.4. Determination of fringe spacing and orientation
5. Experimental results
6. Conclusions
7. Acknowledgments
8. References

1. INTRODUCTION

Double-exposure speckle photography is a well-established whole-field technique for measuring in-plane displacement and surface rotation.^{1,2} In its basic form, the object to be studied is illuminated by a divergent laser beam or white light beam and is imaged by a camera onto film. By double exposure, with the specimen being deformed between the exposures, two speckle patterns are recorded on the same film. In the analysis, the in-plane displacement vector of any point on the object surface can be determined by measuring the separation of the two speckle patterns at the corresponding point on the developed film. This is normally done by probing the photograph with a narrow laser beam. The far-field diffraction pattern consists of speckled cosine-squared fringes similar to Young's fringes, modulated by a diffraction halo (Fig. 1). The fringes are perpendicular to the direction of the displacement vector $d(u, v)$ and have a spacing

S that is inversely proportional to the magnitude of the displacement, i.e.,

$$d = \frac{\lambda L}{M S} \quad (1)$$

where λ is the wavelength of the light source, L is the effective distance between the specklegraph and the diffraction plane, and M is the magnification of the recording system. The displacement components can be obtained from

$$u = d \cos \alpha, \quad v = d \sin \alpha \quad (2)$$

where α is the angle between the displacement vector and the x -axis. Thus, by measuring the spacing and direction of the fringe pattern on a square mesh covering the photograph, one can obtain the complete 2-D displacement field.

The evaluation of these parameters by eye is time consuming and relies heavily on the operator's subjective skills. For this reason, several automatic fringe analysis systems have been developed that combine an optical system with digital signal processing. The main difficulty that any such system must overcome is the noise in the diffraction halo.

There are two basic approaches to this problem. In one, fringes are integrated along the fringe direction, thereby reducing the amount of data considerably, from a 2-D pattern to a 1-D signal.³⁻⁵ Recently, the trend has been toward the development of image processing algorithms capable of detecting and measuring the 2-D image. Some methods achieve good noise immunity and accurate determination of displacement components.⁶⁻⁸ Others apply simple algorithms that remarkably reduce the computing time.⁹⁻¹¹

In this paper, we describe a fully automated technique in which the fringe pattern is first digitized by a video camera and then transmitted into a VAX-11/730 computer. Fringe spacing and orientation are initially estimated by means of a 1-D fast Fourier transform (FFT). An optimum enhancement filter is chosen based on the 2-D spectral analysis of the fringe pattern. For each fringe pattern, a suitable 2-D convolutional Hanning win-

Paper 2619 received July 25, 1988; revised manuscript received April 10, 1990; accepted for publication April 26, 1990. This paper is a revision of paper 954-43, presented at the SPIE conference Optical Testing and Metrology II, June 27-30, 1988, Dearborn, Mich. The paper presented there appears (unreferenced) in SPIE Proceedings Vol. 954.
© 1990 Society of Photo-Optical Instrumentation Engineers.



Fig. 1. Young's fringe pattern.

low is applied according to the initial estimation. A high signal-to-noise ratio (SNR) fringe pattern is achieved, making feasible a high accuracy determination of fringe spacing along two orthogonal scanning directions. Furthermore, because of the varying irradiance of the diffraction halo, fringe shifting has a significant effect on the fringe spacing determination¹²; therefore, we introduce a divisional halo-elimination technique before enhancement. A further reduction in computing time is achieved by modulating each Hanning window into an integral window. The performance of this program is tested on a rigid-body rotation of a disk of 82 mm diameter. High reliability (100% success rate down to fringe visibilities of 10%) and accurate determination of displacement components (0.04 μm and 0.12° in standard deviation) are achieved over a wide range of fringe densities of 3 to 50 per frame pattern. The computing time compares favorably with those of the 2-D averaged autocorrelation and the iterative 2-D FFT methods.^{6,7} Average processing time for each fringe pattern is about 25 s.

2. DESCRIPTION OF THE SYSTEM

The system is shown in Fig. 2. The specklegram is mounted on a 2-D motor-driven stage to be scanned by a laser beam. The impinging laser has a spot size of about 0.6 mm. It produces a diffraction pattern of Young's fringes modulated by a halo function on the viewing screen. The fringe pattern is captured by a TV camera with a frame resolution of 256 \times 256 pixels. The viewing screen is a rectangular glass plate, with the undiffracted part of the laser beam blocked by a central stop on the screen. The light intensity of the fringe pattern is first digitized by an analog-to-digital converter (ADC) and then transmitted into a VAX-11/730 computer through an 8-bit I/O port. Step motor controllers are operated either manually or automatically by the given program through another I/O port of the computer.

3. DIFFRACTION FRINGE PATTERN AND RELATED SPECTRAL ANALYSIS

3.1. Diffraction fringe pattern

Ideally, the intensity of the diffraction fringe pattern has the form²

$$f(x, y) = I_0(x, y) \left[1 + V \cos \left(\frac{2\pi r \cdot d}{\lambda L} \right) \right] \quad (3)$$

where $r = (x, y)$ is the position vector on the diffraction screen, $d = (u, v)$ is the displacement vector on the specklegram, λ is the wavelength of the light source, L is the distance between the specklegram plane and the diffraction plane, V is the fringe visibility of the diffraction pattern, and $I_0(x, y)$ is the diffraction halo, which is theoretically given by²

$$I_0(x, y) = \left\{ \arccos \left(\frac{qr}{DL} \right) - \left(\frac{qr}{DL} \right) \left[1 - \left(\frac{qr}{DL} \right)^2 \right]^{1/2} \right\}^2 \quad (4)$$

where $r = \sqrt{x^2 + y^2}$, D is the diameter of the recording lens, and q is the recording image distance. A sectional view of diffraction Young's fringe pattern is shown in Fig. 3, wherein the unit of the x -coordinate is the discrete pixel.

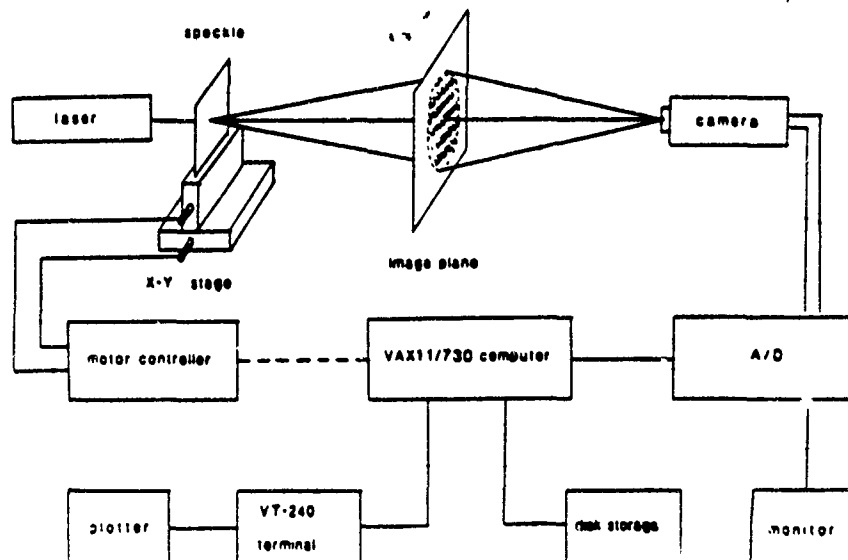


Fig. 2. System setup.

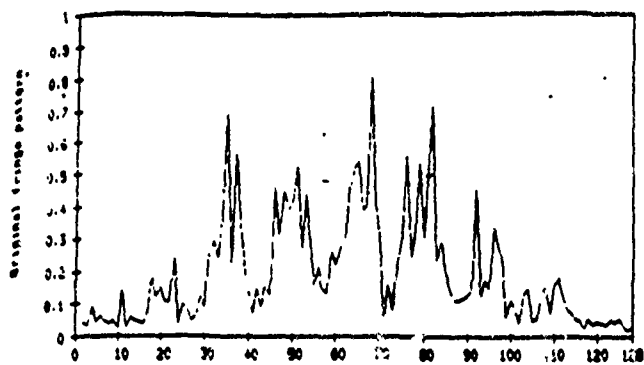


Fig. 3. Sectional view of original fringe.

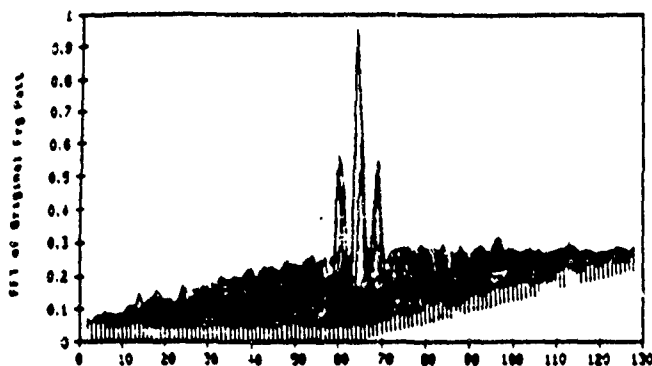


Fig. 4. Spectrum of original fringe.

3.2. Spectral analysis

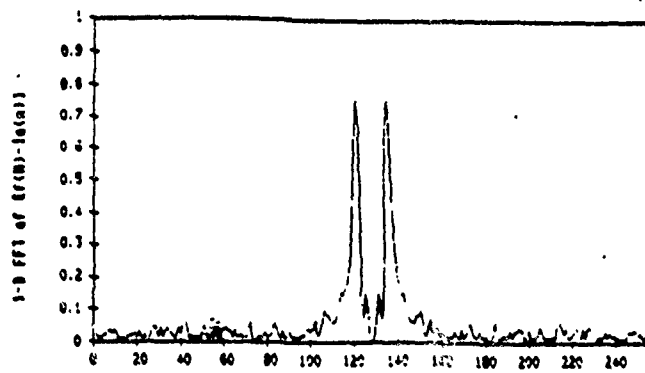
A discrete spectrum of the digitized Young's fringe pattern is carried out from the following discrete Fourier transform by use of the FFT algorithm:

$$F(k_x, k_y) = \sum_{m=1}^N \sum_{n=1}^N f(m, n) \exp \left[-\frac{j2\pi(mk_x + nk_y)}{N} \right], \quad k_x, k_y = 1, 2, \dots, N, \quad (5)$$

where m and n are discrete coordinates on the diffraction pattern, which is referred to as the spatial domain, and k_x and k_y are discrete coordinates on the transformed pattern, which is referred to as the spectral domain. Figure 4 shows the spectrum of the diffraction fringe pattern of Fig. 1. It is seen that the Young's fringe pattern carries a great amount of high frequency noise, which spreads over the whole spectral domain.

4. PROCEDURE OF THE PROGRAM

There are four steps in the processing of the diffraction Young's fringe pattern. First, two 1-D signals along two orthogonal directions are extracted from the 2-D Young's fringe pattern, and the fringe spacing and orientation are roughly estimated based on the peak positions of the spectra of those 1-D signals. Second, we introduce a halo-division technique to eliminate the fringe shifting effect resulting from the halo function. In the third step, the fringe enhancement process is applied by convolving a proper Hanning window to the 2-D fringe pattern. The window size and orientation are determined by the fringe spacing and orientation found in the previous step. The last stage of the program


 Fig. 5. $F(k_x)$, the 1-D FFT of $P(m)$.

is the determination of the fringe spacing and orientation, which is basically a 2-D scanning process over the enhanced fringe pattern.

4.1. Initial estimation of fringe spacing and orientation

An FFT algorithm is applied to the two 1-D signals, obtained from the original fringe pattern along two orthogonal directions, the x -axis and y -axis. The operations are given by

$$F(k_x) = \sum_{m=1}^N f(m) \exp \left(-\frac{j2\pi mk_x}{N} \right), \quad k_x = 1, 2, \dots, N, \quad (6)$$

$$F(k_y) = \sum_{n=1}^N f(n) \exp \left(-\frac{j2\pi nk_y}{N} \right), \quad k_y = 1, 2, \dots, N, \quad (7)$$

where $f(m)$ and $f(n)$ are one-dimensional signals picked up from the diffraction pattern $f(m, n)$ along the x -axis and y -axis and $F(k_x)$ and $F(k_y)$ are 1-D spectra of $f(m)$ and $f(n)$, respectively. It is seen that each of the 1-D spectra, for instance, $F(k_x)$, consists of a broadened delta function at $k_x = 0$ and two conjugate broadened delta functions at fringe frequencies $k_x = k_m$ and $k_x = -k_m$. Because of the undesirable effect of the broadened zero-order peak on the signal peak detection, especially for fringe patterns with lower visibilities, the low frequency peak is removed by subtracting $I_0(m)$ from $f(m)$, and $I_0(n)$ from $f(n)$, before transformation, where $I_0(m)$ and $I_0(n)$ are 1-D signals extracted from the 2-D diffraction halo function $I_0(m, n)$, which is evaluated by smoothing the average of several diffraction halos on different points of a single-exposure photograph. The Fourier transform of $f(m) - I_0(m)$, denoted by $f'(m)$, is shown in Fig. 5. From the detected maximum location on the spectra of the 1-D signals, initial values of fringe orientation and spacing are obtained. Since the detected maximum locations are restricted by the pixel resolution of the digital image, the fringe spacing and orientation thus obtained are not the accurate final solutions. But this rough estimation provides sufficient information about the fringe pattern for the later enhancement process.

In practice, two additional 1-D signals along two other orthogonal axes are used to achieve highly reliable estimation. For example, we consider the fringe pattern in Fig. 1. The maximum values in the corresponding spectrum $F'(k_x)$ are too close to the center of the spectrum to be detected. Therefore, the detected maximum locations along the k_y direction are not reliable. In this case the peak locations in spectrum $F'(k_x)$ are determined by measuring the maximum locations in spectra of two other 1-D signals picked out along axes at $+45^\circ$ and -45° with respect to the x -axis.

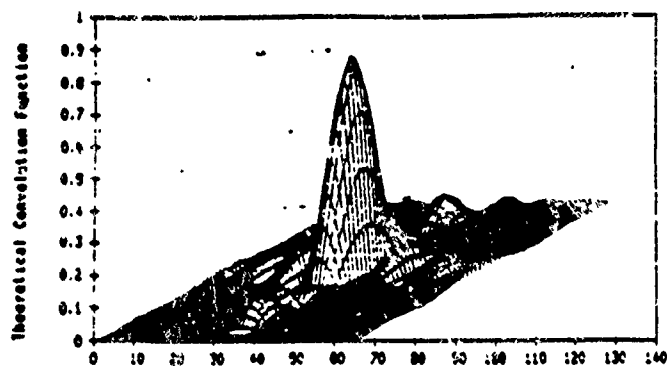


Fig. 6. Theoretical convolution function.

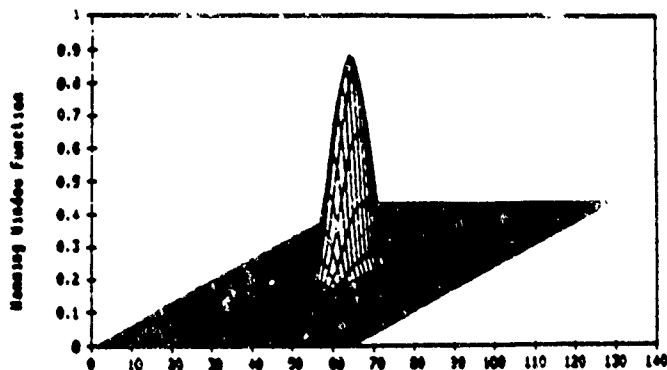


Fig. 7. 2-D Hanning window.

The computation for the above initial estimation is $4N \log_2 N$ multiplications and $N^2 + 4N \log_2 N$ additions, requiring about 3 s. For later use, let the four estimated peak position values of the four 1-D spectra be marked in a decreasing way, i.e., $k_{max1}, k_{max2}, k_{max3}$ and k_{max4} .

4.2. Enhancement: spatial convolution method

It is evident from the spectral distribution (Fig. 4) that the signal is concentrated near the region of $k_x = +k_m, k_y = +k_n$ and $k_x = -k_m, k_y = -k_n$, whereas the noise is spread over the whole spectral domain. Thus, by applying a 2-D rectangular pupil function in the spectral domain, we considerably enhance the SNR. The pupil function can be expressed as

$$H(k_1, k_2) = \begin{cases} 1, & |k_1| \leq K_1, |k_2| \leq K_2, \\ 0, & \text{elsewhere} \end{cases} \quad (8)$$

where k_1 and k_2 are discrete coordinates on the spatial domain along directions of initially estimated k_{max1} and k_{max2} , and

$$K_1 = k_{max1} + \eta k_w, \quad K_2 = k_{max2} + \eta k_w \quad (9)$$

where k_w is the width of the scattered delta functions in the spectral domain (Fig. 4) and η is a proportional constant whose theoretical value is 0.5 and is chosen as 2.5 in our conservative window design. For convenience, we assume that the k_1 and k_2 coordinates are along the k_x and k_y axes, respectively.

As we know, convolving two images in the spatial domain is equivalent to multiplying the Fourier transforms of the two images in the spectral domain; i.e., if

$$F'(k_x, k_y) = F(k_x, k_y)H(k_x, k_y), \quad k_x, k_y = 1, 2, \dots, N \quad (10)$$

then

$$f'(m, n) = \sum_{p=-N/2}^{N/2} \sum_{q=-N/2}^{N/2} f(m-p, n-q)H(p, q) \quad (11)$$

$m, n = 1, 2, \dots, N$

where p and q are formal integral parameters with respect to m and n ; $h(m, n)$ and $H(k_x, k_y)$ are operation functions in the spatial domain and the spectral domain, respectively; $F'(k_x, k_y)$ is the resultant spectrum; and $f'(m, n)$ is the resultant image. $H(k_x, k_y)$, $F(k_x, k_y)$, and $F'(k_x, k_y)$ are spectra of $h(p, q)$, $f(m, n)$, and $f'(m, n)$, respectively, and are given by

$$F(k_x, k_y) = \sum_{m=1}^N \sum_{n=1}^N f(m, n) \exp \left[-\frac{j2\pi(mk_x + nk_y)}{N} \right]$$

$$H(k_x, k_y) = \sum_{m=1}^N \sum_{n=1}^N H(m, n) \exp \left[-\frac{j2\pi(mk_x + nk_y)}{N} \right] \quad (12)$$

$$F'(k_x, k_y) = \sum_{m=1}^N \sum_{n=1}^N f'(m, n) \exp \left[-\frac{j2\pi(mk_x + nk_y)}{N} \right]$$

Theoretically, the Fourier transform of the rectangular pupil function $H(k_1, k_2)$ is given by

$$h(m, n) = \frac{\sin(2\pi m K_1 / N)}{2\pi m K_1 / N} \frac{\sin(2\pi n K_2 / N)}{2\pi n K_2 / N} \quad (13)$$

where $h(m, n)$ is the spectrum of the pupil function (see Fig. 6) and m and n are discrete coordinates in the spatial domain corresponding to the spectral coordinates k_1 and k_2 , respectively.

However, from the figure we know that $h(m, n)$ is a spatially widespread function, which makes the convolution calculation very time consuming. Therefore, we introduce the following 2-D Hanning window (see Fig. 7):

$$h'(m, n) = \begin{cases} \cos^2(\pi m / N_x) \cos^2(\pi n / N_y), & |m| \leq N_x / 2, |n| \leq N_y / 2, \\ 0, & \text{elsewhere} \end{cases} \quad (14)$$

where $N_x = N/K_1$ and $N_y = N/K_2$. This window is an easily evaluated and spatially finite function since N_x and N_y are much less than N when high fringe density patterns are handled, while diminished N_x and N_y are also obtained with low fringe density patterns since a resolution compression technique is applied, as shown in the following paragraph.

In the software, four Hanning windows, oriented along the x - and y -axis and at $+45^\circ$ and -45° with respect to the x -axis, are prepared. The orientation of each selected window depends on the initial estimated fringe orientation. For each fringe pattern, one of the four windows lying in the same orientation as the earlier estimated k_{max} is selected. In dealing with different fringe densities, rather than changing the size of the Hanning windows, a resolution compression technique is used before the enhancement. For 3 to 7 fringes the image is compressed to a resolution of 64×64 ; for 8 to 18 fringes, it is compressed to 128×128 ; and for more than 19 fringes, it remains the original resolution (256×256). It is seen that fringe patterns containing a density of up to 50 fringes can be processed successfully because a higher resolution is selected for higher fringe densities. Furthermore, a high processing speed is achieved since reduced resolutions are selected for lower density fringe patterns, which appear in most cases.

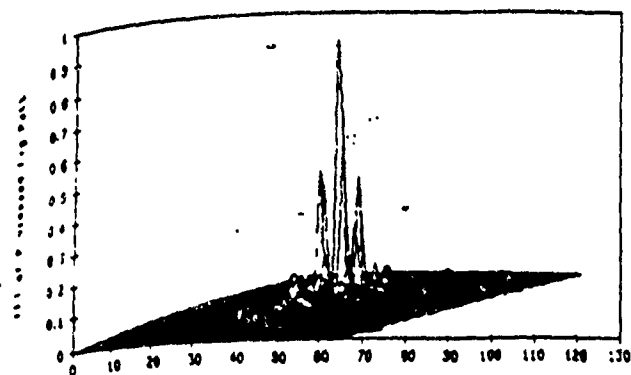


Fig. 8. Spectrum of enhanced fringe.

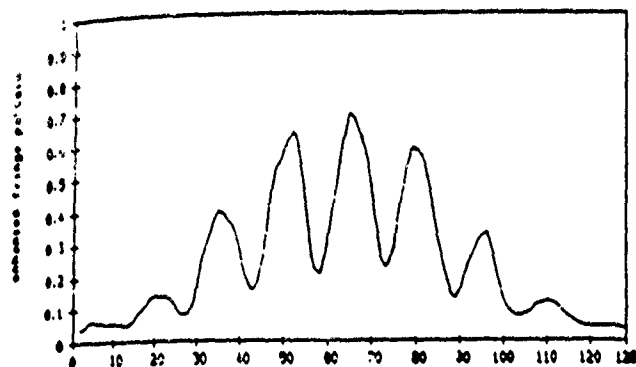


Fig. 9. Sectional view of enhanced fringes.

In practice, since an addition operation is about eight times as fast as a multiplication operation in a minicomputer, a further reduction in computing time is achieved by using integral-modulated Hanning windows. These integral windows make it possible for the convolutional process to be carried out by addition operations. Four matrices of 4×7 are selected as the practical Hanning windows. One of these integral windows, oriented along the x -axis, is as follows:

$$\begin{matrix} 1 & 1 & 2 & 2 & 2 & 1 & 1 \\ 1 & 2 & 3 & 3 & 3 & 2 & 1 \\ 1 & 2 & 3 & 3 & 3 & 2 & 1 \\ 1 & 1 & 2 & 2 & 2 & 1 & 1 \end{matrix} \quad (15)$$

It is seen that only N^2 multiplications and $49N^2$ additions (about 30 s for an image of size 128×128), are employed in the process, which compares favorably with the 2-D autocorrelation method (N^3 multiplications) and the 2-D FFT method ($N^2 \log_2 N$ multiplications in the estimating of $H(k_x, k_y)$ and much more in steps of the Newton-Raphson steepest ascent algorithm in evaluating $H(k_x, k_y)$). The spectrum of the processed fringe pattern is shown in Fig. 8, and a sectional view of the processed fringe pattern is shown in Fig. 9.

4.3. Elimination of the halo effect

As proposed by Jackson and Kaufmann,¹² fringe shifting resulting from varying irradiance of the diffraction halo has a significant effect on the analysis of Young's fringe patterns. For this reason, a halo subtraction technique, as used in the initial estimation in our program, has been used by many authors.^{3,7,8}

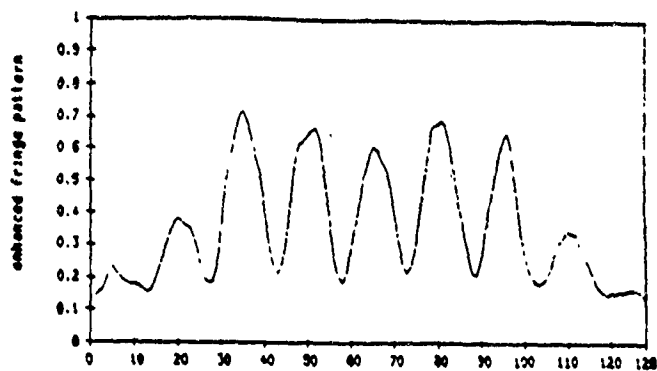


Fig. 10. Sectional view of final fringes.

In practice, however, the halo is not added but modulated with the fringe pattern. Therefore, the subtractive operation cannot alleviate the halo effect completely, although some improved results has been achieved. Others have used a minimum position detection technique that achieves perfect elimination of the halo effect when fringe visibility approaches unity.⁵ In the present program a halo division technique is applied before the enhancement process. The process is in the form

$$I'(m,n) = C \frac{I(m,n)}{I_0(m,n)} \quad (16)$$

where C is a given proper constant. After the enhancement processing on $I'(m,n)$, the final fringe pattern is, theoretically, in the form

$$I''(m,n) = 1 + V \cos\left(\frac{2\pi r \cdot d}{\lambda L}\right) \quad (17)$$

A sectional view of the resulting $I''(m,n)$ is shown in Fig. 10. Compared to Fig. 9, it is in an almost completely halo-free form regardless of the fringe visibility.

4.4. Determination of fringe spacing and orientation

Four scanning sections with different orientations, along the x -axis, the y -axis, and at $+45^\circ$ and -45° with respect to the x -axis, are prepared. Scanning directions are selected according to the initially estimated fringe orientation. For each image two orthogonal directions parallel to the directions of the initially estimated k_{m1} and k_{m2} are employed.

Since good quality fringe patterns have been achieved in the process, the detection of the fringe peaks becomes quite simple. Once an intensity is found to be higher than those intensities of the L_1 pixels before the present point along the scanning direction and not less than those of the L_2 pixels after the present point, a fringe peak is detected. L_1 and L_2 are given proper constants. Figure 11 is a histogram of the detected spacing along a direction of $+45^\circ$ with respect to the x -axis of the above processed fringe pattern. It is seen that the spacing values are in quite good distribution, concentrating on a narrow range. The final fringe spacing along this direction is then evaluated by the weighted average of all of the near-peak spacing values. High accuracy is obtained because rare and far-spread values are not taken into account. Another factor for yielding high accuracy spacing along each scanning direction is that a large number of discrete spacing

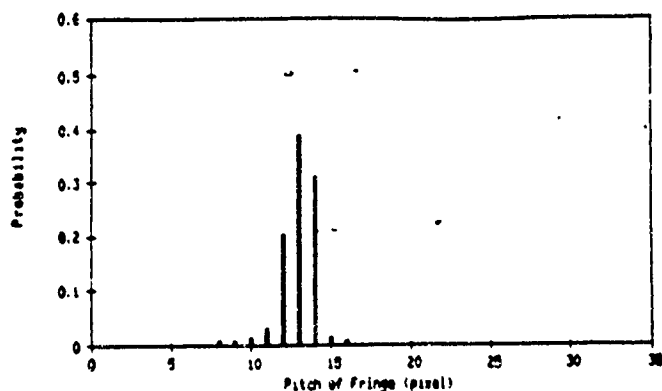


Fig. 11. Histogram of detected spacing values.

values are obtained before averaging (about 100 for 3 fringes, about 1000 for 10 fringes, and about 6500 discrete values for 40 fringes are detected). The larger the fringe number, the larger the number of discrete spacing values achieved. It is reasonable to point out that the accuracy of measured spacing in each scanning direction will increase with the increase of fringe number within an appropriate range of fringe density. It should also be noted that the accuracy of averaged spacing in both scanning directions is not restricted by the image resolution since a small increase in one of the discrete spacings will result in a small decrease in the adjacent spacing, and the average result will not be disturbed.

Fringe spacing and orientation of the fringe pattern are determined by

$$S = \frac{S_i S_j}{\sqrt{S_i^2 + S_j^2}} \quad (18)$$

$$\theta = \arctan\left(\frac{S_j}{S_i}\right) + \theta_{\alpha(i,j)} \quad (19)$$

where θ is the angle between the center fringe and the y -axis, i and j stand for two orthogonal scanning vectors, and $\theta_{\alpha(i,j)}$ is a constant depending on i and j only. The orientation of the displacement vector is perpendicular to the fringes, i.e.,

$$\alpha \approx \theta \quad (20)$$

where α is the angle between the displacement vector d and the x -axis. The amplitude of the displacement vector is determined by

$$d = \frac{\lambda L}{M'} \left(\frac{1}{S_i^2} + \frac{1}{S_j^2} \right)^{1/2} \quad (21)$$

where M' is a magnification factor that is the resultant factor of both the specklegram recording magnification and the digitization magnification.

5. EXPERIMENTAL RESULTS

The performance of the described program was tested experimentally by use of speckle negatives having a distribution of displacement. Rather than attempting to photograph a model

strained in a predetermined way, in which serious errors in the displacement values might be introduced if unpredictable rigid body displacement or rotation occurs, the experiment was carried out by means of rigid body rotation.

In a typical test, a circular plane object with a diameter of 82 mm was mounted on a rotational stage with the object surface normal to the rotation and optical axes. The object was illuminated by a laser beam with $\lambda = 0.632 \mu\text{m}$. The double-exposure speckle photograph was recorded on a 10E75 plate through an objective, 230 mm in focal length, at $f/5.6$ aperture with a magnification of about 1:1. The speckle displacement therefore varied linearly with the distance from the center of rotation. The displacement was evaluated at nine points along a helical track, with an equal separation of 4.4 mm in radius and 22.5° in orientation (Fig. 12). The given rotation of the plate was 0.100° , with displacement varying from $7 \mu\text{m}$ to $70 \mu\text{m}$. Comparison between theoretical and experimental values of the displacement magnitude d is shown in Table I, and the fringe orientation θ is shown in Table II. The maximum error of the measured displacement was $0.094 \mu\text{m}$ in magnitude and 0.192° in orientation. The maximum relative error of the displacement amplitude was 0.22%. The standard deviation of the measurement was $0.04 \mu\text{m}$ in magnitude (about 0.13% of actual displacement) and 0.12° in orientation, respectively.

6. CONCLUSIONS

A fully automatic technique for Young's fringe processing in speckle photography is developed. Good quality fringe patterns are obtained through the enhancement process. The program is written in FORTRAN. The performance of the program is tested by a rotational circular plate. Displacement components with high accuracy are achieved. The system is composed of a VAX-11/730 computer, a Hammamatsu C1000 video camera, and a 2-D motor-driven stage, equipment existing in our facility. The process consumes about 16 s in data acquisition and 25 s in data processing for each fringe pattern. It is suggested that a more compact system may be developed using a personal computer without adding any difficulties. Our further work involves achieving such a system by using an IBM PC-AT and a Spectar Javelin video camera. With the new system, higher processing speed may be achieved.

7. ACKNOWLEDGMENTS

We gratefully acknowledge the financial support provided by the Army Research Office through contract DAAZ0388K0033 (scientific program officer, G. L. Anderson) and the Office of Naval Research through contract N0001482K0566 (scientific officer, Y. Rajapakse). A special thanks is extended to H. S. Don of the Electrical Engineering Department for critical comments on the manuscript. Finally, the authors wish to thank the reviewers, who spent a great deal of time in providing constructive criticism and advice.

8. REFERENCES

1. A. E. Ennos, "Laser speckle and related phenomena," in *Topics in Applied Physics*, Vol. 9, J. C. Dainty, ed., pp. 303-353, Springer-Verlag, New York (1975).
2. F. P. Chiang and D. W. Li, "Diffraction halo functions of coherent and incoherent random speckle patterns," *Appl Opt* 24(14): 2166-2171 (1985).
3. B. Inceken, P. Eglin, and R. Dandluter, "Hybrid optical and electronic

DIGITAL PROCESSING OF YOUNG'S RINGES IN SPECKLE PHOTOGRAPHY

TABLE I. Experimental data of displacement amplitude.

Point Number	1	2	3	4	5	6	7	8	9
d(um)(applied)	7.679	15.359	23.038	30.718	38.397	46.077	53.756	61.436	69.116
d(um)(measured)	7.696	15.378	23.072	30.712	38.413	46.012	53.663	61.460	69.046
δd (um)	+0.017	+0.019	+0.034	-0.006	+0.015	-0.065	-0.094	+0.025	-0.069
$\delta d/d$ (%)	+0.22	+0.12	+0.15	-0.02	+0.04	-0.14	-0.17	+0.04	-0.10

TABLE II. Experimental data of fringe orientation.

Point Number	1	2	3	4	5	6	7	8	9
θ° (theoretical)	0.000	22.500	45.000	67.500	90.000	112.500	135.000	157.500	180.000
θ° (measured)	0.144	22.321	44.808	67.339	90.042	112.500	135.074	157.692	180.000
$\delta\theta^\circ$	+0.144	-0.179	-0.192	-0.161	+0.043	0.000	+0.074	+0.182	0.000

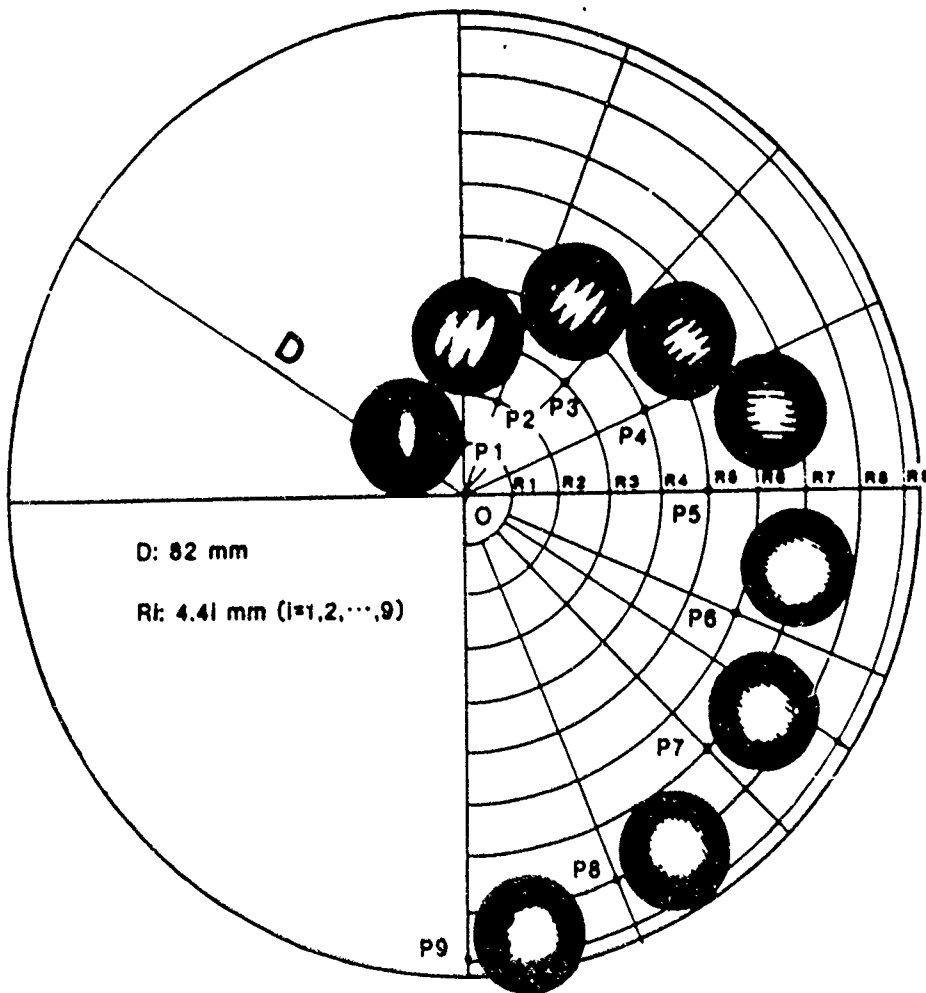


Fig. 12. Diagram of measured points on the specimen and related fringe patterns.

- image processing for strain measurements by speckle photography." *Appl. Opt.* 19(13), 2191-2195 (1980).
4. G. H. Kaufmann, A. E. Ennos, B. Gale, and D. J. Pugh, "Electro-optical readout system for analysis of speckle photographs." *Inst. of Phys.* 13, 579-584 (1980).
 5. G. E. Maddux, R. R. Corwin, and S. L. Moorman, "Improved automated data reduction device for speckle metrology," in *Proc. SESA 1981 Spring Meeting*, pp. 248-258 (1981).
 6. R. Meynart, "Instantaneous velocity field measurements in unsteady gas flow by speckle velocimetry." *Appl. Opt.* 22(4), 535-540 (1983).
 7. J. M. Huntley, "Image processing system for the analysis of speckle photographs." *J. Phys. E*, 19, 43-49 (1986).
 8. J. M. Huntley, "Speckle photography fringe analysis by the Walsh transform." *Appl. Opt.* 25(3), 382-386 (1986).
 9. D. W. Robinson, "Automatic fringe analysis with a computer image-processing system." *Appl. Opt.* 22(14), 2169-2176 (1983).
 10. A. Kirta, C. J. D. Pickering, and N. A. Halliwell, "Particle image velocimetry: a new approach to fringe analysis." *Opt. Lasers Eng.* 7, 137-147 (1986/87).
 11. F. Ansari and G. C. Gupta, "Automated fringe measurement in speckle photography." *Appl. Opt.* 26(9), 1688-1692 (1987).
 12. S. A. Isacson and G. H. Kaufmann, "Two-dimensional digital processing of speckle photography fringes. I: Diffraction halo influence for the noise free case." *Appl. Opt.* 24(2), 185-193 (1985).



D. J. Chen received his BS degree in electrical engineering in 1984 from Xi'an Jiaotong University, China. He is currently a graduate student in the Mechanical Engineering Department of the State University of New York at Stony Brook. His research interests are in the areas of automatic speckle metrology, holographic interferometry, optical information processing, computer image processing, and pattern recognition, as well as structural dynamics and fracture mechanics.



F. P. Chiang obtained his Ph D in mechanics and engineering science from the University of Florida in 1966. He joined the State University of New York at Stony Brook in 1967 and is currently the director of the Laboratory for Experimental Mechanics Research and a leading professor of mechanical engineering. In 1973-74 he was a visiting professor at the Swiss Federal Institute of Technology in Lausanne and in 1980-81 was a senior visiting fellow at the Cavendish Laboratory of the University of Cambridge. His major research interest is in the theory and application of optical methods to mechanics and metrology. He has worked in the areas of photoelasticity, moiré, holography, and laser and white light methods.

Determination of Plastic Strain and
Fatigue in Metal and Metal Composite
Using Laser Speckles*

by

F.P. Chiang
Leading professor of Mechanical Engineering
Director, Laboratory for Experimental Mechanics Research
State University of New York at Stony Brook
Stony Brook, N.Y. 11794-2300

* Invited paper at the Army Symposium in Solid Mechanics: "Mechanics of Engineering Materials and Applications", May 16-18, 1989, Newport, R.I.

Structural failure often is the result of metal fatigue. We have seen in recent years a number of aircraft disasters as a result of fatigue failure. How to detect and measure fatigue damage in metallic materials is therefore an important part of any structure integrity assurance program. It is customary in the field to do inspections visually. While this approach is definitely useful, it nevertheless is very subjective and can only be effective when the damage is substantial. In this paper we explore a non-contact and remote sensing technique using laser speckles which may someday be developed into a portable automated system for field inspection.

Speckle is the result of multiple interference of numerous waves crossing one another. When an optically rough surface is illuminated by a coherent radiation such as that from a laser beam, the reflected wavelets from each and every point mutually interfere to form an pattern of random fringes. These are the speckles: and their properties are related to the surface roughness. It is this relationship that we are exploring in our study.

It is well known that when a material experiences plastic strain, its surface roughness changes. This change can be detected by the spatial spectrum

variation of its resulting laser speckle pattern. And since plastic strain precedes fatigue crack initiation, we can use it to monitor the fatigue process as well.

It has been established that by starting with a well polished specimen, its surface roughness increases (to a certain saturated level) when it experiences increased plastic strain. And the path along which the plastic strain is attained is immaterial as long as the total effective strain is the same. A typical surface roughness and plastic strain relationship is shown in Fig. 1.

When one illuminates a metal surface with a collimated laser beam, the reflected wavelets spread out into a spectrum depicting the spatial frequency content of the surface roughness. It can be shown that the variation of the reflected light intensity may be expressed as

$$\frac{d \langle I \rangle}{\langle I \rangle} \Big|_{v=0} \approx -2g \left(\frac{d\sigma}{\sigma} \right) \quad (1)$$

along the direction of mirror reflection, and

$$\frac{d \langle I \rangle}{\langle I \rangle} \Big|_{v \neq 0} \approx 2 \left(\frac{d\sigma}{\sigma} \right) - \frac{1}{2} V^2 T^2 \left(\frac{dT}{T} \right) \quad (2)$$

along all other directions. In the above two equations $\langle I \rangle$ is the average intensity, V the direction of diffracted light, σ the standard deviation of the surface height function, T the correlation length of the surface height function, and

$$g = \left[\frac{2\pi}{\lambda} (1 + \cos \theta) \sigma \right]^2 \quad (3)$$

where λ is the wavelength of the incident light and θ the incident angle. Typical diffraction patterns from a metal surface at two levels of plastic strain are shown in Figs. 2 and 3.

It is seen from Eq. (2) both the standard deviation σ and the correlation length T of the surface height function appear in the equation. Thus it is more advantageous to use this part of the deflected light for the characterization of surface roughness. This fact is also qualitatively obvious from the two diffraction patterns shown in Figs. 2 and 3.

There are many ways that one may choose to use for extracting information from such a diffraction pattern. The one we selected is a scheme used in the discipline of pattern recognition for extracting texture of an image. We window out a portion of the higher spatial frequency domain of the diffraction spectrum,

digitize it into gray levels, and then calculate the statistical contrast value of the pattern using the following formula:

$$CON = \sum (i - j)^2 P(i, j) \quad (4)$$

where $P(i, j)$ is the probability density function of a pair of gray levels.

We use the CON value to calibrate measured strains of an aluminum specimen (1100-H14) in tension. The result is as shown in Fig. 4. We find that the CON value increases monotonically with respect to total strain until about 1.2% and plastic strain until about 0.9%. Beyond these points the CON values start to decrease with increased strain. From Fig. 1 we know that the surface roughness is still increasing at this level of deformation. Apparently the CON value is no longer sufficient to characterize it. However we shall demonstrate in the following that the CON value nevertheless is quite effective in predicting the onset of fatigue crack initiation and propagation.

We applied the CON method to monitoring plastic strains at nine points surrounding the tip region of a saw-cut notch in an aluminum specimen under cyclic

bending stress. The geometry of the specimen and configuration of the points are as shown in Figs. 5 and 6, respectively.

The specimen was mounted in a fatigue machine capable of rendering cyclic bending load. At predetermined cycles the machine was stopped and a laser beam was used to illuminate the nine selected points sequentially. The resulting diffraction spectrum was received on a ground glass and a window of its high frequency region was digitized. CON value was then calculated for each of these digitized images. And the result was plotted as shown in Fig. 7.

It is interesting to note that along points 2, 5 and 8, which was the eventual crack propagation path, the CON value increased drastically at certain cycles of loading. On the other hand, it stayed almost stationary at all other points, which were merely 0.06 inch away from the path of eventual crack propagation. This experiment was performed on an aluminum specimen with an initial surface roughness of $0.7 \mu\text{m}$.

At the conference we shall also present results from specimens made of metal matrix composites.

Acknowledgement:

Financial support of this work provided by the Army Research Office's Engineering Science Division (Dr. R.E. Singleton, Director) through contract No. DAAL0388K0033 is gratefully acknowledge.

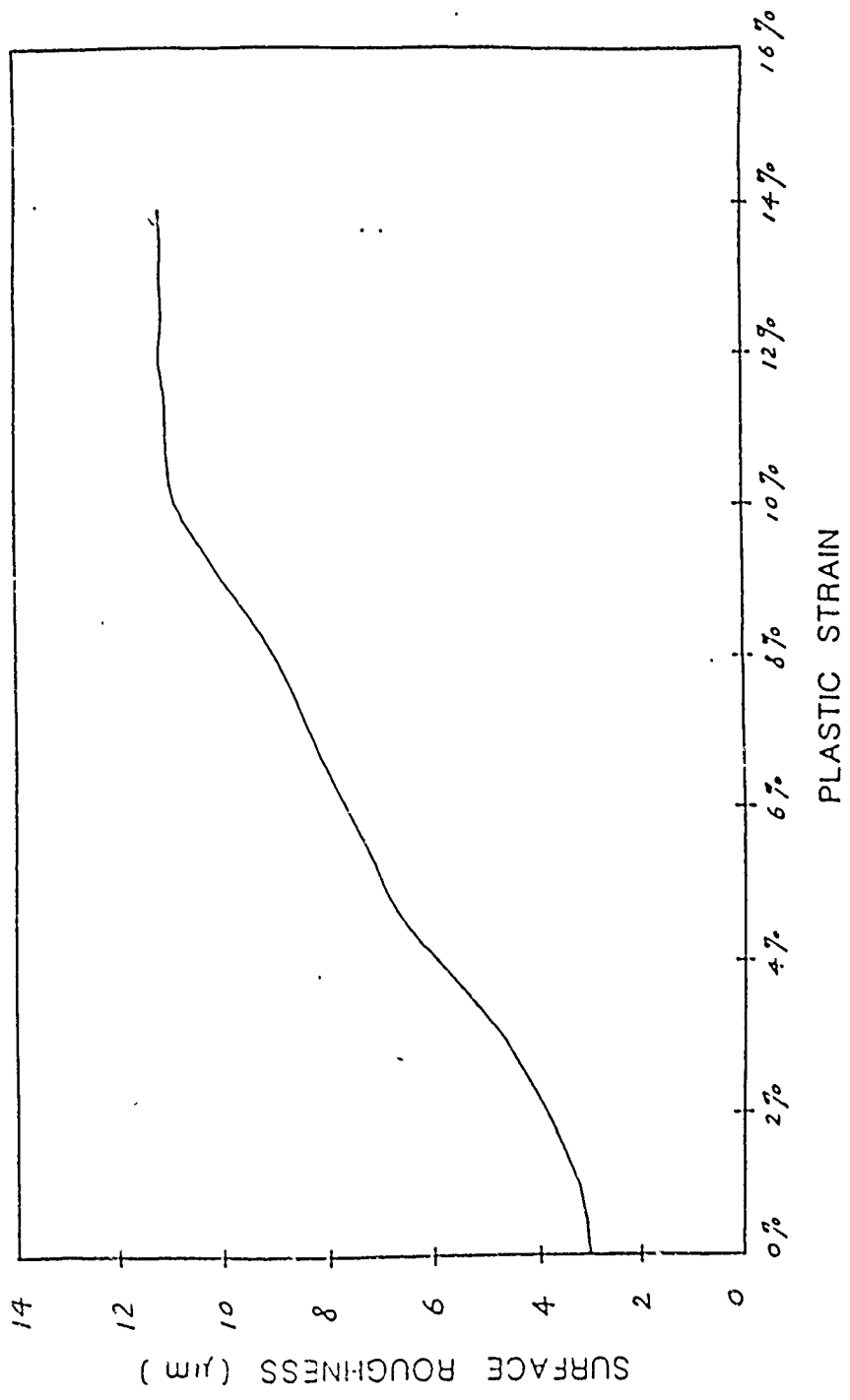
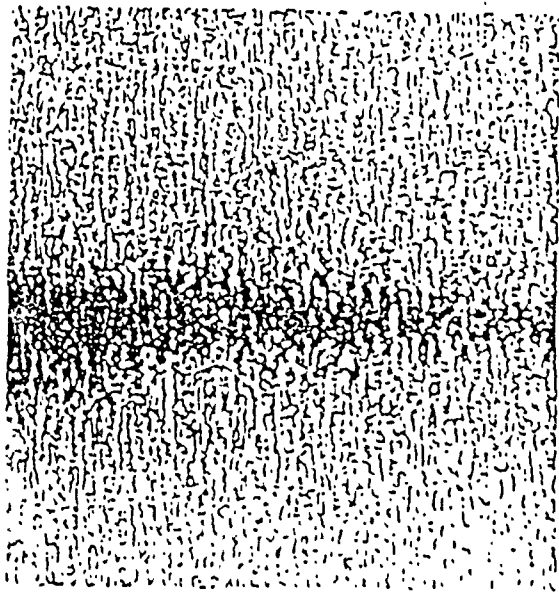
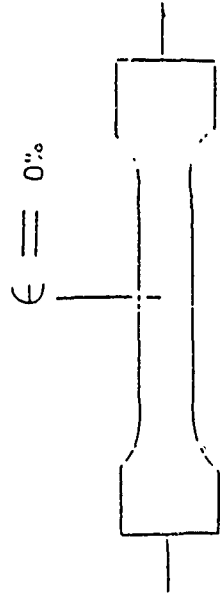


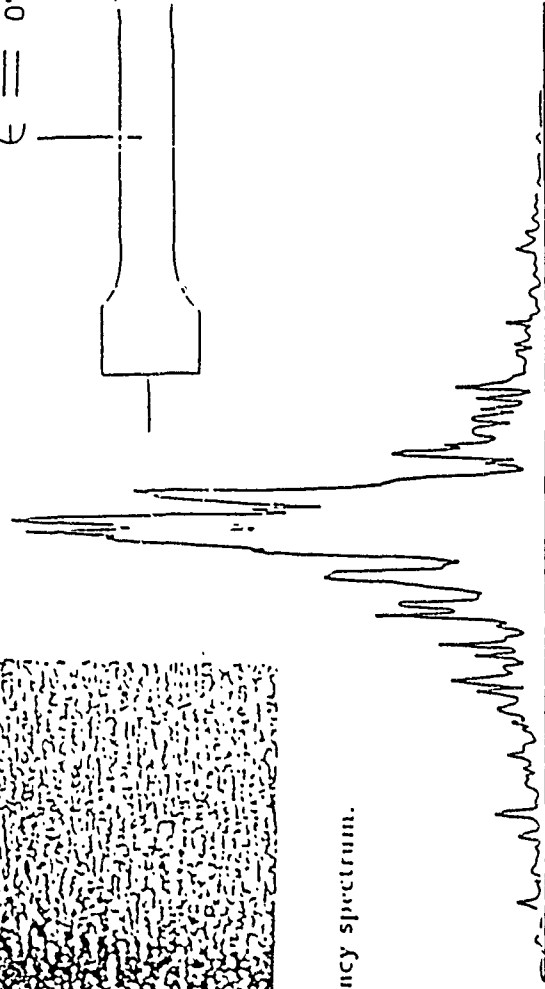
Fig. 1 Plastic strain-surface roughness curve of 1100-O AL material.



plastic strain

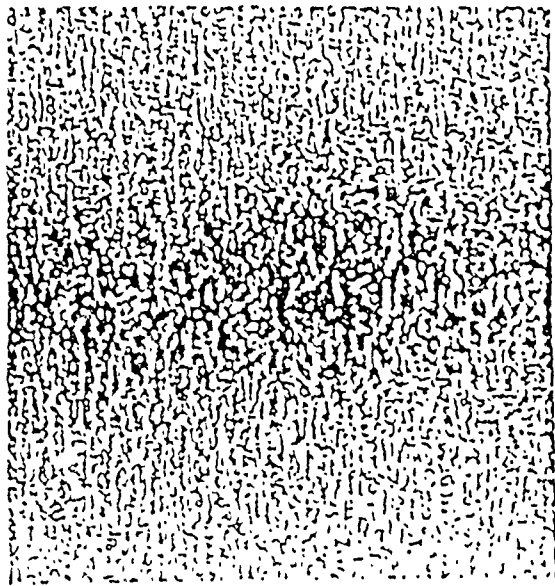


(a) Frequency spectrum.

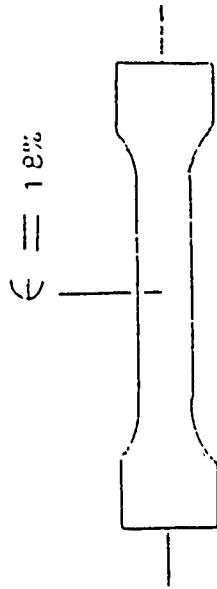


(b) Intensity profile of the frequency spectrum.

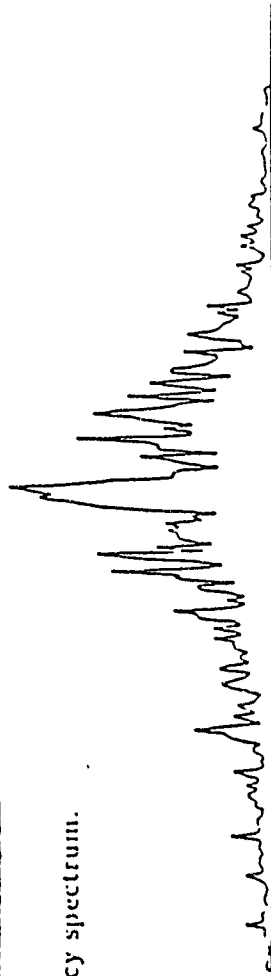
Fig. 2 Frequency spectrum from an 1100-0 AL specimen under the original (as received) state.



plastic strain



(a) Frequency spectrum.



(b) Intensity profile of the frequency spectrum.

Fig. 3 Frequency spectrum from an 1100-O

AL specimen under 1.8% plastic strain.

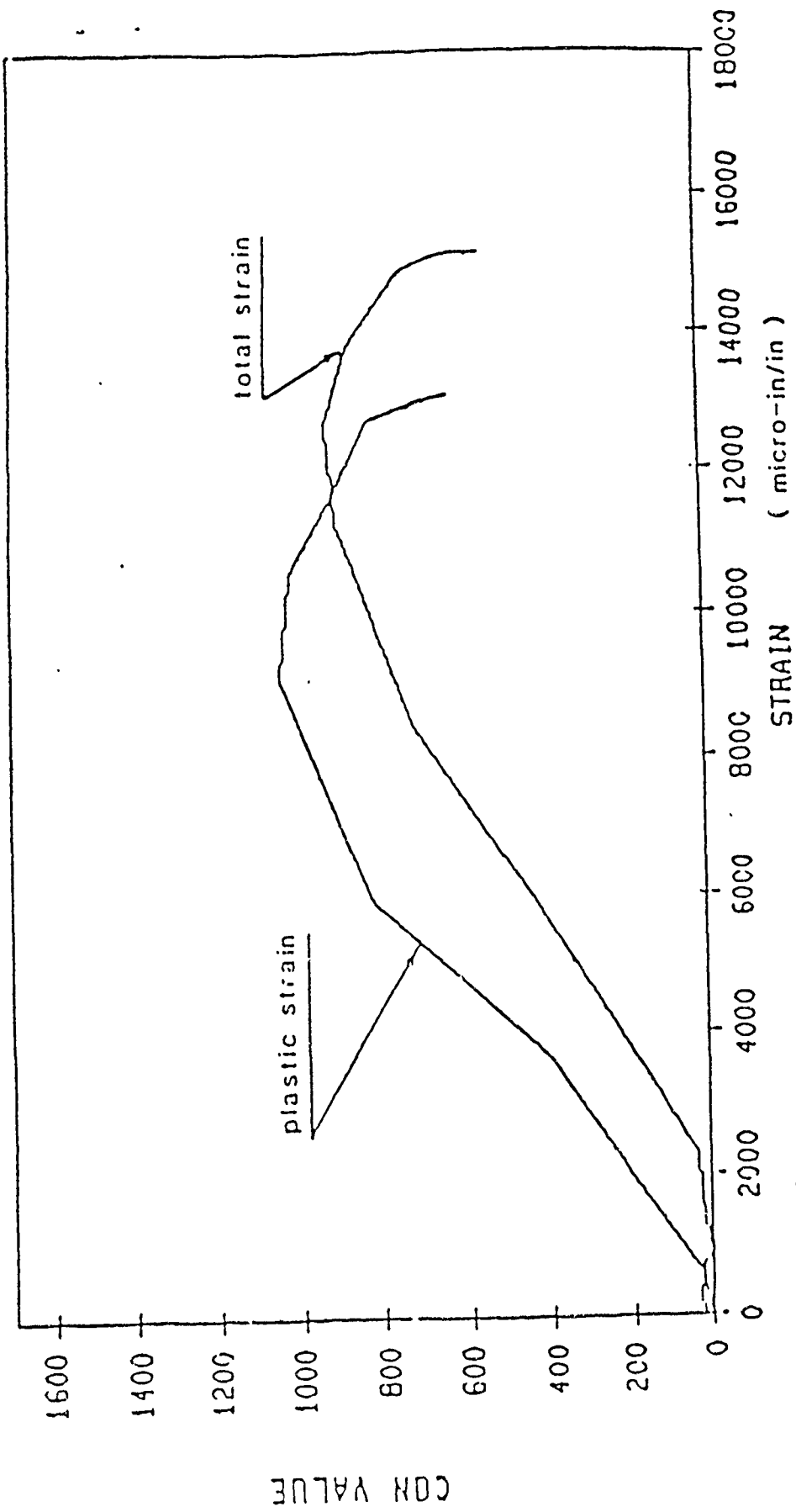


Fig. 4 Variation of strain with CON value to an 1100-H14 AL

specimen with 1 μm initial surface roughness

thickness 0.09

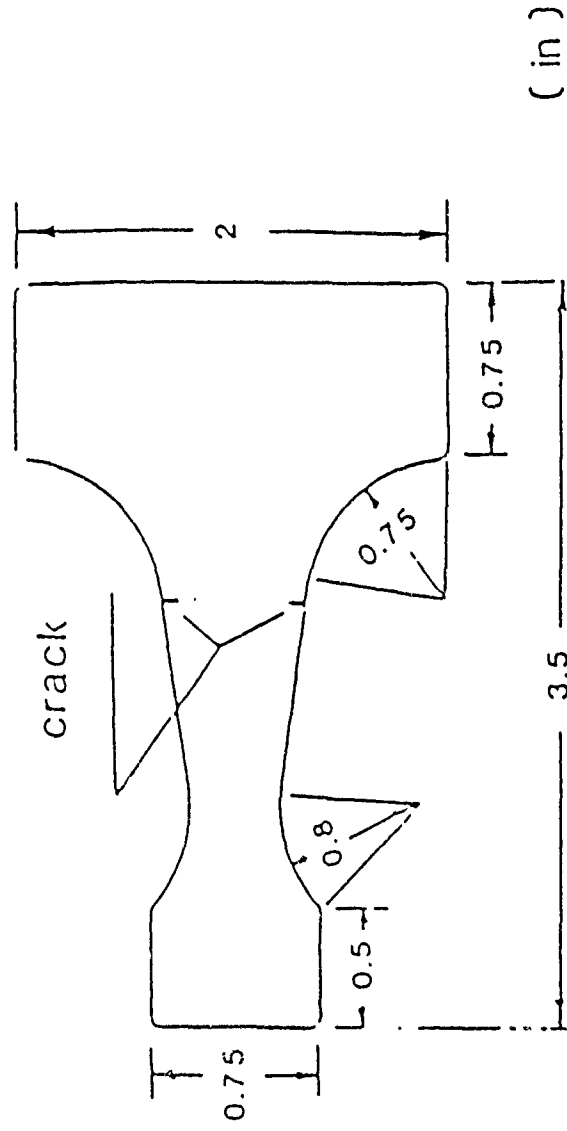


Fig. 5 Fatigue specimen made of 1100-O aluminum with cracks.

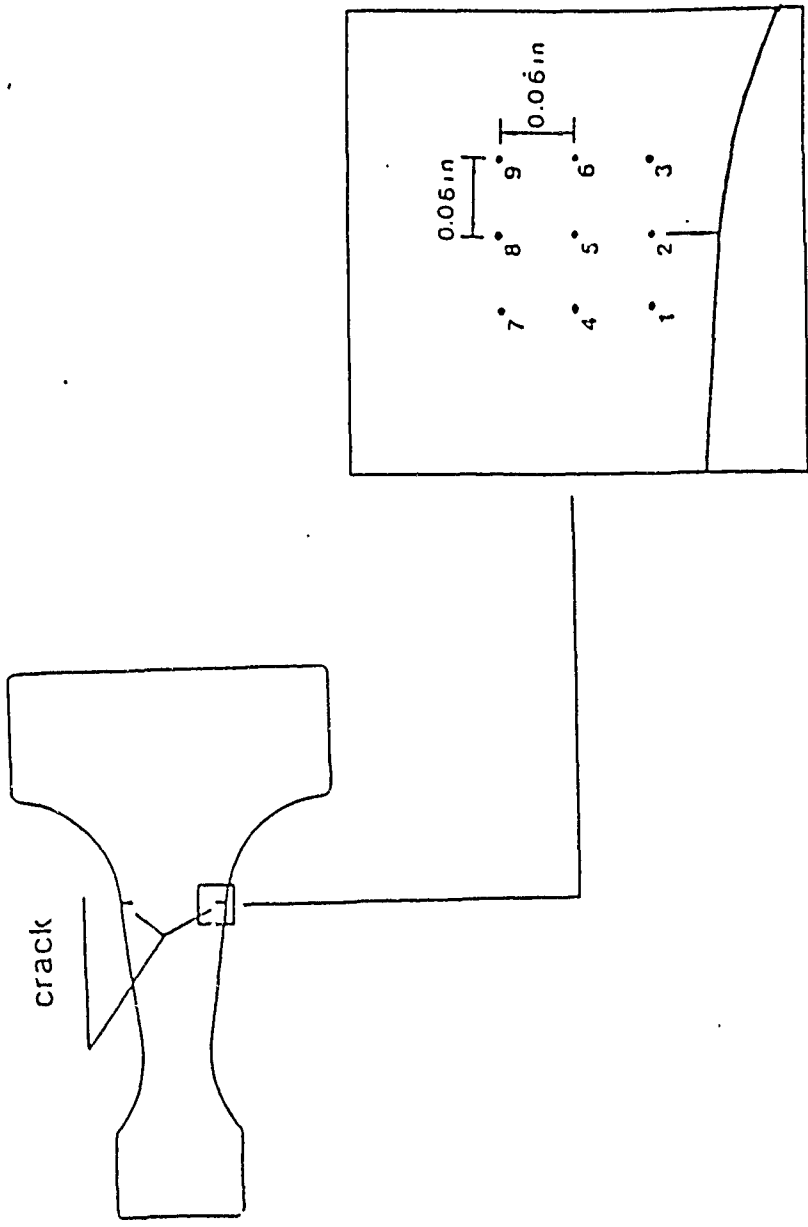


Fig. 6 Fatigue specimen with cracks and the 9 test points are arranged around the tip of a crack.

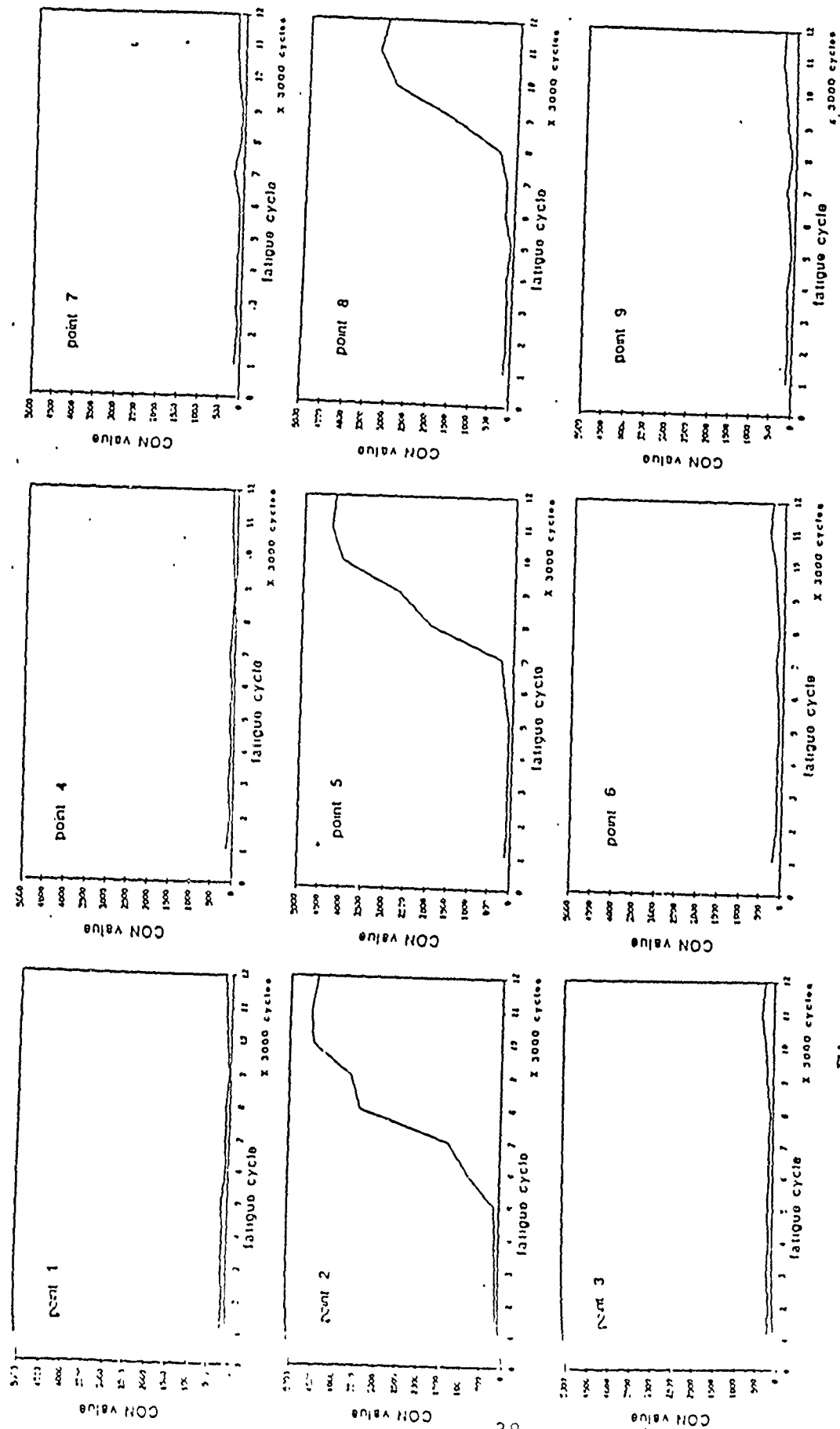


Fig. 7 Statistical contrast value (CON) of the speckle field at different points as a function of fatigue cycle.

Optimal Sampling Resolution and Range of Measurement in Digital Speckle Correlation: Part I. Laser Speckle Method

D.J.Chen and F.P.Chiang

Digital speckle correlation (or referred to as digital image correlation method) is an ideal technique in surface deformation measurement ^(1,2,3,4). In its basic form, two digital speckle patterns of the object under study are taken by a video camera before and after the deformation of the object. By means of digital correlation between subsets chosen from the two images, both displacement components and deformation components in each local region can be deduced ^[3,4].

However, unlike the normal speckle photography in which the resolution of the recording film is generally abundant, the digital resolution of a video camera is limited. Therefore, there is a trade-off between reliability and the capability of measurement. In this paper we explore the optimal sampling resolution of the laser speckle pattern upon which the speckle intensity is sufficiently registered and meanwhile the maximum measurement capability of the equipment is employed. The employed system in data acquisition and image processing is shown in Fig.1. It consists of a TV camera, an A/D converter, a camera controller and a computer. The object is illuminated by a collimated coherent light beam and observed using a video camera with flexible frame resolutions of 256×256 , 512×512 or 1024×1024 pixels. The light intensity of the speckle pattern is digitized by the analog-to-digital converter (ADC) and then transferred into the VAX-11/730 computer through an 8-bit I/O port with discrete intensity of 256 levels. The computer controls the data acquisition system, stores the digitized image data, performs the correlation calculations, and interfaces with the graphic peripherals for a display of results.

A statistical analysis of the subjective speckle pattern is performed with emphasis on spectral investigation. Intensity of one typical subjective speckle is shown in Fig.2, in which D_s , the diameter of the subjective speckle is determined by

$$D_s = 2.44\lambda q/D_L, \quad (1)$$

where λ is the wavelength of the illuminating laser light, D_L is the aperture diameter of the recording

lens and q is the imaging distance of the optical system. Results show that the spectral distribution of the speckle pattern, which determines the requirement of the sampling resolution, is also mainly determined by the lens aperture and the imaging distance of the recording optical system (Fig.3). A theoretical optimal sampling interval of

$$\begin{aligned} S &= \lambda q / (2D_L) \\ &= D_s / 4.88, \end{aligned} \quad (2)$$

is obtained using the Shannon sampling theorem [5]. Furthermore, since a bilinear interpolation is generally used in practice to achieve subpixel resolution for accurate determination of deformation components, an optimal sampling resolution for bilinear interpolation is carried out based on spectral distortion analysis. Capabilities such as accuracy and maximum measurable area are obtained. A method of increasing the upper-limit of measurement is also given by changing the recording arrangement of the optical system.

In the experimental verification of the optimal sampling resolution, a rigid body translation, which involves less decorrelation than all kinds of deformations, is tested using normalized-Product (NProd) method [6]. An area of $8.5 \times 8.5 \text{ mm}^2$ is recorded by a micro-lens of $f = 90 \text{ mm}$ at magnification $M = 1$ and $F_1 = 22$. The typical speckle size is $D_s = 67.94 \mu\text{m}$. The subjective speckle patterns are digitized by frame resolutions of 1024×1024 , 512×512 , 256×256 and 128×128 elements with real pixel distances of $S = 0.125D_s$, $S = 0.25D_s$, $S = 0.50D_s$, and $S = 1.00D_s$, respectively. The dimension of the selected subimage from the undeformed image is 15×15 elements and the dimension the selected searching region on the deformed image is 65×65 elements. Sectional views of interpolated speckle patterns from the above four different sampling resolutions are shown in Fig.4a-d. It is seen that as the sampling resolution decreases, higher signal frequency components are gradually lost and less intensity contrast is preserved. But as long as $S = 0.25D_s$, most of the speckle frequency components and most of the covariance of the speckle pattern are registered.

A mean-elimination process is first applied on both deformed and undeformed images before bilinear interpolation and correlation processes. The effect of this process on correlation evaluation is shown in Fig.5, where Fig.5a presents the correlation of two images without applying mean-elimination and Fig.5b presents the correlation of the same images when mean-elimination is applied. It is shown that in Fig.5b a zero-mean background is obtained and a much higher contrast is achieved between the matching position and the nonmatching positions.

The effect of different sampling resolution on correlation evaluation is shown in Fig.6a-c where

correlations between the deformed and the reference speckle patterns interpolated from different sampling resolutions with pixel distances of $S = 0.125D$, $S = 0.25D$, and $S = 0.50D$, are performed. It is seen that the decrease in sampling resolution results in two effects in the NProd correlation. One is the flattening of the matching hill which reduces the accuracy of measurement, another is the increase of the maximum value in the background which reduces the reliability of measurement. But as long as $S = 0.25D$, the maximum value in the background of the NProd results is less than 0.5, which is much smaller than the matching value of 0.95. Therefore, reliable determination of deformation components can still be obtained. Thus, we choose our practical optimal sampling interval as

$$\begin{aligned} S &= D_s/4 \\ &\simeq 0.6\lambda q/D_L. \end{aligned} \quad (3)$$

It is seen that the digital laser speckle correction method is capable of achieving accurate results in deformation measurement although it generally requires a fine sampling resolution. It is therefore a suitable method for measuring small deformations within a small area. Our further study shows that in the applications of large deformation and large area measurement a white light speckle method is recommended [7].

Acknowledgment:

The work is supported by Army Research Office, Engineering Science Division, Contract No. DAAL0388K0033 (Scientific Officer: Dr. G.L.Anderson) and Office of Naval Research, Solid Mechanics Division, Contract No. N0001482K0566 (Scientific Officer: Dr. Y.Rajapakse).

References:

1. W.H.Peters and W.F.Ranson, "Digital image techniques in experimental stress analysis", *Optical Engineering*, Vol. 21, No.3, pp.427-431 (1982).
2. W.H.Peters, W.F.Ranson, M.A.Sutton, T.C.Chu and J.Anderson, "Application of digital correlation methods to rigid body mechanics", *Optical Engineering*, Vol.22, No.6, pp.738-742 (1983).
3. M.A.Sutton, W.J.wolters, W.H.Peters, W.F.Ranson and S.R.McNeill, "Determination of displacements using an improved digital correlation method". *Image and Vision Computing*, Vol.1, No.3, pp.133-139 (1983).
4. T.C.Chu, W.F.Ranson, M.A.Sutton and W.H.Peters. "Application of digital-image-correlation techniques to experimental mechanics", *Experimental Mechanics*, Vol.25, No.3, pp 232-244 (1985).
5. A.Rosenfeld and A.C.Kak, Digital Picture processing, Vol 1, Academic-Press, New York (1982).
6. H.H.Bailey, F.W.Blackwell, C.L.Lowery and J.A.Ratkovic, "Image correlation: Part I. Simulation

and analysis", A Report Prepared for United States Air Force Project Rand, R-2057/1-PR, Nov (1976).

7. D.J.Chen and F.P.Chiang, "Optimal sampling resolution and range of measurement in digital speckle correlation: Part II. White light speckle method", Submitted for presentation at SEM 1981 Spring Conference on Experimental Mechanics, Cambridge, Massachusetts, USA.

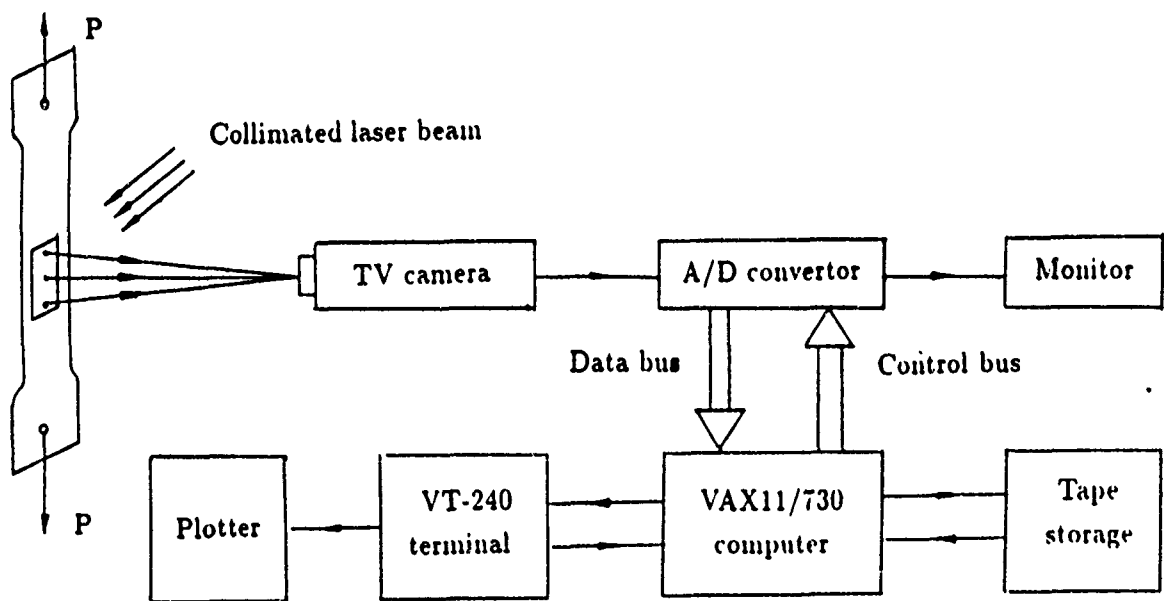


Fig.1 Schematic diagram of data acquisition and image processing system

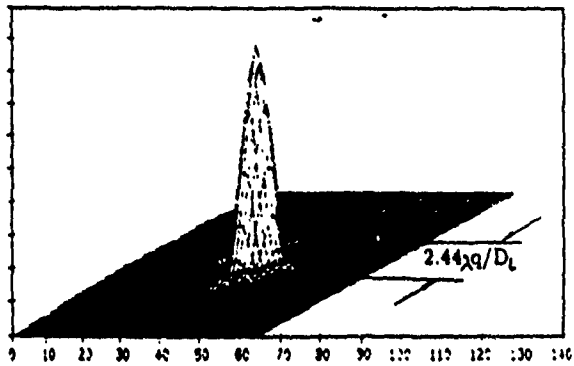


Fig.2 Intensity distribution of a typical subjective speckle

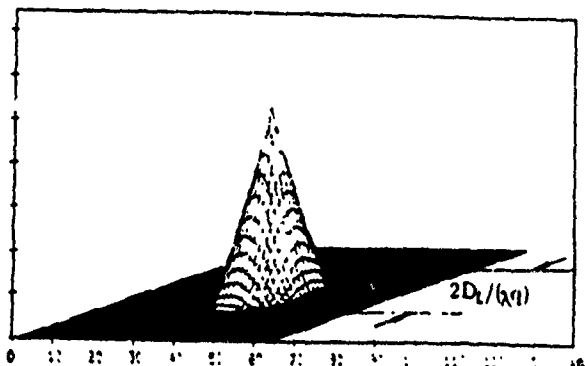


Fig.3 Spectrum of a subjective speckle pattern

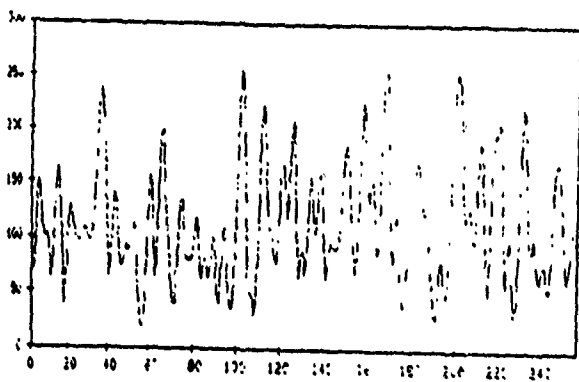


Fig.4a Frame resolution: 1024 × 1024
Sampling interval: $S=0.125D_s$

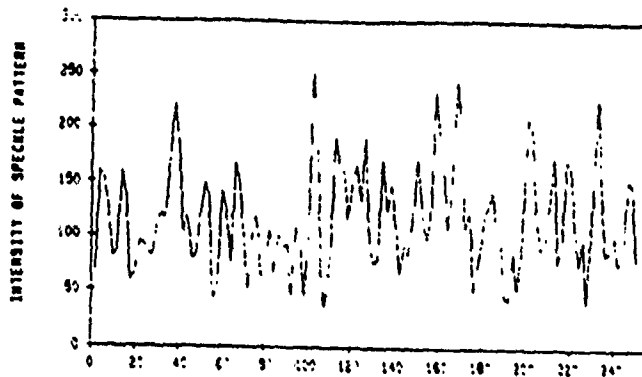


Fig 4b Frame resolution: 512 × 512
Sampling interval: $S=0.25D_s$

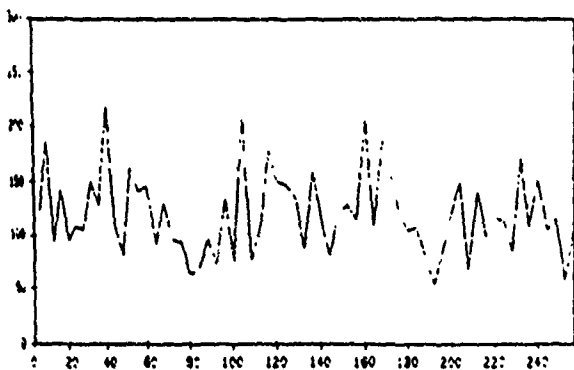


Fig.4c Frame resolution: 256 × 256
Sampling interval: $S=0.50D_s$

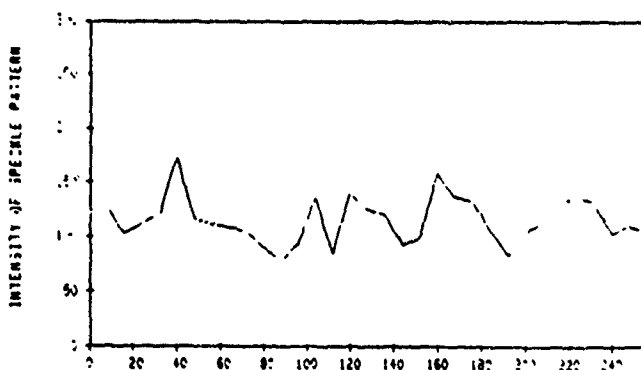


Fig 4d Frame resolution: 128 × 128
Sampling interval: $S=1.00D_s$

Fig.4 Bilinear interpolated speckle patterns from different sampling resolutions
(sectional view)

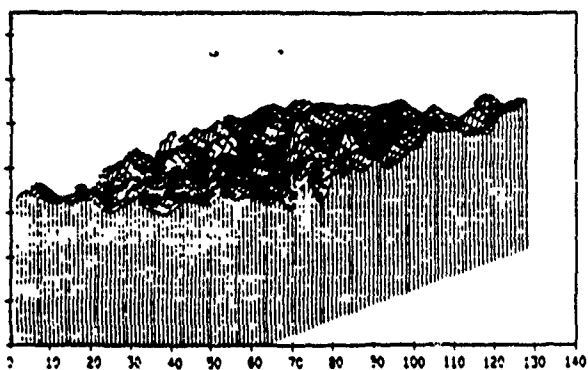


Fig.5a Correlation without applying mean-elimination

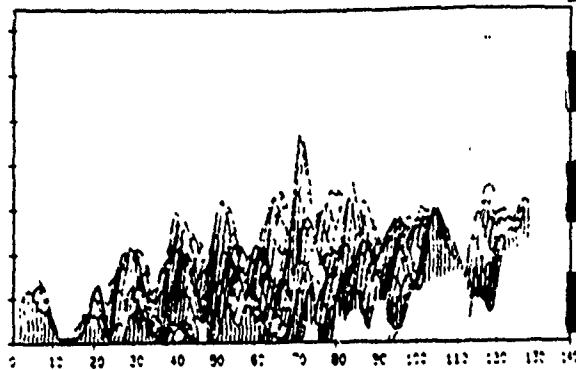


Fig.5b Correlation when mean-elimination is applied

Fig.5 Effect of mean-elimination on correlation evaluation

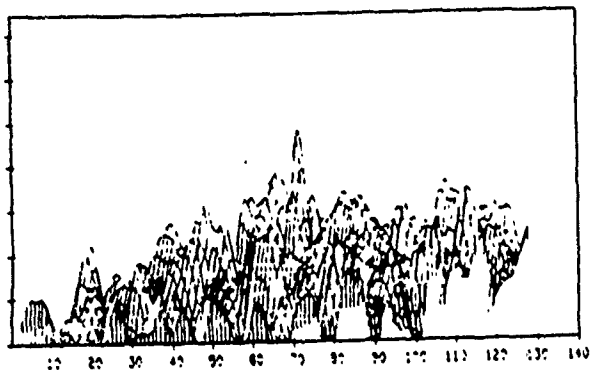


Fig.6a Frame resolution: 1024×1024
Sampling interval: $S=0.125D_s$

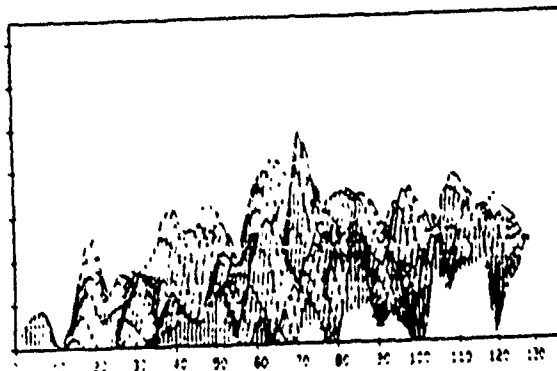


Fig 6b Frame resolution: 512×512
Sampling interval: $S=0.25D_s$

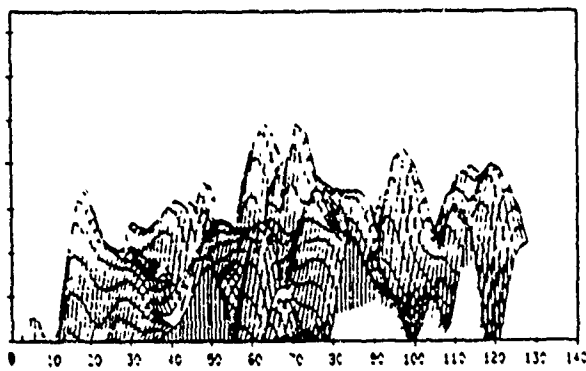


Fig 6c Frame resolution: 256×256
Sampling interval $S=0.50D_s$

Fig 6 Correlations between deformed and undeformed speckle patterns from different sampling resolutions

Optimal Sampling Resolution and Range of Measurement in Digital Speckle Correlation: Part II. White Light Speckle Method

D.J.Chen and F.P.Chiang

In the previous paper [1], we investigated the optimal sampling resolution in the digital speckle correlation technique while a laser speckle method is used. Because of the tiny speckle size, the laser speckle correlation is not capable of solving problems related to either large deformation or large area measurement. In this paper we investigate the optimal sampling resolution of the digital speckle correlation while a white light speckle method is employed and explore the resulted range of measurement.

The employed system in data acquisition and image processing is shown in Fig.1. It consists of a TV camera, an A/D converter, a camera controller and a computer. The object is illuminated by a collimated coherent light beam and observed using a video camera with flexible frame resolutions of 256×256 , 512×512 or 1024×1024 pixels. The light intensity of the speckle pattern is digitized by the analog-to-digital converter (ADC) and then transferred into the VAX-11/730 computer through an 8-bit I/O port with discrete intensity of 256 levels. The computer controls the data acquisition system, stores the digitized image data, performs the correlation calculations, and interfaces with the graphic peripherals for a display of results.

Although they are very similar to each other, the white light speckle and the laser speckle are different in the following aspects: (1) The formation of the subjective laser speckle obeys the principle of coherent image formation while that of the white light speckle obeys the principle of incoherent image formation. (2) The smallest speckle size and the range of spectral distribution of the laser speckle pattern are determined by the lens aperture and imaging distance of the recording system whereas in the white light speckle method they are determined by the smallest size of the artificial speckles, e.g. the spherical bead particles in the retroreflective paint, and the magnification factor of the optical system. (3) Since the artificial speckle size can be made as large as required, white light speckle method has its advantage over laser speckle method in large deformation and large

area measurement.

In this study the surface treatment of the object is basically the same as that used by Chiang and Asundi^[2]. The object is first coated with white paint and then with retroreflective liquid. When the painted surface is illuminated by a white light beam, the light will be reflected back by the beads in the retroreflective paint along the direction approximately same as the illuminating direction. Thus the beads behave like individual point sources resulting in a random speckle pattern. A statistical analysis of the speckle pattern is performed with emphasis on spectral investigation. Results show that the range of spectral distribution of the white light speckle, which dominates the sampling resolution requirement, is mainly determined by the smallest size of the spherical beads in the retroreflective paint and the magnification factor of the optical system. An idealized model of the light distribution of one white-light speckle is employed (Fig.2), thereby a spatial finite approximation is obtained in the spectral domain of the speckle pattern (Fig.3). A theoretical lower-limit of sampling resolution and a theoretical reconstruction function (Fig.4) are obtained using the Shannon sampling theorem^[3]. The theoretical optimal sampling interval is obtained as

$$\begin{aligned} S &= D_s/4.88 \\ &= MD_o/4.88, \end{aligned} \quad (1)$$

where M is the magnification of the optical system, D_s is the smallest speckle diameter on the image plane and D_o is the smallest size of the spherical bead particles. Since a bilinear interpolation is generally used in practice to achieved subpixel resolution for accurate determination of deformation components^[4,5,6], an optimal sampling resolution for bilinear interpolation is carried out through spectral distortion analysis (Fig.5). Upper-limits of accuracy and maximum measurable area are obtained. Methods of increasing the upper-limit are also given by changing the typical size of the spherical bead particles in the retroreflective paint and the magnification factor of the optical system. Experimental data from practical speckle patterns using normalized-Product(NProd)^[7] method is given to verify the results. A practical optimal sampling interval of white-light speckle pattern is obtained as

$$\begin{aligned} S &= D_s/4 \\ &= MD_o/4. \end{aligned} \quad (2)$$

It is seen that a much sparser sampling resolution is generally required for a white light speckle pattern, and the correlation is therefore capable for measurement of larger deformation and larger area compared with the laser speckle method^[1]. The results are consistent with the experience of Chu et al^[6], whereby more reliable results are achieved by the white light speckle method than the laser speckle method.

Acknowledgment:

The work is supported by Army Research Office, Engineering Science Division, Contract No. DAAL0388K0033 (Scientific Officer: Dr. G.L.Anderson) and Office of Naval Research, Solid Mechanics Division, Contract No. N0001482K0566 (Scientific Office: Dr. Rajapakse).

1. D.J.Chen and F.P.Chiang, "Optimal sampling resolution and range of measurement in digital speckle correlation: Part I. Laser speckle method", Submitted for presentation at SEM 1989 Spring Conference on Experimental Mechanics, Cambridge, Massachusetts, USA.
2. F.P.Chiang and A.Asundi, "White light speckle method of experimental strain analysis", Applied Optics, Vol.18, No.4, pp.409-411 (1979).
3. A.Rosenfeld and A.C.Kak, Digital Picture processing, Vol.1, Academic-Press, New York (1982).
4. W.H.Peters, W.F.Ranson, M.A.Sutton, T.C.Chu and J.Anderson, "Application of digital correlation methods to rigid body mechanics", Optical Engineering, Vol.22, No.6, pp.738-742 (1983).
5. M.A.Sutton, W.J.wolters, W.H.Peters, W.F.Ranson and S.R.McNeill, "Determination of displacements using an improved digital correlation method", Image and Vision Computing, Vol.1, No.3, pp.133-139 (1983).
6. T.C.Chu, W.F.Ranson, M.A.Sutton and W.H.Peters, "Application of digital-image-correlation techniques to experimental mechanics", Experimental Mechanics, Vol.25, No.3, pp.232-244 (1985).
7. H.H.Bailey, F.W.Blackwell, C.L.Lowery and J.A.Ratkovic, "Image correlation: Part I. Simulation and analysis", A Report Prepared for United States Air Force Project Rand, R-2057/1-PR, Nov (1976).

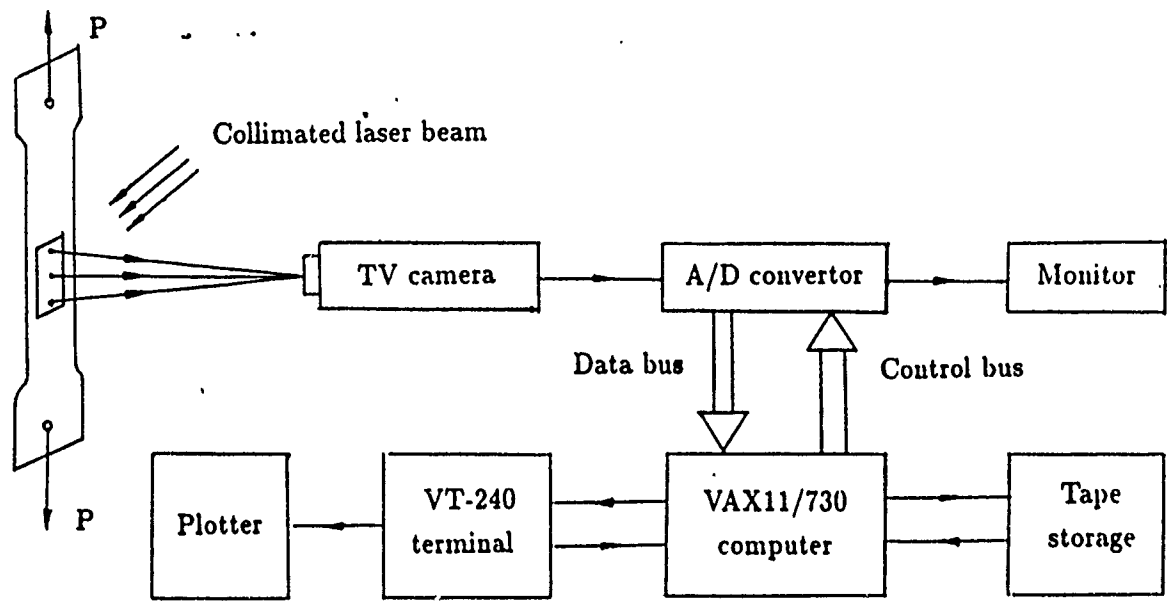


Fig.1 Schematic diagram of data acquisition and image processing system

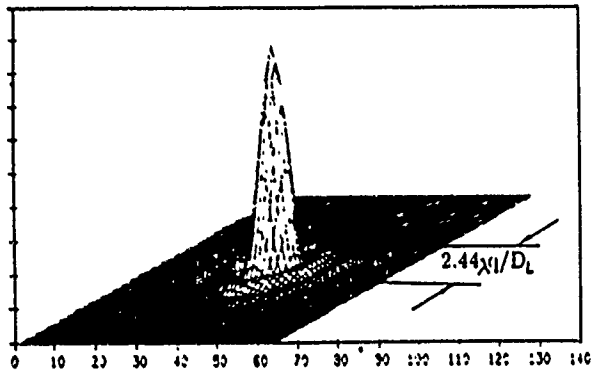


Fig.2 Intensity distribution of a typical subjective speckle

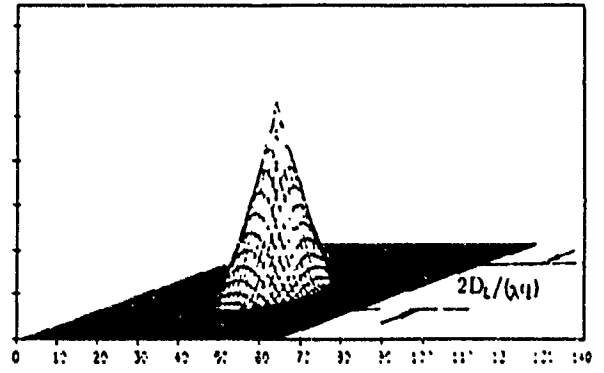


Fig.3 Spectrum of a subjective speckle pattern

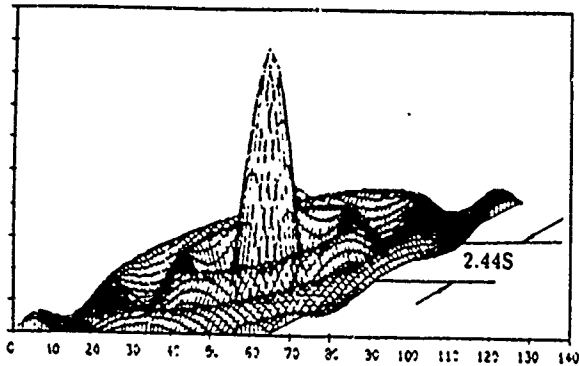


Fig.4 The theoretical reconstruction function

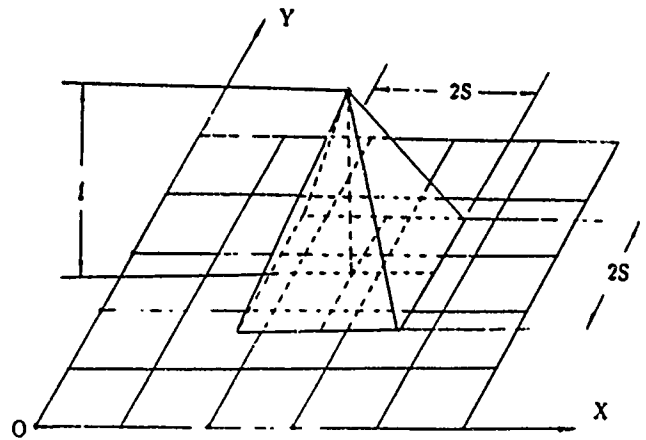


Fig.5 The convolution window of bilinear interpolation process

5

Representation and Recognition of 3-D Curves †

Chong-Huik Lo and Hon-Son Don

Department of Electrical Engineering
State University of New York
Stony Brook, New York 11794

Abstract

The space curves are highly descriptive features for 3-D objects. Invariant representations for space curves are discussed in this paper. We introduce a complex waveform representation for space curves. The waveform is parametrized by arc length. We also propose an invariant representation of space curves using the 3-D moment invariants of their breakpoints. Space curve matching using invariant global features is discussed. An algorithm for matching partially occluded 3-D curves is also presented, in which an association graph is constructed from local matchings. The maximal cliques of the graph will determine the longest portion of the model curves in the scene.

1. Introduction

The space curves (3-D curves) are useful features for object recognition, since they are rich in information and can be easily stored and manipulated. Space curves appear in various computer representation of 3-D objects. The boundary contours of objects or smooth regions in the range data are space curves. On the other hand, the spine axes of the generalized cylinders which represents elongated objects are also space curves [1,10]. The curves are distinct from other geometric entities in that they have an invariant and intrinsic parametrization by their own arc length. Volume and surface do not have such a natural parametrization. This parametrization of space curves provides a natural ordering of points on the curve, and a natural correspondence mapping between two curves. Therefore, in curve matching problem, no more computation is required to find the point correspondence than that to determine the relative offset between two curves.

In this paper, we propose an invariant waveform representation of the 3-D space curves. The waveform is

represented by a complex function. The magnitude of the waveform is the nonnegative curvature function of the space curve and the phase of the waveform depends on the torsion of the space curve. The waveform vanishes wherever the curvature vanishes. The phase of a complex function is allowed to be undefined at a point, where the function vanishes at that point. This invariant waveform is unambiguously defined for any (smoothed) curve that appears in the computer vision problems, and contains the complete 3-D information of the curve. The curves can be decomposed into smooth segments by detecting the maximum curvature points called breakpoints. The feature vector of 3-D moment invariants of the breakpoints is another invariant representation of the curve. The observed curves and the model curves in the library are represented as vectors of invariant global features. Each data curve can be identified with a model curve by using the statistical pattern classification techniques. This recognition scheme is much faster than those using the template matching techniques [5][11]. We will show that the offset parameter can be quickly determined by matching their waveforms. The offset parameter determines the point correspondence mapping between the data curve and model curve. The relative orientation and translation between these two curves can then be computed by the algorithm discussed in Section 3.

When the curves in the scene are partially occluded, their global features can not be used for pattern classification. We propose a template/structural curve matching algorithm for recognizing the partially occluded space curves. More detailed discussion of this approach is given in section 4.

2. Invariant Representation of Space Curves

A regular space curve segment is a vector-value function $x: [a, b] \rightarrow R^3$. The derivatives of x exist and are continuous up through order k ($k \geq 1$) for all t in $[a, b]$. A regular space curve segment can always be reparametrized by the arc length parameter. A curve parametrized by its arc length is called a unit speed curve. In this paper, we shall

† This work was supported by the National Science Foundation under Grant IRI-8710858 and U.S. Army Research Office under Contract DAAL 0388K0033.

assume that all space curves are unit speed curves. The curvature function $\kappa(s)$ of a unit speed curve is the magnitude of the vector field $\frac{dT}{ds}$ along the curve, and $\kappa(s) = \sqrt{(x''^2 + y''^2 + z''^2)}$. The principal normal vector field $\vec{N}(s)$ is defined by the direction of $\frac{dT}{ds}$ at each point on the curve. The binormal vector field is defined as [6]

$$\vec{B}(s) = \vec{T}(s) \times \vec{N}(s); \quad (1)$$

The torsion function is defined by

$$\tau(s) = - \langle \vec{B}'(s), \vec{N}(s) \rangle. \quad (2)$$

The triple $\{\vec{T}, \vec{N}, \vec{B}\}$ forms a local orthonormal system at each point of the space curve. The normal vector, binormal vector and torsion can be unambiguously defined only at those points where the curvatures do not vanish. A straight line in 3-D space has a vanishing curvature function, so its torsion function can not be defined. The evolution of $\{\vec{T}, \vec{N}, \vec{B}\}$ along the curve is governed by the Frenet-Serret equation.

$$\{ \vec{T}'(s), \vec{N}'(s), \vec{B}'(s) \}' = A \{ \vec{T}(s), \vec{N}(s), \vec{B}(s) \}' \quad (3)$$

where A is the following 3×3 matrix

$$A = \begin{bmatrix} 0 & \kappa(s) & 0 \\ -\kappa(s) & 0 & \tau(s) \\ 0 & -\tau(s) & 0 \end{bmatrix} \quad (4)$$

Any regular curve of C^3 with $\kappa > 0$ is completely determined, up to position and orientation, by its curvature and torsion. In computer vision problem, the data of a space curve are available in the form of a list of 3-D points which are the sampled and digitized values of a piecewisely smooth space curve equations. We smooth the curve by passing it through a Gaussian filter of suitable bandwidth, so that the smoothed version of the curve has sufficient differentiability, but still preserves the shape characteristics of the curve. In the following, we shall discuss two representations for curves using the differential geometric descriptors of their smoothed versions.

Complex Waveform Representation

We represent space curve as a complex waveform whose magnitude is the nonnegative curvature function and whose phase depends on its torsion function. The complex function can be defined at every point on the curve, even at those points whose curvatures vanish. The complex function for a (smoothed) space curve is defined as

$$f(s) = \kappa(s) \exp(i 2\pi \tau(s) s) \quad (5)$$

and $f(s) = 0$, whenever $\kappa(s) = 0$. $\lambda(s)$ in (5) is called helicity, which is defined at every point where $\kappa(s) > 0$, by

$$\lambda(s) = \frac{\tau(s)}{\kappa(s)}. \quad (6)$$

Breakpoint Representation

The breakpoints may correspond to local curvature maxima or points where the tangent vectors are discontinuous. The tangent vector discontinuities are significant, because they are the control points for spline reconstruction of the curve. The significant changes in curvature of the smoothed curves are analyzed over a range of spatial scales. Curvature maxima found at multiple scales can be located at the finest scale. The ordered list of the breakpoints is an elegant representation of the space curve.

3. Matching and Recognition of Space Curves

In this section, we shall present algorithms for space curve matching using invariant global features extracted from the complex waveforms or the lists of breakpoints. The curves can be reliably identified by the feature vectors when occlusion does not occur. The arc lengths of the curves depend to some extent on the orientations of the curves, because different amount of digitizing noises are introduced into the curves at different orientations. Therefore, the total length of the curve is a random variable and should be considered as one of the components in the global feature vector. Once a data curve has been identified with a model curve, the relative alignment and Euclidean transformation between the data curve and the corresponding model curve will be determined by a procedure described later in this section. In the following, we shall discuss the global feature extraction for closed contours, open 3-D curves, and the lists of breakpoints.

3.1. Global Feature Extraction

3.1.1. Fourier Descriptor

The boundary contours of objects or smooth regions in the range data are closed curves. The invariant waveforms of closed contours are periodic functions of the arc length parameter. They can be expanded into Fourier series. The data points on the boundary contours are usually extracted by contour following algorithms. Different starting points for contour tracing may be selected from the boundary. The magnitudes of the Fourier coefficients are not affected by the choice of the starting point. Therefore, they can be chosen as the global features. We have found in our experimental study that, using five to eight Fourier features, it is possible to reliably classify various contours in the range data.

3.1.2. Legendre Function Expansion

Open space curve segments often appear in the computer vision problems. Edges in the range data may not form closed curves. Moreover, the spine axes of the generalized cylinder representation of 3-D data are generally open curves. Global shape features of the open curves can be extracted from the orthogonal function expansion. The

invariant waveform of an open space curve can be expressed as the following series

$$I(s) = \sum_{l=0}^{\infty} k_l \sqrt{l + \frac{1}{2}} P_l\left(\frac{2s}{L} - 1\right). \quad (7)$$

where $P_l(x)$'s are the Legendre functions which are orthogonal over $[-1, 1]$. In (7), we have expressed the argument of the Legendre function in terms of the arc length, $x = 2s/L - 1$, so that these functions are orthogonal over the interval $[0, L]$. The complex coefficients k_l are defined by the integrals

$$k_l = \frac{2}{L} \sqrt{l + \frac{1}{2}} \int_0^L I(s) P_l\left(\frac{2s}{L} - 1\right) ds. \quad (8)$$

The feature vector for an open space curve can be constructed from the real and imaginary parts of these coefficients.

3.1.3. 3-D Moment Invariants of Breakpoints

The representation of space curve by an ordered list of breakpoints is not invariant under 3-D translation and rotation. However, the curves can be invariantly represented by scalar functions derived from the 3-D moments of the breakpoints. General and efficient construction rules of the 3-D moment invariants are presented in [7]. Explicit expressions of invariant functions which contain the second order and the third order moments are given in [7]. Experimental study in [7] have shown that the moment invariant features have high discriminative power for noisy data.

The data curve will be identified with a prototype, if its feature vector falls in the region of that prototype in the feature space. We found in the experiments that the minimum distance classifier can be used for space curve classification.

3.2. Coordinate Transform Estimation

In the 3-D space, the input data curve segments can be related to the corresponding model curves by an Euclidean transformation which will minimize the integral

$$\Delta(R, \bar{x}, s_0) = \int_0^L |R\bar{x}(s) + \bar{x} - \bar{y}(s + s_0)|^2 ds. \quad (9)$$

where $\bar{x}(s)$ and $\bar{y}(s + s_0)$ are the data and the corresponding model curve segments, respectively, R is the rotation matrix, \bar{x} is the translation vector, and s_0 is the offset parameter. After some algebraic manipulation, the integral can be expressed as

$$\Delta(R, \bar{x}, s_0) = \int_0^L |\bar{x}(s) - \bar{x}_0|^2 ds + \int_0^L |\bar{y}(s + s_0)|^2 ds - \frac{1}{L} \left| \int_0^L \bar{y}(s + s_0) ds \right|^2 - 2 \text{Tr}(R \Sigma_{yy}(s_0)), \quad (10)$$

where \bar{x}_0 is the average of $\bar{x}(s)$ over $[0, L]$. The cross correlation matrix is defined to be

$$\Sigma_{yy}(s_0) = \int_0^L (\bar{x}(s) - \bar{x}_0) \bar{y}^T(s + s_0) ds. \quad (11)$$

The rotation matrix that maximizes the trace term in (10), will minimize the integral Δ . The trace term will attain its maximum value $\text{Tr}(\Lambda(s_0))$, where $\Lambda(s_0)$ is the singular value matrix of $\Sigma_{yy}(s_0)$ in the following decomposition

$$\Sigma_{yy}(s_0) = U(s_0) \Lambda(s_0) V(s_0)^T, \quad (12)$$

when

$$R(s_0) = V(s_0) U(s_0)^T. \quad (13)$$

Once $R(s_0)$ is determined, the translation vector $\bar{x}(s_0)$ can be calculated as

$$\bar{x}(s_0) = R(1/L) \int_0^L \bar{x}(s) ds - (1/L) \int_0^L \bar{y}(s + s_0) ds. \quad (14)$$

The offset parameter is the value of s_0 that will minimize the following quantity

$$\Delta_{min}(s_0) = \int_0^L |\bar{x}(s) - \bar{x}_0|^2 ds + \int_0^L |\bar{y}(s + s_0)|^2 ds - \frac{1}{L} \left| \int_0^L \bar{y}(s + s_0) ds \right|^2 - 2 \text{Tr}(\Lambda(s_0)). \quad (15)$$

The offset parameter will determine the relative alignment, orientation and position of the data and model curves.

The integrals in (15) will be evaluated at every sampled value of s_0 . Each element in the cross correlation matrix can be calculated by FFT transforms. When the number of sample points, M , in the discrete approximation of (15) is 256 or 512, the total matrix operation for the whole sequence $\Delta(s_0)$, $s_0 = 0, 1, 2, \dots, M-1$ has more or less the same number of computation as that for the evaluation of the cross correlation matrix. By introducing the complex waveform matching, the above searching for s_0 can be greatly speeded up. If the model curve segment which starts at $\bar{y}(s_0)$ and ends at $\bar{y}(L + s_0)$ is the Euclidean transform image of the data curve $\bar{x}(s)$, then the invariant shape waveforms of the two curves will be exactly the same. The offset parameter s_0 which minimizes (9) is one of the zeroes of the functions $D1(s_0)$ or $D2(s_0)$ defined by

$$D1(s_0) = \int_0^L |I_1(s) - I_2(s + s_0)|^2 ds. \quad (16)$$

$$D2(s_0) = \int_0^L |I_1(L - s) - I_2(s + s_0)|^2 ds. \quad (17)$$

where $I_1(s)$ and $I_2(s)$ are invariant shape waveforms of $\bar{x}(s)$ and $\bar{y}(s)$, respectively. $\Delta(s_0)$ is evaluated at each zero of $D1(s_0)$ or $D2(s_0)$ and the offset parameter s_0 is the one that predicts the smallest value of $\Delta(s_0)$. Therefore, we search

for s_0 in the neighborhood of every pronounced local minima of $D1(s_0)$ or $D2(s_0)$. The rotation matrices calculated from the cross correlation matrix of mismatched curves are generally improper orthogonal matrices [2]. By checking the determinants of rotation matrices at various local minima, we shall be able to efficiently reduce the search space of the offset parameter.

If the data are not very noisy and the breakpoints on the curve can be reliably extracted, the rotation transformation between the data and model curves can be more efficiently determined from the relative orientation of the principal axes of their breakpoint sets [7][9].

4. Matching Partially Occluded Curves

When one object is occluded by another object, the boundary curve of the occluded object contains boundaries from both objects. The global features of the boundary curves can not be used for classification. We may have to divide the boundary curve into segments, and match each segment to the curves of the model objects. In general, the scene may contain many discontinuous pieces of various model curves. We shall determine the largest portion of each model curve existing in the scene and its position and orientation relative to its standard position. Our approach is to decompose a given data curve into a set of curve segments $\{z_i(s)\}$. Each $z_i(s)$ is used as a template and matched to various portions of each model curve $y_j(s)$ whose length is not shorter than that of $z_i(s)$. The list of breakpoints will guide the decomposition of data curve into templates. The intersection of boundaries of two objects will generally be a breakpoint. Therefore, templates can be the curve segments between breakpoints. The measure of the mismatch is the expression given in (9), or its equivalent in (15). Appropriate subscripts must be introduced to those equations, because each piece, $z_i(s)$, on the data curve will match to every model curve $y_j(s)$ in the library. Ambiguities often occur when a template can locally match to several model curves, or to various portions of a model curve. These ambiguities can be resolved when the structural information is taken into account. Moreover, the longest segment of a model curve in the scene can be determined, when the maximal set of mutually compatible matches to that model curve is found. The Euclidean transformations of these matches calculated from (13) and (14) decide their structural compatibilities. To find the maximal set of mutually compatible matches, an association graph is constructed as follows. The node of the association graph is the pair of the matched curve segments $(z_i(s), y_j(s_0, s_0+L_i))$, where $z_i(s)$ is a template and $y_j(s_0, s_0+L_i)$ is a segment of the j th model curve with end points at $y_j(s_0)$ and $y_j(s_0+L_i)$. Two nodes are connected by an edge, if the two templates in them match to the same model curve, and the relative Euclidean transformation parameters of these two matches are identical. The maximal cliques in the association graph are determined. Each clique is associated with only one model curve. The part of the model curve which

appears in the scene is the union of all the templates in its clique. Furthermore, if the observable part of one model curve is entirely contained in the observable part of another model curve, we assert that only the latter appears in the scene.

The maximal cliques can be found by recursive procedures [3,4]. They have worst case complexity of exponential time. The template/structural curve matching algorithm may require large amount of computation. The computation can be reduced if we divide each data curves into small number of long curve segments, and the attempted match is performed by matching invariant shape waveforms of the corresponding curve segments. We use long curves in waveform matching, because they can have significant shape characteristics. The waveform matching will efficiently find the corresponding curve segment on the model curve, and their relative offset parameter. The association graphs generally contain relatively small number of nodes, and the cliques can be quickly found by the recursive procedure. More details of this algorithm are given in the paper [8].

5. Experimental Results

In the following, we present the experimental study and computer simulation results of our curve matching algorithm. We first test our algorithm using curves which can be described by parametric equations. In the first experiment, the selected closed curves can be described by $x(t) = r_1 \sin(\theta(t)) \cos(\phi(t))$, $y(t) = r_1 \sin(\theta(t)) \sin(\phi(t))$ and $z(t) = r_2 \cos(\theta(t))$, where $\theta(t) = 2\pi t^2$, $\phi(t) = 2\pi t$ and $0 \leq t \leq 1$. Curves A and B are generated by the equations using parameters $r_1^A = 30$, $r_2^A = 30$, $r_1^B = 50$, and $r_2^B = 30$, $r_1^B = 40$, $r_2^B = 50$, respectively. The curves are rotated to 216 different orientations. At each orientation, they are uniformly sampled, and the Fourier features of the invariant waveforms of these curves are computed. The statistical parameters of these Fourier features are shown in table 1. These features are essentially invariant under rotation and translation. Synthetic curves are also generated by Hermit splines. The tangents of the curves are made to be discontinuous at certain points. The locations of these breakpoints can be determined by tracking the curvature maxima from the coarse to fine scale. The transformation estimation using corresponding feature point sets has been studied, both theoretically and experimentally, using noisy and noise free data. The results are given in [7].

We then test our algorithm using curves determined by the depth discontinuity in the range image. The synthetic range image in Fig.1a is generated by graphics technique. The range image in Fig.1b is obtained by laser range finder. The objects in these images are segmented from the background. Their external boundaries are extracted by contour following procedure using chain coding, and are stored in one dimensional arrays. We rotate and translate the objects such that the same views are preserved. The Fourier features are extracted from the

boundaries of the objects in each position. The means and variances of the features are estimated and shown in table 2. The Gaussian classifiers were constructed using these statistical parameters. We found that the classifiers can always unambiguously distinguish the boundary curves. The synthetic image in Fig.1a is moved to a new position as shown in Fig.2a. Because of aliasing, the boundary contours of these two images look jagged. They contain different number of pixels and their lengths are not the same. The point correspondence and motion parameters are calculated using the algorithm given in section 3. The boundary contour of the image in Fig.2a are transformed and overlapped on the boundary of the image in Fig.1a. The results are shown in Fig.2b. The image in Fig.1b is rotated to a different orientation, and 15% edge points are deleted from its boundary by occlusion, as shown in Fig.3a. This partially obscure contour is matched to the boundary of the image in Fig.1b following the procedure given in section 4. The result is shown in Fig.3b. More experimental results can be found in [8].

6. Conclusions

This paper presents invariant representations and matching algorithms for space curves. The waveform representation allows us to manipulate the segmentation and recognition of 3-D curves similarly to those of 1-D waveform. The latter are simplified and have been very well studied. An algorithm using invariant global features for matching the nonoccluded 3-D curves has been discussed. This approach is fast and robust. A template/structural algorithm is used to recognize partially occluded curves. These curve matching techniques can be used to recognize objects from their 3-D curves.

7. Acknowledgement

We thank Prof. Wei-Chung Lin in the Department of EECS, Northwestern University, for his providing the range image in Fig. 1b.

REFERENCE

- [1] G.J. Agin and T.O. Binford, " Computer description of curved objects," in *Proc. 3rd Int. Joint Conf. Artificial Intelligence*, Stanford, CA, Aug. 20-23, 1973, pp. 629-640.
- [2] K.S. Arun, T.S. Huang, and S.D. Blostein, " Least square fitting of two 3-D point sets," *IEEE Trans. on Pattern Anal. Mach. Intell.*, vol. PAMI-9, 698-700, 1987.
- [3] D.H. Ballard and C.M. Brown, *Computer Vision*. Englewood Cliffs, NJ: Prentice-Hall, 1982.
- [4] R.C. Bolles and R.A. Cain, " Recognizing and locating partially visible objects: The focus feature method," *Int. J. Robotics Research*, vol. 1, no. 3, pp. 57-81, Fall 1982.
- [5] C.M. Bastuscheck et al. " Object recognition by three dimensional curve matching," *Int. J. Intell. Systems*, vol. 1, pp. 105-132, 1986.
- [6] C.C. Hsiung, *A First Course in Differential Geometry*, Wiley-Interscience, New York, 1981.
- [7] C.H. Lo and H.S. Don " 3-D moment forms : their construction and application to object identification and positioning," to appear in *IEEE Trans. on Pattern Anal. Mach. Intell.*
- [8] C.H. L and H.S. Don, " Representation and recognition of 3-D curves," Technical report, Department of Electrical Engineering, SUNY at Stony Brook, 1989.
- [9] Z. Lin, H. Lee and T. S. Huang, " Finding 3-D point correspondences in motion estimation," *The Eighth International Conference on Pattern recognition*,
- [10] R. Nevatia and T.O. Binford, " Description and recognition of complex-curved objects," *Artificial Intell.*, vol. 8, pp. 77-98, 1977.
- [11] J.T. Schwartz and M. Sharir, " Identification of partially obscured objects in two and three dimensions by matching noisy characteristics curves," *Int. J. of Robotics Research*, vol. 6, pp. 29-44, 1987.

Table 1.

c_i	mean (A)	stand. dev.(A)	mean (B)	stand. dev. (B)
c_0	0.153	0.0	0.177	0.0
c_1	0.151	2.0×10^{-4}	0.175	1.2×10^{-4}
c_2	0.137	1.5×10^{-4}	0.164	8.0×10^{-4}
c_3	0.118	2.9×10^{-4}	0.150	0.1×10^{-4}
c_4	0.107	1.7×10^{-4}	0.146	8.9×10^{-4}
c_5	0.098	1.4×10^{-4}	0.134	1.5×10^{-4}
c_6	0.092	0.2×10^{-4}	0.128	0.8×10^{-4}

Table 2.

c_i	mean (a)	stand. dev (a)	mean (b)	stand. dev. (b)
c_0	0.48	0.23	2.60	0.22
c_1	0.09	0.07	0.32	0.18
c_2	0.12	0.10	0.32	0.10
c_3	0.11	0.08	0.39	0.13
c_4	0.09	0.08	0.30	0.09
c_5	0.09	0.06	0.13	0.07
c_6	0.05	0.02	0.20	0.10

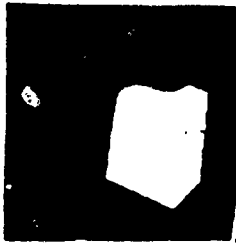


Fig. 1a



Fig. 1b

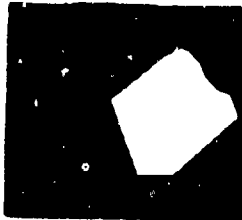


Fig. 2a

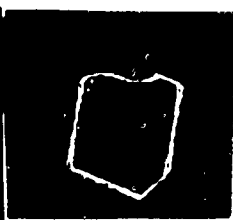


Fig. 2b

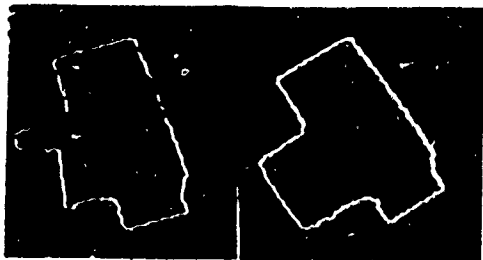


Fig. 3a

Fig. 3b

#6

Determination of Plastic Strain Using Laser Speckles

by

F.P.Chiang and H.H.Qian,

Laboratory for Experimental Mechanics Research

and H.S.Don

Dept. of Electrical Engineering

State University of New York at Stony Brook

N.Y. 11794, U.S.A

ABSTRACT

When a laser beam illuminates an optically rough surface, the reflected wavelets mutually interfere to form a pattern of random dots called speckles which carry the surface roughness information. And when a specimen is strained plastically, its surface roughness changes as a result. Thus, by illuminating a deformed specimen with a smoothly polished surface and monitoring the reflected speckle pattern via a pattern recognition technique, the state of plastic deformation can be ascertained. An example of this technique's application to detecting the initiation and propagation of a fatigue crack is presented.

KEYWORDS

Plastic deformation; laser speckle; fatigue; crack propagation.

Laser Speckle and Surface Roughness

Laser speckle [1] is a result of multiple interference of numerous coherent waves coming from different direction. This condition is created when one illuminates an optically rough surface, i.e. the surface roughness is large compared to the wavelength of light which is of the order of $0.5 \mu\text{m}$. Thus even with collimated illumination the surface will act as an array of point sources creating spherical wavelets from each and every point of the surface. The resulting interference pattern is in the form of random bright and dark dots as shown in Fig. 1. Obviously the speckle pattern carry the surface roughness information. In general, it may be stated that the smoother the surface, the larger the speckles. As the surface roughness increases, the speckle size decreases and its contrast increases until a saturation point is reached. It may be shown that when a collimated laser beam of wavelength λ illuminates an optically rough surface the reflected light intensity $\langle I \rangle$ is distributed as follows

$$\frac{d \langle I \rangle}{\langle I \rangle} \Big|_{v=0} \approx -2g \left(\frac{d\sigma}{\sigma} \right) \quad (1)$$

$$\frac{d \langle I \rangle}{\langle I \rangle} \Big|_{v=0} \approx 2 \left(\frac{d\sigma}{\sigma} \right) - 1/2 V^2 T^2 \left(\frac{dT}{T} \right) \quad (2)$$

where V is the modulus of the directional vector V of the reflected light. $V=0$ represents the direction of mirror reflection. σ and T are the standard deviation and correlation length of the surface roughness, respectively, and g is a nondimensional roughness parameter.

It is seen from the above equation that the intensity variation along the diffraction direction (i.e. $V=0$) contains the information on both the standard deviation and correlation length of the surface roughness. As shall be seen in the following that it is this part of the reflected light that a scheme is developed to extract information which is correlated with surface roughness.

Experimental Arrangement and Data Analysis

The arrangement of the optical and computational system is as shown in Fig. 2. A probing laser beam illuminates the specimen surface which is polished to almost mirror like along one direction. The reflected light distribution is in the form as shown in Fig. 3, together with a trace of the cross sectional intensity distribution. A narrow filtering mask is used to block out the central portion of the diffraction spectrum and the rest (or a portion of) is received by a digital TV camera. From which we select a two dimensional portion of the light spatial frequency signals for digitization and subsequent data extraction. The scheme adopted is a texture discrimination technique [2] used in the field of statistical pattern recognition. Namely, we calculate a statistical contrast measure CON defined as follows.

$$CON = \sum (i - j)^2 P(i, j) \quad (3)$$

where $P(i, j)$ is the probability density of a pair of gray levels occurring at two points separated by a certain distance. This information is then correlated with the surface roughness resulting from plastic deformation. Fig.4 shows the calibration curve between the CON value and the plastic strain and total strain (i.e. plastic and elastic strain) of an 1100-H14 aluminum tension specimen. The total strain was measured with a strain gage at the back of the specimen while it was under load and the plastic strain was the strain when the load was removed. It is seen that a one-to-one (and fairly linear) relationship is maintained between CON value and plastic strain up to about $9,000\mu\epsilon$.

Application to Fatigue Crack Initiation and Propagation

We applied the technique to monitoring fatigue crack initiation and propagation in a specimen with saw cut notches under cyclic bending. The geometry of the specimen is as shown in Fig. 5. The square insert is the region that was monitored by the scheme. After each pre-determined cycle of fatigue loading is finished, the load was released and a laser beam was used to probe sequentially the nine points indicated in the figure. The reflected speckle pattern was digitized and processed to yield the CON values. The results are plotted as shown in Fig 6. It is shown that along lines connecting points 1,4 & 7 and points 3,6,& 9, the CON value remains largely unchanged through

the testing. On the other hand along the crack line connecting points 2,5 & 8, the *CON* value has a drastic change after a certain fatigue cycle. At point 2 which is closest to the notch tip, the *CON* value suddenly increases after 15,000 fatigue cycles. This signals that fatigue generated plastic strain at the notch tip has roughened the polished specimen surface. At point 5 & 8 the *CON* values remain essentially unchanged. At this stage no crack could be seen with the naked eye. Indeed, crack initiation was not apparent until after 30,000 fatigue cycles. This, of course, does not mean that microcracks did not exist. At point 5 the *CON* value has a sudden increase after about 21,000 fatigue cycles and at point 8 it is after 24,000 fatigue cycles. The *CON* values continue to increase at all three points until after about 30,000 fatigue cycles. Eventually it was along this line that the crack propagated.

Acknowledgement

We would like to thank the Army research Office, Engineering Sciences Division for financial support of this project via grant No. DAAL0388K0033

Reference

1. Chiang, F.P. et al, Optical Engineering, vol. No.3, 1982.
2. Haralick, R.M. et al, IEEE Trans. Systems, Man, Cybernetics, vol SMC-3, Nov. 1973.

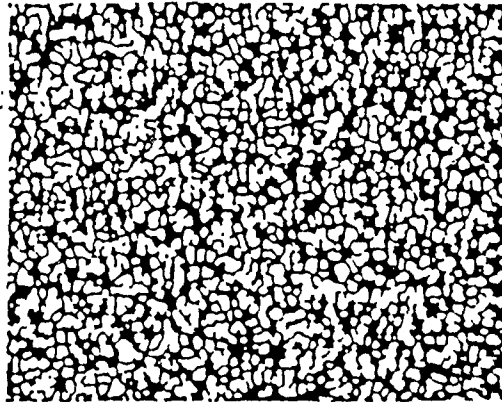


Fig.1 A typical laser speckle pattern.

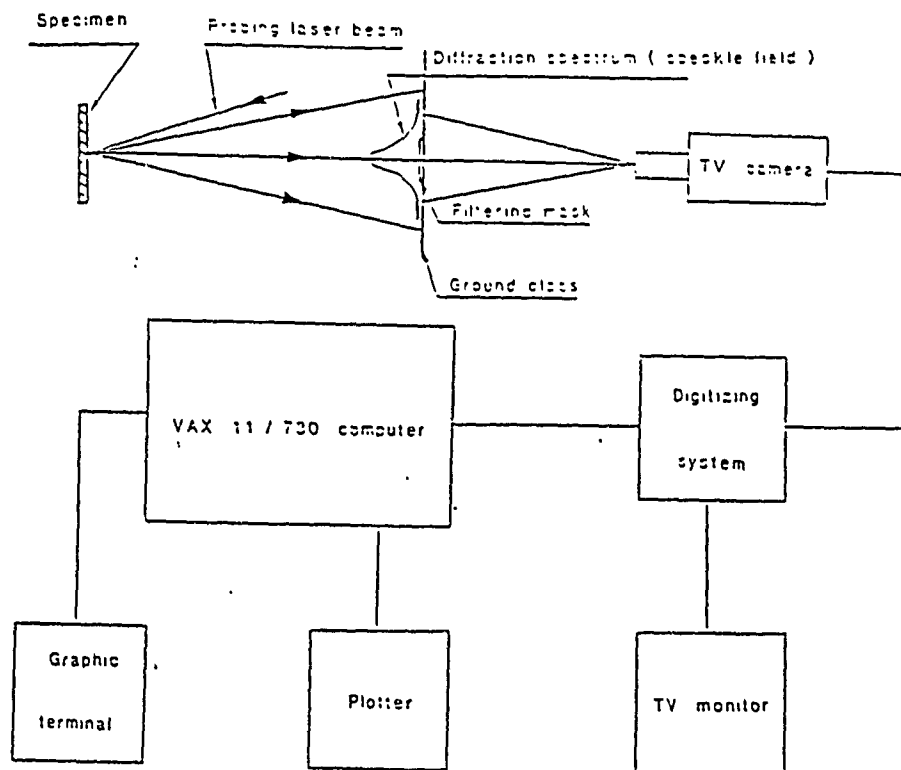


Fig.2 Experimental arrangement.

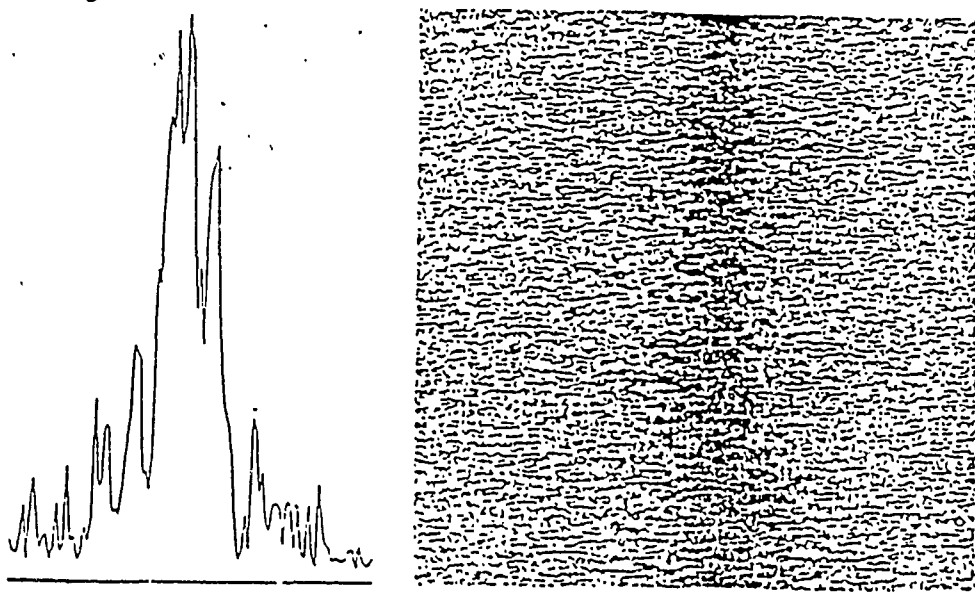


Fig.3 Laser speckle spectrum from a polished aluminum surface.

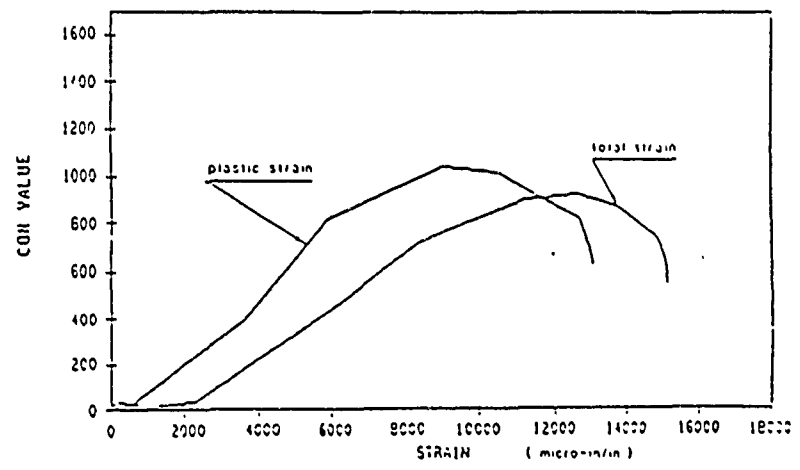


Fig.4 Correlation between CON and strain.

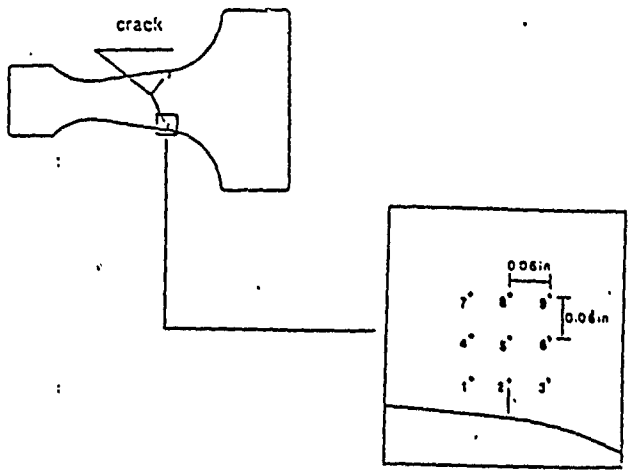


Fig.5 Fatigue specimen.

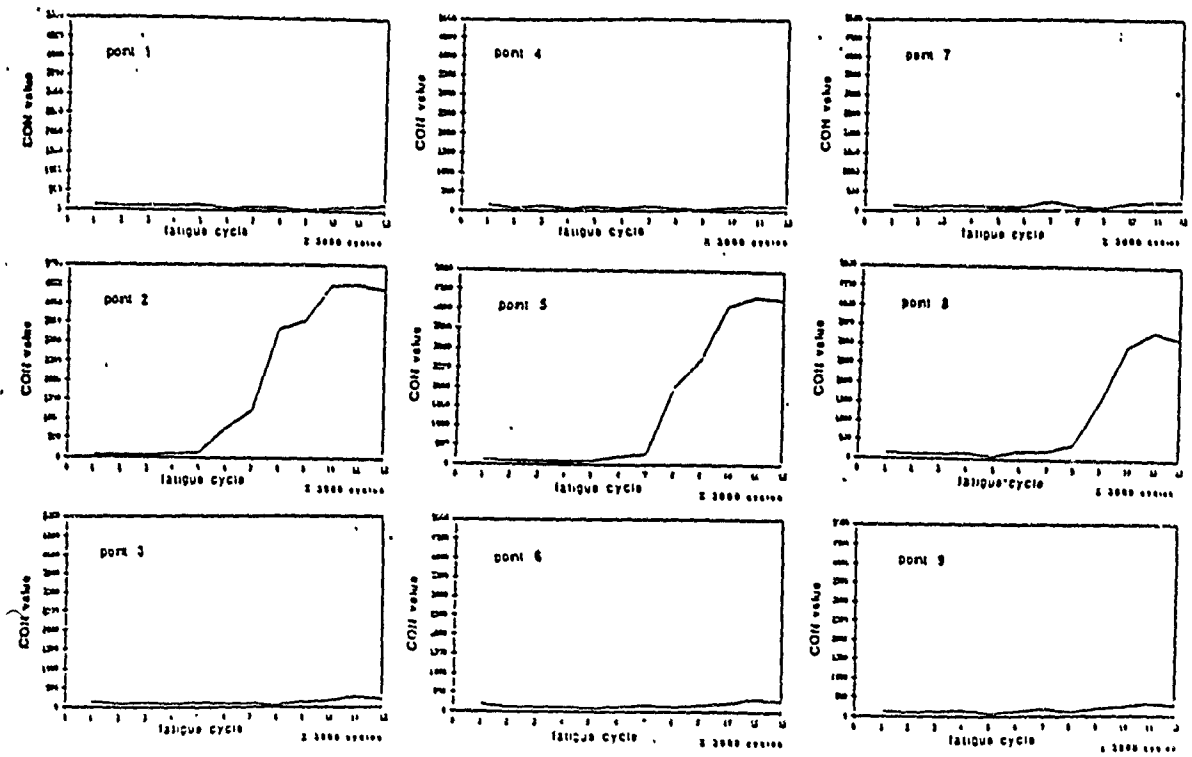


Fig.6 CON values vs fatigue cycles at different points.

LASER SPECKLE INTERFEROMETRY APPLIED TO STUDYING TRANSIENT VIBRATIONS OF A CANTILEVER BEAM

P. K. GUPTA† AND F. P. CHIANG

Department of Mechanical Engineering, State University of New York at Stony Brook, Stony Brook,
New York 11794-2300, U.S.A.

(Received 2 June 1988, and in revised form 14 February 1989)

In this paper, the applicability of laser speckle interferometry for studying the transient vibrations of mechanical structures is demonstrated. The displacement of the free end due to the transverse vibrations following a transverse impact at the tip of a cantilever beam is computed theoretically and compared with that obtained by the experimental method. The two results are found to be in close agreement.

1. INTRODUCTION

There are three essential requirements of a good measurement technique. First, it should be capable of measuring to the desired accuracy in comparison to other available techniques for the same application. Second, besides being sufficiently accurate, it should also be precise, so that when repeatedly applied to make the same measurement the results turn out close enough to be acceptable. Third, the application of the technique itself should not affect or alter the attribute or quantity to be measured. Here, we have demonstrated the capability of double exposure speckle interferometry as a *non-contact, remote sensing, accurate and precise technique* to study transient vibration of mechanical structures.

For the purpose, we selected a simple problem with a well defined solution. We used the cantilever beam as a vehicle to demonstrate the capabilities of double exposure speckle interferometry since the theory of the vibration of a cantilever beam is well established [1]. The experimental results were found to be in close agreement with those predicted by theory.

2. EXPERIMENTAL METHOD

Burch and Tokarski [2] were the first investigators to observe and formulate the basic principles of laser speckle interferometry. This technique was first extended to the measurement of displacement by Burch *et al.* in 1970 and 1972 [3, 4]. Since then a variety of methods based on the same basic principle have been developed. Chiang [5] has presented a unified approach to all of these techniques treating speckle displacement as a general phenomenon.

Khetan and Chiang [6] have derived the governing mathematical relationships for one beam laser speckle interferometry.

2.1. MEASUREMENT OF IN-PLANE DISPLACEMENT BY DOUBLE EXPOSURE SPECKLE INTERFEROMETRY

Double exposure speckle photography can be applied to the measurement of displacement in the plane normal to the line of sight by recording with a camera two superimposed

† Current address: Systron Donner Corp., Inertial Division, Systron Drive, Concord, CA 94518, U.S.A.

images of the surface, one before and one after the surface has moved. The surface is illuminated by coherent light from a laser. The speckle pattern thus generated is recorded on the film. To facilitate the formation of a good speckle field, the surface is sprayed sparingly with flat white paint, giving it a white and dark granular appearance. A slow film such as *Agfa10E75* is used because it has a small grain size, which is essential to capture the fine speckle field and also to differentiate its movement between the two exposures. Because the film grain is fine, a large exposure time or high-intensity illumination is required. For studying displacements following static loading one can conveniently apply a low-power laser and expose the film for a long time. With a 5 mW He-Ne laser a three minute exposure is not uncommon. Obviously, within this time the whole set-up should not move beyond the least detectable limit. Hence, for such experiments, vibration-isolated tables are an essential requirement. On the other hand, for capturing instants while studying dynamic displacement, a pulsed laser emitting in the optical region finds application. Usually a pulsed ruby laser is very befitting for such applications because it emits pulses which are 20 to 30 nanoseconds long and emission is in the red region of the optical spectrum, which is also the region in which *Agfa10E75* is sensitive. If the movement of the image between two exposures is larger than the diameter of the speckles and, furthermore, if the speckles remain correlated with one another, the double exposure specklegram, when illuminated by a narrow laser beam, will scatter it into a diffraction halo, the intensity of which will vary periodically across the field yielding cosine square fringes. These fringes will have an angular spacing α which is related to the wavelength λ of the readout beam, m the demagnification factor of the image and the surface displacement D by the following relation [5, 6]:

$$\sin \alpha = \lambda m / D. \quad (1)$$

The direction of the fringes will be orthogonal to the direction of the displacement.

The camera is placed as normal to the surface as possible to eliminate the effects of the relatively negligible out-of-plane displacements. The selection of the distance at which the camera should be placed from the surface whose displacement has to be determined depends upon the range of displacement to be captured which, in turn, also limits the magnitude of the smallest detectable displacement. To state it simply, higher demagnification results in a larger range of measurable displacement but poorer sensitivity. The speckle diameter on the object (Σ) is [5, 6]:

$$\Sigma = 1.2 m \lambda F. \quad (2)$$

Surface displacements of magnitude greater than this are therefore measurable. The maximum detectable displacement is limited by the ability to resolve or count the fringes in the halo. More than 20 fringes pose a problem unless the fringe quality is excellent.

2.2. EXPERIMENTAL SET-UP AND PROCEDURE

Since we were interested in determining the displacement of the tip of the cantilever beam at successive instants of time from the initial instant of impact by double exposure speckle interferometry, we used a pulsed ruby laser ($\lambda = 0.69 \text{ nm}$) to illuminate the cantilever beam tip.

The optical arrangement in Figure 1 shows the location of the camera with respect to the cantilever beam. Its focal length, f , is 55 mm and it uses a 4 in \times 5 in film. For our experiments we have used the *Agfa10E75* film. The aperture of the camera is set at $f/4.7$ (or $F = 4.7$). The laser beam from the pulsed ruby laser, expanded by using a concave lens, illuminates the area around the beam tip with intensity sufficient to expose the film within its pulse width time of 30 ns. In our case the maximum displacements were of the order of 200 μm . In our set-up a demagnification of around 4.74 was obtained and hence

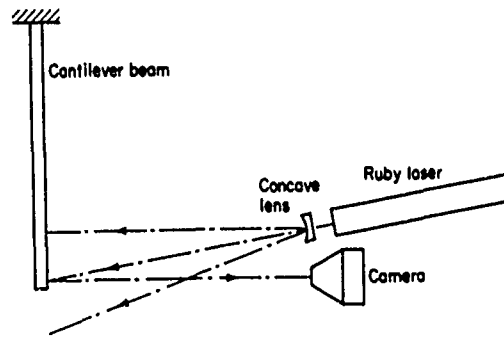


Figure 1. Schematic of the optical arrangement (plan view).

using equation (2) we obtained a sensitivity of about $18.5 \mu\text{m}$ as the smallest detectable displacement. The specklegrams were read by the point-wise technique at the tip of the cantilever beam by using a He-Ne laser. The far field diffraction pattern was observed at a distance $l = 0.79 \text{ m}$ from the specklegram. The halo within a diameter $d = 0.195 \text{ m}$ was observed and n , the number of bright fringes, was counted. The angular spacing of the fringes was estimated by using the equation

$$\sin \alpha = d/nl. \quad (3)$$

Then, by using equation (1), the object displacement per fringe was found.

Figure 2 shows the various components of the experimental set-up. The first exposure was taken before the impact and the second exposure was taken after a certain predetermined delay after the impact, which was increased from 0.1 ms to 20 ms successively in steps of 0.1 ms for each double exposure specklegram.

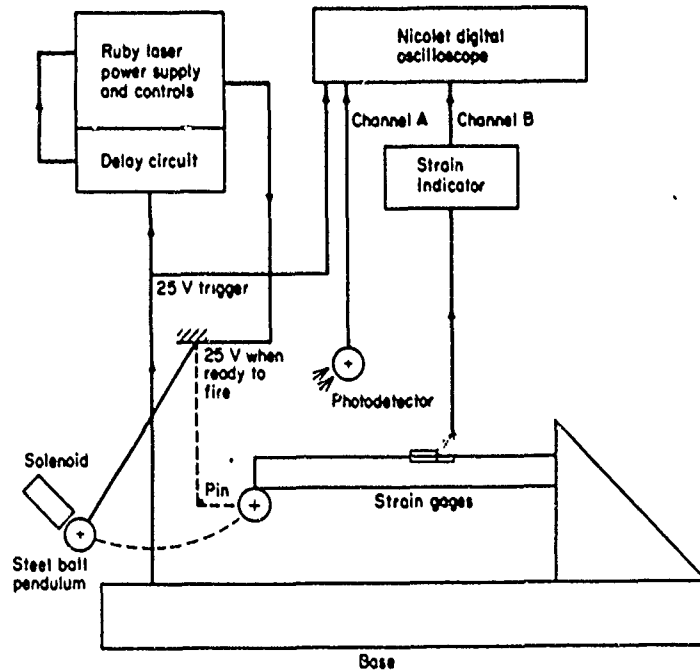


Figure 2. The experimental set-up.

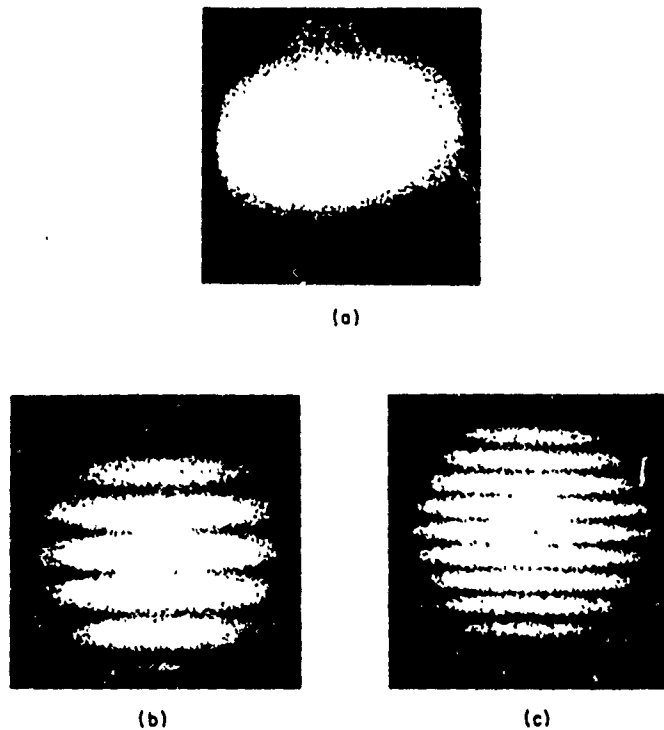


Figure 3. Young's fringes at the tip of cantilever beam at different instants after impact (pointwise filtering). (a) At 10.76 ms; (b) at 4.99 ms, (c) at 3.95 ms.

The signal from the strain gages following the impact was found to be identical for all the 200 repetitions of the process. This confirms the repeatability of the experiment, which is an essential requirement and assumption in applying this technique to construct the displacement *vs.* time curve, point-by-point. Typical fringes obtained after point-wise-readout at the tip of the cantilever beam at different instants of time after impact are shown in Figure 3.

3. THEORETICAL MODEL

In our model, we shall neglect the internal damping offered by the material. We shall consider the external damping due to air. Since we are studying the free transverse vibration *following* the impact, we shall assign a simple form to the impact force.

3.1. STATEMENT OF THE PROBLEM

The cantilever beam is straight with a uniform rectangular cross-section of width b and height h and is of length l .

The y - and z -axes are assumed to be the principal axes of the cross-section. The free end of the beam is located at $x=0$ and the fixed end is welded at $x=l$ to a metal base, as depicted in Figure 4. The welded end is treated as an elastic clamping boundary condition.

External damping is assumed to be proportional to the velocity of the beam at any given section, and the constant of proportionality is c_2 .

Initially, the beam is assumed to be at rest and the ball is assumed to strike the beam normally at the tip in the positive y -direction at a velocity V_0 .

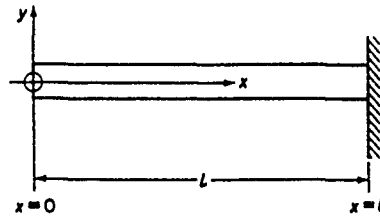


Figure 4. Location of the co-ordinate system.

The problem is to find the deflections and strains in the beam at any given location and thereby verify the experimental results obtained for the deflection at the tip and strains at $x = 24.72$ cm.

3.2. GOVERNING EQUATION FOR THE MOTION OF THE BEAM

The governing equation for the motion of the beam has been developed earlier by Clough and Penzien [1]. The cross-sectional dimensions of the beam are assumed to be smaller in comparison to its length and the effects of shearing force and rotary inertia are neglected. The differential equation, including the effect of external damping is

$$a^2 \partial^4 y / \partial x^4 + r_c \partial y / \partial t + \partial^2 y / \partial t^2 = q(x, t) / \rho, \quad (4)$$

where $a^2 = EI / \rho$ and $r_c = c_2 / \rho$.

3.2.1. Boundary and initial conditions

The boundary conditions can be written as follows:

at the free end ($x = 0$): (a) $y'' = 0$, (b) $y''' = 0$;
at the fixed end ($x = 1$): (c) $EIy'' = -K_4 y'$, (d) $y = 0$.

Here K_4 is the elastic constant of the clamping. K_4 is estimated by measuring the deflection of the elastically clamped cantilever beam tip when a static load is applied at its tip. For such a case, we observed a tip displacement of $31.9 \mu\text{m}$ when a load of 0.5 kg was applied. The displacement was determined by double exposure laser speckle interferometry. Using the theory of deflection of elastic beams developed by Crandall *et al.* [7], we found K_4 to be around 0.6×10^5 Nm.

The initial conditions are taken to be that of zero velocity and zero displacement: i.e., mathematically, $y(x, 0) = 0$, $\dot{y}(x, 0) = 0$.

3.2.2. Formulation of the loading intensity

Following Hoppman [8], the contact force $F(x, t)$ on impact can be written as

$$F(x, t) = G(x)Q(t) \quad (5)$$

with

$$G(x) = \begin{cases} 1 & \text{at } x = 0^+ \\ 0 & \text{elsewhere,} \end{cases}$$

$$Q(t) = \begin{cases} mV_0(1+e)(\pi/2T_c) \sin(\pi t/T_c) & \text{for } 0 < t \leq T_c \\ 0 & \text{for } t \geq T_c \end{cases} \quad (6)$$

Here m is the mass of the steel ball, e is the coefficient of restitution as defined by Newton, T_c is the time of contact, V_0 and V_f are the velocities of the steel ball before and after impact and $Q(t)$ is the temporal component of $F(x, t)$. Since the contact force acts as a point, ($x = 0^+$, in our case because we have a shear-free boundary condition at $x = 0$) it

is independent of x . Hoppman [8] has derived theoretical expressions for e and T_c for steel impacting steel.

The loading intensity $q(x, t)$ can then be written as

$$q(x, t) = g(x)Q(t), \quad (7)$$

where $g(x)$ is a delta function at $x=0^+$. Hence,

$$g(x) = \delta(x - 0^+). \quad (8)$$

3.3. THEORETICAL SOLUTION

We shall first enumerate the values and the units of the various constants involved in our problem (see Table 1). We shall be using the metric system of units.

TABLE 1
Constant parameter values and units

Material of the beam	mild steel	
Young's modulus of the beam, E	200×10^9	(N/m ²)
Height of the beam cross-section, h	25.37×10^{-3}	(m)
Width of the beam cross-section, b	12.68×10^{-3}	(m)
Length of the beam, l	0.3572	(m)
Mass per unit length of beam, ρ	2.53225	(Kg/m)
Impact force contact time, T_c	2.2×10^{-4}	(s)
Initial velocity of ball, V_0	0.931	(m/s)
Coefficient of restitution, e	0.22	
Radius of the ball, r	11.0871×10^{-3}	(m)
Material of the ball	steel	
Density of the ball, d	7.85×10^3	(Kg/m ³)
Coefficient of external damping, c_2	46.29	(Ns/m ²)

The coefficient of restitution has been selected as that derived theoretically by Hoppman [8] for the numerical examples presented in his work.

The coefficient of external damping is calculated from the logarithmic decrement of the tip displacement curve obtained experimentally. Its computation is discussed in section 4.

The impact force contact time is estimated from the strain gage curves recorded on the oscilloscope. It is found to be of the same order as that derived theoretically by Hoppman [8] for similar experimental conditions.

For displacement of the free end of the cantilever beam, by following Clough and Penzien [1], the analytical solution of the above governing equation for the given boundary and loading conditions is found to be as follows:

$$y(0, t) = \sum_{n=1}^{\infty} \frac{X_n^2(0) e^{-\delta_n t}}{p_n \rho \omega_n} \int_0^{T_c} e^{\delta_n \tau} Q(\tau) \sin p_n(t - \tau) d\tau, \quad (9)$$

where, for the first three modes, $k_1 l = 1.67264$, $k_2 l = 4.33152$ and $k_3 l = 7.37983$,

$$X_n(x) = (\sin k_n x + \sinh k_n x) + \alpha_n (\cos k_n x + \cosh k_n x), \quad (10)$$

$$\alpha_n = -(\sin k_n l + \sinh k_n l) / (\cos k_n l + \cosh k_n l), \quad \omega_n = \int_0^l X_n^2 dx, \quad (11, 12)$$

where $p_n^2 = a^2 k_n^4$, $2\delta_n = r_n$, and $p_n = \sqrt{p_n^2 - \delta_n^2}$.

By using these results, the tip displacement was determined as a function of time using a computer. Furthermore, the strain at any given location x on the top surface of the cantilever beam can be obtained by differentiating the result from equation (9) twice with respect to x as follows:

$$\epsilon_x(x, t) = \frac{h}{2} \sum_{n=1}^{\infty} \frac{k_n^2 [(\sin k_n x - \sinh k_n x) + \alpha_n (\cos k_n x - \cosh k_n x)]}{X_n(0)} y_n(0, t). \quad (13)$$

The strain can be numerically computed for any given location x by using this equation.

4. RESULTS

The doubly exposed specklegrams provided the number of fringes from which the displacement at each instant was calculated. For convenience of comparison, we have normalized the amplitude. The fringes obtained were horizontal implying a displacement in the y -direction. The sign of the fringes was inferred from the direction of impact. Since the impact is in the positive y -direction, the fringes in the first half-cycle are assumed to have a positive sign. The sign of the fringes then alternates between positive and negative for each consecutive half-cycle. From the plot of the experimental results (see Figure 5) for the tip displacement, the fundamental frequency of the beam has a time-period of about 7.75 ms. Figure 6 shows the variation of strain at the top surface at $x = 24.72$ cm. The experimental results for tip-displacement and strain have the same fundamental time period. This gives support to the results for displacement obtained experimentally.

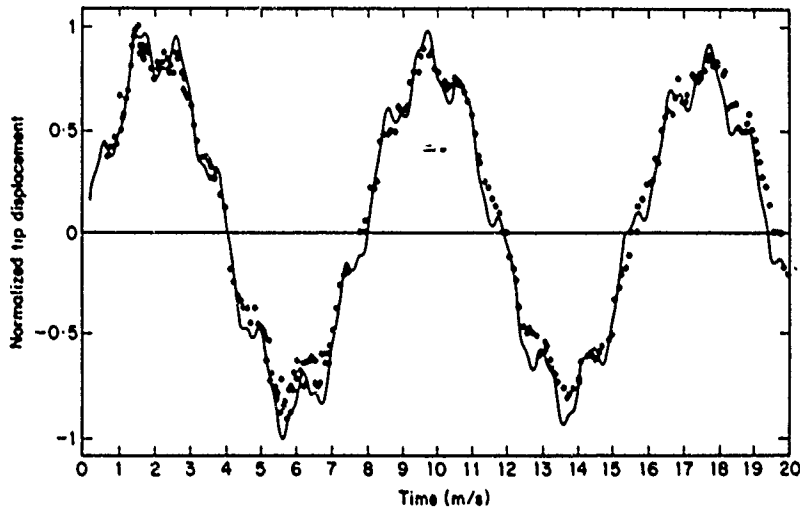


Figure 5. Comparison of theoretical and experimental results for normalized tip displacement vs. time; displacement summed over three modes. ●, Experimental values; —, Theoretical curve.

We shall now discuss the selection of the coefficient of external damping. Since we have considered external damping only, the computation of δ_n involves the damping constant c_2 . To determine c_2 , we first estimated the logarithmic decrement of amplitude from the experimental results. It was found to be 0.1431 over two cycles. Then, one has

$$\log \text{dec} = \ln \{y(0, t)/y(0, t + 2T_1)\} = 2\delta_n T_1,$$

where T_1 is the time period of the fundamental mode which was found to be ≈ 7.75 ms from the experimental results. Using $2\delta_n = r_e$ with $r_e = c_2/\rho$, one finds that $c_2 \approx 46.7$.

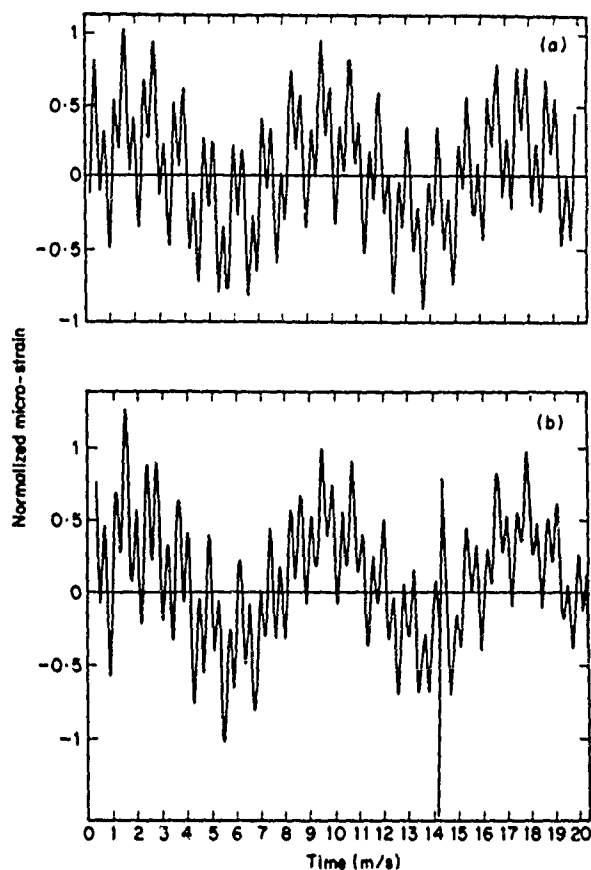


Figure 6. Comparison of theoretical and experimental results for strain. (a) Theoretical results for strain $x = 24.72$ cm, summed over three modes; (b) strain recorded from strain gages at $x = 24.72$ cm.

By using the results from equations (6), (9) and (13), and the values of the constants involved, the values of the tip displacement and strain at $x = 24.72$ cm at successive instants of time were computed. The summation was performed for $n = 1, 2$ and 3 . The contribution of terms corresponding to $n \geq 4$ was found to be negligible. The results were then plotted by using a graphics software.

Figure 5 shows the tip-displacement obtained theoretically superimposed by the experimental results for the same. As can be readily seen, the two are in close agreement. Figure 6 shows a typical experimental recording and the theoretical results for the axial strain obtained at the top surface at $x = 24.72$ cm. The spike in the strain gage output at $t = 14.2$ ms is due to the firing of the laser at this instant. The two results for strain are in close agreement as well, thereby giving validity to our theoretical model.

5. CONCLUSION

From a comparison of the experimental and theoretical results, we have found that double exposure speckle interferometry can be successfully applied to determine the displacement curves for vibrating structures. But it is also recognized that to determine the curve totally, one must ensure that the motion is repeatable. Difficulty can also arise if tilt is present.

The accuracy of measurement is limited by the range of measurement, the aperture size of the camera and the wavelength of the light used. To cover a larger range it becomes imperative to sacrifice upon the accuracy. In other words, the two are inversely related.

The results thus obtained can be used to check finite-element codes which in turn can be used for analyzing complex structures where analytical or experimental results are difficult to obtain.

ACKNOWLEDGMENT

Financial support provided by the Army Ballistic Laboratory through a contract to S & D Dynamics, Inc. is gratefully acknowledged. We would like to thank J. Pilcher of BRL and Dr M Soifer of S & D Dynamics Inc., for encouragement throughout the project. Efforts for the modification and correction of the project at its later stage were supported in part by the Army Research Office, Engineering Science Division, through contract No. DAAL0388K0033.

REFERENCES

1. R. W. CLOUGH and J. PENZIEN 1975 *Dynamics of Structures*. New York: McGraw-Hill.
2. J. M. BURCH and J. M. J. TOKARSKI 1968 *Optica Acta* 15 (2), 101-111. Production of multiple beam fringes from photographic scatterers.
3. E. ARCHBOLD, J. M. BURCH and A. E. ENNOS 1970 *Optica Acta* 17(12), 883-898. Recording of in-plane surface displacement by double exposure speckle photography.
4. E. ARCHBOLD and A. E. ENNOS 1972 *Optica Acta* 19(4), 253-271. Displacement measurement from double exposure laser photographs.
5. F. P. CHIANG 1978 *SM Archives* 3 (Issue 1), 1-32. A new family of 2D and 3D experimental stress analysis techniques using laser speckles.
6. R. P. KHETAN and F. P. CHIANG 1976 *Applied Optics* 15(9), 2205-2215. Strain analysis by one-beam laser speckle interferometry. 1: single aperture method.
7. S. H. CRANDALL, N. C. DAHL and T. J. LARDNER 1978 *An Introduction to the Mechanics of Solids* (second edition) New York: McGraw-Hill.
8. W. H. HOPPMAN 1948 *Journal of Applied Mechanics, Transactions of the American Society of Mechanical Engineers* 70(A), 125-136. Impact of a mass on a damped elastically supported beam.

3-D Moment Forms: Their Construction and Application to Object Identification and Positioning

CHONG-HUAH LO, STUDENT MEMBER, IEEE, AND HON-SON DON, MEMBER, IEEE

Abstract—The 3-D moment method has been applied to object identification and positioning. A general theory of deriving 3-D moment invariants is proposed in this paper. The notion of complex moments is introduced. Complex moments are defined as linear combinations of moments with complex coefficients. They are collected into multiplets such that each multiplet transforms irreducibly under 3-D rotations. Using the group-theoretic technique, various invariant scalars are extracted from compounds of complex moments via Clebsch-Gordon expansion. Twelve moment invariants consisting of the second and third order moments are explicitly derived in the paper. They can be used as feature vectors for automatic identification of 3-D objects and CAT images in statistical pattern recognition technique. Vectors which consists of the third order moments can be derived in a similar manner. They can be used to solve the problem of two-way ambiguity in defining the principal axes of a 3-D object, so that the rigid body rotation can be unambiguously determined from the relative orientation of principal axes in two frames. The vector moment forms can also be used in the tensor algorithm for motion estimation.

Index Terms—Complex moments, irreducible representation, motion estimation, object identification, 3-D moment invariants.

I. INTRODUCTION

OBJECT recognition and positioning are two important aspects in scene analysis. The moment method of image analysis has applications to both aspects. To achieve position invariant pattern recognition, each pattern is represented by a feature vector of moment functions, which are invariant under Euclidean or affine transformation. These moment functions are called moment invariants. Historically, the first significant paper on the use of moment invariants on the 2-D pattern recognition was published by Hu [7]. In the following years, the method was demonstrated by character recognition [7], automatic aircraft identification [3], and scene matching [21]. Hu used the theory of algebraic forms to derive invariant forms. Reddi proposed a simpler construction method using radial and angular moments [15]. Teague extended Hu's idea to the concept of orthogonal moment sets [18]. Abu-Mostafa and Psaltis introduced the notion

of complex moments and used them to derive moment invariants [1]. Each complex moment depends only on one coefficient of the circular expansion of the 2-D image function. This fact determines which features of the image contribute to the value of each complex moment, and readily answers that how much information is lost when the moment sequence is truncated. They also derived analytic formula which relates the correlation functions of the complex moments to the statistic parameters of noise in the stochastic process. In fact, their paper provides a theoretical framework to predict the discrimination power and robustness of the moment invariants in any 2-D pattern recognition problem. Recently, Teh and Chin reported a comparative study of image analysis using geometric, orthogonal and complex moments [19]. Fundamental issues, such as image representation capability and noise sensitivity were discussed in their paper. The extension of 2-D moment invariants to 3-D moment invariants was attempted by Sadjadi and Hall [16]. However, only second-order moment invariants have been explicitly derived. Moreover, their derivation of the second-order moment invariants was based on the theory of conic surfaces, so it may not be easily generalized to derive higher order moment invariants. Since 3-D information of objects can be obtained by computer tomographic reconstruction, passive 3-D sensors or active range finders, algorithms of systematic derivation of 3-D moment invariants should be developed for 3-D object identification.

Moments have also been used to find the coefficients of the coordinate transformation that relates the arbitrary and the standard positions of an object. Using the moment tensors of two image patterns which result from the orthogonal projections of a rigid planar patch at two orientations, Cyganski and Orr proposed a method to derive the affine transformation which relates these two image patterns [2], [14]. Euler angles are subsequently recovered from the parameters of the 2-D affine transformation. The four parameters of the general linear transformation matrix are the solutions of four linear algebraic equations whose coefficients are the components of the two vectors (unit rank tensors) evaluated at two different orientations. These two vectors are derived from higher order moments by contracting their tensor indices with those of antisymmetric permutation tensors. Cyganski and

Manuscript received October 12, 1987, revised November 10, 1988. Recommended for acceptance by C. Brown. This work was supported by the National Science Foundation under Grant IRI-8710856 and U S Army Research Office under Contract DAAL 0388K0033.

The authors are with the Department of Electrical Engineering, State University of New York, Stony Brook, NY 11794.

IEEE Log Number 8929338.

Orr also developed a tensor-based algorithm for object identification. They used their procedure to "normalize" the unknown pattern as well as the library pattern with respect to the same standard, so that the matching between the unknown pattern and the library pattern can be done in a view independent way. Tensor method has been extended to study 3-D medical images by Faber and Stokely [5]. They applied the method to two binary images, a geometric object and a thresholded spect (single photon emission computed tomography) study of a cardiac blood pool. Their derivation of vector forms utilized complicated tensor algebra. Moreover, the results may be sensitive to noise, because their vectors contain higher order moments which may be more vulnerable to noise contamination.

The orientation of a rigid body can be determined by its second-order 3-D central moments. The second-order moments are calculated from the 3-D data of a set of feature points on the object, if the coordinates of these feature points are known in both original and rotated frames. The eigen decomposition of the moment matrix will result in four sets of principal axes, due to the two-way ambiguity of each axis. Therefore, four possible rotation matrices can be derived from measuring the relative orientations of the principal axes in the original and rotated frames. Lin *et al.* [12] proposed that the correct rotation matrix could be selected by exact matching feature points. On the other hand, we construct a vector which consists of the third order 3-D moments of an object. The correct rotation matrix is determined by matching the values of this vector form. More detailed discussion of our method will be given in Section IV of this paper. The eigenvalues of the central moment matrix are moment invariants. In fact, they are solutions of the third-order characteristic polynomial whose coefficients are the second order moment invariants given in [16]. Using a perturbation formula, we can show that the noise contamination effect is suppressed, if the eigenvalues are very distinct from each other. Therefore, we may use the second-order moment invariants to predict the performance of the algorithm in a noisy environment. It is also possible to estimate the rotation matrix by the perturbation formula, if values of the parameters in the formula are determined by noise models or other assumptions.

In this paper, we propose a group-theoretic method to derive the 3-D Euclidean moment invariants. In a coordinate system whose origin coincides with the centroid of a 3-D pattern, a moment of order p is transformed into a linear combination of moments of order p under the rotation transformation of coordinates. The coefficients of the linear combination are the matrix elements of a reducible representation of three-dimensional rotation group. The dimension of the reducible representation is $(p+1)(p+2)/2$. Our procedure is to decompose the reducible representation into a direct sum of irreducible representations using the Clebsch-Gordon expansion [4], [8]. We also derive complex moments of order p as linear combinations of moments of the same order, such that they are transformed by the irreducible representations in the Clebsch-Gordon series. The coefficients of the linear

combinations may be complex. The irreducible representations of rotation group are indexed by a positive integer or a half integer [4], [8]. Only representations with integer indices will appear in this paper. Complex moments which are transformed by an irreducible representation D are collected in the vector $\nu_l = [\dots \nu_l^m \dots]$ and $-l \leq m \leq l$. The vector ν_l has all the transformation properties as those of irreducible tensor operators [4], [20]. In fact, we shall derive moment invariants and various complex moment forms using the calculus of irreducible tensor operator. Composite complex moment forms which are transformed according to certain irreducible representations D' can be derived as linear combinations of the product ν_l and $\nu_{l'}$ where the coefficients in the linear combinations are the vector coupling coefficients [4]. The moment invariants are those moment forms which are transformed by the identity representation corresponding to $j = 0$. Higher rank moment invariants can be obtained from the combination of three or more complex moment multiplets ν_l by using the analogous procedures and combining them two at a time. We shall demonstrate our method by deriving the second- and third-order moment invariants. Three second-order invariants have been derived and shown to be equivalent to the results of [16]. Moreover, we have obtained another nine invariants. There are two quadratic third-order moment invariants, which are bilinear forms of the third-order moment; four quartic third-order moment invariants, which are fourth-order homogeneous polynomials of the third-order moments; and three moment invariants which contain both the second order and third-order moments. We have also derived three vector forms using only the second- and third-order moments. They can be used to determine translation and orientation parameters of a given object by the tensor method. Our expression for the unit rank tensors are much simpler than those in [5]. Furthermore, our results may be less sensitive to noise because we do not use moments with order higher than three in the derivation.

The paper is organized as follows. In Section II, we define complex moments and describe our procedure for deriving various 3-D moment forms using complex moments. In Section III, we construct the second- and third-order moment invariants. In Section IV, we discuss the application of vector moment forms to various algorithms for 3-D motion parameter estimation. The computer verification of our formulae and empirical study of the digitizing errors and noise effect are given in Section V. Section VI contains our concluding remarks.

II. 3-D MOMENTS, COMPLEX MOMENTS, AND GROUP-THEORETIC METHOD

The 3-D moments of order $l + m + n$ of a 3-D density function $\rho(x_1, x_2, x_3)$ are defined by the Riemann integrals

$$M_{lmn} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x_1^l x_2^m x_3^n \rho(x_1, x_2, x_3) dx_1 dx_2 dx_3. \quad (1)$$

If the density function is piecewisely continuous and bounded in a finite region in R_3 space, then moments of all order exist. In this case, the characteristics function of

the density function can be defined

$$M(u_1, u_2, u_3) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \exp^{j(u_1x_1 + u_2x_2 + u_3x_3)} \cdot \rho(x_1, x_2, x_3) dx_1 dx_2 dx_3. \quad (2)$$

It can be expanded into a power series,

$$M(u_1, u_2, u_3) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \sum_{p=0}^{\infty} \frac{j^p}{p!} (u_1x_1 + u_2x_2 + u_3x_3)^p \cdot \rho(x_1, x_2, x_3) dx_1 dx_2 dx_3. \quad (3)$$

Interchanging the integration and summation in (3), and using the definition of moments, we can express the characteristic function as an infinite series of homogeneous polynomials of $u_1, u_2,$ and $u_3,$

$$M(u_1, u_2, u_3) = \sum_{p=0}^{\infty} \frac{j^p}{p!} H_p(u_1, u_2, u_3) \quad (4)$$

where

$$H_p(u_1, u_2, u_3) = \sum_{l,m,n=0}^p \frac{p!}{l!m!n!} M_{lmn} u_1^l u_2^m u_3^n. \quad (5)$$

For later convenience, we define a vector u whose components are the monomials $u_1^l u_2^m u_3^n$ and a vector m whose components are the coefficients of $u_1^l u_2^m u_3^n$ in $H_p,$ that is,

$$u = [u_1^0, u_2^0, \dots, u_1^l u_2^m u_3^n, \dots]^T, \\ m = \left[M_{p00}, M_{0p0}, \dots, \frac{p!}{l!m!n!} M_{lmn}, \dots \right]^T. \quad (6)$$

so that H_p can be written as a scalar product of m and $u.$ The components in u and m may be arranged in different order, so long as their scalar product remains to be $H_p.$ The centroid of the density function can be determined from the zeroth and the first-order moments by

$$\bar{x}_1 = \frac{M_{100}}{M_{000}}, \quad \bar{x}_2 = \frac{M_{010}}{M_{000}}, \quad \bar{x}_3 = \frac{M_{001}}{M_{000}}. \quad (7)$$

Moments with their origin at the centroid of the density function are called central moments. The central moments are invariant under translation. From now on, we assume that the origin of the coordinate system is always at the centroid, so all the moments are central moments, and the characteristic function is also referred to central moments.

The phase of the characteristic function is invariant under rotation, so both the spatial frequency vector $\vec{u} = (u_1, u_2, u_3)$ and the coordinate vector $\vec{x} = (x_1, x_2, x_3)$ are transformed by the same orthogonal matrix under rotation of the coordinate system, i.e.,

$$\vec{x}' = R\vec{x}, \quad \vec{u}' = R\vec{u}. \quad (8)$$

The transformation property of the moments can be determined from that of the spatial frequency \vec{u} by the following equation:

$$\sum_{l,m,n=0}^p \frac{p!}{l!m!n!} M_{lmn} u_1^l u_2^m u_3^n \\ = \sum_{l,m,n=0}^p \frac{p!}{l!m!n!} M'_{lmn} u_1'^l u_2'^m u_3'^n. \quad (9)$$

where M'_{lmn} and u_i' are the corresponding quantities after transformation. Since monomials $u_1^l u_2^m u_3^n$ are transformed among themselves under rotation, moments of the same order are also transformed among themselves. The (absolute) moment invariant is a homogeneous polynomial of moments which retain its value after rotations. Moment invariants are valuable for recognition of 3-D objects independent of their positions and orientations. In [16], it has been shown that there are at most $((p + 1)(p + 2)/2) - 3$ absolute moment invariants of order $p.$ In the following, we propose an algebraic method to construct the moment invariants. The basic principle of our method is the representation theory of rotation group. A comprehensive and elegant introduction of the group representation theory can be found in [4], [8].

The first step of our approach is to decompose the symmetric tensor space into a direct sum of invariant subspaces on which the rotation group operates irreducibly. We choose the new basis of the tensor space to be the harmonic polynomials $y_l^m (= \rho^p Y_l^m),$ where ρ is the magnitude of the vector $\vec{u},$ and Y_l^m is the spherical harmonics. The spherical harmonics are solutions of the Laplace differential equation. They form a complete orthonormal basis for the continuous real valued functions defined on a unit sphere. It is well known that a spherical harmonics Y_l^m is transformed by rotation of coordinate system into a function which can be expressed as a sum of spherical harmonics with the same l but with m running over the range $-l \leq m \leq l.$ Therefore, the harmonic polynomials y_l^m span an invariant irreducible subspace of dimension $2l + 1.$ Expressing each monomial basis $u_1^l u_2^m u_3^n$ as a linear combination of the harmonic polynomials $y_l^m,$ we achieve the decomposition of the symmetric tensor space into a direct sum of invariant irreducible subspaces. The symmetric tensor space of rank $p,$ can split into invariant subspaces of dimensions $2p + 1, 2p - 3, 2p - 7, 2p - 11,$ etc. For example, a second rank symmetric tensor splits into two invariant subspaces of dimension five and one; a third rank symmetric tensor splits into two invariant subspaces of dimension seven and three; a fourth rank symmetric tensor splits into three invariant subspaces of dimension nine, five, and one.

The transformation matrix between the monomial basis and the harmonic polynomial basis is generally complex. We shall use the hermitian conjugate of this matrix to transform the moments into complex moments. Let y be the vector of harmonic polynomials which are the basis of the symmetric tensor space of rank $p.$

$$y = [y_p^0, \dots, y_p^{-p}, y_{p-2}^0, \dots, y_{p-2}^{+2}, \dots]^T. \quad (10)$$

There exists a nonsingular complex matrix $A,$ such that

$$u = Ay, \quad (11)$$

where the vector u is given in (6). The rows of A are indexed by the three indexes lmn of moments and the columns of A are indexed by the two indexes l and m of spherical harmonics. We define the vector of complex

moments ν by

$$\nu = A^* m \tag{12}$$

where A^* is the hermitian conjugate of A . The element of ν is of the form ν_l^m , with $l = p, p - 2, p - 4, \dots$, and $-l \leq m \leq l$. If one permutes the components of u , then the components of m and the columns of A matrix will be permuted correspondingly. However, the complex moments calculated from (12) is invariant under this permutation transformation. Note that H_p can also be expressed as scalar products of y and ν . Under rotation transformation, y is transformed by the block diagonal matrix D of the form

$$D = \begin{bmatrix} D^p & & & & \\ & D^{p-2} & & & \\ & & \dots & & \\ & & & D^l & \\ & & & & \dots \end{bmatrix}, \tag{13}$$

where the matrix D^l on the diagonal of D belongs to the irreducible representation l of the rotation group, and $l = p, p - 2, p - 4, \dots$. The block diagonal structure of D implies that the complex moments ν_l^m of a given l are also transformed by the irreducible representation matrix D^l . A different derivation of complex moments is given by Kanatani [10], [11]. He used the generating function of spherical harmonics to transform moments to complex moments. Let ν_l be the vector $[\nu_l^l, \nu_l^{l-1}, \dots, \nu_l^{-l}]$. The tensor product of ν_l and $\nu_{l'}$ is transformed by the reducible representation $D^l D^{l'}$. This reducible representation can be decomposed into a direct sum of irreducible representations by the Clebsch-Gordon series [4], [8]

$$D_{m'm}^l D_{n'n}^{l'} = \sum_{j=|l-l'|}^{j=l+l'} \langle l, m', l', n' | l, l', j, m' + n' \rangle \cdot D_{m'+n', m+n}^{l'} \langle l, m, l', n | l, l', j, m + n \rangle \tag{14}$$

where $\langle l, m, l', n | l, l', j, m + n \rangle$ is the vector coupling or Clebsch-Gordon coefficient [4]. "Composite" complex moment forms which are transformed according to the irreducible representations in the Clebsch-Gordon series can be constructed as linear combinations of $\nu_l^m \nu_{l'}^{k-m}$

$$\nu(l, l')_j^k = \sum_{m=-l}^{m=l} \langle l, m, l', k - m | l, l', j, k \rangle \nu_l^m \nu_{l'}^{k-m}. \tag{15}$$

The $\nu(l, l')_j^0$ is a moment invariant. Moment invariants which are high-order homogeneous polynomials of moments can be obtained by using (15) to construct $j = 0$ invariant scalars from the tensor product of $\nu(l, l')_j^k$'s and $\nu_{l'}^m$'s. In particular, the "norm" of the complex moment function $\nu(l, l')_j^k$

$$\nu^2(l, l')_j = \sum_{k=-j}^{k=j} (-1)^{j-k} (2j + 1)^{-1/2} \cdot \nu(l, l')_j^k \nu(l, l')_j^{-k} \tag{16}$$

is a moment invariant. Unit rank tensors can also be obtained from (15) with $j = 1$. These tensors are useful in the recently developed tensor method for orientation estimation and object identification [2], [14].

III. MOMENT INVARIANTS

In the following, we shall demonstrate our method by constructing the second- and third-order moment invariants.

A. Construction of Second-Order Moment Invariants

The vector u of the monomial basis of the second rank symmetric tensor is given explicitly as

$$u = [u_1^2, u_2^2, u_3^2, u_1 u_2, u_1 u_3, u_2 u_3]^T. \tag{17}$$

The monomial $u_1^2, u_1 u_2, \dots$ can be expressed as linear combinations of y_0^m and y_2^m ($m = -2, \dots, 2$). The transformation formulas are given in [4]. The corresponding A matrix in (11) is found to be

$$A = \sqrt{\pi} \begin{bmatrix} \sqrt{\frac{2}{15}} & 0 & -\frac{2}{3\sqrt{5}} & 0 & \sqrt{\frac{2}{15}} & \frac{2}{3} \\ -\sqrt{\frac{2}{15}} & 0 & -\frac{2}{3\sqrt{5}} & 0 & -\sqrt{\frac{2}{15}} & \frac{2}{3} \\ 0 & 0 & \frac{4}{3\sqrt{5}} & 0 & 0 & \frac{2}{3} \\ -j\sqrt{\frac{2}{15}} & 0 & 0 & 0 & j\sqrt{\frac{2}{15}} & 0 \\ 0 & -\sqrt{\frac{2}{15}} & 0 & \sqrt{\frac{2}{15}} & 0 & 0 \\ 0 & j\sqrt{\frac{2}{15}} & 0 & j\sqrt{\frac{2}{15}} & 0 & 0 \end{bmatrix}. \tag{18}$$

Using (12), the complex moments are calculated to be

$$\nu_2^2 = \sqrt{\frac{2\pi}{15}} (M_{200} - M_{020} + j2M_{110}), \quad (19.a)$$

$$\nu_2^1 = \sqrt{\frac{2\pi}{15}} (-2M_{101} - j2M_{011}), \quad (19.b)$$

$$\nu_2^0 = \frac{2\sqrt{\pi}}{3\sqrt{5}} (2M_{002} - M_{200} - M_{020}), \quad (19.c)$$

$$\nu_2^{-1} = \sqrt{\frac{2\pi}{15}} (2M_{101} - j2M_{011}), \quad (19.d)$$

$$\nu_2^{-2} = \sqrt{\frac{2\pi}{15}} (M_{200} - M_{020} - j2M_{110}), \quad (19.e)$$

$$\nu_2^0 = \frac{2}{3} \sqrt{\pi} (M_{200} + M_{020} + M_{002}). \quad (19.f)$$

The complex moments in (19) agrees with those obtained by Kanatani [10]. We immediately observe that ν_2^0 is a linear moment invariant. The quadratic moment invariant is the norm of ν_2^m

$$\nu(2, 2)_0^0 = (5)^{-1/2} (2\nu_2^2\nu_2^{-2} - 2\nu_2^1\nu_2^{-1} + (\nu_2^0)^2). \quad (20)$$

The cubic moment invariant of the second-order moments can be obtained by first constructing a second rank spherical tensor η_2^m as follows:

$$\eta_2^m = \sum_{i=-2}^{i=2} \langle 2, i, 2, m-i | 2, 2, 2, m \rangle \nu_2^i \nu_2^{m-i}, \quad (21)$$

and then taking the scalar product of η_2^m and ν_2^m to get

$$\eta\nu = (5)^{-1/2} (\eta_2^2\nu_2^{-2} - \eta_2^1\nu_2^{-1} + \eta_2^0\nu_2^0 - \eta_2^{-1}\nu_2^1 + \eta_2^{-2}\nu_2^2). \quad (22)$$

The three second-order moment invariants given in [16] are

$$J_1 = M_{200} + M_{020} + M_{002}, \quad (23.a)$$

$$J_2 = M_{200}M_{020} + M_{200}M_{002} + M_{020}M_{002} - M_{110}^2 - M_{011}^2, \quad (23.b)$$

$$J_3 = M_{200}M_{020}M_{002} - M_{002}M_{110}^2 + 2M_{110}M_{101}M_{011} - M_{020}M_{101}^2 - M_{200}M_{011}^2. \quad (23.c)$$

Our results can be expressed as algebraic equations of J_1 , J_2 , and J_3 ,

$$\nu_2^0 = \frac{2\sqrt{\pi}}{3} J_1, \quad (24.a)$$

$$\nu(2, 2)_0^0 = \frac{2\pi}{15\sqrt{5}} \left(\frac{8}{3} J_1^2 - 8J_2 \right), \quad (24.b)$$

$$\eta\nu = \frac{2\pi\sqrt{2\pi}}{15\sqrt{35}} \left(\frac{-32}{9} J_1^3 + 16J_1J_2 - 48J_3 \right). \quad (24.c)$$

Therefore, it is evident that the two results are equivalent to each other.

B. Construction of Third-Order Moment Invariants

The complex third-order moments can be similarly calculated, they are given below

$$\nu_3^3 = \sqrt{\frac{\pi}{35}} [(-M_{300} + 3M_{120}) + j(M_{030} - 3M_{210})], \quad (25.a)$$

$$\nu_3^2 = \sqrt{\frac{6\pi}{35}} [(M_{201} - M_{021}) + j2M_{111}], \quad (25.b)$$

$$\nu_3^1 = \frac{\sqrt{3\pi}}{5\sqrt{7}} [(M_{300} + M_{120} - 4M_{102}) + j(M_{030} + M_{210} - 4M_{012})], \quad (25.c)$$

$$\nu_3^0 = \frac{2\sqrt{\pi}}{5\sqrt{7}} [2M_{003} - 3M_{201} - 3M_{021}], \quad (25.d)$$

$$\nu_3^{-1} = \frac{\sqrt{3\pi}}{5\sqrt{7}} [(-M_{300} - M_{120} + 4M_{102}) + j(M_{030} + M_{210} - 4M_{012})], \quad (25.e)$$

$$\nu_3^{-2} = \sqrt{\frac{6\pi}{35}} [(M_{201} - M_{021}) - j2M_{111}], \quad (25.f)$$

$$\nu_3^{-3} = \sqrt{\frac{\pi}{35}} [(M_{300} - 3M_{120}) + j(M_{030} - 3M_{210})], \quad (25.g)$$

and

$$\nu_1^1 = \frac{\sqrt{6\pi}}{5} [(-M_{300} - M_{120} - M_{102}) - j(M_{030} + M_{210} + M_{012})], \quad (26.a)$$

$$\nu_1^0 = \frac{2\sqrt{3\pi}}{5} [M_{003} + M_{201} + M_{021}], \quad (26.b)$$

$$\nu_1^{-1} = \frac{\sqrt{6\pi}}{5} [(M_{300} + M_{120} + M_{102}) - j(M_{030} + M_{210} + M_{012})]. \quad (26.c)$$

The quadratic moment invariants are the norms of ν_3^m and ν_1^m ,

$$\nu(3, 3)_0^0 = (7)^{-1/2} [2\nu_3^3\nu_3^{-3} - 2\nu_3^2\nu_3^{-2} + 2\nu_3^1\nu_3^{-1} - (\nu_3^0)^2], \quad (27.a)$$

$$\nu(1, 1)_0^0 = (3)^{-1/2} [2\nu_1^1\nu_1^{-1} - (\nu_1^0)^2]. \quad (27.b)$$

Moment invariant which is a cubic form of third-order moments vanishes, due to the symmetry property of the vector coupling coefficients in (A1.a) of Appendix A. The quartic moment invariants can be obtained by following

the general procedures outlined in Section II. "Composite" complex moment forms $\nu(l, l')_j^m$ with $l, l' = 1, 3$ and $j = \text{odd integer}$, also vanish identically because of the symmetry property in (A1.a). Other complex moment forms can be calculated by using (15). The expressions of those complex moment forms which are relevant to the construction of moment invariants, are given in Appendix B. There are only four independent quartic moment invariants,

$$\nu^2(3, 3)_2 = (5)^{-1/2} \sum_{m=-2}^2 (-1)^{(2-m)} \cdot \nu(3, 3)_2^m \nu(3, 3)_2^{-m}, \quad (28.a)$$

$$\nu^2(3, 1)_2 = (5)^{-1/2} \sum_{m=-2}^2 (-1)^{(2-m)} \cdot \nu(3, 1)_2^m \nu(3, 1)_2^{-m}, \quad (28.b)$$

$$\nu(3, 3)_2 \nu(3, 1)_2 = (5)^{-1/2} \sum_{m=-2}^2 (-1)^{(2-m)} \nu(3, 3)_2^m \cdot \nu(3, 1)_2^{-m}, \quad (28.c)$$

$$\nu(3, 1)_2 \nu(1, 1)_2 = (5)^{-1/2} \sum_{m=-2}^2 (-1)^{(2-m)} \cdot \nu(3, 1)_2^m \nu(1, 1)_2^{-m}. \quad (28.d)$$

The norms of $\nu(3, 3)_2^m$ and $\nu(3, 3)_4^m$ are linear combinations of $\nu^2(3, 3)_2$ and the square of the quadratic moment invariants $\nu(3, 3)_0^0$. The scalar product of $\nu(3, 3)_2^m$ and $\nu(1, 1)_2^m$ is a linear combination of the quartic invariant $\nu^2(3, 1)_2$ and the product of quadratic moment invariants $\nu(1, 1)_0^0 \nu(3, 3)_0^0$. The scalar product $\nu(3, 3)_4^m$ and $\nu(3, 1)_4^m$ is proportional to the moment invariants $\nu(3, 3)_2 \nu(3, 1)_2$. The norm of $\nu(1, 1)_2^m$ is proportional to the square of quadratic moment invariants $\nu(1, 1)_0^0$. It is possible to form moment invariants which contain both the second- and third-order moments. In fact, the scalar product of ν_2^m in (19) with any of the three second rank composite complex moment forms in Appendix B is a moment invariants. We write down these expressions below:

$$\nu(3, 3)_2 \nu_2 = (5)^{-1/2} \sum_{m=-2}^2 (-1)^{(2-m)} \nu(3, 3)_2^m \nu_2^{-m}, \quad (29.a)$$

$$\nu(3, 1)_2 \nu_2 = (5)^{-1/2} \sum_{m=-2}^2 (-1)^{(2-m)} \nu(3, 1)_2^m \nu_2^{-m}, \quad (29.b)$$

$$\nu(1, 1)_2 \nu_2 = (5)^{-1/2} \sum_{m=-2}^2 (-1)^{(2-m)} \nu(1, 1)_2^m \nu_2^{-m}. \quad (29.c)$$

Under the similitude transformation, i.e., the change of size, the moment, M_{lmn} , is transformed as

$$M_{lmn} \rightarrow \lambda^{p+3} M_{lmn}, \quad (30)$$

where $l + m + n = p$ and λ is the scale factor. It is obvious that the following quantity:

$$\mu_{lmn} = \frac{M_{lmn}}{M_{000}^{(p/3)+1}}, \quad (31)$$

is a similitude invariant moment. Similitude-Euclidean moment invariants can be easily derived by substituting μ_{lmn} for M_{lmn} in the equations of moment invariants. On the other hand, the authors of [16] divided various moment invariants by suitable powers of ν_0^0 , and derived similitude-invariant quantities. We follow their procedure and obtain

$$I_{22}^2 = \frac{\nu(2, 2)_0^0}{(\nu_0^0)^2}, \quad (32.a)$$

$$I_{222}^2 = \frac{\eta \nu}{(\nu_0^0)^3}, \quad (32.b)$$

$$I_{33}^3 = \frac{\nu(3, 3)_0^0}{(\nu_0^0)^{12/5}}, \quad (32.c)$$

$$I_{11}^3 = \frac{\nu(1, 1)_0^0}{(\nu_0^0)^{12/5}}, \quad (32.d)$$

$$I_{333}^{2,3} = \frac{\nu(3, 3)_2 \nu_2}{(\nu_0^0)^{17/5}}, \quad (32.e)$$

$$I_{123}^{2,3} = \frac{\nu(3, 1)_2 \nu_2}{(\nu_0^0)^{17/5}}, \quad (32.f)$$

$$I_{112}^{2,3} = \frac{\nu(1, 1)_2 \nu_2}{(\nu_0^0)^{17/5}}, \quad (32.g)$$

$$I_{3333}^3 = \frac{\nu^2(3, 3)_2}{(\nu_0^0)^{24/5}}, \quad (32.h)$$

$$I_{1333}^3 = \frac{\nu(3, 3)_2 \nu(3, 1)_2}{(\nu_0^0)^{24/5}}, \quad (32.i)$$

$$I_{1133}^3 = \frac{\nu^2(3, 1)_2}{(\nu_0^0)^{24/5}}, \quad (32.j)$$

$$I_{1113}^3 = \frac{\nu(3, 1)_2 \nu(1, 1)_2}{(\nu_0^0)^{24/5}}. \quad (32.k)$$

Under the improper transformation, $\vec{x} \rightarrow -\vec{x}$, moments (complex moments) of odd order will obtain a negative sign, but moments (complex moments) of even order will remain unchanged. The eleven moment invariants in (32) are invariant under spatial inversion. The numerical study of these quantities, including roundoff effects, will be given in Section V.

IV. MOMENT FORMS AND MOTION ESTIMATION OF 3-D OBJECTS

Moments have also been used to estimate the affine transform parameters of 2-D and 3-D images. Cyganski

and Orr developed a tensor-based technique to recover the affine transform of a 2-D image resulting from the orthogonal projection of a rigid planar patch [2], [14]. Their method has been extended to affine transform determination for 3-D objects by Faber and Stokely [5]. Faber and Stokely calculated a set of central moments, and then derived several unit rank tensors (vectors) by contracting the tensor indexes of the moments with the tensor indexes of the totally antisymmetric tensor ϵ^{ijk} . However, the unit rank tensors derived by their method may often contain the fourth- and the fifth-order central moments. In the tensor approach of motion estimation, the linear part of the affine transform is determined by solving the following system of linear equations for r

$$o = Pr, \tag{33}$$

where the array o contains the x, y, z components of three vectors b', v', w' , which are derived from the central moments of the object to be identified. The vector components are arranged in o as follows:

$$o = [b'_x, v'_x, w'_x, b'_y, v'_y, w'_y, b'_z, v'_z, w'_z]^T. \tag{34}$$

The array r contains the matrix elements of the affine transform.

$$r = [R_{11}, R_{12}, R_{13}, R_{21}, R_{22}, R_{23}, R_{31}, R_{32}, R_{33}]^T. \tag{35}$$

The 9×9 matrix P is a block diagonal matrix with three identical 3×3 matrices on the diagonal. This 3×3 matrix has the following expression:

$$\begin{bmatrix} b_x & b_y & b_z \\ v_x & v_y & v_z \\ w_x & w_y & w_z \end{bmatrix} \tag{36}$$

where $b_x, b_y, \dots, v_x, v_y, \dots, w_x, \dots$, are the components of the corresponding vectors b, v, w , which are derived from the central moments of the prototype. Once the linear part of the affine transform is determined, the translation parameters can be determined from the transformation property of the first-order noncentral moments.

Practically, we may assume that the linear part of the affine transform is simply a product of a constant and an orthogonal matrix which specify the relative size and orientation of the object and the prototype, respectively. In this case, the calculation of vectors using (15) can save computation effort. In fact, v_1^m in (26) is the already available vector, and we identify it with b . Using (15), v can be calculated from the product of v_2^m and v_1^m , and w can be calculated from the product of v_3^m and v_2^m . Their expressions are given in Appendix C. In the expressions of v and w , there is no term which contains central moments of order higher than three. Moreover, each component of the vector is an accumulative sum of products of two complex moments. This makes tensor technique for motion estimation easier to be implemented on digital computer. We use the following equations to define the x, y, z components of a vector in terms of its components in the

spherical harmonic basis

$$\begin{aligned} a_x &= [a(-1) - a(1)]/\sqrt{2} \\ a_y &= j[a(-1) + a(1)]/\sqrt{2} \\ a_z &= a(0), \quad (a = b, v, w). \end{aligned} \tag{37}$$

Because, only in rectangular coordinate system, the rotation transformation matrix for a rank one tensor is a real orthogonal matrix.

There are other algorithms which use moments for 2-D and 3-D object positioning. The position and orientation of a 2-D image can be determined from the first- and the second-order moments of the image. The eigenvectors of the second-moment matrix are principal axes. The orientation of the 2-D image is defined to be the orientation of the principal axis of the least inertia which is also known as axis of elongation [9]. The 2-D rotation that rotates one image to the other is easily determined by measuring the angle between the elongation axes of these two images. This approach can be generalized to find 3-D motion parameters. An algorithm for 3-D orientation estimation using principal axes is discussed in the rest of this section. The basic assumption is that the 3-D coordinates of a set of feature points on a rigid body are known at two frames. Moments are derived from the 3-D data of these feature points. The parameters in the rotation matrix can be recovered from these moments. Let Q_1 and Q_2 be the matrices of the second-order central moments in the first and second frames, respectively,

$$Q_i = \begin{bmatrix} M(i)_{200} & M(i)_{110} & M(i)_{101} \\ M(i)_{110} & M(i)_{020} & M(i)_{011} \\ M(i)_{101} & M(i)_{011} & M(i)_{002} \end{bmatrix}, \quad i = 1, 2. \tag{38}$$

The second-order moments in the matrix are calculated by (1), where the density function is a sum of delta functions located at each feature point. The origin of the coordinate system is at the centroid of the density function. From (1) and (8), the following relation between Q_1 and Q_2 can be obtained

$$Q_2 = RQ_1R^T. \tag{39}$$

The real symmetric moment matrices can be decomposed as

$$\begin{aligned} Q_1 &= AA^T, \\ Q_2 &= B\Lambda'B^T \end{aligned} \tag{40}$$

where Λ and Λ' are the diagonal eigenvalue matrices of Q_1 and Q_2 , respectively. A and B are orthogonal matrices whose column vectors are the corresponding eigenvectors of Λ and Λ' , respectively. In the noise free case, Q_1 and Q_2 have common eigenvalues, so $\Lambda = \Lambda'$. The eigenvalues are determined by the following characteristic equation:

$$f(\lambda) = \lambda^3 - J_1\lambda^2 + J_2\lambda - J_3 \tag{41}$$

where J_1, J_2, J_3 are the second-order moment invariants given in (23). The rotation matrix R is determined from the correlation of the principal axes in the two frames,

$$R = BA^T = \sum_{i=1}^3 b_i a_i^T \quad (42)$$

where a_i and b_i are eigenvectors of Q_1 and Q_2 , respectively. Owing to the uncertainties in the signs of the column vectors in B , four possible rotation matrices can be constructed. Traditionally, in the 2-D case, the "two way ambiguity" of principal axes is solved by computing the values of higher order moments [7], [9]. We shall also use the values of higher order moment forms to select the correct rotation matrix. Our procedure is given in the following. The third-order moments of the feature sets are first calculated. The Cartesian coordinates of the vector form b is then constructed from the third-order moments using (26) and (37). The vector b is the same one used in (34) and (36). Let b and b' be the values of the vector form b evaluated in the first and the second frames, respectively. b is transformed by the four rotation matrices. The transformed vectors are matched with the vector b' to determine the correct rotation matrix. Even if the data are noisy, this algorithm can still choose the rotation matrix which is the closest to the correct one.

We must point out that the principal axes algorithm is quite similar to the one which is recently proposed by Lin *et al.* [12]. Equations (39)-(42) all appeared in [12], even though the terms of moments, moment invariants and principal axes are never mentioned in their paper. Moreover, in [12], the correct rotation matrix cannot be determined, until the exact point correspondences of the feature set have been established. In their approach, each feature point in the first frame has to be transformed by four possible rotations. Heuristic searching and matching between these transformed points and the feature points in the second frame will discover the point correspondence. Our algorithm can save a lot of computing effort, especially when large number of feature points are used. The computer simulation results in [12] indicate that the algorithm may not correctly predict the motion parameters when noise is strongly present in the data. In this case, this algorithm can still be used to determine feature point correspondences [12]. The rotation and translation parameters will be those that minimize

$$\text{Min} \sum_{i=1}^N \|R x_i + t - x'_i\|.$$

The solution of the minimization problem can be found in [6], [17].

If only small amount of noise is present in the data, the orthogonal matrices A , B , and R can all be calculated by the perturbation method [13]. Let \tilde{Q}_i , ($i = 1, 2$), be the central moment matrix of the noisy feature points in the i th frame. The \tilde{Q}_i can be represented by $Q_i + V_i$, where Q_i is the second-order central moment matrix when the noise is absent from the i th frame, and V_i is the sym-

metric perturbation matrix due to noise. The matrix elements of V_i are small. The eigenvectors of \tilde{Q}_i can be calculated from the eigenvectors of Q_i by the perturbation formula [13]

$$\begin{aligned} a_i &= N_{1i} \left[a_i - \sum_{k \neq i} \frac{a_k^T V_1 a_i}{\tilde{\lambda}_{1i} - \tilde{\lambda}_{1k}} a_k + \dots \right] \\ b_i &= N_{2i} \left[b_i - \sum_{k \neq i} \frac{b_k^T V_2 b_i}{\tilde{\lambda}_{2i} - \tilde{\lambda}_{2k}} b_k + \dots \right] \\ i, k &= 1, 2, 3 \end{aligned} \quad (43)$$

where a_i, b_i, \tilde{a}_i , and \tilde{b}_i are eigenvectors of Q_1, Q_2, \tilde{Q}_1 , and \tilde{Q}_2 , respectively. N_{1i} and N_{2i} are the normalization constants of the corresponding eigenvectors. We can express the rotation matrix R in terms of the matrix \tilde{R} which is calculated from (42) by substituting \tilde{a}_i for a_i and \tilde{b}_i for b_i in that equation, and the outer products of eigenvectors \tilde{b}_i and \tilde{a}_i .

$$\begin{aligned} R &= \tilde{R} - \sum_{i=1}^3 \sum_{k \neq i} \left(\frac{b_k^T V_2 b_i}{\tilde{\lambda}_{2i} - \tilde{\lambda}_{2k}} \right) \tilde{b}_k \tilde{a}_i^T \\ &\quad - \sum_{i=1}^3 \sum_{k \neq i} \left(\frac{a_k^T V_1 a_i}{\tilde{\lambda}_{1i} - \tilde{\lambda}_{1k}} \right) \tilde{b}_i \tilde{a}_k^T + \dots, \end{aligned} \quad (44)$$

the omitted terms are higher order terms of $(a_k^T V_1 a_i)/(\tilde{\lambda}_{1i} - \tilde{\lambda}_{1k})$ and $(b_k^T V_2 b_i)/(\tilde{\lambda}_{2i} - \tilde{\lambda}_{2k})$. It is known that, in the nondegenerate perturbation theory, the diagonal matrix element of V_i is the first-order correction to the eigenvalue. For example, $b_i^T V_2 b_i = \tilde{\lambda}_{2i} - \lambda_i$, where $\tilde{\lambda}_{2i}$ and λ_i are the i th eigenvalues of \tilde{Q}_2 and Q_2 , respectively [13]. These eigenvalues are moment invariants determined from (41), where J_1, J_2, J_3 are evaluated using both noisy and noise-free data. On the other hand, the off diagonal elements $a_k^T V_1 a_i$ and $b_k^T V_2 b_i$ are not related to the moment invariants. Note that $(a_k^T V_1 a_i)/(\tilde{\lambda}_{1i} - \tilde{\lambda}_{1k})$ and $(b_k^T V_2 b_i)/(\tilde{\lambda}_{2i} - \tilde{\lambda}_{2k})$ are the first-order approximations of the scalar products of unit vectors a_i, a_k and b_i, b_k ($i \neq k$). Therefore, their magnitudes should be less than one. If noise model is available and the perturbation matrix elements can be estimated, we can use (44) to estimate the rotation matrix R . The denominators of $(a_k^T V_1 a_i)/(\tilde{\lambda}_{1i} - \tilde{\lambda}_{1k})$ and $(b_k^T V_2 b_i)/(\tilde{\lambda}_{2i} - \tilde{\lambda}_{2k})$ are predominantly determined by the geometric distribution of the feature points on the object. It is obvious that the noise contamination effect is highly suppressed if the eigenvalues are very distinct from each other. Following this perturbative approach, we can use the second-order moment invariants to predict the performance of this algorithm when the numerical roundoff effect and perturbative noise are present in the data.

V. COMPUTER SIMULATION AND NUMERICAL RESULTS

We have implemented our algorithm of deriving various 3-D moment forms in Fortran 77. The invariance of the 3-D moment invariants has been confirmed exper-

mentally. A rectangular solid was used in the experiment. It was rotated to various orientations, with one of its corner fixed at the origin of the coordinate system. At each orientation, we calculated the second- and the third-order moment invariants and the vectors b , v , and w . In order to reduce the data storage and increasing the processing speed, the moments were calculated using a set of feature points. In general, the contours, corners and vertices are the choices of feature sets. The only restriction on the selection of feature sets is that they preserve the essential geometric information of the object. We chose the corners of the rectangular solid as the feature set and expressed the density function as a sum of delta functions. The moment invariants calculated at various orientations agreed exactly with each other. This checked our equations for moment invariants. A similar experiment using pyramid was also done and the result also confirmed the theory. The rotation matrix was then recovered by using (33), where the three vectors b , v , w were evaluated at two different orientations of the object. The calculated matrix elements agreed, in the first three decimal places, with those in the orthogonal matrix used to rotate the geometric object to the new orientation.

The impact of digitizing errors to the moment invariants was experimentally studied in both fine and coarse sampling cases. In case A, the rectangular solid was sampled into $30 \times 45 \times 75$ volume pixels. In case B, the same solid was sampled into $6 \times 9 \times 15$ volume pixels. Each corner point was first rotated by the orthogonal matrix to its new position, and then rounded to the nearest volume pixel. These pixel values were used to calculate the moment invariants. The invariants exhibited small variations at different orientations. Using large number of data, we obtained the means and standard deviations of those similitude-Euclidean moment invariants. Their values are given in Table I. The moment invariants were also calculated from the discrete data of a pyramid with the same two resolutions. Case C was for the better resolution, and case D was for the poor resolution. The statistical parameters of moment invariants in both cases were calculated and given in Table II. Note that different moment invariants may have different order of magnitude. Therefore, they should be normalized by their sample variance before they are used to construct feature space. The results in the tables indicate that the variances have weak dependence on the shape of the geometric objects. The quadratic moment invariants are less sensitive to digitizing noise. On the other hand, the cubic and quartic moment invariants have relatively large standard deviations in coarse sampling. Nevertheless, we can see from our experiment results that the digitizing noise does not significantly reduce the discriminating power of the moment invariant features.

The effect of digitizing noise to the tensor algorithm for motion estimation was experimentally studied. The rectangular solid were sampled as in case A and case B. The vectors b , v , w were calculated at various orientations in each case. The transformation matrix was calculated by (33). After performing a large number experiments, we found that, the weak digitizing noise of fine sampling only

TABLE I
MEANS AND STANDARD DEVIATIONS OF MOMENT INVARIANTS EVALUATED USING THE CORNER POINTS OF A RECTANGLE

invariant	mean (A)	stand. dev (A)	mean (B)	stand. dev (B)	exact
$I_2 (10^{-4})$	69.92	0.15	70.00	3.75	69.57
$I_{20} (10^{-7})$	-13.62	0.01	-13.70	0.70	-13.53
$I_{21} (10^{-4})$	-39.95	0.04	-40.10	0.87	-39.42
$I_{22} (10^{-4})$	-2.41	0.01	-2.43	0.09	-2.10
$I_{23} (10^{-4})$	4.53	0.06	4.51	0.29	4.52
$I_{24} (10^{-4})$	-4.75	0.04	-4.73	0.22	-4.74
$I_{30} (10^{-4})$	19.70	0.10	19.60	0.70	19.60
$I_{31} (10^{-4})$	3.63	0.03	3.68	0.26	3.59
$I_{32} (10^{-4})$	-3.18	0.03	-3.40	0.30	-3.15
$I_{33} (10^{-4})$	3.33	0.03	3.36	0.20	3.30
$I_{34} (10^{-4})$	-13.40	0.04	-13.32	0.94	-13.30

TABLE II
MEANS AND STANDARD DEVIATIONS OF MOMENT INVARIANTS EVALUATED USING THE CORNER POINTS OF A PYRAMID

invariant	mean(C)	stand. dev (C)	mean(D)	stand. dev (D)	exact
$I_2 (10^{-4})$	53.43	0.29	53.18	2.28	53.00
$I_{20} (10^{-4})$	-0.02	0.13	0.84	0.45	0.00
$I_{21} (10^{-4})$	-6.45	0.05	-6.32	0.42	-6.10
$I_{22} (10^{-4})$	-51.45	0.47	-51.81	1.48	-51.32
$I_{23} (10^{-4})$	8.09	0.15	7.65	1.51	7.88
$I_{24} (10^{-4})$	-5.52	0.05	-5.41	0.57	-5.42
$I_{30} (10^{-4})$	15.70	0.40	15.87	1.33	15.48
$I_{31} (10^{-4})$	13.97	0.37	13.06	3.97	13.43
$I_{32} (10^{-4})$	-10.37	0.14	-9.82	1.92	-10.07
$I_{33} (10^{-4})$	8.62	0.08	8.43	0.86	8.16
$I_{34} (10^{-4})$	-7.17	0.54	-6.92	2.92	-6.50

produced 0.3-0.8 percent error to each vector component; and the strong digitizing noise of coarse sampling produced 1-3 percent error to each vector component. The calculated transformation matrices using data of high resolution were only approximately orthogonal, and each matrix element had approximately 10 percent error from the exact value. On the other hand, the calculated transformation matrices using coarsely sampled data were very different from correct ones. Our finding seems to be contradictory to the results of [5], where they claimed that no significant numerical roundoff effect had been observed. The explanation of this discrepancy may be that they used the whole binary image, not a few feature points, to calculate the moments. Moreover, the resolution of their digital image might be fine, so the digitizing errors were not significant.

The empirical study on the noise sensitivity in principal axes algorithm for orientation estimation is given as fol-

lows. Twenty feature points were randomly chosen on the surfaces which are described by the parametric equations

$$\begin{aligned}x &= a \cos \phi \sin \theta, \\y &= b \sin \phi \sin \theta, \\z &= c \cos \theta.\end{aligned}\quad (45)$$

Four sets of values were given to the parameters a , b , c . In case E, $a = 5$, $b = 17$, $c = 32$, the object is an ellipsoid. In case F, $a = b = c = 10$, the object is a sphere. In case G, $a = b = 5$, $c = 32$, the object has a shape of a cigar. In case H, $a = 5$, $b = c = 20$, the object has a shape of a plate. In each run, the feature points on the four surfaces were determined by the twenty pairs of polar angles (ϕ, θ) which were randomly generated. The zero mean and unit variance Gaussian noise was added to the first and the second frames. The moment matrices were calculated in both frames. The rotation matrices of these four objects were computed using (40) and (42). This experiment was repeated many times, and the results of every run were at least in gross agreement with the theoretical prediction. We present, as a representative, the eigenvalues and the estimated rotation matrices of one experiment run. The estimated rotation matrices are given in (46), and the eigenvalues are given in Table III.

Case E (ellipsoid)

$$\begin{bmatrix}-0.618 & 0.186 & 0.764 \\0.533 & -0.615 & 0.581 \\0.577 & 0.767 & 0.281\end{bmatrix}$$

Case F (sphere)

$$\begin{bmatrix}-0.664 & 0.147 & 0.733 \\0.589 & -0.500 & 0.635 \\0.460 & 0.853 & 0.246\end{bmatrix}$$

Case G (cigar)

$$\begin{bmatrix}-0.601 & -0.217 & 0.769 \\0.798 & -0.229 & 0.558 \\0.005 & 0.949 & 0.311\end{bmatrix}$$

Case H (plate)

$$\begin{bmatrix}-0.624 & 0.200 & 0.755 \\0.547 & -0.578 & 0.606 \\0.558 & 0.791 & 0.251\end{bmatrix}$$

The correct rotation matrix

$$\begin{bmatrix}-0.622 & 0.143 & 0.769 \\0.548 & -0.622 & 0.559 \\0.559 & 0.769 & 0.309\end{bmatrix}\quad (46)$$

Note that, in each frame, the feature points on the "cigar" have two "almost degenerate" eigenvalues of the moment matrix. The third eigenvalue is much larger than these two. The two terms with small denominators which are the differences of these "almost degenerate" eigen-

TABLE III
THE EIGENVALUES ($\lambda_1, \lambda_2, \lambda_3$) OF THE SECOND-ORDER MOMENT MATRICES CALCULATED USING 3-D DATA OF FEATURE POINTS ON VARIOUS SURFACES IN EITHER OF THE TWO FRAMES. E-1 REPRESENTS THE FEATURE SET ON THE ELLIPSOID (CASE E) IN THE FIRST FRAME. G-2 REPRESENTS THE FEATURE SET ON THE CIGAR SHAPE SURFACE (CASE G) IN THE SECOND FRAME, ETC.

Data	λ_1	λ_2	λ_3
E-1	122.53	1082.39	9998.08
E-2	153.91	1037.63	9985.47
F-1	293.29	588.76	1195.95
F-2	281.52	593.79	1152.89
G-1	89.17	165.66	9984.17
G-2	102.34	156.39	10013.27
H-1	122.18	1485.36	3998.48
H-2	137.73	1421.46	3959.81

values dominate (44). These two terms strongly enhanced the noise effect, and the estimated rotation matrix deviates very much from the correct one. We also found that the randomly generated feature points on the ellipsoid whose three axes are all different, often have very nondegenerate geometric distribution.

VI. CONCLUDING REMARKS

In this paper, a general theory of 3-D moment invariants is proposed. The application of 3-D moment method to motion estimation is also discussed. The complex moments are defined. They are grouped into multiplets, such that each multiplet of complex moments transforms irreducibly under 3-D rotations. Using the group-theoretic technique, various 3-D moment invariants can be constructed from the irreducible complex moment multiplets. The moment invariants have the capability of classifying and identifying 3-D objects irrespective of the frame of reference. A set of the second order and third-order moment invariants are explicitly derived in the paper. These invariant features have good discriminating power for machine identification of 3-D objects or CAT images using statistical pattern recognition method. Vectors which consist of the third order moments can be similarly derived from complex moments. The vector form can be used to remove the two-way ambiguity in defining the direction of the principal axis, so that the rotation matrix can be unambiguously determined from the relative orientation of the principal axes of the original and rotated frames. Moreover, based on a perturbation formula, we have shown that the second-order moment invariants can be used to predict whether the estimation using noisy data is reliable or not. The new derivation of vector forms also facilitates the calculation of motion estimation in tensor approach.

APPENDIX A

A brief summary of the symmetry properties of the vector coupling coefficients which have been used in the paper are given below.

$$\begin{aligned} \langle l, m, l', n | l, l', j, m+n \rangle \\ = (-1)^{l+l'-j} \langle l', n, l, m | l, l', j, m+n \rangle \end{aligned} \quad (A1.a)$$

$$\begin{aligned} \langle l, m, l', n | l, l', j, m+n \rangle \\ = (-1)^{l+l'-j} \langle l, -m, l', -n | l, l', j, -m-n \rangle \end{aligned} \quad (A1.b)$$

$$\begin{aligned} \langle l, m, l', n | l, l', j, m+n \rangle \\ = (-1)^{l'+n} \sqrt{\frac{2j+1}{2l+1}} \\ \cdot \langle l', -n, j, m+n | l', j, l, m \rangle. \end{aligned} \quad (A1.c)$$

Other symmetry relations, and the recursive relations for evaluation of vector coupling coefficients can be found in [4].

APPENDIX B

The third-order complex moment forms which are relevant to the construction of quartic moment invariants are given below.

$$\nu(3, 3)_2^2 = 2\sqrt{\frac{5}{42}} \nu_3^3 \nu_3^{-1} - 2\sqrt{\frac{5}{21}} \nu_3^2 \nu_3^0 + \sqrt{\frac{2}{7}} (\nu_3^1)^2 \quad (B1.a)$$

$$\nu(3, 3)_2^1 = \frac{5}{\sqrt{21}} \nu_3^3 \nu_3^{-2} - \sqrt{\frac{5}{7}} \nu_3^2 \nu_3^{-1} + 2\sqrt{\frac{1}{42}} \nu_3^1 \nu_3^0 \quad (B1.b)$$

$$\nu(3, 3)_2^0 = \frac{5}{\sqrt{21}} \nu_3^3 \nu_3^{-3} - \sqrt{\frac{3}{7}} \nu_3^2 \nu_3^{-1} + \frac{2}{\sqrt{21}} (\nu_3^0)^2 \quad (B1.c)$$

$$\nu(3, 3)_2^{-1} = \frac{5}{\sqrt{21}} \nu_3^3 \nu_3^{-3} - \sqrt{\frac{5}{7}} \nu_3^{-2} \nu_3^1 + \frac{2}{\sqrt{42}} \nu_3^{-1} \nu_3^0 \quad (B1.d)$$

$$\nu(3, 3)_2^{-2} = 2\sqrt{\frac{5}{42}} \nu_3^{-3} \nu_3^1 - 2\sqrt{\frac{5}{21}} \nu_3^{-2} \nu_3^0 + \sqrt{\frac{2}{7}} (\nu_3^{-1})^2 \quad (B1.e)$$

$$\nu(3, 1)_2^2 = \frac{-1}{\sqrt{105}} \nu_3^2 \nu_1^0 + \frac{1}{\sqrt{35}} \nu_3^2 \nu_1^{-1} + \frac{1}{5\sqrt{21}} \nu_3^1 \nu_1^1 \quad (B2.a)$$

$$\nu(3, 1)_2^1 = \sqrt{\frac{2}{105}} \nu_3^2 \nu_1^{-1} + \frac{1}{5\sqrt{7}} \nu_3^0 \nu_1^1 - \frac{2}{5} \sqrt{\frac{2}{21}} \nu_3^1 \nu_1^0 \quad (B2.b)$$

$$\nu(3, 1)_2^0 = \frac{-1}{5} \sqrt{\frac{3}{7}} \nu_3^0 \nu_1^0 + \frac{1}{5} \sqrt{\frac{2}{7}} \nu_3^1 \nu_1^{-1} + \frac{1}{5} \sqrt{\frac{2}{7}} \nu_3^{-1} \nu_1^1 \quad (B2.c)$$

$$\nu(3, 1)_2^{-1} = \sqrt{\frac{2}{105}} \nu_3^{-2} \nu_1^1 + \frac{1}{5\sqrt{7}} \nu_3^0 \nu_1^{-1} - \frac{2}{5} \sqrt{\frac{2}{21}} \nu_3^{-1} \nu_1^0 \quad (B2.d)$$

$$\nu(3, 1)_2^{-2} = \frac{-1}{\sqrt{105}} \nu_3^{-2} \nu_1^0 + \frac{1}{\sqrt{35}} \nu_3^{-3} \nu_1^1 + \frac{1}{5\sqrt{21}} \nu_3^{-1} \nu_1^{-1} \quad (B2.e)$$

$$\nu(1, 1)_2^2 = \frac{1}{5} (\nu_1^1)^2 \quad (B3.a)$$

$$\nu(1, 1)_2^1 = \frac{\sqrt{2}}{5} \nu_1^0 \nu_1^1 \quad (B3.b)$$

$$\nu_2^0 = \frac{1}{5} \sqrt{\frac{2}{3}} \{ (\nu_1^0)^2 + \nu_1^1 \nu_1^{-1} \} \quad (B3.c)$$

$$\nu_2^{-1} = \frac{\sqrt{2}}{5} \nu_1^0 \nu_1^{-1} \quad (B3.d)$$

$$\nu_2^{-2} = \frac{1}{5} (\nu_1^{-1})^2 \quad (B3.e)$$

APPENDIX C

Vector forms which are derived from the second- and third-order complex moments are given below.

$$\nu(1) = \sqrt{\frac{1}{10}} \nu_2^0 \nu_1^1 - \sqrt{\frac{3}{10}} \nu_2^1 \nu_1^0 + \sqrt{\frac{3}{5}} \nu_2^2 \nu_1^{-1} \quad (C1.a)$$

$$\nu(0) = \sqrt{\frac{3}{10}} \nu_2^{-1} \nu_1^1 - \sqrt{\frac{2}{5}} \nu_2^0 \nu_1^0 + \sqrt{\frac{3}{10}} \nu_2^1 \nu_1^{-1} \quad (C1.b)$$

$$\nu(-1) = \sqrt{\frac{3}{5}} \nu_2^{-2} \nu_1^1 - \sqrt{\frac{3}{10}} \nu_2^{-1} \nu_1^0 + \sqrt{\frac{1}{10}} \nu_2^0 \nu_1^{-1} \quad (C1.c)$$

$$\begin{aligned} w(1) = & \sqrt{\frac{1}{35}} \nu_3^{-1} \nu_3^2 - \sqrt{\frac{3}{35}} \nu_3^0 \nu_3^1 + \sqrt{\frac{6}{35}} \nu_3^1 \nu_3^0 \\ & - \sqrt{\frac{2}{7}} \nu_3^2 \nu_2^{-1} + \sqrt{\frac{3}{7}} \nu_3^1 \nu_2^{-2} \end{aligned} \quad (C2.a)$$

$$\begin{aligned} w(0) = & \sqrt{\frac{1}{7}} \nu_3^{-2} \nu_3^2 - 2\sqrt{\frac{2}{35}} \nu_3^{-1} \nu_3^1 + 3\sqrt{\frac{1}{35}} \nu_3^0 \nu_3^0 \\ & - 2\sqrt{\frac{2}{35}} \nu_3^1 \nu_2^{-1} + \sqrt{\frac{1}{7}} \nu_3^2 \nu_2^{-2} \end{aligned} \quad (C2.b)$$

$$w(-1) = \sqrt{\frac{3}{7}} \nu_3^{-2} \nu_2^2 - \sqrt{\frac{2}{7}} \nu_3^{-2} \nu_2^1 + \sqrt{\frac{6}{35}} \nu_3^{-1} \nu_2^0 \\ - \sqrt{\frac{3}{35}} \nu_3^0 \nu_2^{-1} + \sqrt{\frac{1}{35}} \nu_3^{-1} \nu_2^{-2} \quad (C2.c)$$

REFERENCE

- [1] Y. S. Abu-Mostafa and D. Psaltis, "Recognitive aspects of moment invariants," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. PAMI-6, pp. 698-706, Nov. 1984.
- [2] D. Cyganski and J. A. Orr, "Applications of tensor theory to object recognition and orientation determination," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. PAMI-7, pp. 662-673, Nov. 1985.
- [3] S. Dudani, K. J. Breeding, and R. B. McGhee, "Aircraft identification by moment invariants," *IEEE Trans. Comput.*, vol. C-26, no. 1, pp. 39-46, Jan. 1977.
- [4] A. R. Edmonds, *Angular Momentum in Quantum Mechanics*. Princeton, NJ: Princeton Univ. Press, 1974.
- [5] T. L. Faber and E. M. Stokely, "Affine transform determination for 3-D objects: A medical imaging application," *IEEE Conf. Comput. Vision Pattern Recogn.*, 1986, pp. 440-445.
- [6] O. D. Faugeras and M. Hebert, "The representation, recognition, and positioning of 3-D shapes from range data," in *Three-Dimensional Machine Vision*, T. Kanade, Ed. New York: Kluwer Academic, 1986.
- [7] M. K. Hu, "Visual pattern recognition by moment invariants," *IEEE Trans. Inform. Theory*, vol. IT-8, pp. 179-187, Feb. 1962.
- [8] M. Hamermesh, *Group Theory and Its Application to Physical Problems*. New York: Addison-Wesley, 1962.
- [9] B. K. P. Horn, *Robot Vision*. New York: McGraw-Hill, 1986.
- [10] K.-I. Kanatani, "Distribution of directional data and fabric tensors," *Int. J. Eng. Sci.*, vol. 22, no. 2, pp. 149-164, 1984.
- [11] —, "Stereological determination of structural anisotropy," *Int. J. Eng. Sci.*, vol. 22, no. 5, pp. 531-546, 1984.
- [12] Z. Lin, H. Lee, and T. S. Huang, "Finding 3-D point correspondences in motion estimation," *The Eighth Int. Conf. Pattern Recogn.*, 1986, pp. 303-305.
- [13] H. Margenau and G. M. Murphy, *Mathematics in Physics and Chemistry*. Princeton, NJ: Van Nostrand, 1964.
- [14] Z. Pinjo, D. Cyganski and J. A. Orr, "Determination of 3-D object orientation from projections," *Pattern Recogn. Lett.*, vol. 3, no. 5, pp. 351-356, Sept. 1985.
- [15] J. J. Reddi, "Radial and angular moment invariants for image identification," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. PAMI-3, pp. 240-242, Mar. 1981.
- [16] F. A. Sadjadi and E. L. Hall, "Three dimensional moment invariants," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. PAMI-2, pp. 127-136, Mar. 1980.

- [17] J. T. Schwartz and M. Sharir, "Identification of partially obscure objects in two dimensions by matching of noisy characteristic curves," *Int. J. Robot. Res.*, vol. 6, no. 2, pp. 29-44, 1987.
- [18] M. R. Teague, "Image analysis via the general theory of moments," *J. Opt. Soc. Amer.*, vol. 70, pp. 920-930, Aug. 1980.
- [19] C.-H. Teh and R. T. Chin, "On image analysis by the method of moments," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 10, pp. 496-513, July 1988.
- [20] E. Wigner, *Group Theory and Its Application to Quantum Mechanics of Atomic Spectra*. New York: Academic, 1959.
- [21] R. Y. Wong and E. L. Hall, "Scene matching with invariant moments," *Comput., Graph., Image Processing*, vol. 8, pp. 16-24, 1978.



Chong-Huah Lo (S'89) received the B.S. degree from National Cheng Kung University, Taiwan in 1973, the M.S. degree from National Tsing Hua University, Taiwan, in 1975, and the Ph.D. degree from Syracuse University, NY, in 1983, all in physics. He received the M.S. degree in electrical engineering from State University of New York at Stony Brook in 1988.

He has been a postdoctoral research associate in the Theoretical Physics Institute of University of Alberta, Canada, and published several research papers on particle physics and field theory. Currently, he is a full time student in the Department of Electrical Engineering, SUNY at Stony Brook, New York, working toward the Ph.D. degree in electrical engineering. His research interests include pattern recognition, computer vision, digital signal processing, and physics.

Dr. Lo is a member of the IEEE Computer Society.



Hon-Son Don (S'82-M'85) received the B.S. degree from National Taiwan University, Taipei, Taiwan, in 1977, the M.S. degree from University of Texas, Austin, TX, in 1981, and the Ph.D. degree from Purdue University, West Lafayette, IN, in 1985, all in electrical engineering.

He joined the Department of Electrical Engineering, SUNY at Stony Brook in 1985 and is currently an Assistant Professor there. He received an Outstanding Paper Award from the Pattern Recognition Society in 1985. His research interests include computer vision, pattern recognition, image processing, computer graphics, artificial intelligence, and robotics.

Dr. Don is a member of the IEEE Computer Society.

Investigation of Optimal Sampling Resolution in Digital Laser Speckle Correlation

D.J.Chen and F.P. Chiang¹

ABSTRACT—A statistical analysis of the subjective laser speckle pattern is performed with emphasis on its spectral distribution. An optimal sampling resolution of the laser speckle pattern is obtained using the Shannon sampling theorem. Experiments show that the displacement-only correlation using laser speckles is a reliable measurement as long as the sampling resolution requirement is satisfied. Upper limits of measurement such as accuracy, measurable area and measurable deformation are discussed. A method of changing the measurement capabilities by adjustment of the speckle recording arrangement is also given.

¹D.J. Chen (SEM Student Member) is Graduate Research Assistant, and F.P. Chiang (SEM Fellow) is Leading Professor, Department of Mechanical Engineering, State University of New York, Stony Brook, NY 11794-2300.

Primary paper was presented at the 1989 SEM Conference on Experimental Mechanics held in Cambridge, MA on May 28-June 1.

1 Introduction

Laser speckle photography is an ideal technique in experimental stress analysis. It is noncontact, nondestructive and remote sensing; has a high sensitivity; reveals whole-field deformation; is applicable to static as well as dynamic, 2-D as well as 3-D problems.^{1,2} Its basic process consists of image recording, film development and fringe pattern analysis. A more straightforward approach for deformation measurement is the digital speckle correlation approach.³⁻¹¹ This technique keeps all the advantages of the conventional speckle photography but requires neither photograph development nor fringe pattern analysis. It was first created for displacement searching of specimen surface^{3,4} and later on well developed in the measurement of both displacement components and displacement gradients.⁵⁻⁸ It has been employed not only in optical speckle method but also in acoustical speckle method,⁹ with applications found both in solid mechanical and fluid mechanical problems.¹⁰

However, unlike the conventional speckle photography in which the frame resolution of recording film is generally abundant, the digital resolution of a video camera is limited. Although a finer sampling of the speckle patterns will result in a higher matching coefficient in the correlation evaluation, but the measured real area will be reduced for a given video camera. Experience of previous researches show that the results obtained from laser speckle method are not as satisfactory as those from white-light speckle method.⁶ This observation is explained by the insufficient sampling of the small speckles and by the easy decorrelation of the laser speckles during

the surface deformation. Our pioneer experiments also revealed that unimodal result was easily achieved in the two-dimensional searching of the displacement components, but not in the six-dimensional searching of both the displacement components and the displacement gradients. ⁷ Therefore high quality images are required to achieve accurate displacement gradients even in the white-light speckle method. Moreover, the six-parameter searching correlation is a time consuming process. Several improved numerical approaches have been developed to reduce the computation. ⁵⁻⁸ Effects of subpixel image restoration on digital correlation error in the white-light speckle method have also been estimated. ¹¹

As known, laser speckle method has its own advantage over white-light speckle approach because of the unnecessary of surface treatment. In this paper, we explore the utilization of the displacement-only correlation using laser speckles. A statistical investigation of subjective speckle patterns is performed with emphasis on spectral analysis. Results show that the spectral distribution of a laser speckle pattern, which determines the sampling resolution requirement, is controlled by the optical imaging magnification and the lens-aperture of the recording system. A theoretical lower-limit of sampling resolution is obtained using the Shannon sampling theorem, in which both the speckle intensity is sufficiently registered and the maximum measurement capability of the equipment is employed. Capabilities of measurement such as accuracy of displacement components, upper limits of measurable area and measurable deformation are discussed. A method of changing the measurement capabilities is also

given by adjustment of the recording optical arrangement. Experimental data from practical speckle patterns using the normalized product (*Nprod*) correlation metric is given to verify the results. Experiments show that the displacement-only correlation using laser speckles is a reliable measurement as long as the optical imaging system is properly arranged and the sampling resolution requirement is satisfied. Computing time is significantly reduced since a two-parameter searching process is employed instead of a six-parameter searching. Though not directly obtained from the correlation process, the necessary strain and rotation components can be obtained by numerical derivation of the resulted displacement components.

2 Displacement-only digital speckle correlation

The system used in the data acquisition and image processing is shown in Fig.1, which consists of a TV camera, an A/D converter, a camera controller, and a micro-computer. The object surface is illuminated by a collimated laser beam. The image of the object is captured by a video camera with flexible frame resolutions of 256×256 , 512×512 or 1024×1024 pixels. The light intensity of the speckle pattern is digitized by the analog-to-digital converter (ADC) and then transmitted to the VAX-11/730 computer through an 8-bit I/O port with discrete intensity of 256 levels. The computer controls the data acquisition system, stores the digitized image data, performs the correlation calculations, and interfaces with the graphic peripherals for displaying the results.

In the speckle correlation process, two images of the interested object are registered by the data acquisition system, one before and another after the specimen deformation. The digital correlation is then performed by relating the deformed image with the undeformed one. Consider a small subset of the digital image centered at $P_o(x_o, y_o)$ on the undeformed image, as shown in Fig.2. During the deformation, the subset center moves to a new point

$$P_o^*(x_o^*, y_o^*) = P_o(x_o + u, y_o + v), \quad (1)$$

and the subset deforms. If the subset is small, one may assume that straight lines on the undeformed subset remain straight after deformation. A linear mapping technique can be applied to the relationship between the two subsets. Thus, an arbitrary point $P(x, y)$ on the undeformed subimage moves to a new point

$$P^*(x^*, y^*) = P \left(x + u + \frac{\partial u}{\partial x} \delta x + \frac{\partial u}{\partial y} \delta y, y + v + \frac{\partial v}{\partial x} \delta x + \frac{\partial v}{\partial y} \delta y \right), \quad (2)$$

where (x, y) are coordinates on the undeformed image and (x^*, y^*) are coordinates on the deformed image, (u, v) are displacement components of the subimage center and

$$\delta x = x - x_o \quad (3)$$

$$\delta y = y - y_o,$$

are distant components from the subimage center $P_o(x_o, y_o)$ to the arbitrary point $P(x, y)$. Iterative searching of all the displacement components (u, v) and the displacement gradients $\left(\frac{\partial u}{\partial x}, \frac{\partial u}{\partial y}, \frac{\partial v}{\partial x}, \frac{\partial v}{\partial y} \right)$ yields the six-parameter correlation process.⁵⁻⁸

If both the displacement gradients and the subset size are small, one may further assume that all points on the subimage undergo the same displacement during the

deformation. Thus, an arbitrary point $P(x, y)$ on the undeformed subimage moves to a new location of

$$P^*(x^*, y^*) = P(x + u, y + v). \quad (4)$$

Searching of the two displacement components (u, v) yields the process of the two-parameter correlation, or displacement-only correlation.

The correlation algorithm used is the normalized product (N_{prod}) metric which is more noise resistant than many other existing approaches.^{13,14} This metric is defined as

$$N_{prod}(u, v) = \frac{\sum_{\Delta} g_1(x, y)g_2(x^*, y^*)}{[\sum_{\Delta} g_1^2(x, y)]^{1/2} [\sum_{\Delta} g_2^2(x^*, y^*)]^{1/2}}, \quad (5)$$

where

$$\begin{aligned} x^* &= x + u \\ y^* &= y + v, \end{aligned} \quad (6)$$

$g_1(x, y)$ denotes the undeformed subimage, $g_2(x^*, y^*)$ denotes the deformed subimage and Δ denotes the subimage region. By varying the values of u and v in (5), those values that maximize the N_{prod} are found to be the local displacement components of the subimage.

3 Statistical analysis of subjective laser speckle

3.1 Image formation

Fig.3 shows the optical arrangement of the subjective laser speckle recording system.

A collimated laser beam illuminates the object plane (α, β) . The speckle field of the object is recorded by the recording lens at a distance $p (= \overline{OO'})$ from the object surface and registered onto the camera sensor in the observation plane (x, y) at a distance $q (= \overline{O'O''})$ behind the lens. Plane (x, y) is the conjugate of plane (α, β) in the imaging system. The image at the observation plane is captured twice by the CCD camera with the specimen deformed between exposures.

Denoting the random light intensity distribution on the surface of the specimen (i.e., the objective speckle pattern) by $I(\alpha, \beta)$, we can express the complex amplitude of the objective speckle field as

$$f(\alpha, \beta) = \sqrt{I(\alpha, \beta)} \exp[j\phi(\alpha, \beta)]. \quad (7)$$

The point-spread function of the optical system, which is the response of a point source at $O(0, 0)$ on the object plane, is

$$h(x, y) = \iint_{-\infty}^{+\infty} P(\xi, \eta) \exp \left\{ -j \frac{2\pi}{\lambda} \left[\left(\frac{\alpha}{p} + \frac{x}{q} \right) \xi + \left(\frac{\beta}{p} + \frac{y}{q} \right) \eta \right] \right\} d\xi d\eta, \quad (8)$$

where $P(\xi, \eta)$ is a volume-normalized pupil function of the recording lens-aperture defined by¹⁵⁻¹⁷

$$P(\xi, \eta) = \begin{cases} \frac{4}{\pi D^2}, & \text{when } \sqrt{\xi^2 + \eta^2} \leq D/2 \\ 0, & \text{elsewhere,} \end{cases} \quad (9)$$

with D the diameter of the lens-aperture. Let

$$\begin{aligned}\frac{1}{\lambda} \left(\frac{a}{p} + \frac{z}{q} \right) &= r \cos \theta \\ \frac{1}{\lambda} \left(\frac{b}{p} + \frac{y}{q} \right) &= r \sin \theta,\end{aligned}\tag{10}$$

and

$$\begin{aligned}\xi &= \rho \cos \phi \\ \eta &= \rho \sin \phi,\end{aligned}\tag{11}$$

we achieve the integral in eq (8) as

$$\begin{aligned}h(x, y) &= \frac{4}{\pi D^2} \int_0^{\frac{D}{2}} \int_0^{2\pi} \exp[-j2\pi\rho r \cos(\phi - \theta)] \rho d\phi d\rho \\ &= 2\pi \int_0^{\frac{D}{2}} \rho J_0(2\pi\rho r) d\rho \\ &= \frac{2J_1(\pi D r)}{\pi D r}.\end{aligned}\tag{12}$$

Thereafter the point-spread function of the imaging system becomes (Fig.4)

$$h(x, y) = \frac{2J_1 \left[\frac{\pi D}{\lambda q} \sqrt{x^2 + y^2} \right]}{\frac{\pi D}{\lambda q} \sqrt{x^2 + y^2}},\tag{13}$$

where $J_1()$ is the first order Bessel function of the first kind.

Theoretically, the aberrationless image of the objective speckle field, which would be recorded using an infinite lens-aperture, is in a form of $f_i(x, y) = \frac{1}{M^2} f\left(-\frac{x}{M}, -\frac{y}{M}\right)$, where $M (= q/p)$ is the magnification factor of the recording system. The complex amplitude of the subjective speckle pattern, which is recorded by the practical finite lens-aperture, can be obtained by convolving the point-spread function with the theoretical aberrationless image, i.e.,

$$\begin{aligned}
f_i(x, y) &= \frac{1}{M^2} \iint_{-\infty}^{+\infty} f\left(-\frac{\alpha_1}{M}, -\frac{\beta_1}{M}\right) \frac{2J_1\left[\frac{\pi D}{\lambda_f} \sqrt{(x-\alpha_1)^2 + (y-\beta_1)^2}\right]}{\frac{\pi D}{\lambda_f} \sqrt{(x-\alpha_1)^2 + (y-\beta_1)^2}} d\alpha_1 d\beta_1 \\
&= \iint_{-\infty}^{+\infty} f(\alpha, \beta) \frac{2J_1\left[\frac{\pi D}{\lambda_f} \sqrt{(x+M\alpha)^2 + (y+M\beta)^2}\right]}{\frac{\pi D}{\lambda_f} \sqrt{(x+M\alpha)^2 + (y+M\beta)^2}} d\alpha d\beta.
\end{aligned} \tag{14}$$

Thereby the light intensity of the subjective laser speckle pattern is

$$\begin{aligned}
g(x, y) &= C |f_i(x, y)|^2 \\
&= C \left| \iint_{-\infty}^{+\infty} f(\alpha, \beta) \frac{2J_1\left[\frac{\pi D}{\lambda_f} \sqrt{(x+M\alpha)^2 + (y+M\beta)^2}\right]}{\frac{\pi D}{\lambda_f} \sqrt{(x+M\alpha)^2 + (y+M\beta)^2}} d\alpha d\beta \right|^2,
\end{aligned} \tag{15}$$

where C is a constant.

3.2 Spectral analysis

Neglecting the constant C , eq (15) is rewritten as

$$g(x, y) = f_i(x, y) f_i^*(x, y), \tag{16}$$

where $f_i^*(x, y)$ is the conjugate of the complex amplitude field $f_i(x, y)$. Let $F_i(\omega_x, \omega_y)$ be the spectrum of the complex amplitude of the objective speckle field. From the Fourier transform property^{16,17}, we obtain the spectrum of the subjective speckle pattern

$$G(\omega_x, \omega_y) = F_i(\omega_x, \omega_y) \otimes F_i^*(-\omega_x, -\omega_y), \tag{17}$$

where $F_i(\omega_x, \omega_y)$ is the Fourier transform of $f_i(x, y)$, $*$ denotes complex conjugate and \otimes denotes convolution. In regard of the fact that eq (14) is equivalent to

$$f_i(x, y) = \frac{1}{M^2} f\left(-\frac{x}{M}, -\frac{y}{M}\right) \otimes h(x, y), \tag{18}$$

we obtain the spectrum of the complex amplitude field as

$$F_s(\omega_x, \omega_y) = F(-M\omega_x, -M\omega_y) \cdot H(\omega_x, \omega_y), \quad (19)$$

where $H(\omega_x, \omega_y)$ denotes the spectrum of the point-spread function which is geometrically similar to the lens-aperture of the recording system, i.e.,

$$H(\omega_x, \omega_y) = \begin{cases} 1, & \text{when } \sqrt{\omega_x^2 + \omega_y^2} \leq D/(2\lambda q) \\ 0, & \text{elsewhere.} \end{cases} \quad (20)$$

Substituting eq (19) into eq (17), the spectrum of the subjective speckle pattern becomes

$$G(\omega_x, \omega_y) = [F(-M\omega_x, -M\omega_y) \cdot H(\omega_x, \omega_y)] \otimes [F^*(M\omega_x, M\omega_y) \cdot H^*(-\omega_x, -\omega_y)]. \quad (21)$$

In practice, the cutoff frequency of a typical optically rough surface is much larger than $D/(2\lambda q)$.¹⁵ A good simplification can be made by taking the amplitude profile of $F(M\omega_x, M\omega_y)$ as a constant (=unity) for

$$\sqrt{\omega_x^2 + \omega_y^2} \leq D/(2\lambda q). \quad (22)$$

Therefore, the amplitude profile of the spectrum of the subjective speckle field is

$$G(\omega_x, \omega_y) = H(\omega_x, \omega_y) \otimes H^*(-\omega_x, -\omega_y), \quad (23)$$

or

$$G(\omega_x, \omega_y) = H(\omega_x, \omega_y) \otimes H(\omega_x, \omega_y), \quad (24)$$

since $H(\omega_x, \omega_y)$ is a real and symmetric function. Substituting eq (20) into eq (24), one obtains the amplitude profile of the spectrum of the subjective speckle pattern

as

$$G(\omega_x, \omega_y) = \begin{cases} \frac{D^2}{2\lambda^2 q^2} \left[\cos^{-1} \left(\frac{\lambda q \sqrt{\omega_x^2 + \omega_y^2}}{D} \right) - \left(\frac{\lambda q \sqrt{\omega_x^2 + \omega_y^2}}{D} \right) \sqrt{1 - \left(\frac{\lambda q \sqrt{\omega_x^2 + \omega_y^2}}{D} \right)^2} \right], & \text{when } \sqrt{\omega_x^2 + \omega_y^2} \leq \Omega \\ 0, & \text{elsewhere,} \end{cases} \quad (25)$$

where $\Omega = D/(\lambda q)$ is the maximum radius of the spectrum as shown in Fig.5. It is seen that the spectrum of the subjective image vanishes outside a circular region of

$$\sqrt{\omega_x^2 + \omega_y^2} = \Omega. \quad (26)$$

From eq (23), the intensity of a typical subjective speckle is

$$g_0(x, y) = |h(x, y)|^2, \quad (27)$$

where $h(x, y)$ is the point-spread function of the optical system. Substituting eq (13) into eq (27), one obtains the intensity distribution of a typical subjective speckle (Fig.6)

$$g_0(x, y) = \left\{ \frac{2J_1 \left[\frac{\pi D}{\lambda q} \sqrt{x^2 + y^2} \right]}{\frac{\pi D}{\lambda q} \sqrt{x^2 + y^2}} \right\}^2. \quad (28)$$

From Fig.6 it is seen that the radius of the first zero crossing, which is usually referred to as the typical speckle size,¹ is

$$S = 1.22 \frac{\lambda q}{D}. \quad (29)$$

4 Optimal sampling resolution of laser speckle

4.1 Theoretical analysis

Recall the fact that the spectrum of the subjective laser speckle vanishes over all but a bounded region of

$$\sqrt{\omega_x^2 + \omega_y^2} \leq \Omega. \quad (30)$$

From the Shannon Sampling theorem,¹² the original unsampled subjective speckle pattern can be exactly reconstructed from its sampled values taken over a lattice of points (mT, nT) by the following interpolation

$$g_b(x, y) = \sum_{m=-\infty}^{+\infty} \sum_{n=-\infty}^{+\infty} g_s(mT, nT) \frac{2J_1\left[\frac{\pi}{T}\sqrt{(x-mT)^2 + (y-nT)^2}\right]}{\frac{\pi}{T}\sqrt{(x-mT)^2 + (y-nT)^2}}, \quad (31)$$

provided that the sampling interval is sufficiently small. Wherein $g_s(mT, nT)$ is the sampled speckle pattern and

$$b(x, y) = \frac{2J_1\left(\frac{\pi}{T}\sqrt{x^2 + y^2}\right)}{\left(\frac{\pi}{T}\sqrt{x^2 + y^2}\right)} \quad (32)$$

is the theoretical reconstruction function (Fig.7).

The upper-limit of the sampling interval T , which is referred to as Nyquist sampling interval, is determined by the spectral distribution of the speckle pattern. From the Fourier transform property, the spectrum of the sampled image is a periodic replication of the spectrum of the original image, i.e.,

$$G_s(\omega_x, \omega_y) = \sum_{m=-\infty}^{+\infty} \sum_{n=-\infty}^{+\infty} G(\omega_x - \frac{m}{T}, \omega_y - \frac{n}{T}), \quad (33)$$

where $G_s(\omega_x, \omega_y)$ is the spectrum of the sampled speckle pattern and $G(\omega_x, \omega_y)$ is the spectrum of the analog (unsampled) speckle pattern (see Fig.8). It is easy to prove that the spectrum of the reconstruction function $b(x, y)$ is a circular pupil function given by (Fig.9)

$$B(\omega_x, \omega_y) = \begin{cases} 4T^2/\pi, & \text{when } \sqrt{\omega_x^2 + \omega_y^2} \leq 1/(2T) \\ 0, & \text{elsewhere.} \end{cases} \quad (34)$$

Fig.8 shows the amplitude profile of the spectrum of a sampled speckle pattern and Fig.9 shows the spectrum of the theoretical reconstruction function.

From eq (31) it is seen that the reconstructed image is a convolution of the reconstruction function with the sampled image $g_s(mT, nT)$. Therefore, the spectrum of the reconstructed image is equivalent to the product of the spectrum of the reconstruction function with the spectrum of the sampled image, i.e.,

$$G_b(\omega_x, \omega_y) = B(\omega_x, \omega_y) \sum_{m=-\infty}^{+\infty} \sum_{n=-\infty}^{+\infty} G(\omega_x - \frac{m}{T}, \omega_y - \frac{n}{T}), \quad (35)$$

where $G_b(\omega_x, \omega_y)$ is the spectrum of the reconstructed image and $B(\omega_x, \omega_y)$ is the spectrum of the reconstruction function. From Fig.8 and Fig.9, it is evident that if

$$T^{-1} > 2\Omega, \quad (36)$$

then aliasing problem will not occur in the spectral domain of the reconstructed image. Hence the original speckle pattern can be faithfully reconstructed. Recalling $\Omega = D/(\lambda q)$, we obtain the theoretical lower-limit sampling resolution of the subjective

speckle pattern

$$\begin{aligned} T &= \lambda q / (2D) \\ &= 0.41S, \end{aligned} \tag{37}$$

where S is the conventional speckle size defined in eq (29).

4.2 Experimental verification

In the experimental verification of the optimal sampling resolution, a rigid translation of $200 \mu m$ of a aluminum plate was tested. An area of $8.5 \times 8.5 mm^2$ was recorded by a micro-lens of $f = 90 mm$ at magnification of $M = 1.18$ and aperture number of $F_1 = 22$. The typical speckle size was $S = 37.0 \mu m$. The subjective speckle patterns were digitized by frame resolutions of 1024×1024 , 512×512 , 256×256 and 128×128 pixels with real pixel intervals of $T = 0.264S$, $T = 0.528S$, $T = 1.06S$ and $T = 2.11S$, respectively. The dimension of the selected subimage from the undeformed speckle pattern was 15×15 elements and the dimension of the selected searching region on the deformed speckle pattern was 65×65 elements. Fig.10 shows the sectional views of a speckle pattern sampled by the four different resolutions. It is seen that upon the diminishing of the sampling resolution, high signal frequency components are lost gradually and less intensity contrasts are preserved. But as long as $T = 0.528S$, most of the speckle frequency components and the speckle intensity variance are registered.

A mean-elimination process was applied on each of the undeformed subimages and deformed searching regions before the correlation evaluation. The elimination is

expressed as

$$g_{s1}(mT, nT) = g_s(mT, nT) - \frac{1}{N_o^2} \sum_{i=1}^{N_o} \sum_{j=1}^{N_o} g(iT, jT), \quad (38)$$

where $N_o = 15$ for each subimage and $N_o = 65$ for each searching region. The effect of the mean-elimination on the correlation evaluation is shown in Fig.11. It shows that the correlation between the mean-eliminated images preserves much more contrast between the matching position and all the nonmatching positions because a zero-mean background is achieved in computing the N_{prod} .

The effect of different sampling resolutions on the correlation evaluation is shown in Fig.12a-c. Correlations between the deformed and reference speckle patterns using different sampling intervals of $T = 0.264S$, $T = 0.528S$, and $T = 1.06S$ were performed. It is seen that the decrease in sampling resolution results in two effects in the N_{prod} correlation. One is the flattening of the matching hill, which reduces the sensitivity or accuracy of the measurement. The other is the increase of the background noise, which reduces the reliability of measurement. However, at $T = 0.528S$, the maximum value of the background of N_{prod} is less than 0.5, which is much smaller than the matching coefficient 0.95. Therefore, a reliable determination of displacement components is obtained at $T = 0.528S$.

5 Capability of measurement

5.1 Upper-limit of measurable area

Recall that the sampling requirement is

$$T = \lambda q / (2D). \quad (39)$$

Since

$$q = (1 + M)DF_1, \quad (40)$$

where M is the imaging magnification of the optical system, and F_1 is the aperture number, eq (39) becomes

$$T = \lambda F_1 (1 + M) / 2. \quad (41)$$

This equation shows that in regard of the sampling resolution requirement the optical imaging magnification and the aperture number must be selected consistently.

In practice, one may increase the measurable area by choosing the largest aperture number. The minimum applicable imaging magnification is then determined by

$$M = \frac{2T}{\lambda F_1} - 1. \quad (42)$$

Equation (42) implies the dependence of the applicable imaging magnification on the selected recording aperture number (F_1) for a fixed sensor pixel interval T (see

Fig.13). Thus, the maximum measurable area on the specimen is

$$\begin{aligned} A_o &= A_s/M^2 \\ &= A_s / \left(\frac{2T}{\lambda F_1} - 1 \right)^2, \end{aligned} \tag{43}$$

where A_s is the real area of the sensor and A_o is the measured area on the specimen. Some applicable values of the recording magnification and the measurable area in our facility are shown in Table I, with different resolution modes and aperture numbers.

5.2 Upper-limit of accuracy

The accuracy of the displacement components obtained from the direct correlation of the two sampled images is limited by the pixel interval of the sensor array. For accurate measurement, a subpixel detection of the displacement components is necessary. In the six-parameter white-light speckle correlation approach, direct interpolation of original sampled speckle patterns has been used before the correlation process.⁶⁻⁸ In this paper, an alternate interpolation approach is employed. The direct correlation between the two subimages is performed on the original sampled lattices only. A coarse estimation of the displacement components is achieved by detecting the maximum coefficient on the correlation lattice. A fine interpolation is then performed in the local region of the maximum coefficient on the correlation lattice and more accurate determination of displacement components is obtained by further searching of maximum coefficient in the local interpolated region.

The correlation coefficient at any subpixel location (u,v) in the correlation domain

is obtained by a two-dimensional classic cardinal reconstruction from the correlation lattice as ¹⁸

$$N_{prod}(u, v) = \sum_{m=-K}^K \sum_{n=-K}^K N_{prod}(mT, nT) \frac{\sin \left[\frac{\pi}{T}(u - mT) \right]}{\frac{\pi}{T}(u - mT)} \cdot \frac{\sin \left[\frac{\pi}{T}(v - nT) \right]}{\frac{\pi}{T}(v - nT)}. \quad (44)$$

Since both the undeformed subimage and the deformed searching region are sufficiently sampled, aliasing does not occur in the spectrum of the correlation lattice. Therefore, interpolation of the correlation lattice by the cardinal reconstruction function will result in exactly the same coefficient at any subpixel location of the correlation domain as that obtained from the direct correlation between the two corresponding analog (unsampled) speckle patterns. In practice, a finite reconstruction window containing only the mainlobe and the first sidelobe on each side of the cardinal function is employed. The coefficient at a subpixel location is therefore merely a nonlinear interpolation from the sixteen (4×4) nearest coefficients of the correlated lattice. Such an interpolation and maximum searching process results in a fast correlation speed because direct correlation is done only on the sampled lattices rather than on all the fine interpolated points. Further improvement of speed may be obtained by using the Newton-Raphson climbing algorithm in the interpolation and maximum searching process.

Although an interpolation can be made as fine as required, experiments show that the minimum measurable displacement components can not be less than a certain limit for a given speckle pattern. This is because the hill of the correlation function becomes flat when the interpolation is very fine. To the authors' experience, a reliable

subpixel detection of the displacement components as small as 2% of the sampling interval can be obtained. Thus, the unmeasurable distance on the image is $0.02T$, or $0.0082S$. Using eq (42), the uncertainty of the displacement components on the specimen is

$$\begin{aligned}\delta_o &= 0.02T/M \\ &= 0.02T / \left(\frac{2T}{\lambda F_1} - 1 \right).\end{aligned}\tag{45}$$

Equation (45) shows that the uncertainty of the measured displacement components also depends on the selected lens-aperture number at the optimal sampling resolution.

It is seen from eq (41) and eq (45) that the only way to reduce the uncertainty of the displacement components without disturbing the sampling resolution requirement is to reduce the number of the lens-aperture and increase the imaging magnification of the recording system. Such an arrangement will obviously reduce the measurable area. But it is practical in a small scale deformation measurement to achieve accurate results. Therefore, there is a tradeoff between the measurable area and the accuracy of the measurement. Generally, the aperture number and the imaging magnification must be selected by considering both factors.

In our facility, the sensor size is $L = 10.00 \text{ mm}$. When the frame resolution mode of 1024×1024 is selected, the sensor pixel interval is $T = 9.77 \text{ }\mu\text{m}$. At aperture number of $F_1 = 22$, the minimum applicable magnification is $M = 0.403$; the maximum measurable area is $A_o = 615.9 \text{ mm}^2$; and the uncertainty of the measured displacement components is $\delta_o = 0.485 \text{ }\mu\text{m}$. Some more configurations are listed in Table I at other possible resolution modes and aperture numbers.

5.3 Upper-limit of measurable deformation

Since specimen deformation always results in decorrelation between the two registered speckle patterns, it is an essential task to seek the upper deformation limit measurable by the displacement-only correlation method. Denote the real dimension of the selected subimage by L_s and its discrete array size by N_s . First, we consider the case that tensile component $\frac{\partial u}{\partial x}$ is the only nonzero displacement gradient (Fig.14a). After deformation, the subimage is stretched into a length of $L_s + \delta L_s$, with each end extending $\delta L_s/2$. When the extension of each boundary approaches the speckle size S , the expected correlation value contributed from the boundary elements becomes zero. Therefore, the deformed subimage varies from completely correlated (at center) to completely uncorrelated (at ends) to the undeformed subimage. For a reliable measurement, we select the measurable extension limit on both ends to be half of the speckle size, i.e.,

$$\delta L_s \leq S. \quad (46)$$

In regard of $S = 2.44T$ and $L_s = 15T$, one obtains

$$\begin{aligned} \frac{\partial u}{\partial x} &= \frac{\delta L_s}{L_s} \\ &\leq 0.16. \end{aligned} \quad (47)$$

To include the cases of both extension and compression, eq (46) is rewritten as

$$\left| \frac{\partial u}{\partial x} \right| \leq 0.16. \quad (48)$$

Thus the upper limit of the measurable uniaxial tensile strain is ± 0.16 .

Next, let's consider the case that shear component $\frac{\partial u}{\partial y}$ is the only nonzero displacement gradient (Fig.14b). Similar to the tensile case, we select the shear deformation limit as

$$\delta L_s \leq S. \quad (49)$$

Therefore, the limit of the gradient is

$$\left| \frac{\partial u}{\partial y} \right| \leq 0.16. \quad (50)$$

Hence, the upper-limit of the measurable single shear gradient is also ± 0.16 .

In general, more than one nonzero displacement gradients occur. We select the upper-limit of the maximum displacement gradient in a general deformation to be half of the single gradient cases, i.e.,

$$\text{Max} \left(\left| \frac{\partial u}{\partial x} \right|, \left| \frac{\partial u}{\partial y} \right|, \left| \frac{\partial v}{\partial x} \right|, \left| \frac{\partial v}{\partial y} \right| \right) \leq 0.08. \quad (51)$$

An alternately upper-limit may be obtained in terms of strains and rotation,

$$\text{Max} (|\epsilon_{xx}|, |\epsilon_{yy}|, |\epsilon_{xy}|, |\omega_{xy}|) \leq 0.08, \quad (52)$$

where

$$\begin{aligned} \epsilon_{xx} &= \frac{\partial u}{\partial x} \\ \epsilon_{yy} &= \frac{\partial v}{\partial y} \\ \epsilon_{xy} &= \frac{1}{2} \left(\frac{\partial v}{\partial x} + \frac{\partial u}{\partial y} \right) \\ \omega_{xy} &= \frac{1}{2} \left(\frac{\partial v}{\partial x} - \frac{\partial u}{\partial y} \right). \end{aligned} \quad (53)$$

Therefore, the maximum measurable strain or rotation is ± 0.08 in a general deformation.

Another decorrelation factor may come from a rigid translation of the specimen. To the authors' experience, however, decorrelation from a rigid translation of 5 mm is a small effect in a collimated illumination system. Therefore, in most practical cases we may assume the measurable rigid translation is unlimited and the decorrelation from the rigid translation is negligible.

6 Conclusion and discussion

An optimal sampling resolution of the laser speckle pattern is obtained using the Shannon sampling theorem and verified by experimental results. The normalized product metric, which is a noise resistant approach, is employed in the correlation process. Reliable results of displacement components from displacement-only correlation of laser speckle patterns are obtained. Fast computation speed is achieved because a two-parameter searching process is used and direct correlation is performed only on the integral sites of the sampled lattice. High accuracy of displacement components is achieved by the cardinal interpolation and subpixel maximum searching in the correlation domain. The upper limits of the accuracy and measurable area are analyzed and a method of changing these limits is given by adjusting the imaging configuration. The maximum measurable specimen translation is 5 mm in a collimated illumination system, and the maximum measurable strain and rotation is up to 0.08 in a general deformation. Though not directly obtained from the correlation process, necessary strain and rotation components can be obtained by numerical derivation

of the resulted displacement field. Further improvement in processing speed may be obtained by using the Newton-Raphson climbing algorithm in the coarse-to-fine interpolation and maximum searching process.

7 Acknowledgments

This work was supported by the Army Research Office through Contract DAA20388K0033 (Scientific Program Officer: Dr.G.L. Anderson) and the Office of Naval Research through Contract N0001482K0566 (Scientific Officer: Dr. Y.Rajapakse). These supports are gratefully acknowledged. We would also like to thank Prof. H.S. Don of the Electrical Engineering Department for his fruitful criticism and advice.

8 References

1. F.P.Chiang, "A new family of 2D and 3D experimental stress analysis techniques using laser speckles", SM Archives, Vol.3, No.1, pp.27-58 (1978).
2. Y.Y.Hung, "Shearography: a new optical method for strain measurement and nondestructive testing", Optical Engineering, Vol.21, No.3, pp.391-395 (1982).
3. W.H.Peters and W.F.Ranson, "Digital image techniques in experimental stress analysis", Optical Engineering, Vol. 21, No.3, pp.427-431 (1982).
4. W.H.Peters, W.F.Ranson, M.A.Sutton, T.C.Chu and J.Anderson, "Application of digital correlation methods to rigid body mechanics", Optical Engineering, Vol.22,

No.6, pp.738-742 (1983).

5. M.A.Sutton, W.J.wolters, W.H.Peters, W.F.Ranson and S.R.McNeill, "Determination of displacements using an improved digital correlation method", Image and Vision Computing, Vol.1, No.3, pp.133-139 (1983).

6. T.C.Chu, W.F.Ranson, M.A.Sutton and W.H.Peters, "Application of digital-image-correlation techniques to experimental mechanics", Experimental Mechanics, Vol.25, No.3, pp.232-244 (1985).

7. M.A.Sutton, Mingqi Cheng, W.H.Peters, Y.J.Chao and S.R.McNeill, "Application of an optimized digital correction method to planar deformation analysis", Image and Vision Computing, Vol.4, No.3, pp.143-150 (1986).

8. H.A. Bruck, S.R. McNeill, M.A. Sutton and W.H. Peters, " Digital image correlation using Newton-Raphson method of partial differential correlation ", Experimental Mechanics, Vol. 29, No. 3, pp. 261-267 (1989).

9. M.A. Hamed, " Object-motion measurements using pulse-echo acoustical speckle and two-dimensional correlation ", Experimental Mechanics, Vol.27, No. 3, pp. 250-254 (1987).

10. Z.H.He, M.A.Sutton, W.F.Ranson and W.H.Peters, "Two-dimensional fluid-velocity measurements by use of digital-speckle correction techniques", Experimental Mechanics, Vol.24, No.2, pp.117-121 (1984).

11. M.A. Sutton, S.R. McNeill, J. Jang and M. Babai, " Effects of subpixel image restoration on digital correlation error estimates", Optical Engineering, Vol. 27, No.

10, pp. 870-877 (1988).

12. A.Rosenfeld and A.C.Kak, Digital Picture processing, Vol.1, Academic-Press, New York (1982).

13. H.H.Bailey, F.W.Blackwell, C.L.Lowery and J.A.Ratkovic, "Image correlation: Part I. Simulation and analysis", A Report Prepared for United States Air Force Project Rand, R-2057/1-PR, Nov (1976).

14. H.W.Wessely, "Image correlation: Part II. Theoretical basis", A Report Prepared for United States Air Force Project Rand, R-2057/2-PR, Nov (1976).

15. R.P.Khetan and F.P.Chiang, " Strain analysis by one-beam laser speckle interferometry. 1: Single aperture method", Applied Optics, Vol.15, No.9, pp. 2205-2215 (1976).

16. J.W. Goodman, Introduction to Fourier Optics, McGraw-Hill, San Francisco (1968).

17. Francis T.S.Yu, Optical Information Processing, John Wiley and Sons, New York, 1983.

18. F.J.Harris, Multirate FIR Filters for Interpolating and Desampling, in Handbook of Digital Signal Processing, Engineering Applications, D.F. Elliott, ed. pp.173-287, Academic Press, New York (1987).

Captions of Illustrations:

Table 1-Measurable area and uncertainty at different system arrangements

Fig. 1- Schematic of data acquisition and image processing system

Fig. 2- Local deformation of a subimage

Fig. 3- Optical image formation of laser speckle pattern

Fig. 4- Point-spread function of optical system (cross section)

Fig. 5- Amplitude profile (normalized) of the spectrum of a subjective
laser speckle pattern

Fig. 6- Intensity distribution of a typical subjective laser speckle

Fig. 7- Theoretical reconstruction function

Fig. 8- Amplitude profile of the spectrum of a sampled subjective laser
speckle pattern

Fig. 9- Spectrum of theoretical reconstruction function

Fig.10- A speckle pattern sampled by different resolutions (sectional view).

(a) Frame resolution: 1024×1024 ; sampling interval: $T = 0.264S$.

(b) Frame resolution: 512×512 ; sampling interval: $T = 0.528S$.

(c) Frame resolution: 256×256 ; sampling interval: $T = 1.06S$.

(d) Frame resolution: 128×128 ; sampling interval: $T = 2.11S$.

Fig.11- Effect of mean-elimination on correlation evaluation.

(a) Without mean-elimination. (b) After mean-elimination

Fig.12- Correlations between deformed and undeformed speckle patterns sampled

by different resolutions.

(a) Frame resolution: 1024×1024 ; sampling interval: $T = 0.264S$

(b) Frame resolution: 512×512 ; sampling interval: $T = 0.528S$

(c) Frame resolution: 256×256 ; sampling interval: $T = 1.06S$

Fig.13- Magnification and uncertainty as a function of recording aperture

Fig.14- Decorrelation effect of specimen deformation

(a) Uniaxial tension. (b) Pure shear.

Table 1 Measurable area and uncertainty at different system arrangements

$A_s(\text{mm}^2)$	Resolution ($N \times N$)	F_s	$T(\mu\text{m})$	M	$A_o(\text{mm}^2)$	$\delta_o(\mu\text{m})$
100.0 (10.00 × 10.00)	512 × 512	11	19.53	4.61	4.71(2.17 × 2.17)	0.085
		22	19.53	1.81	30.52(5.52 × 5.52)	0.216
	1024 × 1024	11	9.77	1.81	30.52(5.52 × 5.52)	0.108
		22	9.77	0.403	615.9(24.8 × 24.8)	0.485

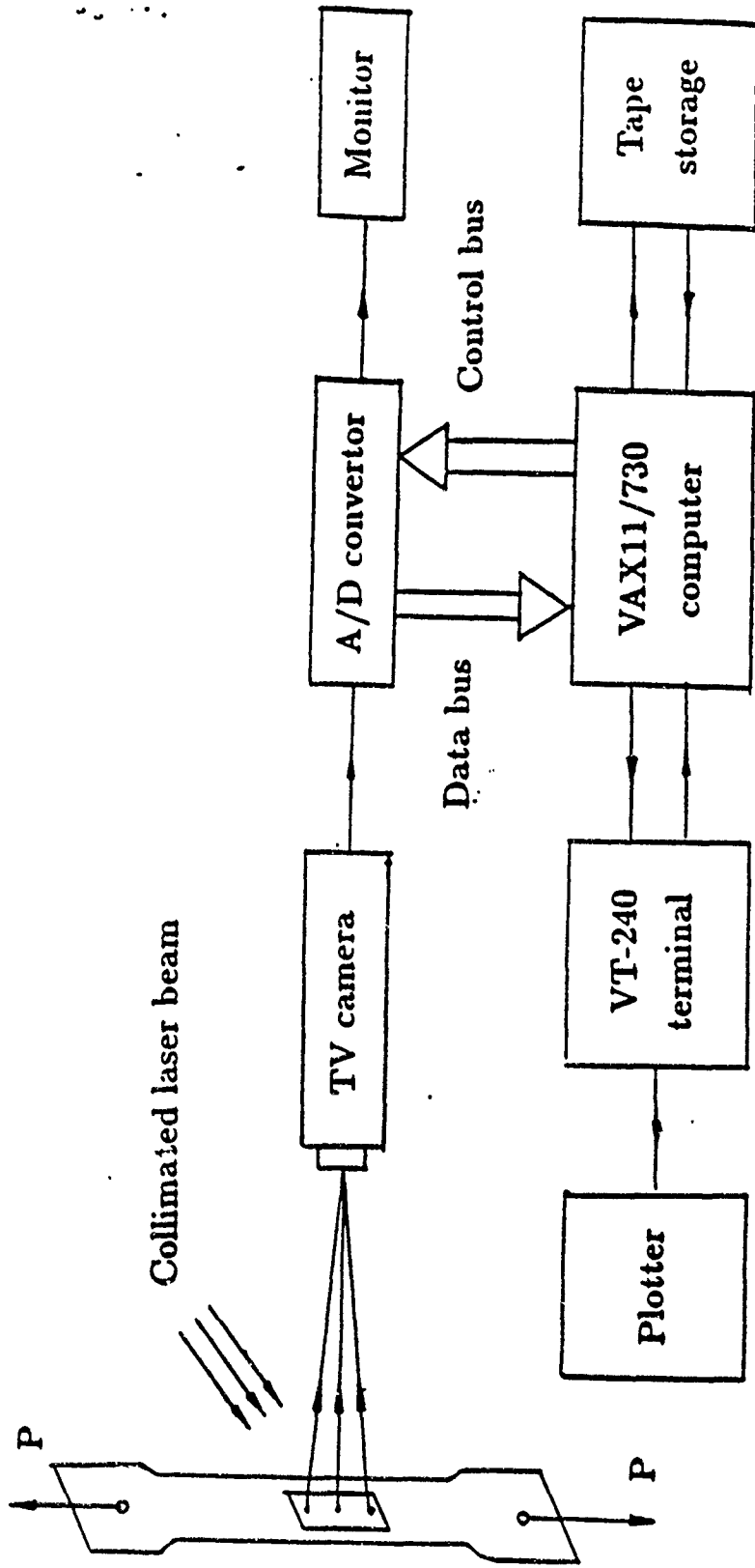


Fig.1 Schematic of data acquisition and image processing system

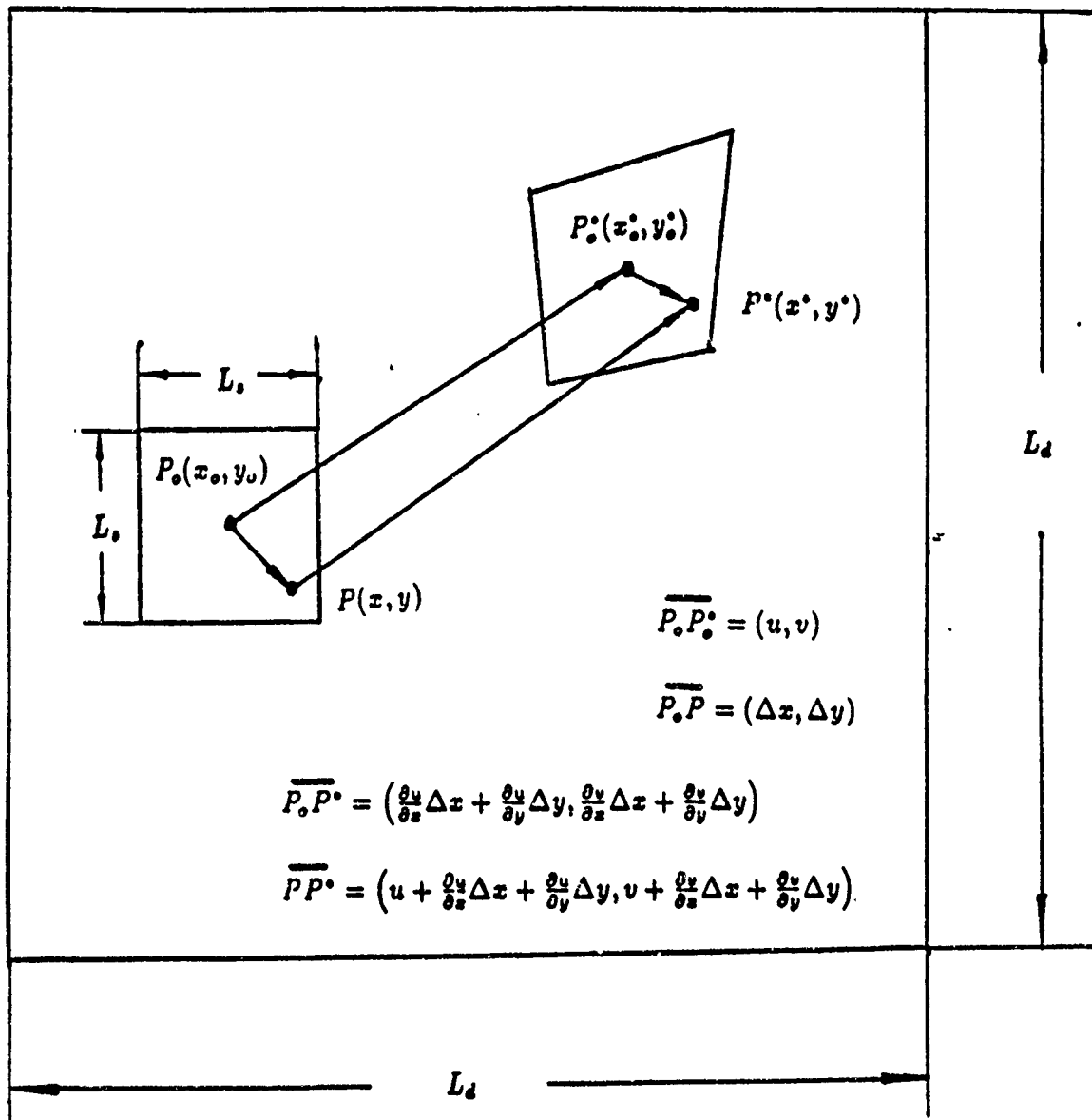


Fig.2 Local deformation of a subimage

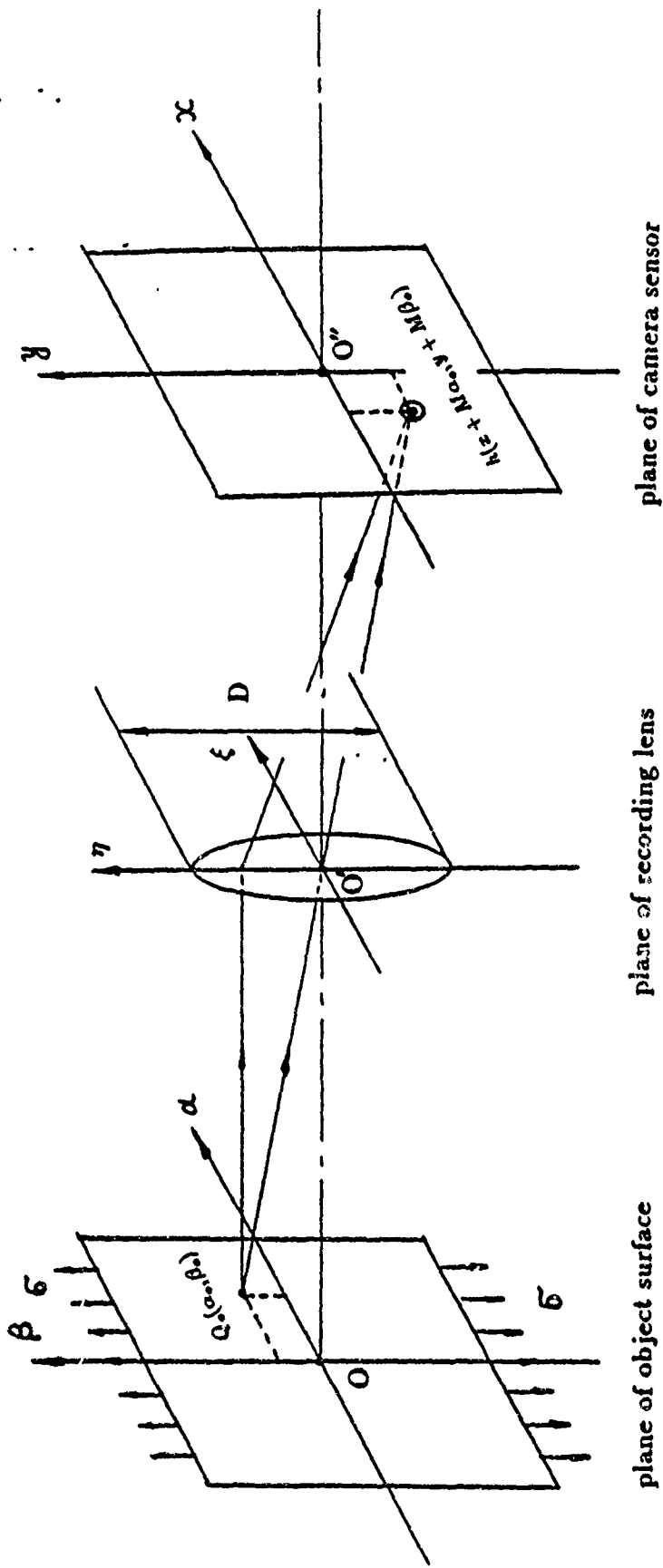


Fig.3 Optical image formation of laser speckle pattern

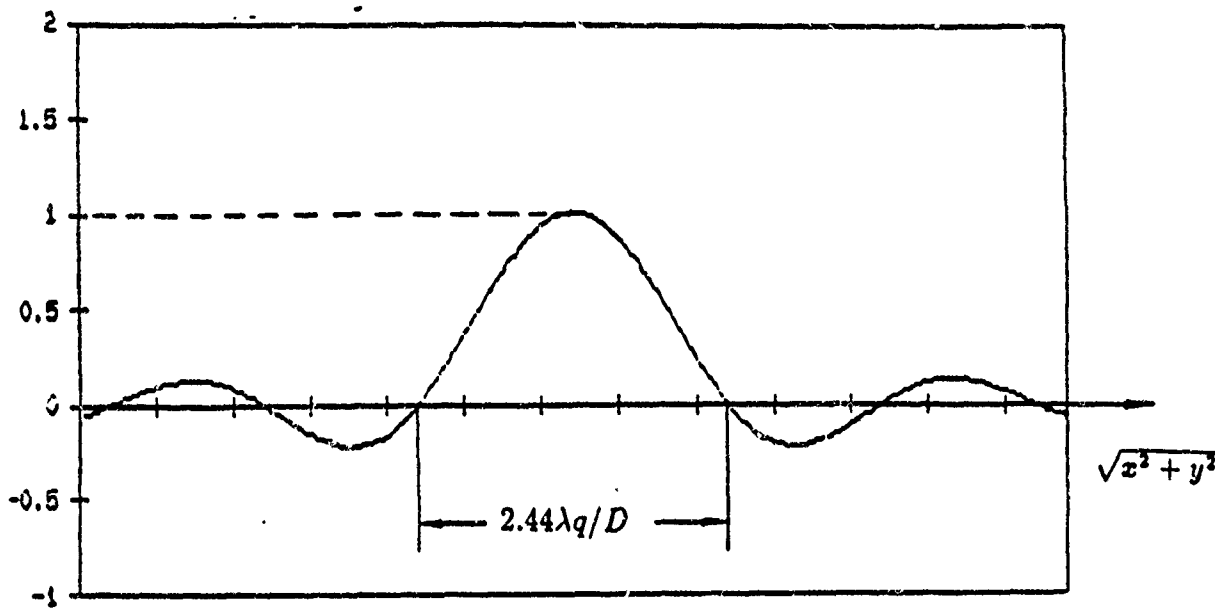


Fig.4 Point-spread function of optical system (cross section)

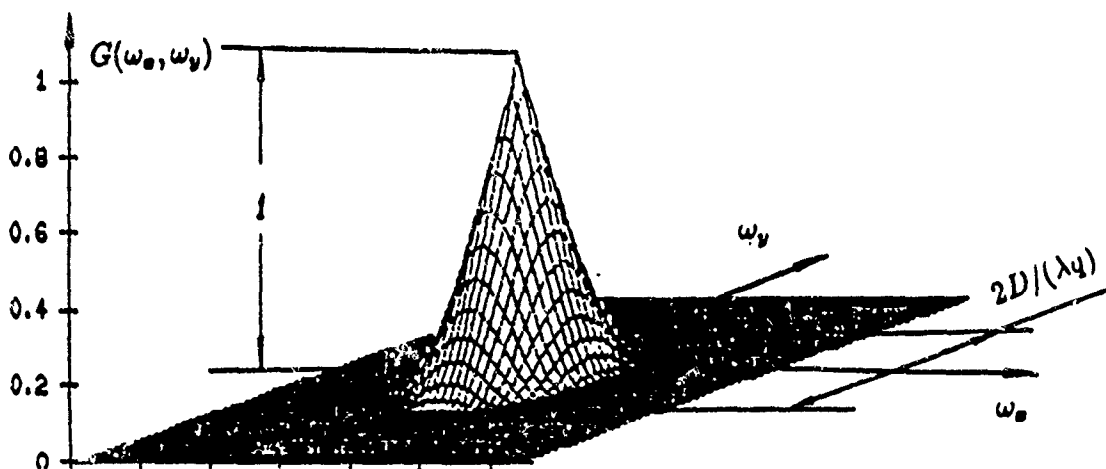


Fig. 5-Amplitude profile (normalized) of the spectrum of a subjective laser speckle pattern

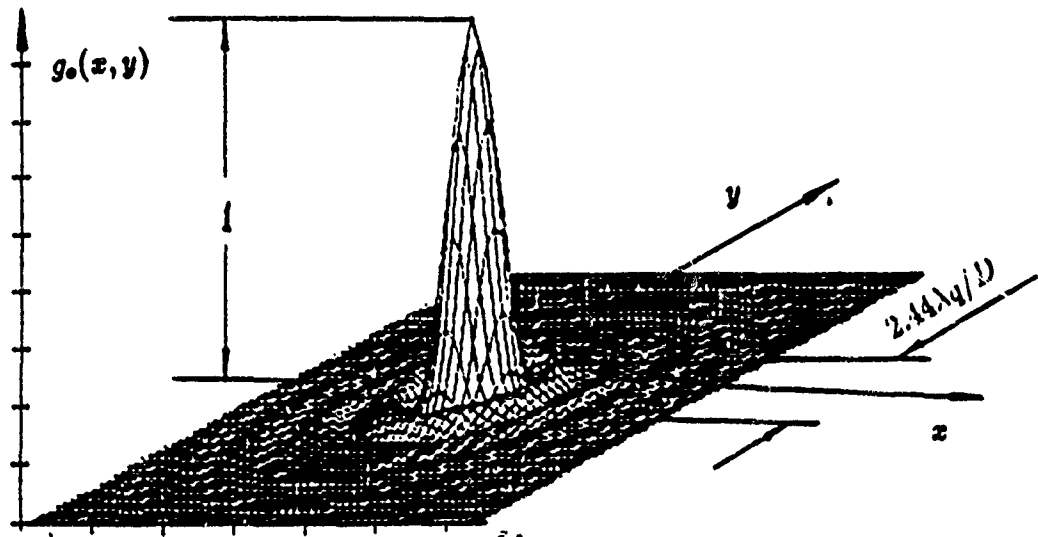


Fig.6 Intensity distribution of a typical subjective laser speckle

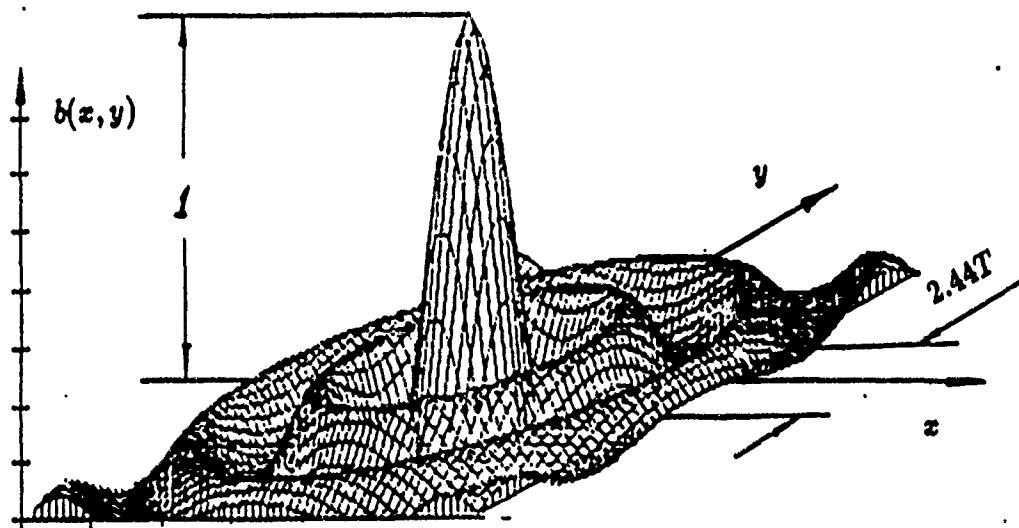


Fig.7 Theoretical reconstruction function

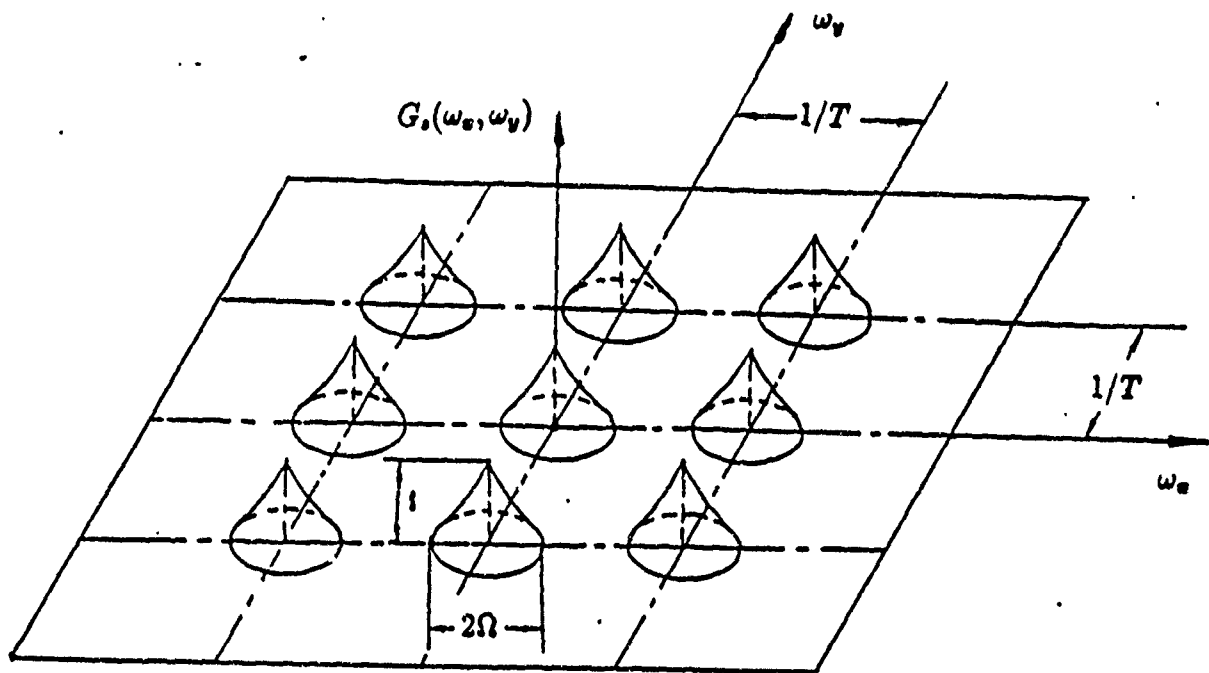


Fig. 8—Amplitude profile of the spectrum of a sampled subjective laser speckle pattern

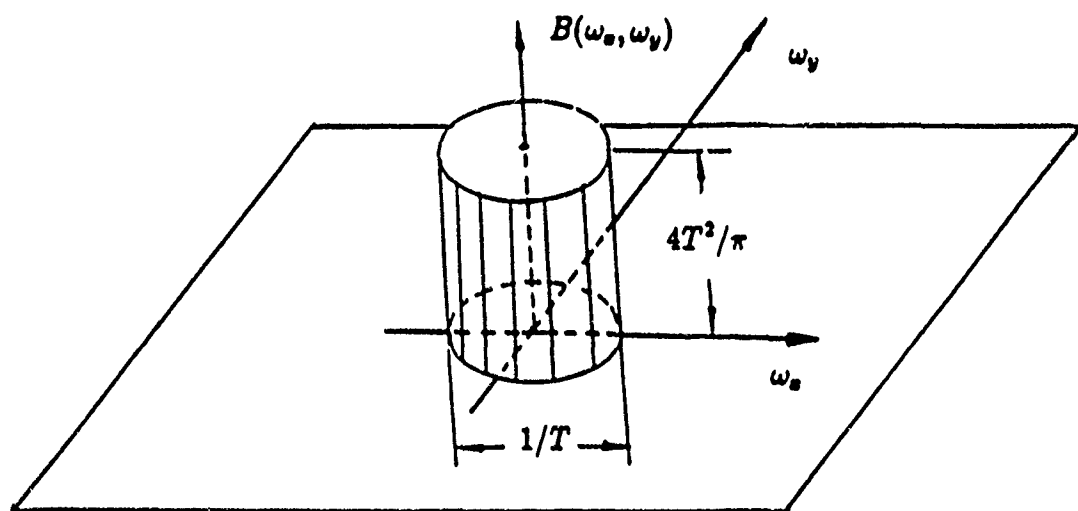


Fig.9 Spectrum of theoretical reconstruction function

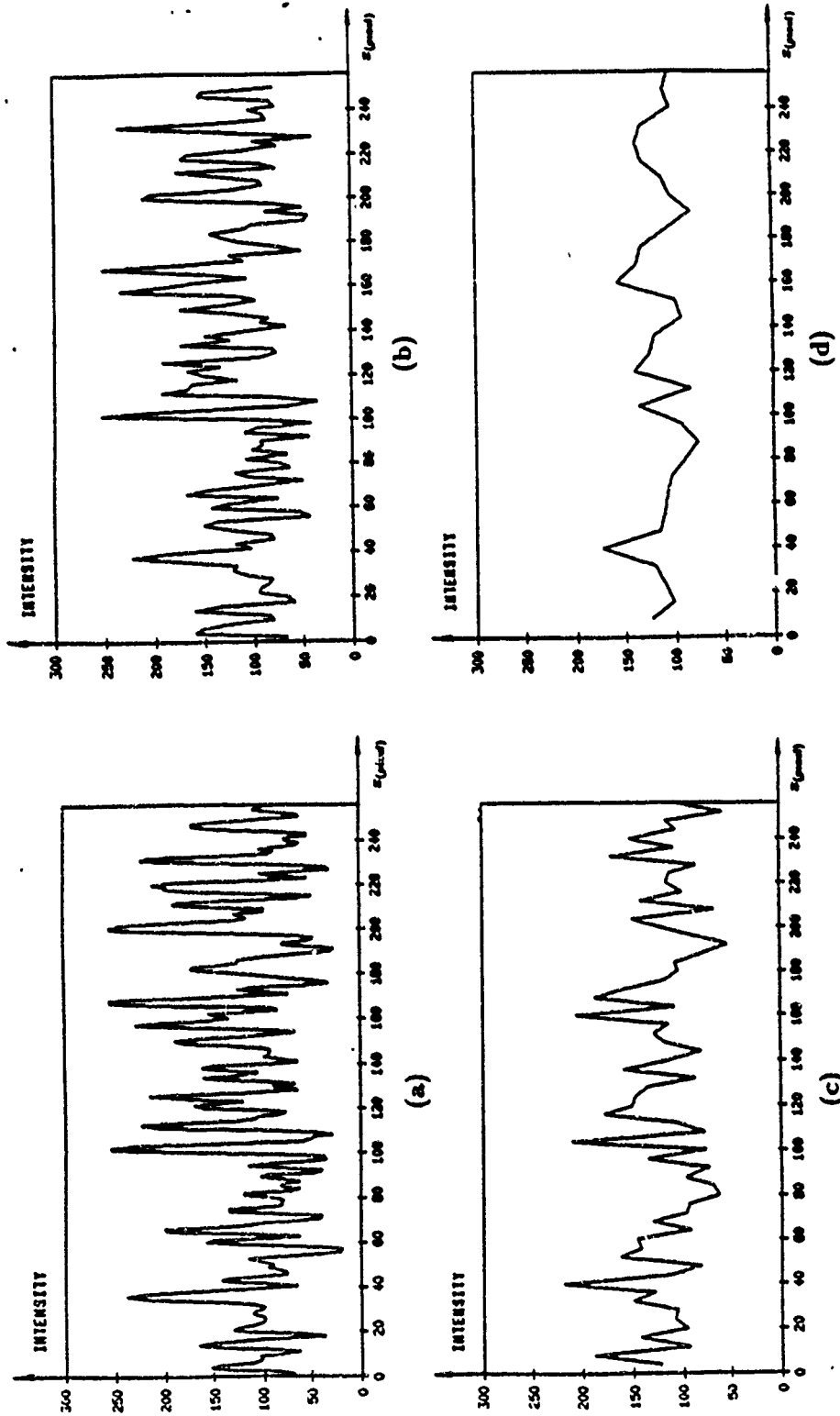


Fig.10- A speckle pattern sampled by different resolutions (sectional view).

- (a) Frame resolution: 1024×1024 ; sampling interval: $T = 0.264S$.
- (b) Frame resolution: 512×512 ; sampling interval: $T = 0.528S$.
- (c) Frame resolution: 256×256 ; sampling interval: $T = 1.06S$.
- (d) Frame resolution: 128×128 ; sampling interval: $T = 2.11S$.

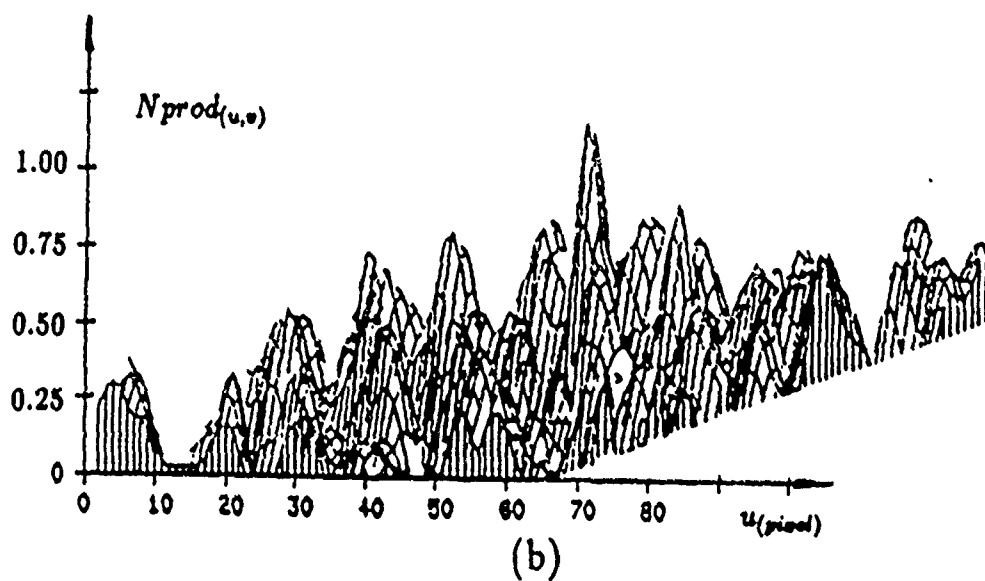
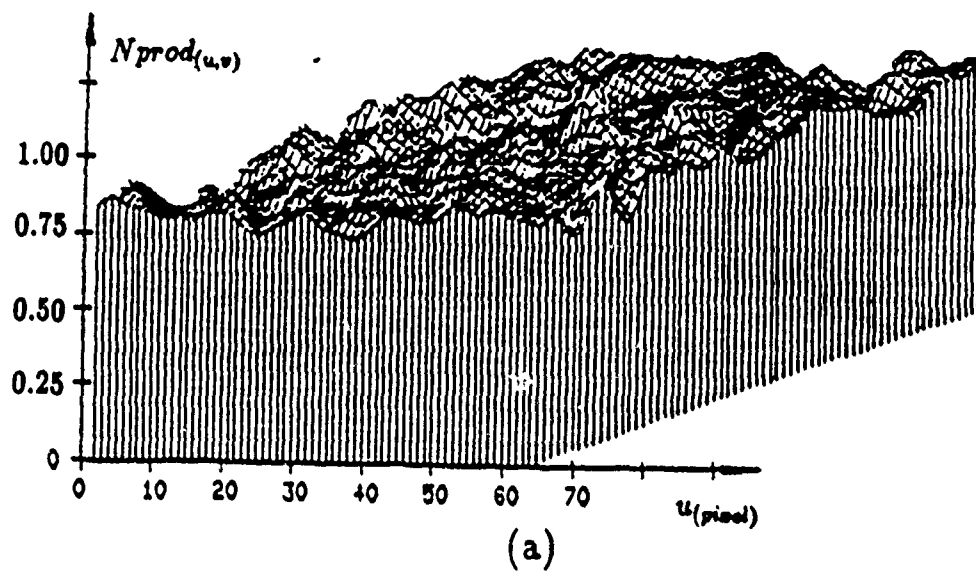
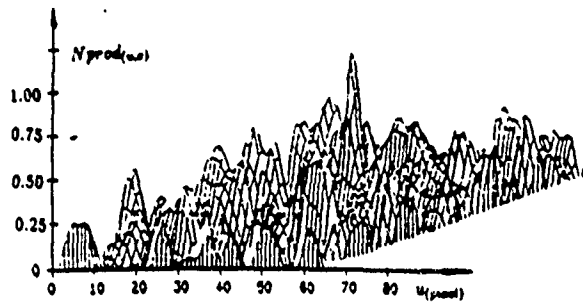
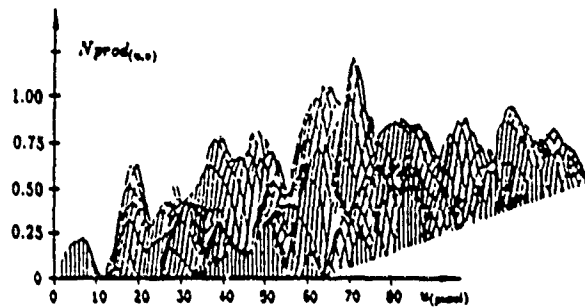


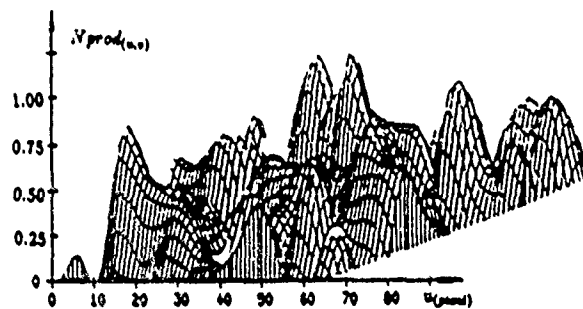
Fig.11-Effect of mean-elimination on correlation evaluation.
 (a) Without mean-elimination. (b) After mean-elimination



(a)



(b)



(c)

Fig.12- Correlations between deformed and undeformed speckle patterns sampled by different resolutions.

- (a) Frame resolution: 1024×1024 ; sampling interval: $T = 0.264S$
- (b) Frame resolution: 512×512 ; sampling interval: $T = 0.528S$
- (c) Frame resolution: 256×256 ; sampling interval: $T = 1.06S$

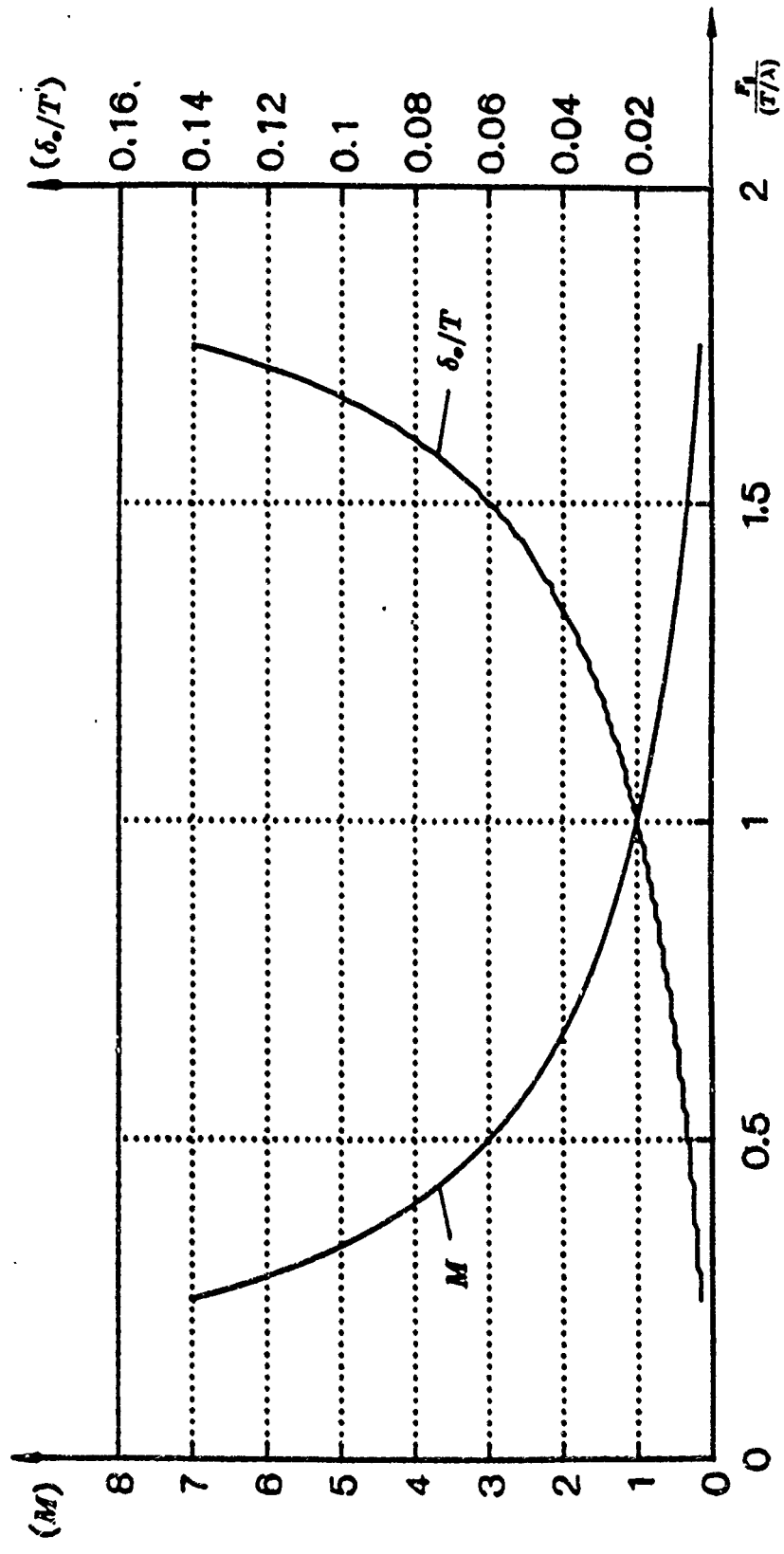


Fig.13 magnification and uncertainty as a function of recording aperture

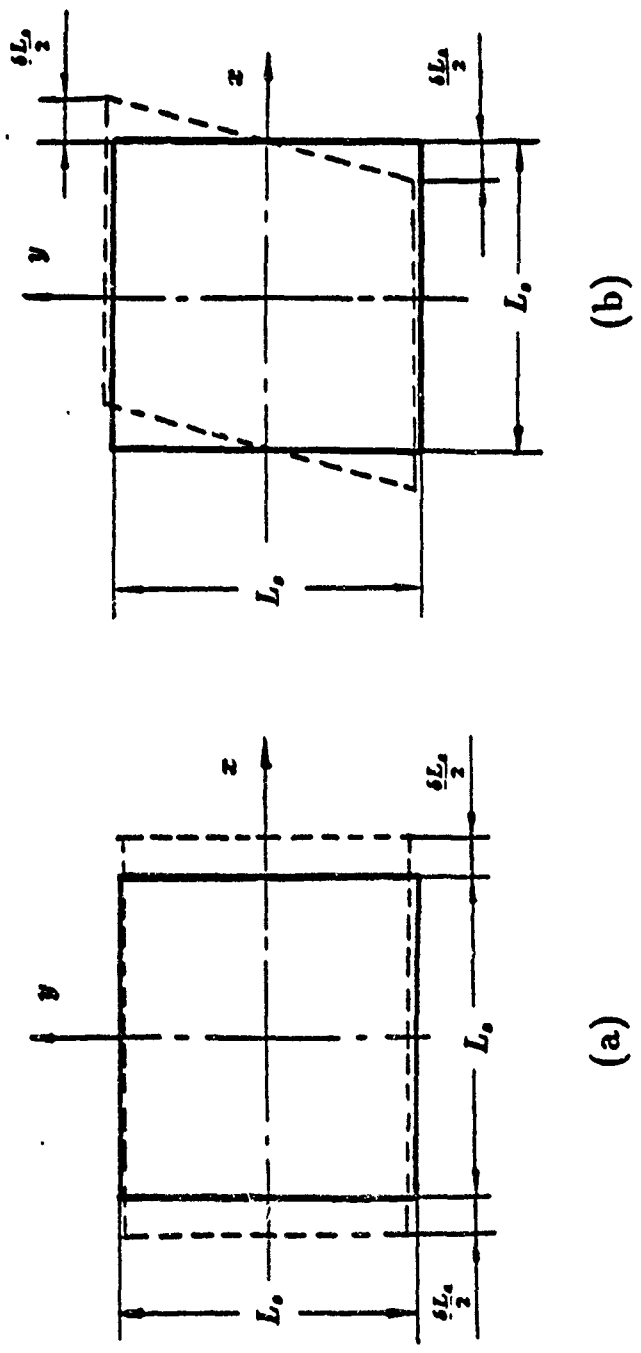


Fig.14-Decorrelation effect of specimen deformation

(a) Uniaxial tension. (b) Pure shear.

a reprint from Applied Optics

Resolution of resultant displacement into components in double exposure speckle photography

Piyush K. Gupta and Fu-Pen Chiang

An equation to resolve the resultant displacement, recorded using double exposure laser speckle photography, at the surface of a cylinder suffering radial expansion, axial twist, and transverse displacement, into these three components is presented.

I. Introduction

Double exposure laser speckle photography is a well-established optical technique for measuring in-plane displacement.¹⁻⁵ Applying this technique to measuring in-plane displacement of planar surfaces is quite straightforward. Here we extend the principles of laser speckle movement to resolve the resultant displacement of a cylinder suffering radial expansion, axial twist, and transverse displacement, recorded using double exposure laser speckle photography, into the three components. The analysis also gives an insight into the proper design of the experiment to obtain the transverse component directly.

II. Resolution of Resultant y -Displacement into Components

Consider a cylinder with its axis lying transversely along the x -axis as shown in Fig. 1. The camera is placed in the x - z plane and focused on a plane parallel to the x - y plane. The three motions at the surface of the cylinder that contribute toward the resultant speckle displacement in the y -direction observed at the plane at which the camera is focused are (1) transverse motion y_t , (2) radial expansion of the cylinder y_r , and (3) surface displacement due to the twisting of the cylinder about the x -axis y_s . If θ is the angular displacement in radians, measured positive counterclockwise, $y_s = R\theta$, where R is the radius of the cylinder.

We derive below the relation governing the contributions of all these factors toward resultant speckle displacement in the y -direction at the plane at which the camera is focused. We consider four possible imaging conditions separately and show that, under certain assumptions, they can be reduced to a unique governing equation.

Let d_{fp} be the distance between the x -axis and the plane at which the camera is focused, as depicted in Figs. 2-5.

Consider the points (such as a) on the surface of the cylinder lying within the region that can be imaged by the camera. Point a is displaced to a_t due to the transverse displacement y_t . It then displaces to a_r as a result of the radial expansion y_r , and finally to a_s after suffering a radial twist of angle θ . The net effect of these displacements on the speckles generated on the focus plane by light scattered from point a is a transverse displacement at the focus plane. It is important to note that, although a pure translation component of the displacement of point a yields equal transverse displacement of the speckles, the twisting or rotation of point a by an angle θ results in rotation of the speckles by an angle 2θ and with an arm equal to the distance between the focus plane and point a . Assuming that the displacements and rotation are small so that the speckle fields before and after the total displacement remain correlated, for the four different imaging conditions we can write the following equations which relate the total transverse displacement y of the speckles in the focus plane to the transverse (y_t), radial (y_r), and twist (y_s) components of the displacement suffered by points such as a .

Case 1. For points on the surface of the cylinder lying to the left of the focus plane and above the origin, using Fig. 2 we can write

$$y = y_t + y_r \sin \phi - (R + y_r)\theta \cos \phi + \{(R + y_r) \cos \phi + (R + y_r)\theta \sin \phi - d_{fp}\} \cdot 2\theta \quad (1)$$

which rearranges to

Piyush Gupta is with Systron Donner Corporation, Inertial Division, 2700 Systron Drive, Concord, California 94518, and Fu-Pen Chiang is with SUNY at Stony Brook, Department of Mechanical Engineering, Stony Brook, New York 11794-2300.

Received 17 March 1989.

0003-6935/90/111642-04\$02.00/0.

© 1990 Optical Society of America.

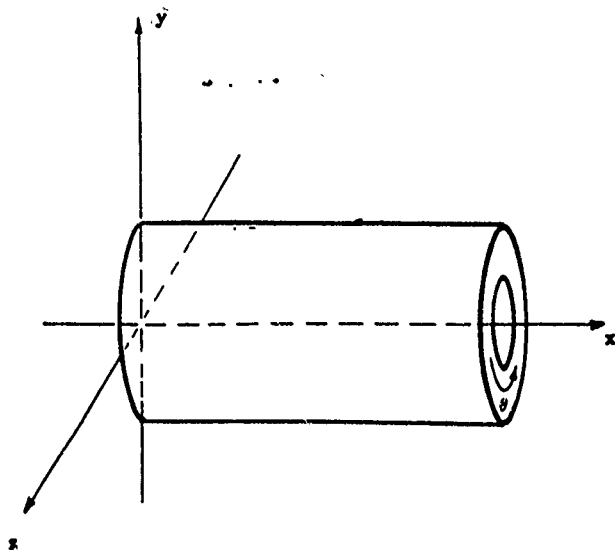


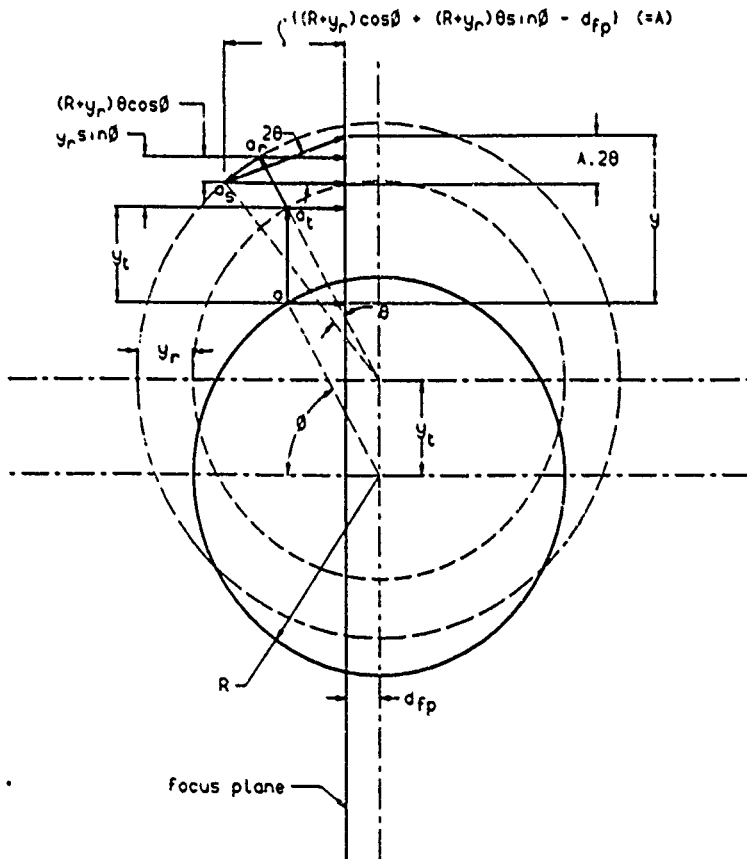
Fig. 1. Cylinder as viewed by the camera.

$$y = y_t + y_r \sin\phi + (R + y_r)\theta \cos\phi + (R + y_r) \sin\phi \cdot 2\theta^2 - d_{fp} 2\theta. \quad (2)$$

Case 2. For points lying above the origin and to the right of the plane at which the camera is focused, using Fig. 3 we have

$$y = y_t + y_r \sin\phi - (R + y_r)\theta \cos\phi - [d_{fp} - (R + y_r) \cos\phi - (R + y_r)\theta \sin\phi] \cdot 2\theta, \quad (3)$$

which rearranges to



$$y = y_t + y_r \sin\phi + (R + y_r)\theta \cos\phi + (R + y_r) \sin\phi \cdot 2\theta^2 - d_{fp} 2\theta. \quad (4)$$

Case 3. For points lying below the origin and to the left of the focus plane, using Fig. 4 we can write

$$y = y_t - y_r \sin\phi' - (R + y_r)\theta \cos\phi' + [(R + y_r) \cos\phi' - (R + y_r)\theta \sin\phi' - d_{fp}] \cdot 2\theta, \quad (5)$$

which rearranges to

$$y = y_t - y_r \sin\phi' + (R + y_r)\theta \cos\phi' - (R + y_r) \sin\phi' \cdot 2\theta^2 - d_{fp} 2\theta. \quad (6)$$

Case 4. For points below the origin and to the right of the plane at which the camera is focused, using Fig. 5 we have

$$y = y_t - y_r \sin\phi' - (R + y_r)\theta \cos\phi' - [d_{fp} - (R + y_r) \cos\phi' + (R + y_r)\theta \sin\phi'] \cdot 2\theta. \quad (7)$$

which rearranges to

$$y = y_t - y_r \sin\phi' + (R + y_r)\theta \cos\phi' - (R + y_r) \sin\phi' \cdot 2\theta^2 - d_{fp} 2\theta. \quad (8)$$

But since $\phi' = 2\pi - \phi$, we have $\cos\phi' = \cos\phi$ and $\sin\phi' = -\sin\phi$. If $y_r \ll R$, then $(R + y_r) \approx R$, where R is the radius of the outer surface of the cylinder. Then, we can rewrite Eqs. (2), (4), (6), and (8) as

$$y = (y_t - 2d_{fp}\theta) + \sin\phi(y_r + 2R\theta^2) + \cos\phi(R\theta). \quad (9)$$

The camera records an image in a plane parallel to the x - y plane and the specklegram is read out at dis-

Fig. 2. Speckle displacement for points above the origin and to the left of the focus plane.

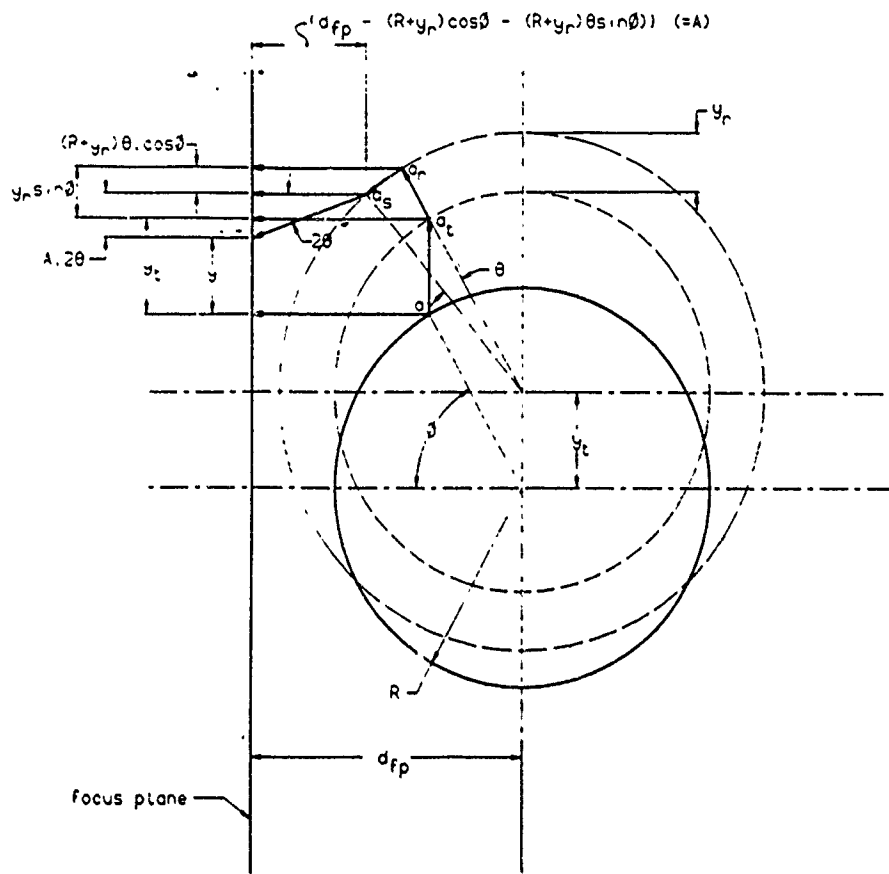


Fig. 3. Speckle displacement for points above the origin and to the right of the focus plane.

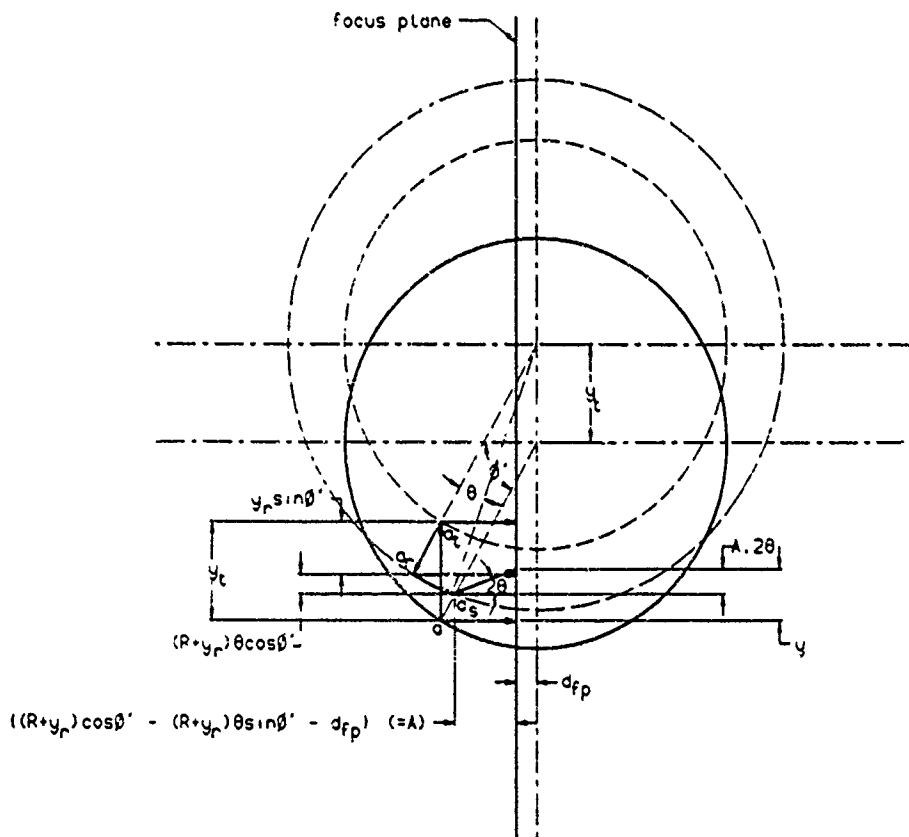


Fig. 4. Speckle displacement for points below the origin and to the left of the focus plane.

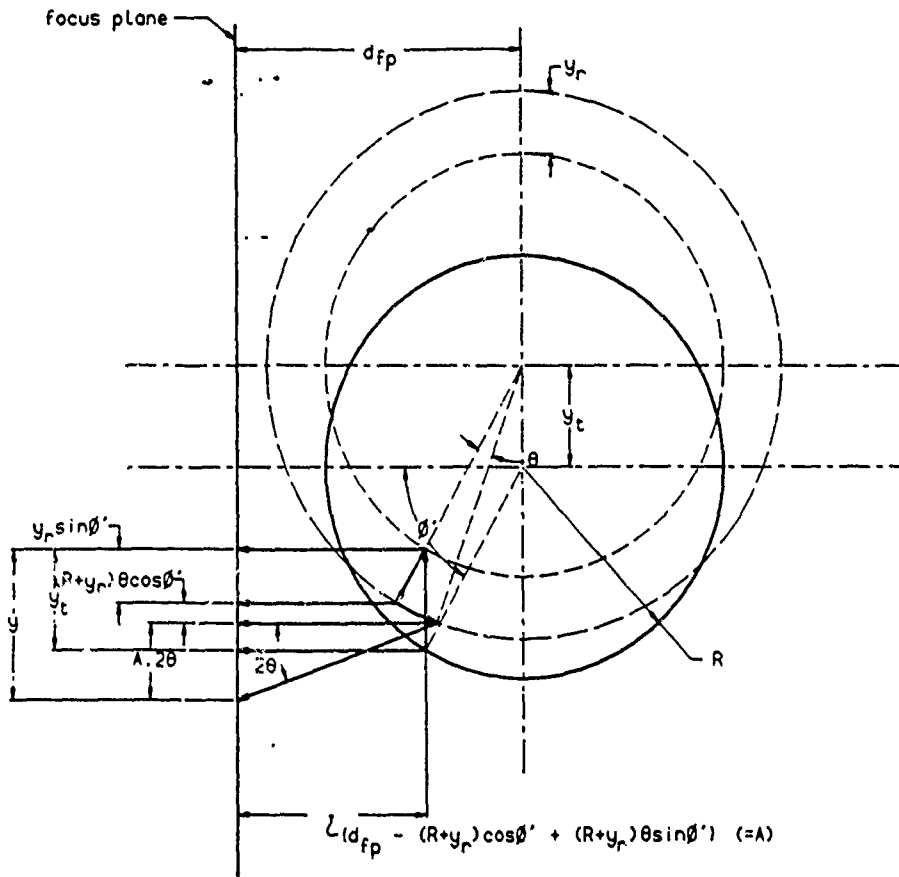


Fig. 5. Speckle displacement for points below the origin and to the right of the focus plane.

crete points (i, j) , where i is the x -coordinate and j is the y -coordinate of the point being analyzed using the pointwise technique. Equation (9) governs the resultant y -displacement recorded on the specklegram at any point (i, j) . The angle ϕ is independent of i . For a given i , we can select three suitable values of j (which yield minimum error) to obtain a system of three linear equations with the three coefficients shown within parentheses in Eq. (9) as the three unknowns. By solving this system of three linear equations, we obtain the twist effect $y_t = R\theta$ as the coefficient of the cosine term in Eq. (9). Then, knowing $R, R\theta$ (thus θ as well), and $(y_r + 2R\theta^2)$ as the coefficient of the sine term, we evaluate y_r . Further, if the camera is focused at a distance d_{fp} , knowing θ , we evaluate y_t . Note that, if $d_{fp} = R/2$, at $\phi = 0, y = y_t$. Hence, if we analyze the specklegrams at $\phi = 0$, we can directly obtain the transverse displacement component when $d_{fp} = R/2$.

III. Conclusion

To apply double exposure laser speckle photography to problems involving cylindrical surfaces undergoing radial expansion, axial twist as well as transverse displacement, the camera should be preferably focused at a plane which is at a distance of $R/2$ ahead of the

vertical plane containing the axis. With such a setup, the three components can be readily obtained from a single double exposure specklegram, with the transverse component available directly without the need to solve the simultaneous equations.

Financial support provided by the Army Ballistic Laboratory through a contract to S&D Dynamics, Inc. and partial support from the Army Research Office, Engineering Science Division, through contract DAAL0388K0033 is gratefully acknowledged.

References

1. P. K. Gupta and F. P. Chiang, "Laser Speckle Interferometry Applied to Studying Transient Vibrations of a Cantilever Beam," *J. Sound Vib.* 133, 251-259 (1989).
2. E. Archbold, J. M. Burch, and A. E. Ennos, "Recording of In-Plane Surface Displacement by Double Exposure Speckle Photography," *Opt. Acta* 17, 883-898 (1970).
3. E. Archbold and A. E. Ennos, "Displacement Measurement from Double Exposure Laser Photographs," *Opt. Acta* 19, 253-271 (1972).
4. F. P. Chiang, "A New Family of 2D and 3D Experimental Stress Analysis Techniques Using Laser Speckles," *SM Archives* 3, 1-32 (1978).
5. R. P. Khetan and F. P. Chiang, "Strain Analysis by One-Beam Laser Speckle Interferometry. 1: Single Aperture Method," *Appl. Opt.* 15, 2205-2215 (1976).

Segmentation of Bilevel Images Using Mathematical Morphology

Jin-Chang Cheng and Hon-Son Don
Department of Electrical Engineering
State University of New York
Stony Brook, NY 11794

Abstract

This paper presents the results of a study on the use of morphological skeleton transformation to segment grayscale images into bilevel images. When bilevel image is digitized, the result is a grayscale image due to the point spread function of digitizer, non-uniform illumination and noise. Our method can recover the original bilevel image from the grayscale image. The theoretical basis of the algorithm is the physical structure of the skeleton set. A connectivity property of the grayscale skeleton transformation is used to separate and remove the background terrain. The object pixels can then be obtained by applying a global threshold. Experimental results are given.

Keywords: segmentation, mathematical morphology, thresholding

1. Introduction

Assume that the input gray scale image is intended to have only two levels, such as the printed texts and line drawings. The most commonly used method in extracting object region from background is "thresholding". Thresholding classifies the pixels of a given image into two groups, that is, object and background. If the object is clearly distinguishable from the background, the gray scale histogram will be bimodal and the threshold for segmentation can be easily chosen as the bottom of the valley. However, gray scale histograms are not always bimodal. Several methods have been proposed to solve this problem so that the valley seeking technique can still be applied. Some other publications on automatic selection of thresholding values include iterative method [9], illumination-independent contrast measure [11], moment-preserving method [13], and methods based on entropy of the histogram [3, 8], on local estimation of uniform error analysis [1] and on other image statistics. Some aforementioned algorithms use simple global threshold while others use (multiple) local thresholds. However, simple global thresholding of such images produces poor results due to the nonuniformity

of the light distribution across the images. Local thresholding techniques need multiple passes over the data. This causes fast data throughput difficult. Recently, Pavlidis and Wolberg [6] proposed a method based on a model of distortion of the images and tried to invert the distortion process. Another iterative method was proposed by Pérez and Gonzalez [7]. They used Taylor series expansion method on an illumination-reflectance model. Like many iterative algorithms, the result of their method depends on the selection of initial values.

Mathematical morphology was first introduced by Matheron[5] and Serra[10]. It is characterized by its robustness, speed, accuracy and flexibility. It can extract information about the geometrical structure of an object by transforming it with another smaller and simpler object, called structuring element. Many applications of mathematical morphology have been proposed in literature, such as noise removal, feature detection (edges, holes, corners, ..., etc), thickenings and thinnings. Skeletonization by using binary mathematical morphological operations was introduced by Serra[10].

Our major topic in this paper, is the segmentation of bilevel image. It is done by first extracting the skeleton of the given gray scale image by the gray scale morphological operations and then using the connectivity property of the skeleton to segment the given image into bilevel image. For many kinds of image, like machine parts, printed texts and line drawings, the object pixels and the noises can only be inscribed by spheres of small size, while the pixels in the relatively smooth background can be inscribed by spheres of much larger size. The centers of the inscribing spheres (i.e., skeleton points) of these different kinds of regions usually are quite distance apart. Therefore they can be easily separated. After properly separating spheres of small size from large ones, we retain only the foreground (object or noise) pixels and remove the background pixels. A global thresholding can then be applied to extract the object pixels.

In our experiments, the algorithm is compared against global thresholding scheme, a local thresholding scheme reported by White and Rohrer[14] as well as a segmentation algorithm reported by Pavlidis and Wolberg[6].

In Section 2, the basic morphological set transformations are summarized. In Section 3, we discuss the morphological skeleton representation of grayscale images. In Section 4, we propose the algorithms for bilevel image segmentation. Experimental results and conclusions are given in Section 5 and Section 6, respectively.

2. Basic Morphological Set Operations

In the following, we summarize some basic morphological transformations. The details of these formula can be found in [2].

Let E^N denote the set of N -tuple integers, ... B be subsets of E^N , x be a vector in E^N .

Definition 1: The binary dilation \oplus of A by B is defined by

$$A \oplus B = \{ c \in E^N \mid c = a + b, \text{ for some } a \in A \text{ and } b \in B \}$$

Definition 2: The binary erosion \ominus of A by B is defined by

$$A \ominus B = \{ x \in E^N \mid x + b \in A, \text{ for every } b \in B \}$$

Definition 3: The binary opening \circ of A by B is defined by

$$A \circ B = (A \ominus B) \oplus B$$

Definition 4: The binary closing \bullet of A by B is defined by

$$A \bullet B = (A \oplus B) \ominus B$$

Let f and g be functions defined on E^N , x and y are vectors in E^N .

Definition 5: Let $A \subseteq E^{N+1}$ and $F = \{ (x, y) \in E \mid (x, y) \in A \}$. The top (or top surface) of A is defined by

$$T[A](x) = \max \{ y \mid (x, y) \in A \}$$

Definition 6: The umbra of f , denoted by $U[f]$ is defined by

$$U[f] = \{ (x, y) \mid y \leq f(x), x \in E^N, y \in E \}$$

Definition 7: The grayscale dilation of f by B is defined by

$$f \oplus B = T[U[f] \oplus U[B]]$$

Proposition 1:

$$(f \oplus B)(x) = \max_b \{ f(x-b) + B(b) \}. \quad (1)$$

Definition 8: The grayscale erosion of f by B is defined by

$$f \ominus B = T[U[f] \ominus U[B]]$$

Proposition 2:

$$(f \ominus B)(x) = \min_b \{ f(x+b) - B(b) \}. \quad (2)$$

Definition 9: The grayscale opening of f by B is defined by

$$f \circ B = (f \ominus B) \oplus B. \quad (3)$$

Definition 10: The grayscale closing of f by B is defined by

$$f \bullet B = (f \oplus B) \ominus B. \quad (4)$$

Proposition 3: Umbra Homomorphism Theorem

$$U[f \oplus B] = U[f] \oplus U[B] \quad (5)$$

$$U[f \bullet B] = U[f] \bullet U[B] \quad (6)$$

Other formulas can be found in [2, 5, 10].

3. Morphological Skeleton Transformation of Gray Scale Images

The skeleton $S(f)$ of a discrete N -dimensional binary image f is defined as the set of the centers of the maximal N -dimensional spheres, inscribable inside the image f . An N -dimensional sphere is *maximal* if it is not properly contained in any other spheres which are totally included in f . Hence, a maximal sphere must touch the boundary of f at least at two

different points. The morphological skeleton transformation of a discrete binary image has been studied in [5,10]. For gray scale images, the maximal sphere can be similarly defined and the gray scale skeleton transformation can be derived by using the concept of umbra. Let B be a N -dimensional gray scale structuring element with semi-sphere shape of flat bottom. We define the skeleton $S(f)$ of a discrete N -dimensional gray scale image f as the set of the centers of B which the umbra $U[B]$ of B is the maximal umbra inscribable inside the umbra $U[f]$ of f . An umbra of a semi-sphere is maximal if it is not properly contained in any other umbra of semi-sphere totally included in $U[f]$. Therefore, the gray scale skeleton can be derived as follows:

$$S_n(f) = (U[f] \ominus U[nB]) - (U[f] \ominus U[nB]) \circ U[B] \quad (7)$$

and

$$S(f) = \bigcup_{n=0}^M S_n(f)$$

where $S_n(f)$ denotes the n^{th} skeleton subset of $U[f]$, B is the N -dimensional gray scale structuring element and nB is defined as the gray scale dilation of B with itself n times:

$$nB = B \oplus B \oplus \dots \oplus B \quad (n \text{ times}).$$

By $S_0(f) = U[f] - U[f] \circ U[B]$, we can see that the 0^{th} skeleton subset of $U[f]$ consists of those points in $U[f]$ and they can not be inscribed or touched by any unit-size maximal sphere B . Similarly, $S_n(f)$ consists of points which are the translated centers of size- n sphere nB when the translation of $U[nB]$ is inscribable inside $U[f]$ but that of $U[(n+1)B]$ is not. To find the skeleton by using gray scale erosions and openings, we need the following property.

Proposition 4:

$$S_n(f) = U[f \ominus nB] - U[(f \ominus nB) \circ B] \quad (8)$$

proof: By (6), $U[f] \ominus U[nB] = U[f \ominus nB]$. Let $h = f \ominus nB$, then by (3)(5), $(U[f] \ominus U[nB]) \circ U[B] = U[h] \circ U[B] = (U[h] \ominus U[B]) \oplus U[B] = U[h \ominus B] \oplus U[B] = U[(h \ominus B) \oplus B] = U[h \circ B] = U[(f \ominus nB) \circ B]$. Therefore, $S_n(f) = U[f \ominus nB] - U[(f \ominus nB) \circ B]$.

The morphological skeleton transformation is invertible. The discrete gray scale image f can be exactly reconstructed as the finite union of its $M+1$ skeleton subsets dilated by the structuring element B of proper size.

$$f = T\left[\bigcup_{n=0}^M S_n(f) \oplus U[nB]\right] \quad (9)$$

The structuring element B in (8) can be considered as an unit-distance probe. Skeleton point in $S_n(f)$ has the minimum distance of n to the boundary of image f . The distance measure can be interpreted by using city-block, chess-board, rectangular-solid distance, Euclidean distance or any other distance measures. Note that the first three distances satisfy the equation $nB = B \oplus B \oplus \dots \oplus B$ (n times). The Euclidean distance does not satisfy this equation. Therefore the Euclidean distance is not suitable to be used as the

unit-size structuring element for the iterative skeleton transformation (8). Although some structuring elements do not look like spheres in shape, we still adapt the term "maximal sphere" to the structuring elements not spherical in shape but are maximal.

4. Segmentation of Bilevel Image

In this section, we present our algorithm for bilevel image segmentation. The basic idea of the segmentation algorithm as well as its computational complexity are discussed in Section 4.1. In Section 4.2, the theoretical analysis of our algorithm are given. The selection of the structuring element and the robustness of the algorithm are discussed in Section 4.3. Finally, a comparison with the Top Hat Transformation [10] is given in Section 4.4.

4.1. Basic Approach

Given an ideal bilevel image f_i whose gray values at pixel (x,y) are $f_i(x,y)=a$ for object pixels, and $f_i(x,y)=0$ for background pixels. Let its blurred image f be denoted by

$$f(x,y) = f_i(x,y) + \text{noise}(x,y) + \text{offset}(x,y),$$

where $\text{noise}(x,y)$ and $\text{offset}(x,y)$ are the added noise and the background offset caused by non-uniform lighting distribution, respectively. Our aim is to find a base surface $b(x,y)$ which is the estimation of the offset, i.e., $b(x,y)=\text{offset}(x,y)$. Then the bilevel image f_i can be reconstructed by thresholding $f(x,y)-b(x,y)$.

To find the base surface b for a given gray scale image f , we begin with the skeletonization of the input image f by using the gray scale skeleton transformation (8). Recall that the skeleton point is the center of the maximal inscribed spheres. We will call the skeleton point which inscribes the object (background) pixels as *object (background) center*. The skeleton of a connected discrete image may be disconnected. We will explain in the following that the disconnection occurs in the intersection of object pixels and background pixels. Within object pixels, the radius change is continuous and the distance between two neighboring skeleton points are less than or equal to two in noise-free case (slightly larger than two in noisy case). The magic number two is due to the possible increment of the size by one between two consecutive maximal spheres and the one pixel downshift of the larger sphere. Since the sizes of the spheres which inscribe the background pixels are relatively much larger than those of the spheres which inscribe the object pixels. Therefore, in the intersection, the distance between the object center and its successive background center will be much greater than two. This big jump of distance between two successive centers causes the disconnection of the skeleton. By detecting this distance change, we can easily locate those background centers.

Two skeleton points are *connected* if the distance between them is less than or equal to two. Otherwise, we say that they are *disconnected*. For disconnected skeleton points, the skeleton associated with the larger radius is a background

center. Any skeleton point connected to a background center is a background center and any skeleton point connected to an object center is an object center.

After the background centers are found, the base surface b can then be constructed from those background centers by using the inverse grayscale skeleton transformation (9). The base surface b is the top of the union of those umbras which are the translation of the umbra of nB by Q , for every background center Q in $S_b(f)$. Here, we list our segmentation algorithm as follows:

Algorithm:

Input: a grayscale image f .

Output: a segmented bilevel image.

Steps:

0. Initialize the base surface $b(x,y)=0$.
1. Using (8), find the skeleton subsets $S_n(f)$, for $0 \leq n \leq M$.
2. Construct b by the top of the union of those umbras which are the translation of the umbra of nB by Q , for every background center Q in $S_n(f)$:

$$b = T \left[\bigcup (U[nB])_{Q} \right]$$
3. Remove the background offset by subtraction: $g(x,y)=f(x,y)-b(x,y)$.
4. Thresholding the background-removed image $g(x,y)$.

The value of M in step 1 can be the size of the minimum square mask of M by M pixels which can not be completely contained inside any object in the image. The computation time is dominated by the skeletonization process in step 1. Assume we choose the structuring element B as an 1-D unit-size rectangle (see Fig. 1a) for 1-D line scan process and unit-size rectangular solid for 2-D process. Then, by (2), the grayscale erosion(dilation) of f by B needs 2 parallel min(max) operations in 1-D or 8 parallel min(max) operations in 2-D by using a parallel computer. Since the grayscale erosion(dilation) of f by nB can be done by successively eroding f by B n times, $0 \leq n \leq M$, step 1 needs $\frac{1}{2}M(M+1)$ grayscale erosions by B plus $M+1$ grayscale erosions by B and M dilations by B for the opening operations in (8), which amounts to a total of (M^2+5M+2) grayscale erosion/dilation operations. Therefore, step 1 totally takes (M^2+5M+2) and $4(M^2+5M+2)$ parallel min/max operations on a parallel computer for 1-D and 2-D fashion, respectively, for the selected structuring element B . Note that the 2-D fashion takes 4 times longer than 1-D line scan fashion.

4.2. Theoretical Analysis

For the simplicity of description, we assume the image is a one-dimensional image which can be the row- (or column-) version of a two-dimensional image. All the proper-

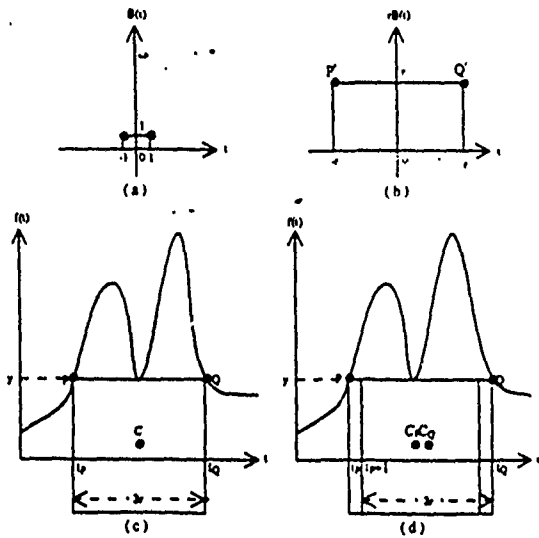


Fig. 1: (a) 1-D rectangular unit-size structuring element, (b) 1-D rectangular size-r structuring element, (c) A maximal inscribed size-r sphere with center at C, touches the local minimum point and two points P and Q on the discrete signal f , when $t_p - t_q$ is even, (d) Two neighboring maximal inscribed spheres with centers at C_1 and C_2 , respectively, when $t_p - t_q$ is odd.

ties derived below can be naturally and easily extended to two-dimensional image. We also assume the pixels on the image boundary belong to the background. In the continuous image case, a maximal sphere inscribed in f must touch the top of f at least at two points. These two points which have maximum distance among those touched points are called *two farthest points*. While in discrete image case, due to the nature of discrete skeleton transformation (8), the discrete skeleton may be of width up to two. The two farthest points may be touched by one maximal sphere or by two neighboring maximal spheres of the same size. In either case, without loss of generality, we still say that these two farthest points are *touched by a maximal sphere*. First, we give the following observation:

Proposition 5: Assume the structuring element B is a 1-D unit-size rectangle (see Fig. 1a), that is, $B(t)=1$ for $|t| \leq 1$, and $B(t)$ is undefined for $|t| > 1$. Let $P=(t_p, f(t_p))$ and $Q=(t_q, f(t_q))$ be two farthest points which are touched by a maximal sphere rB of radius r , $t_p \leq t_q$, then

- (i) $f(t_p)=f(t_q)$.
- (ii) $r = \frac{1}{2} |t_q - t_p|$, and $t_q - t_p = 2r$, if $t_q - t_p$ is even; $t_q - t_p = 2r + 1$, if $t_q - t_p$ is odd.
- (iii) if and only if $t_q = t_p$ or $t_q = t_p + 1$ then $r=0$.
- (iv) If $t_q - t_p$ is even, then the inscribed maximal sphere is centered at $(\frac{t_p + t_q}{2}, f(\frac{t_p + t_q}{2}))$, otherwise P and Q are touched by two neighboring maximal spheres which are centered at $(t_p + r, f(t_p) - r)$ and $(t_p + r + 1, f(t_p) - r)$, respectively.

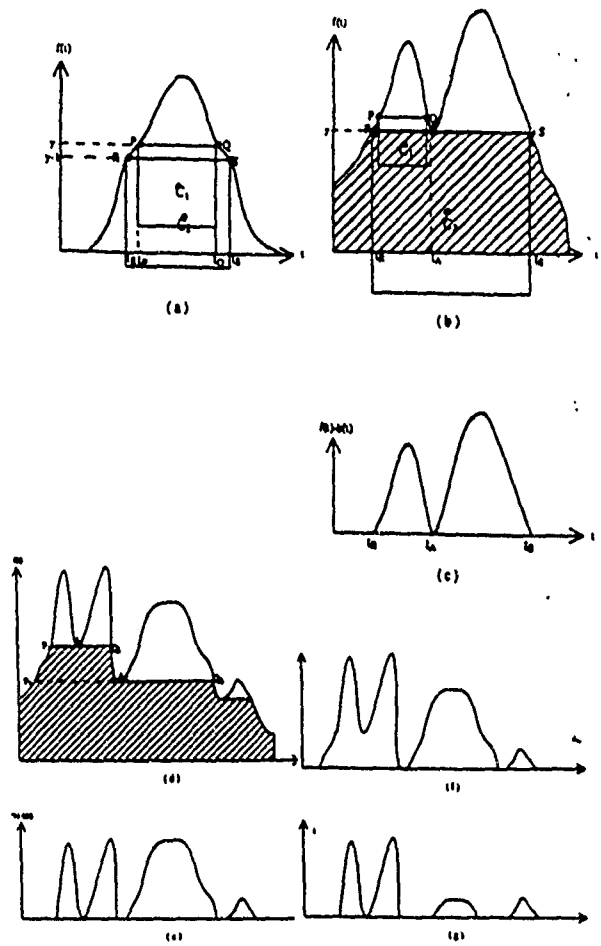


Fig. 2: (a) 1-D representation of bilevel one-peak signal, marked with two consecutive and connected skeleton points C_1 and C_2 , (b) A two-peak signal, marked with two consecutive but disconnected skeleton points C_1 and C_2 , respectively. Pixels in shaded area are background pixels, (c) Result of the signal in Fig. 2b after the removal of the shaded area (background). Note that all signals are plotted with continuous line, they actually represent discrete signals.

proof: Since B is an unit size rectangular and $rB = B \oplus B \oplus \dots \oplus B$ (r times), therefore rB is a size- r rectangular, i.e., $rB(t)=r$ for $|t| \leq r$ and $rB(t)$ is undefined for $|t| > r$ (see Fig. 1b). When $t_q - t_p$ is even (see Fig. 1c), umbra of rB slides around and fit within umbra of f , the two corner points P' and Q' in rB will match f at two farthest points $P=(t_p, f(t_p))$ and $Q=(t_q, f(t_q))$, respectively. Since $rB(-r)=rB(r)=r$, we get $t_q - t_p = |P - Q| = |P' - Q'| = 2r$, and the skeleton point is at $(\frac{1}{2} |P - Q|, f(\frac{t_p + t_q}{2})) = (t_p + r, f(t_p) - r)$. When $t_q - t_p$ is odd (see Fig. 1d), P will be touched by the maximal sphere rB of centered at $(t_p + r, f(t_p) - r)$ and matched P' while Q will be touched by the maximal sphere rB of centered at

$(t_p+r+1, f(t_p)-r)$ and matched Q' . So, $t_Q-t_p=2r+1$.
(iii) is followed obviously by (ii).

In the following, we fix our structuring element as the unit-size rectangle for 1-D process, and unit-size rectangular solid for 2-D process.

Proposition 6: If the shape of a simple signal has only one peak, that is, no local minimum points, then every two consecutive skeleton points are at most of distance 2.

proof: Let $P=(t_p, f(t_p))$ and $Q=(t_Q, f(t_Q))$, $t_p < t_Q$, are two farthest points touched by a maximal sphere of radius r (see Fig. 2a). For simplicity, assume t_Q-t_p is even and let $y=f(t_p)$. Then, by proposition 5, $y=f(t_p)=f(t_Q)$, $r = \frac{1}{2}[t_Q-t_p]$ and the inscribed maximal sphere is centered at $C_1=(t_p+r, y-r)$. Let two neighboring points of P and Q are $R=(t_R, y-1)$ and $S=(t_S, y-1)$ respectively. Then t_R must be equal to t_p or t_p-1 and t_S must be equal to t_Q or t_Q+1 . Let the inscribed maximal sphere which touches two farthest points R and S be of radius r' , and centered at C_2 . In $t_R=t_p-1$ and $t_S=t_Q+1$ case, we get $r' = \frac{1}{2}|R-S| = \frac{1}{2}[t_S-t_R] = r+1$ and $C_2=(t_p+r', y-1-r')=(t_p+r, y-r-2)$, therefore $|C_1C_2| = \max((t_p+r)-(t_p+r), (y-r)-(y-r-2)) = 2$. In other cases, $t_R=t_p$ or $t_S=t_Q$, which imply $|C_1C_2| \leq 2$.

Proposition 6 says that none of the skeleton points in a simple (one peak) signal, which has no local minimums, are background centers. Thus, the base surface will be zero. Therefore we have shown that a simple signal will be preserved and not be affected by our segmentation algorithm. In the following, we discuss the more general cases.

Proposition 7: For objects with more than one local minimum points, any skeleton point of the inscribed maximal sphere which touches one of the local minimum points must be a background center.

proof: Let $A=(t_A, f(t_A))$ be an local minimum point of f , $y=f(t_A)$, and the inscribed maximal sphere which touches A is of radius r , centered at C_2 , and touches f at two farthest points $R=(t_R, f(t_R))$ and $S=(t_S, f(t_S))$, $t_R+1 < t_A < t_S-1$, respectively (see Fig. 2b). By proposition 5, we have $f(t_R)=f(t_S)=f(t_A)=y$, $r = \frac{1}{2}[t_S-t_R]$ and the center $C_2=(t_R+r, f(t_R)-r)$. Assume points $P=(t_p, f(t_p))$ and $Q=(t_Q, f(t_Q))$ are two respective neighbors of points R and A , and the inscribed maximal sphere which touches points P and Q is of radius r' and centered at C_1 , then we have $f(t_p)=f(t_Q)=f(t_A)+1$, $t_p=t_R$ or $t_p=t_R+1$, $t_Q < t_A$, $r' = \frac{1}{2}[t_Q-t_p]$ and center $C_1=(t_p+r', f(t_p)-r')$. Therefore, $|C_1C_2| = \max(|t_R+t_p-r'|, |f(t_A)-r-f(t_p)+r'|) = r-r'+1$. Since $r = \frac{1}{2}[t_S-t_R]$, $r' = \frac{1}{2}[t_Q-t_p]$, we immediately get $|C_1C_2| > 2$, and therefore C_2 is a background center.

By proposition 7, every skeleton points C_2 touches one of the local minimum points must be a background center. The translation of rB to C_2 is nothing but the inscribed maximal sphere with center at C_2 . Its top surface makes a horizontal baseline RS which passes the minimum point and inter-

sects the signal $f(t)$ at two points R and S . Furthermore, all the successive skeleton points of C_2 must be background centers since they are farther away from C_1 than C_2 . Therefore, every pixel underneath RS (the shaded area on Fig. 2b) will be removed by the segmentation algorithm. Fig. 2c shows the result after background removal of Fig. 2b. Fig. 2d shows a more general case which has more than one local minimum points. The shaded area in Fig. 2d are detected as background and are to be removed by step 2 of the segmentation algorithm. Fig. 2e shows the result of Fig. 2d after subtracting the background offset by step 3 of the segmentation algorithm.

4.3. About the Structuring Element

The reason why we chose the rectangular-solid distance in the previous section is simply because it is suitable to be used as mentioned in Section 3, and because it is flat on top and, therefore, easy to analyze. However, the computation time as well as the segmentation results will be the same by using any one of the three distance measures.

By using these kind of structuring elements, the connectivity of the skeleton may be biased and the spatial localization may be inaccurate. However, our goal is the segmentation of bilevel images, not the representation of the image structures by using the skeleton. These two problems will not affect the results of segmentation, since the algorithm will transform back to the original signal shape by (9) after the background are removed. The shape of the original signal above the background terrain will be preserved no matter which one of the structuring elements is used.

Another nice property of this algorithm is that it uses a fixed unit-sized structuring element for all kinds of input images. This property makes the automation possible. In addition, most morphological hardware architecture implementations are limited to a fix-sized structuring element. However, the choices (of the shape, size and/or orientation) of the structuring element are varying in most applications using mathematical morphology, like feature extraction, edge detection, etc, depending heavily on the image to be processed. For example, the choice of the structuring element to extract features depends on the size of the interested features and the requirement of how accurate the boundary localization is. If the size of the structuring element is too small, noises as well as unnecessary details may be extracted. If it is too large, some small features may not be extracted. Since our segmentation algorithm uses the structuring element only as a distance measure, the choice of structuring element is relatively not critical and is independent of the size and the orientation of the images. Any small size structuring element should be appropriate. No a priori knowledge of the size of the object and the orientation of the light source(s) is necessary. Therefore, this segmentation algorithm is robust.

4.4. Comparison with Top Hat Transformation

Top Hat Transformation (THT) [10] was originally proposed by Meyer and was applied to the extraction of

chromatin in cell nuclei. The transformation has been used in the past for the same kind of bilevel image segmentation as our algorithm. It is made up of the residual from the opening by rB and f , i.e. $f - f \circledast Q_1 \circledast rB$. It is also a kind of background removal process since $f \circledast Q_1 \circledast rB$ represents the background portion swept out by all the translations of rB under the umbra of f . However, in THT, a simple opening operation with a fixed size structuring element is usually not good enough for the background removal. The difficulties are not only that the size of the structuring element is not easy to determine (usually it depends on the size of the object), but also a fixed size structuring element can not separate the background properly. For example, consider the 1-D signal $f(t)$ shown in Fig. 2d. Let A_1 and A_2 be two local minimum points of $f(t)$, the inscribed maximal sphere which touches A_1 has two farthest points P_1 and Q_1 , and the inscribed maximal sphere which touches A_2 has two farthest points P_2 and Q_2 . Assume $|A_2Q_2| > |P_1Q_1|$. Then the size of the fixed size structuring element is either \geq or $< \frac{1}{2}|P_1Q_1|$. In the first case, if the size is $\geq \frac{1}{2}|P_1Q_1|$, then the two peaks in-between P_1A_1 and A_1Q_1 can not be separated because the structuring element can not touch the minimum point A_1 (see Fig. 2f). On the other hand, if the size of the structuring element is less than $\frac{1}{2}|P_1Q_1|$, then the base under the peak in-between A_2Q_2 will touch the small structuring element too much, and will thus be removed too much. As a result, the peak in-between A_2Q_2 will be narrowed and distorted (see Fig. 2g). Therefore the background can not be removed completely or the signal may be altered in either case by using a fixed size structuring element.

On the contrary, our segmentation algorithm uses a fixed, unit size structuring element, which will preserve the original shapes of all peaks (see Fig. 2e).

5. Experimental Results

In this section, we present some experimental results to show the performance of our segmentation algorithm. We used a simulated checkerboard image under two nonuniform illuminations and two real images for the experiments. For the purpose of comparison, the images in all experiments were segmented by 1) our algorithm, 2) a global threshold selected interactively to yield "optimum" visual results, 3) the White-Rohrer algorithm [14], and 4) the Pavlidis-Wolberg algorithm [6]. The White-Rohrer algorithm has been used successfully in character extraction applications.

The simulated image is a checkerboard of size 256×256 pixels. The light regions have gray scale intensity 40 and the dark regions have gray scale intensity 0. To simulate the nonuniform illumination effect, illumination patterns were added to this perfect checkerboard. Fig. 3a shows the checkerboard image superimposed an illumination pattern which changes linearly in the x and y directions plus a jump in the middle of the pattern, i.e.,

$$\text{offset}(x,y) = k_1x + k_2y + k_3u$$

where k_i 's are constants and $u=0$, if $x \leq x_0$ and $y \leq y_0$; $u=2$, if $x > x_0$ and $y > y_0$; otherwise $u=1$. The jump simulates the situation that some part of the image are under shadow. Fig. 3b

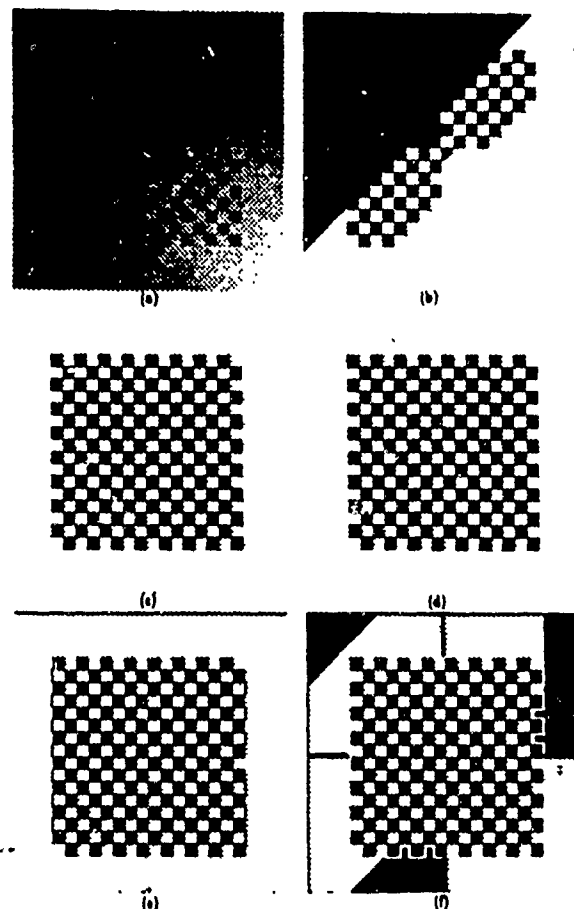


Fig. 3: (a) Simulated image 1, (b) Result of best thresholding of (a), (c)-(d) Results of our segmentation algorithm applied to (a), by using 2-D and 1-D scan fashion, respectively. (e) Results of White-Rohrer[14] algorithm, (f) Results of Pavlidis-Wolberg[6] algorithm.

shows the result of thresholding the image in Fig. 3a. The difficulty of thresholding the image is obvious. The image shown in Fig. 3a is easily segmented by our segmentation algorithm. We applied it in both 2-D fashion and 1-D row scan fashion, respectively, and got the same good results (Fig. 3c and Fig. 3d). Note that the nonuniform illumination effect is completely removed.

Fig. 4a shows the checkerboard image added to a Gaussian illumination pattern, defined by

$$\text{offset}(x,y) = \exp \left[\frac{-[(x-x_0)^2 + (y-y_0)^2]}{\sigma^2} \right]$$

with $\sigma=62$. Fig. 4b shows the result of thresholding the image in Fig. 4a. Using our segmentation algorithm in both 2-D fashion and 1-D row scan fashion, respectively, we got the same good results (Fig. 4c and Fig. 4d).

Fig. 5a shows a text image of size 192×512 pixels. The data were obtained by a vidicon camera with a nonuniform lighting condition. Thresholding result of the text image is also shown in Fig. 5b. Fig. 5c and Fig. 5d show the results

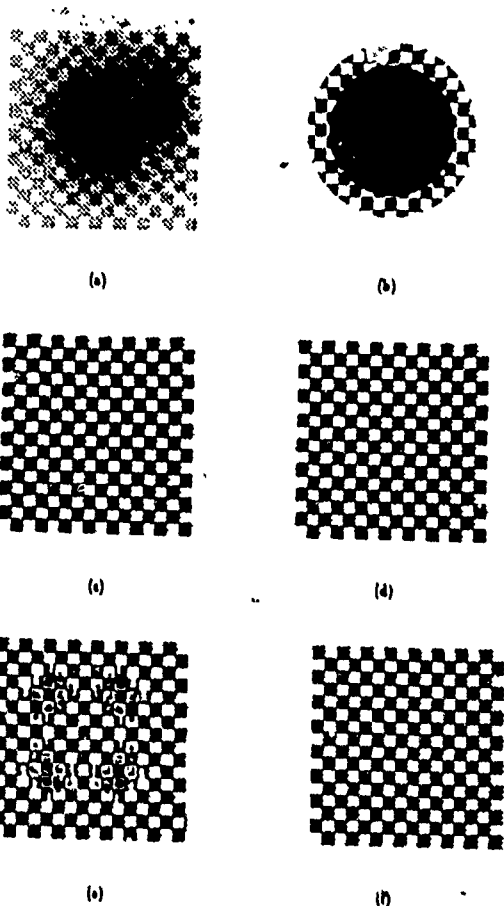


Fig. 4: (a) Simulated image 2, (b)-(f) Results of algorithms as in Fig. 3.

after segmentation by our algorithm using 2-D fashion and 1-D scan fashion, respectively. Fig. 5e and Fig. 5f show the segmented results by the White-Rohrer algorithm and the Pavlidis-Wolberg algorithm, respectively.

Experimental results using tool images are shown in Fig. 6 and Fig. 7. Fig. 6a and 7a show two tool images of size 256×256 pixels. Fig. 6b and 7b show the results obtained by global thresholding. Fig. 6c and 7c show the results after segmentation by using our algorithm in 2-D scan fashion and Fig. 6d and 7d show the results in 1-D scan fashion. Fig. 6e and 7e show the results by the White-Rohrer algorithm. Fig. 6f and 7f show the results by the Pavlidis-Wolberg algorithm.

6. Discussion and Conclusions

We have proposed a bilevel image segmentation algorithm by using gray scale morphological operations. The morphological skeleton transformations on binary images are extended to the gray scale images, which forms the basis of our segmentation algorithm. The segmentation algorithm skeletonizes the input gray scale image first, then check for discontinuity of skeleton set. The background due to the nonuniform illumination effect is then removed according to the discontinuity information. Finally, a global thresholding

SUNY at Stony Brook
Stony Brook, New York

Department of Electrical Engineering
SUNY at Stony Brook
Stony Brook, New York

Department of Electrical Engineering
SUNY at Stony Brook
Stony Brook, New York

Department of Electrical Engineering
SUNY at Stony Brook
Stony Brook, New York

Department of Electrical Engineering
SUNY at Stony Brook
Stony Brook, New York

Department of Electrical Engineering
SUNY at Stony Brook
Stony Brook, New York

Fig. 5: (a) Text image, 192×512 pixels, (b)-(f) Results of algorithms as in Fig. 3.

is applied to get the proper segmentation result. The algorithm is also shown to be better than the best manually selected global threshold under different illumination conditions. It is also better than a local threshold algorithm proposed by White and Rohrer and another algorithm proposed by Pavlidis and Wolberg. A visual comparison from the experimental results leaves the impression that this segmentation algorithm running in 1-D scan fashion can give as good results as that running in 2-D fashion. However, the computation time is four times faster by using 1-D scan fashion.

References

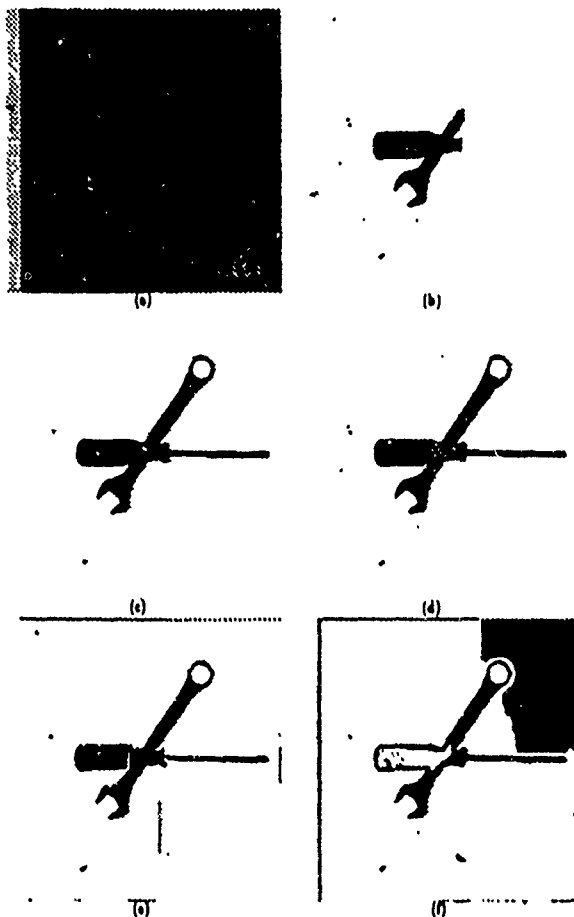


Fig. 6: (a) Tool image 1, 256×256 pixels, (b)-(f) Results of algorithms as in Fig. 3.

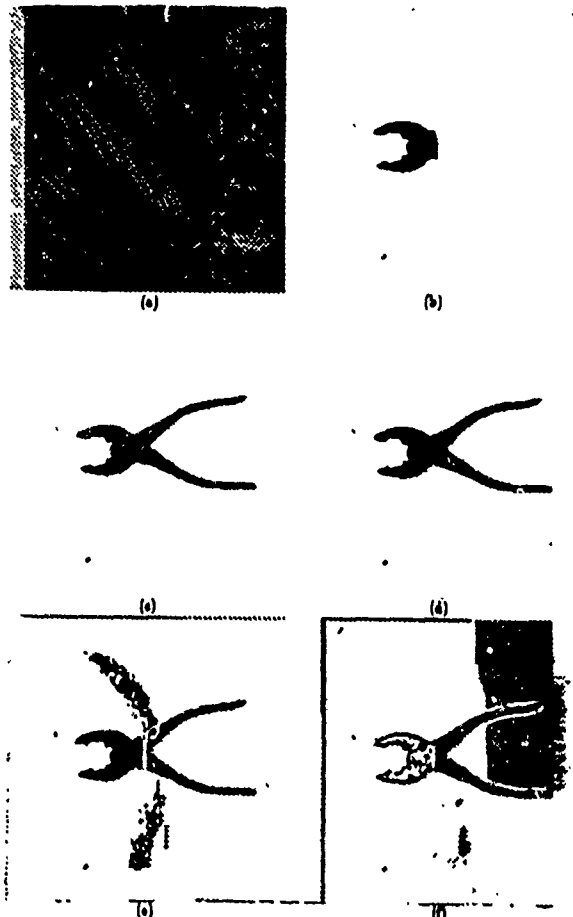


Fig. 7: (a) Tool image 2, 256×256 pixels, (b)-(f) Results of algorithms as in Fig. 3.

- [1] S. M. Dunn, D. Harwood, and L. S. Davis, "Local estimation of the uniform error threshold", *IEEE Trans. Pattern Anal. Machine Intell.*, vol. PAMI-6, no.6, November 1984.
- [2] R.M. Haralick, S.R. Sternberg, and X. Zhuang, "Image analysis using mathematical morphology", *IEEE Trans. Pattern Anal. Machine Intell.*, vol. PAMI-9, no.4, July 1987.
- [3] J. N. Kapur, P. K. Sahoo, and A. K. C. Wong, "A new method for gray-level picture thresholding using the entropy of the histogram", *Computer Vision, Graphics and Image Processing*, vol.29, no.3, pp.273-285, 1985.
- [5] G. Matheron, "Random sets and integral geometry", New York, Wiley, 1975.
- [6] T. Pavlidis and G. Wolberg "An algorithm for the segmentation of bilevel images", *Proc. CVPR-86*, Miami Beach, Florida, pp.570-575, June, 1986
- [7] A. Pérez and R. C. Gonzalez, "An iterative thresholding algorithm for image segmentation", *IEEE Trans. Pattern Anal. Mach. Intell.* vol. PAMI-9, no.6, pp.742-751, 1987.
- [8] T. Pun, "Entropic thresholding: A new approach", *Computer Graphics Image Processing*, vol.16, no.3, pp.210-239, 1981.
- [9] T. W. Ridler and S. Calvard, "Picture thresholding using an iterative selection method", *IEEE Trans. Systems Man Cybernet*, vol. SMC-8, no.8, pp.630-632, 1978.
- [10] J. Serra, "Image analysis and mathematical morphology", Academic Press, New York, 1982.
- [11] A. Shio, "An Automatic Thresholding Algorithm Based on an Illumination-Independent Contrast Measure", *Proc. CVPR-89*, pp.632-637, June 1989.
- [12] S. R. Sternberg, "Biomedical image processing", *Computer*, vol.16, no.1, pp.22-34, January 1983.
- [13] W. H. Tsai, "Moment-preserving thresholding: A new approach", *Computer Vision, Graphics and Image Processing*, vol.29, no.3, pp.377-393, 1985.
- [14] J. M. White and G. D. Rohrer, "Image thresholding for optical character recognition and other applications requiring character image extraction", *IBM J. Res. Develop.*, vol.27, no.4, pp.400-411, 1983.

APPLICATION OF SCATTERING THEORY TO PLASTIC STRAIN ESTIMATION

Y.Z. Dai and F.P. Chiang

Laboratory for Experimental Mechanics Research
State University of New York at Stony Brook
Stony Brook, NY 11794-2300

ABSTRACT

Light scattering theory and plastic deformation induced surface roughening were studied in an effort to provide some theoretical background to a nondestructive, non-contact plastic strain evaluation method. Experimental investigation shows that plastic strain on the surface of a metallic specimen is linearly proportional to surface root-mean-square roughness and inversely proportional to surface profile correlation length. Utilizing these relations along with Beckmann's light scattering theory, we obtained an expression which describes the scattered light intensity distribution in terms of plastic strain. This expression was applied to plastic strain evaluation and favorably verified by some experiments on aluminum alloy and brass materials.

INTRODUCTION

The phenomenon that plastic deformation changes metallic surface roughness as well as the scattered light intensity distribution has been investigated in the field of experimental mechanics for developing a plastic strain evaluating method.¹⁻³ However, to the authors' knowledge, all the related work in the published literature are experimental and there has been little theoretical analysis relating plastic strain with scattered light intensity distribution.

Obviously, surface roughness is the key factor relating plastic strain with scattering. Therefore the approach should be on two aspects: the plastic deformation induced surface roughening and the roughness caused light scattering.

In the present work, the relation between plastic strain and surface profile wavelength was experimentally studied; Then these relations were combined with Beckmann's light scattering theory⁴ such that the scattered light intensity distribution was expressed in terms of plastic strain. It is hoped that this paper will provide some better understanding on plastic deformation induced light scattering variation and its application to plastic strain evaluation.

PLASTIC STRAIN & SURFACE ROUGHNESS

Study of plastic deformation induced surface roughening was carried out on tensile specimens made of aluminum alloy and brass materials. Specimens were cut from stock sheets along rolling direction and then polished to an initial surface roughness of approximately $R_q = 0.05\mu m$, where R_q is root-mean-square (RMS) roughness. After being loaded on a Tinius Olsen 1000 testing machine to a certain amount of plastic deformation (measured by strain gage and moiré method) the specimen surface profile was recorded by a system as shown in Fig.1, where the tip radius of the stylus of the profilometer is $5\mu m$ and the surface profile was digitized at a frequency of 1000 points/mm. The surface profiles of aluminum 1100-00 at different amount of plastic deformation are shown in Fig.2. It is seen that the surface height variation increases while the profile wavelength decreases with plastic strain.

Plastic Strain & RMS Roughness

The zero-mean surface profiles were processed to yield the RMS roughness by the following expression:

$$R_q = \sqrt{\frac{1}{L} \sum_{i=1}^N f^2(i)} \quad (1)$$

where N is the number of data point; $f(i)$ is the discrete profile height at i th point; and L is the sampling length.

Five measurements of surface profile were taken at each plastic strain level. After the maximum and minimum values were eliminated, the rest data were averaged and plotted versus plastic strain as shown in Fig.3 in which the vertical bars represent the standard deviation of the data. The relation between plastic strain and RMS roughness may be expressed by a linear function as:

$$\sigma = \alpha \gamma \quad (2)$$

where α is a constant. This result confirms some early experimental work.^{5,6}

Plastic Strain & Profile Correlation Length

The profile correlation length is defined as the lag length T at which the autocorrelation of the profile is $1/e$ (0.368) of the zero-lag autocorrelation value. The value of T should satisfy the following condition:

$$\frac{\sum_{i=1}^{N-T} f(i) \times f(i+T)}{\sum_{i=1}^N f^2(i)} = \frac{1}{e} \quad (3)$$

The obtained profile correlation lengths were processed by the same method as described in the above session and the results are shown in Fig.4. The relation between plastic strain and surface profile may be approximated by:

$$T = \frac{\beta}{\gamma} \quad (4)$$

where β is a constant.

The constants α, β in the above equations can be determined by using the least square method to fit the experimental data. Tests carried out on brass tensile specimens show that $\alpha \approx 7\mu\text{m}$, and $\beta \approx 0.75\mu\text{m}$.

In general, Eq.2 and Eq.4 should take the form

$$\sigma = \alpha\gamma + c \quad \text{and} \quad T = \frac{\beta}{\gamma + a} + b \quad (5)$$

indicating that the surface roughness σ and profile correlation length T are finite at zero plastic strain due to the limitation of polishing process.

Plastic Strain & Substructure Size

Profile correlation length in general is proportional to profile wavelength, the magnitude of which is observed to be of the order of tens of microns. It might be worthwhile to point out that another plastic deformation induced surface substructure, which is of the order of less than ten microns, is related to plastic strain in a similar way as Eq.4.

According to Holt's derivation^[7], the dislocation substructure size d is related to dislocation density ρ by:

$$d \approx K_c \rho^{-1/2} \quad (6)$$

Applying linear regression to Chiem and Duffy's experimental data⁸ on shear strain γ and dislocation density ρ , we get:

$$\gamma = 1.85\sqrt{\rho} - 11.52 \quad (7)$$

Substitute Eq.7 into Eq.6 and write the resulting expression in a more general form as:

$$d \approx \frac{pK_c}{\gamma} + q \quad (8)$$

where p, q, K_c are constants. Eq.8 indicates that shear strain γ is inversely proportional to substructure size d which is an analog to Eq.4.

BECKMANN'S LIGHT SCATTERING THEORY

Beckmann⁴ discussed extensively the scattering of electromagnetic waves from rough surfaces. The solution describing the scattered light intensity distribution, also called power-spectral-density (PSD) for some circumstances, from a square area is:

$$\langle \rho\rho^* \rangle = e^{-g} \left(\rho_0^2 + \frac{\pi T^2 F}{A} \sum_{m=1}^{\infty} \frac{g^m}{m!m} e^{-g^2 T^2 / 4m} \right) \quad (9)$$

where the left hand side term is the ensemble average of light intensity $\rho\rho^*$ which was originally described as a ratio proportional to the mean scattered power; A is the area of the surface being illuminated; T is the profile correlation length and the other variables are defined as:

$$\sqrt{g} = k\sigma(\cos\theta_1 + \cos\theta_2) \quad (10)$$

$$\rho_o = \frac{\sin(v_x X) \sin(v_y Y)}{v_x X v_y Y} \quad (11)$$

$$F = \frac{1 + \cos \theta_1 \cos \theta_2 - \sin \theta_1 \sin \theta_2 \cos \theta_3}{\cos \theta_1 (\cos \theta_1 + \cos \theta_2)} \quad (12)$$

$$v_{xy} = k \sqrt{\sin^2 \theta_1 - 2 \sin \theta_1 \sin \theta_2 \cos \theta_3 + \sin^2 \theta_2} \quad (13)$$

where $k = 2\pi/\lambda$; λ is the wavelength of the light source and θ_1 is the light incident angle and θ_2, θ_3 are observation angles (see Ref.[4] for detail).

In practice, normal incidence ($\theta_1 = 0$) is often used which makes $F = 1$ and $v_{xy} = k \sin \theta_2$. Then Eq.9 can be rewritten as:

$$\langle \rho \rho^* \rangle = e^{-g} \left(\rho_o^2 + \frac{\pi T^2}{A} \sum_{m=1}^{\infty} \frac{g^m}{m!m} e^{-(kT \sin \theta_2)^2 / 4m} \right) \quad (14)$$

This expression can be extended to the case of a circular illuminated area with radius r , then the area A is πr^2 and

$$\rho_o = \frac{2J_1(Z)}{Z} \quad (15)$$

where $Z = kr\theta$ and $J_1(Z)$ is the first order Bessel function of the first kind. Utilizing Eq.2 and Eq.4, Eq.14 can be written in terms of plastic strain γ as:

$$\langle \rho \rho^* \rangle = e^{-g} \left(\rho_o^2 + \left(\frac{\beta}{r\gamma} \right)^2 \sum_{m=1}^{\infty} \frac{g^m}{m!m} e^{-(k\beta \sin \theta_2)^2 / 4m\gamma^2} \right) \quad (16)$$

where

$$\sqrt{g} = k\alpha\gamma(1 + \cos \theta_2) \quad (17)$$

Therefore, the PSD function has been expressed in terms of shear plastic strain. Once this distribution and the necessary constants are known, then the plastic strain can be evaluated by Eq.16; or, given a plastic strain value, its corresponding PSD distribution can be predicted (Fig.5).

APPLICATION TO PLASTIC STRAIN EVALUATION

In this part of the discussion, Eq.16 shall be used to predict the relation between PSD width and plastic strain followed by an experimental verification.

Analytical Approach

PSD width is defined as the width of the scattered light intensity distribution when its value is a certain fraction of its maximum. This width is approximately proportional to the scattered angle θ . In the following discussion, θ will be used to represent PSD width. Denote the intensity ratio as W , we have:

$$W = \frac{\langle \rho \rho^* \rangle_{\theta_2=\theta}}{\langle \rho \rho^* \rangle_{\theta_2=0}} = \frac{\rho_o^2 + \left(\frac{\beta}{r\gamma} \right)^2 \sum_{m=1}^{\infty} \frac{[k\alpha\gamma(1+\cos\theta)]^{2m}}{m!m} e^{(-k\beta \sin\theta)^2 / 4m\gamma^2}}{1 + \left(\frac{\beta}{r\gamma} \right)^2 \sum_{m=1}^{\infty} \frac{(2k\alpha\gamma)^{2m}}{m!m}} \quad (18)$$

The term $e^{-k\sigma(1+\cos\theta)}$ has been omitted because the error it introduces is less than 1.0% for $\theta \leq 4^\circ$, which is the maximum scattering angle observed in experiments. Fig.6

shows the PSD width at both $W = 1/2$ and $W = 1/e$ versus plastic strain. It can be seen that the theoretically predicted relation is approximately linear.

Experimental Verification

The linear relation between the width of the PSD distribution and plastic strain was verified by a system as described elsewhere.³ An experimentally obtained PSD distribution for brass material is shown in Fig.7.* As can be seen, it takes quite a similar form compared to the one (Fig.5) predicted by the light scattering theory. The scattering in experimentally obtained PSD is caused by speckle effects and is believed to vanish by taking an ensemble average on macroscopically similar but microscopically different surfaces undergone the same amount of plastic deformation.

Fig.8 shows some experimental results on the relation between plastic strain and PSD width ($1/e$). A linear relation, which is in support of the analytical result as shown in Fig.6, can be clearly seen.

DISCUSSION & CONCLUSION

It has been shown that metallic material surface roughens as the specimen undergoes plastic deformation in such a way that surface RMS roughness increases linearly while surface profile correlation length decreases inversely proportional to plastic strain for the material tested. The range of validity of this result is 10% plastic deformation even though a much wider range was reported.^{5,6} The initial surface roughness plays an important role in the determination of the constants in Eq.2 and Eq.4.

No explanation can be offered at this point to the similarity between Eq.4 and Eq.8. The knowledge of the internal relationship of these two equations may provide us with a better understanding of the plastic deformation induced metallic surface roughening mechanism.

The expression for describing PSD distribution as a function of plastic strain seemed to be correct. Comparing the analytical solution with the experimental data, we find a good agreement for PSD width method. However, we also see some discrepancy in the early stages of plastic deformation which may be improved by using a more general model like Eq.5. The accuracy of using this technique to estimate plastic strain can be improved by taking an ensemble average of PSD distributions.

Further experimental development and theoretical investigation on plastic deformation induced surface roughening mechanism may provide us with a practical non-contact, nondestructive method for evaluating plastic strain. The method has a potential to be used under severe environmental conditions such as high temperature and/or high strain rate.

ACKNOWLEDGMENT

Financial support provided by the Army Research Office, Engineering Service Division through contract No. DAAL0388K0083 is acknowledged.

REFERENCES

1. K. Miyagawa, A. Azushima, Measurement of plastic deformation by means of a laser beam, 9th World Conference on Nondestructive Testing, Australia, (1979)
2. A. Azushima, M. Miyagawa, Measurement of plastic zone around a fatigue crack of a stainless steel structure member with a laser beam, Proceedings of the Joint Conference on Experimental Mechanics Hawaii, (1982)
3. F.P. Chiang, Y.Z. Dai, B.Q. Xu, A. Kato, Study of surface roughening under different stress modes by correlation, to appear in Proceedings of SEM Fall Conference, Baltimore, (1990)
4. P. Beckmann, A. Spizzichino, "The scattering of electromagnetic waves from rough surfaces", Pergamon Press, (1963)
5. M. Fukuda, K. Yamaguchi, N. Takakura, Y. Sakano, Roughening phenomenon on free surface of products in sheet metal forming, J. Jap. Soc. Technol. Plast., (10), 994, (1974)
6. P.F. Thomson, B.V. Shafer, The roughening of free surfaces during plastic working, Int. J. Mach. Tool Des. Res., (22), No.4, 261, (1982)
7. D.L. Holt, Dislocation cell formation in metals, Journal of Applied Physics, (41) No.8, July, (1970)
8. C.J. Chiem, J.D. Duffy, Strain rate history effects and observations of dislocation substructure in aluminum single crystals following dynamic deformation, Brown University Report MRLE-137, Rhode Island, (1981)

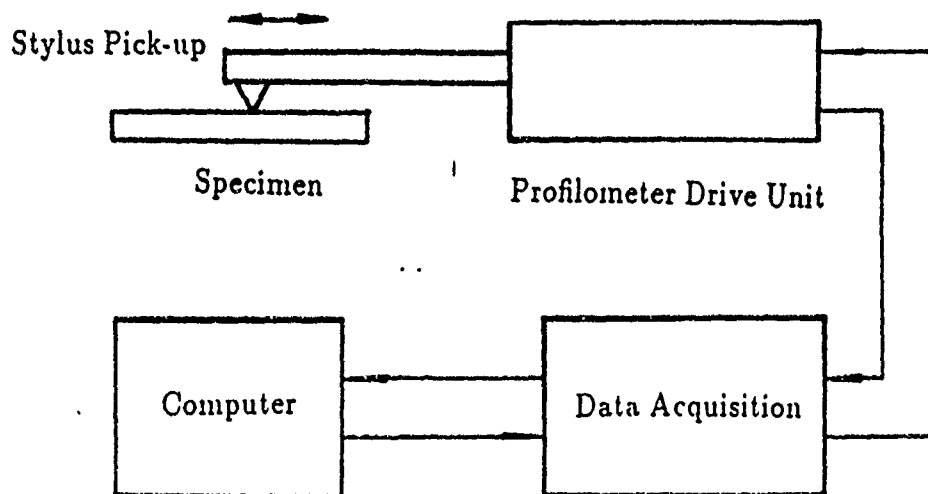


Fig.1 Surface profile measuring system

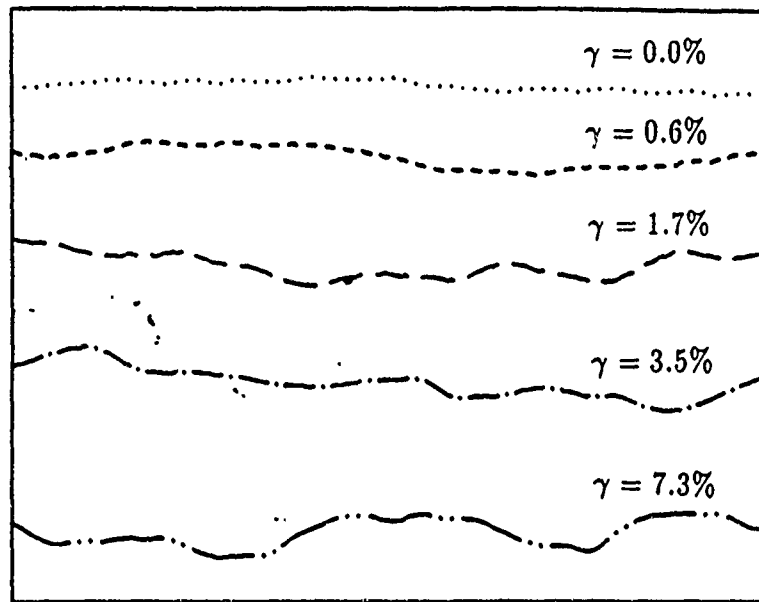


Fig.2 Surface profiles at different plastic strain levels

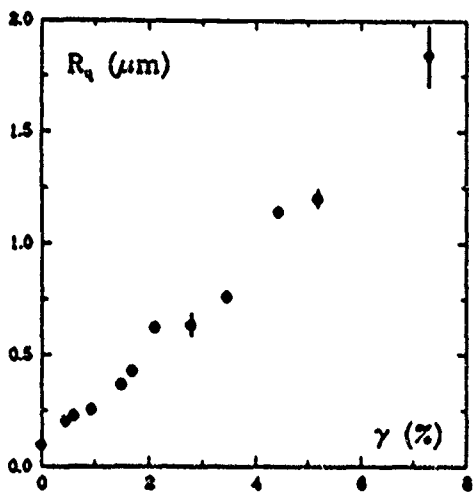


Fig.3 Plastic strain induced surface RMS roughness change

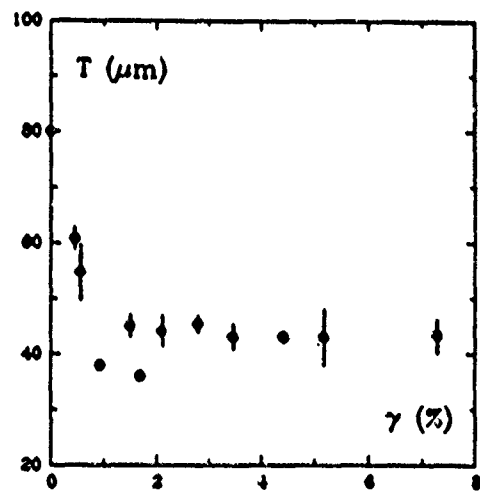


Fig.4 Plastic strain induced surface correlation length variation

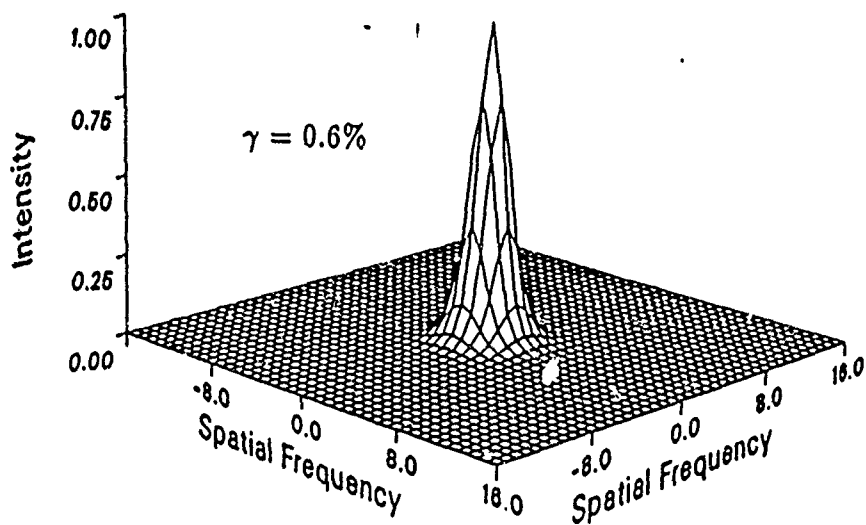


Fig.5 Analytical PSD distribution ($\gamma = 0.6\%$)

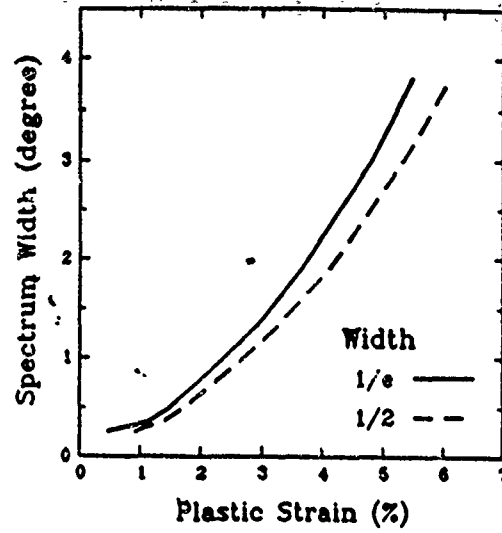


Fig.6 PSD width vs. strain (theoretical)

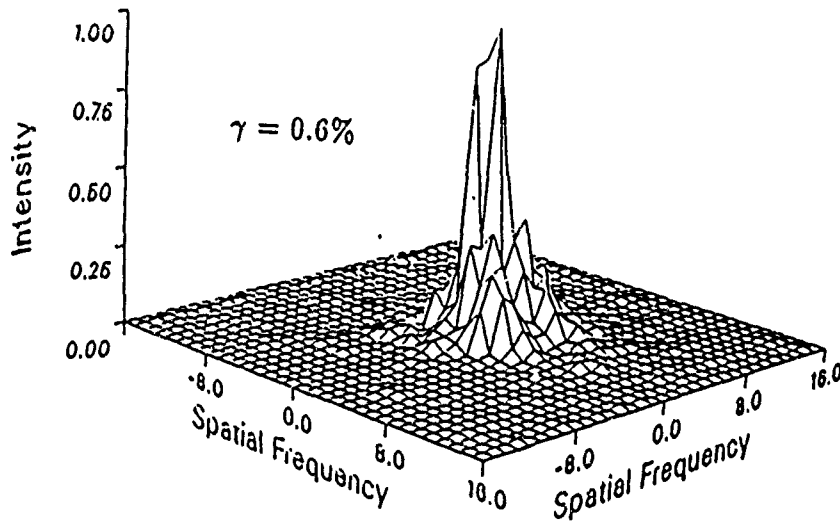


Fig.7 Experimental PSD distribution ($\gamma = 0.6\%$)

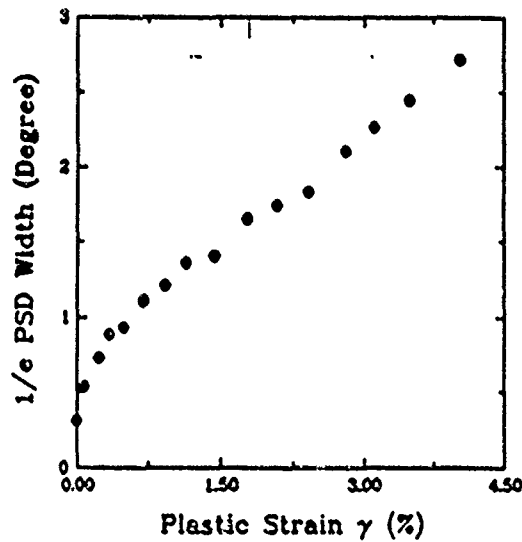


Fig.8 PSD width vs. strain (experimental)

DETERMINATION OF ELASTIC-PLASTIC BOUNDARY BY SPECKLE PATTERN CORRELATION

F.P. Chiang

C.J.Tay

Y.Z.Dai

Lab. for Experimental
Mechanics Research
SUNY at Stony Brook

Dept. of Mech. & Indus.
Nat. Univ. of Singapore
Singapore

Dept. of Mech. Eng.
SUNY at Stony Brook
NY 11794-2300

ABSTRACT

An optical-numerical correlation technique has been developed for detecting the elastic-plastic boundary on the surface of a material which has undergone plastic deformation. Laser speckle pattern produced from a surface is measured before and after plastic deformation which produces a change in their cross correlation. The speckle images are recorded and processed using a computer based vision system. Experimental results obtained for aluminum alloys have demonstrated the high sensitivity of this method. A comparison with theoretical and finite element results is presented.

INTRODUCTION

It has been known for sometime that the surface roughness of material increases with increase in surface strain ¹⁻³. This is mainly due to the process of dislocation movement. This phenomenon suggests methods to relate plastic strain quantitatively with variation in surface roughness. Surface roughness can be measured by a stylus-type profilometer. But this method tends to damage the material surface under examination. There are optical techniques to quantify surface plastic strain by measuring surface texture parameters. Some ^{4,5} employed coherent light source and obtained information about surface texture from the light beam reflected from the test surface. Others ^{6,7} used a white light source and a computer vision system to determine the plastic strain.

In this paper a nondestructive, non-contacting computer based method is described for the determination of plastic deformation. In particular the technique is applied to the determination of the elastic-plastic boundary around a central hole in a plate of finite width. The method employs a converging laser beam which detects the change in surface roughness introduced by plastic deformation. The diffracted images are recorded and digitized using a computer based vision system and a correlation technique is employed to analyze the digitized images.

THE SPECKLE PATTERN CORRELATION TECHNIQUE

The basic principles of the speckle pattern correlation techniques is based on the phenomenon that the surface of a piece of metallic specimen roughens as it is subjected to sufficient amount of plastic deformation. Consequently the resulting speckle pattern changes accordingly. If $g(x)$ is a function which describes the surface of a given area before deformation and $f(x)$ describes the surface after the deformation, the zero shift cross correlation $C_o(g, f)$, which is a measure of the resemblance between the two functions, is defined as:

$$C_o(g, f) = \int_{-\infty}^{\infty} g(x) \times f(x) dx \quad (1)$$

Normalization of C_o yields the cross correlation coefficient:

$$C_c(g, f) = \frac{\int_{-\infty}^{\infty} g(x) \times f(x) dx}{\sqrt{\int_{-\infty}^{\infty} g^2(x) dx \times \int_{-\infty}^{\infty} f^2(x) dx}} \quad (2)$$

For digitized images, $g(x)$ and $f(x)$ take discrete form and the above expression becomes ⁸:

$$C_c(g, f) = \frac{\sum_{i=1}^N g(i) \times f(i)}{\sqrt{\sum_{i=1}^N g^2(i) \times \sum_{i=1}^N f^2(i)}} \quad (3)$$

For two dimensional images the expression for cross correlation coefficient becomes:

$$C_c(g, f) = \frac{\sum_{i=1}^M \sum_{j=1}^N g(i, j) \times f(i, j)}{\sqrt{\sum_{i=1}^M \sum_{j=1}^N g^2(i, j) \times \sum_{i=1}^M \sum_{j=1}^N f^2(i, j)}} \quad (4)$$

where $g(i, j)$, $f(i, j)$ are the discrete intensity levels at points i, j of image g and f , respectively and M and N are the dimensions of the arrays $g(i, j)$ and $f(i, j)$.

EXPERIMENTAL PROCEDURES AND DATA ACQUISITION

Plate specimens of 3.2mm thick with a central hole were cut from aluminum alloy 6061-T6 in the rolling direction. The elastic modulus is 72GPa; the poisson's ratio is 0.32 and the yield strength is 275MPa. The geometry of the specimens are shown in Fig.1 where the $h = 76mm$, $a = 6.4mm$ and $R = 12.7mm$. One side of the specimen surface was polished until an initial roughness value of approximately $R_a = 0.05\mu m$ was obtained, where R_a is the arithmetic average roughness.

The specimen was loaded axially in a universal testing machine at $\sigma = 0.765\sigma_o$, where σ is the remote normal stress along the X axis, and σ_o is the yield strength of the material. The loaded specimen was then mounted on a X - Y translation stage which is capable of translation in each direction in 0.006mm increments.

The point of interest on the test object was illuminated with a 20mw He-Ne laser. To increase the resolution of the technique, the diameter of the laser beam was reduced through a system of optical arrangement consisting of a spatial filter, an aperture, a collimating lens and a converging lens as shown in Fig.2. The actual size of the laser beam directed on the test piece was estimated to be 0.05mm. A piece of ground glass was placed at a distance of 0.8m from the test object such that the diffraction pattern was observed in the Fraunhofer diffraction zone ⁹. The diffracted speckle pattern was recorded by a digital camera placed at 0.36m from the ground glass. The camera (resolution of 256 × 256 pixels, each pixel has 256 gray levels) was connected to a TV monitor and supported by a computer which together with the developed software provided full

image processing and analysis capabilities.

RESULTS AND DISCUSSION

Fig.3 shows the diffraction images of the test piece along the vertical axis (section *C-C*). Fig.3a is the diffraction image in the elastic zone. As the light source moves towards the central hole, the strain level increases and as illustrated in Fig.3b & 3c there is a continuous drop in the sharpness of the images and the image boundary becomes less distinguishable. The observation suggests that the correlation level of each subsequent image will decrease.

Fig.4 shows a plot of the cross correlation coefficient along section *C-C*. As can be seen, the initial images correlate well with that of the elastic zone at *B* and the correlation coefficient remains fairly constant over a distance of approximately 5.4 mm. However, beyond this point (as indicated by point *P*) the correlation coefficient begins to decrease and reaches a minimum value at the edge of the hole where the maximum plastic strain occurs. The decrease of the correlation coefficient (point *P*) would therefore provide an indication of the incipient yielding and hence the location of the elastic-plastic boundary.

The threshold value of cross correlation at which the elastic-plastic boundary is to be determined should be chosen according to the cross correlation-plastic strain calibration curve and the definition of elastic-plastic boundary for each specific problem. Suppose the elastic-plastic boundary is defined at where the effective residual strain is 1%; while from the cross correlation-plastic strain relation we find 1% effective strain corresponds to 0.7 cross correlation value, then 0.7 should be used as the threshold to determine the boundary.

Fig.5 shows the elastic-plastic boundary identified using the method described. The results shown are those obtained in region *A* of the specimen as indicated in Fig.5. The results obtained using the finite element method is also included. A numerical calculation which assumed a state of plane stress and adopted the von Mises yield criterion was carried out using a finite element package (ABAQUS version 4-7-21). The solid line represents a computed effective strain of 0.01%. As can be seen both the experimental and computed results agree reasonably well. An attempt was also made to compare the above results with those obtained using the Kirsch's solution for the stress distribution around a small circular hole in an infinite flat plate subject to uniform tension¹⁰ and the von Mises yield criterion. As can be seen the theoretical results (broken line) underestimated the size of plastic zone. The discrepancy is maybe due to the simplification and assumptions made in the theoretical solution.

CONCLUDING REMARKS

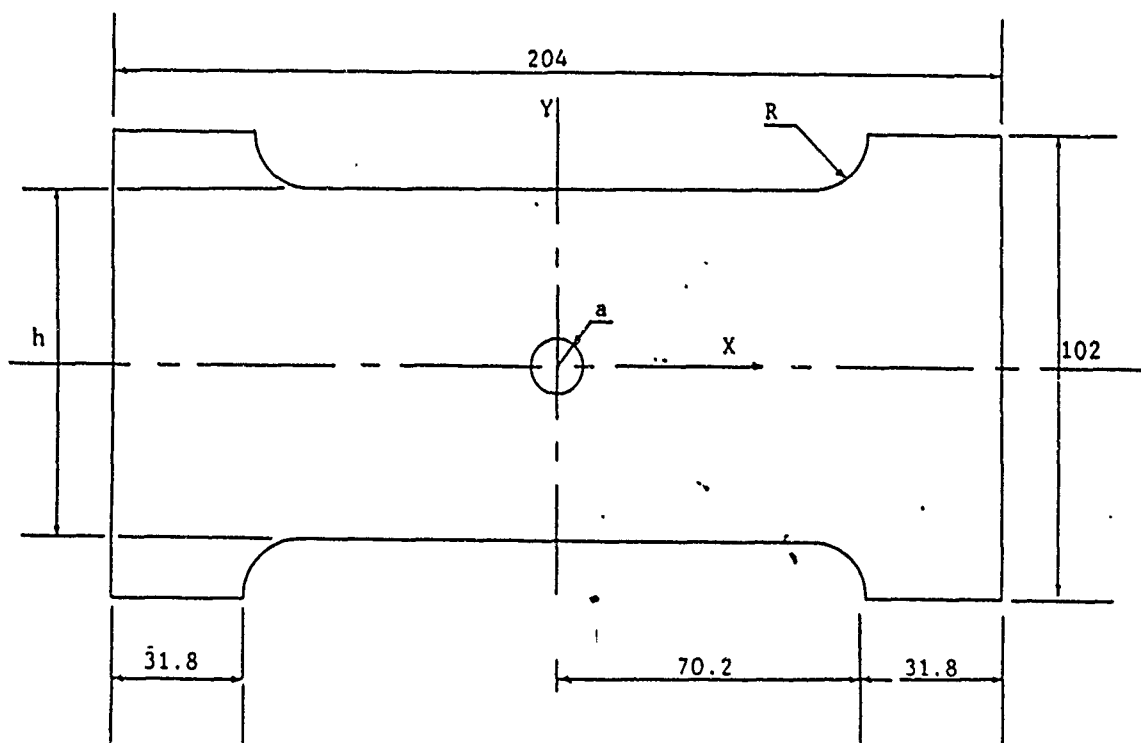
In this paper, it has been demonstrated that speckle pattern cross correlation technique can be applied to the detection of the incipience of plasticity. In a material which had undergone elastic-plastic deformation, the method described was able to detect plastic strain of the order of 0.01%. The technique was supported by a computer vision system via which the recorded images were analyzed and processed. With further automating the technique, it has the potential of being developed into a practical tool.

ACKNOWLEDGMENT

The authors wish to thank the Army Research Office, Engineering Service Division for providing the financial support through contract No. DAAL0388K0033.

REFERENCES

1. K. Yamaguchi, P.B. Mellor, Thickness and grain size dependence of limit strains in sheet metal stretching, Int. J. Mech. Sci., 18, pp.85-90, 1976
2. P.F. Thomson, B.V. Shafer, The roughening of free surface during plastic working, Int. J. of Mach. Tool Des. Res., 22,(4), pp.261-264, 1982
3. J.H. Dautzenberg, and J.A.G. Kals, Stress state and surface roughness, in: "Advances Technology of Plasticity", Japan, 1984
4. M. Miyagawa, A. Azushima, Measurement of plastic deformation by means of a laser beam, in: "Proceedings of the 9th World Conference on Nondestructive Testing", Session 4a-4, pp.1-7, Melbourne, Australia, Nov.,1979
5. D. Sigler, W.L. Haworth, Strain measurement by optical correlation, Journal of Nondestructive Evaluation, 2,(2), pp.125-132, 1981
6. C. Lee, Y.J. Chao, W.H. Peters, M.A. Sutton, Surface texture and plastic deformation, in: "Proceedings of the 13th Southeastern Conference on Theoretical and Applied Mechanics", pp.275-283, Columbia, South Carolina, 1986
7. C. Lee, Y.J. Chao, M.A. Sutton, et al., Determination of plastic strain at notches by image-processing methods, Experimental Mechanics, 29,(2), pp.214-220, 1989
8. A. Rosenfeld, C.A. Kak, "Digital Picture Processing", Academic Press, pp.19, 1982
9. M. Francon, "Laser Speckle and Related Applications in Optics", Academic Press, pp.21-24, 1979
10. D.J.P. Hartog, "Advanced Strength of Materials", Mc-Graw Hill Book Co., 1952



All dimensions in mm

Fig.1 Dimensions and geometry of the specimen

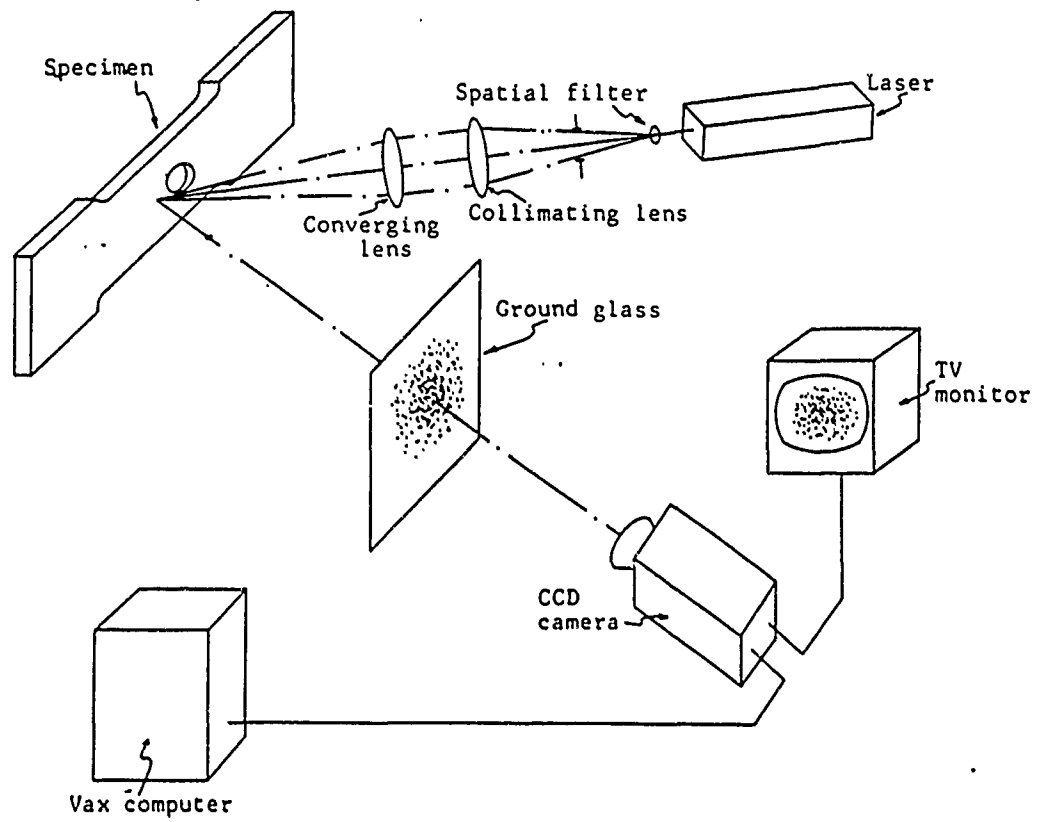


Fig.2 Optical-numerical set-up



Fig.3a Diffraction image of point *B*

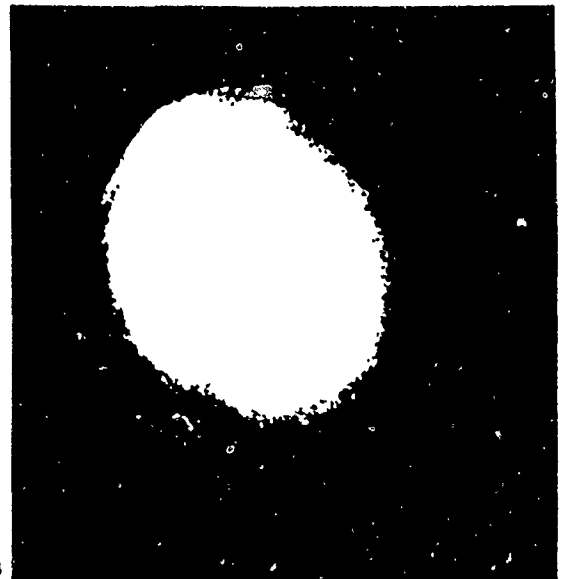


Fig.3b Diffraction image of point *P*



Fig.3c Diffraction image of point *Q*

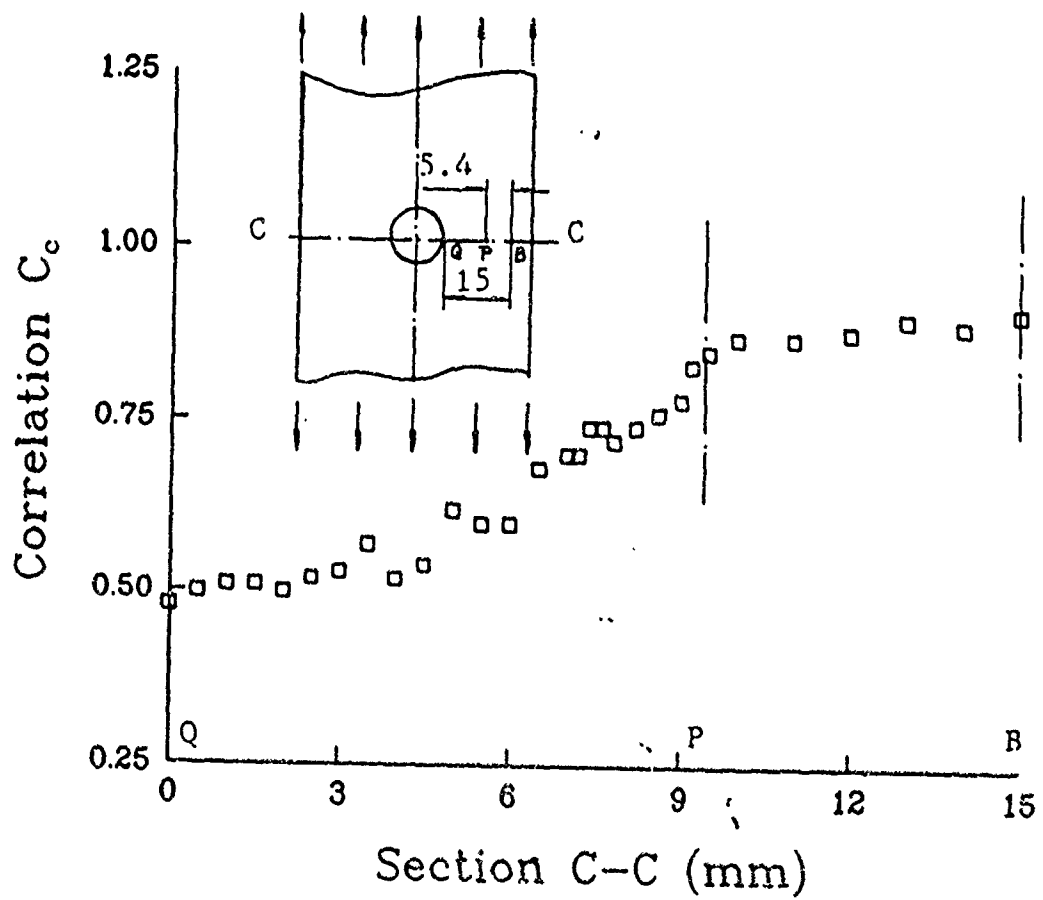


Fig.4 Correlation coefficient across the plastic zone (along section *C-C*)

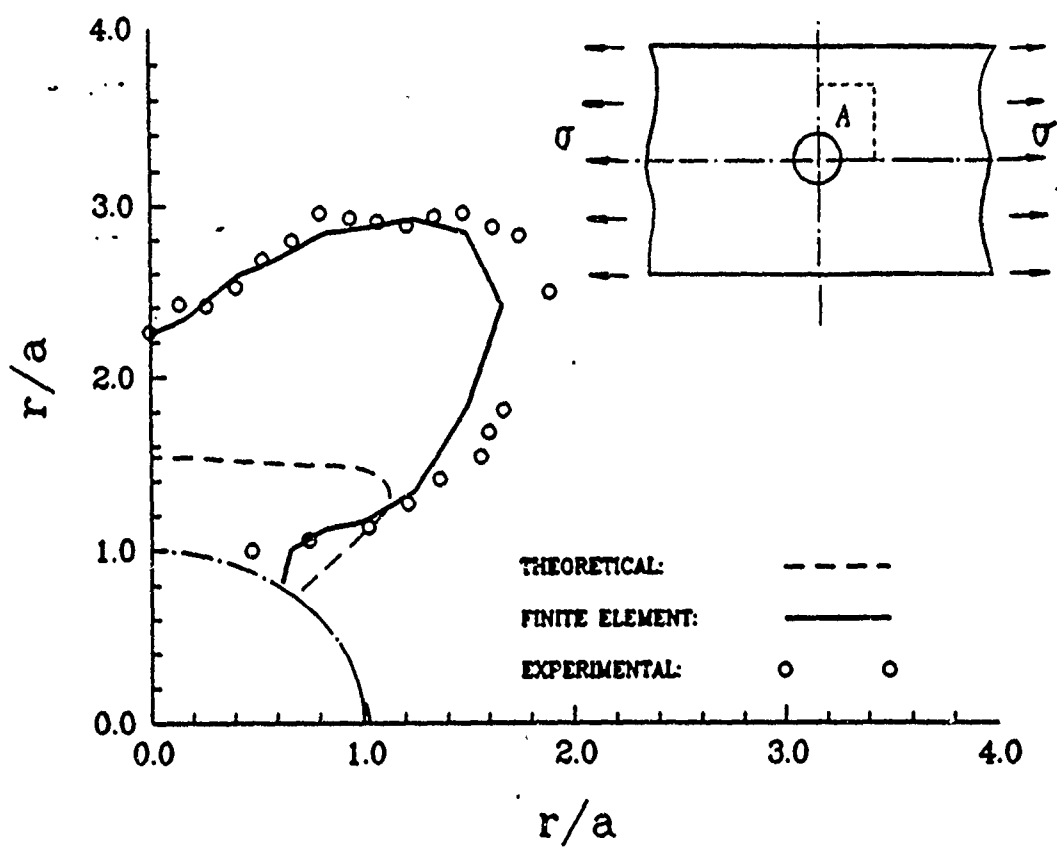


Fig.5 Elastic-plastic boundary at $\sigma = 0.765\sigma_0$.

Pattern Recognition Using 3-D Moments

Chong-Huah Lo and Hon-Son Don

Department of Electrical Engineering
State University of New York
Stony Brook, New York 11794

Abstract

A 3-D moment method of object identification and positioning is proposed. Moments are computed from 3-D CAT image functions, 2.5-D range data, space curves and discrete 3-D points. Objects are recognized by their shapes via moment invariants. Using an algebraic method, scalars and vectors are extracted from compound of moments using Clebsch-Gordon expansion. The vectors are used to estimate position parameters of the object. Moment features of range data can be used in the view-independent object recognition when the 3-layer perceptron encodes the feature space distribution of the object in the weights of the network. Objects are recognized from an arbitrary viewpoint by the trained network.

1. Introduction

Moments have been used in computer vision for 30 years. Moment invariant features of images are useful in pattern recognition. Various methods of computing 2-D moment invariants have been proposed [5][12]. The evaluation of moment integrals requires a large amount of computation. However, the computation complexity will not hinder the use of the moment method in the real time computer vision system. The moment integrals can be parallelly computed. Moreover, VLSI chips [1] and hardware [13] have been designed to compute 2-D moments in real time.

In this paper, we present formulae for computing moments of a 3-D grey-level function, of a range image, of a 3-D curve, and of a set of discrete feature points. Range images are commonly used in robotic vision applications. We shall show that the 3-D moment M_{lmn} of an object in the range image can be considered the 2-D moment M_{lm} of an image function which depends on the range image and the derivatives of that range image. Hardware similar to those in [1][13] may be built so that 3-D moment feature extraction from the range data can be done in real time. The invariant and vector moment functions can be constructed using our proposed method [7]. Our method is based on the Clebsch-Gordon expansion in group representation theory [6][7]. Moments are first expressed in the basis

of spherical harmonic polynomials. They are called complex moments. Vectors and scalars are extracted from the compounds of complex moments via Clebsch-Gordon expansion. Higher order moment invariants can be derived in this way. They represent the fine spatial details on the objects. Feature vectors containing higher order moment invariants have high discriminative power in pattern classification. The vector moment functions can also be used in the algorithms of position estimation [2][9].

The range image is a multiview representation of 3-D objects. The moment invariants of an object in the range image only represent a particular view of the object. Even so, the set of invariant features still provides an elegant description of an object in each viewpoint. A convex polyhedron can be represented by a number of Gaussian peaks in the moment invariant feature space. Each peak represents a characteristic view of the convex polyhedron. A smooth convex object can always be approximated by a convex polyhedron, so its feature space distribution can be approximated as a superposition of Gaussian distributions. As for non-convex objects, one part of the object may be occluded by another part of the object. Therefore, non-convex objects generally have rather complicated feature space distributions. Patterns with multimodal non-Gaussian distributions may form arbitrarily complex disconnected decision regions. The recently proposed 3-layer perceptron network has the capability of classifying patterns with arbitrarily complex decision regions [8][10]. This network can be trained by the back-propagation algorithm [10]. We have conducted experiments on both convex and non-convex curved objects. The 3-D moment features of objects from all viewpoints were extracted. These features were used as the training samples for the inputs of a 3-layer perceptron network. Our experimental results show that objects coded by moment invariant features can always be satisfactorily classified by the 3-layer perceptron network.

The paper is organized as follows. In Section 2, we present the computation of 3-D geometric moments of 3-D grey-level functions, range-data, space curves and discrete feature points. In Section 3, we present the derivation of invariant and vector moment functions using Clebsch-Gordon expansion. The 3-D moment method for pattern recognition and position estimation is given in Section 4. The experimental results are given in Section 5. Section 6 concludes the paper with a summary of results and a few remarks.

The ten third-order moments are decomposed into two complex moment multiplets. Seven of the third-order complex moments, denoted by $[v_3^3, v_3^2, \dots, v_3^2, v_3^3]$, corresponding to $L=3$. Another three third-order complex moments, denoted by $[v_1^3, v_1^2, v_1^3]$ corresponding to $L=1$. v_3^m and v_1^m are linear combinations of the third order moments. These expressions are given in [7].

Three moment invariants containing the second-order moments are given below.

$$v\theta = \frac{2}{3}\sqrt{\pi}(M_{200} + M_{020} + M_{002}). \quad (7.a)$$

$$v(2,2)\theta = (5)^{-\frac{1}{2}}(2v_2^2v_2^2 - 2v_2^1v_2^1 + (v\theta)^2). \quad (7.b)$$

$$\eta v = (5)^{-\frac{1}{2}}(\eta_2^2v_2^2 - \eta_2^1v_2^1 + \eta_2^0v_2^0 - \eta_2^{-1}v_2^{-1} + \eta_2^{-2}v_2^{-2}), \quad (7.c)$$

where η is a second rank spherical tensor derived from the products of v_2^m in (6). $\eta_2^m = \sum_{i=2}^{i=2} \langle 2, i, 2, m-i | 2, 2, 2, m \rangle v_2^i v_2^{m-i}$. ηv is a cubic polynomial of second order moments. In [7], we have shown that our results in (7) agree with those obtained by Sadjadi and Hall [11]. Two quadratic moments invariants containing the third-order moments are given below

$$v(3,3)\theta = (7)^{-\frac{1}{2}}(2v_3^3v_3^3 - 2v_3^2v_3^2 + 2v_3^1v_3^1 - (v\theta)^2), \quad (8.a)$$

$$v(1,1)\theta = (3)^{-\frac{1}{2}}(2v_1^3v_1^3 - (v\theta)^2). \quad (8.b)$$

A complete set of moment invariants containing the second and the third order moments have been derived. Three vector moment functions which are used in motion estimation have also been computed. Explicit expressions of these moment functions can be found in [7]. The advantages of our approach is that (1) It is a general and systematic way of deriving higher order moment invariants, (2) sufficient number of moment vector functions for motion estimation can be derived from the second and the third order moments. Vectors constructed from higher order moments are too expensive to compute and very vulnerable to noise.

When each invariant moment feature is divided by suitable power of $v\theta$, it becomes invariant under changing of size. Those similitude-invariant quantities derived from the moment invariants in (7) and (8) are given below

$$I_2 = \frac{v(2,2)\theta}{(v\theta)^2}, \quad I_{22} = \frac{\eta v}{(v\theta)^3},$$

$$I_3 = \frac{v(3,3)\theta}{(v\theta)^3}, \quad I_{11} = \frac{v(1,1)\theta}{(v\theta)^3}.$$

Others derived from higher order moment invariants are given in [7].

4. Object Identification and Position Estimation

4.1. Object Recognition using Moment Invariants

The moment invariants of the range image can be used in the view-independent 3-D object recognition. Each view of a convex polyhedron can be represented by a Gaussian peak in its moment feature space. The feature space distribution of a convex polyhedron is a superposition of Gaussian distributions centered at a number of positions. The distribution is the superposition of Gaussian peaks.

$$P(x_1, \dots, x_n) = \sum_{j=1}^N P_j(x_1, \dots, x_n) P_j \quad (9)$$

where $P_j(x_1, \dots, x_n)$ is the Gaussian distribution whose parameters are the mean and variance of the features. P_j is the a priori probability of the j th prototype view of the object and $\sum_{j=1}^N P_j = 1$.

A curved convex object can be approximated by a convex polyhedron to any precision. Thus its feature space distribution can also be approximated by (9). The non-convex objects are self-occluded. In different viewpoint, the self-occlusion occurs in different degrees. The feature space distributions of non-convex curved objects are generally more complicated than those in (9). The Gaussian classifiers can not classify patterns whose feature space distributions are non-Gaussian. Nonparametric techniques had been suggested to construct statistical pattern classifier for Non-Gaussian data [3]. However, such techniques are severely restricted by the dimensionality of the feature space.

The recently proposed multi-layer feed forward neural network has the capability of classifying patterns with non-Gaussian feature space distributions. A three-layer perceptron, shown in Fig.1, can classify patterns of arbitrarily complex decision regions [8][10]. The learning rule for perceptron using sigmoidal threshold function is called back-propagation training algorithm [10]. The algorithm uses the iterative gradient descend method to find the weights in the network. The rules can be summarized in three equations.

$$w_{ji}(n+1) = w_{ji}(n) + \eta \delta_j O_i + \alpha(w_{ji}(n) - w_{ji}(n-1)) \quad (10)$$

where w_{ji} is the weight and η is the step size. The last term of the above equation is a momentum term which will improve the convergence of the training process. The error signals at the output nodes are given by

$$\delta_j = (t_j - O_j) f_j'(net_j) \quad (11)$$

where t_j is the desired signal at the output node j . The error signals at nodes in the hidden layer and the input layer are given by

$$\delta_j = f_j'(net_j) \sum_{\text{node } k \text{ in the upper layer}} \delta_k w_{kj} \quad (12)$$

The inputs to the system are the moment invariant features extracted from the range data of the object. The output signals of the system are associated with the pattern classes of the prototypes. Our experimental study indicates that the three layer perceptron can successfully classify patterns with non-Gaussian distributions when these distribution profiles do not overlap.

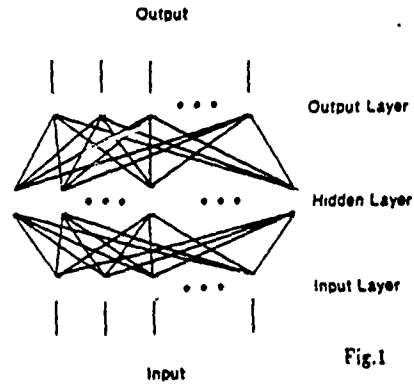
4.2. Position Estimation using Vector Moment Functions

Moments have also been used to estimate the affine transform parameters of 2-D and 3-D images. Cyganski and Orr developed a tensor-based technique to recover the affine transform of a 2-D image resulting from the orthogonal projection of a rigid planar patch [2]. Their method has been extended to affine transform determination for 3-D objects by Faber and Stokely [4]. In the tensor approach of estimating rigid body motion parameters, the rotation parameters is determined by solving a system of linear equations containing three vector moment functions. The translation parameters can then be determined from the transformation property of the first order geometric moments.

The 3-D orientation can be estimated from the principal axes of feature points in two frames. The basic assumption is that the 3-D coordinates of a set of feature points on a rigid body are known at

International Conference on Pattern recognition, pp. 303-305, 1986.

- [10] D. Rumelhart and J. McClelland *Parallel Distributed Processing* Vol. 1 M.I.T Press, Cambridge, MA, 1986.
- [11] F. A. Sadjadi and E. L. Hall, "Three dimensional moment invariants," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. PAMI-2, pp. 127-136, Mar. 1980.
- [12] M. R. Teague, "Image analysis via the general theory of moments," *J. Opt. Soc. Amer.*, vol. 70, pp. 920-930, Aug. 1980.
- [13] M.F. Zakaria, L.F. Vroomen, P.J.A. Zsombor-Murray and J.M.H.M. Van Kessel. "Fast algorithm for the computation of moment invariants", *Pattern Recognition* vol. 20, no. 6. pp 639-643, 1987.



invariant	Fig.2a	Fig.2b	Fig.2c	Fig.2d
$I_{1,0}$	0.150	0.150	0.150	0.150
$I_{1,1}$	-0.561	-0.561	-0.561	-0.560
$I_{1,2}$	-0.033	-7.700	-6.940	-1.036
$I_{1,3}$	-66.900	-66.540	-51.731	-0.037
$I_{2,0}$	2.070	2.271	2.129	0.304
$I_{2,1}$	-1.262	-1.347	-1.411	-0.218
$I_{2,2}$	3.301	3.490	3.090	0.653
$I_{2,3}$	27.099	33.537	26.743	0.627
$I_{2,4}$	-18.070	-19.722	-18.650	-0.443
$I_{2,5}$	10.043	12.302	12.560	0.331
$I_{2,6}$	-27.400	-28.000	-34.947	-0.041

invariant	Fig.3a	Fig.3b	Fig.3c	Fig.3d
$I_{1,0}$	0.000	0.000	0.004	0.007
$I_{1,1}$	-0.017	-0.010	-0.010	0.009
$I_{1,2}$	-0.131	-0.130	-0.130	-0.250
$I_{1,3}$	-0.010	-0.000	-0.030	-0.100
$I_{2,0}$	0.017	0.017	0.017	0.070
$I_{2,1}$	-0.010	-0.010	-0.010	-0.000
$I_{2,2}$	0.000	0.000	0.000	-0.000
$I_{2,3}$	0.010	0.010	0.010	0.000
$I_{2,4}$	-0.000	-0.004	-0.004	-0.000
$I_{2,5}$	0.004	0.004	0.004	0.000
$I_{2,6}$	-0.000	-0.000	-0.000	0.001

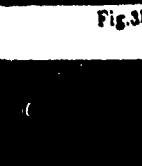
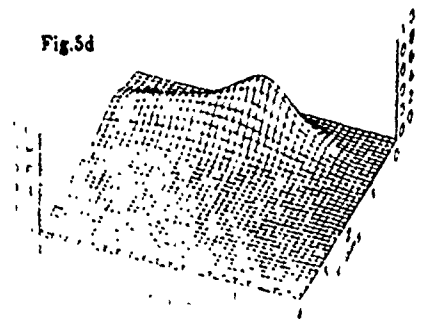
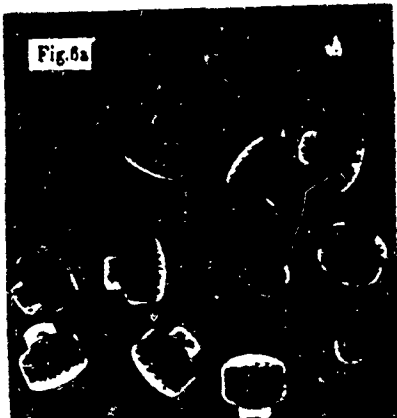
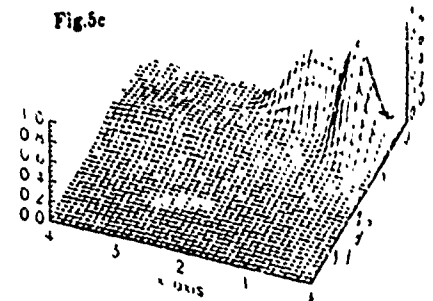
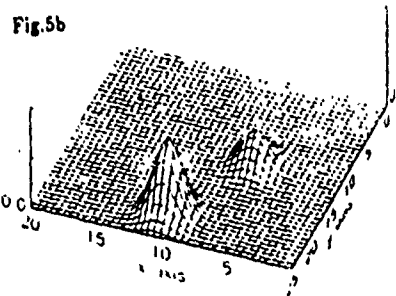
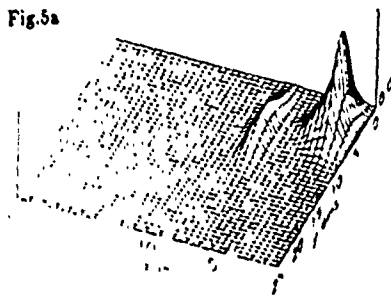
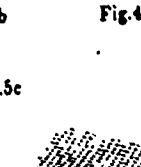
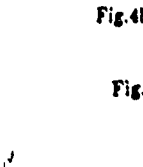
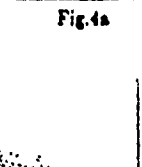


Table I Moment invariants evaluated using the curves in Fig.4. $I_{2,5}^{2,5}$, ... are defined in [7]

Table II Moment invariants of the surface data on the range image in Fig.2b and Fig.3. $I_{2,5}^{2,5}$, ... are defined in [7].



Estimation of plastic strain by fractal

Y.Z.Dai and F.P.Chiang

Laboratory for Experimental Mechanics Research
SUNY at Stony Brook, NY 11794-2300

ABSTRACT

An attempt was made to explore the possibility of utilizing fractal dimension to estimate surface roughness and plastic strain of metallic specimens. The roughness and strain related characteristics of specimen surface were revealed by laser speckle technique and related to fractal dimension in this paper. Some preliminary experimental results show that fractal could be used for the evaluation of plastic strain and surface roughness.

1. INTRODUCTION

Permanent strain in metals generally changes surface features of the specimen. This phenomenon has been used for evaluating plastic strain^{1,2,3,4,5}. The present work is designed to extend the investigation on surface roughness and plastic strain measurement by introducing the fractal dimension analysis method.

Fractal, which abstracts some essential features of an object and characterizes it by a simple parameter, has a history of only one decade and it has already generated wide interests in many scientific fields. In this study, fractal dimension is applied to abstract the surface roughness and plastic strain related information in the power spectrum of specimen surface by laser speckle technique. The power spectrum, which is the diffraction speckle pattern produced by an optical configuration as shown in Fig.1, of a specimen surface reflects the change in surface textures caused by plastic deformation. It is hoped that fractal dimension can be used as a measure to detect this change and hence to estimate surface roughness and plastic strain.

2. FRACTAL DIMENSION

Fractal is defined as a 'set for which the Hausdorff-Besicovitch dimension strictly exceeds the topological dimension'⁶ or 'a shape made of parts similar to the whole in some way'⁷. Fractal, in a sense, represents some essential characteristics of a somewhat random curve or a surface.

The essential part of fractal analysis is to define a measure, which could be length for curves, area for a surfaces or volume for three dimensional objects, and its relationship with a scale with which the measurement is taken. Suppose we measure the length of a curve by covering the curve with the minimum number of squares, obviously the measured length of the curve depends on the size of the square and the smaller the square size, the longer the measured curve length. If the

length of the measured curve length is denoted as L ; the side length of the square is δ and the minimum number of squares needed to cover the curve is $N(\delta)$, then the measured curve length is:

$$L = N(\delta) \cdot \delta \quad (1)$$

If we measure the curve length with different sized squares, a series of curve length will be obtained. Experiments have shown that the measured length is nicely approximated by the formula:

$$L(\delta) = a \cdot \delta^{1-D} \quad (2)$$

where D is the so called fractal dimension for curves.

For example, the fractal dimension of the coast of Britain is $D \approx 1.3$; while the fractal dimension of the coast of southern part Norway is $D \approx 1.52$. These two coast lines could then be identified by their fractal dimensions.

Fractal has found wide applications in describing the nature such as landscape, coast line, clouds and so on. It can also be used to generate objects which look just like natural ones. A speckle pattern is somewhat random and may be described by fractal as well.

3. POWER SPECTRUM AND FRACTAL

When an optically rough plate (x, y plane) is illuminated by a narrow laser beam as shown in Fig.1, the power spectrum of the complex field immediately in front of the surface being illuminated would be formed in the Fourier transform plane (u, v). If a piece of film or a video camera is put in this plane, only the intensity, i.e. the power spectrum, will be recorded. Fig.2 shows a power spectrum digitized and recorded via a video camera and Fig.3 is a plot of the power spectrum along the central cross section for specimens which have undergone different amount of plastic deformation.

Fractal dimension may be evaluated in different ways depending on what kind of measure-scale pair is used. Based on the fact that landscapes have the highest amplitudes at the lowest frequencies, a Fourier transform method was introduced⁸. The Fourier amplitude C_f of a landscape, which is the specimen surface illuminated by laser for our study, is related to its corresponding spatial frequency f by⁸:

$$C_f \propto f^{-\beta} \quad (3)$$

The parameter β is related to fractal dimension D simply by⁸:

$$D = 2 - \beta \quad (4)$$

This means that a Fourier transform on specimen surface features is needed and this is exactly what has been done to the specimen surface by the optical configuration shown in Fig.1. The complex amplitude $A(u, v)$ at the u, v plane, assuming it is sufficient far away, is the Fourier transform of the amplitude $O(x, y)$ immediately in front of the object plane (x, y)⁹. i.e.

$$A(u, v) = \frac{C(u, v, L)}{j\lambda L} \int \int_{-\infty}^{\infty} O(x, y) e^{-j\frac{2\pi}{\lambda L}(ux+vy)} dx dy \quad (5)$$

where λ is wavelength; L is the distance between the observation plane and object surface; and $C(u, v, L)$ is a complex value independent of x, y coordinates. The coordinates u, v in the observation plane are related to spatial frequencies f_x, f_y by:

$$\begin{aligned} f_x &= u/\lambda L \\ f_y &= v/\lambda L \end{aligned} \quad (6)$$

What can be easily recorded is the intensity, which is what we record on either photographic film or video camera at u, v plane. This recorded intensity is the power spectrum of the surface. i.e.

$$I(u, v) = A(u, v) * A^*(u, v) \quad (7)$$

where $A^*(u, v)$ is the complex conjugate of $A(u, v)$.

The light intensity on (u, v) plane can then be expressed as a function of f_x, f_y :

$$I(u, v) = F(f_x, f_y) \quad (8)$$

This two dimensional equation can be further simplified to one dimensional by averaging the intensity $I(u, v)$ along circles ($f = \sqrt{f_x^2 + f_y^2}$) from the mirror reflection point which is called the centroid of the speckle pattern. This averaging yields spatially averaged intensity G_f and is based on the assumption that the surface roughness is isotropic.

The complex magnitude C_f is replaced by intensity G_f . The G_f, f relation can be easily obtained by digitizing the power spectrum, processing it and then using Eq.3 to fit the thus obtained distribution curve by the least square method. The proportional constant in Eq.3 is irrelevant indicating that the fractal parameter is independent of the intensity of the light source used.

4. PLASTIC STRAIN EVALUATION

Fractal analysis was carried out on uni-axial tensile specimens made of aluminum 1100-0 and 6061-T6. The specimen surface was polished such that the surface scratches produced during the manufacturing process were removed and the surface profile looked approximately isotropic with an initial surface roughness $R_a \approx 0.1\mu m$.

Surface roughness was measured by SurfTest-402 (Mitutoyo) profilometer with a tip radius of $5\mu m$. The spectrum was digitized by a C1000 Hamamatsu video camera whose resolution was set at 256×256 pixels each with 256 gray levels. Plastic strain was measured by a moiré method. A line grating with a density of 300 lpi was printed on one side of specimen surface while the other side was reserved for surface roughness and speckle pattern measurement. Specimen was loaded on Instron 1332 testing machine to cause a certain amount of plastic deformation then released from it for surface roughness measurement, strain evaluation and laser speckle pattern recording and

processing.

The centroid of the speckle pattern, where the spatial frequency is zero, was found by the moment method. i.e.

$$X_c = \frac{\sum_{i=1}^{256} \sum_{j=1}^{256} g_{i,j} \times i}{\sum_{i=1}^{256} \sum_{j=1}^{256} g_{i,j}} \quad (9)$$

and

$$Y_c = \frac{\sum_{i=1}^{256} \sum_{j=1}^{256} g_{i,j} \times j}{\sum_{i=1}^{256} \sum_{j=1}^{256} g_{i,j}} \quad (10)$$

where $g_{i,j}$ is gray level at coordinates i, j and X_c, Y_c are the coordinates of centroid.

The image was then re-sorted in such a way that gray levels were averaged along circles radiating from the centroid. The gray levels along each individual circle were averaged to yield G_f :

$$G_f(k) = \frac{\sum_{r=k}^N g_{i,j}}{N} \quad (11)$$

where k represents the distance from the centroid and is proportional to spatial frequency. The gray level $g_{i,j}$ was selected such that i, j satisfies:

$$k < \sqrt{i^2 + j^2} < k + 1 \quad (12)$$

Fig.4 is a series of plots of Fourier amplitude intensity versus spatial frequency for different amount of plastic strain. They indeed take the form as described by Eq.3. As the plastic strain increases, the spectrum distribution becomes flatter resulting in a decrease in β and an increase in fractal dimension D (Eq.4).

Fig.5 is a plot of the unified fractal value versus surface roughness of an aluminum 1100-0 specimen. Fig.6 shows the unified fractal value versus plastic strain of an aluminum 6061-T6 specimen. Fractal dimension D increases fairly rapidly with respect to the increase of surface roughness R_a in the beginning and becomes almost a constant after R_a becomes larger than $1\mu m$ for aluminum 1100-0 specimen. The curve of fractal dimension D versus plastic strain ϵ_p is pretty much the same except that it does not increase as rapidly as the previous one in the beginning and it saturates at a plastic strain of about 80% of the necking plastic strain for an aluminum 6061-T6 specimen.

5. SUMMARY

It has been shown that the power spectrum of specimen surface in the form of laser speckle pattern can be characterized by fractal dimension which reflects the change in surface texture due to plastic strain. Fractal dimension method could be used to evaluate surface roughness and plastic strain. Among the many measure-scale pairs we tried for this study, the intensity-frequency pair is the best as far as relating fractal to plastic strain or surface roughness is concerned.

6. ACKNOWLEDGMENTS

We would like to thank the Army Research Office, Engineering Service Division for providing the financial support to this work through grant No. DAAL0388K0033.

7. REFERENCES

1. Miyagawa, K., Azushima, A., 'Measurement of Plastic Deformation by Means of a Laser Beam', *Proceedings of the 9th World Conference on Non-destructive Testing*, pp. 1-7, Australia, 1979
2. Azushima, A. Miyagawa, M., 'Measurement of Plastic Zone around a Fatigue Crack of a Stainless Steel Structure Member with a Laser Beam', *Proceedings of the 1982 Joint Conference on Experimental Mechanics*, pp. 709-713, Hawaii, 1982
3. Sigler, D., Haworth, W.L., 'Strain Measurement by Optical Correlation', *Journal of Nondestructive Evaluation*, Vol. 2, No.2, pp. 125-132, 1981
4. Wu, X.P., Qian, H.H., Chiang, F.P., 'Study of Fatigue Damage and Crack Initiation by Laser Speckle Using a Digital Camera', *Proceedings SEM 1986 Spring Conference*, pp. 781, New Orleans, 1986
5. Lee, C., Peters, W.H., Sutton, M.A., Chao, Y.J., 'A study of Plastic Zone Formation by Digital Image Processing', *Proceedings of the SECTAM XIII Conference*, pp. 275-283, Columbia, South Carolina, 1986
6. Mandelbrot, Benoit B., 'The Fractal Geometry of Nature', pp. 15, W.H. Freeman, New York, 1982
7. Mandelbrot, Benoit B., 'Private Communication', 1987
8. Feder, J., 'Fractals', pp. 217, Plenum Press, New York, 1988
9. Goodman, J.W., 'Fourier Optics', pp. 61, McGraw-Hill Book Company, 1968

Recognition of Range Data Using Distance Measure of Attributed Graphs

Chong-Huah Lo and Hon-Son Don

Department of Electrical Engineering
State University of New York
Stony Brook, New York 11794

Abstract

Objects in the range image are represented by attributed graphs. The nodes of the graphs are smooth surface patches on the objects, which are attributed by the invariant features extracted from the surface of these patches. A scalar graph distance measure for the semantic and structural differences between the graphs are then computed. The distance measure guides the object inference and groups surfaces in the range image into objects. A minimum distance classification scheme is proposed for the recognition of objects in the range image.

those elementary objects. We shall illustrate by examples that the graph distance is very effective in object inferencing and grouping visible surface patches into objects.

This paper is organized as follows. In Section 2, we discuss our range data segmentation algorithm. In Section 3, we discuss our method of quadric surface parameter estimation. In Section 4, we discuss the algorithm of attributed graph matching and the calculation of graph distance measure. Experimental results are presented in Section 5. Section 6 contains our conclusions and remarks.

1. Introduction

Object recognition is important in Computer Vision. In many practical situations, the objects to be recognized have complex structures. A useful representation of such objects is to decompose it into simpler subpatterns. Each subpattern is considered a primitive entity which is described by its features. The structural relationships among those subpatterns are represented by certain data structure. In this paper, we use the attributed graphs to represent object prototypes in the range data. Euclidean invariant features extracted from surfaces are used to describe the nodes in the graph. The proposed object representation is a multiview representation. Objects in the scene are recognized by matching their attributed graphs with those of models. Our graph matching algorithm will compute a scalar distance measure which represents the semantic and structural differences between graphs [1][11]. The decision making in pattern recognition is based on the graph distances. It classifies objects into model categories and inspect the defects and distortions on the surfaces of the objects.

2. Range data segmentation

The range data are first segmented into uniform regions. There are edge detectors which can label the edge pixels on the range images [3][5]. However, complicated edge linking procedure must be used to group edge pixels into region boundaries. We use a region based segmentation approach [8]. Our proposed range data segmentation algorithm has four steps. These are split-and-merge, region grouping, large region growing/small region elimination and postprocessing. The algorithm is described by a flow chart in Fig.1. The range image is initially represented by a quad tree. Each node of the quad tree is a square image block marked by the coordinates of the lower left corner and the size of the square. The image data in each block is assumed to be a square patch of the bivariate polynomial surface. The order of the polynomial surface depends on the size of the block. The root mean squared error between the image data in the block and the interpolated value at each sampling point is computed and stored in the node of the quadtree. The merging phase precedes the splitting phase. During merging process, four sibling nodes are merged into a parent node, if the image data in the parent node can be approximated by a bivariate surface. We then recursively split each nonuniform image block into four quadrants until each quadrant is either uniform or its size is less than a predetermined threshold value. We subsequently merge non-sibling nodes in the quad tree. A block will merge into an adjacent block, if both blocks are uniform and there is no discontinuity of depth along and across their borders. Moreover, the surface curvatures and orientations of these two blocks should be compatible. The region adjacent graph (RAG) is generated after this region grouping operation. The large regions in the segmented image are first contracted and then expanded. Small regions in RAG which usually reside on edges of objects will merge into the nearby large regions. The adjacency relation of the nodes in RAG is updated in the process of small region elimination.

In the literature, various structural pattern recognition methods have been proposed for range data recognition. Some used rigidity constraints to guide the matching of primitive features of the objects [2][4][10]. Others constructed the attributed hypergraph representation for objects and determine graph monomorphism between the hypergraphs for object recognition [13]. Our object description is insensitive to coordinate transformation and stable against changing of viewpoints. Our recognition system, which uses an informative graph distance measure for objects inference and classification, is robust. In particular, it provides useful information when there are defects on the visible surfaces.

The segmented images usually look too fragmentary. A smooth surface patch in the range image is often segmented into several pieces. The RAG which contains many nodes and adjacency relations is not readily suitable for structural pattern recognition. The over segmentation is due to the fact that the bivariate polynomial approximation $Z(x,y) = \sum_{m,n=0}^N A_{mn} Q_m(x) Q_n(y)$ used in the uniform predicate is only a local approximation of curved surfaces. The curved surfaces are globally approximated by the quadric surface. The bivariate polynomial approximation is used to measure the uniformity of each image block in the recursive split-and-merge process, because it has low computation complexity. The RAG representing the over segmented range image is further processed by nonrecursive region mergings. The quadric parameters in each labeled regions are estimated. Adjacent regions with similar quadric parameters are merged. As a result, the processed RAG has less number of nodes and adjacency relations. Finally, the labeling of each pixel on the region boundary is determined by computing the curvature property in the neighborhood of

In our approach, the range data are first segmented into smooth regions which are enclosed by labeled edges. The smooth regions on the surfaces of the object form the nodes of the graph. The quadric surface parameter are estimated from the range data of each smooth region. The nodes are attributed by invariant features computed from these parameters. They are recognized by statistical pattern recognition method. The attributed graph representing a prototype view of a model object is then constructed by connecting adjacent region nodes which are visible from that view.

The attributed graph of range data representing a scene is derived by a different procedure. A graph is first constructed by connecting adjacent regions in the range data separated by convex crease edges. The connected components of this graph are then extracted. We examine each connected component and connect region nodes which are separated by concave crease edges. The resulting connected attributed graph represents a 3-D view of an elementary object in the range data. We shall merge adjacent elementary objects into a composite object, if each of them matches to the same prototype and the composite object is closer to the prototype than

that pixel. The concave crease edges correspond to the positive curvature extrema. The convex crease edges correspond to the negative curvature extrema.

3. Surface Patch Description

In this section, we shall give a detailed discussion of the surface parameter estimation. Our surface parameter estimation is represented by the diagram in Fig.2. A planarity test is first performed on each smooth region. A region is planar, if the smallest eigenvalue of the variance matrix

$$D = \sum_{i=1}^N (\vec{x}_i - \bar{x})(\vec{x}_i - \bar{x})^T \quad (1)$$

is less than a threshold value. In (1), \bar{x} is the mean of the vectors \vec{x}_i , and N is the number of sample points on the patch. The surface normal, \vec{V} , is parallel to the eigenvector corresponding to the smallest eigenvalue of D and pointing toward the viewer. The distance between the origin and the plane, d , is determined by

$$d = -\vec{V}^T \bar{x} \quad (2)$$

The area, the second order 3-D moment invariants [6][7] and the normal vector of the planar patch are used as its attributes. If a region is classified as nonplanar, then it is assumed to be quadric. The quadric surface is described by the following quadratic equation which contains ten parameters.

$$h(x, y, z) = x^T A x + v^T x + d = 0, \quad (3)$$

where

$$A = \begin{bmatrix} a_{11} & a_{12}\sqrt{2} & a_{13}\sqrt{2} \\ a_{12}\sqrt{2} & a_{22} & a_{23}\sqrt{2} \\ a_{13}\sqrt{2} & a_{23}\sqrt{2} & a_{33} \end{bmatrix} \quad (4)$$

The quadric parameters can be derived from the sampled data by minimizing the sum $\sum_{i=1}^N h(x_i, y_i, z_i)$ subjected to the invariant constraints on these parameters. The constraints will eliminate the undesired solution where all parameters are zero identically. Parameter estimation using quadratic constraint $Tr(AA) = 1$ is robust.

The six coefficients in matrix A corresponding to the smallest eigenvalue of the 6×6 matrix $B - CD^{-1}C^T$, where B, C, D are the $6 \times 6, 6 \times 3, 3 \times 3$ matrices in the correlation matrix M defined as follows.

$$M = \sum_{i=1}^N Q_i Q_i^T = \begin{bmatrix} B & C \\ C^T & D \end{bmatrix} \quad (5)$$

$$Q_i = [x_i^2 - \bar{x}^2, y_i^2 - \bar{y}^2, z_i^2 - \bar{z}^2, \sqrt{2}(x_i y_i - \bar{x}\bar{y}), \sqrt{2}(x_i z_i - \bar{x}\bar{z}), \sqrt{2}(y_i z_i - \bar{y}\bar{z}), x_i - \bar{x}, y_i - \bar{y}, z_i - \bar{z}]$$

The 3×3 matrix D is the variance matrix in (1). Once the coefficients in A are known, the linear and the constant coefficients of the quadric surface can be determined. The principal values, which are the eigenvalues of the symmetric matrix A , are Euclidean invariant features. Quadric surfaces can be classified by their principal values.

Each node of the attributed graph, which represents a surface patch, contains the following information.

- (1) surface_type;
- (2) invariant_features;
- (3) Surface orientation; (normal vector for planar surface, symmetry axis for cylindrical surface, principal axes for general quadric surfaces);

The surface equation of a region containing too small number of pixels can not be correctly estimated. Such small regions can be textures or defects on the objects. Nodes corresponding to small surfaces will not be attributed by numerical features. They are only characterized by the surface type "small". Nodes corresponding to disjoint regions on a sphere or a cylinder will be described by the same set of surface parameters. These nodes will be distinguished by their adjacency relations to neighboring regions.

When a non-symmetric 3-D object is in two different orientations, the rotation parameters can be determined by the normal vectors or symmetry axes of surface patches on the object. However, if the object is symmetric, there can be more than one candidate rotation matrix. There is no need to establish feature point correspondence in orientation estimation using surface parameters. Therefore, this approach is computational economic. The

computed rotation matrix will impose orientation constraints in the later graph matching process.

Each surface on the model objects will be associated with a discriminant function of invariant features. A surface patch in the scene will be identified with a surface on the models, if the value of the discriminant function of that prototype surface is larger than those of all others and greater than a predetermined threshold value. However, in order to handle occlusion, we identify a planar surface in the range data which can not be recognized by discriminant functions, with all planar surfaces on the model whose area is larger than its. The graph matching procedure, using orientation constraint and adjacency relation of nodes, will recognize the partially occluded planar surfaces. Moreover, when large number of training samples are available, the parameters in the discriminant function can be determined by the means and variances of the features.

4. Graph Matching and Scalar Distance Measure

Informative distance measure of attributed graphs is useful for computer recognition of 3-D objects. In the following, we shall formulate the graph matching as a goal state searching process. A state space representation is constructed from two given graphs. The "goal" state represents an optimal subgraph matching between these two graphs. The distance measure is systematically computed by a state space searching algorithm [9]. Distance between unmatched graphs is infinite. The distance between two matched graphs is finite and is a quantitative measure of the semantic and structural differences between the graphs. The general expression of the finite distance measure is given as

$$D(G, G') = \sum_{n=1}^N d_n f_n(x_n, x'_n) + \sum_{ij} w_{ij}(x_i, x'_j) + \sum_{ij} w_{ij'}(x_i, x'_j) \quad (6)$$

where N is the number of matched nodes in the two graphs. The matched nodes must be of the same surface type and described by the same set of numerical features. Their primitive deformation measure $f_n(x_n, x'_n)$ is a symmetric function of the feature vectors of the nodes. If the statistical discriminant functions are used in surface classification, the corresponding primitive deformation measures are defined by the following equation

$$f_n(x_n, x'_n) = c_n \left[1 - \exp \left\{ - \frac{\sum_i (x_i - x'_i)^2}{2\sigma_i^2} \right\} \right] \quad (7)$$

where c_i 's are the standard deviations of the invariant features, and c_n is a normalization constant. The positive weight constant d_n in (6) depends on the degree of the n th matched node. Nodes with larger degrees are more heavily weighted. The weights w_{ij} and $w_{ij'}$ are the contributions from the unmatched branches x_i, x'_j and x_i, x'_j' in G and G' , respectively. The structural distortion measure of each unmatched branch depends on the surface types and the attributes of the two nodes of the branch. The distortion measure of an unmatched branch connecting two large surface patches has to be a large number (its order of magnitude is ten or hundred, if $c_n = 1$). On the other hand, the distortion measure of an unmatched branch connecting small surface to a large surface or another small surface will be a small number (its order of magnitude is one tenth).

For any given node of an attributed graph, we define the basic attributed graph (BAG) at this node to be a one level tree, whose root is the given node, and whose leaves are nodes adjacent to it [1]. We denote the BAG at node n_i as G_i . Our approach of matching two AG's G and G' , is to progressively reconstruct the graphs from their BAG's. This process of reconstruction has a state space representation [1][12]. A labeled state of the state space denotes the reconstruction of a subgraph from graph G and a subgraph from graph G' , as well as the matching of their respective BAG's. The initial state of the space contains a pair of empty graphs. Each successor of the initial state contains a pair of matched BAG's whose roots are not small surfaces. The matching condition of two BAG's is the following. The surfaces of the root nodes in the BAG's must be of the same type and they are identified to be the same surfaces by their numerical features using discriminant functions. Moreover, when the surfaces are not spherical, the relative orientation of these surface patches are computed from their principal axes. The surfaces of the leaves in the BAG's are then pairwise matched. In addition, the relative orientation of the matched leaf nodes should be consistent with that of the root nodes. All leaves whose surface types are not "small" must have their matched partners. Exceptions are those in one of the BAG which contains more leaves representing large surfaces than the other. The distance between G_k and G'_k is defined as

$$d(G_k, G_k') = \sum_{i=1}^N f(x_i, x_i') + \sum_j w(x_j, x_j') + \sum_j w(x_j', x_j) \quad (8)$$

where x_1 is the root node of G_k and x_1' is the root node of G_k' . A general state in the space contains an ordered list of root nodes of the matched BAG. $\{(n_1, n_1'), (n_2, n_2'), \dots, (n_k, n_k')\}$, which is called the core of the state. The state also stores the leaves of G_k and G_k' whose roots are the last elements of the ordered list in the core, i.e. (n_k, n_k') . These leaves are called the terminals of the state.

The rule of state expansion is as follows. A pair of nodes, denoted by (n_i, n_i') , are selected from the terminals of the state, which are not small surfaces. If (n_i, n_i') is not in the core and their BAG's G_i, G_i' are matched, then a successor of the present state is constructed by appending the pair (n_i, n_i') to the core, and replacing the old terminals by the leaves of G_i, G_i' . A cost of state transition is associated with the arc which connects the parent state and the successor. The cost is the distance measure $d(G_i, G_i')$ in (8). The data structure of each state node and the state transition rule ensure that the topology of the state space is a tree. We define the goal states to be those whose cores contain all the nodes of either G or G' , or both of them. Two graphs G and G' are matched, if there is a directed path in the search tree that connects the initial state to the goal state. The graph distance is defined to be the minimum cost of such paths. On the other hand, if G and G' are not matched, the graph distance is infinite.

The graph distance can be efficiently calculated by the following data-directed state space searching algorithm.

Algorithm : Graph Distance

Input:

- (1) Attributed graph G of the model
- (2) Attributed graph G' of an object in the scene
- (3) Parameters for the discriminant functions and deformation measures

Output:

Graph distance between G and G'

Method:

- [1] Identify node_s in G with nodes in G' using discriminant functions.
- [2] Put the initial state on a list called HEAP, and set graph_distance=0.
- [3] If HEAP is empty, exit and return graph_distance = INFINITE.
- [4] Remove the state from HEAP whose graph_distance is the smallest. Call this node n .
- [5] If n is a goal state, exit and return graph_distance.
- [6] Expand node n , generating all of its successors. If there are no successors, exit and return graph_distance = INFINITE. For each successor n_i , compute its graph_distance by graph_distance = graph_distance + $D(G_i, G_i')$, where G_i, G_i' are the most recently imbedded BAG's. Put these successors on HEAP, and associate the graph_distance with them.
- [7] Go to [2].

The list HEAP is implemented as a priority queue. The computed graph distance has the expression in (6) where $d_n = (\text{degree of the } n\text{th matched node}) + 1$. The entire matching processing is described by the ordered list of nodes in the core of the goal state. If there are large number of models, we can increase the speed of matching by first screening the models through special features (such as number of planar patches and spherical surfaces) of the objects in the scene.

5. Experimental results

In this section, we present our experimental results of the surface parameter estimation and the graph matching algorithm. The algorithms were implemented in "C" language and ran on VAX 11/780 under UNIX system. The IMSL Fortran routines LINV2F and LSVDF were called by the "C" programs to perform matrix inversion and eigen decomposition in surface parameter estimation. The threshold parameter in the planarity test was defined by the empirical formula $\eta = 0.5 \cdot N$, where N is the number of data points used in the estimation.

The graph distance measure was used in object inference and recognition. Both range images and synthetic images were used in the experiments. Images in Fig.3(a) and Fig.3(b) are the range data of a synthetic 3-D objects in two different orientations. The image in Fig.3(a) was used as the

model and the image in Fig.3(b) was used as the scene. The model was segmented into five surface patches. The segmented image and its attributed graph whose nodes are labeled by numbers are shown in Fig.4(a). The range data in Fig.3(b) was segmented into three elementary objects according to the labelings of the edges. They are shown in Fig.4(b). The statistical parameters of the invariant surface features were computed from the digitized data of the synthetic object which was transformed by coordinate transformations. The Gaussian discriminant functions and the deformation measures are constructed from these parameters. The graphs of the elementary objects were matched to the model graph. We assumed that the normalization constants in all the primitive deformation measures are one, i.e., $c_n = 1.0$. These distances contain structural distortion measures and are large. The merging of the graphs of the elementary objects will reduce the contribution of structural distortion to the graph distance. This suggests that the three components in Fig.4(b) should merge into one. Our algorithm can successfully recognize objects in a complex scene. The scene in Fig.5(a) contains synthetic objects of a sphere, two cylinders and a polyhedron (a wedge on top of a block). Three planar surfaces of the polyhedral objects

are occluded by the sphere and cylinders. The model polyhedron is given in Fig.5(b). The occluded surfaces have different moment invariant features from those of the model surfaces. They could not be classified by the discriminant functions. Each of them was tentatively identified with any planar surfaces on the model object, whose area is larger than that of the occluded surface. In Fig.5(a), the surface 'c' was identified with surface '2' in Fig.5(b). Both surfaces 'd' and 'e' in Fig.5(a) were identified with surfaces '2', '3', and '4' in Fig.5(b). Our graph matching program used orientation and the adjacency relation constraints to eliminate the false matches between the occluded surfaces and the model surfaces. It computed a finite distance between the occluded polyhedron and the model polyhedron.

Real range images are given in Fig.6(a) and Fig.7(a), respectively. Their segmented images were labeled and shown in Fig.6(b) and Fig.7(b), respectively. The Euclidean distance discriminant functions are used. Based on the segmented images, the surface parameters and other invariant features of each smooth region in the range data were computed. The object in Fig.6(a) was used as the model. After the object inference procedure, the attributed graphs of the two objects are shown in Fig.8. Each graph contains three connected components. The graph distances between the corresponding connected subgraphs are calculated. Objects are identified correctly. In all experiments, we found that our graph matching algorithm worked well and fast. However, the feature extraction procedure often took a little more time.

6. Conclusions

In this paper, we present an attributed graph representation for the range data and a method of analyzing segmented range data for object recognition. The nodes of the attributed graphs are the symbolic representation of the smooth regions. Nodes representing planes are described by their areas and moment invariant features. On the other hand, nodes representing quadric surfaces are described by their principal values. The object identification is to match the attributed graphs of objects in the scene with those representing prototype views of the models. The matching uses both statistical and structural pattern recognition techniques. The matching of the structures of the graphs is formulated as a searching problem in a state space. The graph distance between two graphs is calculated by the uniform cost algorithm. The graph distance is useful for object classification. It is also used in the spatial inference to direct the merging of objects in the range data into a compound object.

Reference

- [1] M.A. Eshera and K.S. Fu, "A graph distance measure for image analysis," *IEEE Trans. Syst., Man, Cybern.*, vol. SMC-14, no. 3, pp. 398-408, 1984.
- [2] O.D. Faugeras and M. Hebert, "The representation, recognition, and positioning of 3-D shapes from range data," in *Three Dimensional Machine Vision*, edited by T. Kanade, Kluwer Academic Publishers, 1987.
- [3] T.J. Fan, G. Medioni, and R. Nevatia, "Segmented Descriptions of 3-D Surfaces," *IEEE Trans. Robot. and Auto.*, vol. RA-3, pp. 527-538, Dec. 1987.
- [4] T.J. Fan, G. Medioni, and R. Nevatia, "Recognizing 3-D objects using surface descriptions," *IEEE Trans. Pattern Anal. Machine*

- [5] S. Inokuchi, T. Nita, F. Matsuday, and Y. Sakurai, "A three dimensional edge-region operator for range pictures," in *Proc. 6th Int. Conf. Pattern Recognition*, Munich, West Germany, Oct 19-22, 1982, pp. 918-920.
- [6] C.H. Lo and H.S. Don "3-D moment forms : their construction and application to object identification and positioning," in *IEEE Trans. on Pattern Anal. Mach. Intell.*, vol. PAMI-11, pp. 1053-1064, Oct. 1989.
- [7] C. H. Lo and H. S. Don. "Pattern recognition Using 3-D moments", to appear in *Proc. of the 10th International Conf. on Pattern Recognition (ICPR)*, Atlantic City, New Jersey, June 16-21. 1990.
- [8] C. H. Lo, "Range data recognition using 3-D invariant features", Ph.D dissertation. Dept. of Elect. Engin. SUNY at Stony Brook. 1990.
- [9] N.J. Nilsson *Problem Solving Methods in Artificial Intelligence*. New York: McGraw-Hill, 1971.
- [10] M. Oshima and Y. SHirai, "Object recognition using three-dimensional information", *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 3, no. 4, pp. 353-361, 1983.
- [11] A. Sanfeliu and K.S. Fu, "A distance measure between attributed relational graphs for pattern recognition," *IEEE Trans. Syst., Man, Cybern.*, vol. SMC-13, no. 3, pp. 353-362. 1983.
- [12] W. Tsai and K.S. Fu, "Subgraph error-correcting isomorphisms for syntactic pattern recognition", *IEEE Trans. Syst., Man, Cybern.*, vol. SMC-13, no. 1, pp. 48-62. 1983.
- [13] A. Wong, S.W. Lu and M. Rioux. "Recognition and shape synthesis of 3-D objects based on attributed hypergraphs". *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 11, no. 3, pp. 279-290. 1989.

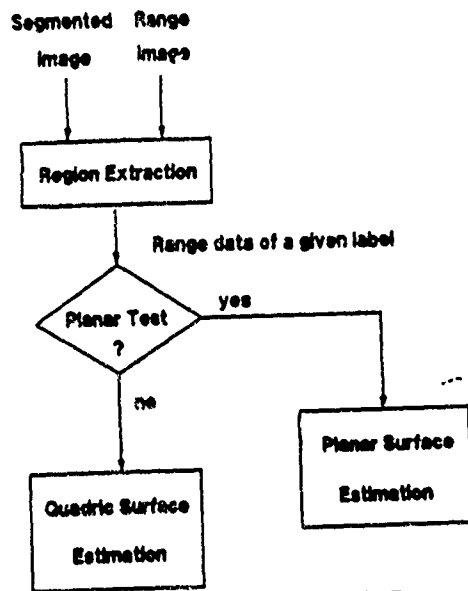


Fig.2



Fig.3(a)



Fig.3(b) :



Fig.4(a)

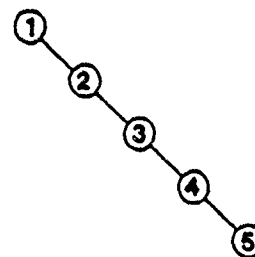


Fig.4(b)

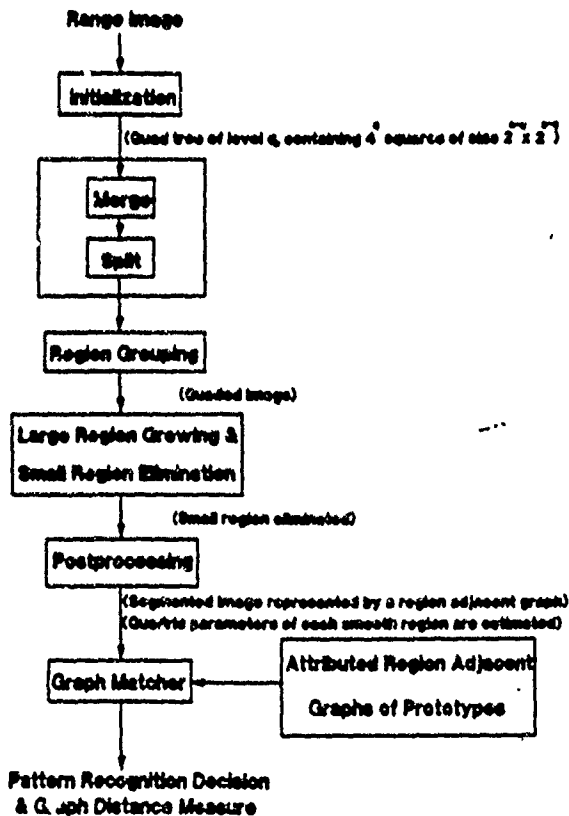


Fig.1

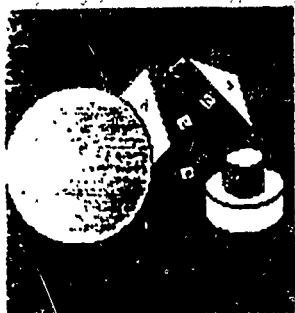


Fig.5(a)

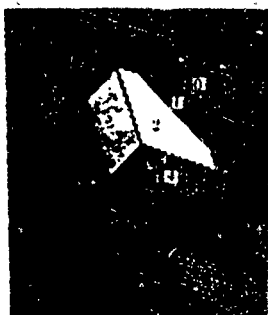


Fig.5(b)

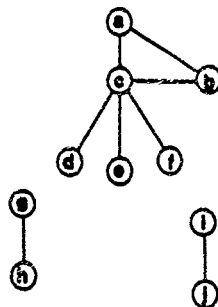


Fig.8(a)

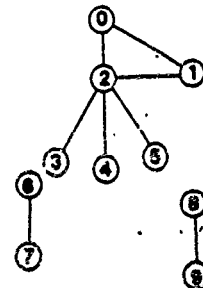


Fig.8(b)



Fig.6(a)

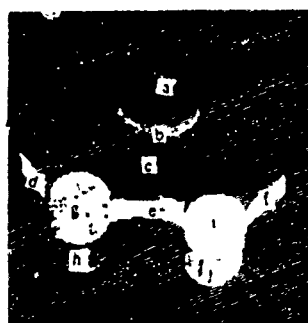


Fig.6(b)



Fig.7(a)

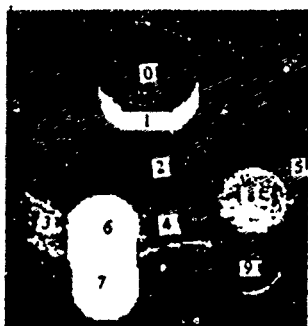


Fig.7(b)

The Application of Laser Speckle Interferometry to Measuring Strain and Strain Rate Under Dynamic Loading Conditions

J.F. Emslie¹
J.L. Green²
S.C. Chou²
P.K. Gupta¹
F.P. Chiang¹

1 Abstract

This report investigates the application of speckle interferometry for the measurement of strain when a material is subjected to various loading rates and elevated temperature conditions. In Stage I of the experiments, an Olsen testing machine at SUNY Stony Brook is used to conduct uniaxial tension tests at strain rates on the order of 10^{-3} sec^{-1} . The Medium Strain Rate Facility at the Army's Materials Technology Laboratory in Watertown, Massachusetts is used to conduct uniaxial tension tests at strain rates of 10^{-2} sec^{-1} to 10^{-1} sec^{-1} , temperatures up to 250° F and heating rates of 250° F/sec in the experiments of Stage II. Strain is measured by laser speckle interferometry technique and by strain gages. The results of both methods are compared. The laser speckle interferometry results are in agreement with the strain gage results.

Laser speckle interferometry is also used for the measurement of strain at large deformation, i.e., necking region of a tensile specimen. This technique also indicates that laser speckle interferometry will be an excellent non-contact localized strain measuring device for adverse conditions.

2 Introduction

There are many applications where a material is subjected to high strain rates, high heating rates and large deformation, for eg., forging, ballistic impact, and penetration. To properly analyze and design systems, these conditions must be taken into consideration. Therefore, it is necessary to acquire material relationships under conditions of high temperature, high heating rates, and various loading rates up through large strain. Measurement of large strain when subjecting a material to strain rates greater than 0.0001 sec^{-1} and high temperatures has always posed severe limitations. Many investigators have developed contact and non-contact extensometers for measuring strain in hostile environments [6]. These extensometers

are often limited in strain rate and total strain they can measure. One non-contact method for measuring strain under quasi-static load conditions is the laser speckle interferometry technique [3] [1] [4]. This optical technique of double exposure laser speckle photography is well established for measuring in-plane displacement. This technique has been applied to measure the transverse displacement of a cantilever beam following tip impact [5].

For this investigation, a test plan is formulated to determine the feasibility of using *lsr* to make strain measurements between strain rates of 0.0001 sec^{-1} and 0.01 sec^{-1} and at elevated temperatures. To accomplish this, the displacement field along the axial direction of a specimen is measured. Under uniaxial tension, axial strain can be found by calculating the slope of the axial displacement data versus the gage length of the specimen. Due to the nonuniformity of the strain field at the ends, only the displacement data from the central region of the specimen are used. The displacement at various points along the axis and at various instants of time is obtained by using laser speckle interferometry.

3 Experimental Setup and Procedure

By using double exposure speckle photography the displacement of a surface in a plane normal to the line of sight may be measured. This is done by using a camera to record two superimposed images of the surface, one before and one after deformation. The optical setup of figure 1 shows the location of the camera and the ruby laser with respect to the specimen. The laser beam from the pulsed ruby laser is expanded by using a concave lens. This allows the entire surface of the specimen facing the camera to be illuminated. The intensity of the pulse is sufficient to expose the film within its pulse width time of 30 nanoseconds, when in Q-switching mode, and generate the speckle pattern on the film. Flat white paint is sprayed on the transparent Plexiglas specimens of stage one to make them reflective and so that they produce a good speckle field. The natural reflectiveness of the metals tested in stage two helps them produce a good speckle field.

The pulsed ruby laser allows one to capture instants in dynamic experiments due to its short pulse time. If the move

1 - Department of Mechanical Engineering, State University of New York at Stony Brook, Stony Brook, N.Y. 11794-2300
2 - Army's Material Technology Laboratory, 405 Arsenal Street Watertown, MA 02172
3 - Current Address: Systron Donner Corporation, Inertial Division, 2700 Systron Drive, Concord, CA 94518

ment of the object's surface between the two exposures is larger than the diameter of the speckles recorded by the camera onto the film and if these speckle remain correlated with one another, then the image will scatter a beam of laser light into a diffraction halo. The intensity of the light in this halo varies periodically across the field yielding cosine square fringes. This is shown in figure 2. These fringes will have an angular spacing, α , given by

$$\sin \alpha = \frac{\lambda m}{D} \quad (1)$$

where λ is the wavelength of the readout beam, m is the demagnification of the image and D is the surface displacement [3]-[4]. The fringes are perpendicular to the direction of displacement.

The specklegrams are read by using a He-Ne laser to illuminate different points across the gage length of the specimen. The farfield diffraction pattern is observed at a distance l from the specklegram and the distance, S , between dark fringes is measured. Equation 2 relates the angular spacing, α , to these quantities

$$\sin \alpha = l/S \quad (2)$$

By combining equations 1 and 2, object displacement is obtained.

$$D = \frac{m\lambda l}{S} \quad (3)$$

The demagnification factor, m , cancels out when strain is calculated due to the fact that the gage length used for calculating the strain is measured from the image on the film rather than from the specimen itself,

$$\epsilon = \frac{D_2 - D_1}{l_{21}} = \frac{m\lambda l}{m l_{21}} \left(\frac{1}{S_2} - \frac{1}{S_1} \right) = \frac{\lambda l}{l_{21}} \left(\frac{1}{S_2} - \frac{1}{S_1} \right) \quad (4)$$

where l_{21} is the distance traversed on the image by the He-Ne laser and D_2 and D_1 are the displacements at image locations 2, i_2 , and 1, i_1 , respectively. Strain rate is calculated by dividing equation (4) by the time between exposures, Δt ,

$$\dot{\epsilon} = \frac{\lambda l}{l_{21} \Delta t} \left(\frac{1}{S_2} - \frac{1}{S_1} \right). \quad (5)$$

Figure 3 shows the heating apparatus used to perform the soak and high heating rate tests done in Stage II of this work. (Stage I experiments used an Olsen testing machine to load Plexiglas specimens. Otherwise, the equipment used in Stage I and Stage II is identical.) The heating apparatus utilizes the resistive heating technique. The resistive heating technique passes current through the specimen using the specimen's resistance to convert the current to heat. The apparatus consists of an autotransformer (controlling the power output of the apparatus), a step down transformer (converting low current high voltage to high current low voltage) and a high voltage relay (allowing the autotransformer to be controlled by the computer). The computer triggers the high voltage relay by way of the test program, allowing current to flow to the specimen.

For the soak tests, the autotransformer voltage is adjusted and manually controlled so that the specimen maintains a constant temperature. The temperature, measured with a thermocouple, is monitored throughout the soak tests. For the high heating rate tests, the voltage is set and the temperature

and its time history are measured with a quick response thermocouple attached to the specimen. Then this setting, on an auto-transformer, is used on all the subsequent tests for that particular heating rate.

The loading for the experiments of Stage II is performed by the servo-hydraulic machine also shown in figure 3. This machine has the capability of pneumatic operation.

The electrical schematic of the experimental setup is shown in figure 4. The first exposure is taken before loading the specimen and the second exposure is taken after a certain predetermined load level (strain level in Stage I tests). The load level is increased from 4480 to 8064 N, depending on the heating and loading conditions to allow for measurable strain data using laser speckle. (For Stage I tests, the strain level is increased from 300 to 700 $\mu\epsilon$.) These measurable strain values lie between 900 to 2500 $\mu\epsilon$ for a particular specklegram. For values of strain less than 900 $\mu\epsilon$, displacement is not larger than the speckle diameter recorded with the camera aperture fully open and a demagnification equal to one. Conversely, with these camera conditions, strains greater than 2500 $\mu\epsilon$ make very difficult to distinguish the fringes and eventually lead to decorrelation of the speckle pattern.

The triggering of the ruby laser for the second exposure at any desired load level is achieved by comparing the load level signal to a desired reference voltage using a comparator circuit.

Other equipment used in Stage I and II include a high speed photodiode, for the determination of when the laser is fired, and a Nicolet oscilloscope, to store the photodiode, strain gage and load cell readings. The photodiode used in stage two tests has a rise and fall time of about 12 picoseconds and allows for accurate determination of when the laser fires. The Nicolet oscilloscope is able to record and store the voltage from the diode, load cell and strain gage for later analysis.

The tests performed have two purposes;

1. To demonstrate the applicability of laser speckle interferometry for obtaining strain rate curves (Stage I) and strain curves (Stage II).
2. To apply it to situations where other strain measuring devices fail, such as in severe environments and necking regions of a tensile specimen (Stage II).

To accomplish the first objective the strain rate and strain data of laser speckle interferometry is compared to that of strain gages. The two are found to agree, just like the tests done without the use of a strain gage as a check on the speckle data.

The situations where strain gages and clip gages fail were explored include elevated temperatures and the high strain in the necked regions of tensile specimens.

4 Results

A typical plot of axial displacement versus the gage length of the specimen is depicted in figure 5. Assuming a uniform strain field, the slope of the points is equal to the instantaneous strain.

The goal of Stage I tests is to determine strain rate and strain at various instants of time for the three different strain rates obtained by laser speckles as well as the strain rate

plotted versus time. The signal from the strain gages is verified to be identical for all repetitions of the process for a given strain rate. This confirms the repeatability of the experiment which is an essential requirement and assumption in applying this technique to construct the strain versus time curve, point by point. The strain rate is then estimated by the slope of the best fitting straight line through these points. A typical result is presented in figure 6. A total of 3 curves are obtained. For the first and third case, the percentage error is around 5%. For the second case, a percentage error of 16% is observed. The sources of error include the relatively significant electrical noise in the strain gage recordings which affect the accuracy of the strain gage results and the difficulty in obtaining a clean "t = 0" for all repetitions of the experiments.

Figure 7 shows one of the stress-strain curves obtained in the Stage II tests at the strain rates of 10^{-1}sec^{-1} . In each of the 3 cases performed at $\dot{\epsilon} = 10^{-1} \text{sec}^{-1}$, the speckle elastic modulus is in agreement with the strain gage modulus. The error in each curve is 7.69%, 4.3%, and 2.34%, respectively. Absolute speckle strain is found by adding the strain gage value at the time of the first pulse to the strain increase given by laser speckle. The error bars of the strain gage and speckle (approximately 6% for each specklegram) fall within this error range. The source of the error in the curves are due to the significant electrical noise in the strain gage and load cell voltages when the laser fires. Interpolation to the exact load and strain voltages when the laser fires is more difficult at the strain rate of 10^{-1}sec^{-1} , as compared to 10^{-3}sec^{-1} and static tests.

The last 2 figures demonstrate the applicability of laser speckle to measuring strain in situations where it was not possible to use strain gages or extensometers. In figure 8, the stress-strain curve for aluminum at room temperature measured with a load cell and strain gage is shown. The speckle data, obtained by adding the strain increment given by speckle to the strain gage value at the time of the first laser pulse, is found to lie slightly above the strain gage data. With error included, the speckle data does agree with the strain gage data.

Figure 9 shows the strain measurement capability of laser speckle in the regions of large strain, particularly the necking region. The specklegram and subsequent fringes observed, are obtained by first heating the aluminum to a very high temperature. Then it is pulled as slowly as possible until necking is visible. A first exposure is taken at this time. The second exposure is taken an instant after heating the aluminum to a very high temperature and allowing it to creep. Figure 9 shows that the strain increment in the necking region, due to the loading, is greater than that over the whole length of the specimen.

As in all previous cases, absolute strain measurement is not possible with speckle derived from the ruby laser in single pulse mode. Even in double pulse mode, absolute strain measurement is not possible. However, it is shown that speckle can measure increments of strain in large strain areas.

5 Conclusion

The applicability of laser speckle to measure strain increments has been demonstrated. It has been proven to apply in conditions of elevated strain rate, heating rate, temperature and areas of large strain. With a powerful chopped laser and a camera recording images at the same time of the laser pulses, absolute strain data versus time can be obtained in the severe conditions stated above. Subsequently, the limitation that it measures only increments of strain can be eliminated.

6 Acknowledgment

Financial support provided by the Army Research Office (Contract No. DAA20388K0033) is gratefully acknowledged.

References

- [1] Archbold, E., Burch, J.M. and Ennos, A.E., "Recording of In-plane Surface Displacement by Double Exposure Speckle Photography", *Optica Acta*, Vol. 17, No. 12, pp 883-898, 1970.
- [2] Archbold, E. and Ennos, A.E., "Displacement Measurement from Double Exposure Laser Photographs", *Optica Acta*, Vol. 19, No. 4, pp 253-271, 1972.
- [3] Chiang, F.P., "A New Family of 2D and 3D Experimental Stress Analysis Techniques Using Laser Speckles", *SAE Archives*, Vol. 3, Issue 1, pp 1-32, Feb 1978.
- [4] Khetan, R.P. and Chiang, F.P., "Strain Analysis by One-beam Laser Speckle Interferometry. I: Single Aperture Method", *Applied Optics*, Vol. 15, No. 9, pp 2205-2215, 1976.
- [5] Gupta, P.K. and Chiang, F.P., "Laser Speckle Interferometry Applied to Studying Transient Vibrations of a Cantilever Beam", *Journal of Sound and Vibration*, Vol. 133, No. 2, pp 251-259, 1989.
- [6] Chou, S.C. and Green, J.L., "Mechanical behavior of SiC/2014 Al Composite At High Temperatures", *Proceedings of an International Conference*, Taipei, May 19-23, 1986.

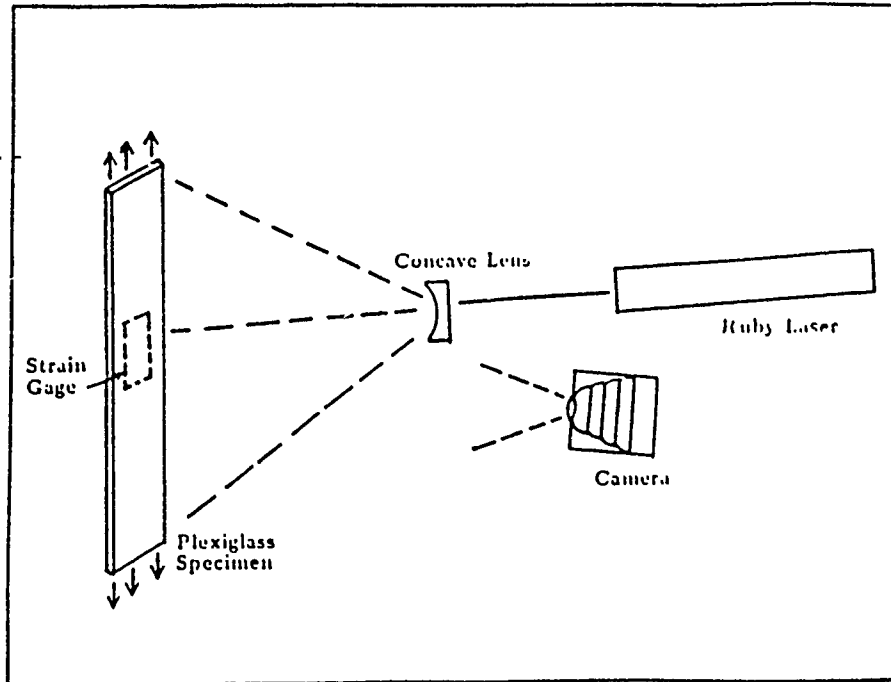


Figure 1: Optical Arrangement of Experimental Setup (Stage I and II)

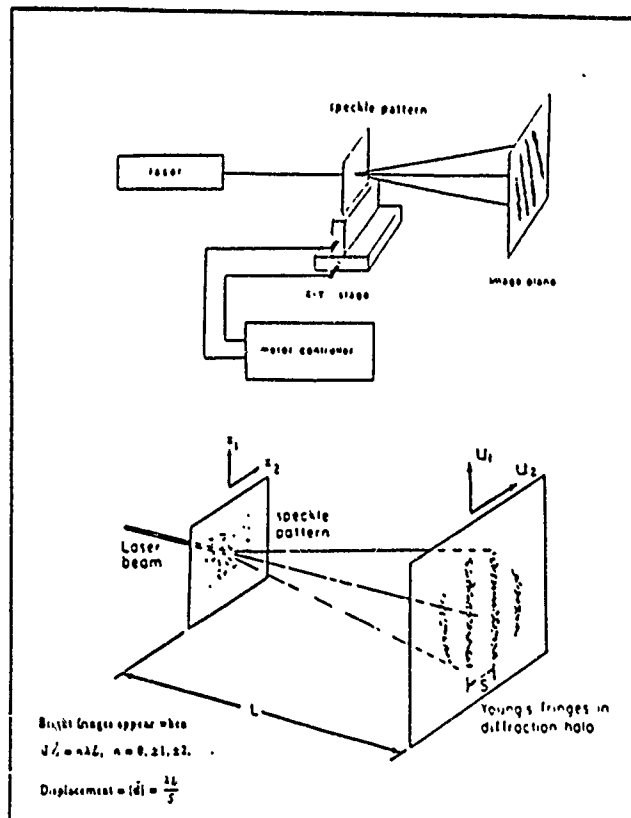


Figure 2: Measurement of Displacement by Laser Speckle Photography

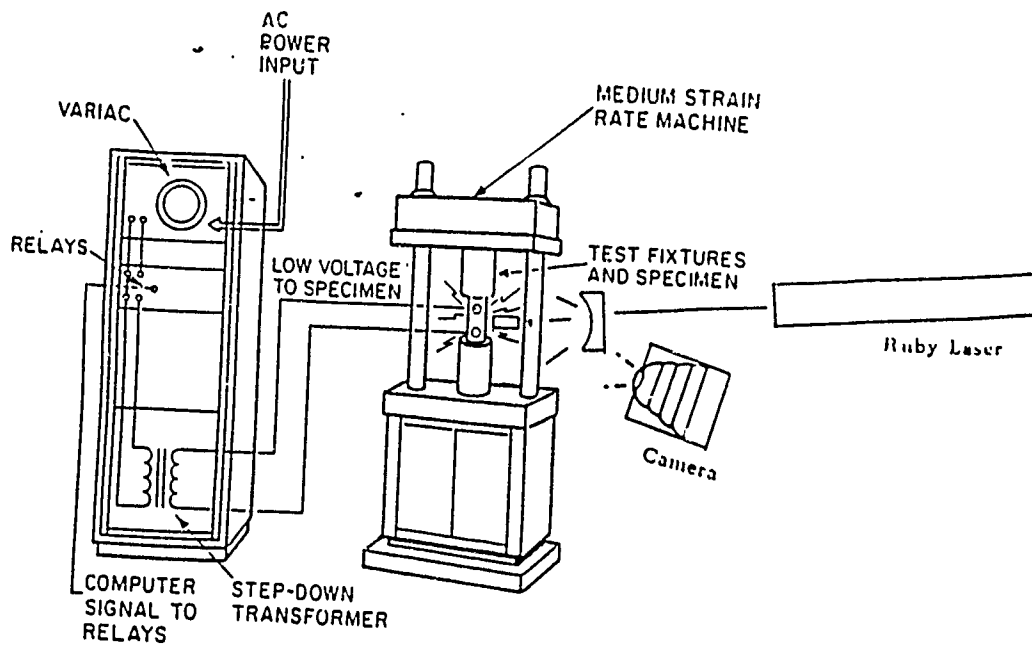


Figure 3. Loading and Heating Apparatus of the Experimental Setup

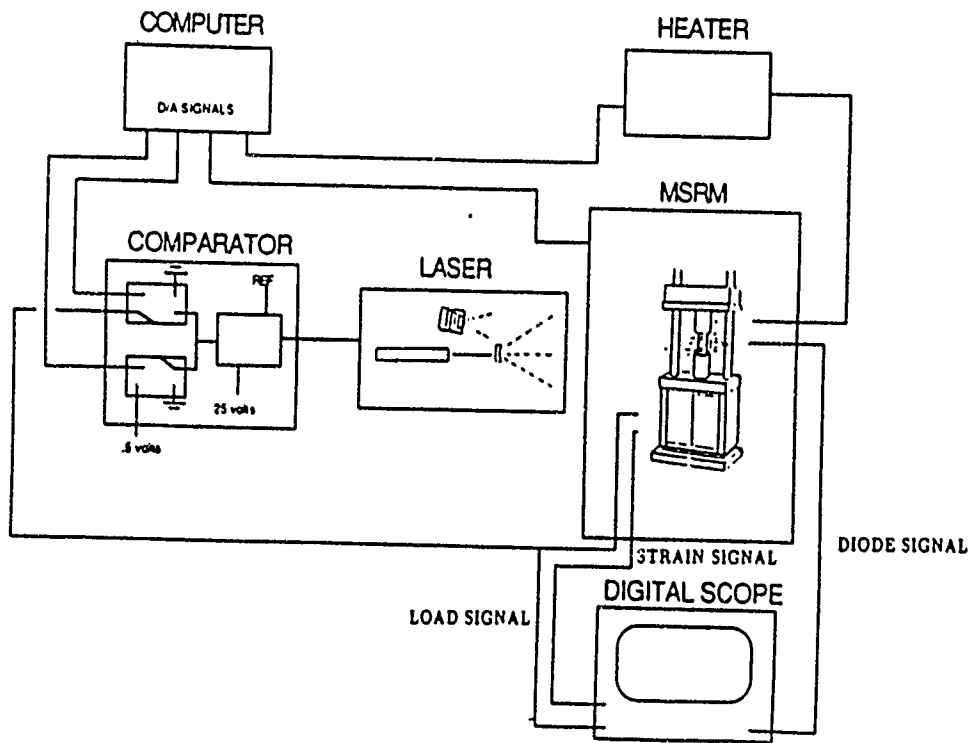


Figure 4. Electrical Schematic of Experimental Setup (Stage II)

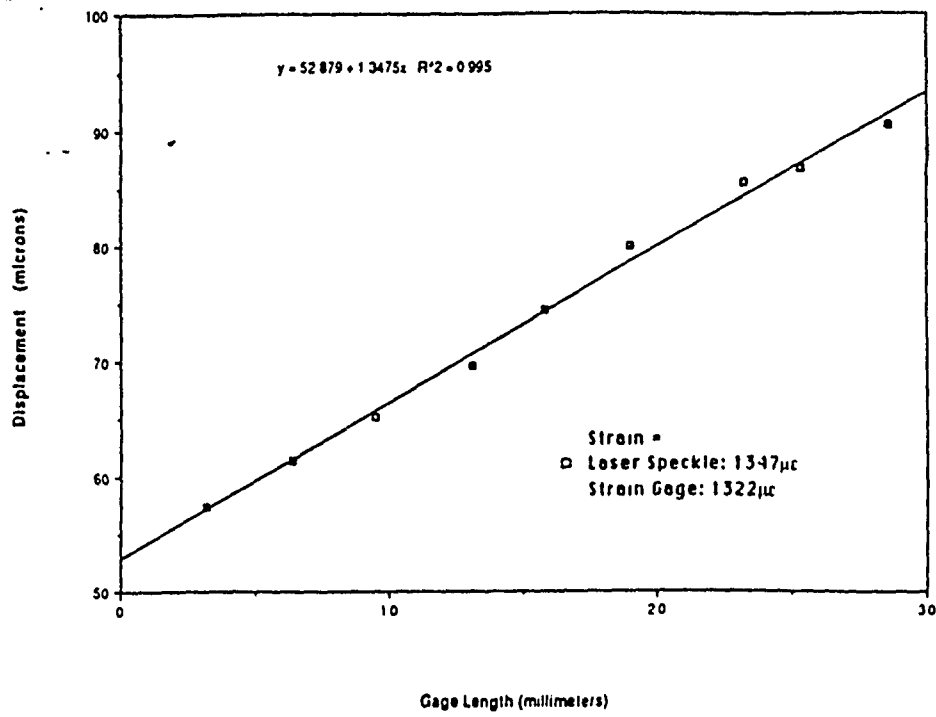


Figure 5: Axial Displacement versus Length of Image

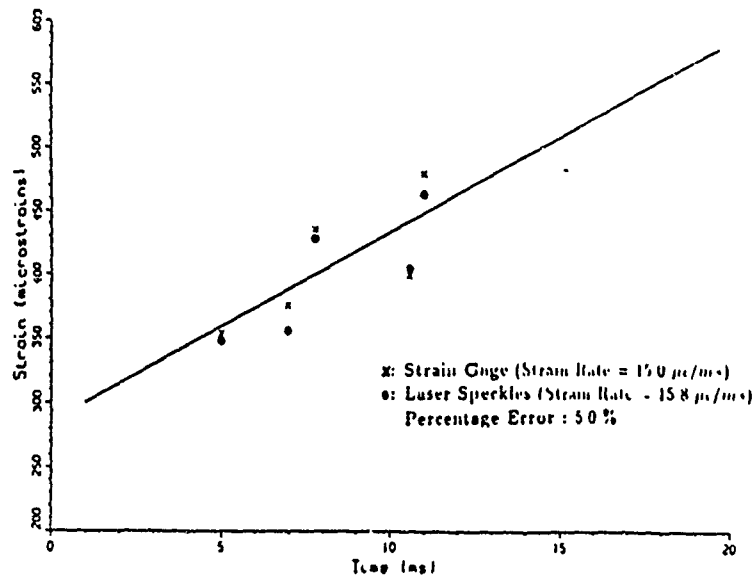


Figure 6. Strain versus Time; Plexiglas, $\epsilon = 15.0 \times 10^{-3} \text{sec}^{-1}$

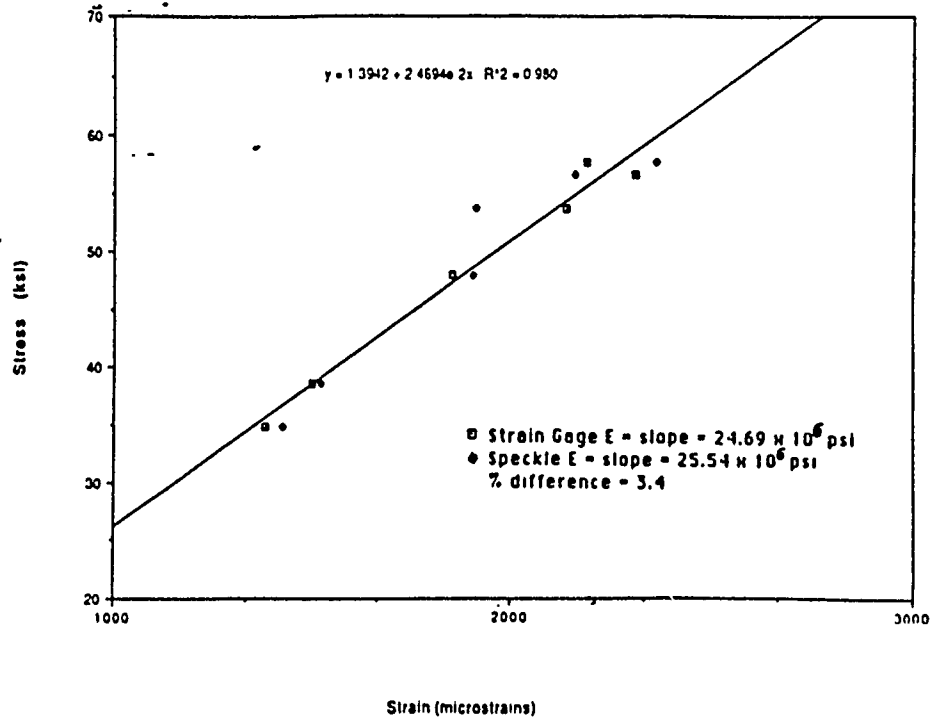


Figure 7: Stress versus Strain; Steel, $250^\circ F/sec$, $\dot{\epsilon} = 10^{-1} sec^{-1}$

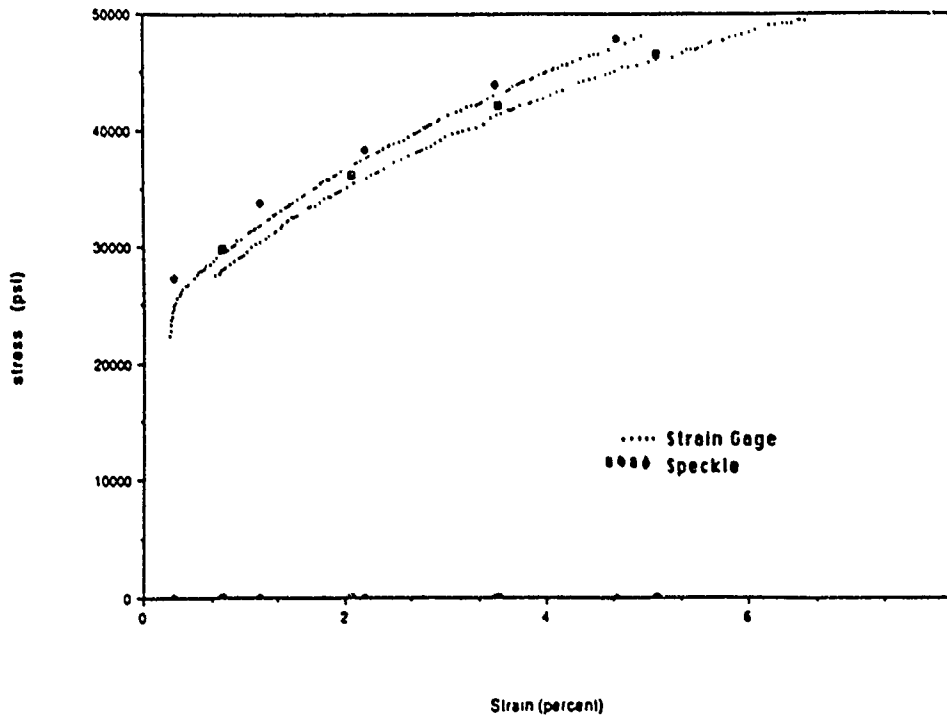


Figure 8: Stress versus Strain; Aluminum, Room Temperature, Static

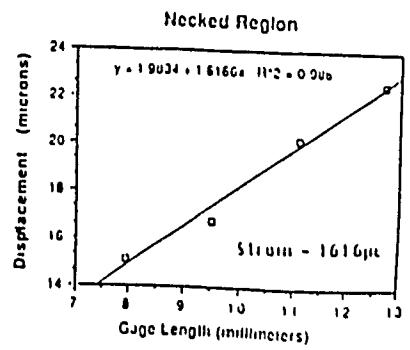
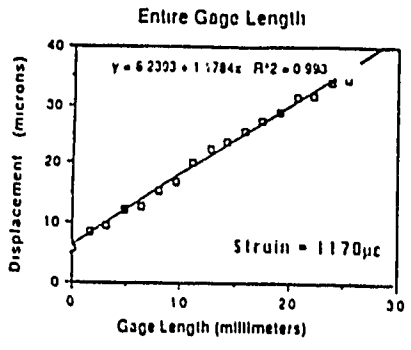
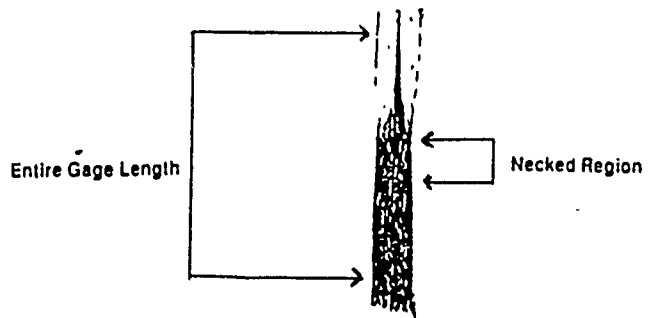


Figure 9- Necked Aluminum Specimen with Displacements
Plotted over Entire Gage Length and over the Necked Region

Study of Surface Roughening under Different Stress Mode by Correlation

F.P. Chiang, Y.Z. Dai; B.Q. Xu; A. Kato

Abstract

This paper presents some experimental results on the plastic deformation induced free surface roughening of aluminum and copper alloy materials under different stress modes by a computer based optical-numerical correlation analysis system. It is intended to discuss some fundamental issues for a non-contact nondestructive plastic strain evaluation technique that utilized surface roughening phenomenon in combination with laser speckle method. Experimental results for the materials tested and loading conditions used in the present work show that metallic material surface roughens isotropically and the magnitude of surface roughness is independent of material properties and stress mode but dependent of equivalent plastic strain. The validity of this statement on a more general base will enable the above mentioned plastic strain evaluation technique be calibrated on a certain material by simple loading method and then applied to general problems in practice.

1 Introduction

The phenomenon that metallic material surface roughens as it is subjected to sufficient amount of plastic deformation has been utilized in the development of techniques for plastic strain evaluation and fatigue life monitoring.^(1,2,3,4) These techniques all applied laser light to illuminate the area of interest on the specimen surface and correlated the diffracted light intensity distribution in the form of laser speckle pattern to either plastic strain or fatigue cycles which roughen the specimen surface and hence change the diffracted light intensity distribution accordingly. Most of the work was done on uniaxial tensile specimens made of a certain material with the expectation that the result could be applied to any other materials under any different stress modes. Obviously, before these techniques could be applied to general problems in practice it is necessary that some fundamental issues such as the influence of material properties and stress modes on plastic deformation induced roughening as well as the isotropy of roughening, be studied satisfactorily.

The papers^(5,6,7) on plastic deformation induced roughening available in published literature discussed the influence of material property, grain size, specimen thickness as well as stress systems in terms of principal strain ratio. The major conclusion is that free surface roughening is dependent of equivalent plastic strain but independent of material property and stress system. However, because all the researchers used stylus type profilometer, which only yields surface profile information along a single cross section, it naturally raises the question as to whether the obtained results are true along any other directions on the specimen surface. Besides, the influence of stress mode in terms of principal stress ratio on roughening has not been studied which motivated the authors of this paper to make further investigation along this line.

Realizing the drawbacks of stylus type profilometer, such as skid may leave a mark on specimen surface making successive measuring unreliable, and the fact that it is only an approximation due to the radius of the stylus, we used diffraction pattern correlation analysis. The diffraction pattern of the specimen surface profile obtained by illuminating the specimen with a laser beam, reflects both the two dimensional profile correlation length and the height variation.⁽⁸⁾ It is digitized and processed by a computer based image processing system and used to correlate the surface roughness in a non-contact, remote way.

2 Experimental Procedure

Uniaxial tensile specimens made of four different materials (aluminum alloy 1100-00, 2024-T3, 6061-T6 and copper alloy C26800-H01) were cut from stock sheet along the rolling direction. The specimen used for biaxial loading was made of copper alloy C26800-H01 and machined to a disk shape as shall be discussed in the following session. The material properties and gauge dimensions are listed in Table 1.

One side of the specimen surface was first polished by cloth buffer wheel and then by aluminum powder to an initial surface

¹F.P. Chiang is professor and Y.Z. Dai is graduate student with SUNY at Stony Brook

²B.Q. Xu is professor with the China Institute of Science and Technology, China

³A. Kato is professor with Chubu University, Japan

roughness of about $R_a = 0.05 \mu\text{m}$, where R_a is the arithmetic mean of surface profile from the mean line. The specimen surface was intended to be polished isotropically but the existence of some very fine directional scratches still remained. The other side of the specimen was reserved for plastic strain measurement by mounting strain gage or preparing moiré gratings on it.

Specimens were loaded step by step to different plastic deformation levels in a testing machine. Between every two loading steps, the specimen was released for plastic strain assessment, diffraction pattern recording and surface roughness measurement. Plastic strain was evaluated by strain gauge and moiré (or grid) method for small and large scale plastic deformation, respectively. The diffraction pattern was digitized, recorded and processed by a configuration as described in Ref. [8] to yield a parameter which was used as a measure of the surface roughness. For comparison, surface roughness was also measured by a stylus with a skid tip radius of $5 \mu\text{m}$.

3 Correlation Analysis

Two kinds of correlation analysis^[9] were made: The first one is a cross correlation analysis which calculates the degree of correlation between two diffraction patterns from surfaces with and without plastic deformation; The second is an autocorrelation analysis which computes the necessary shifting length at which the correlation of a diffraction pattern with its shifted image drops to a certain level.

3.1 Cross Correlation

Two images are involved in a cross correlation analysis. For an one dimensional case such as curves $g(x)$ and $f(x)$ as shown in Fig. 1a, the zero shift cross correlation is defined as:

$$C'(g, f) = \int_{-\infty}^{\infty} g(x) \times f(x) dx \quad (1)$$

The normalization of C' yields the cross correlation coefficient:

$$C_c(g, f) = \frac{\int_{-\infty}^{\infty} g(x) \times f(x) dx}{\sqrt{\int_{-\infty}^{\infty} g^2(x) dx \times \int_{-\infty}^{\infty} f^2(x) dx}} \quad (2)$$

The corresponding expression for digitized discrete images is:

$$C_c(g, f) = \frac{\sum_{i=1}^N g(i) \times f(i)}{\sqrt{\sum_{i=1}^N g^2(i) \times \sum_{i=1}^N f^2(i)}} \quad (3)$$

For two dimensional discrete images the expression becomes:

$$C_c(g, f) = \frac{\sum_{i=1}^M \sum_{j=1}^N g(i, j) \times f(i, j)}{\sqrt{\sum_{i=1}^M \sum_{j=1}^N g^2(i, j) \times \sum_{i=1}^M \sum_{j=1}^N f^2(i, j)}} \quad (4)$$

where $f(i, j)$ and $g(i, j)$ are discrete gray levels at coordinates i, j ; M and N are the limits for i and j , respectively. Both M and N are 256 in our experiments.

It is seen that cross correlation counts the common area shared by the two objects whose correlation is to be calculated. For none overlapping objects, the cross correlation coefficient is zero; while for two identical objects, it is unity

3.2 Autocorrelation

Autocorrelation involves only one object. The way to find autocorrelation is by shifting the object a distance δ then use the shifted object as a second image to do cross correlation analysis with the original one (Fig. 1b). Therefore the unified autocorrelation coefficient is expressed as:

$$C_a(g; \delta) = \frac{\sum_{i=1}^N g(i) \times g(i - \delta)}{\sum_{i=1}^N g^2(i)} \quad (5)$$

For two dimensional images the expression for calculating autocorrelation coefficient is:

$$C_a(g; \delta i, \delta j) = \frac{\sum_{i=1}^M \sum_{j=1}^N g(i, j) \times g(i - \delta i, j - \delta j)}{\sum_{i=1}^M \sum_{j=1}^N g^2(i, j)} \quad (6)$$

For a non-periodical image, the correlation coefficient decreases as the shift length increases. Obviously, if the shift length δ is zero, then the autocorrelation coefficient is unity and it becomes zero if the image has been shifted out of the range of the original one. For the same amount of shift, different objects will have different autocorrelation coefficients. This is often used in the other way around for object characterization, i.e. to find the correlation length at which the correlation coefficient drops to a certain level, for example 80%.

If, for some reason, the intensity of laser light source changes when the diffraction patterns were recorded, neither the cross correlation nor the autocorrelation coefficient will be affected if the diffracted intensity is linearly proportional to the illuminating light intensity. The reason is this: the change of illuminating light will multiply the intensity with a factor which will appear in both the numerator and denominator of the above equations and hence is canceled. Therefore the correlation coefficients are independent of the intensity of light source.

4 Experimental Results

The diffraction pattern recorded for each loading level was digitized into a discrete image with 256×256 pixels. Cross correlation analysis was made in two ways: the first was on whole image (two dimensional); the second on the spatially averaged intensity distribution^[4] (one dimensional). We found that the latter, as can be seen from Fig. 2, gave better result. Similar result was obtained for autocorrelation analysis. A comparison shows that autocorrelation analysis on the spatially averaged diffraction intensity distribution gives the best result as far as correlating surface roughness and equivalent plastic strain concerned. In the following, unless otherwise specified, all experimental results are given by the autocorrelation length obtained for spatially averaged intensity distribution of the diffraction patterns.

Fig.3 is a plot of correlation length-versus surface roughness R_a . It can be seen that correlation length of the diffraction pattern correlate with surface roughness R_a well even though it should be emphasized again that the correlation coefficient reflects both the surface profile wavelength and height variation. In the following discussion correlation analysis shall be applied to study roughening isotropy, influence of material properties and stress modes

4.1 Isotropy of Roughening

The isotropy of roughening was studied by finding the surface roughness along different cross sections on specimen surface. It was done by finding correlation coefficients along different cross sections in the diffraction pattern obtained. The procedure is as follows:

1. Record and store a diffraction pattern in computer;
2. Find the center of the pattern by moment method^[4];
3. Select pixels in lines which pass through the center;
4. Compute the angle of each line with respect to a reference line;
5. Calculate autocorrelation length along each direction;

Fig 4 is autocorrelation length along different cross sections for two images, one with and the other without plastic deformation. It is seen that the one without plastic deformation has different correlation length and hence different surface roughness along different directions indicating the surface roughness was not isotropic; While the one with 1.5% plastic deformation becomes nearly isotropic.

This method was applied to aluminum alloys 1100-00, 2024-T3, 6061-T6 and copper alloy C26800-H01 materials under different stress modes and similar results were obtained. It seems safe to say that for the plastic strain range studied metallic material surface roughens isotropically and the initial anisotropy of the surface due to the manufacturing or polishing process will vanish after plastic deformation has reached a certain value. This value may vary depending on the way the specimen surface is polished. It is about 1% for cloth buffer wheel and aluminum powder polished specimens ($R_a \approx 0.05\mu m$) in our experiment.

4.2 Influence of Material Properties

In order to study the influence of material properties on surface roughening, tensile specimens made of four different materials as listed in Table 1 were prepared and tested under the same procedure. Fig.5 is a plot of autocorrelation length at correlation level of 80% versus plastic strain for averaged intensity distribution of different materials. It can be seen that autocorrelation length increases rapidly in the beginning of plastic deformation and almost stops increasing after about 5% equivalent plastic strain. However, there is no apparent difference in autocorrelation length among those four materials for the plastic strain range studied. The scattering of data is partially caused by, we believe, the difference in initial surface roughness and partially by the speckle noise in the diffraction pattern.

4.3 Influence of Stress Modes

In this study only tensile loading was used to generate different principal stress ratios through the use of a specially designed disk specimen^[10] as shown in Fig.6. The specimens were made of copper alloy C26800-H01 (its mechanical properties are listed in Table 1). As the loading angle α changes from -15 to 30 degrees, which is achieved by loading the specimen through different pair of holes along different directions, the ratio of principal stress σ_1/σ_2 at the central area of the specimen varies from -0.577 to -3.732 according to Ref [10]. Fig.7 is a plot of correlation length with equivalent plastic strain for copper alloy material. We see that the surface roughening does not appear to be sensitive to stress modes.

5 Conclusion

Experimental results suggested that correlation analysis can be applied to correlate surface roughness and equivalent plastic strain. For the material tested and experimental apparatus used, it was found that plastic deformation induced surface roughening is isotropic, independent of material properties and of stress mode, but dependent of equivalent plastic strain. It should be noted that the current work only tested fcc materials and their alloys and the tested stress modes are also limited. Further experiments will be focused on the validity of this method on a more general basis including bcc materials.

6 Acknowledgment

We would like to thank the Army Research Office, Engineering Science Division for financial support through contract No.DAA03-88-K-0033.

References

- [1] Miyagawa, M., Azushima, A., "Measurement of plastic deformation by means of a laser beam". Proceedings of the 9th World Conference on Nondestructive Testing, pp.1-7, Australia, Nov., 1979
- [2] Azushima, A., Miyagawa, M., "Measurement of plastic zone around a fatigue crack of a stainless steel structure member with a laser beam". Proceedings of the Joint Conference on Experimental Mechanics, pp.709-713, Hawaii, 1982
- [3] Chiang, F P, Qian, H.H., Don, H.S., "Determination of plastic strain using laser speckles". presented at The 2nd Int Symposium on Plasticity & Its Current Applications, Japan, July, 1989

- 4) Dai, Y.Z., Chiang, F.P., "Estimation of plastic strain by fractal". to appear in proceedings of SPIE. (1332)85, 1990
- 5) Fukuda, M., Yamaguchi, K., Takakura, N., Sakano, Y., "Roughening phenomenon on free surface of products in sheet metal forming". J. Jap. Soc. Technol. Plast., (15)167, pp 994-1002, 1974
- 6) Thomson, P.F., Shafer, B.V., "The roughening of free surface during plastic working". Int. J. of Mach. Tool Des. Res., (22)4, pp.251-264, 1982
- 7) Dautzenberg, J.H., Kals, J.A.G., "Stress state and surface roughness". Advances Technology of Plasticity, (1), pp.186-191, Japan, 1984
- 8) Dai, Y.Z., Chiang, F.P., "Application of scattering theory to plastic strain estimation". presented at the Fourth International Symposium on Nondestructive Characterization of Materials, Maryland, June, 1990
- 9) Rosenfeld, A., Kak, C.A., "Digital picture processing". Academic Press, pp.19, 1982
- 10) Arcan, M., Hashin, Z., Voloshin, A., "A method to produce uniform plane-stress states with applications to fiber-reinforced materials". Experimental Mechanics, (18)4, pp.141-146, 1978

Table 1. Mechanical Properties and Dimensions of Specimens

	Al 1100-00	Al 2024-T3	Al 6061-T6	C26800-H01
Yield Strength (MP _a)	36	340	275	264
Tensile Strength (MP _a)	92	480	310	370
Length (mm)	135	107.5	107.5	107.5
Width (mm)	12.7	12.7	12.7	12.7
Thickness (mm)	2.2	3.2	1.6	1.6



Fig. 1a: Cross Correlation



Fig. 1b: Autocorrelation

Fig. 1 Correlation Analysis

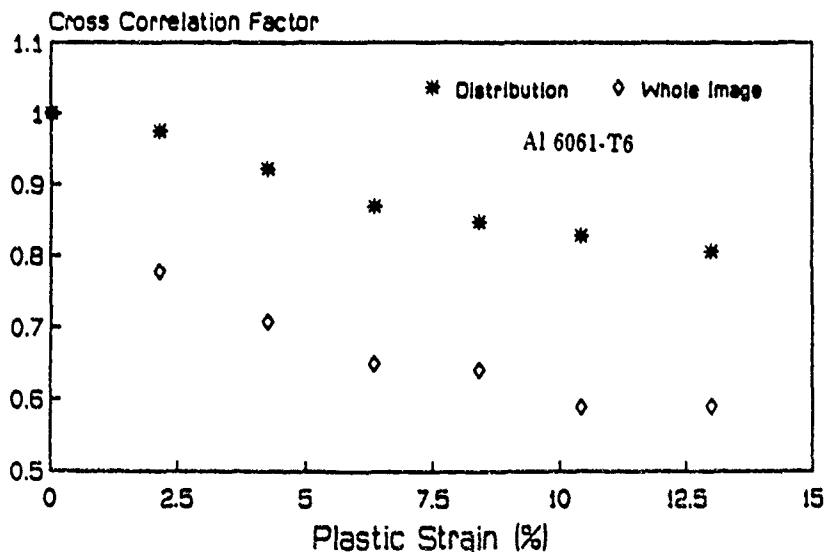


Fig. 2 Cross Correlation for Whole Image and Spatially Averaged Distribution

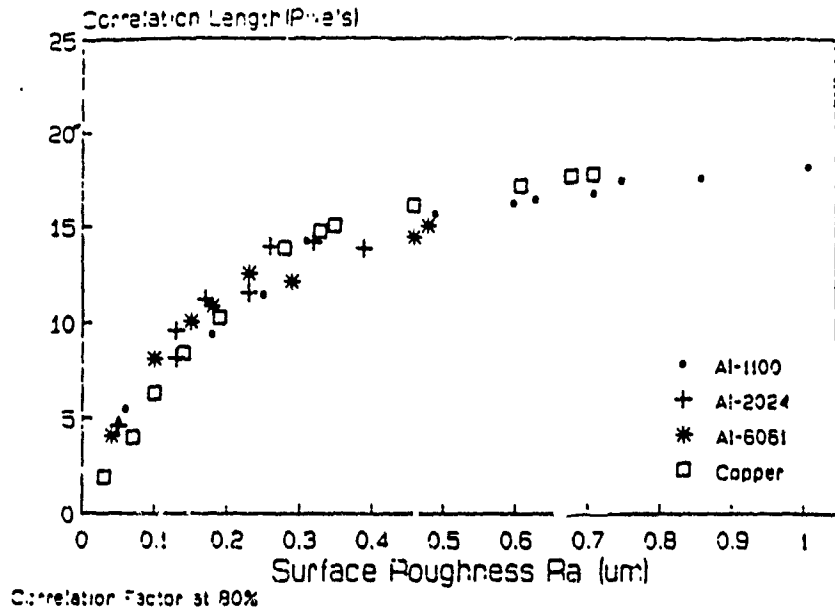


Fig.3 Correlation Length versus Surface Roughness

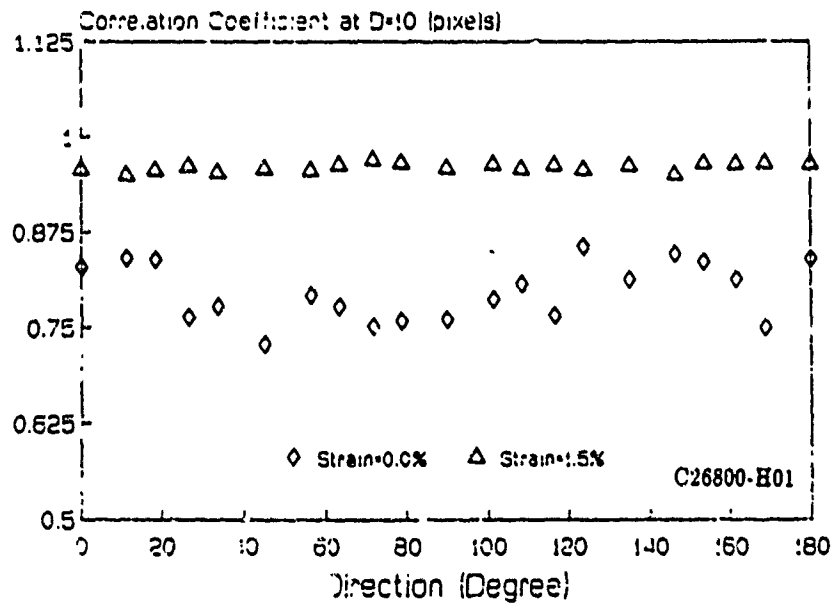


Fig.4 Isotropic Roughening of Material by Plastic Deformation

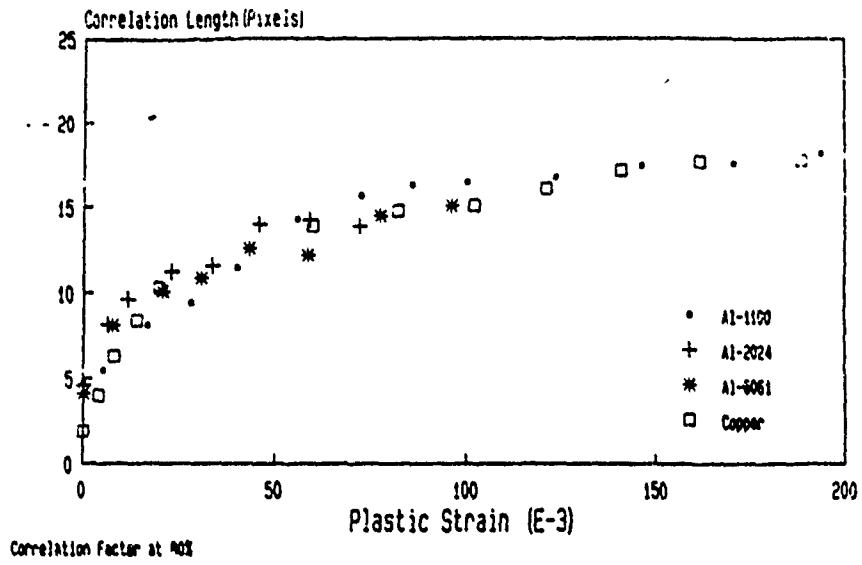


Fig.5 Influence of Material Property on Roughening

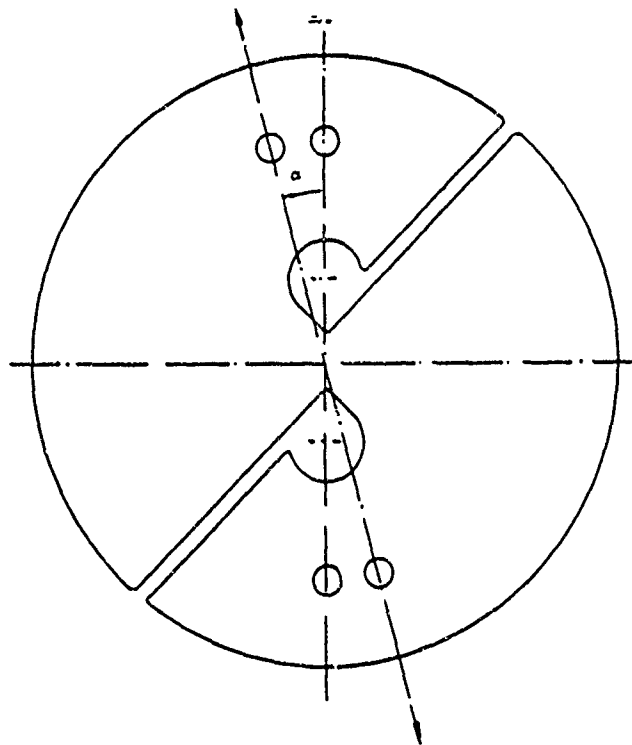


Fig.6 Disk Specimen for Producing Different Stress Modes

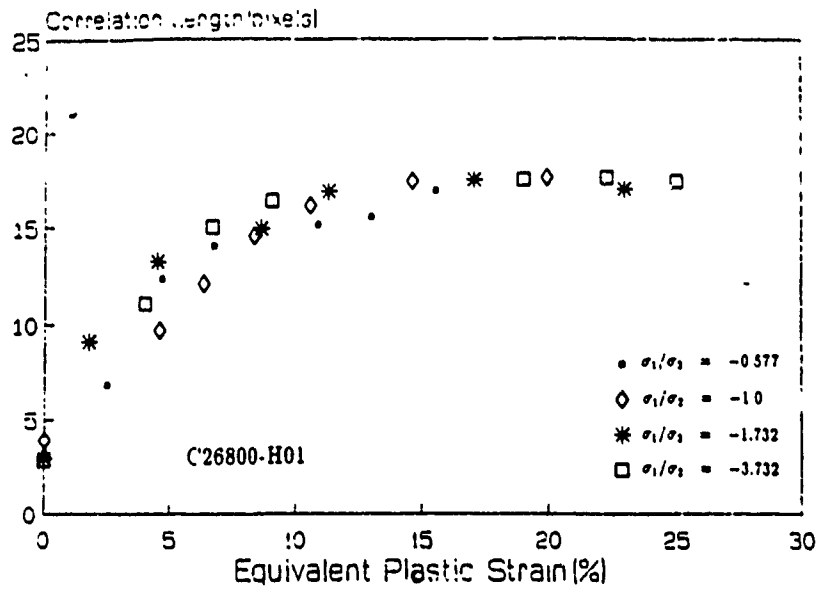


Fig.7 Influence of Stress Mode on Roughening

DAMAGE MONITORING OF COMPOSITE MATERIAL BY IMAGE PROCESSING

by Y.Z. Dai and F.P. Chiang

Like most of engineering materials, composite material deforms plastically when it is subjected to extensive loading. The surface of composite material roughens under a sufficient amount of plastic deformation. The change of surface roughness has been used for the evaluation of plastic strain or fatigue life since the 1970s.^{1,3} In addition to these methods, the techniques developed for surface roughness measurement^{4,5} can be extended to plastic strain measurement as well. However, all these methods either require sophisticated optical set-up or need film developing and fringe contrast analyzing which makes them somewhat impractical.

In this project, instead of measuring surface topography of a specimen directly, we made use of its diffraction pattern. A laser beam was directed to the area of interest on the specimen surface and the diffraction patterns of the surface profile were observed on a piece of ground glass, G, digitized by a digital camera and then processed by a computer (Fig. 1). The light intensity distribution of these diffraction patterns at different plastic strain levels differs from one another (Fig. 2) indicating the feasibility of measuring the plastic strain or monitoring damage development in a mechanical component by the difference in the diffraction patterns. This difference was quantified by the cross-correlation method through an image processing system and utilized as a criterion for damage monitoring.

SILICON CARBIDE REINFORCED ALUMINUM

The silicon carbide reinforced aluminum material used in the test has eight silicon carbide layers, of which six are running perpendicularly to the other two layers as shown in Fig. 3. Uniaxial tensile testing was carried out for both 0-deg and 90-deg directions on an Instron 1332 testing machine. The tensile strength was found to be 310 MPa along the 90-deg direction and 825 MPa along the 0-deg direction, respectively, for monotonic tensile loading.

Y.Z. Dai (SEM Member) is Graduate Student, Department of Mechanical Engineering and F.P. Chiang (SEM Fellow) is Director, Laboratory for Experimental Mechanics Research, SUNY at Stony Brook, NY

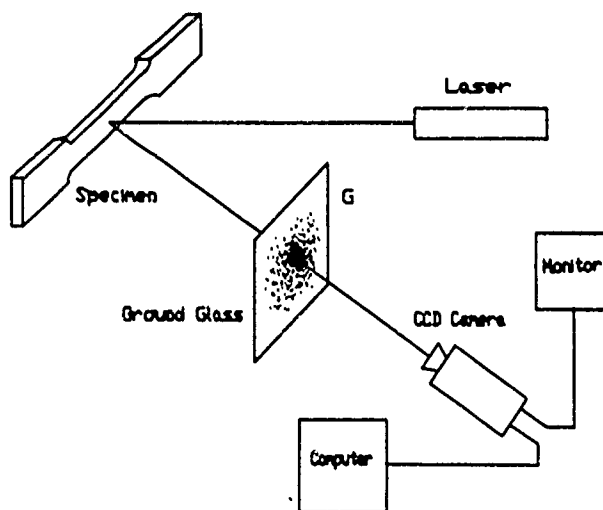


Fig. 1—Optical configuration for damage monitoring of composite material

Surface roughness was measured by Mitutoyo Surftest 420 profilometer with a stylus tip radius of 5 mm and a cut-off frequency of 0.25 mm⁻¹. A plot of surface roughness R_r versus applied load is shown in Fig. 4. There is a drastic change in surface roughness when the stress level reaches about 70 percent of the fracture stress and remains fairly constant until fracture. The sudden change of surface roughness, we believe, is due to the fracture of some very closely located silicon carbide fibers, which manifests itself by emitting pip noise during the loading process.

SPECKLE PATTERN CORRELATION

Correlation techniques have been applied to scientific researches such as pattern recognition and displacement determination. The correlation analysis used in this paper is known as cross correlation.⁶

The cross correlation involves two objects. It refers to the product of these two objects in their overlapping region. For the one dimensional case, the correlation is given by:

$$C_r(g, f) = \int_{-\infty}^{\infty} g(x) \times f(x) dx \quad (1)$$

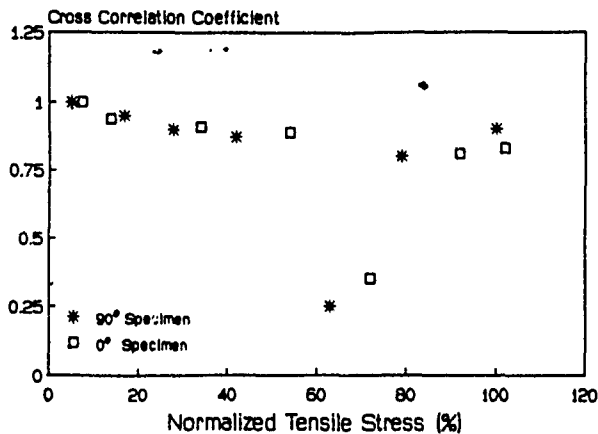


Fig. 5—Correlation for damage monitoring. Correlation factor versus loading along 0 deg and 90 deg directions

with the point where the surface roughness R_s changes drastically (Fig. 4). Thereafter the correlation factor started to increase due to the fact that the specimen surface profile wavelength caused by the silicon carbide fiber becomes larger than the diameter of the laser beam and hence contributes little to the diffraction pattern. This phenomenon, however, is quite useful for the determination of material life: once the correlation factor starts to increase after a big drop, a critical point would have been reached and the usage of this silicon carbide reinforced material should be terminated.

Experimental results obtained so far suggest that cross correlation analysis of diffraction patterns is effective in monitoring damage or material life of silicon carbide reinforced aluminum material. Further work using the proposed cross correlation method will be along the line of plastic strain and fatigue measurement of composite material under high temperature and/or at high strain rate

ACKNOWLEDGMENT

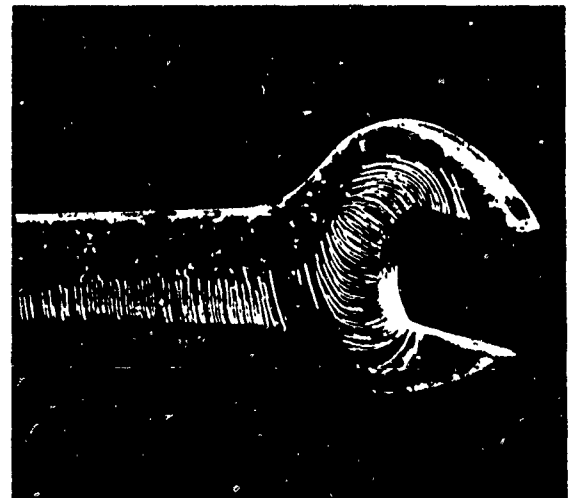
Funding was provided by the Office of Naval Research, Solid Mechanics Division, through contract No N0001482K0566, and Army Research Office, Engineering Service Division through contract No DAAL0388K0083

REFERENCES

1. Sigler, D. and Haworth, W.L., "Strain Measurement by Optical Correlation," *J. of Nondestructive Evaluation*, 2 (2), (1981)
2. Miyagawa, M. and Azushima, A., "Measurement of Plastic Deformation by Means of a Laser Beam," *Ninth World Conf. on Nondestructive Testing, Australia* (1979).
3. Bond, R.E., Beissner, R.E., Lankford, J. Jr and Bradshaw, W.W., "Optical Correlation Evaluation of Surface Deformation due to Fatigue," *Holography and Optical Filtering, NASANSP-299* (1973).
4. Leger, D. and Perrin, J.C., "Real-Time Measurement of Surface Roughness by Correlation of Speckle Patterns," *J. of Optical Soc. of Am.*, 66 (11), (Nov 1976).
5. Fujii, H., Asakura, T. and Shundo, Y., "Measurement of Surface Properties by Using Image Speckle Contrast," *J. of Optical Soc. of Amer.*, 66 (11), (Nov 1976)
6. Rosenfeld, A. and Kak, C.A., "Digital Picture Processing," Academic Press (1982)

Stresscoat®

FOR FAST, ACCURATE & ECONOMICAL STRESS ANALYSIS



Patterns reveal stress points on wrench coated with Stresscoat

With Stresscoat, weaknesses and flaws are quickly detected. This allows engineering test procedures to be intelligently accomplished without costly over-building or delays.

Stresscoat is a scientifically developed brittle coating that is easily applied like a paint to the surface or object to be analyzed. Under the stress of impact, normal or abnormal loading, or testing, stress cracks appear in the coating that can be quickly analyzed to calculate areas of weakness that may impair the object's performance.

Stresscoat can help you achieve fast, accurate and economical stress analysis
Contact Stresscoat Today!

THERE IS ONLY ONE STRESSCOAT!

Stresscoat.

International Sales (714) 946-4977
FAX (714) 946-6267 TELEX USA 704-980
Domestic Sales: Electrix Industries Inc.
P.O. Box F, Grayslake, IL 60030
(708) 740-5500 FAX (708) 740-5506



Computer Speckle Interferometry (CSI)

D.J. Chen and F.P. Chiang¹

Abstract—A fully automatic speckle metrology technique is developed. Two speckle patterns of a specimen, one before and one after specimen deformation, are captured by the video camera. A resultant "double exposure" speckle pattern is obtained by superposing the two digital images. The superposed speckle pattern is segmented into a group of small subimages. A fast Fourier transform (FFT) is applied to each subimage and a computer-generated Young's fringe pattern is obtained. A further FFT of the Young's fringe pattern is performed to analyze the fringe pattern in its spectral domain. Accurate characterization of the local displacement components is obtained by a cardinal interpolation and a maximum searching near the signal hill in the spectral domain. An artificial rigid shifting between the two speckle images is introduced for the cases of very large or very small displacements. The range of measurable displacement is unlimited as long as the two speckle patterns remain correlated.

1 Introduction

Laser speckle interferometry is an ideal technique for the measurement of surface as well as internal deformation.¹ Being nondestructive and remote sensing, it is applicable to high temperature environment, static as well as dynamic deformations.² Its basic process involves speckle recording, specklegram developing, fringe pattern generation and analysis. Several automatic fringe pattern analysis methods have been developed in both the pointwise and the whole-field analysis of the specklegram.³⁻⁶ However, specklegram developing is still an unavoidable procedure in the whole process. The well developed electronic speckle pattern interferometry (ESPI) is an automatic technique for both out-of-plane and in-plane displacement measurements.⁷ Further improvement of the ESPI is achieved in the phase value evaluation of the fringe patterns using a phase-shifting technique.^{8,9} Another straightforward approach is the digital speckle correlation.¹⁰⁻¹² Reliable results of both in-plane displacement components and

displacement gradients have been reported using the white-light speckle approach in a pointwise analysis fashion. Recently, this digital correlation technique has been applied to the laser speckle approach. The digital sampling requirement of laser speckle patterns was performed and reliable results of laser speckle correlation using a displacement-only searching algorithm have been obtained.¹³

In this work, we develop a new approach of automatic laser speckle interferometry. Fully computer-aided processes, including speckle pattern registration, fringe pattern generation and fringe pattern analysis, provide a whole-field survey of displacement components in a pointwise fashion. The system is composed of simple optical as well as electrical setup. It retains all the advantages of the conventional optical laser speckle interferometry and provides an extended range of measurable displacement using a rigid image-shifting technique.

2 General procedures

The system used for data acquisition and image processing is shown in Fig.1. The specimen is illuminated by a collimated laser beam. The image of the object is captured by a video camera. The light intensity of each speckle pattern is digitized into an array of 1024 × 1024 pixels by the analog-to-digital converter (ADC) and then transmitted into the VAX-11/730 mini-computer. The computer controls the data acquisition system, stores the digitized image, performs image processing, and interfaces with the graphic peripherals for a display of the results.

Basic processes of the technique involve data acquisition and image processing. In the data acquisition, two speckle patterns of the specimen, one before and one after the specimen deformation, are captured by the video camera and stored in the computer.

The image processing procedure consists of four stages. First, a resultant "double exposure" speckle pattern is obtained by superposing the two digital images (Fig. 2), and then segmented into a series of small subimages. Second,

¹D.J. Chen is a Graduate Research Assistant. F.P. Chiang is a Leading Professor with the Department of Mechanical Engineering, State University of New York, Stony Brook, NY 11794-2300

a fast Fourier transform (FFT) is applied to each segmented subimage and a computer-generated Young's fringe pattern is obtained. Third, a further FFT is applied to the Young's fringe pattern and the local displacement vector of the specimen is roughly determined by detecting one of the two signal peaks of the spectral amplitude. Moreover, a cardinal interpolation process is applied to the local region around the detected signal hull and an accurate characterization of the displacement components is obtained by a further maximum searching in the interpolated region. An artificial rigid shift between the two original subimages is performed in case the displacement is too large or too small. As a result, the range of measurable displacement magnitude is unlimited as long as the two speckle patterns remain correlated.

3 Theory

3.1 Intensity fringe pattern method

3.1.1 Fringe pattern generation

The first step 2D FFT operation in Fig. 3 shows the fringe pattern generation process. Considering the case that both the subimage size and the displacement gradients are small, one may assume all the points on a subimage undergo the same displacement, say $\vec{d} = (u, v)$, during the deformation. Let's denote the two original subimages by $h(x, y)$ and $h(x - u, y - v)$, respectively. The superposed subimage can be expressed as

$$f(x, y) = h(x, y) + h(x - u, y - v). \quad (1)$$

Denoting the spectrum of the single subimage $h(x, y)$ by $H(\omega_x, \omega_y)$, i.e.,

$$H(\omega_x, \omega_y) = \iint_{\Delta} h(x, y) \exp[-j2\pi(x\omega_x + y\omega_y)] dx dy, \quad (2)$$

where Δ is the subimage region, then the complex spectrum of the superposed subimage is in a form of

$$\begin{aligned} F(\omega_x, \omega_y) &= \iint_{\Delta} f(x, y) \exp[-j2\pi(x\omega_x + y\omega_y)] dx dy \\ &= \iint_{\Delta} [h(x, y) + h(x - u, y - v)] \exp[-j2\pi(x\omega_x + y\omega_y)] dx dy \\ &= H(\omega_x, \omega_y) \{1 + \exp[-j2\pi(u\omega_x + v\omega_y)]\} \\ &= 2H(\omega_x, \omega_y) \exp[-j\pi(u\omega_x + v\omega_y)] \cos[\pi(u\omega_x + v\omega_y)]. \end{aligned} \quad (3)$$

Therefore, the spectral intensity of the superposed subimage is

$$I_s(\omega_x, \omega_y) = |F(\omega_x, \omega_y)|^2 = 4I_h(\omega_x, \omega_y) \cos^2(\pi \vec{d} \cdot \vec{\omega}), \quad (4)$$

where $I_h(\omega_x, \omega_y) = |H(\omega_x, \omega_y)|^2$ is the intensity diffraction halo function and $\vec{\omega} = (\omega_x, \omega_y)$ is the spectral domain coordinate vector. To be distinguished from the spectral domain

(ξ, η) in a later section of fringe analysis, (ω_x, ω_y) is hereafter referred to as the fringe pattern domain. The profile of $I_h(\omega_x, \omega_y)$ has been studied as^{13,14}

$$I_h(\omega_x, \omega_y) = \left[\cos^{-1} \left(\frac{\omega q \lambda}{D} \right) - \frac{\omega q \lambda}{D} \sqrt{1 - \left(\frac{\omega q \lambda}{D} \right)^2} \right]^2$$

where λ is the wavelength of the illuminating light, D is the aperture diameter and q is the imaging distance in the speckle recording setup (Fig. 4).

From Eq. (4) it is seen that the diffraction intensity of the superposed subimage is the intensity halo function modulated by a series of uniformly spaced cosine square fringes whose spacing is inversely proportional to d ($= |\vec{d}|$), and whose direction is normal to \vec{d} . Indeed, if $I_s(\omega_x, \omega_y)$ is displayed on a monitor, one will see dark fringes when

$$\vec{d} \cdot \vec{\omega} = n + \frac{1}{2}, \quad n = 0, \pm 1, \pm 2, \dots \quad (5)$$

and bright fringes when

$$\vec{d} \cdot \vec{\omega} = n, \quad n = 0, \pm 1, \pm 2, \dots \quad (7)$$

Similar to those in the conventional speckle interferometry, these fringes may be customarily referred to as Young's fringes. To distinguish them from those in the later amplitude fringe pattern approach, we denote these fringes as intensity Young's fringes. By noting the fact that the displacement vector is normal to the Young's fringes, Eqs. (6) and (7) can be combined into a single equation as

$$d = |\vec{d}| = \frac{1}{S}, \quad (8)$$

where $S = |\vec{\omega}|/n$ is the fringe spacing.

In regard of the decorrelation effect between the two original speckle patterns, the fringe visibility may be less than uniform. A more general form of the intensity Young's fringe pattern is expressed as

$$I_s(\omega_x, \omega_y) = 2I_h(\omega_x, \omega_y) [1 + V \cos(2\pi \vec{d} \cdot \vec{\omega})], \quad (9)$$

where V (≤ 1) is the fringe visibility. An actual computer-generated intensity Young's fringe pattern from a superposed and mean-eliminated speckle subimage is shown in Fig. 5, where a subimage of size 64×64 pixels is used.

3.1.2 Fringe pattern analysis

Since the quality of the so obtained Young's fringe pattern is poor, direct analysis of the fringe pattern is difficult. Instead, an efficient recognition of the fringe pattern can be obtained in its spectral domain. The second step 2D FFT in Fig. 3 shows the fringe pattern analysis process.

Let's denote the spectrum of the intensity diffraction halo function by $G_{I_h}(\xi, \eta)$, i.e.,

$$G_{I_h}(\xi, \eta) = \iint_{\Omega} I_h(\omega_x, \omega_y) \exp[-j2\pi(\omega_x \xi + \omega_y \eta)] d\omega_x d\omega_y \quad (10)$$

where Ω is the fringe-pattern domain. From Eq. (9), the spectrum of the fringe pattern is

$$\begin{aligned} G_{I_f}(\xi, \eta) &= \iint_{\Omega} I_f(\omega_x, \omega_y) \exp[-j2\pi(\omega_x \xi + \omega_y \eta)] d\omega_x d\omega_y \\ &= 2 \iint_{\Omega} I_h(\omega_x, \omega_y) \{1 + V \cos[2\pi(u\omega_x + v\omega_y)]\} \\ &\quad \exp[-j2\pi(\omega_x \xi + \omega_y \eta)] d\omega_x d\omega_y \\ &= 2G_{I_h}(\xi, \eta) + V^* G_{I_h}(\xi - u, \eta - v) + V G_{I_h}(\xi + u, \eta + v). \end{aligned} \quad (11)$$

It is seen that the spectrum of the intensity fringe pattern is composed of the spectrum of the intensity diffraction halo at $\xi = 0, \eta = 0$ and the two anti-directional shifted spectra of the diffraction halo at $\xi = u, \eta = v$ and $\xi = -u, \eta = -v$. Fig. 6 shows the spectral magnitude of a computer-generated intensity Young's fringe pattern. Since the noise is spread fairly uniformly over the entire spectral domain while the signal is concentrated on the two shifted hills, the signal-to-noise ratio is considerably enhanced in the spectral domain. By detecting the peak position of one of the shifted signal hills, local displacement vector (u, v) between the two superposed speckle subimages can be obtained.

3.2 Amplitude fringe pattern method

3.2.1 Fringe pattern generation

Let us rewrite the complex spectrum of the superposed subimage in Eq. (3) as

$$F(\omega_x, \omega_y) = 2H(\omega_x, \omega_y) \exp[-j\pi(u\omega_x + v\omega_y)] \cos[\pi(u\omega_x + v\omega_y)], \quad (12)$$

and denote the amplitude of the complex spectrum of the single speckle subimage by $A_h(\omega_x, \omega_y) (= |H(\omega_x, \omega_y)|)$. From Eq. (5), the profile of $A_h(\omega_x, \omega_y)$ is

$$A_h^2(\omega_x, \omega_y) = \cos^{-1} \left(\frac{\omega q \lambda}{D} \right) - \frac{\omega q \lambda}{D} \sqrt{1 - \left(\frac{\omega q \lambda}{D} \right)^2}. \quad (13)$$

Similar to the intensity diffraction halo, $A_h(\omega_x, \omega_y)$ can be referred to as the amplitude diffraction halo function. Thus, the spectral amplitude of the superposed speckle subimage is

$$\begin{aligned} A_f(\omega_x, \omega_y) &= 2A_h(\omega_x, \omega_y) |\cos[\pi(u\omega_x + v\omega_y)]| \\ &= \frac{4}{3} A_h(\omega_x, \omega_y) \left\{ 1 + \frac{2}{3} \cos[2\pi(u\omega_x + v\omega_y)] - \frac{1}{18} \cos[4\pi(u\omega_x + v\omega_y)] \right. \\ &\quad \left. + \dots + (-1)^{n+1} \frac{1}{(2n-1)!(2n+1)!} \cos[2n\pi(u\omega_x + v\omega_y)] + \dots \right\}. \end{aligned} \quad (14)$$

Neglecting the higher order harmonic terms in Eq. (14), we obtain

$$A_f(\omega_x, \omega_y) = \frac{16}{3\pi} A_h(\omega_x, \omega_y) \left\{ \frac{1}{4} + \cos^2[\pi(u\omega_x + v\omega_y)] \right\}. \quad (15)$$

Thus, we see that the spectral amplitude of the superposed subimage is the amplitude diffraction halo modulated by a series of uniformly spaced cosine square fringes whose spacing is inversely proportional to d , and whose direction is normal to \vec{d} . On a monitor display, one will see dark fringes when

$$\vec{d} \cdot \vec{\omega} = n + \frac{1}{2}, \quad n = 0, \pm 1, \pm 2, \dots \quad (16)$$

and bright fringes when

$$\vec{d} \cdot \vec{\omega} = n, \quad n = 0, \pm 1, \pm 2, \dots \quad (17)$$

These fringes may be customarily referred to as amplitude Young's fringes. By noting the fact that the displacement vector is normal to the Young's fringes, Eqs. (16) and (17) can be combined into a single equation as

$$d = |\vec{d}| = \frac{1}{S}, \quad (18)$$

where $S = |\vec{\omega}|/n$ is the fringe spacing.

In Eq. (15), the ideal fringe visibility is 2/3. In regard of speckle decorrelation effect, the real fringe visibility may be less than 2/3. A more general form of the amplitude Young's fringe pattern is

$$A_f(\omega_x, \omega_y) = \frac{4}{\pi} A_h(\omega_x, \omega_y) [1 + V \cos(2\pi \vec{d} \cdot \vec{\omega})], \quad (19)$$

where $V (\leq 2/3)$ is the fringe visibility. An actual computer-generated amplitude Young's fringe pattern from a superposed and mean-eliminated speckle subimage is shown Fig. 7.

3.2.2 Fringe pattern analysis

The so obtained amplitude Young's fringe pattern can also be recognized in its spectral domain (cf. Fig. 3). Denote the spectrum of the amplitude diffraction halo by $G_{A_h}(\xi, \eta)$, i.e.,

$$G_{A_h}(\xi, \eta) = \iint_{\Omega} A_h(\omega_x, \omega_y) \exp[-j2\pi(\omega_x \xi + \omega_y \eta)] d\omega_x d\omega_y. \quad (20)$$

From Eq. (19), the spectrum of the amplitude fringe pattern is

$$\begin{aligned} G_{A_f}(\xi, \eta) &= \iint_{\Omega} A_f(\omega_x, \omega_y) \exp[-j2\pi(\omega_x \xi + \omega_y \eta)] d\omega_x d\omega_y \\ &= \frac{4}{\pi} \iint_{\Omega} A_h(\omega_x, \omega_y) \{1 + V \cos[2\pi(u\omega_x + v\omega_y)]\} \\ &\quad \exp[-j2\pi(\omega_x \xi + \omega_y \eta)] d\omega_x d\omega_y \\ &= \frac{4}{\pi} G_{A_h}(\xi, \eta) + \frac{2V}{\pi} G_{A_h}(\xi - u, \eta - v) + \frac{2V}{\pi} G_{A_h}(\xi + u, \eta + v). \end{aligned} \quad (21)$$

Therefore, the spectrum of the amplitude fringe pattern is composed of the spectrum of the amplitude diffraction halo at $\xi = 0, \eta = 0$ and the two anti-directional shifted spectra of the diffraction halo at $\xi = u, \eta = v$ and $\xi = -u, \eta = -v$. Fig. 8 shows the spectral magnitude of a computer-generated amplitude Young's fringe pattern. While the noise is spread fairly uniformly over the entire spectral domain, high signal-to-noise ratio is achieved in the spectral domain. By detecting the peak position of one of the shifted signal hills, local displacement vector (u, v) between the two superposed subimages can be obtained.

3.3 Advantages of amplitude fringe pattern method

We have analyzed both the intensity and the amplitude Young's fringe pattern methods. As a comparison, we realize that the amplitude fringe pattern approach has several advantages over the intensity approach. (i) Fringe domain square operation, which consumes N_x^2 multiplications for an $N_x \times N_x$ subimage, is not required in the amplitude approach, but is required in the intensity approach. (ii) Experiments show that the amplitude approach yields apparent signal hills in the spectral domain of the Young's fringe pattern even without the mean-elimination of the superposed subimage (Fig.9b), while the intensity approach does not provide clear signal hills in the spectral domain without the mean-elimination (Fig.9a). This is because the constant background of the subimage contributes to a large pulse at the origin of the fringe domain. A square operation in the fringe domain greatly magnifies this pulse and results in a higher background noise in the spectral domain. Therefore, the mean-elimination is not a necessary process in the amplitude approach, but it is necessary in the intensity approach. (iii) Most importantly, the resulted spectral domain signal hills from the the amplitude fringe pattern approach are much sharper than those from the intensity approach (see Fig.10). This is because the intensity diffraction halo is the square of the amplitude diffraction halo (Fig.11) and the spectrum of the intensity diffraction halo is the auto-correlation of the spectrum of the amplitude diffraction halo, i.e.,

$$G_{I_A}(\xi, \eta) = G_{A_A}(\xi, \eta) * G_{A_A}(\xi, \eta). \quad (22)$$

where $*$ denotes correlation. Therefore, more precise determination of the local displacement components can be obtained from the amplitude approach.

The amplitude approach also has disadvantages. (i) The fringe visibility of the amplitude fringe pattern is slightly smaller than that of the intensity fringe pattern. (ii) The theoretical amplitude fringe pattern contains unexpected higher order harmonics. However, to the authors' experience, the higher order harmonics are sufficiently weak and the first harmonic is always sufficiently strong for a reliable determination.

Overall, the amplitude approach advances over the intensity approach. We use the amplitude approach in our practical measurement.

3.4 Extended range of measurement using image-shifting

Let's denote the real dimension of the selected subimage by L , and its discrete array size by N_x . First, we consider the case that u is the only nonzero displacement component (Fig.

12a). When the displacement on the subimage approaches $L/2$, the theoretical visibility of the amplitude fringe pattern drops to 1/3. More critically, the two signal hills will move to the boundaries of the spectral domain at $\xi = -N_x/2$ and $\xi = N_x/2$. Obviously, more displacement value will result in aliasing in the spectral domain.¹⁵ Therefore, the range of the measurable single displacement component on the specimen is

$$u = L, (2M),$$

where M is the speckle recording magnification.

Then we consider the general two-dimensional displacement case (Fig. 12b). When $u = L/2$ and $v = L/2$, the theoretical visibility of the amplitude fringe pattern drops to 1/6. More critically, the two signal hills will move to the corners of the spectral domain at $\xi = N_x/2, \eta = N_y/2$ and $\xi = -N_x/2, \eta = -N_y/2$. Thus, the range of the measurable 2D displacement on the specimen is

$$\text{Max}(|u|, |v|) \leq L, (2M). \quad (24)$$

In practice, with regard to both the processing speed and measurement reliability, a subimage size of 32×32 is found to be suitable for most applications. When this subimage size is used, the upper-limit is

$$\text{Max}(|u|, |v|) = \sqrt{6}T/M, \quad (25)$$

where T is the sampling interval of the video camera.

In case that the specimen translation is very large, an intentional close-together rigid shifting between the two speckle patterns can be made before the image segmentation. On the other hand, when the displacement is very small, the signal peak in the spectral domain is also unmeasurable because it is too close to the center hill (Fig.13a). In such a case, we apply a depart-from-each-other rigid shifting between the speckle patterns (Fig.13b). In both cases, the real displacement vector of the subimage can be obtained by subtracting the shifted distant vector from the fine searched displacement vector. Thus, both the upper and the lower limit displacement in the traditional speckle interferometry are removed and the measurable displacement is unlimited as long as the two speckle patterns remain correlated.

4 Experimental results

The performance of the described amplitude fringe pattern method was tested experimentally by an rigid rotation of an aluminum disk. The disk was mounted on a rotational stage with the object surface normal to the rotational and optical axes. The object was illuminated by a laser beam with wavelength of $\lambda = 0.6328 \mu\text{m}$. The sensor array of the camera consists of 1024×1024 pixels, with a real area of

10×10 mm². The speckle patterns were recorded through an objective, 90 mm in focal length, at f-number of $F_1 = 22$. The recording magnification was 1:2.5, with a real area of $625 \times 25 = 25 \mu\text{m}^2$ covered on the specimen. The typical speckle size¹⁴ was $23.55 \mu\text{m}$ and the sampling interval of the sensor was $9.77 \mu\text{m}$.

The rotation was applied in steps, with angular increments of 0.100° , 0.300° , and 0.900° , respectively. Four speckle patterns were recorded, before, between, and after the rotations. The vertical displacement between each two succeeding speckle patterns was evaluated at sixteen points on the horizontal diameter of the disk. The subimages were of 32×32 pixels and the interval between each two adjacent subimages was 64 pixels, with a real separation of 1.563 mm on the specimen.

Since the displacement near the center region of the specimen was very small, a rigid shifting of 2 pixels in the vertical direction between each pair of images was performed before the superposition and segmentation. For the evaluations of the first and the second step rotations, all the 16 subimages were obtained from the superposition of the shifted images. While for the third step rotation, only the 4 subimages near the rotational center were obtained from the superposition of the shifted images, with all the rests obtained from direct (unshifted) superposition of the related images. A rough estimation of the local displacement components was achieved first

by detecting the maximum magnitude in the discrete spectral domain. A cardinal interpolation of the spectral magnitude was then performed in the local region of the maximum discrete coefficient,¹⁵ and a more accurate determination of the displacement components was obtained using a further maximum searching in the interpolated region. The final vertical displacements were obtained by subtracting the shifted distance from the searched vertical displacements. Figure 14 shows the measured vertical displacement distribution on the disk diameter. A least-square linear fitting of the three displacement data sets were performed and a standard deviation of about $0.4 \mu\text{m}$ was obtained in all the three sets. The slopes obtained from the linear fitting are 0.001801, 0.003283, and 0.01579, which correspond to angular rotations of 0.1032° , 0.3028° , and 0.9046° , respectively. The experiment showed that the measured vertical displacement component and the estimated angular rotations agree well with the applied rotations.

5 Conclusion and discussion

A computer speckle interferometry capable of precise measurement of whole-field displacement is developed. The pre-process involves quick image superposition and segmentation. The main process is composed of fringe pattern generation

and analysis. Both the intensity and amplitude fringe pattern methods have been attempted and a preference of the amplitude method is found in practical measurement. A subimage size of 32×32 pixels is found suitable in most applications. A rough determination of displacement components is achieved by detecting the maximum magnitude in the discrete spectral domain, and a precise characterization is obtained by a cardinal interpolation and a further maximum searching near the signal hill. Both the upper and the lower limits of displacement in the traditional optical speckle method have been removed using an image-shifting technique. Good agreement between the theoretical and the measured results is obtained in an experimental verification. Since only one exposure is required at each deformation state, applications of such a technique to vibration and transient problems may be obtained by use of a high speed camera. Further economization of computation may be achieved using an improved FFT algorithm in the fringe pattern generation and analysis, and a hierarchical interpolation process in the spectral domain peak-position searching.

6 Acknowledgments

This work was supported by the Army Research Office through Contract DAA20388K0033 (Scientific Program Officer: Dr. C. L. Anderson) and the Office of Naval Research through Contract N0001482K0566 (Scientific Officer: Dr. Yapa D. S. Rajapakse). These supports are gratefully acknowledged. We would also like to thank Prof. H.S. Don of the Electrical Engineering Department for his constructive criticism and advice.

7 References

1. F.P. Chiang, "A new family of 2D and 3D experimental stress analysis techniques using laser speckles", SM Archives, Vol.3, No.1, pp.27-38 (1978).
2. F.P. Chiang, J. Adachi, R. Anastasi and J. Beatty, "Subjective laser speckle method and its application to solid mechanics problems", Optical Engineering, Vol.21, No.3, pp.379-390 (1982).
3. D.W. Robinson, "Automatic fringe analysis with a computer image processing system", Applied Optics, Vol.22, No.14, pp.2169-2176 (1983).
4. J.M. Huntley, "Speckle photography fringe analysis by the Walsh transform", Applied Optics, Vol.25, No.3, pp.383-386 (1986).
5. D.J. Chen and F.P. Chiang, "Digital processing of Young's fringes in speckle photography", in Optical Testing and Metrology II, C.P. Grover, ed., Proceedings of SPIE, Vol.954, pp.310-

320 (1988).

6. G.T.Reid, "Automatic fringe pattern analysis: a review", *Optics and Lasers in Engineering*, Vol.7, no.1, pp.37-68 (1986/7)

7. C.Wykes, "Use of electronic speckle pattern interferometry (ESPI) in the measurement of static and dynamic surface displacements", *Optical Engineering*, Vol.21, No.3, pp.400-408 (1982)

8. K.C'reath, "Phase shifting speckle interferometry", *Applied Optics*, Vol.24, No.18, pp.3053-3058 (1985).

9. D.W.Robinson and D.C.Williams, "Digital phase shifting speckle interferometry", *Optics Communications*, Vol.57, No.1, pp.26-30 (1986).

10. W.H.Peters and W.F.Ranson, "Digital image techniques in experimental stress analysis", *Optical Engineering*, Vol. 21, No.3, pp.427-431 (1982).

11. T.C.Chu, W.F.Ranson, M.A.Sutton and W.H.Peters, "Application of digital-image-correlation techniques to experimental mechanics", *Experimental Mechanics*, Vol.25, No.3, pp.232-244 (1985).

12. H.A. Bruck, S.R. McNeill, M.A. Sutton and W.H. Peters, "Digital image correlation using Newton-Raphson method of partial differential correlation", *Experimental Mechanics*, Vol. 29, No. 3, pp. 261-267 (1989).

13. D.J.Chen and F.P.Chiang, "Investigation of optimal sampling resolution in digital laser speckle correlation", submitted for publication in *Experimental Mechanics*.

14. R.P.Khetan and F.P.Chiang, "Strain analysis by one-beam laser speckle interferometry. 1: Single aperture method", *Applied Optics*, Vol.15, No.9, pp. 2205-2215 (1976)

15. A.Rosenfeld and A.C.Kak, *Digital Picture Processing*, Vol.1, Academic-Press, New York (1982).

16. F.J.Harris, "Multirate FIR Filters for Interpolating and Decimating", in *Handbook of Digital Signal Processing*, Engineering Applications, D.F. Elliott, ed., pp.173-257, Academic Press, New York (1987).

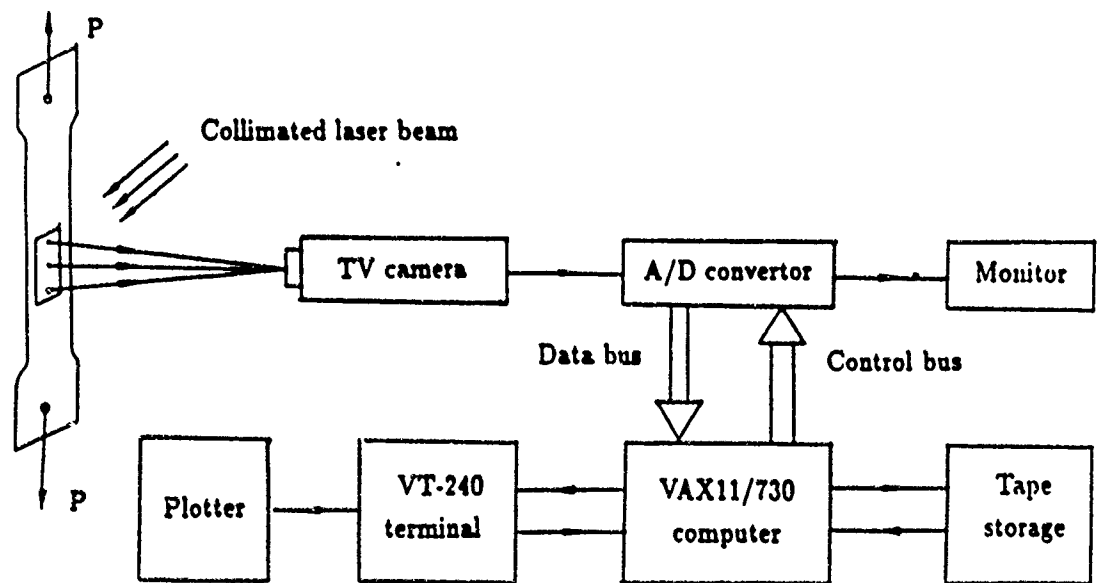


Fig.1 Schematic of data acquisition and image processing system

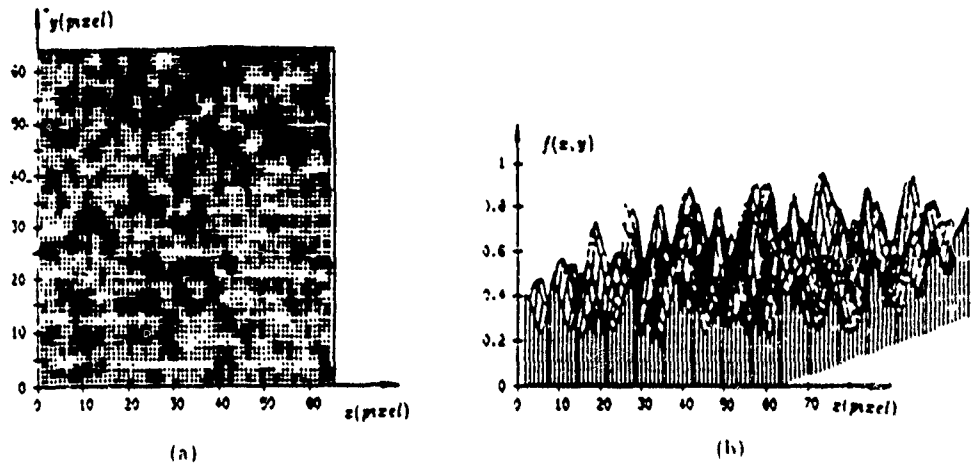


Fig 2 Intensity distribution of superposed speckle pattern:
 (a) half-tone displav, (b) 3D displav

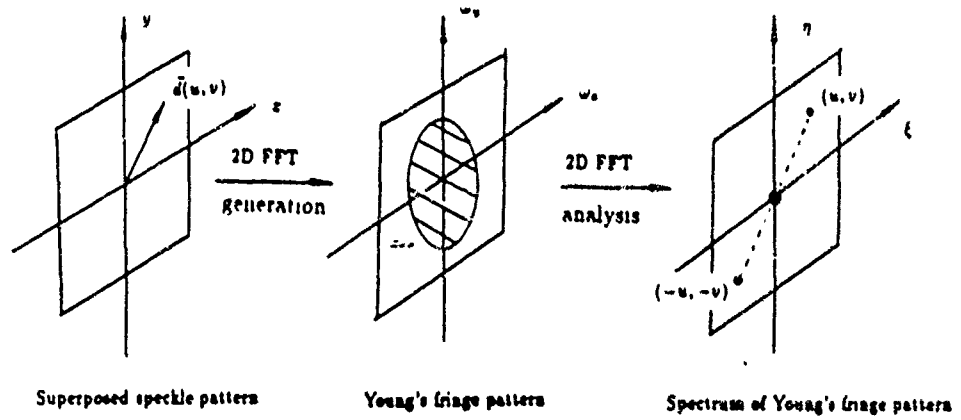


Fig.3 Schematic of fringe pattern formation and recognition

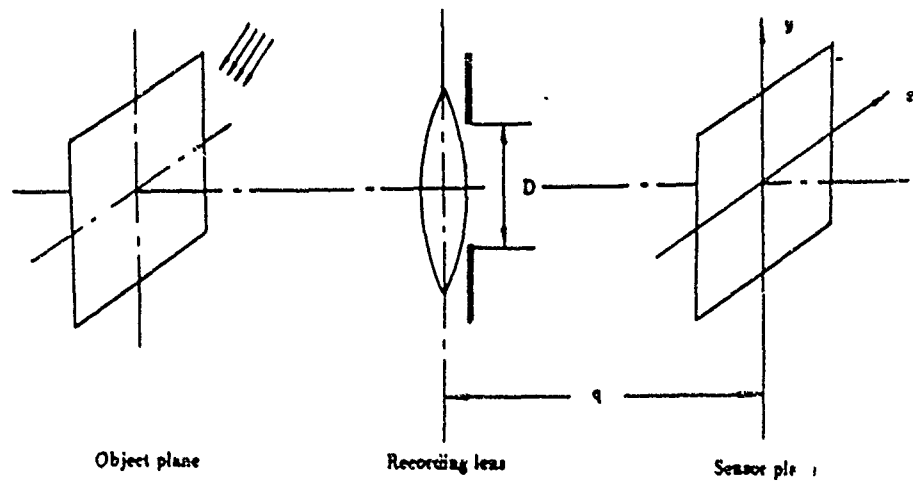


Fig 4 Optical image formation of laser speckle pattern

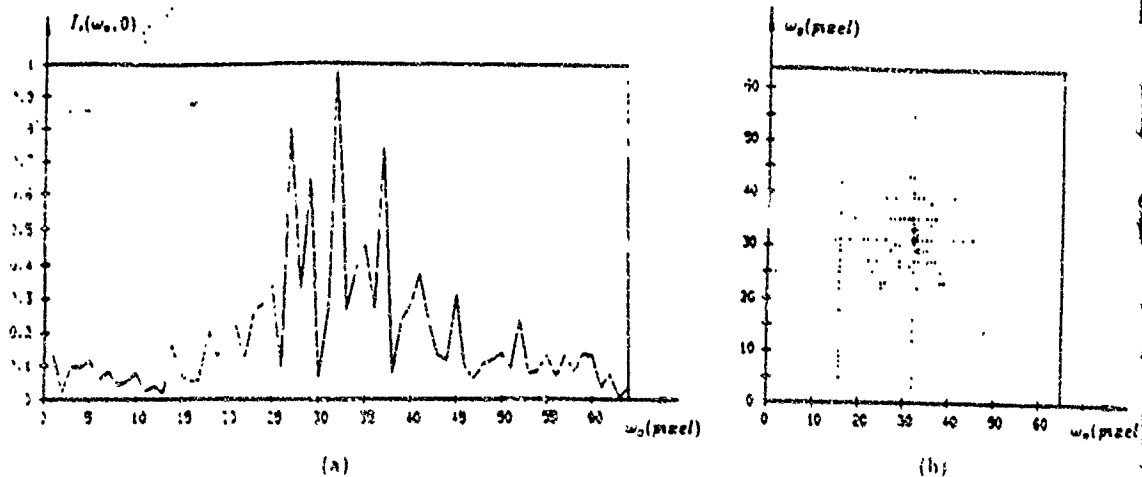


Fig 5 Computer generated intensity Young's fringe pattern. (a) sectional view. (b) halftone display

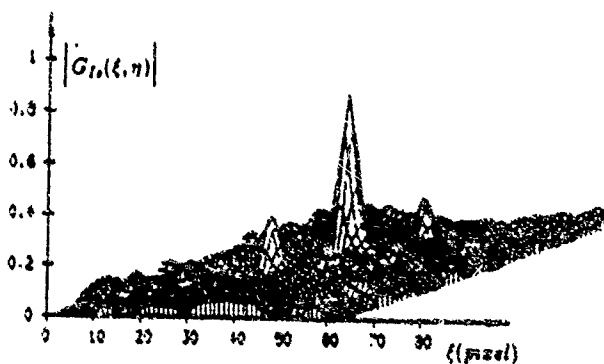


Fig.6 Spectral magnitude of intensity Young's fringe pattern

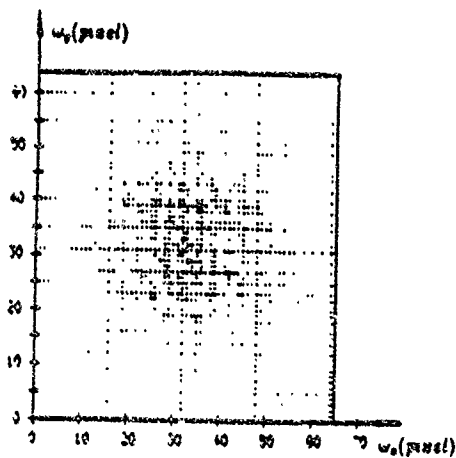


Fig.7 Computer generated amplitude Young's fringe pattern

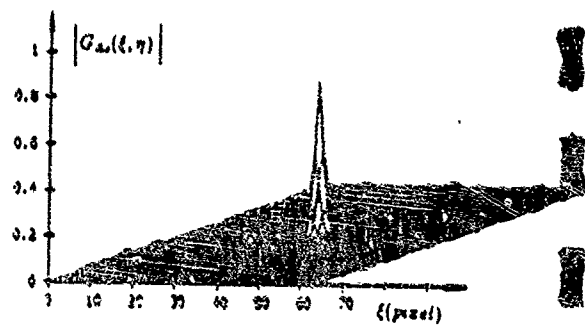


Fig 8 Spectral magnitude of amplitude Young's fringe pattern

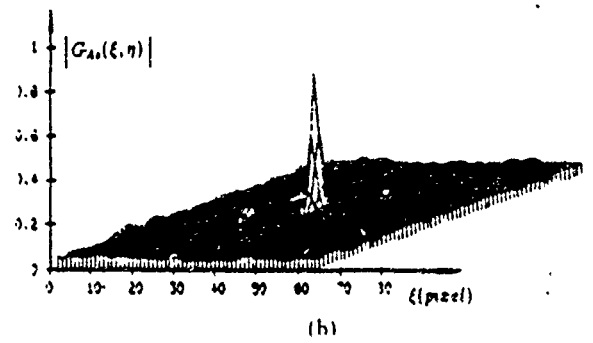
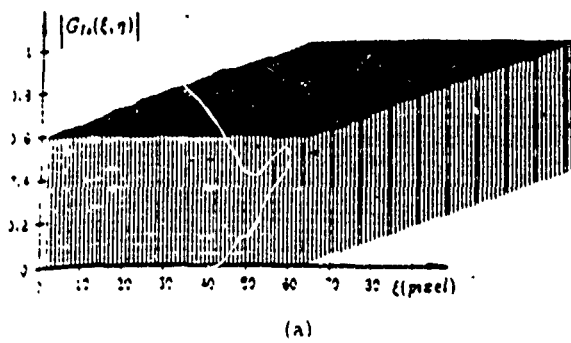


Fig.9 Spectral magnitudes of Young's fringe patterns without mean-elimination.

(a) from intensity fringe pattern. (b) from amplitude fringe pattern.

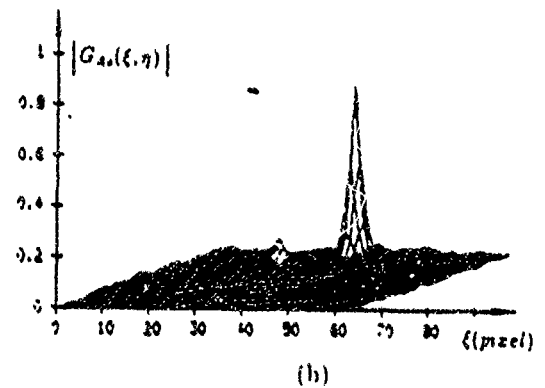
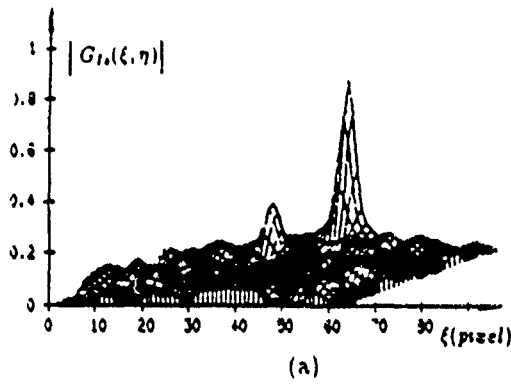


Fig.10 Spectral magnitudes of Young's fringe patterns:

(a) from intensity fringe pattern. (b) from amplitude fringe pattern.

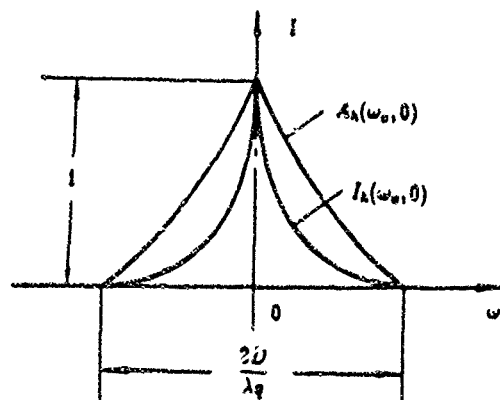


Fig.11 Intensity and amplitude diffraction halos (cross section)

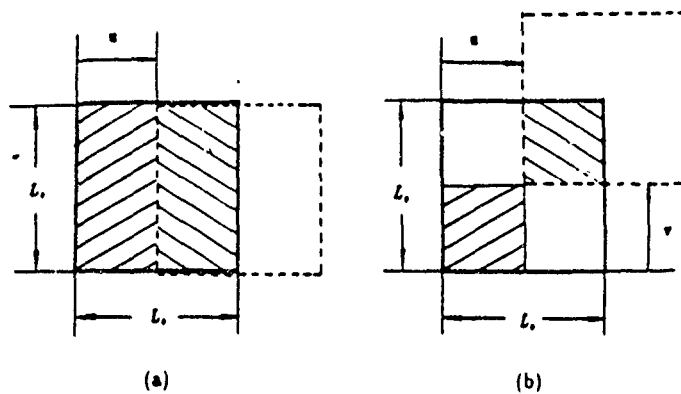


Fig 12 Decorrelation of displacements (▨ correlated area in first speckle pattern, ▩ correlated area in second speckle pattern):
 (a) when $u=L_x/2, v=0$. (b) when $u=L_x/2, v=L_y/2$

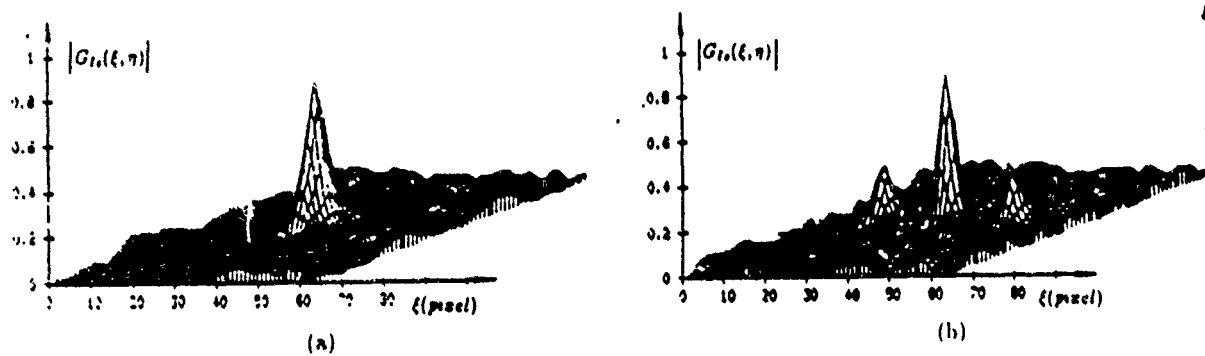


Fig.13 Measurement of very small displacement: (a) spectrum before image shifting.
 (b) Spectrum after image shifting

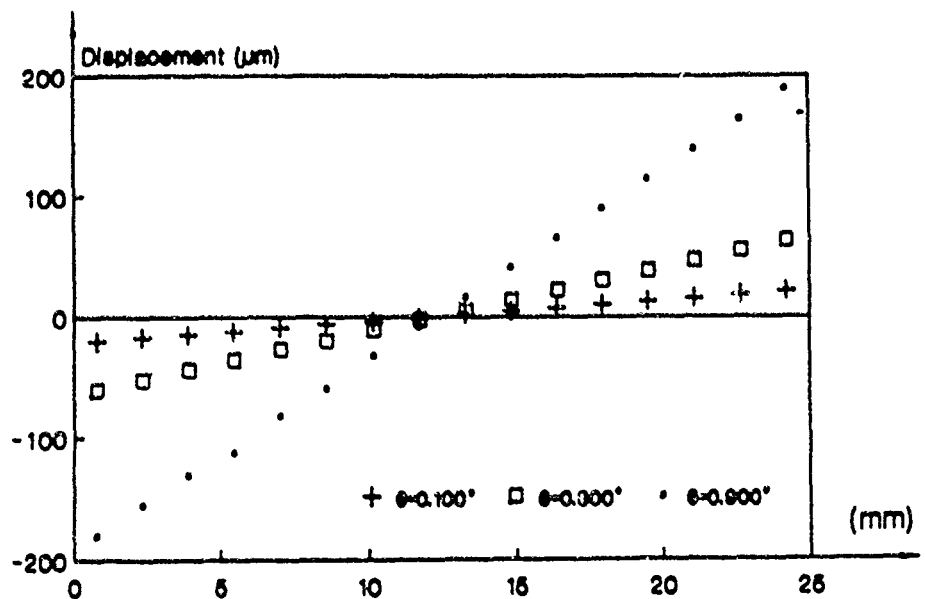


Fig 14 Measured vertical displacements on the horizontal diameter.

A New Approach to Evidential Reasoning

by A Potential Model[†]

Chua-Chin Wang and Hon-Son Don[‡]
Department of Electrical Engineering
State University of New York
Stony Brook, NY 11794
Tel: (516) 652-8404

Abstract

A method of belief combination based on a potential model is proposed, in which a belief function associated with an evidence is modeled as a mass distribution function and each point mass thereon generates a potential on the goal hypothesis. The resulted belief combination at the goal hypothesis is a cumulative integral of the potentials generated by all point masses on the evidences. This model is motivated by the Newton's law of gravity. This model can handle both discrete belief functions and continuous belief functions. Also it has resolved the confictions resulting from either the mutual dependency relationship among evidences or the structural dependency in an inference network due to various combination orders of evidences. A belief combination procedure is given, which can process an arbitrary number of evidences without any confliction.

1 Introduction

Belief combination is one task of evidential reasoning which is referring to combination of relevant evidences for or against hypotheses, and is the core of many rule-based systems that will help people to process pattern recognition, decision making, and diagnosis. There are three major frameworks of belief combination of evidential reasoning in the literature, i.e., the Dempster-Shafer theory of evidence [5], the fuzzy set theory [7], and the Bayesian probability theory [2]. The advantages and weaknesses of these three frameworks have been discussed in [1], [6], and [7]. The application of Shafer's belief function to manage uncertainty of information in a rule-based system has attracted much attention in artificial intelligence research. The Shafer's model uses numerical value in the interval [0, 1] to represent the degree of incompleteness of information. The nonrobustness of this model has been discussed in [4], [8]. Besides this drawback, the basic probability assignment (BPA) of a belief function, as used in all of the previous works, is in a form of discrete type function which can not always provide a precise description of an evidence for all the situations. For instance, if an evidence has the property that the closer it is to the truth, the stronger it is. It can be conveniently modeled by a linear continuous belief function $Bel(\theta) = k \cdot \theta$, where θ is in the interval [0, 1], and k is a constant. We can hardly find any significant thresholds in this kind of evidence to quantize the associated belief function into a discrete form so that it can be handled by the Dempster-Shafer theory, fuzzy set theory, or Hau's method [4].

A conflict problem, [2], [4] occurs when there are many uncertain facts, say 3 or more, to be combined. In a sequential programming style of evidence combination, two of them have to be combined first, then the result will be combined with the other one, and so on. If these facts are partially or totally dependent upon each other, then conflict usually appears. The conflict is that different order of evidence combination will result in different degrees of support to the hypothesis, which is against the intuition

of human reasoning.

Even though lots of effort have been spent on belief combination, the uncertainty management of the continuous belief functions and the conflict due to the dependency of evidences are still not solved. In this paper, we propose a novel method of belief combination based on a potential model. It can achieve the conflict resolution of belief combination resulting from dependency of evidences, also can handle the information aggregation based upon the continuous belief functions.

2 Theory of Potential Model

2.1 Representation of Potential Model

In this subsection, we present the proposed model of the belief function of an evidence and how these functions are combined. The model uses the notion of "mass" and "potential", which is motivated by Newton's law of gravity.

Definition 1: A fact or evidence in a rule-based system is represented by a frame of discernment A , and a belief function associated with A is represented by a probability distribution function $p_A(\theta)$, where $\theta \in [\bar{A}, A]$. The interval $[\bar{A}, A]$ stands for an interval of belief degree from \bar{A} to A , where \bar{A} denotes the complement of A . θ is a numerical value representing a degree of partial truth of an evidence.

For the sake of simplicity, we can use 0 to denote the complement of the authenticity of an evidence, and 1 to denote the truth of the evidence. Hence, the interval $[\bar{A}, A]$ can be replaced by the interval [0, 1]. This type of belief function can be transformed into a dichotomous belief function by assigning two bounds to the interval.

The $p_A(\theta)$ represents the distribution of belief over the interval $[\bar{A}, A]$, which can be continuous or discrete. Furthermore, we can easily see that

$$\int_{\bar{A}}^A p_A(\theta) d\theta = 1 \quad (1)$$

The probability density function, $p_A(\theta)$, can also be considered as a mass distribution, in which each $p_A(\theta)$ is treated as a point mass located at θ . Our model of evidence combination is motivated by the Newton's law of gravity, in which each point mass on an evidence will generate a potential at a distant location on the goal hypothesis. The resulted value of the belief function at a certain location on the goal hypothesis will be the cumulative integral of all the potentials generated by all point masses on all the evidences. The details of our model will be given in the following. We also have to consider the definition of belief combination which refers to the belief conjunction of several evidences supporting the same hypothesis.

[†]This work was supported by the U.S. National Science Foundation under Grant IRI-8710856 and U.S. Army Research Office under Contract DAAL 0338K0033.

[‡]All correspondences should be sent to the second author, Hon-Son Don.

Definition 2: Belief combination refers to the deduction of the belief associated with $((A \rightarrow C) \cap (B \rightarrow C))$ from the belief associated with evidences A and B , respectively, where C is the hypothesis supported by A and B . That is, given two frames of discernment Θ_A and Θ_B , a compatibility relation between Θ_A and Θ_B is the Cartesian product of them, which is represented as

$$\Theta_A \times \Theta_B \rightarrow \Theta_C \quad (2)$$

Assume a belief function of an evidence A is represented by a mass distribution on an interval, e.g. $[\bar{A}, A]$, on a plane. A hypothesis C to be supported by this evidence is located parallelly to the line segment of this interval at a distance. The relationship between the evidence A and the hypothesis C is shown in Fig. 1. We can assume that the effect from a point of the evidence on a point of hypothesis is proportional to its mass magnitude $q_A(\theta)$ of the point of an evidence, and is inversely proportional to the distance r_{AC} between these two points. This assumption of human reasoning is motivated by the Newton's law of gravity. Hence, the notions of the potential law of gravity can be borrowed to model the effect of an evidence supporting a hypothesis. To the desired hypothesis C in Fig.1, the potential $V(\theta)$ at every point θ thereon is affected, respectively, by each point mass $q_A(t)$ on the evidence. This relationship can be described as

$$V(\theta) \propto \frac{q_A(t)}{r_{AC}} = \frac{q_A(t)}{k \cdot r_{AC}} \quad (3)$$

where $r_{AC} = \sqrt{(t - \theta)^2 + R_{AC}^2}$, R_{AC} is the distance between A and C , and k is a constant. Hence, the overall contribution of the evidence, A , to a specific degree of belief θ of the hypothesis C can be formulated as the following,

$$V_C(\theta) = \int_{\bar{A}}^A \frac{q_A(t)}{k \cdot \sqrt{(t - \theta)^2 + R_{AC}^2}} dt. \quad (4)$$

The physical meaning expressed in the last equation is that the mass of every point of an evidence will project its own effect, which is called potential, on every individual point of the hypothesis. As a result, the total effect on a point of the hypothesis is the cumulative sum of the potential from every point of the evidence. This model has some advantages over the previous models. First, it can be extended to the combination of many evidences. In Fig. 2, a model of the combination of n belief functions is shown. In this case, the resulted value of the belief function of the goal hypothesis $V_H(\theta)$ is given as

$$V_H(\theta) = \sum_{i=1}^n V_{H_i}(\theta), \quad (5)$$

where $V_{H_i}(\theta)$ is the potential generated by evidence E_i , $i = 1, \dots, n$. Another advantage is that it provides two parameters, as will be explained later, for the evidence combination, which allow the mutual dependencies among the evidences and the hypothesis to be easily coped with. Furthermore, it can handle both the conventional discrete probability assignment and the continuous probability assignment of a belief function. We will discuss these advantages in more details later.

Definition 3: The distance from an evidence to the desired hypothesis is called its *absolute dependency*, which represents how important this evidence is with respect to the hypothesis.

For example, R_{AC} in Fig. 1 is the absolute dependency from A to C . Two special cases are discussed in the following.

Case 1: $R_{AC} \rightarrow \infty$

This special case implies that the evidence "mass" is placed at a infinite distance from the hypothesis, which contributes zero potential to the hypothesis. According to the above equations,

$$\lim_{r \rightarrow \infty} V = \lim_{r \rightarrow \infty} \frac{q}{k \cdot r} = 0 \quad (6)$$

This is an expected result. Its physical meaning is that the evidence has nothing to do with the hypothesis, i.e., they are totally

independent.

Case 2: $R_{AC} \rightarrow 0$

In this case, intuitively the physical meaning is that the evidence projects itself onto the hypothesis. This can be shown as follows. Referring to Fig. 1, the ratio of the potentials generated by two mass points, $q_A(t)$ and $q_A(\theta)$, where $t \neq \theta$, onto the position θ on C is

$$\begin{aligned} \lim_{R_{AC} \rightarrow 0} \frac{V_t(\theta)}{V_\theta(\theta)} &= \lim_{R_{AC} \rightarrow 0} \frac{\frac{q_A(t)}{\sqrt{(t-\theta)^2 + R_{AC}^2}}}{\frac{q_A(\theta)}{R_{AC}}} \\ &= \lim_{R_{AC} \rightarrow 0} \frac{q_A(\theta)}{q_A(t)} \frac{R_{AC}}{\sqrt{(t-\theta)^2 + R_{AC}^2}} = 0 \end{aligned}$$

This shows that the potential $V_t(\theta)$ is totally determined by the point mass at the position θ of evidence A . Also if there are other evidences with nonzero values of absolute dependencies, their effects on the hypothesis comparatively can be ignored. Hence, the hypothesis is totally dependent upon the first evidence A .

Definition 4: The ratio of the absolute dependencies of two beliefs is called their *relative dependency ratio*, which represents the relative importance of the two evidences with respect to the hypothesis.

For example, the relative dependency ratio of the two evidences E_1 and E_2 in Fig.2 can be expressed as following,

$$\rho_{12} = \frac{R_1}{R_2} \quad (7)$$

Let E_3 be another evidence, then it is easy to show that the following is also true.

Lemma 1: Relative dependency is transitive, i.e.,

$$\rho_{13} = \rho_{12} \cdot \rho_{23}. \quad (8)$$

As an example, consider the case shown in Fig.3, which has two evidences A and B . Suppose the relative dependency ratio ρ_{AB} is given. We can set the value of the absolute dependency R_{AC} based on the desired effect of A onto C . Then the total potential distribution of hypothesis C is given by

$$V_C(\theta) = \int_{\bar{A}}^A \frac{q_A(t)}{k \cdot \sqrt{(t - \theta)^2 + R_{AC}^2}} dt + \int_{\bar{B}}^B \frac{q_B(t)}{k \cdot \sqrt{(t - \theta)^2 + R_{BC}^2}} dt \quad (9)$$

where $R_{BC} = \rho_{BA} \cdot R_{AC}$. Because $V_C(\theta)$ only shows a relative degree of the cumulative potential strength of each point on the hypothesis C , it has to be normalized to become a belief function. The resulted belief function can then be used in another stage of belief combination process in an inference network.

Another factor to be taken into consideration is the length of a belief on which a mass function is distributed. Referring to Fig. 3, if a ratio of the length of the belief A , i.e., from A to \bar{A} , to its absolute dependency R_{AC} is comparatively large, then the potential generated from one end of the evidence A on the other end of the hypothesis C will be small. Apparently, the final potential distribution on C will depend on this length. This length forms the second parameter of the model in addition to the first parameter, i.e., the absolute dependency.

Definition 5: The ratio of the length of the interval, where a belief function distributes on, with respect to the absolute dependency of the belief function is called *belief length ratio* (B.L.R.).

This brings up an argument about what a good belief length ratio is in order that a good evidence combination result can be obtained. We have performed some simulations in order to find a reasonably good range of this ratio, which will be presented in Section 4. The following result can be easily proved.

Lemma 2: If the belief length ratio approaches infinite, the potential value of each point on the hypothesis will converge to a constant.

Proof: The potential value of the hypothesis at θ is given by

$$V_C(\theta) = \lim_{L \rightarrow \infty} \int_{-L}^L \frac{q_A(t)}{k \cdot \sqrt{(t-\theta)^2 + R_{AC}^2}} dt$$

Let

$$W(t) = \frac{q_A(t)}{k \cdot \sqrt{(t-\theta)^2 + R_{AC}^2}} \leq \frac{1}{k \cdot \sqrt{(t-\theta)^2 + R_{AC}^2}}$$

When $t \rightarrow \infty$, $W(t) < \frac{1}{kt}$ as long as $R_{AC} \neq 0$. By elementary calculus, the $V_C(\theta)$ will converge.

2.2 A Generalized Procedure for Belief Combination

Based on the above potential model, we propose in the following a procedure for computing the combination of n evidences, which can resolve the dependency conflict problem in an inference network. Assume there are n evidences to be combined, which are expressed as following,

$$q_i(\theta) \Big|_{inf}^{sup}, \quad i = 1 \dots n$$

where $q_i(\theta)$ represents a mass density function distributed on the interval $[inf, sup]$, and inf represents the complement of sup . The relative dependency ratio between the i th and j th beliefs is denoted by ρ_{ij} .

Procedure of combination of n evidences

- (i) Choose one of the n beliefs as a basic belief, say j , which is supposed to the strongest evidence supporting the desired hypothesis; Set $R_{j,hypo}$ to a desired value.
- (ii) Compute the potential on the hypothesis by

$$V_{hypo}(\theta) = \sum_{i=1}^{i=n} \int_{inf}^{sup} \frac{q_i(\theta)}{k \cdot \sqrt{(t-\theta)^2 + (\rho_{ij} R_{j,hypo})^2}} dt \quad (10)$$

where $R_{j,hypo}$ is the absolute dependency of $q_j(\theta)$ to the hypothesis, and $(sup - inf) = \text{belief length}$.

- (iii) Normalize $V_{hypo}(\theta)$ to a probability density function form.

Note that a rule of thumb of selecting the basic belief is to choose the evidence having the strongest absolute dependency with respect to the hypothesis.

3 Examples of Simulation

In this section, the feasibility and consistency of the proposed potential model are demonstrated by some examples.

Example 1.

This example shows the capability of the proposed model to manage the belief combination of discrete type belief functions. Suppose the uncertainty of a belief function is distributed on the $[0, 1]$ interval, where 0 denotes the complete false, and 1 the complete true. Assume two evidences, A and B , have the relative dependency $\rho_{BA} = 0.6$, which implies the evidence A is the stronger one of the two evidences with respect to the hypothesis C . Let

$$\begin{aligned} p_A(\theta) &= 0.36(\theta - 0.0) + 0.46(\theta - 0.5) + 0.36(\theta - 1.0) \\ p_B(\theta) &= 0.26(\theta - 0.0) + 0.46(\theta - 0.5) + 0.46(\theta - 1.0) \end{aligned}$$

which are illustrated in Fig. 4. Following the procedure given in the previous section, the absolute dependency R_{AC} is chosen as a unit, i.e. $R_{AC} = 1$. Hence, the absolute dependency R_{BC} is determined by Lemma 1,

$$R_{BC} = \frac{1.0}{0.6} = 1.666$$

The hypothesis C is then derived by applying the procedure

of Section 2.2. The resulted credibility and plausibility of belief function $p_C(\theta)$ of the hypothesis C are tabulated in Table 1.

B. L. R.	Cr_C	Θ_C	$1 - Pl_C$
1	0.328649	0.346598	0.324753
2	0.325444	0.360448	0.314108
5	0.325997	0.377129	0.296874
10	0.329302	0.385782	0.284916
20	0.332534	0.391732	0.275735
50	0.335258	0.396312	0.268430
100	0.336332	0.398081	0.265587
200	0.336903	0.399021	0.264076
500	0.337258	0.399603	0.263138
1000	0.337377	0.399801	0.262821

Table 1: the result of proposed model applied to Fig. 4.

Here, we have employed two conventional notations which are the Cr , *credibility*, referring to the lower bound of the belief function defined in [4] and [5], and the Pl , *plausibility*, referring to the upper bound, for the sake of comparison. From Table 1, we can see that when the *belief length ratio* approaches infinite, the belief function of hypothesis C will converge to a stable value, which is predicted by Lemma 2. We can also see that if the B.L.R is small, which indicates that the two ends of belief function are close to each other, then the mutual interaction of point masses is strong, and vice versa. From the simulation result, a reasonable choice of a belief length ratio seems to be about 20 to 100.

In [4], one very crucial problem has been pointed out, which is that in a sequential programming rule-based system, the structural dependency problem can hardly be avoided. Referring to Fig.5, suppose there are three uncertain facts to be combined. In a sequential programming style, two of them have to be combined first, then the result of this combination combines with the third one. Two possible structures of combination are shown in Fig. 6. If all of these three facts are totally independent with each other, there will be no conflict in the final result for the two structures in Fig. 6. But if these three facts are partially dependent upon one another, then conflict will happen. The final results obtained from the two different structures will be inconsistent with each other.

Example 2.

Consider the case in Fig. 5 and the two different structures of combination shown in Fig. 6. Assume

$$\begin{aligned} Cr(M) &= 0.98, \quad Pl(M) = 0.99, \quad \rho_{MK} = \rho_{KN} = 0.5, \\ Cr(N) &= 0.01, \quad Pl(N) = 0.02, \quad \rho_{NM} = 0.1, \\ Cr(K) &= 0.01, \quad Pl(K) = 0.99, \quad \rho_{KM} = 0.9. \end{aligned}$$

Based on Hau's approach [4], the results listed in Table 2 will be obtained.

	Cr_A	Θ_A	$1 - Pl_A$
Case 1	0.006903	0.012857	0.980239
Case 2	0.003064	0.033357	0.963579

Table 2: The results of Hau's approach applied to Fig. 6.

There are some obvious inconsistencies in the results of Table 2. From the assumption, evidence B is the strongest one to support the hypothesis A , the other two evidences C and D are less important than B . According to Table 2, the resulted credibility of case 1 is more than twice of that of case 2. On the contrary, the plausibility, $(Cr_A + \Theta_A)$ of case 1 is only half of that of case 2. These results indicate that Hau's method is easily subject to the combination order of the evidences, which is not consistent with the intuition of human reasoning. To a human, if these three evidences are given, a belief function associated with A should be dominated by B , since the relative dependency ratios of C and D indicate their less influence on A , and because D has a stronger dependency to B than C , B should have the dominant impact on A . Therefore, we conclude that the results given by Hau's method are not consistent with the human reasoning. The reason why the conflict appears in this example is that the mutual dependency relationship will propagate through the inference network. Our

model, on the other hand, can resolve this problem and provide a much more reasonable result than the conventional approaches. As a comparison, we use our model to compute the belief function of hypothesis C in the following.

We use the same data as in the above and the same combination structure as shown in Fig. 5. We apply the proposed procedure given in Sec. 2.2. The results are listed in Table 3.

B. L. R.	Cr_A	Θ_A	$1 - Pl_A$
1	0.352041	0.348043	0.299916
2	0.375143	0.363695	0.261162
5	0.414667	0.391576	0.193757
10	0.441216	0.414554	0.144230
20	0.460307	0.431718	0.107975
50	0.477805	0.441651	0.080544
100	0.485792	0.444197	0.070011
200	0.490254	0.445354	0.064391
500	0.493070	0.446039	0.060891
1000	0.494030	0.446260	0.059702

Table 3: The results of proposed model applied to case of Fig. 5.

Based on Table 2 and Table 3, a comparison is given as follows. Evidence B is the strongest one to support the hypothesis A , the other two evidences C and D are less important than B . To a human, if these three evidences are given, a belief function associated with A should be dominated by B , since the relative dependency ratios of C and D indicating their less influence to A . Also the result should be consistent, despite the arrangement of the inference network. Referring to Table 3, no matter what the belief length is, the credibility Cr_A and plausibility ($Cr_A + \Theta_A$) of hypothesis A are strongly influenced by B , and only slightly perturbed by C and D . This is close to what we expect from the human reasoning.

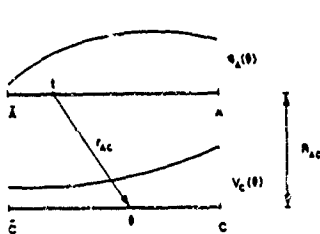


Fig. 1 The relationship between an evidence A and the hypothesis C.



Fig. 2 The relationship between n evidences, $E_i, i = 1, \dots, n$, and the hypothesis H.

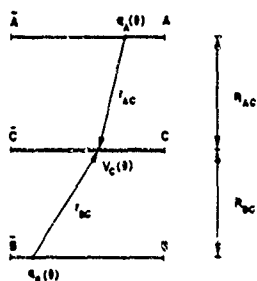


Fig. 3 The relationship between two evidences A, B and the hypothesis C.

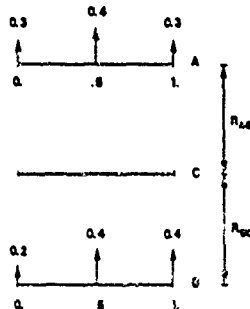


Fig. 4 An example to demonstrate the proposed belief combination model.

4 Conclusion

Our method offers several advantages over previous methods. First, the conflict due to the mutual dependency relationship among the evidences in an inference network is solved. Second, not only the discrete belief functions, but also the arbitrary continuous belief functions can be processed, which has not been explored up to date. The merit of a continuous belief function is that it can represent the vagueness of a human concept more appropriately than a conventional discrete one. Third, given different belief length ratios, different potentials of the goal will be obtained, which provides a variety of options. Fourth, the complexity of computing Dempster's belief function approach or Hau's approach in an inference network is significantly reduced because the proposed model avoid the inherent structural dependency problem of a lattice structure.

References

- [1] Y. Cheng and R. L. Kashyap, "A study of associative evidential Reasoning," *IEEE Trans. Pattern Anal. Machine Intell.*, vol.11, no.6, June 1989.
- [2] R. O. Duda, P. E. Hart, and N. Nilsson, "Subjective Bayesian methods for rule-based inference systems," *Proc. AFIPS*, 1976.
- [3] M. L. Ginsberg, "Nonmonotonic reasoning using Dempster's rule," *Proc. National Conf. Artificial Intell.* 1984, pp. 126-129.
- [4] H. Y. Hau and R. L. Kashyap, "Belief combination and propagation in a lattice-structured inference network," *IEEE Trans. Systems Man Cybernet.*, vol. 20, no. 1, Jan./Feb. 1990.
- [5] G. Shafer and J. Pearl, *Readings in uncertain reasoning*, Morgan Kaufmann Publishers, Inc., San Mateo, California, 1990
- [6] H. Tahani and J. M. Keller, "Information fusion in computer vision using the fuzzy integral," *IEEE Trans. Systems Man Cybernet.*, vol. 20, no.3, May/June 1990.
- [7] J. Yen, "Generalizing the Dempster-Shafer theory to fuzzy sets," *IEEE Trans. Systems Man Cybernet.*, vol. 20, no.3, May/June 1990.
- [8] L. A. Zadeh, "A simple view of the Dempster-Shafer theory of evidence and its implication for the rule of combination," *Fuzzy Set Syst.*, vol. 11, 1983

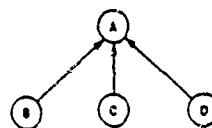


Fig. 5 Case 1: combination of three evidences B, C, and D to support the hypothesis A.

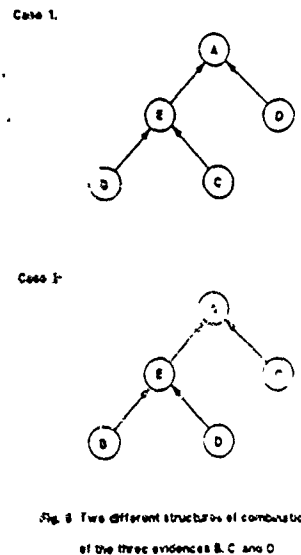


Fig. 6 Two different structures of combination of the three evidences B, C, and D.

Fatigue monitoring by laser speckle

Y.Z. Dai, A. Kato and F.P. Chiang

A non-contact, non-destructive remote fatigue damage monitoring technique is described. This technique employs a laser beam that illuminates the surface of a cyclically loaded specimen, and an image processing system that extracts fatigue damage related information in the speckle pattern scattered from the surface. The spectrum width extracted from the laser speckle pattern increases as a function of the number of loading cycles, indicating the possibility that it may be utilized for monitoring fatigue damage development. The numerical process for obtaining the spectrum width is discussed in detail followed by an experimental demonstration on a tension-tension fatigue study of the aluminium alloy 6061-T6.

Key words: fatigue monitoring; lasers; speckle pattern

Fatigue damage monitoring and fatigue life prediction have long been a challenge to scientists in mechanical engineering practice. A high-sensitivity, more practical technique is still the goal of many experimental engineers. The present work describes an improved spectrum width method employing a laser beam and an image processing system.

Fatigue damage generates slip bands, twinning, intrusion, extrusion and some other topographic features on specimen surfaces where cracks are believed to initiate.¹ All these lead to the roughening of the specimen surface as measured by the surface roughness parameters such as height deviation and correlation length. When illuminated by a narrow laser beam, the fatigue-induced surface roughness in turn causes changes in the reflected and scattered light intensity distributions, that is a laser speckle pattern. This phenomenon renders itself as a possible means for monitoring the development of fatigue damage. This possibility has been explored and methods such as optical correlation,² spectrum width,³ double peak,⁴ and speckle contrast⁵ have been developed to measure plastic-strain-induced and fatigue-induced material damage. While having the advantage of high sensitivity, the optical correlation technique is limited in its practical applications because it is sensitive to rigid-body movement and it needs a time consuming wet photographic process. The spectrum width method has been applied only to plastic strain contour measurements and its behaviour as a function of loading cycles is as yet unknown. The double-peak phenomenon may require a sufficiently large relative rotation between the illuminated areas separated by a wide enough crack. This means that the sensitivity of this method is low and it can be used only after the crack has developed to a substantial size. The speckle contrast method suffers from the fact that its relation with plastic strain is non-monotonic complicating the interpretation of the experimental results.

Compared with the above mentioned methods, the spectrum width technique is advantageous because it is sensitive to surface changes prior to fatigue crack initiation and the experimental set-up is relatively simple. However, the spectrum width method as it was originally introduced has two drawbacks. Firstly, it uses only one cross section of a speckle pattern rendering the experimental results different

with different choices of scanning direction; secondly, the determination of the spectrum width could be quite erratic, especially for severely roughened surfaces because of the inherent limitations of the system used. In order to overcome these two drawbacks, we used a digital camera and an image processing system instead of a single CdS diode and a potential meter. With the help of this apparatus, numerical processing such as smoothing, normalization and optimization can be performed on digitized speckle patterns. As a result, the whole speckle pattern recorded within a certain spatial frequency range is being utilized. In addition, the determination of the spectrum width becomes more accurate and its correlation with the fatigue damage more reliable.

Principle of the method

Fatigue processes may roughen the specimen surface (Fig. 1) by increasing the surface height deviation and decreasing the surface correlation length. When a fatigued specimen surface is illuminated by a laser beam (Fig. 2), the reflected and scattered light intensity distribution, in the form of a speckle pattern, carries surface roughness and hence fatigue damage related information.⁶ Before the speckle pattern is applied to the study of fatigue damage, it is necessary that the relation between a rough surface and the speckle pattern scattered from it be discussed briefly.

Suppose that a laser illuminates a specimen surface in the X - Y plane and the light scattered from the surface is observed on a piece of ground glass on the U - V plane as shown in Fig. 2. If the surface slopes are small and depolarization, multiple scattering and shadowing are negligible, then the electromagnetic field $O(x,y)$ in the X - Y plane immediately to the right of the specimen surface is related to the surface height variation $h(x,y)$ by⁷

$$O(x,y) = RP(x,y) \exp\left\{i\frac{2\pi}{\lambda}(1 + \cos \theta)h(x,y)\right\} \quad (1)$$

where R is the average reflectivity, $P(x,y)$ is the complex field incident on the surface, λ is the wavelength of the incident light and θ is the angle formed by the incident light and the surface normal.

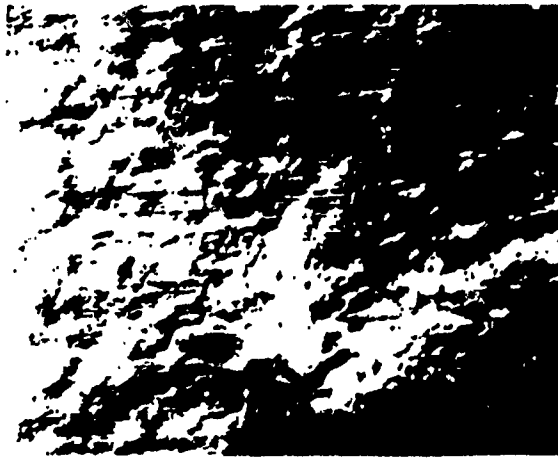


Fig. 1 An optical microscopic photograph showing fatigue-induced roughening (aluminium 6061-T6; 33840 tension-tension loading cycles; 100 \times)

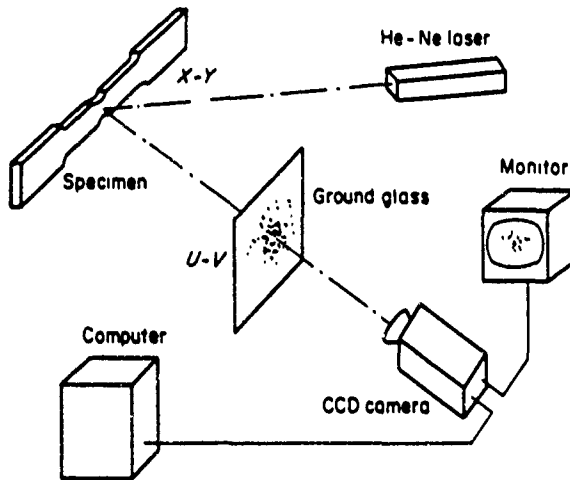


Fig. 2 A schematic illustration of the experimental set-up

If the distance between the specimen surface ($X-Y$) and the observation plane ($U-V$) is great enough that the Fraunhofer approximation can be used, the complex amplitude $A(u,v)$ at the $U-V$ plane is⁴

$$A(u,v) = \frac{C}{i\lambda L} \int_{-x}^x \int_{-y}^y O(x,y) \times \exp\left(-i\frac{2\pi}{\lambda L}(ux + vy)\right) dx dy \quad (2)$$

where C is a complex factor, and L is the distance between the observation plane and the specimen surface. Aside from the factors preceding the integral, Equation (2) is the Fourier transform of the complex field $O(x,y)$, evaluated at the spatial frequencies

$$\begin{aligned} f_x &= u/\lambda L \\ f_y &= v/\lambda L \end{aligned} \quad (3)$$

where f_x and f_y are the spatial frequencies along the X - and Y -directions, respectively, and u , v are the coordinates in the $U-V$ plane.

It should be pointed out that the term spatial frequency used in this discussion refers to the spatial frequency of the

complex field $O(x,y)$ rather than that of the surface profile variation $f(x,y)$. They are approximately the same only for extremely smooth surfaces. This requires $2\pi\sigma/\lambda \ll 1$, where σ is the surface height deviation or root mean square roughness. The polishing process in the present study yields surfaces with $2\pi\sigma/\lambda \approx 1$, whereas the fatigue damage roughens surfaces to $2\pi\sigma/\lambda > 1$.

The light intensity, which is what is being recorded on either a photographic film or a CCD camera, on the $U-V$ plane is

$$I(u,v) = A(u,v)A^*(u,v) \quad (4)$$

where $A^*(u,v)$ is the complex conjugate of $A(u,v)$. On substituting Equation (2) into the above expression, we obtain

$$I(u,v) = \frac{1}{\lambda^2 L^2} \left| \int_{-x}^x \int_{-y}^y O(x,y) \times \exp\left(-i\frac{2\pi}{\lambda L}(ux + vy)\right) dx dy \right|^2 \quad (5)$$

Aside from a constant, this intensity maps the power spectrum of the complex field $O(x,y)$ immediately in front of the illuminated surface. It is a function of, among many other factors, the surface height variation $h(x,y)$. Fatigue damage changes $h(x,y)$ and hence alters its corresponding speckle pattern. For surfaces with low initial surface roughness values (for instance, $\sigma \approx 0.1 \mu\text{m}$), the fatigue damage development tends to increase the high spatial frequency contents while decreasing the low spatial frequency contents in the speckle pattern. This leads to the speckle pattern spreading out and hence the increasing of the spectrum width. This is the fundamental principle of the proposed technique.

Method

Figure 3 shows some experimentally obtained speckle patterns at different stages of fatigue damage development of a cyclically loaded (tension-tension) specimen. The specimen was made of the aluminium alloy 6061-T6 with two edge notches of radius 6.35 mm as shown in Fig. 2. A laser beam was pointed at the central region of the specimen where it was free of the cracks that were generated from the notch roots by the loading cycle. Figures 3(a), (b) and (c) are the speckle patterns of such a point. It can be seen that before the specimen is loaded the speckles are distributed around the central spike and have a preferred orientation owing to the polishing processes (Fig. 3(b)). As the number of loading cycles increases, the speckles spread out and the speckle patterns (Fig. 3(a), (c)) appear to be circular till failure. This implies that, under the test conditions, fatigue damage roughens the surface of the specimen isotropically. The speckle pattern of the fatigue crack (Fig. 3(d)), on the other hand, is not circular. The difference in speckle patterns of points at and away from a fatigue crack may be used to detect the location of the crack. However, for the purpose of fatigue life prediction, only the speckle patterns taken at a certain distance (about 3 mm) away from the crack tip were studied.

As shown in Fig. 3, a speckle pattern contains high-frequency speckles and a low-frequency envelope. Speckles by themselves carry surface feature information only for very smooth surfaces ($\sigma < 0.15\lambda$),⁷ while the envelope can be used for a much wider range of surfaces.⁸ The half spectrum width method, as it was originally proposed, made use of this envelope, but only along a cross section. Therefore its application is limited to orthotropic or isotropic surfaces

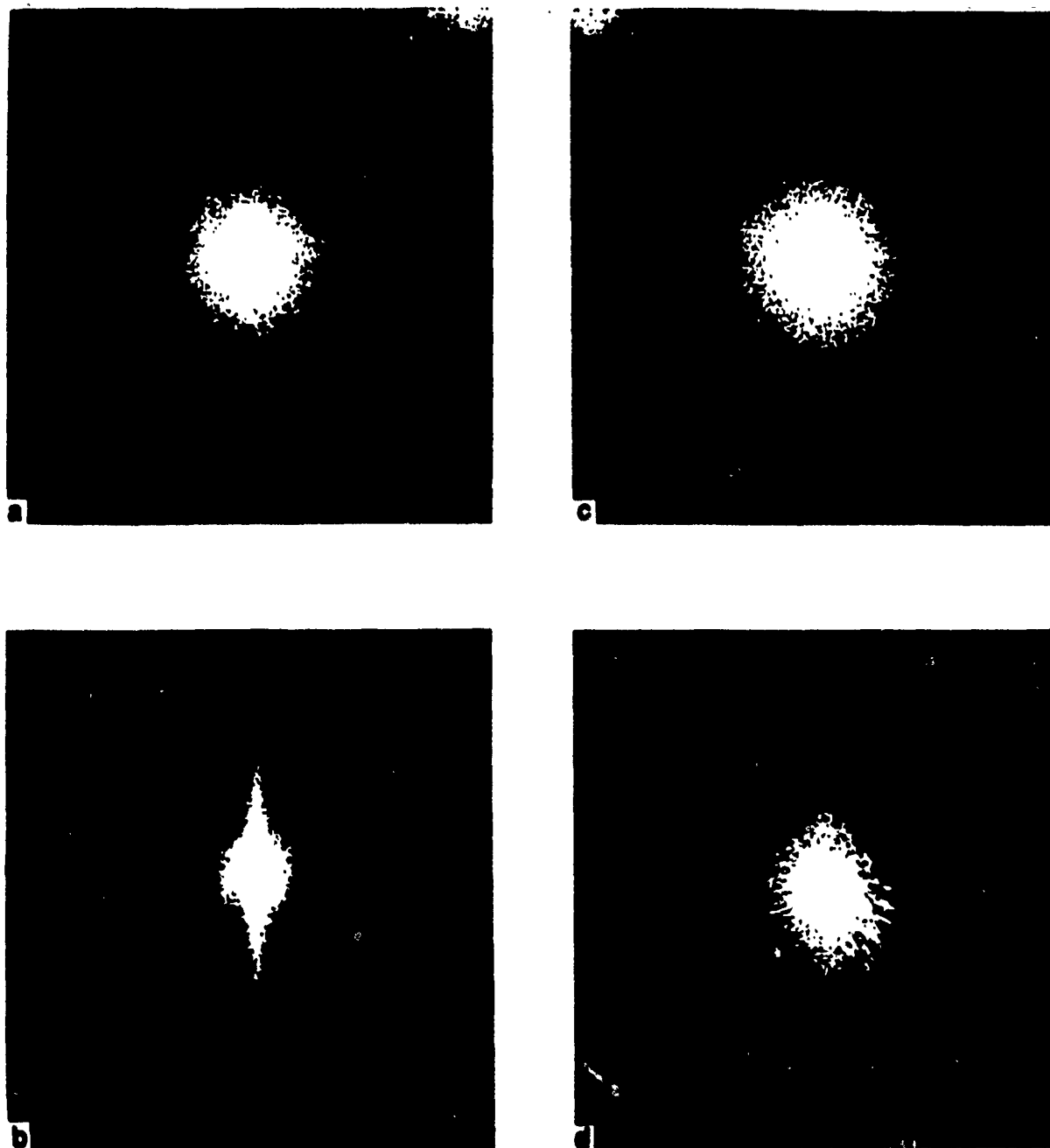


Fig. 3 Speckle patterns at different fatigue damage levels for a specimen of aluminium 6061-T6 with a notch radius of 6.35 mm: (a) $N_f = 23400$, (b) $N_f = 0$, (c) $N_f = 33840$, (d) $N_f = 33840$ at the fatigue crack

Even for isotropic surfaces, the light intensity distribution along a particular cross section of the speckle pattern would vary wildly because of the random nature of the speckles. This scattering makes the half spectrum width vary along different directions. As will be shown in the following discussion, some numerical treatments can greatly depress the speckle noise with the help of an image processing system.

Smoothing of the scattering

The scattering in the light intensity distribution can be reduced by taking the ensemble average, the spatial average, or by smoothing. Among these methods, ensemble averaging is the best because it does not alter the shape of the envelope by factors not related to surface features. However, ensemble

averaging requires taking many speckle patterns from surfaces that are statistically the same. This process needs excessive processing time rendering it impractical.

Because of the symmetry properties of the resulting speckle patterns, the spatial average can be used instead of the ensemble average. This may be done using the following steps: (i) finding the centroid of a speckle pattern, (ii) drawing circles from the centroid with incremental radii, (iii) grouping points according to their distance from the centroid, and (iv) averaging the intensities of the sorted points within each of two neighbouring circles.

Let the digitized intensity distribution of the speckle pattern be $g(i, j)$ and the coordinates of its centroid (\bar{i}, \bar{j}) . The determination of the centroid will be discussed in the following section, then its spatial average \bar{g} from the centroid is

$$\bar{g}(r_i) = \frac{1}{n_i} \sum_{i-1 < r < i} g(i, j) \quad (6)$$

where $r_i = [(i - i_c)^2 + (j - j_c)^2]^{1/2}$, and n_i is the number of points within the ring $i - 1 < r_i < i$.

Such a spatial averaging may not be sufficient to depress the speckle noise, especially at regions near the centroid, since the number of points averaged is small when r_i is small. Therefore, the spatially averaged intensity distribution needs to be further smoothed by a smooth averaging method using the following equation:

$$\bar{g}_s(r_i) = \frac{1}{2k + 1} \sum_{i-k}^{i+k} \bar{g}(r_i) \quad (7)$$

The number of averaging points $2k + 1$ should be chosen with care because it may degrade the sharpness of the central spike in a speckle pattern for smooth surfaces and hence may alter the spectrum width. Other than that, it should be determined by the average size of the speckles such that $2k + 1 > 2k_d$, where k_d is the average number of pixels occupied by a speckle.

For the purpose of demonstrating the effects of the numerical processing, the speckle patterns of a tensile specimen loaded uniaxially to different degrees of plastic deformation were utilized. Figure 4 was obtained by plotting the smoothed average intensity against the radius r_i . The definition of the spectrum width is also given in this figure. It is seen that the curve has become quite smooth making the determination of the spectrum width much easier.

Centroid

The centroid of a speckle pattern is taken to be the specular reflection point from the surface in question. This point is clearly seen for speckle patterns of smooth surfaces ($\sigma < 0.1 \mu\text{m}$) when using a He-Ne laser ($\lambda = 0.6328 \mu\text{m}$). On the other hand, only an area, where the spectral reflection point is supposed to reside, could be observed for rougher surfaces ($0.1 \mu\text{m} < \sigma < 0.5 \mu\text{m}$). For even rougher surfaces ($\sigma > 0.5 \mu\text{m}$), it is next to impossible to determine the centroid without the aid of appropriate instruments. In this study the specular reflection point is determined numerically by a moment method using the following equation:

$$j_c = \frac{\sum_{i=1}^m \sum_{j=1}^n g(i, j) j}{\sum_{i=1}^m \sum_{j=1}^n g(i, j)} \quad (8)$$

and

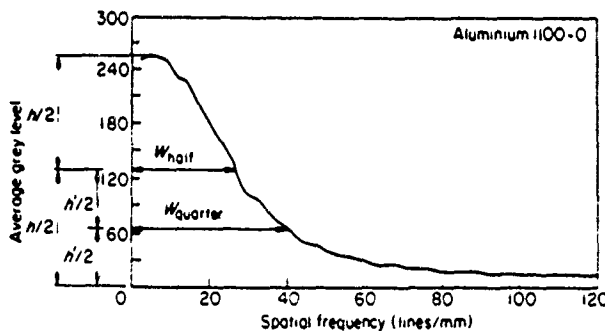


Fig 4 Smoothed intensity distribution and definition of spectrum width

$$j_c = \frac{\sum_{i=1}^m \sum_{j=1}^n g(i, j) j}{\sum_{i=1}^m \sum_{j=1}^n g(i, j)}$$

where m and n are the number of pixels along the i and j directions, respectively. They both took the value of 256 for the uniaxial tension test.

Generally speaking, the moment method does not yield an accurate result because of the limited recording range of the spatial frequencies. Suppose the light intensity distribution of a speckle pattern assumes a conical shape and thereafter its cross section across the centroid is triangular. Two such triangles I and II for the same recording system are shown in Fig. 5. The distance b is the maximum length that could be recorded by the system. It is related to the maximum recording spatial frequency by Equation (2). We discuss two cases:

- 1) If the centre lines cc' for both triangles coincide, respectively, with the centre lines cc' of the recording lens, then the moment method gives the correct coordinates of the centroid.
- 2) More often than not, they do not coincide with each other. For the case of triangle I, the moment method still yields the correct results as long as the whole triangle remains in region $-b$ to b . However, for the case of triangle II, the moment method does not give the correct location of the centroid.

In order to overcome this problem, an approximate centroid was obtained by the moment method, and spatially averaged intensity distributions were computed for a number of points around the approximate centroid. Then the summation of the first ten average intensities for each distribution was taken. The point corresponding to the maximum summation of the first ten average intensities was taken as the centroid. Figure 6 shows a comparison of averaged intensity distribution curves computed from different centre points. Curve 1 was based on the approximate centroid calculated by the moment method and the other four curves were computed at points that were three pixels, equivalent to 4.5 lines/mm of spatial frequency, away from the approximate centroid in four different directions. It is seen that for this specific case the distributions are quite different and this difference could affect the measurement of the spectrum width if the above procedure is not followed. Point 4 was chosen as the centroid based on the method described above. Points 1, 2, 3 and 5 all violate the assumption that the intensity (average grey level) should be a maximum at the zero spatial frequency.

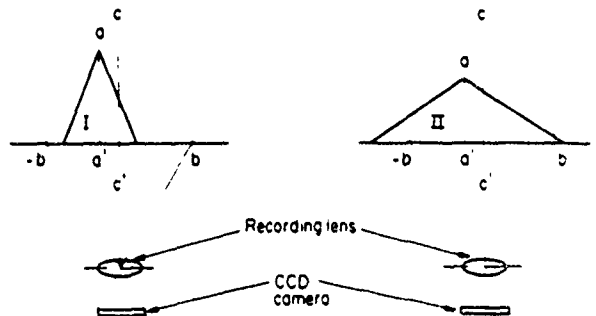


Fig 5 Effect of bandwidth of recorded spatial frequency

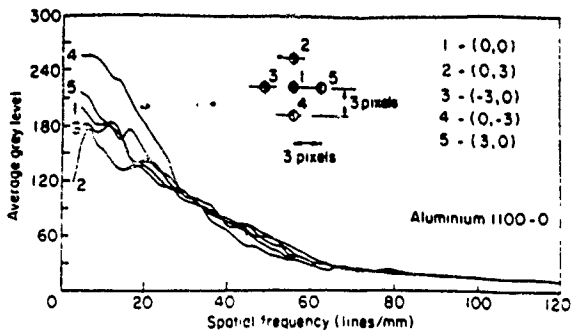


Fig 6 Effect of different centroids on the intensity distribution

Normalization

Figure 7 is a plot of the average intensity distributions at different levels of the plastic deformation. It is seen that for distribution curves corresponding to a large plastic strain, such as curve 5 with 10.42% plastic strain, it becomes difficult to determine the spectrum width because the curve is quite flat.

This problem was solved by normalizing the intensity distribution by taking the average of the first ten intensities from the centroid of the distribution. The reason for taking the average of the first ten intensities instead of the intensity at the centroid is again to reduce the error caused by speckle. Denoting this average by \bar{g}_{10} , then

$$\bar{g}_{10} = \frac{1}{10} \sum_{i=1}^{10} \bar{g}(r_i) \tag{10}$$

and the averaged light intensity is normalized by the following expression

$$\bar{g}'(r_i) = 256 \bar{g}_i(r_i) / \bar{g}_{10} \tag{11}$$

Thus normalized, the curves in Fig. 7 were replotted as shown in Fig. 8. As can be seen, the determination of the spectrum width has become easier and more accurate.

Figure 8 also shows that the portion of the curve at high spatial frequencies is smoother than that at low spatial frequencies. The reason is that towards the higher spatial frequencies the radii are becoming larger and covering more points and the average is taken of more and more points. This implies that the spectrum width defined on the lower portion of the curve, such as the quarter spectrum width, in general, offers a better signal-to-noise ratio than that defined on the higher portion, for example the half spectrum width (Fig. 4).

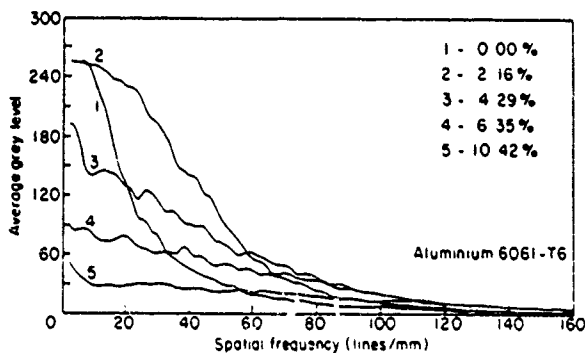


Fig 7 Intensity distributions at different plastic deformation levels

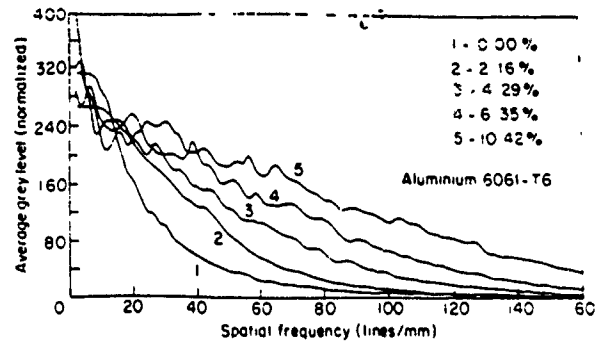


Fig 8 Normalized intensity distributions at different plastic deformation levels

Results

Tension-tension fatigue tests were carried out on specimens made of aluminium alloy 6061-T6 (tensile yield strength $\tau_y = 465$ MPa). The specimens were cut from stock sheet along the rolling direction and their surfaces were polished using a cloth buffer wheel with an aluminium dioxide polishing powder to an initial surface roughness of $\sigma = 0.1 \mu\text{m}$ on the area of interest. Two symmetrical semicircular notches were cut on each side of the specimen. The radii of the notches were 1.5875 mm, 6.35 mm and 9.525 mm for the three specimens, respectively. The dimensions of the specimens are given in Fig. 9. The thickness of the specimens was 3.175 mm. The specimens were fatigued in a testing machine under a stress-controlled cyclic tension-tension loading condition. The maximum nominal tensile fatigue stress was $\tau_{max} = 372$ MPa and the minimum nominal tensile fatigue stress was $\tau_{min} = 46.5$ MPa, which are 80% and 10%, respectively, of the tensile strength of the tested material. The mean cyclic loading stress was $\tau_m = 209.25$ MPa and the magnitude of the cyclic loading was $\tau_c = 162.75$ MPa. The frequency of the cyclic loading was 6 Hz.

A set-up, as schematically illustrated in Fig. 2, was used to demonstrate the applicability of the technique. A 15 mW He-Ne laser with a beam radius of about 1 mm was used as the illuminating light source. A high-speed digital camera was used for digitizing the laser speckle patterns at different loading cycles. The speed of the camera was set at 200 frames per second and its resolution at 238×192 pixels, each varying from 0 to 64. The laser beam was directed at the centre portion of the specimen between the two notches. Because the speckle pattern change induced by the elastic response is negligible compared with that caused by the fatigue damage,

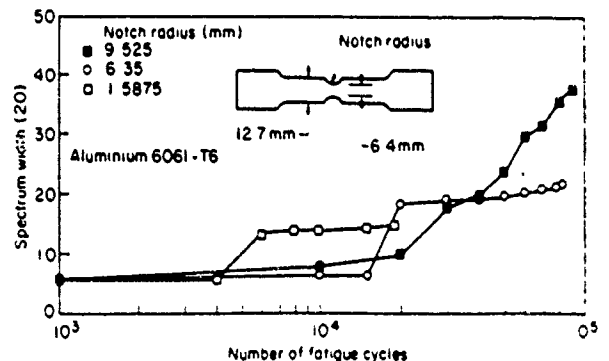


Fig 9 Spectrum width plotted against fatigue life for specimens with different notch radii

the speckle patterns were recorded while the specimen was being cyclically loaded without stopping the test machine. Several speckle patterns were recorded at different loading cycles. The digitized speckle patterns were later processed on a computer numerically by the procedures described in the previous sections.

Figure 9 shows some experimental results for the aluminium specimens of different notch radii. The spectrum width was defined at the point where the intensity was 20 (the maximum intensity was 64). It is seen that the spectrum width does not change much at early stages of cyclic loading. However, it jumps drastically when the specimens have been cyclically loaded to about one quarter of their total fatigue life. This sudden change in spectrum width reflects the drastic surface roughness change that may be the result of the development of slip bands. It also shows that the smaller the notch radius, the earlier the drastic change in the spectrum occurs. This may be because of the fact that a smaller notch radius results in a greater stress concentration and hence earlier slip bands and fatigue crack development. It is interesting to note that for small enough radii, the spectrum width becomes almost constant after the drastic change at about 25% fatigue life, implying that the surface roughness does not change much afterwards. In contrast, for specimens with a sufficiently large notch radius, the spectrum width increases constantly with the number of loading cycles.

Conclusions

An improved spectrum width method has been developed and shown to be effective in a fatigue-induced surface roughening study and in fatigue life monitoring. It is found, for specimens made of aluminium alloy 6061-T6 under tension-tension cyclic loading, that the surfaces of the specimens are roughened isotropically. For fatigue life monitoring, the spectrum width was observed to increase with the number of loading cycles. The increasing rate strongly depends on the dimension of the notch radius, which results in different magnitudes of the stress concentration. A drastic change in spectrum width occurs at about 25% of the fatigue life, which may be the result of the development of persistent slip bands. This drastic change in the spectrum width may be used as a criterion for the determination of the fatigue life.

This technique, with the merits of being a non-contact and non-destructive method, has the potential of being applied to homogeneous and composite material tests at high temperatures. It is believed that such studies, which can not

be investigated by most of the conventional methods, will be of major importance in the years to come.¹⁰

Acknowledgement

The financial support provided by the US Army Research Office, Engineering Science Division through contract No DAALC388K0083 is gratefully acknowledged.

References

1. Kocanda, S. *Fatigue Failure of Metals* (Sijthoff & Noordhoff, Warsaw, 1978)
2. Chuang, K.C. 'Application of the optical correlation measurement to detection of fatigue damage' *Mater Eval* (1968) pp 116-119
3. Azushima, A. and Miyagawa, M. 'Measurement of plastic zone around a fatigue crack of a stainless steel structure member with a laser beam' *Proc of Joint Conf on Experimental Mechanics, Hawaii, 1982* (Society for Experimental Mechanics, 1982) pp 709-713
4. Pernick, B.J. and Kennedy, J.R. 'Fatigue crack detection on rough surfaces by optical Fourier analysis' *Appl Opt* 24 12 (1985) pp 1764-1769
5. Chiang, F.P., Qian, H.H. and Don, H.S. 'Determination of plastic strain using laser speckles' *Proc of 2nd Int Symp on Plasticity and Its Current Applications, Japan, 1989*
6. Dai, Y.Z. and Chiang, F.P. 'Application of scattering theory to plastic strain estimation' *Proc of Fourth Int Symp on Nondestructive Characterization of Materials, Annapolis, Maryland, 1990* (Plenum, New York) p 58
7. Goodman, J.W. 'Statistical properties of laser speckle patterns' *Topics in Applied Physics* Ed J.C. Dainty (Springer, Berlin, 1975) Vol 9 pp 9-75
8. Goodman, J.W. *Introduction to Fourier Optics* (McGraw-Hill, New York, 1968)
9. Coffin, L.F. 'Some perspective on future directions in low cyclic fatigue' *Low Cycle Fatigue, ASTM, STP 942* (American Society for Testing and Materials, 1988)
10. Manson, S.S. 'Future directions on low cycle fatigue' *Low Cycle Fatigue, ASTM, STP 942* (American Society for Testing and Materials, 1988)

Authors

Y.Z. Dai and F.P. Chiang are with SUNY at Stony Brook, NY 11794, USA. A. Kato is with the Chubu University in Japan, and was a visiting scholar at SUNY when the work reported here was done.

Scattering from plastically roughened surfaces and its applications to strain assessment

YuZhong Dai

Brown & Sharpe Manufacturing
Company

Department of Research & Development
North Kingstown, Rhode Island 02852

F. P. Chiang

SUNY at Stony Brook
Department of Mechanical Engineering
Stony Brook, New York 11794

Abstract. Statistical properties of plastically roughened metallic surfaces are studied together with the relation between the angular distribution of the reflected-plus-scattered intensity from such a surface and the plastic deformation it has experienced. The height probability density distribution of the plastically roughened surfaces is shown to be Gaussian. The surface autocorrelation function is neither Gaussian nor exponential, even though the former fits the experimental data better. The combination of a scalar scattering theory and the plasticity-induced surface roughening empirical relations yields some new expressions. These expressions relate scattered light with plastic deformation for surfaces with Gaussian or exponential autocorrelation. The applicability, convergence, and limitation of the thus obtained expressions are also studied. A theoretical-numerical analysis provides background to some existing strain measuring methods, such as the maximum intensity, half width, and correlation coefficients. The present approach is verified favorably by some experiments on aluminum alloys and is in support of the experimental results in literature.

Subject terms: scattering; plasticity-induced roughening; surface roughness; statistical parameters; strain assessment.

Optical Engineering 30(9), 1269-1276 (September 1991)

CONTENTS

- 1 Introduction
- 2 Properties of plastically roughened surfaces
 - 2.1 Statistical characteristics
 - 2.2 Plastic strain and surface roughness
- 3 Plastic strain caused scattering
 - 3.1 Scalar scattering theory
 - 3.2 Plastic strain and scattering
 - 3.3 Applicability
- 4 Analysis of strain-measuring methods
 - 4.1 Intensity-related methods
 - 4.2 Correlation coefficients
 - 4.3 Half-width method and its verification
- 5 Conclusion
- 6 Acknowledgment
- 7 References

1. INTRODUCTION

Plastic deformation roughens the surface of metallic materials by increasing its height variation and decreasing its profile wavelength and thereby changes the angular distribution of the reflected-plus-scattered intensity, which shall be called scattered field in the following discussion, from the surface. This phenomenon has generated interest in the fields of both physics and mechanics.

Various theories have been developed for relating the scattered field with statistical parameters of the rough surface as discussed in an excellent review by Bennett and Mattsson.¹ One

of them is a vector scattering theory² that is developed primarily for very fine surfaces such as optical surfaces. Another is a scalar scattering theory^{3,4} that can be applied to a wider range of surfaces. The emphasis of these studies is placed on the understanding of the scattering electromagnetic waves from a rough surface. The applications of these theories have been mainly on the determination of surface statistical characteristics.⁵

In the field of mechanics, the change in the scattered field from a metallic surface experiencing plastic deformation has been used as a measure of plastic strain. Many methods, such as the maximum intensity,⁶ half width,⁷ intensity at a certain distance from the specular reflection point,⁸ contrast,⁹ correlation coefficients,¹⁰ etc. have been proposed. These works are empirical in nature and contributes little to the understanding of the fundamental relation between scattered field and plastic deformation. In a recent study,¹¹ plastic deformation caused scattering is investigated by means of a scalar scattering theory. However, some key issues, such as the height distribution and the form of autocorrelation function of the plastically roughened surfaces, and the applicability of the scattering theory, are not discussed. Such discussions are essential for the proper application of the scalar scattering theory to plastic deformation assessment.

The purpose of the present work is therefore: (a) to study the statistical properties of the plastically roughened surfaces, (b) to express the scattered field in terms of plastic strain for surfaces with Gaussian and exponential autocorrelation, (c) to study the applicability and limitation of these expressions, (d) to provide a theoretical background to some existing plastic strain evaluating parameters, and (e) to verify the present approach experimentally.

Paper 2869 received Feb 12, 1990, revised manuscript received Feb 26, 1991, accepted for publication March 7, 1991.

© 1991 Society of Photo-Optical Instrumentation Engineers.

2. PROPERTIES OF PLASTICALLY ROUGHENED SURFACES

Surface statistical properties such as height distribution and autocorrelation function were studied on uniaxial tensile specimens made of aluminum 2024 alloy. Specimens were polished by cloth buffer wheel and alumina powder to an initial surface rms roughness of about $0.05 \mu\text{m}$. Then the specimens were loaded on a testing machine to different amount of plastic deformation and surface roughness. Plastic deformation is quantified by shear strain γ , which by definition is the angular change between two line segments that are originally perpendicular.

A stylus-type profilometer was used to measure height variation by mechanically scanning the specimens' surface. The radius of the stylus tip was $5 \mu\text{m}$. The sampling interval between each two adjacent points was $2 \mu\text{m}$, and the sampling length was 2mm in length. These parameters met the requirements that the sampling interval be less than 10% of profile correlation length and that the sample length be about 60 times the correlation length.¹² The surface profile signal from the profilometer was digitized by a data acquisition system and then processed by a computer. Ten trials were made with a total of 10,000 measured profile data for each specimen, and each trial was made along a direction perpendicular to the loading direction. Unlike most machined surfaces, the roughness of plastically deformed surface of polycrystalline material is isotropic.¹⁰ This means that surface height variation measured along one direction may be regarded as the surface statistical properties of the specimen along any other directions.

2.1. Statistical characteristics

The measured profile height variation along each trial was first leveled using the least square method to make the mean zero. The height probability density distribution was made on all the 10,000 data points for each specimen. Figure 1 shows some of the height histograms at different plastic deformation levels. The solid line is a Gaussian function, which was obtained by fitting the data using the least square method. It is safe to say that the height distribution of plastically roughened surfaces may be approximated by a Gaussian curve.

The autocorrelation of a profile $h(x)$ is defined as

$$C(\tau) = \frac{\sum_{i=1}^n h(x_i) \times h(x_i + \tau)}{\sum_{i=1}^n h^2(x_i)} \quad (1)$$

where τ is the lag length. The lag length τ at which the value of the autocorrelation function drops to $1/e$ (≈ 0.368) of that at $\tau = 0$ is denoted as T and called correlation length.

The autocorrelation function $C(\tau)$ may take two simple forms: Gaussian and exponential. The Gaussian autocorrelation with correlation length T is

$$C(\tau) = \exp(-\tau^2/T^2) \quad (2)$$

The exponential autocorrelation with correlation length T is

$$C(\tau) = \exp(-|\tau|/T) \quad (3)$$

The autocorrelation of each specimen was obtained by taking the average of the autocorrelation functions computed for ten trials. Figure 2(a) shows the surface profile at $\gamma = 3.0\%$ and its corresponding autocorrelation function. Figure 2 shows the

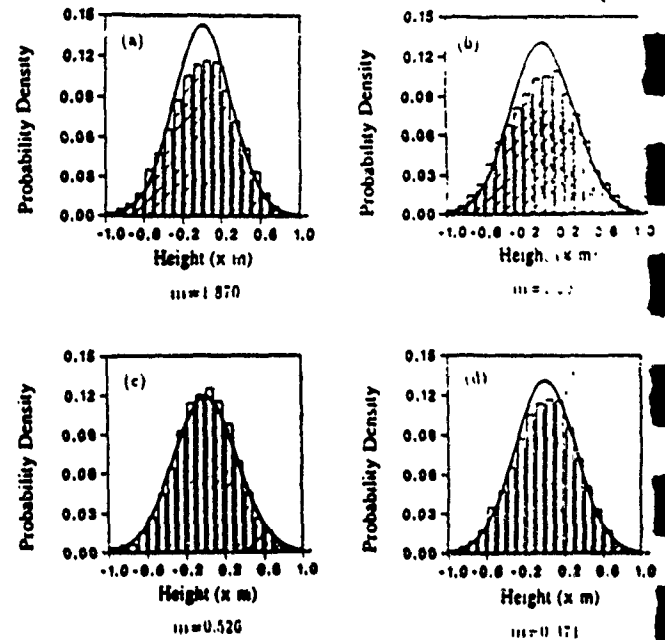


Fig. 1. Histogram of profile height at different strain levels (based on 10,000 data points) (a) $\gamma = 2.00\%$, $\sigma = 0.17 \mu\text{m}$, $T = 44.0 \mu\text{m}$; (b) $\gamma = 4.32\%$, $\sigma = 0.33 \mu\text{m}$, $T = 35.6 \mu\text{m}$; (c) $\gamma = 8.25\%$, $\sigma = 0.67 \mu\text{m}$, $T = 30.4 \mu\text{m}$; and (d) $\gamma = 11.12\%$, $\sigma = 0.87 \mu\text{m}$, $T = 42.0 \mu\text{m}$.

autocorrelation distributions for surfaces with height histogram as shown in Fig. 1. The experimental data were fitted both by Gaussian and exponential functions. The quality of fitting is measured by the deviation which is defined as

$$s = \sum_{i=1}^n [C_f(\tau_i) - C(\tau_i)]^2 \quad (4)$$

where $C_f(\tau_i)$ is the value of fitting function at point i . The smaller the value s , the better is the fitting. The fitted curves are also shown in Fig. 2. It is seen that the experimental data appear to be neither Gaussian nor exponential, even though Gaussian function is a better approximation (s is smaller) for most cases. Exponential fitting is quite off especially for small plastic strain. As the plastic strain increases, however, the experimentally obtained autocorrelation function tends to approach exponential distribution as indicated by the change in the deviation s shown in the figures.

It should be pointed out that the surface height measuring apparatus described earlier has its limitations posed by the radius of the stylus, sampling interval, and sampling length. Nevertheless, as shall be experimentally demonstrated later in the discussion, Gaussian function seems to be a better approximation to the autocorrelation of the surfaces roughened by plastic deformation.

2.2. Plastic strain and surface roughness

Surface roughness is generally characterized by rms roughness σ and profile correlation length T . Some experimental observations on plastic strain induced roughening for aluminum 2024-T3 alloy are shown in Table 1. The empirical relation between σ , T , and shear strain γ are obtained by fitting these experimental data using the least square method. The empirical relation between σ and γ is

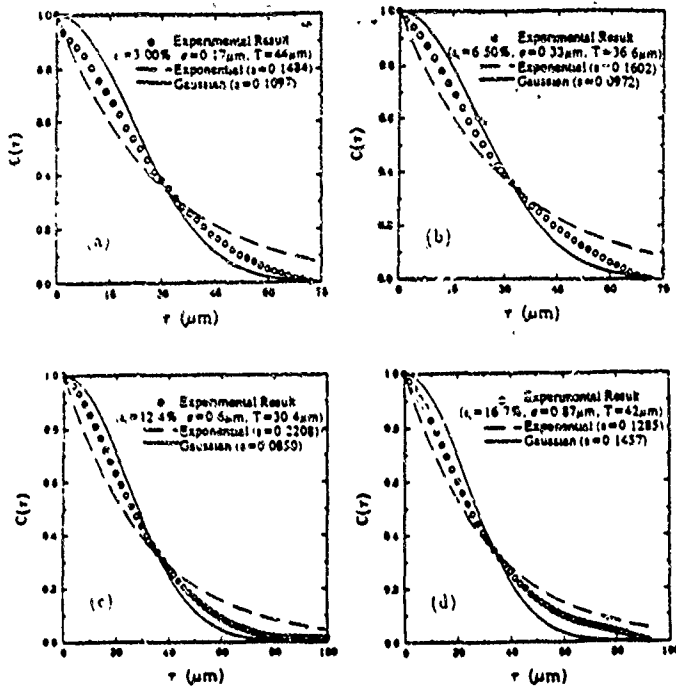


Fig. 2. Autocorrelation distribution at different strain levels (based on 10 trials with a total of 10,000 data points). (a) $\gamma = 2.00\%$, $\sigma = 0.17 \mu\text{m}$, $T = 44.0 \mu\text{m}$; (b) $\gamma = 4.32\%$, $\sigma = 0.33 \mu\text{m}$, $T = 35.6 \mu\text{m}$; (c) $\gamma = 8.25\%$, $\sigma = 0.60 \mu\text{m}$, $T = 30.4 \mu\text{m}$; and (d) $\gamma = 11.12\%$, $\sigma = 0.87 \mu\text{m}$, $T = 42.0 \mu\text{m}$.

$$\sigma = 7.21\gamma + 0.02 (\mu\text{m}) \quad (5)$$

where γ by definition is unitless and the two constants carry the same unit as that of σ , i.e., μm .

The empirical relation between T and γ is

$$T = \frac{0.66}{\gamma + 0.008} + 20.43 (\mu\text{m}) \quad (6)$$

where the constant 0.008 is unitless and the other two constants carry the unit of μm .

These two empirical relations will be used for relating plastic shear strain with the scattered field in the following discussion.

3. PLASTIC STRAIN CAUSED SCATTERING

3.1. Scalar scattering theory

Beckmann⁴ discussed extensively the scattering of electromagnetic waves from rough surfaces based on the Kirchhoff solution of the Helmholtz integral. In order to simplify the derivation, the following assumptions are made:⁴

1. The surface is perfectly conducting.
2. The surface height variation $\xi(x,y)$ is a stationary random process.

3. The incident wave is plane and linearly polarized.
4. The point of observation is in the Fraunhofer region of diffraction.
5. The profile curvature is much larger than the wavelength of the incident light.
6. Shadowing and multiple scattering may be neglected.

Under these conditions, the solution describing the mean scattered field $\langle \rho\rho^* \rangle$ is⁴

$$\langle \rho\rho^* \rangle = \frac{2\pi F^2}{A} \int_0^\infty J_0(u_{xy}\tau) \chi_2(u_x - u_x;\tau) \tau d\tau \quad (7)$$

where $\langle \dots \rangle$ denotes ensemble average; A is the area being illuminated; J_0 the Bessel function of the first kind of order zero; τ the distance between two points on the rough surface; $\chi_2(u_x^2 - u_x;\tau)$ is the 2-D Fourier transform associated with the probability density distribution of $\xi(x,y)$; and F, u_x, u_{xy} are defined as follows⁴

$$F = \frac{1 + \cos\theta_1 \cos\theta_2 - \sin\theta_1 \sin\theta_2 \cos\theta_3}{\cos\theta_1(\cos\theta_1 + \cos\theta_2)} \quad (8)$$

$$u_x = k(\cos\theta_1 + \cos\theta_2) \quad (9)$$

$$u_{xy} = k\sqrt{\sin^2\theta_1 - 2\sin\theta_1 \sin\theta_2 \cos\theta_3 + \sin^2\theta_2} \quad (10)$$

where $k = 2\pi/\lambda$; λ is the wavelength of the incident light beam; and the definitions of the angles $\theta_1, \theta_2, \theta_3$ are given in Fig. 3.

Equation (7) is a general solution that can be applied to surfaces with any autocorrelation functions. If, however, $\xi(x,y)$ is a normal process with mean zero, standard deviation σ , and autocorrelation $C(\tau)$, then it can be evaluated by a series expansion or by saddle-point integration. The series expansion of Eq. (7) for Gaussian autocorrelation [Eq. (2)] is⁴

$$\langle \rho\rho^* \rangle = \exp(-g) \left[\rho_0^2 + \frac{\pi T^2 F}{A} \sum_{m=1}^\infty \frac{g^m}{m!m} \exp(-u_x^2 T^2 / 4m) \right] \quad (11)$$

where ρ_0^2 is proportional to the intensity of the diffracted field of the illuminated area A . For example, it is the Airy pattern for a circular area A . The scaled roughness g is defined as⁴

$$\sqrt{g} = k\sigma(\cos\theta_1 + \cos\theta_2) \quad (12)$$

It should be noted that there is no restriction on the value of σ in Eq. (11) even though the series will converge too slowly to be of much practical use when σ and, hence, g become large. For strong scatterers ($g \gg 1$) with Gaussian autocorrelation, an alternative close form expression can be obtained⁴ by a direct saddle-point integration of Eq. (7)

Table 1. Plastic strain and surface roughness (aluminum alloy 2024-T3).

$\gamma(\%)$	0.40	1.2	2.3	3.6	5.2	7.3	8.8	10.2
$\sigma(\mu\text{m})$	0.06	0.10	0.19	0.23	0.36	0.50	0.63	0.78
$T(\mu\text{m})$	85	46	38	32	30	32	31	28

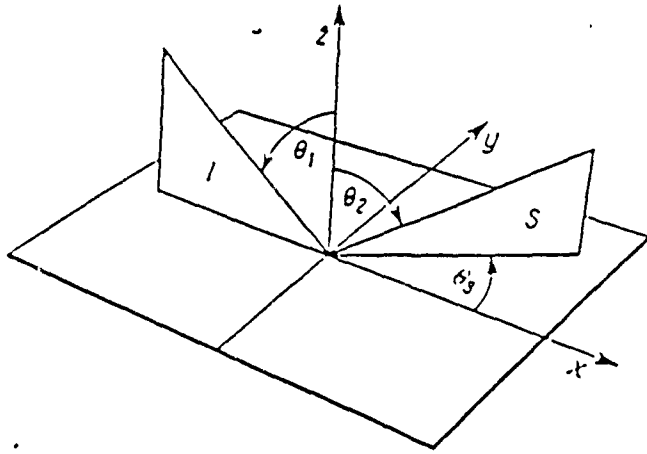


Fig. 3. Light-scattering geometry. (I is the plane of incidence and S the scattering plane.)

$$\langle \rho\rho^* \rangle = \frac{\pi F^2 T^2}{A v_z^2 \sigma^2} \exp\left(-\frac{v_z^2 T^2}{4 v_z^2 \sigma^2}\right) \quad (g \gg 1) \quad (13)$$

The saddle-point integration of Eq (7) for exponential autocorrelation [Eq. (3)] yields⁴

$$\langle \rho\rho^* \rangle = \frac{2\pi F^2 T^2}{A v_z^2 \sigma^2 (1 + v_z^2 T^2 / v_z^2 \sigma^2)^{3/2}} \quad (g \gg 1) \quad (14)$$

For the present study, the incident light is normal to the specimen surface making $\theta_1 = 0$, $F = 1$, $v_x = k \sin\theta_2$, and $v_z = k(1 + \cos\theta_2)$. We notice that θ_3 disappears in the above equations, indicating that the scattered field is circularly symmetric.

3.2. Plastic strain and scattering

According to the definitions^{1,13}, the rms roughness is equivalent to the standard deviation. Therefore they are all denoted as σ in this discussion. Substituting Eqs. (5) and (6) into Eqs. (11), (13), and (14), we can express the scattered field in terms of shear strain. With the normal incidence condition applied, and the constants in Eqs. (5) and (6) replaced by *a, b, c, d, p, q*, Eq. (11), the summation form for Gaussian autocorrelation, may be rewritten as

$$\langle \rho\rho^* \rangle = \exp(-g) \left(\rho_0^2 + \frac{\pi[a/(\gamma+b)+c]^2}{A} \sum_{m=1}^{\infty} \times \frac{g^m}{m!m} \exp\{-\sin^2\theta_2[a/(\gamma+b)+c]^2/4m\} \right) \quad (15)$$

where $\sqrt{g} = k(p\gamma + q)(1 + \cos\theta_2)$.

Similarly, Eq. (13), the close form for Gaussian autocorrelation, may be rewritten as

$$\langle \rho\rho^* \rangle = \frac{\pi[a/(\gamma+b)+c]^2}{Ak^2(1+\cos\theta_2)^2(p\gamma+q)^2} \times \exp\left(-\frac{\sin^2\theta_2[a/(\gamma+b)+c]^2}{4(1+\cos\theta_2)^2(p\gamma+q)^2}\right) \quad (16)$$

Equation (14), the close form for exponential autocorrelation, may be rewritten as

$$\langle \rho\rho^* \rangle = \frac{2\pi[a/(\gamma+b)+c]^2/[Ak^2(1+\cos\theta_2)^2(p\gamma+q)^2]}{(1+\sin^2\theta_2[a/(\gamma+b)+c]^2/[(1+\cos\theta_2)^2(p\gamma+q)^2])^{3/2}} \quad (17)$$

Three new expressions have been obtained, each describes the scattered field in terms of shear plastic strain γ . The choice on the use of each individual expression depends on the magnitude of strain and the form of autocorrelation function of the surface under investigation. In this study, however, all three expressions will be applied in parallel. Given a plastic strain value and the form of autocorrelation function, then the scattered field may be evaluated by the above expressions. A series of these distributions at different plastic shear strain levels are plotted in Fig. 4. As shown, the maximum intensity decreases and the spread of the scattered field increases as the plastic strain increases.

3.3. Applicability

A few points should be discussed before the above expression can be applied to plastic strain assessment, first the applicability of these three equations and then the convergence of Eq (15).

The conditions (Sec. 3.1) under which Eqs. (15) through (17) are derived should be satisfied. These requirements are in general met for application to plastic strain assessment of metallic materials with a He-Ne laser ($\lambda = 0.6328 \mu\text{m}$). Generally speaking, metals are conducting (condition 1); plastically roughened surface may be considered as a stationary random process (condition 2); laser beam is plane and polarized wave (condition 3), the point of observation is sufficiently far from the specimen surface or a lens is being used such that it is in the Fraunhofer zone of diffraction (condition 4); and the profile wavelength is much larger than the incident light wavelength (conditions 5, 6).

The validity of Eqs. (15) through (17) can be investigated by comparing the scattered field computed by these expressions and

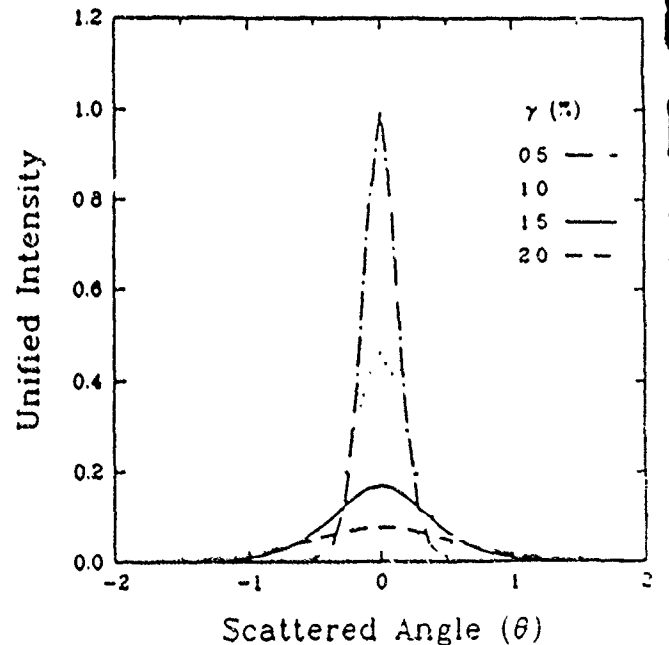


Fig. 4. Scattered fields at different strain levels (based on Eq (15))

that by measurements. Because the solutions from the scattering theory are based on the ensemble average on macroscopically similar but microscopically different surfaces, the average over many scattered fields is needed in order to make the comparison. One hundred scattered fields from a uniformly deformed specimen were taken by a system schematically shown in Fig. 5. Figure 6 shows a comparison between the two scattered fields where Fig. 6(a) is the computed one by Eq. (15) while Fig. 6(b) is the measured result. The degree of similarity between these two distributions, as evaluated by the cross correlation factor [Eq. (20)] is more than 99%. Figure 6(c) shows the relation between the number of averaged scattered fields and the cross correlation of the averaged with the computed results.

For the plastic strain range studied ($\gamma \leq 10\%$), σ ranges from $0.05 \mu\text{m}$ to $0.75 \mu\text{m}$. For a He-Ne laser with a wavelength of $\lambda = 0.6328 \mu\text{m}$, g varies from 0.986 to 193.23. This means that the close-form solutions Eqs. (16) and (17) cannot be used for the whole plastic strain range mentioned earlier because g is not sufficiently large for small strains. Therefore, Eq. (15) has to be used. Because it is in a series form, it is necessary that its convergence be studied.

By applying D'Alambert principle, we can show that Eq. (15) is convergent for any values of strain γ . However, as γ gets larger, convergence will be slow. A numerical analysis was made to study the convergence of Eq. (15). The computation result is shown in Fig. 7 where the horizontal axis is the number of terms added, and the vertical axis is the unified summation over N terms. It is seen that more than 200 terms are needed for Eq. (15) to converge when $\gamma = 10\%$.

4. ANALYSIS OF STRAIN-MEASURING METHODS

Almost all the methods used so far by various investigators for relating scattered field to plastic strain are based on the observations similar to those shown in Fig. 4. Namely, the intensity at the specular direction decreases and the spread of scattered field increases as the plastic strain increases. Several methods, such as the maximum intensity, half width, intensity at a certain distance from the specular reflection point, and correlation coefficients, have been introduced. Figure 8 gives the definition of the parameters used in some of the methods. This section will be devoted to providing a theoretical-numerical analysis to the relation between these parameters and plastic strain followed by an experimental verification for the half width method.

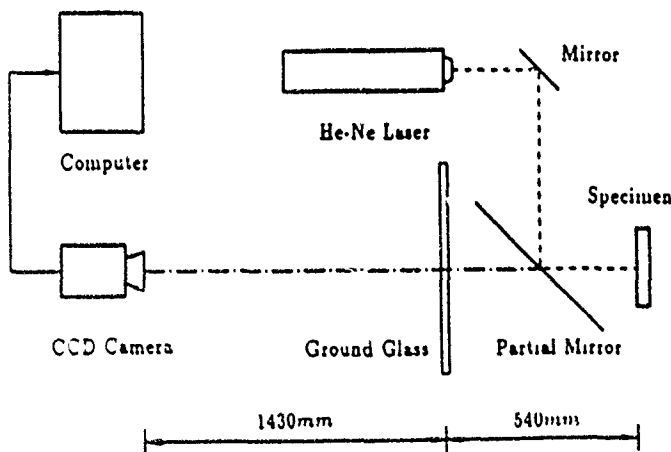


Fig 5 A schematic plot of experimental setup for scattered-field recording and processing.

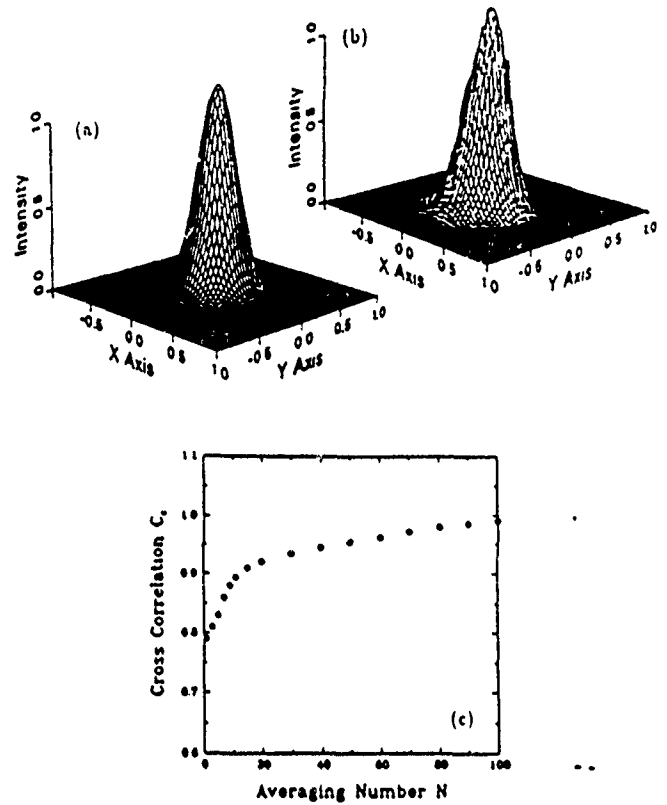


Fig. 6. Verification of the scalar scattering theory by ensemble average ($\sigma = 0.66 \mu\text{m}$, $T = 39.50 \mu\text{m}$). (a) Computed scattered field [based on Eq. (15)]; (b) Experimental scattered field (average over 100 measurements); and (c) Effect of the number of scattered fields averaged.

4.1. Intensity-related methods

The maximum intensity implies the intensity at the specular direction along which $\theta_2 = 0$, $v_{xy} = 0$ and $v_z = 2k$. Substituting these conditions into Eqs. (15) to (17), we shall get expressions of the maximum intensity in terms of strain γ for each case correspondingly. For example, the maximum intensity for Eq. (15) is

$$\langle \rho \rho^* \rangle_0 = \exp(-g_0) \left(1 + \frac{\pi[a/(\gamma + b) + c]^2}{A} \sum_{m=1}^{\infty} \frac{g_0^m}{m!m} \right) \quad (18)$$

where

$$\sqrt{g_0} = 2k(p\gamma + q) \quad (19)$$

Figure 9 is a plot of the normalized maximum intensity versus plastic strain γ evaluated by both the summation and close form expressions for Gaussian autocorrelation [Eqs. (15) and (16)]. The maximum intensity for exponential autocorrelation [Eq. (17)] is not plotted because it differs from that of Eq. (16) only by a constant factor of two. The trend predicted by the approach is similar to the experimental result.⁶ (An exact comparison is not made because of the insufficiency of information provided in the work cited.) It can be seen that the maximum intensity is very sensitive to the change in plastic strain in the early stages of plastic deformation and becomes saturated after strain becomes larger than 2%.

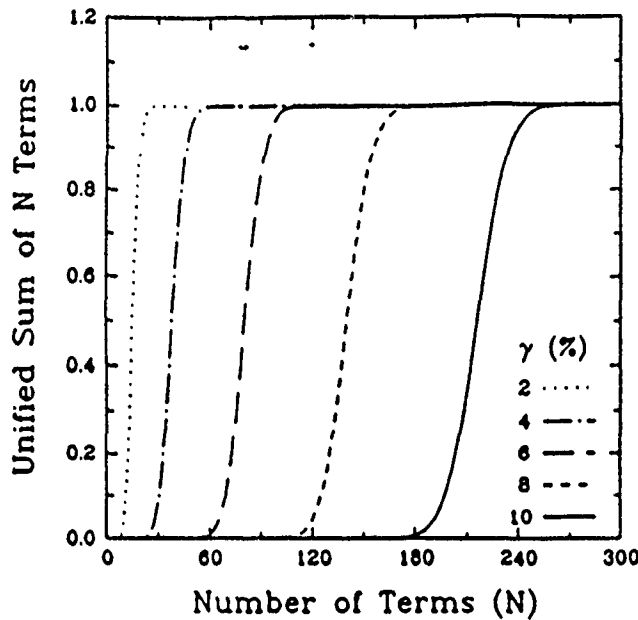


Fig. 7. Convergence of the summation form for Gaussian autocorrelation [Eq. (15)]. $\gamma = 2.0\%$: $\sigma = 0.31 \mu\text{m}$, $T = 34.2 \mu\text{m}$; $\gamma = 4.0\%$: $\sigma = 0.45 \mu\text{m}$, $T = 30.1 \mu\text{m}$; $\gamma = 6.0\%$: $\sigma = 0.60 \mu\text{m}$, $T = 27.9 \mu\text{m}$; $\gamma = 8.0\%$: $\sigma = 0.74 \mu\text{m}$, $T = 26.5 \mu\text{m}$.

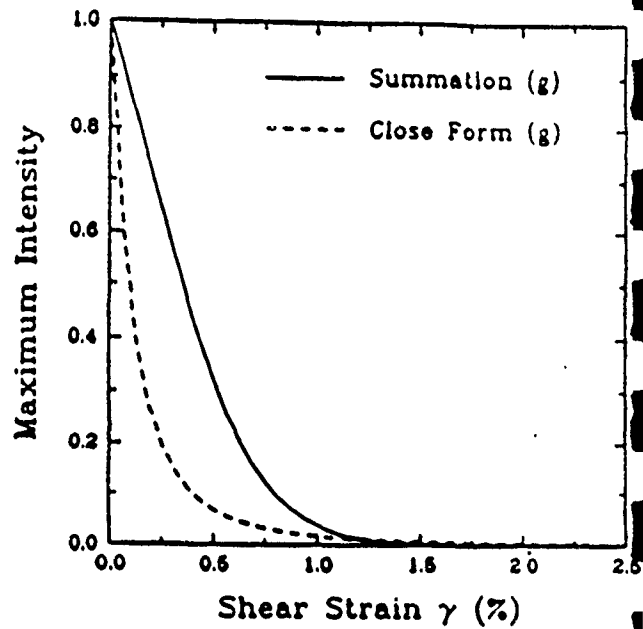


Fig. 9. Maximum intensity versus strain.

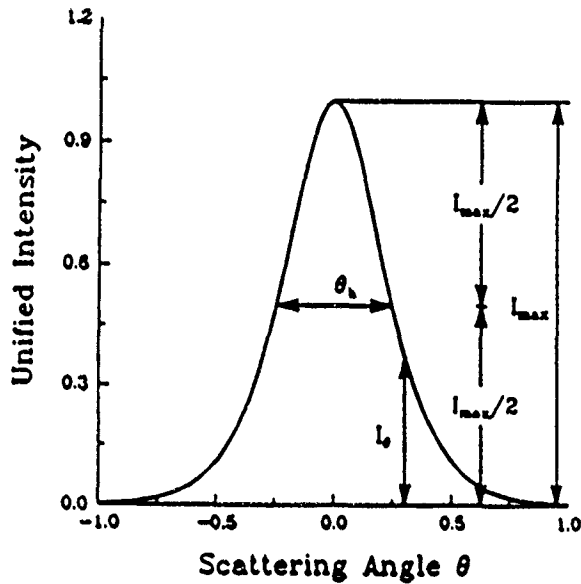


Fig. 8. Definitions of some strain measuring parameters. I_{max} : the maximum intensity; I_{θ} : intensity along angle θ ; θ_h : half width.

The second intensity-related parameter was originally defined as the intensity at a certain distance from the specular point on the observation plane.⁸ For generality, intensity at a certain angle θ_2 denoted as I_{θ} is used (Fig. 8). Let θ be a certain angle along which a sensor is placed for measuring intensity, I_{θ} can then be evaluated by Eqs. (15) to (17). Figure 10 plots I_{θ} versus strain at four different angles based on Eq. (15). It can be seen that the threshold of this method depends on the angle θ . For $\theta = 0.6$ deg, the value of this threshold is about 0.5%, while for $\theta = 1$ deg, it is 1%. In other words, putting the sensor at different spots, one will get different I_{θ} , γ relation. Another disadvantage of this parameter is that the measurable range of strain is very much limited (about 1%). Besides, the relation between this

parameter and plastic strain is nonmonotonic when strain becomes larger than a certain value (about 1.5% for $\theta = 0.6$ deg). Similar results were reported for the contrast parameter.⁹ This nonmonotonic behavior makes the determination of strain not practical for situations when the range of strain is unknown. However, because of the high sensitivity of the intensity-related parameters to early stage plastic deformation, they may be applied to detect the onset of plastic deformation. Such applications may include the elasto-plastic boundary determination and fatigue detection.

4.2. Correlation coefficients

Unlike the conventional correlation analysis used in statistics, the term *correlation coefficient* used in this discussion is adopted from pattern recognition technique. Two kinds of correlation coefficients were performed: cross correlation and autocorrelation. Cross correlation coefficient is a measure of the degree of similarity between two different objects. Autocorrelation is a measure of the spread of an object. Correlation analysis can be performed on both 2-D images and 1-D curves. Because of the circular symmetry of Eqs. (15) to (17), only 1-D correlation analysis was made in the present study, while an extension to 2-D problems is straightforward.

A 1-D zero-shift, cross-correlation coefficient C_c is defined as:

$$C_c = \frac{\int_{-\infty}^{\infty} f(x)g(x) dx}{\sqrt{\int_{-\infty}^{\infty} f^2(x) dx \int_{-\infty}^{\infty} g^2(x) dx}}$$

where $f(x)$, $g(x)$ are functions of variable x . One of them, for example, $f(x)$, is the scattered field $\langle pp^* \rangle$ at zero strain while $g(x)$ is the one at any other strain level. As the strain increases, the corresponding scattered field will be less and less correlated with that at the zero strain, causing the cross correlation level to drop.

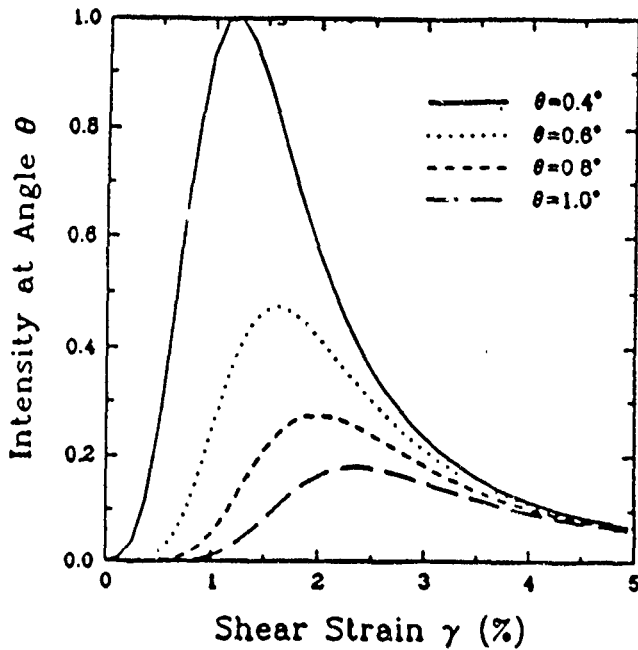


Fig. 10. Intensity along angle θ versus strain.

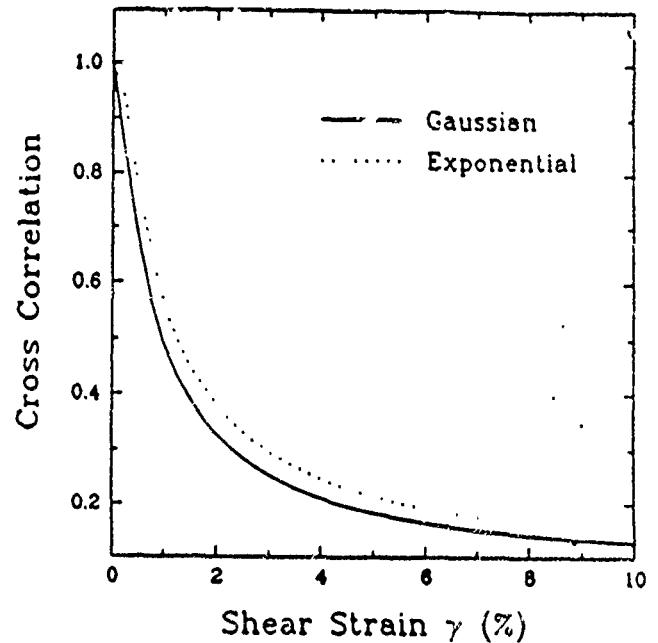


Fig. 11. Cross-correlation coefficient versus strain.

One-dimensional autocorrelation coefficient C_a is defined as

$$C_a(\tau) = \frac{\int_{-\infty}^{\infty} f(x)f(x - \tau) dx}{\int_{-\infty}^{\infty} f^2(x) dx} \quad (21)$$

where τ is the lag length shifted between two copies of function $f(x)$. This definition is similar to that discussed in Eq. (1) except that it is performed on the scattered field rather than on the rough surface height variation.

Correlation coefficients were evaluated by using Eqs. (20) and (21) on the intensity distribution curves computed by Eqs. (15) to (17). Figure 11 is a plot of cross correlation coefficient versus strain. Figure 12 shows the autocorrelation analysis result where the vertical axis is the lag length at 75% correlation level. The relation between lag length τ and strain is linear for $\gamma \leq 6\%$. These relations support the experimental results.¹⁰

4.3. Half-width method and its verification

Half width was originally defined as the spreading distance u_h on the observation plane of the scattered field when its value is one half of its maximum. For generality, we use the spread angle θ_h instead. These two parameters are related by $u_h = L \times \tan \theta_h$. For small scattering angles, which are a few degrees for the current application, θ_h is approximately proportional to u_h . The expression for, for example, Eq. (16) is

$$\frac{\langle pp^* \rangle_{\theta_2 = \theta_h}}{\langle pp^* \rangle_{\theta_2 = 0}} = \frac{1}{2} = \frac{4}{(1 + \cos \theta_h)^2} \times \exp\left(-\frac{\sin^2 \theta_h [a/(\gamma + b) + c]^2}{4(1 + \cos \theta_h)^2 (p\gamma + q)^2}\right) \quad (22)$$

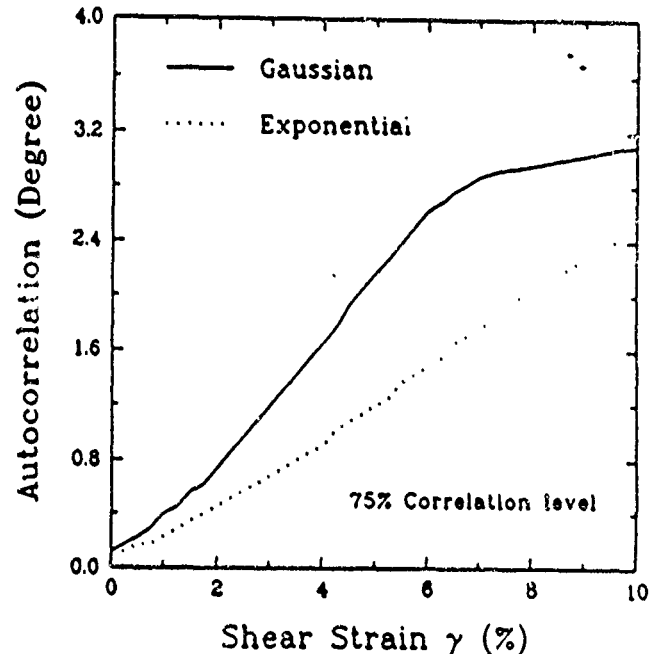


Fig. 12. Autocorrelation analysis: lag length versus strain for $C_a = 75\%$.

Figure 13 is a plot of θ_h versus plastic strain γ evaluated by expressions corresponding to different cases. The measured roughness values T and σ were used. Experimentally measured half widths are also included for comparison. Apparently, the experimental results agree well with that predicted by both the summation and close-form expressions corresponding to Gaussian autocorrelation. This result confirms the conclusion made earlier (Sec. 2.1) that Gaussian function approximates autocorrelation of plastically roughened surface better.

5. CONCLUSION

Based on the earlier discussions, we may conclude that the height distribution of the plastically roughened metallic surface is Gaus-

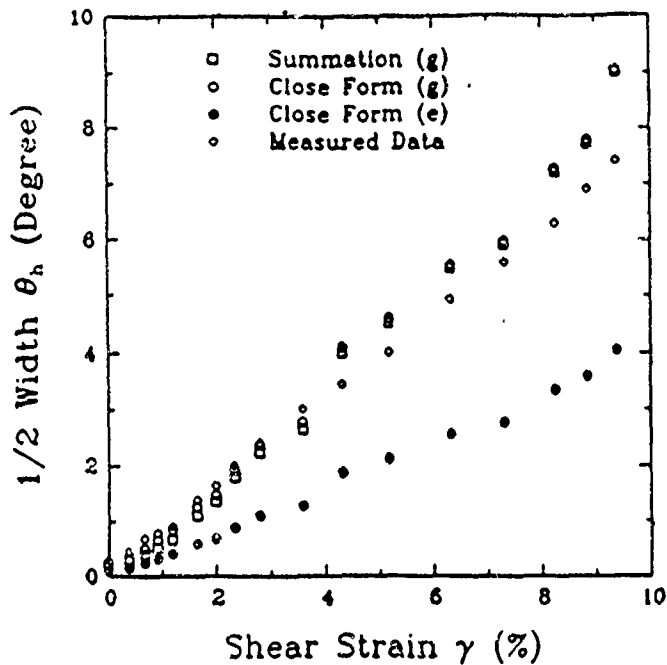


Fig. 13. Half width versus strain with experimental verification.

sian and the autocorrelation is neither Gaussian nor exponential. However, Gaussian function in general fits the experimental data better. The experimentally obtained autocorrelation becomes closer to exponential distribution as the strain gets larger.

The combination of plastic deformation induced roughening models with scalar scattering theory has provided expressions that relate the scattered field to plastic strain directly. For the strain range studied, these new expressions have been shown to be valid by an ensemble average study and some numerical analysis.

The analysis shows that the intensity-related strain-measuring methods, such as the maximum intensity and the intensity at a certain scattering angle, are not as suitable as distribution related ones such as the half-width and correlation coefficients. Comparing the theoretical-numerical predictions with the experimental data, we find a good agreement for all the parameters. However, considering the factors such as the simplicity of the relation, the measurable range of plastic strain, we recommend autocorrelation and half-width methods rather than the others. The current approach also shows that the range of plastic shear strain measurable by using the scattered field via the earlier described methods is about 10%.

6. ACKNOWLEDGMENT

Financial support provided by Army Research Office, Engineering Science Division, through contract No. DAAL0388K0083 is gratefully acknowledged.

7. REFERENCES

1. J. M. Bennett and L. Mattsson, *Introduction to Surface Roughness and Scattering*, Optical Society of America, Washington, DC (1989).
2. J. M. Elson and J. M. Bennett, "Vector scattering theory," *Opt. Eng.* 18(2), 116-124 (1979).
3. H. Davis, "The reflection of electromagnetic waves from a rough surface," *Proc. Inst. Elec. Eng.* 4(101), 209-214 (1954).
4. P. Beckmann and A. Spizzichino, *The Scattering of Electromagnetic Waves from Rough Surfaces*, Pergamon, London (1963).
5. E. L. Church, H. A. Jenkinson, and J. M. Zavada, "Relationship between surface scattering and microtopographic features," *Opt. Eng.* 18(2), 125-136 (1979).
6. K. Miyagawa and A. Azushima, "Measurement of plastic deformation by means of a laser beam," 9th World Conference on Nondestructive Testing, 1-7, Australia (1979).
7. A. Azushima and M. Miyagawa, "Measurement of plastic zone around a fatigue crack of a stainless steel structure member with a laser beam," *Proc. Joint Conf. on Exper. Mechanics*, 709-713, Hawaii (1982).
8. B. Q. Xu and X. P. Wu, "Measuring plastic deformation on metal surface and determining the elasto-plastic boundary from the change of laser produced diffraction patterns," *Acta Mech. Sinica* 16(2), 175-183 (1984).
9. F. P. Chiang, H. H. Qian, and H. S. Don, "Determination of plastic strain using laser speckles," 2nd. Int. Symposium on Plasticity & its Current Applications, Japan (1989).
10. F. P. Chiang, Y. Z. Dai, B. Q. Xu, and A. Kato, "Study of surface roughening under different stress modes by correlation," *Hologram Interferometry and Speckle Metrology*, 36-42, Baltimore (1990).
11. Y. Z. Dai and F. P. Chiang, "Application of scattering theory to plastic strain estimation," presented at the 4th Int. Symposium on Nondestructive Characterization of Materials, Annapolis (1990).
12. A. Ogilvy and J. R. Foster, "Rough surfaces: Gaussian or exponential statistics?" *J. Phys. D* 22(9), 1243-1251 (1989).
13. P. E. Pfeiffer and D. A. Schum, *Introduction to Probability*, Academic Press, New York (1973).



Y. Z. Dai obtained his Ph.D. in mechanical engineering from the State University of New York at Stony Brook in 1991, and his MS and BS degrees from Tianjin University, China, in 1981 and 1984, respectively. He joined Brown & Sharpe Manufacturing Company in 1991. This work was done while he was a graduate student at Stony Brook.



F. P. Chiang obtained his Ph.D. in mechanics and engineering science from the University of Florida in 1966. He joined the State University at Stony Brook in 1967 and is currently the director of the Laboratory for Experimental Mechanics Research and a leading professor of mechanical engineering. His major research activity is in the theory and application of optical methods to mechanics and metrology. He has worked in photoelasticity, moiré, holography, and speckle methods.

Assessment of Flow Stress and Plastic Strain by Spectrum Analysis

by Y.Z. Dai and F.P. Chiang

ABSTRACT—The relation between the angular distribution of the reflected-plus-scattered light intensity (scattered field) from a metallic surface and the flow stress, plastic strain the material has experienced is experimentally and theoretically investigated. A scattered field, which is obtained by illuminating a specimen surface using a laser beam, carries surface-feature-related information. Experimental evidence suggests that surface correlation length of a material decreases in proportion to the flow stress and plastic strain that the material experiences. A theoretical derivation based on Huygens-Fresnel principle, Fraunhofer approximation, and Wiener-Khinchine theorem shows that the correlation length may be obtained by performing a Fourier transform to the scattered field from the surface. This leads to the development of a noncontact, nondestructive, and remote technique for measuring flow stress and plastic strain.

Introduction

Loading metallic materials beyond the yield strength will change their microstructural features due to the movement of dislocations. This change will be reflected by surface-roughness parameters such as root-mean-square (*rms*) roughness and correlation length. The surface-roughness variation in turn will affect the angular distribution of the reflected-plus-scattered intensity, which shall be called scattered field in the following discussion, from the surface. This phenomenon has been investigated experimentally in an effort to develop a technique for plastic-strain measurement.¹⁻⁴ Methods such as the maximum intensity,¹ spectrum width,² intensity at a certain scattering angle,³ contrast,⁴ moment,⁵ and correlation coefficient⁶ have been introduced. These methods are empirical in nature and the intrinsic relation between a scattered field and plastic deformation is yet to be revealed. The revealing of such a relation lies on the understanding of: (1) the plastic-deformation-induced surface roughness and (2) the retrieval of surface-roughness parameters from a scattered field.

Experimental observations show that the surface correlation length decreases in proportion to plastic strain and flow stress. The implication is that surface correlation length, once measured, may be used as a measure of flow stress and plastic strain. Correlation length represents the horizontal property of a surface profile which may be obtained by using instruments such as a mechanical

profilometer. The light scattering method shall be used in this study because of its merits of being noncontact, non-destructive, and two dimensional. A theoretical derivation based on the Huygens-Fresnel principle, Fraunhofer approximation, and Wiener-Khinchine theorem shows that the Fourier transform of a scattered field is approximately proportional to the autocorrelation function of the surface profile for optically smooth surfaces. This means that surface correlation length and hence flow stress and plastic strain may be evaluated from the scattered field.

The purpose of the present work is therefore: (1) to give a rigorous mathematical derivation on the relation between a scattered field and surface correlation length and (2) to use the scattered field for flow stress and plastic strain measurement. Furthermore, the relation between plastic-deformation-induced surface macroscopic and microscopic features is discussed.

Stress, Strain and Correlation Length

Figure 1 is a microscopic photo of a plastically deformed surface. A surface profile is shown in Fig. 2. The horizontal property of a one-dimensional surface profile may be characterized by surface correlation length which is defined as⁷

$$\frac{\int_T h(x) \times h(x-T) dx}{\int_0^l h^2(x) dx} = \frac{1}{e} \quad (1)$$

where $h(x)$ is the surface profile; l the sampling length; and T the correlation length as shown in Fig. 2; $1/e$ is equal to 0.3678.

Experimental observations conducted on copper alloy (C26800-H01, yield strength: 203 MPa, Poisson's ratio: 0.33) show that correlation length T decreases in proportion to flow stress τ and plastic strain γ as shown in Fig. 3. These relations may be approximately expressed as

$$\tau = \frac{\beta_\tau}{T + c_\tau} + \alpha_\tau \quad (2)$$

and

$$\gamma = \frac{\beta_\gamma}{T + c_\gamma} + \alpha_\gamma \quad (3)$$

where β_τ , c_τ , α_τ , β_γ , c_γ , α_γ are constants. The validity of both eq (2) and eq (3) requires that the materials be linear hardening. Otherwise, different forms of expressions may be needed.

Interestingly enough, these relations are analogous to those relating flow stress and plastic strain to cell size in a much smaller scale. Plastic deformation in metals produces dislocations and dislocation bands which cluster into dis-

Y.Z. Dai (SEM Member) is Senior Engineer, Research and Development Department, Brown and Sharpe Manufacturing Co., North Kingstown, RI 02852. F.P. Chiang (SEM Fellow) is Professor, Department of Mechanical Engineering, State University of New York at Stony Brook, NY 11794.

Original manuscript submitted, May 1990. Final manuscript received, March 5 1991.

location cell walls. The applied stress is the driving force in creating dislocation and cell structures. Once a microstructure is stabilized, a higher stress is generally required for activating new slip lines and producing more cells. The dislocation cell size d may be calculated by⁹

$$d = \sqrt{\frac{6A}{\pi N}} \quad (4)$$

where A is the area in which the cell size is to be determined and N is the number of cells in the area.

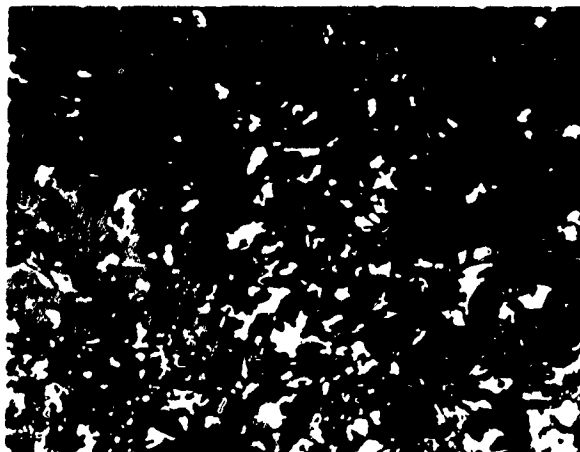


Fig. 1—Microstructures of a plastically deformed metallic surface

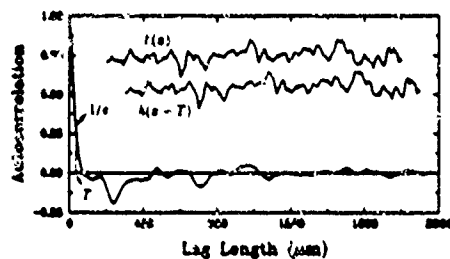


Fig. 2—A surface profile and its autocorrelation function

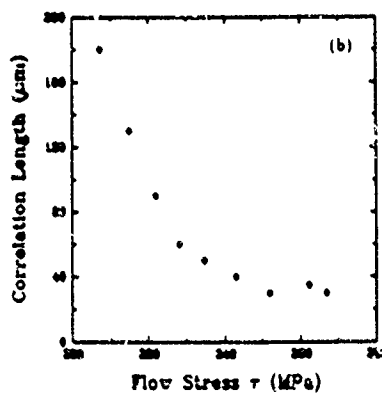
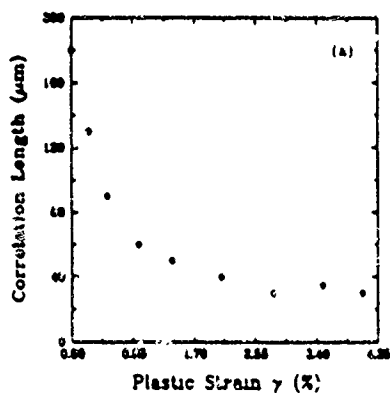


Fig. 3—Relation between flow stress, plastic strain and correlation length. (a) flow stress and correlation length, (b) plastic strain and correlation length

The cell size d has been shown to be related to flow stress τ by⁹

$$d = aK_0 \frac{Gb}{\tau - \tau_0} \quad (5)$$

where a, K_0 are constants; b the Burgers vector; G the shear modulus and τ_0 the friction stress. Equation (5) may be written in a more general form as

$$\tau = \frac{\beta'_1}{d + c'_1} + \alpha'_1 \quad (6)$$

The dislocation cell size d is related to dislocation density ρ by⁹

$$d = K_1 / \sqrt{\rho} \quad (7)$$

Relations expressed in eqs (5) and (7) have been shown to be valid for aluminum, copper and iron. For mechanical-engineering applications, plastic strain is a much more popular quantity than dislocation density and hence there is a need to find the relation between plastic strain and dislocation density. Some observations were made on the relation between shear strain γ and dislocation density ρ of an aluminum single crystal.¹⁰ The results are as shown in Fig. 4. (It was plotted versus the dislocation density originally.) Linear function appears to be a good approximation to the experimental data.

$$\gamma = \alpha \sqrt{\rho} + \beta \quad (8)$$

where α, β are constants.

Substituting eq (7) into eq (8), we can write the resulting equation in a more general form:

$$\gamma = \frac{\beta'_1}{d + c'_1} + \alpha'_1 \quad (9)$$

where $\beta'_1, c'_1, \alpha'_1$ are constants. Equations (6) and (9) indicate that both flow stress τ and shear strain γ decrease in proportion to cell size d . This is analogous to the relation between flow stress/plastic strain and surface correlation length, although the magnitude of surface correlation length is about one order higher than that of the cell size.

Scattered Field and Correlation Length

The scattered field from a rough surface can be obtained using a configuration as schematically shown in

Fig. 5. The plane immediately in front of the rough surface is denoted as $X-Y$ and the observation plane $U-V$. The phase variation $\psi(x, y)$ on the $X-Y$ plane caused by the surface height variation $h(x, y)$ is¹¹

$$\psi(x, y) = \frac{2\pi(1 + \cos \vartheta)}{\lambda} h(x, y) \quad (10)$$

where λ is the wavelength of the incident light; ϑ is the angle between the incident light and surface normal.

The standard deviation of the phase variation $\psi(x, y)$ and the surface height variation $h(x, y)$ are related by

$$\sigma_\psi = \frac{2\pi(1 + \cos \vartheta)}{\lambda} \sigma \quad (11)$$

where σ_ψ , σ are the standard deviations of $\psi(x, y)$ and $h(x, y)$, respectively.

If surface slopes are small and depolarization, multiple scattering, and shadowing are negligible, the electromagnetic field $f(x, y)$ in the $X-Y$ plane is related to the surface height variation $h(x, y)$ by¹¹

$$f(x, y) = RP(x, y) \exp \left[i \frac{2\pi}{\lambda} (1 + \cos \vartheta) h(x, y) \right] \quad (12)$$

where R is the average reflectivity; $P(x, y)$ is the complex field incident on the surface. This equation can be further simplified if the specimen surface is perfectly reflecting ($R = 1$), the incident light is normal to the surface ($\vartheta = 0$), and the light is uniform and polarized [$P(x, y) = 1$]. Then eq (12) can be written as

$$f(x, y) = \exp \left[-\frac{4\pi}{\lambda} h(x, y) \right] = \exp [i\psi(x, y)] \quad (13)$$

If the observation plane is far enough from the specimen surface so that the Fraunhofer approximation can be applied, the electromagnetic field $\rho(\xi, \eta)$ in the $U-V$ plane is¹¹

$$\rho(\xi, \eta) = C \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) e^{-i2\pi(\xi x + \eta y)} dx dy \quad (14)$$

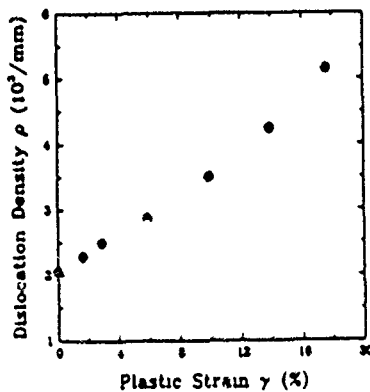


Fig. 4—Relation between plastic strain and dislocation density

where ξ, η are spatial frequencies which are related to the coordinates u, v in the $U-V$ plane by

$$\xi = \frac{u}{L\lambda} \quad \eta = \frac{v}{L\lambda} \quad (15)$$

where L is the distance between the rough surface and the $U-V$ plane. Aside from the multiplicative factor C preceding the integral, eq (14) is simply the Fourier transform $F(\xi, \eta)$ of $f(x, y)$.

$$F(\xi, \eta) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) e^{-i2\pi(\xi x + \eta y)} dx dy \quad (16)$$

The light intensity on the observation plane recorded either on a piece of film or a CCD sensor is

$$\rho(\xi, \eta) \rho^*(\xi, \eta) = \left| C \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) e^{-i2\pi(\xi x + \eta y)} dx dy \right|^2 \quad (17)$$

where $\rho^*(\xi, \eta)$ is the complex conjugate of $\rho(\xi, \eta)$. Using eq (16), we can rewrite eq (17) as

$$\rho(\xi, \eta) \rho^*(\xi, \eta) = C^2 |F(\xi, \eta)|^2 \quad (18)$$

The term $|F(\xi, \eta)|^2$ is the power-spectral-density function of $f(x, y)$. Equation (18) indicates that a scattered field is proportional to the power-spectral-density function of the electromagnetic field immediately in front of the surface. According to the Wiener-Khinchine theorem,¹¹ the autocorrelation function $C_f(p, q)$ of the electromagnetic field $f(x, y)$ is the inverse Fourier transform of the power-spectral-density function of $f(x, y)$, i.e.,

$$C_f(p, q) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |F(\xi, \eta)|^2 e^{i2\pi(\xi p + \eta q)} d\xi d\eta \quad (19)$$

where p, q are the lag lengths of the autocorrelation function along the x, y directions, respectively.

Using eq (18), we may express the autocorrelation function $C_f(p, q)$ as

$$C_f(p, q) \propto \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \rho(\xi, \eta) \rho^*(\xi, \eta) e^{i2\pi(\xi p + \eta q)} d\xi d\eta \quad (20)$$

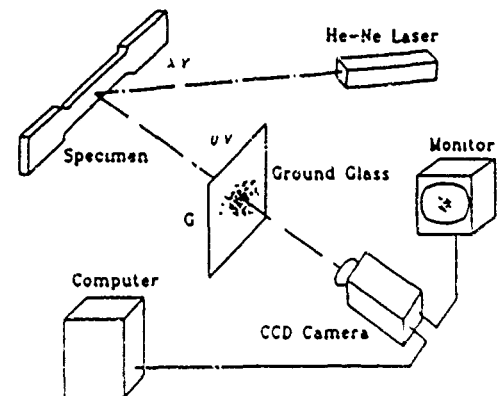


Fig. 5—A schematic setup for scattered field processing

The implication of eq (20) is that the Fourier transform of a scattered field maps the autocorrelation function of the electromagnetic field immediately in front of a surface. By definition the autocorrelation function $C_f(p, q)$ of $f(x, y)$ is¹⁴

$$C_f(p, q) = \langle f(x, y) f(x-p, y-q) \rangle \quad (21)$$

Substituting eq (13) into the right-hand-side term of eq (21), we get

$$C_f(p, q) = \langle \exp(i[\psi(x, y) + \psi(x-p, y-q)]) \rangle \quad (22)$$

If a surface is optically smooth, i.e., the surface height variation caused phase variation in near field is less than 2π , the condition $\langle \psi^2(x, y) \rangle = \sigma_\psi^2 \ll 1$ is satisfied. Then the above equation can be simplified by taking only the first three terms in a polynomial expansion of the right-hand-side term:

$$C_f(p, q) = 1 - \sigma_\psi^2 - \langle \psi(x, y) \psi(x-p, y-q) \rangle \quad (23)$$

where the zero mean condition $\langle \psi(x, y) \rangle = 0$ has been applied. The last term $\langle \psi(x, y) \psi(x-p, y-q) \rangle$ in eq (23) is the autocorrelation function $C_\psi(p, q)$ of the phase variation $\psi(x, y)$. Therefore eq (23) may be re-

written as

$$C_f(p, q) = 1 - \sigma_\psi^2 - C_\psi(p, q) \quad (24)$$

On the other hand, the autocorrelation function $C_h(p, q)$ of the phase variation $\psi(x, y)$ is proportional to that of the surface height variation $h(x, y)$ according to eq (10):

$$C_f(p, q) = \left(\frac{4\pi}{\lambda}\right)^2 C_h(p, q) \quad (25)$$

Hence, the autocorrelation function $C_h(p, q)$ of $h(x, y)$ is related to $C_f(p, q)$ by

$$C_h(p, q) = \frac{\lambda^2}{(4\pi)^2} [1 - \sigma_\psi^2 - C_f(p, q)] \quad (26)$$

For a given rough surface illuminated by a certain coherent light beam, the term $1 - \sigma_\psi^2$ and the multiplicative factor preceding it are constants. Therefore, eq (26) indicates that the autocorrelation function $C_h(p, q)$ of surface height variation $h(x, y)$ is proportional to that of the electromagnetic field $f(x, y)$ immediately in front of the surface. To apply eq (20), eq (26) may be rewritten as

$$C_h(p, q) \propto \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \rho(\xi, \eta) \rho^*(\xi, \eta) e^{i2\pi(p\xi + q\eta)} d\xi d\eta \quad (27)$$

The inverse Fourier transform of a scattered field has been shown to be approximately proportional to the autocorrelation function of the surface height variation. The surface correlation length can therefore be evaluated.

Spectrum Width and Stress, Strain

As a first approximation, a scattered field is assumed to be normally distributed.¹⁵ For a normal function, its autocorrelation function is also a normal function. For example, if the one-dimensional intensity distribution of a scattered field takes the form

$$\rho(\xi) \rho^*(\xi) = \exp(-\xi^2/u_s^2) \quad (28)$$

where u_s is the spectrum width at $1/e$ of the maximum spectrum intensity along the ξ directions, then autocorrelation function is

$$C_s \propto \frac{u_s}{\sqrt{2}} \exp(-p^2 u_s^2/4) \quad (29)$$

The surface correlation length T is then related to u_s by

$$T = \frac{2}{u_s} \quad (30)$$

The relation between spectrum width u_s and flow stress and plastic strain can be obtained by substituting eq (30) into eqs (2) and (3) yielding

$$\tau_s \approx \frac{\beta_s}{2} u_s + \alpha_s \quad (31)$$

and

$$\gamma_s \approx \frac{\beta_\gamma}{2} u_s + \alpha_\gamma \quad (32)$$

It is seen that both flow stress and plastic strain are proportional to spectrum width u_s . Once u_s is measured

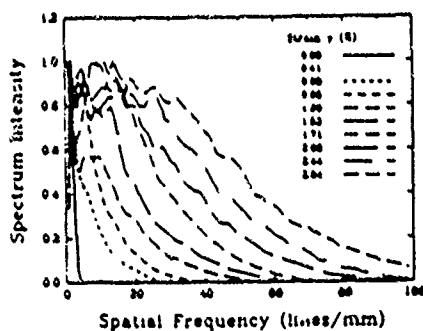


Fig. 6—One-dimensional distributions of scattered fields

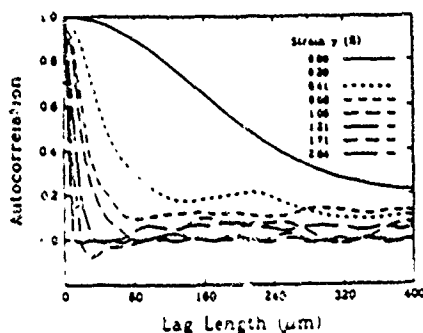
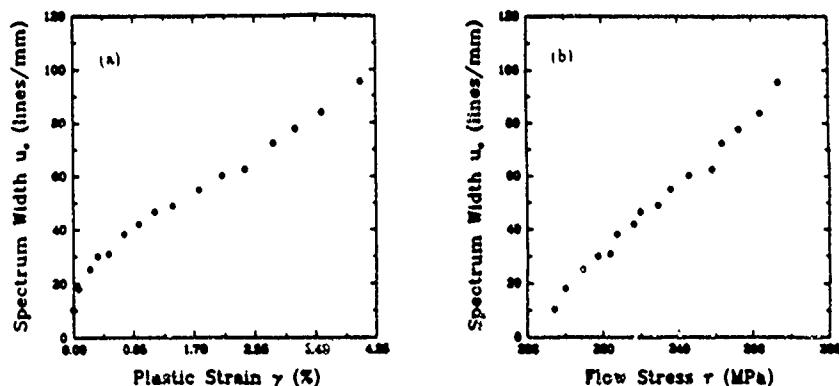


Fig. 7—Normalized autocorrelation-function distributions

Fig. 8—Relation between flow stress, plastic strain and spectrum width, (a) plastic strain versus spectrum width, (b) flow stress versus spectrum width



and constants α_1 , β_1 , α_2 , β_2 are obtained by calibration, the flow stress and plastic strain can readily be evaluated.

Experimental Verification

The validity of eqs (31) and (32) was experimentally verified on tensile specimens made of copper alloy cut from stock sheet along the rolling direction. The specimens were polished by a cloth buffer wheel to $\sigma = 0.03 \mu\text{m}$. The specimens were uniaxially loaded on a testing machine to plastic deformation and then unloaded for scattered field recording. Plastic (residual) strain was measured by a strain-gage rosette and in-plane moiré method after unloading. A He-Ne laser with a wavelength of $0.6328 \mu\text{m}$ was used as the illuminating light source. The distance between the specimen surface and the observation plane is 610 mm. In order to reduce computing time and suppress speckle noise, one-dimensional distribution of the scattered field is used. Figure 6 is a plot of the spatially averaged, smoothed, and normalized one-dimensional distribution of scattered fields at different strain levels. Figure 7 shows the Fourier transform of some of the curves in Fig. 6.

The spectrum width u_s may be measured using the autocorrelation functions as shown in Fig. 7. Figure 8 shows the experimental results on the relations between flow stress, plastic strain and spectrum width in terms of spatial frequency. The experimental results in general agree well with the theoretical predictions [eqs (31) and (32)], confirming the feasibility of the proposed method for surface correlation length and plastic strain/flow stress evaluation for metallic materials.

Discussion and Conclusion

For the material and plastic deformation range studied, experimental evidence suggests that surface correlation length decreases in proportion to flow stress and plastic strain. This relation is found to be analogous to that between flow stress, plastic strain and microscopic cell size.

The theoretical derivation shows that the Fourier transform of the scattered field is proportional to the autocorrelation function of surface height variation and hence can be used for measuring surface correlation length, flow stress, and plastic strain. The merits of this method are noncontact, nondestructive, and remote. Experimental results in general support the analytical prediction.

It seems that the experimental results do not agree well with the theoretical prediction in the early stages of plastic

deformation ($\gamma < 0.25$ percent), which may have been caused by the initial surface conditions due to polishing. Because a calibration process is needed, the imperfection in initial surface conditions may influence the value of the constants involved making this technique sensitive to the surface-preparation process.

Acknowledgment

Financial support provided by the Army Research Office, Engineering Science Division through Contract No. DAAL0388K0083 is gratefully acknowledged.

References

1. Miyagawa, K. and Azushima, A., "Measurement of Plastic Deformation by Means of a Laser Beam," 9th World Conf. on Non-destructive Testing, Australia, 1-7 (1979).
2. Azushima, A. and Miyagawa, M., "Measurement of Plastic Zone Around a Fatigue Crack of a Stainless Steel Structure Member with a Laser Beam," Proc. 1982 Joint Conf. on Exp. Mech., Hawaii, 709-713 (1982).
3. Xu, B.Q. and Wu, X.P., "Measuring Plastic Deformation on Metal Surface and Determining the Elasto-plastic Boundary from the Change of Laser Produced Diffraction Patterns," Acta Mechanica Sinica, 16 (2), 175-184 (1984).
4. Chiang, F.P., Qian, H.H. and Don, H.S., "Determination of Plastic Strain Using Laser Speckles," 2nd Int. Symp. on Plasticity & Its Current Applications, Japan (1989).
5. Kato, A., Dai, Y.Z. and Chiang, F.P., "Damage Monitoring of Aluminum by Laser Speckle," Proc. 9th Int. Conf. on Exp. Mech., Durumak, 1940-1948 (1990).
6. Chiang, F.P., Dai, Y.Z., Xu, B.Q. and Kato, A., "Study of Surface Roughening Under Different Stress Modes by Correlation," Proc. 1990 SEM Spring Conf. on Exp. Mech., 36-42 (1990).
7. Bennett, J.M. and Matzson, L., Introduction to Surface Roughness and Scattering, Opt. Soc. of Amer., Washington, D.C. (1989).
8. Dieter, G.E., Mechanical Metallurgy, McGraw-Hill Book Co., Inc (1968).
9. Holl, D.L., "Dislocation Cell Formation in Metals," J. Appl. Phys., 41 (8), 3197-3201 (1970).
10. Chien, C.Y. and Duffy, J.D., "Strain Rate History Effects and Observations of Dislocation Substructure in Aluminum Single Crystals Following Dynamic Deformation," Brown Univ. Rep. MRLE-137 (1981).
11. Goodman, J.W., "Statistical Properties of Laser Speckle Patterns," Topics in Applied Physics, ed. J.C. Dainty, 9 (64), (1973).
12. Goodman, J.W., Introduction to Fourier Optics, McGraw-Hill (1968).
13. Rosenfeld, A. and Kak, A.C., Digital Picture Processing, 2nd Ed., Academic Press Inc. (1982).
14. Rasigni, G. and Rasigni, M. et al., "Statistical Parameters for Random and Pseudo-random Rough Surfaces," J. Opt. Soc. of Amer. (A), 5 (1), 99-103 (1988).
15. Dai, Y.Z. and Chiang, F.P., "The Application of Light Scattering to Plastic Strain Estimation," presented at 4th Int. Symp. on Non-destructive Characterization of Materials, Annapolis (1990).
16. Dai, Y.Z. and Chiang, F.P., "Fatigue Monitoring by Laser Speckle," accepted for publication, Int. J. of Fatigue (1991).

Strain Path and Surface Roughness

Y.Z. Dai* F.P. Chiang†

August 10, 1991

Introduction

Different aspects of the roughening of free surfaces due to plastic deformation have been reported.¹⁻⁴ The major conclusion is that the free surface roughening is: 1. dependent of equivalent plastic strain, average grain size, temperature, and; 2. independent of hydraulic pressure, material property and stress mode. The present work is designed to experimentally study the influence of strain path on the roughening of free surfaces.

Two kinds of specimens made of copper alloy (C26800-H01, yield strength: $264MP_a$, elastic modules: $110GP_a$) were used. One was $6.5mm$ thick for both tension and compression loadings. The other was $1.65mm$ thick for producing different strain paths through uniaxial tensile load(Fig.1). Prior to loading both kinds of specimens were polished to a final surface finish of $0.03\mu m$ in

*Senior Engineer with Research & Development Department, Brown & Sharpe Manufacturing Co., North Kingstown, RI 02852

†Leading Professor with Department of Mechanical Engineering, SUNY at Stony Brook, Stony Brook, NY 11794

terms of *rms* roughness. Strain gage rosettes were used to monitor surface strain of the specimens. The specimens were loaded at an average strain rate of $25\mu\epsilon/sec$. The three strain components ϵ_{45} , ϵ_{90} , ϵ_{135} were measured and used in calculating the principle strain ϵ_1 , ϵ_2 and the effective strain ϵ_e at any instant during loading.

Surface roughness was measured by means of a stylus type mechanical profilometer to yield the surface roughness characteristics. Surface roughnesses of different specimens plastically deformed to an effective strain of 2.3% were compared. Some of the specimens were deformed to this effective strain value by a single loading step, while others were first deformed 1.5% first and then to the final effective strain value. The principal strain ratio was fixed for the former case, whereas it varied for the latter.

Results

In order to find out the dependence of surface roughness on strain path, a simple tension-compression test was carried out. The idea is to compare the surface roughness of a specimen at the same plastic deformation level reached via two strain paths. The tension-compression specimen was first loaded by tension to an effective plastic strain of 0.75%. It was then plastically deformed by tension to an effective plastic strain of 2.4%. Finally, the specimen was compressed back to a final effective strain of 0.75%. Surface roughness was measured twelve times at each effective strain level. The av-

eraged surface roughness is given in Table 1 where the symbol s represents the standard deviation of the data. It is seen that the surface *rms* roughness are quite different. The surface *rms* roughness at 0.75% plastic strain level after the tension-compression process is larger than that obtained directly by tension. It is, however, smaller than that at 2.4% effective plastic strain.

It should be noted that a reverse yielding occurred during the tension-compression process. Most of the disk specimens (Fig.1) were used to test the surface roughening by different strain path without reverse yielding. As the loading angle α changed, which was achieved by loading the specimen through different pairs of holes in the specimen along different directions, the increasing rate of the three strain components varied. Figure 2 shows the ϵ_1 versus ϵ_2 corresponding to different loading angles. As the loading angle α changed from -15 to 115 degrees, the ratio of principal strain ϵ_1/ϵ_2 varied from -0.922 to 0.233 according to experimental observations. Therefore different strain paths were achieved. The final effective strain was obtained by either a one-step loading such that the principal strain ratio ϵ_1/ϵ_2 was constant or a two-step loading such that the principal strain ratio varied. The final effective strain was set at 2.3% and the intermediate effective strain was set at 1.5%. The elastic strain component was found to be less than 5% of the total strain and therefore was neglected.

The surface *rms* roughness and its standard deviation are shown in Table

2 for specimens plastically deformed via different strain paths. It is seen that the surface *rms* roughness is very much the same for all the strain paths except the one via the loading direction 115° to -15° . The reverse yielding occurred for this particular strain path as indicated in Fig.2.

Conclusion

Through the experimental study for the material studied under the experiment conditions, we may conclude that the roughening of free surfaces subject to plastic deformation is dependent of strain paths; however, that the influence of strain path on the surface roughness is not obvious if no reverse yielding occurs.

Acknowledgment

We would like to thank the Army Research Office, Engineering Science Division for financial support through contract No.DAA03-88-K-0033.

References

- [1] Osakada, K. and Oyane, M. (1971), "On the roughening of free surface in deformation process", *Bulletin of the JSME* 14, 171.
- [2] Thomson, P.F. and Shafer, B.V. (1982), "The roughening of free surface during plastic working", *Int. J. of Mach. Tool Des. Res.* 22, 261.

- [3] Azushima, A. and Miyagawa, M. (1986), "Effect of working factors and metallurgical factors on roughening phenomenon on free surface of sheets", *J. of the Japan Soc. for Tech. of Plasticity* 27, 1261.
- [4] Chiang, F.P., Dai, Y.Z., Xu, B.Q. and Kato, A.(1990), Study of surface roughening under different stress modes by correlation, in: *Proceedings of Hologram Interferometry and Speckle Metrology*, Baltimore, 36.

Table 1. Surface Roughness Comparison
(the tension-compression specimen)

ϵ_e (%)	0	0.75	2.4	0.75
<i>rms</i> (μm)	0.03	0.06	0.18	0.11
<i>s</i>	0.005	0.015	0.048	0.031

Table 2. Surface Roughness Comparison
(the disk specimens)

α (Degree)	90,0	0,90	30,60	40	0	90	75	75,30	-15,45	115,-15
<i>rms</i> (μm)	.147	.151	.154	.143	.151	.142	.148	.143	.133	.180
<i>s</i> (μm)	.026	.022	.011	.027	.014	.017	.020	.021	.056	.029

Captions .

Figure 1 Schematic configuration for strain path effect study

Figure 2 Principal strain paths

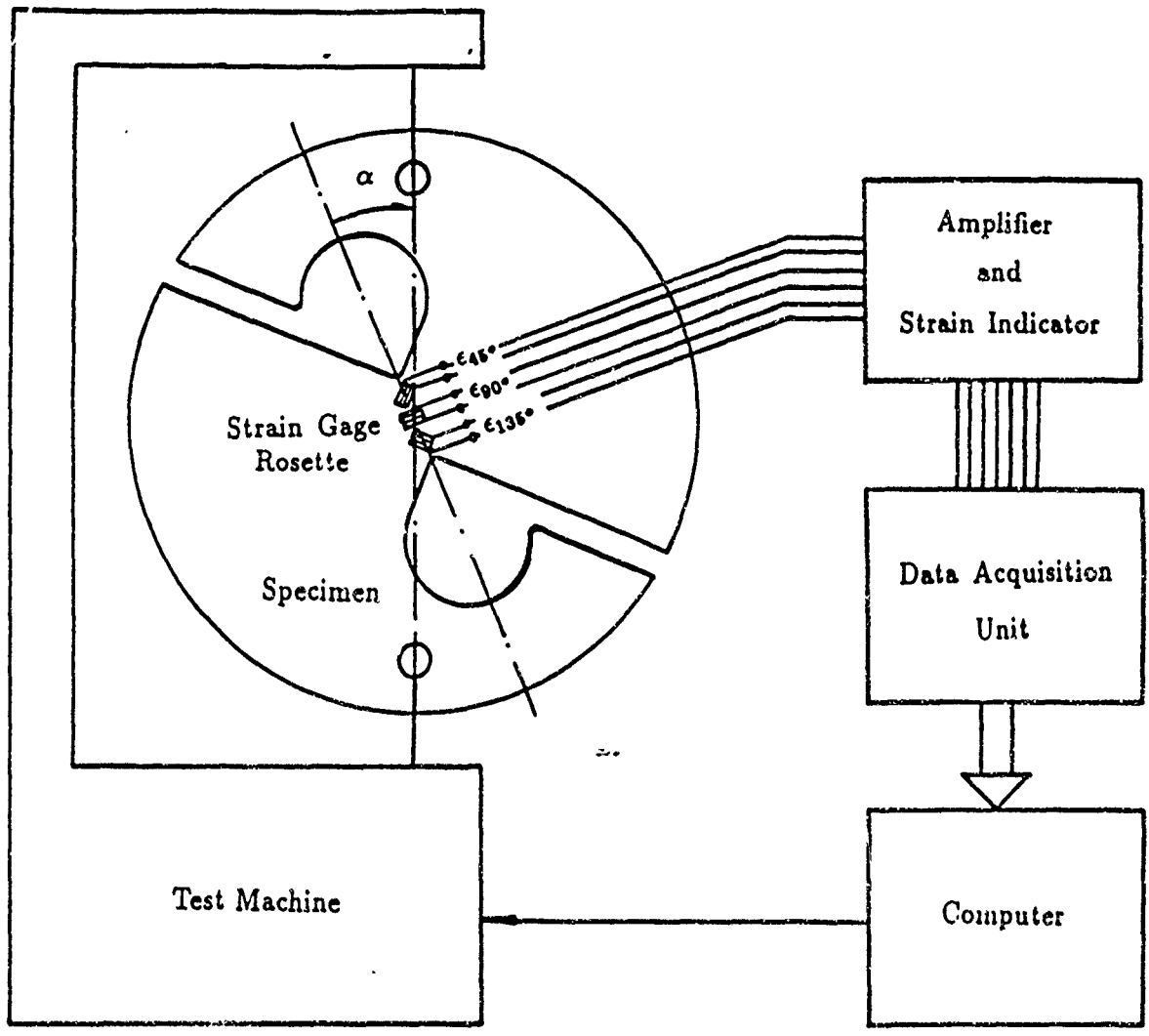


Figure 1 Schematic configuration for strain path effect study

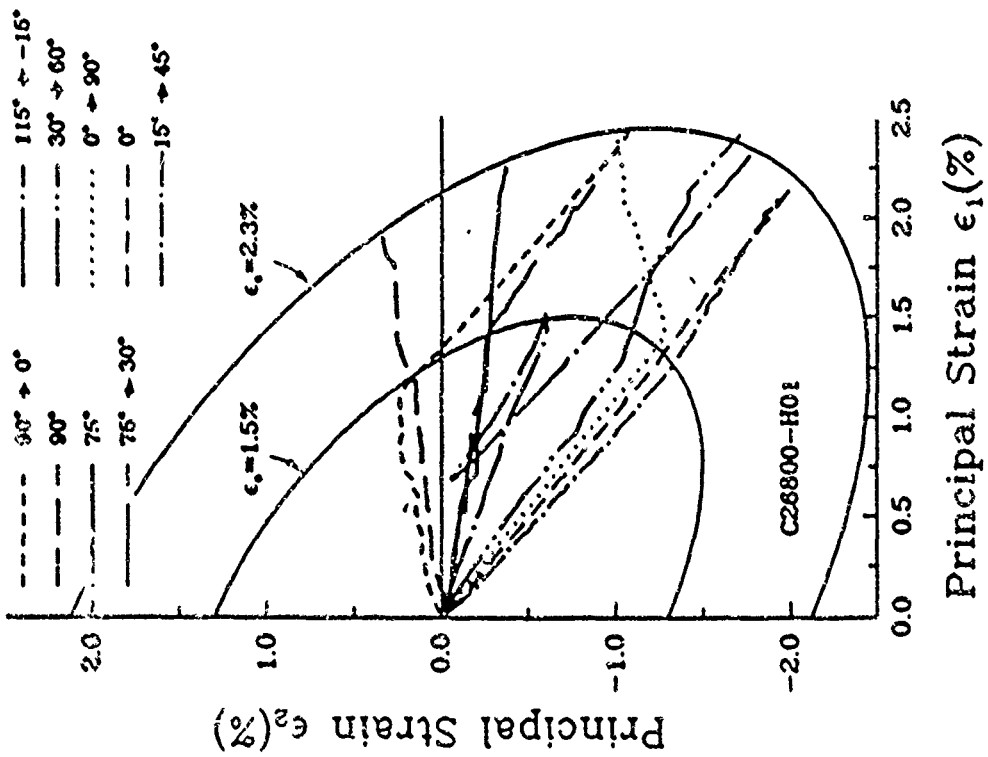


Figure 2 Principal strain paths

Contouring by Moiré Interferometry

by Y.Z. Dai and F.P. Chiang

ABSTRACT—A double-exposure moiré-interferometry technique for topographic contour measurement of an arbitrarily curved object is presented. A curved surface coated with light-sensitive material is exposed twice in a volume of virtual gratings formed by the interference of two coherent light beams split from a laser. An adequate rotation of the curved surface relative to the virtual grating between the two exposures produces moiré fringes which reveal topographic contour, or contour under some conditions, of the surface. The advantage of the present method in comparison with others is that it offers both reasonably good fringe quality and easily adjustable high sensitivity. The sensitivity of the technique is shown to be from the order of micrometer to that of millimeter depending on the frequency of the virtual grating and the amount of the relative rotation. This technique was successfully applied to the topographic contour measurement of a cylindrical shell with and without a diametrical point loading.

The principle of this paper and some early results were presented at the SPIE conference held at Dearborn, MI on June 27-30, 1988 and appeared in its proceedings.¹

Introduction

Some of the existing methods for measuring topographic contours include the shadow moiré method,^{2,3} the projection moiré method,⁴ the holographic method⁵ and the laser-speckle method.⁶ The main disadvantage of shadow and projection moiré methods is the lack of sensitivity; of the holographic method, the sophisticated optical arrangement; of the laser-speckle method, the poor fringe quality. In the present work, a double-exposure moiré-interferometry technique is described for topographic contour measurement which offers both high sensitivity and good fringe quality.

Moiré interferometry⁷ as a highly sensitive displacement measuring method is widely used in experimental-mechanics research. However, it has been limited so far to in-plane problems except when it is combined with the holographic method.^{8,9} The disadvantage of such a combination is twofold: First, it makes the corresponding optical arrangement very sophisticated; Secondly, it cannot offer good quality fringes compared with those by classic moiré or moiré interferometry methods because of the speckle

effect. This drawback can be overcome by the present technique which shall be called 'Double-exposure Moiré Interferometry' in the following discussion.

Like any other moiré method, double-exposure moiré interferometry also needs two sets of gratings to form moiré fringes. These two sets of gratings are all formed on a specimen surface, which is coated with a light sensitive medium, by exposing the specimen to a volume of virtual gratings formed by the interference between two intersecting coherent light beams split from a laser. If the specimen rotates an adequate angle relative to the virtual grating between the two exposures, then the two sets of gratings formed on the specimen surface will interfere to form moiré fringes which reveal the topographic contour of the object. Specimen deformation between the exposures will produce moiré fringes too, making the deformation also measurable by this technique.

The scope of this paper will include discussion on the fringe-forming mechanism; topography-evaluating method; the sensitivity and range of measurement; the condition under which this method provides contour; and some experimental demonstrations.

Fringe-forming Mechanism

When two collimated coherent light beams split from a laser meet in space as shown in Fig. 1, they interfere constructively and destructively such that the light-intensity distribution is uniform in planes with surface normals in the plane formed by these two beams and perpendicular to line \overline{OC} . Therefore, a volume of dark and light sheets in the common space of those two beams is formed and called a virtual grating which will cast gratings on any objects in the volume. According to the geometry shown, the pitch p of such a grating in an arbitrary plane π is

$$p = \frac{\lambda}{\sin i_A + \sin i_B} \tag{1}$$

or

$$p = \frac{\lambda}{2 \sin \alpha \cos \theta} \tag{2}$$

where p is the grating pitch in the π plane, λ is the wavelength of the illuminating light source; α is half the

Y.Z. Dai (SEM Member) is Graduate Student and F.P. Chiang (SEM Fellow) is Professor, State University of New York at Stony Brook, Department of Mechanical Engineering, Stony Brook, NY 11794-2300
Original manuscript submitted: November 11, 1989 Final manuscript received: January 10, 1991

$$\epsilon_{\theta i}'' = \frac{1}{E} [\sigma_{\theta i}'' - \nu(\sigma_{x i}'' + \sigma_{z i}'')]$$

$$u_i'' = -\frac{D_m}{2} \epsilon_{\theta i}''; \quad \Delta u_i'' = u_{i+1}'' - u_i''$$

Appendix II

If H_k can be considered proportional to F_k , eq. (5) gives

$$G_1 \xi_{i+1} + G_2 \xi_i + G_3 \xi_{i-1} = G_4 \quad (5)$$

where

$$\xi_i = \frac{\sum_{k=1}^i F_k}{F}$$

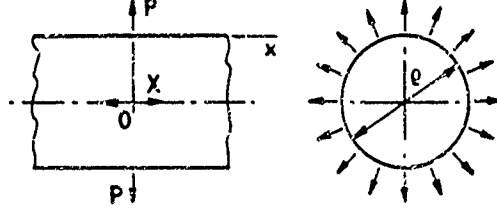


Fig. 4—Cylinder loaded by radial force distributed along a circular section

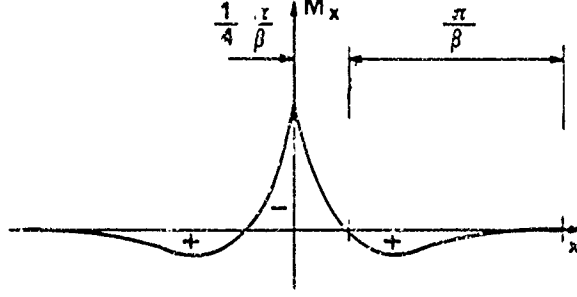


Fig. 5—Bending moment due to radial loading distributed along a circular section

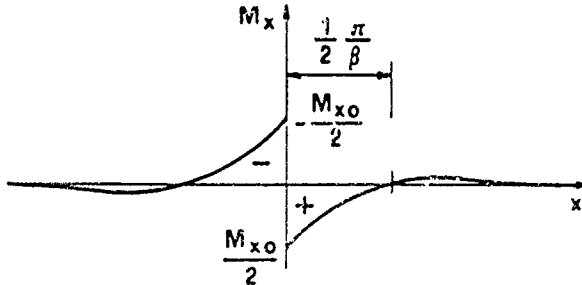


Fig. 6—Bending moment due to a moment distributed along a circular section

$$G_1 = \frac{\tan \gamma D_m}{2E} \left[\nu \left(\frac{1}{A'} + \frac{1}{A''} \right) - \frac{2 \tan \gamma}{\pi D_m l} \frac{D_m^2}{D_2^2 - D_m^2} \right] - 2h$$

$$G_2 = \frac{2 \tan \gamma}{\pi D_m l} \frac{D_m^2}{D_2^2 - D_m^2} \left(\frac{\tan \gamma D_m}{E} - \nu \frac{l}{E} \right) + \left(\frac{1}{A'} + \frac{1}{A''} \right) \cdot \left(\frac{l}{E} - \nu \frac{\tan \gamma D_m}{2E} \right) + 4h$$

$$G_3 = -\frac{2 \tan \gamma}{\pi D_m l} \frac{D_m^2}{D_2^2 - D_m^2} \left(\frac{\tan \gamma D_m}{2E} - \nu \frac{l}{E} \right) - 2h$$

$$G_4 = \frac{l}{EA'}$$

$$h = \frac{H_k}{F_k}$$

Equation (6) is a finite-difference equation with constant coefficients. It may be integrated taking into account the following conditions.

$$\xi_0 = 0 \quad \text{when } K = 0$$

$$\xi_n = 1 \quad \text{when } K = n$$

The F_k/F ratios can be evaluated from the ξ_k values.

Appendix III

Consider an indefinite cylinder loaded by radial forces uniformly distributed along a circular section (see Fig. 4). According to the general theory of shells:²³

$$M_x = -\frac{P}{4\beta} e^{-\beta x} (\cos \beta x - \sin \beta x) \quad (7)$$

where

$$\beta^4 = \frac{12(1-\nu)^2}{\rho^2 s^3}$$

In the above equations P is the radial load per unit length, s is the wall thickness, ρ the mean diameter; x , always positive, is the distance along the axis of symmetry from the cross section where the load is applied; M_x is the bending moment per unit length on the cross section at x , positive when it produces compression in the outside surface of the cylinder and tension in the inside surface (see Fig. 5).

In the same way, if the load is a bending moment M_{x0} uniformly distributed along a circular section, assuming that in this section radial displacement is nil (see Fig. 6):

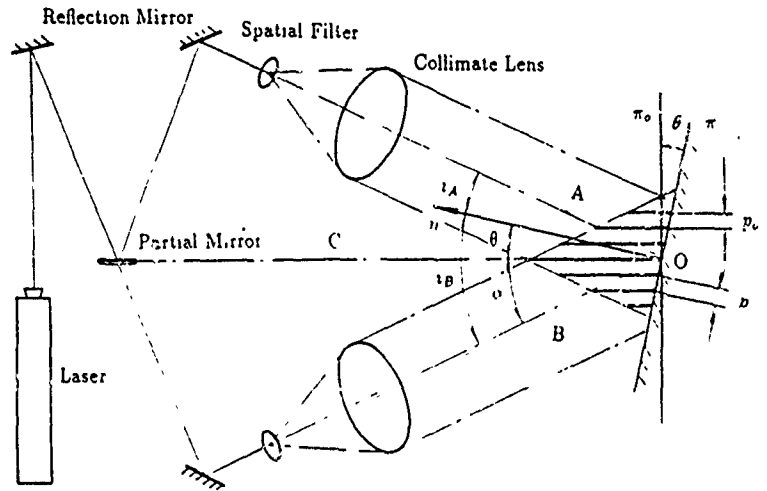
$$M_x = \pm \frac{M_{x0}}{2} e^{-\beta x} \cos \beta x \quad (8)$$

where the sign is positive for the right side and negative for the left side of the cylinder.

In both cases it can easily be seen²³ that all stress components and displacements almost vanish for $x \leq \pi/\beta$ ($x \geq 25$ mm for the female member wall). Therefore, if the loads are far enough from the ends of a finite cylinder, eqs (7) and (8) can be used.

By superimposing the effects of all radial loads P_k and moments M_{0k} , the bending moments at the $S_1 \dots S_n$ cross sections were determined applying the above solutions.

Fig. 1—An optical configuration for moiré interferometry



illuminating angle between the two light beams; and θ is the angle between the normal direction of the specimen surface and the bisecting direction OC of the two light beams. When $\theta = 0$, which corresponds to symmetric illumination ($i_A = i_B = \alpha$), we have

$$p = p_0 = \frac{\lambda}{2 \sin \alpha} \quad (3)$$

where p_0 is the grating pitch in the π plane (surface normal OC) and it is a constant once the optical system is fixed. Equation (2) can be simplified by utilizing the relation given by eq (3), yielding

$$p = \frac{p_0}{\cos \theta} \quad (4)$$

The corresponding grating frequency is

$$f = \frac{1}{p} = \frac{\cos \theta}{p_0} \quad (5)$$

Equations (4) and (5) show that grating pitch p and frequency f in the π plane will change as a function of angle θ , indicating that two sets of gratings with a different number of gratings could be registered on a certain region of an object if there is a change in θ between the two gratings' registration. The change in θ could be due to a rotation of the specimen relative to the virtual grating.

If a curved surface coated with a light-sensitive medium undergoes the same procedure, then the two sets of gratings will interfere to form a fringe pattern on the specimen surface. This fringe pattern reveals the surface topographic contour via some mathematical expressions which shall be derived in the later sessions. This is the principle of topographic contour measurement by double-exposure moiré interferometry.

Measurement of Uniform Deformation

Double-exposure moiré interferometry may be used for the measurement of uniform radial expansion or shrinkage of cylindrical surfaces even without any rotation between the two exposures. A cylindrical surface is exposed to the virtual grating as shown in Fig. 2. Suppose there are n gratings formed on arc $\hat{c}f$. Then, for some reason, the

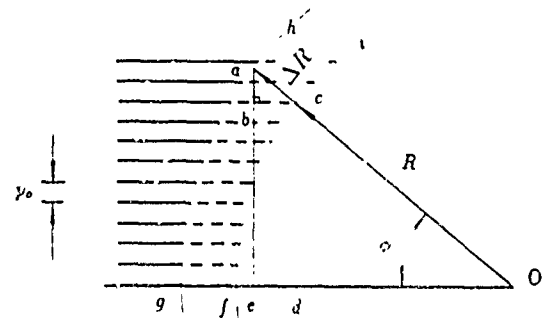


Fig. 2—Uniform radial expansion of a cylindrical shell

shell expands from $\hat{f}ci$ to $\hat{g}ah$ with point f moving to g and point c to a . Because of the expansion, arc $\hat{f}c$ which has already been exposed to a total number of n gratings now is expanded to arc $\hat{g}a$ and becomes capable of registering a total number of m gratings. The difference in grating numbers will create fringes which are related to the amount of shell expansion.

Obviously, the grating pitch on the shell surface is a function of angle θ . However, the projection of surface-grating pitch on the vertical line $\hat{e}a$ is the same, i.e., pitch p_0 . By following a similar analysis as that for the shadow-moiré technique,¹⁰ we can write the number of fringes formed on arc $\hat{g}a$ as

$$Np_0 = mp_0 - np_0 = \overline{ea} - \overline{dc} = \Delta R \sin \phi \quad (6)$$

Dividing both sides of the above equation by $\sin \phi$ yields

$$\Delta R = \frac{Np_0}{\sin \phi} \quad (7)$$

At a given point a , p_0 is a constant and N can be determined from the fringe pattern obtained. In order to evaluate ΔR , we must know the value of ϕ which depends on how the fringe pattern is being recorded. If the optical axis of the recording lens coincides with line $\hat{g}dO$, then arc $\hat{g}a$ is recorded as line $\hat{e}a$ which is defined as x . The following

relation exists

$$\sin \phi = \frac{x}{R + \Delta R} \quad (8)$$

Solving for ΔR from the above two equations, we get

$$\Delta R = \frac{RNp_o}{x - Np_o} \quad (9)$$

For the case of uniform shell shrinkage, the initial radius is $R = \overline{Oa}$, $x = \overline{dc}$ and $\sin \phi = x/R$. Substitute these into eq (7), we get

$$\Delta R = \frac{NRp_o}{x} \quad (10)$$

N in eqs (9) and (10) is ordered in such a way that it is zero when $x = 0$ and it increases for shell expansion while decreases for shrinkage, making ΔR bear the right sign. To evaluate ΔR , one just has to choose a fringe, number it, measure its x coordinate, then substitute them along with the values of R and p_o into eq (10).

Because x can be almost as large as R and N can be as small as one, ΔR has nearly the same magnitude as that of pitch p_o . While the magnitude of the latter can be as small as that of the wavelength of the illuminating light source indicating that the sensitivity of this method is quite high. The lower limit of sensitivity of this technique depends on the requirements on the density of fringes and the radius of the shell. Suppose $N/x = 10$ (fringe spacing is 0.1 mm) and $R = 100$ mm, then the lower limit is $100p_o$. Noting that the value of p_o is generally about one micrometer, the measurable shell radial deformation range is approximately between one micrometer to one millimeter.

The above derivation was made under the condition that the fringe pattern was recorded when the shell is at the deformed state. The same analyzing method can also be applied to the case that the fringe pattern is recorded when the shell has resumed its original dimension after the two exposures, only to yield slightly different expressions.

Measurement of Surface Topography

Double-exposure moiré interferometry can be used directly for measuring topographic contour. A specimen coated with photosensitive material is first exposed to a volume of virtual grating and then it is exposed to the same virtual grating again after either the optical system or the specimen is rotated a certain angle. These two exposures will make two sets of gratings on the specimen surface. They will in general interfere with each other to form a moiré fringe pattern which is related to the topography of the specimen through certain mathematical expressions.

The form of such expressions depends on the way the optical system or specimen is rotated and the setup of the coordinate system. Three cases and the corresponding topography-evaluation expressions will be studied in the coming sessions followed by a discussion.

Rotating the Optical System

Suppose the optical system is rotated clockwise an angle α in the $X-W$ plane (Fig. 3) between the two exposures. Without losing generality, point e is chosen to be the origin of the coordinate system and a is the point

of interest. Using the conventional moiré fringe analyzing method and denoting np_o by x , we get

$$Np_o = mp_o - np_o = x \cos \alpha + w \sin \alpha - x \quad (11)$$

Solving eq (11) for w ,

$$w = \frac{Np_o + x(1 - \cos \alpha)}{\sin \alpha} \quad (12)$$

where N is the fringe order at point a and it is zero when $x = w = 0$.

Applying the same analysis for the left part of the X axis, we get the following expression.

$$w = \frac{Np_o - x(1 - \cos \alpha)}{\sin \alpha} \quad (13)$$

Note that x is the length of mp_o , and should be positive for both sides of the X axis. We conclude that the w evaluation expression is slightly different for the different sides of the X axis. This difference can be negligible because the term $x(1 - \cos \alpha)$ is very small compared with Np_o for small α , making the fringes represent contours of equal w coordinates, which will be discussed later.

If the fringe pattern is recorded via a lens with its optical axis coinciding with the W axis, then x in the above expressions can be measured from the photographed fringe pattern. The values of p_o and α are also known from the optical setup and the amount of rotation. Therefore, w can be evaluated.

For the case that the optical system rotates counterclockwise, the w evaluation expressions will be the same as eqs (12) and (13) except that the sign of the term $x(1 - \cos \alpha)$ will be just the opposite.

Rotating the Specimen

More often than not, it is easier to rotate a specimen rather than the optical system between the two exposures. Because of the nature of relative motion, we do expect to see the same expressions as eqs (12) and (13). However, we will go through the derivation for rotating a specimen about an arbitrary point in the following discussion.

Curve \widehat{adc} on an arbitrary object rotates an angle α counterclockwise about point O to $\widehat{a'b'c'}$, as shown in Fig. 4. The difference in grating numbers yields the following relation.

$$Np_o = mp_o - np_o = x - \overline{cb} \sin(\theta - \alpha) \quad (14)$$

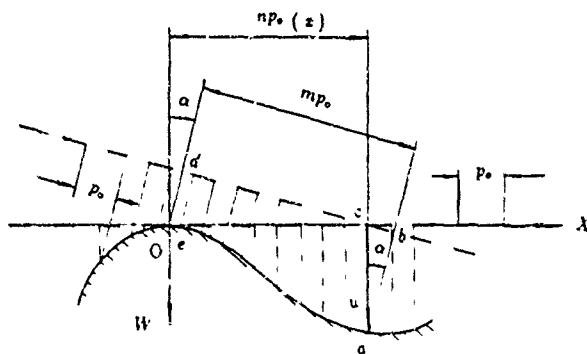


Fig 3—Rotating the optical system

Noting that $\sin(\theta - \alpha) = \sin \theta \cos \alpha - \cos \theta \sin \alpha$; $\overline{cb} = c'b'$, and $cb \sin \theta = x$, $\overline{cb} \cos \theta = w$, we can rewrite eq (14) as

$$w = \frac{Np_0 - x(1 - \cos \alpha)}{\sin \alpha} \quad (15)$$

The term $x(1 - \cos \alpha)$ bears a negative sign because the rotation of the specimen is counterclockwise and the coordinate system is fixed at the final position of the object (Fig. 4). If we change the direction of rotation or the side of the X axis under investigation, we shall get the same result as that discussed in the previous sessions.

The fact that the expressions derived for the general case are the same as those for special cases indicates that the rotating object and its center of rotation have no effect on the evaluation of topography. While the rotating direction or the change in the part of the X axis will only change the sign of a term in the expression.

Using a Periphery Camera

So far only the conventional camera has been utilized for fringe-pattern recording. The disadvantages of using such a camera for fringe-pattern recording on a curved surface is that not all points could be focused sharply because w and hence the objective distance changes from

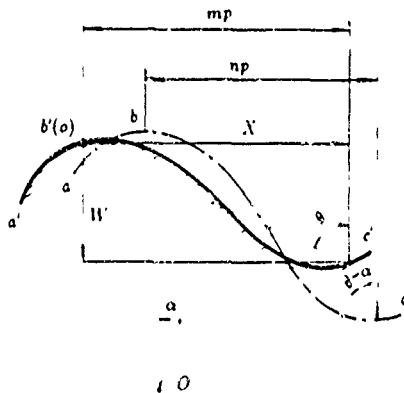


Fig 4—Rotating the specimen about an arbitrary point

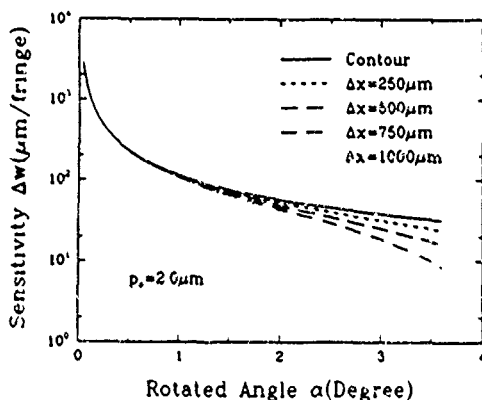


Fig 5—Sensitivity versus rotated angle at various fringe spacings

point to point making the determination of the x coordinate erroneous. Besides, w variation induced perspective effect also introduces error. Errors like these may be negligible for slightly curved surfaces but must be taken into consideration for cylindrical or conical surfaces with small radii.

In order to reduce this kind of error, a so called periphery camera¹¹ may be used for fringe-pattern recording on cylindrical, conical or near cylindrical, near conical surfaces. A periphery camera is capable of developing the circumference of a cylindrical or a conical surface into a plane. In doing so, x in the above expressions will be replaced by another variable l which is a direct measure from the fringe pattern photographed by the periphery camera. For a cylindrical surface, $l = R\phi$, $x = R \sin \phi$ and $\sin \phi = \sin(l/R)$. Thus, eq (15) can be rewritten in terms of l as

$$w = \frac{Np_0 - R \sin(l/R)(1 - \cos \alpha)}{\sin \alpha} \quad (16)$$

In the following discussion, all the fringe patterns were recorded by such a periphery camera.

Further Discussion

Equations (12), (13) and (16) describe fringe equations for the topographic contours instead of contours because they are dependent on the x coordinate. They represent contour only if the second terms in the numerators are negligible compared with the first terms Np_0 , yielding

$$w = \frac{Np_0}{\sin \alpha} \quad (17)$$

As shall be seen, such an approximation is valid for small rotating angles and small fringe spacing.

The sensitivity of this technique is determined by the difference in w between two adjacent fringes. Taking eq (15) as an example, the sensitivity is given by

$$\Delta w = w(N+1) - w(N) = \frac{p_0 - \Delta x(1 - \cos \alpha)}{\sin \alpha} \quad (18)$$

where Δx is the difference in x between two adjacent fringes (fringe spacing in the X direction).

Apparently, p_0 and α all influence the sensitivity. Pitch p_0 is normally about the order of micrometer and takes a fixed value once the optical arrangement is fixed. Then the most easily adjustable factor is angle α . Figure 5 shows the relation between Δw (logarithmic scale) and α numerically calculated according to eq (18) where the grating pitch p_0 is taken to be 2.0 μm . The solid line represents the sensitivity for w contour [eq (17)] defined as

$$\Delta w = \frac{p_0}{\sin \alpha} \quad (19)$$

The sensitivity Δw evaluated for different values of fringe spacing Δx is also plotted in Fig. 5. It is seen that for small α values the sensitivity is almost independent of Δx . When $\alpha \leq 1$ deg, for $\Delta x \leq 1$ mm, the error in sensitivity caused by using contour expression [eq (19)] instead of topographic contour expression [eq (18)] is less than eight percent.

Figure 5 also shows that the sensitivity can be adjusted to the order of micrometer by making the rotation angle be about 3.5 deg for $\Delta x = 1$ mm. The lower limit of the

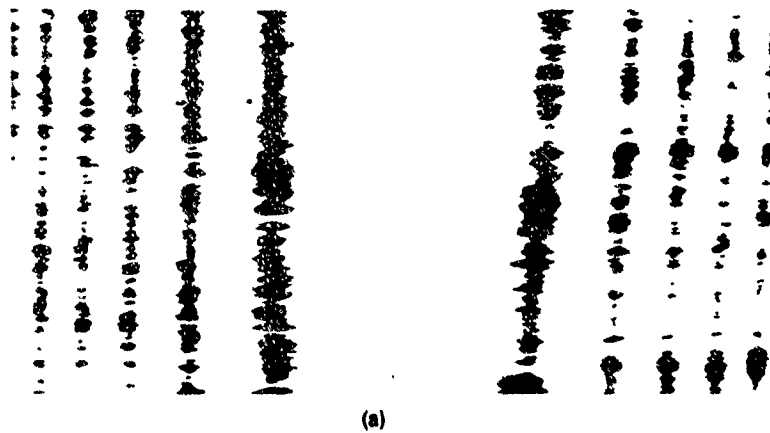
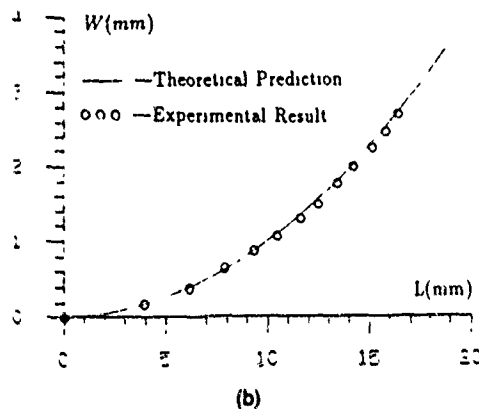


Fig. 6—Topographic contour study of a cylinder. (a) Topographic contour fringe pattern ($R = 25.4$ mm, $p_s = 0.7$ μ m, $\alpha = 0.17$ deg); (b) comparison



sensitivity, as shown in Fig. 5, can go as much as a few millimeters for sufficiently small angle rotations.

Equation (15) can be rewritten as

$$w = N \frac{p_s - (1 - \cos \alpha)x/N}{\sin \alpha} \quad (20)$$

where x/N is the average fringe spacing. Noting that eq (20) is analogous to eq (18), we can conclude that Fig. 5 and the result of the above error analysis discussion can be readily applied to the comparison between topographic contour [e.g., eq (16)] and w contour [eq (17)] as well. Thus, the fringes obtained by the present technique approximately represent contours of equal w coordinates for small angle rotations ($\alpha \leq 1$ deg).

For a flat plane, Δw is zero and $p_s = p \cos \alpha$, eq (18) becomes

$$\Delta x = \frac{p_s}{1 - \cos \alpha} = \frac{p_s p}{p - p_s} \quad (21)$$

This is exactly the same as the expression for fringe-spacing evaluation of a mismatch fringe pattern of classic moiré method.¹⁰

Experimental Investigation

The validity of the above derivation was verified by some experiments carried out on cylindrical shells with and without diametrical loading. The specimens were prepared according to a procedure described elsewhere¹² except that no aluminum coating was applied to the present work.

Figure 6(a) ($R = 25.4$ mm, $p_s = 0.7$ μ m, $\alpha = 0.17$ deg, $\Delta w = 230$ μ m) is the fringe pattern of a uniform cylinder recorded by a periphery camera where the horizontal axis is no longer x but rather arc length l . The small circles in Fig. 6(b) are w coordinates calculated by eq (16) based on the fringe pattern obtained [Fig. 6(a)]; while the solid line represents w coordinates calculated according to the measurement of shell geometry by

$$w = R(1 - \cos \theta) \quad (22)$$

Apparently, the experimental results agree well with the calculated one.

Figures 7 and 8 show some experimental results for nonuniform cylindrical shells. Figure 7 is the moiré fringe pattern of a cylindrical shell with some initial imperfection



Fig. 7—Topographic fringe pattern of a shell with some imperfection

tion. Figure 8(a) ($R = 50.2$ mm, $p_0 = 1.66$ μ m, $\alpha = 0.6$ deg, $\Delta w = 157$ μ m) is the fringe pattern of a shell under diametrical point loading.

Applying this double-exposure moiré-interferometry technique twice, once before and once after specimen is deformed, one will get two sets of moiré fringe patterns respectively. Comparing these two, one can find the w displacement due to the load alone. Figure 8(b) is such a comparison where the dots represent the w coordinates of a shell with the diametrical load and the solid line represents that without. The difference in w gives the displacement due to the loading alone. Therefore, surface-topographic contour or w displacement, no matter it is uniform or not, may be measured by double-exposure moiré interferometry.

Conclusion

An experimental technique, along with some theoretical derivation and experimental verification, for the determination of topographic contour, or contours of equal w coordinates for small angle rotations, based on a double-exposure moiré-interferometry technique, is presented. Its applications to the accurate determination of shell uniform radial deformation and of topography of arbitrarily curved surfaces are demonstrated.

The sensitivity of the technique is shown to be from the order of micrometers to that of millimeters. It can be easily adjusted by controlling the amount of rotation of the specimen relative to the optical system between the two exposures. Within a certain rotating limit, which is about four deg for the test conditions in the present work, the larger the rotated angle, the higher the sensitivity and *vice versa*.

The w evaluation expression is independent of the rotating object and its rotating center between the exposures. However, the change in rotating direction or the side of the X axis will alter the sign of a term in the expression. This difference is negligible for small angle rotations.

The fringe pattern produced by this method in general yields surface-topographic-contour because the coordinate x is involved in the expressions for the w coordinate evaluation. However, for very small angle relative rotations, this involvement is negligible indicating that the moiré fringes obtained by this technique represent contours of equal w coordinates.

The major drawback of this technique is that sometimes it is difficult to record the fringe pattern with the same contrast on the whole specimen surface because of the grating diffraction effect. Besides, *in situ* adjustment of fringe density such as that of the shadow moiré method is not possible with the present technique. This means that either a rough knowledge about the curved surface or a trial and error process is needed before a satisfactory fringe pattern can be obtained.

Acknowledgments

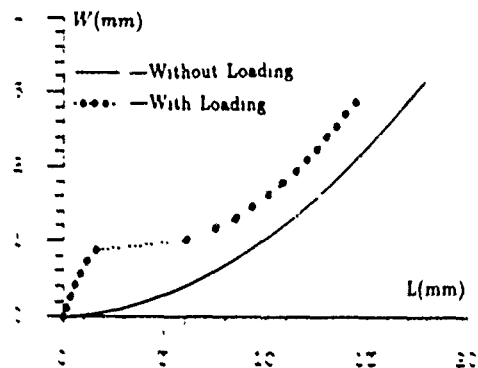
Financial supports provided by the Office of Naval Research, Solid Mechanics Division through contract No. N0001482K0566, and Army Research Office, Engineering Science Division through contract No. DAAL0388K0083 are acknowledged.

References

1 Dai, Y.Z. and Chiang, F.P., "Moiré Interferometry Applied to Topographic Contour Measurement," *Proc SPIE*, 954, 153-159 (1988)



(a)



(b)

Fig. 8—Topographic contour study of a shell under diametrical loading. (a) Topographic contour fringe pattern ($R = 50.2$ mm, $p_0 = 1.66$ μ m, $\alpha = 0.6$ deg); (b) comparison

2. Meadows, D.W., Johnson, W.O. and Allen, J.B., "Generation of Surface Contours by Moiré Patterns," *Appl. Opt.*, 9 (4), 942-947 (1970)
3. Takasaki, H., "Moiré Topography," *Appl. Opt.*, 9 (6), 1467-1472 (1970).
4. Khetan, R.P., "Theory and Application of Projection Moiré Methods," *PhD Thesis*, State University of New York at Stony Brook (1975).
5. Hildebrand, B.P. and Haines, K.A., "Multiple-Wavelength and Multiple-Source Holography Applied to Contour Generation," *J Opt Soc. of Amer.*, 57 (2), 155-162 (1967).
6. Jaishigh, G.K. and Chiang, F.P., "Contouring by Laser Speckle," *Appl Opt.*, 20 (19), 3385-3387 (1981)
7. Post, D., "Development in Moiré Interferometry," *Opt Eng.*, 21 (3), 458-467 (1982).
8. Basehore, M.L. and Post, D., "Moiré Method for In-plane and Out-of-plane Displacement Measurement," *EXPERIMENTAL MECHANICS*, 21 (9), 321-328 (1981).
9. Beranek, W.J. and Bruinsma, A.J.A., "Determination of Displacement and Strain Fields Using Dual-beam Holographic-moiré Interferometry," *EXPERIMENTAL MECHANICS*, 20 (9), 317-323 (1982)
10. Chiang, F.P., "Moiré Methods of Strain Analysis," *Manual on Experimental Stress Analysis, Third Edition Chapt 6*, SESA 51-69 (1979)
11. Chiang, F.P., Juang, R.M. and Dai, Y.Z., "Application of Moiré and Speckle Techniques to Curved Surfaces Using a Periphery Camera," *Proc SPIE*, 814, 249-256 (1987).
12. Anastasi, R.F., Dai, Y.Z. and Chiang, F.P., "Simplified Procedure for Obtaining High-Frequency, Highly Reflective Specimen Gratings for Moiré Interferometry," *EXPERIMENTAL TECHNIQUES*, 12 (9), 16-17 (1988)

Small-crack Closure Measurements in Titanium Alloys

by J.M. Larsen and J.R. Jira

ABSTRACT—An automated interferometric displacement gage was used to monitor crack-mouth-opening behavior of naturally initiated small surface cracks in a series of titanium alloys having a range of microstructures and deformation characteristics. Findings indicate that the transient development of crack closure plays a significant role in the early propagation of small fatigue cracks.

Introduction

After several years of active research to study the propagation of small fatigue cracks, a significant body of literature dealing with this subject has emerged (see Refs. 1-11). It has been shown that, when described by the linear-elastic stress-intensity-factor range, ΔK , small cracks may propagate significantly faster than nominally equivalent large cracks and that small cracks often do not exhibit the normal threshold stress-intensity-factor range, ΔK_{th} , that limits the propagation of large cracks. Small cracks often have been categorized by one or more descriptive terms which further classify them as being: mechanically, microstructurally, physically, or chemically small.¹² Mechanically small cracks are defined as having a length of the order of the crack-tip plastic-zone size, while microstructurally small cracks have a length of the order of the dominant microstructural dimension. Due to its size, a crack that is simply physically small may behave differently than an otherwise equivalent large crack and, finally, a small crack may respond differently to a chemically aggressive environment than a corresponding large crack. Although there are numerous factors that may contribute to the disparity between small- and large-crack growth rates, recent evidence indicates that differences in crack-closure behavior of small versus large cracks are often of primary importance.

Crack closure, which has been the subject of a number of review articles¹³⁻¹⁵ and a symposium,¹⁶ may be briefly

described as a wedging action that occurs in the wake of a fatigue crack, interfering with the normal opening and closing motion of the crack. A crack that experiences the remotely applied value of $\Delta K = K_{max} - K_{min}$, may actually experience an effective stress-intensity-factor range, $\Delta K_{eff} = K_{max} - K_{cl}$, where K_{cl} is the stress-intensity factor at crack closure. Thus, for situations where K_{cl} is substantially greater than K_{min} , the crack-growth behavior may be significantly affected by crack closure. Since closure operates in the wake of a fatigue crack, a finite crack length must exist before the mechanism can develop fully. It has been speculated that the absence of fully developed closure for small cracks is at least partially responsible for their 'anomalously' fast growth. The development of crack closure may be quite complex, however, since closure is produced by three primary mechanisms [associated with fracture-surface plasticity, oxides, and roughness (asperities)], which may operate simultaneously to various degrees depending on loading, material, environment, temperature, and crack size.

Experimental measurement of the closure behavior of small cracks is extremely difficult, however, which accounts for the very limited number of such data available in the literature. Recently, this difficult task has been accomplished using an interferometric displacement gage (IDG) developed by Sharpe.¹⁷ Larsen *et al.*¹⁸ demonstrated the utility of the IDG for small-crack testing on an advanced titanium alloy and have since applied this technique to a number of other alloys.¹⁹ Lee and Sharpe²⁰ and Su and Sharpe²¹ have used the IDG to investigate small-crack behavior in aluminum alloys. This paper will discuss the characteristics and attributes of the computerized interferometric displacement gage and present direct experimental measurements of the development of crack closure in surface flaws in a series of titanium alloys possessing a range of microstructures and deformation characteristics.

Experimental Procedure

The procedure for small-crack testing was specifically developed to allow small cracks to initiate naturally on a residual-stress-free specimen surface. All small-crack testing was performed on electropolished axial specimens containing a mild notch ($K_t = 1.037$), which served to localize crack initiation.²² The specimens were cycled at a

J.M. Larsen is Materials Research Engineer and J.R. Jira is Materials Engineer, U.S. Air Force - Materials Laboratory, WRDC/MLLN, Wright-Patterson AFB, OH 45433

Paper was presented at the 1988 SEM Spring Conference on Experimental Mechanics held in Portland, OR on June 6-10

Original manuscript submitted June 2, 1989. Final manuscript received January 14, 1991

Computer-Aided Speckle Interferometry (CASI): Part II. An Alternate Approach Using Spectral Amplitude and Phase Information

D.J. Chen, F.P. Chiang

Laboratory for Experimental Mechanics Research

Department of Mechanical Engineering

State University of New York

Stony Brook, NY 11794-2300

Y.S. Tan

Department of Mechanical Engineering

Xi'an Jiaotong University

Xi'an 710049, P.R.China

H.S. Don

Department of Electrical Engineering

State University of New York

Stony Brook, NY 11794-2350

ABSTRACT

An alternate approach of fully-automatic speckle interferometry is developed. Two speckle patterns of a specimen, one before and one after deformation, are registered by a video camera. These digital speckle patterns are segmented into many very small subimages. Analysis of corresponding subimage pairs is performed pointwise at various locations from both speckle patterns. Basic data process involves two-step fast Fourier transform (FFT). Local displacement vector is revealed by detecting a signal hill in the secondary spectral domain. Accurate characterization of displacement components is achieved by a biparabolic fitting and maximum-searching near the signal hill. Whole field displacement distribution is deduced by analysis of all subimage pairs. An artificial rigid shift between the two speckle patterns is introduced in case of very large displacement. The range of measurable displacement is unlimited as long as the two speckle patterns remain correlated. Incremental deformations in a dynamic loading process can be measured by registering more speckle patterns at consecutive deformation stages and analyzing each pair of successive speckle patterns. The technique has been applied to the study of crack tip deformation fields.

1 Introduction

Laser speckle interferometry is an ideal technique for the measurement of surface as well as internal deformation.¹ Being nondestructive and remote sensing, it is applicable to high temperature environment, static as well as dynamic deformations.² Its basic process involves speckle recording, specklegram developing, fringe pattern generation and analysis. Many automatic fringe pattern analysis methods have been developed in both the pointwise and the whole field analysis of the specklegram.³⁻¹⁰ However, specklegram developing is still a manual procedure which contradicts full automation of the whole process. There-

also exist unavoidable upper and lower displacement limits. The well developed electronic speckle pattern interferometry (ESPI) is an automatic technique for out-of-plane or in-plane displacement measurements by using different optical arrangements.¹¹ Phase value evaluation of the fringe patterns in the ESPI has been obtained using a phase-stepping technique.¹²⁻¹⁵ While high sensitivity is achieved in a whole field survey, the system is intolerant of environmental disturbances, measurable displacement is restricted in several wave lengths of the illuminating light, and the system is effective only for detection of one displacement component unless a particular treatment is given. Another automatic approach is the digital speckle correlation.¹⁶⁻²⁵ Reliable results of both in-plane displacement components and displacement gradients have been reported using the white-light, laser or acoustical speckles in a pointwise analysis fashion over wide range of deformation. The drawback is such a process is the large amount of computation required in the correlation evaluation, although some efforts have been taken to improve the computation.

In 1990, a new computer-aided speckle interferometry (CASI) was developed.²⁶ It retains all advantages of the conventional optical speckle interferometry and provides an extended range of measurement in cases of very large or very small displacements. Comparing with ESPI, it utilizes much simpler optical setup, provides complete 2-D displacement fields in a much larger deformation range, and is noise-resistant in industrial environments. While providing comparable range with digital correlation method, it achieves 20 times faster computing speed.

In this work, we develop a modified approach to the problem. Besides all the highlights of the previous approach, the new system requires no image-shifting between two speckle patterns even in case of very small displacements. Neither is the shifting required to determine signs of displacement components. Since only one signal hill is generated in the secondary spectrum, the whole spectral domain is more efficiently used in displacement estimation. The range of measurable deformation is about the same of the previous approach. While the amount of computation is 1.5 times of the first approach.

2 System and Procedures

The system used for data acquisition and image processing is shown in Fig. 1. The specimen is illuminated by a white-light or laser beam so that laser or white-light speckles are generated on its surface. Speckle patterns of the object are imaged by the objective and registered onto the sensor of a video camera. The light intensity of each speckle pattern is digitized into an array of 512×480 pixels with 8 bits gray levels by the DT-2861 Frame Grabber and stored into frame buffers or diskettes. The IBM-PC/AT controls the data acquisition, performs image processing, and interfaces with a laser printer.

Basic process of the technique involves data acquisition and data processing. In data acquisition, two speckle patterns of the specimen, one before and one after deformation, are captured by the video camera and stored in the computer.

Fig.2 depicts general procedures of the data processing in the program. Two registered speckle patterns are first segmented into 240 very small subimages (32×32 pixels in size). Each two corresponding subimages from both large images constitute a subimage pair and are to be analyzed to yield local displacement vector at the subimage location. A FFT is applied to both subimages to generate two complex spectra and a resultant spectrum is obtained by taking square-root of the product of the two spectral amplitudes as amplitude field and the phase difference between the two spectra as phase field. A second-step FFT is then applied to the resultant spectrum and a signal hill is generated in the secondary spectrum. The local displacement vector of the specimen at the subimage location is uniquely determined by detecting peak position of the signal hill in the secondary spectral domain. By analyzing all subimage pairs of the two large speckle patterns, 2-D displacement fields of the specimen can be deduced completely. More steps of incremental deformation in a dynamic loading process can be revealed by registering more speckle patterns at different deformation stages and analyzing each pair of successive registered speckle patterns. Strain fields can be obtained by differentiating the resulting u - and v -displacement fields.

3 Theoretical Basis

Suppose we have registered two speckle patterns of a specimen into the computer, one before and one after deformation (see Fig.2). These two speckle patterns are first segmented into two groups of very small (32×32 pixels) subimages. As a typical process in the pointwise recurring analysis, we examine one subimage pair, $h_1(x, y)$ and $h_2(x, y)$, from the two large images respectively. Since the subregion is very small, one can assume that displacement inside the subimage is uniform. Then if represent the first subimage by $h(x, y)$, the second subimage is a shifted function of $h(x, y)$, plus an uncorrelated random noise term, i.e.,

$$\begin{aligned} h_1(x, y) &= h(x, y), \\ h_2(x, y) &= h(x - u, y - v) + n(x, y), \end{aligned} \quad (1)$$

where (u, v) is local displacement vector, $n(x, y)$ is uncorrelated noise. By applying Fourier transforms to both subimages, two spectra of them, $H_1(f_x, f_y)$ and $H_2(f_x, f_y)$, respectively, are obtained as,

$$\begin{aligned} H_1(f_x, f_y) &= |H_1(f_x, f_y)| \exp[j\phi_1(f_x, f_y)] \\ &= \int \int_{\Delta} h(x, y) \exp[-j2\pi(xf_x + yf_y)] dx dy \\ &= |H(f_x, f_y)| \exp[j\phi(f_x, f_y)], \\ H_2(f_x, f_y) &= |H_2(f_x, f_y)| \exp[j\phi_2(f_x, f_y)] \\ &= \int \int_{\Delta} [h(x - u, y - v) + n(x, y)] \exp[-j2\pi(xf_x + yf_y)] dx dy \\ &= |H(f_x, f_y)| \exp\{j[\phi(f_x, f_y) - 2\pi(uf_x + vf_y)]\} + N(f_x, f_y), \end{aligned} \quad (2)$$

where (f_x, f_y) are spectral domain coordinates, Δ denotes the subimage region, $|H(f_x, f_y)|$ and $\phi(f_x, f_y)$ are, respectively, spectral amplitude and phase fields, and $N(f_x, f_y)$ is spectrum of the uncorrelated noise. From the two resulting spectra a new complex spectrum can be organized as

$$F(f_x, f_y) = H_1(f_x, f_y) H_2^*(f_x, f_y) / \sqrt{|H_1(f_x, f_y) H_2(f_x, f_y)|}, \quad (3)$$

or,

$$F(f_x, f_y) = \sqrt{|H_1(f_x, f_y) H_2(f_x, f_y)|} \exp\{j[\phi_1(f_x, f_y) - \phi_2(f_x, f_y)]\}. \quad (4)$$

In practice, the noise spectrum is a weak function. For simplification of analysis, we neglect $N(f_x, f_y)$ hereafter. By substituting Eq.(2) into Eq.(4), the resultant spectrum becomes

$$F(f_x, f_y) \approx |H(f_x, f_y)| \exp[j2\pi(u f_x + v f_y)]. \quad (5)$$

From Eq.(5) it is seen that the phase field of the resultant spectrum is a plane wave whose orientation characterize the local displacement vector (u, v) , and whose amplitude field equals the amplitude of both individual spectra.

To obtain a more apparent function of displacement vector (u, v) , $F(f_x, f_y)$ is again Fourier transformed into a secondary spectral domain (ξ, η) . The resulting secondary spectrum is

$$\begin{aligned} G(\xi, \eta) &= \int \int_{\Delta_f} F(f_x, f_y) \exp[-j2\pi(f_x \xi + f_y \eta)] df_x df_y \\ &\approx \int \int_{\Delta_f} |H(f_x, f_y)| \exp\{-j2\pi[f_x(\xi - u) + f_y(\eta - v)]\} df_x df_y \\ &= G_1(\xi - u, \eta - v) \end{aligned} \quad (6)$$

where Δ_f denotes the spectral domain, and $G_1(\xi - u, \eta - v)$ is an expanded impulse function located at (u, v) in the secondary spectral domain (ξ, η) . By detecting peak position of this impulse, local displacement vector can be uniquely determined.

Two typical signal hills obtained in the secondary spectral domain, one at zero displacement and one at nonzero displacement, is shown in Fig. 3. It is seen that the signal hill appears at the exact center of the secondary spectral domain at zero displacement (Fig.3a), while it moves out of the center when the local displacement is not zero (Fig.3b).

4 High accuracy by bipolarabolic fitting

The accuracy of the displacement components from direct detection of the signal hill in the discrete spectral domain is limited by the pixel size of the sensor array. For more accurate characterization, a subpixel detection of the displacement is necessary. In this work we employ a bipolarabolic least-square fitting algorithm near the signal hill.

Suppose $P_o(\xi_o, \eta_o)$ is the detected on-pixel maximum while $P_m(\xi_m, \eta_m)$ is the subpixel fine maximum position to be estimated (Fig.4). For simple description we select a new coordinate (ξ', η') as

$$\begin{aligned} \xi' &= \xi - \xi_o, \\ \eta' &= \eta - \eta_o. \end{aligned} \quad (7)$$

The fine maximum position in the new coordinate is $P_m(\xi'_m, \eta'_m)$, with

$$\begin{aligned} \xi'_m &= \xi_m - \xi_o, \\ \eta'_m &= \eta_m - \eta_o. \end{aligned} \quad (8)$$

The continuous function to be fitted near the signal hill can be interpolated by a Taylor's series around $P_m(\xi'_m, \eta'_m)$ as

$$G_f(\xi', \eta') = C_0 + C_1(\xi' - \xi'_m) + C_2(\eta' - \eta'_m) + C_3(\xi' - \xi'_m)^2 + C_4(\eta' - \eta'_m)^2 + C_5(\xi' - \xi'_m)(\eta' - \eta'_m) + C_6(\xi' - \xi'_m)^3 + C_7(\eta' - \eta'_m)^3 + C_8(\xi' - \xi'_m)^2(\eta' - \eta'_m) + C_9(\xi' - \xi'_m)(\eta' - \eta'_m)^2 + H.O.T., \quad (9)$$

where *H.O.T.* denotes higher order terms. In practice, the signal hill can be approximately considered as a symmetric function about $P_m(\xi'_m, \eta'_m)$ (see Fig.4). Thus several terms in Eq.(9) vanish. i.e.,

$$C_1 = C_2 = C_5 = C_6 = C_7 = C_8 = C_9 = 0, \quad (10)$$

$$C_3 = C_4.$$

In regard to the symmetry and neglecting *H.O.T.* in Eq.(9), the Taylor series is reduced to

$$G_f(\xi', \eta') = C_0 + C_3[(\xi' - \xi'_m)^2 + (\eta' - \eta'_m)^2], \quad (11)$$

or

$$G_f(\xi', \eta') = c_0 + c_1\xi' + c_2\eta' + c_3(\xi'^2 + \eta'^2), \quad (12)$$

where

$$c_0 = C_0 + C_3(\xi_m'^2 + \eta_m'^2)$$

$$c_1 = -2C_3\xi'_m$$

$$c_2 = -2C_3\eta'_m$$

$$c_3 = C_3. \quad (13)$$

Constants $c_0 \sim c_3$ can be estimated by least-square-error method. The square-error summation between the discrete spectrum $G(\xi', \eta')$ and the continuous fitting function $G_f(\xi', \eta')$ in a square region of $(2K + 1) \times (2K + 1)$ pixels around the signal hill is

$$\delta = \sum_{\xi'=-K}^K \sum_{\eta'=-K}^K [c_0 + c_1\xi' + c_2\eta' + c_3\xi'^2 + c_3\eta'^2 - G(\xi', \eta')]^2, \quad (14)$$

where K is a appropriate positive integer. At $\delta = \text{minimum}$, we obtain four constraint equations

$$\partial\delta/\partial c_i = 0, \quad \text{for } i = 0 \sim 3. \quad (15)$$

Let

$$M_{kl} = \sum_{\xi'=-K}^K \sum_{\eta'=-K}^K \xi'^k \eta'^l, \quad \text{for } k, l = 0 \sim 4. \quad (16)$$

and note that

$$M_{kl} = 0, \quad \text{when } k \text{ or } l \text{ is odd}, \quad (17)$$

Eq.(15) becomes

$$\begin{bmatrix} M_{00} & 0 & 0 & M_{20} + M_{02} \\ 0 & M_{20} & 0 & 0 \\ 0 & 0 & M_{02} & 0 \\ M_{20} + M_{02} & 0 & 0 & M_{40} + 2M_{22} + M_{04} \end{bmatrix} \begin{bmatrix} c_0 \\ c_1 \\ c_2 \\ c_3 \end{bmatrix} = \begin{bmatrix} q_0 \\ q_1 \\ q_2 \\ q_3 \end{bmatrix} \quad (18)$$

where $q_0 \sim q_3$ are constants estimated from the local region of the discrete spectrum, i.e.,

$$\begin{bmatrix} q_0 \\ q_1 \\ q_2 \\ q_3 \end{bmatrix} = \begin{bmatrix} \sum_{\xi'=-K}^{+K} \sum_{\eta'=-K}^{+K} G(\xi', \eta') \\ \sum_{\xi'=-K}^{+K} \sum_{\eta'=-K}^{+K} \xi' G(\xi', \eta') \\ \sum_{\xi'=-K}^{+K} \sum_{\eta'=-K}^{+K} \eta' G(\xi', \eta') \\ \sum_{\xi'=-K}^{+K} \sum_{\eta'=-K}^{+K} (\xi'^2 + \eta'^2) G(\xi', \eta') \end{bmatrix} \quad (19)$$

Note that the square matrix in Eq.(18) is a sparse matrix. coefficients $c_0 \sim c_3$ can be readily found by.

$$\begin{bmatrix} c_0 \\ c_1 \\ c_2 \\ c_3 \end{bmatrix} = \begin{bmatrix} (M_{40} + 2M_{22} + M_{04})/\Delta & 0 & 0 & -(M_{20} + M_{02})/\Delta \\ 0 & 1/M_{20} & 0 & 0 \\ 0 & 0 & 1/M_{02} & 0 \\ -(M_{20} + M_{02})/\Delta & 0 & 0 & M_{00}/\Delta \end{bmatrix} \begin{bmatrix} q_0 \\ q_1 \\ q_2 \\ q_3 \end{bmatrix} \quad (20)$$

where

$$\Delta = M_{00}(M_{40} + 2M_{22} + M_{04}) - (M_{20} + M_{02})^2. \quad (21)$$

When a local region of 5×5 pixels is used in fitting, we have $K = 2$, $M_{00} = 25$, $M_{20} = M_{02} = 50$, $M_{22} = 100$, $M_{40} = M_{04} = 170$ and $\Delta = 3500$ in Eq.(20). The solution of $c_0 \sim c_3$ is

$$\begin{bmatrix} c_0 \\ c_1 \\ c_2 \\ c_3 \end{bmatrix} = \begin{bmatrix} 27/175 & 0 & 0 & -1/35 \\ 0 & 1/50 & 0 & 0 \\ 0 & 0 & 1/50 & 0 \\ -1/35 & 0 & 0 & 1/140 \end{bmatrix} \begin{bmatrix} \sum_{\xi'=-2}^{+2} \sum_{\eta'=-2}^{+2} G(\xi', \eta') \\ \sum_{\xi'=-2}^{+2} \sum_{\eta'=-2}^{+2} \xi' G(\xi', \eta') \\ \sum_{\xi'=-2}^{+2} \sum_{\eta'=-2}^{+2} \eta' G(\xi', \eta') \\ \sum_{\xi'=-2}^{+2} \sum_{\eta'=-2}^{+2} (\xi'^2 + \eta'^2) G(\xi', \eta') \end{bmatrix} \quad (22)$$

From Eq.(13), the maximum subpixel position of the signal hill can be found as

$$\begin{aligned} \xi'_m &= -c_1/(2c_3), \\ \eta'_m &= -c_2/(2c_3). \end{aligned} \quad (23)$$

And the local displacement components are

$$\begin{aligned} u &= \xi_m = \xi_0 + \xi'_m, \\ v &= \eta_m = \eta_0 + \eta'_m. \end{aligned} \quad (24)$$

5 Results and discussion

In a simple experiment, a single edge-cracked aluminum specimen of length 225 mm, width 75 mm and thickness 3.4 mm under mode I loading condition was tested. The pre-fatigued crack was 25 mm in length. One surface of the specimen was slightly painted by white and black sprays alternately and illuminated by a white-light bulb after it was mounted onto a INSTRON loading cell. An area of about $20 \times 20 \text{ mm}^2$ around the crack tip was monitored by the video camera. A tensile loading was applied in two steps. Three digital speckle patterns were obtained at zero load, first and second loading levels, respectively (Fig. 5a-c). The u- and v-displacement distributions are obtained at 240 points in the inspected area at both loading levels. Result at each point was obtained by analyzing the corresponding two subimages of 32×32 pixels.

Displacement fields at load level I were obtained by processing the first and the second speckle pattern (see contours in Fig.6). And that at load level II were obtained by processing the first and the third speckle pattern (contours in Fig.7). Fig.6 and 7 show that the near-tip displacement fields are in good agreement with our earlier work done by moire. Strain fields may be obtained by differentiating the u- and v-displacement fields. And comparison of displacement and strain distributions with theoretical elastic and plastic crack tip solutions may be performed. These analyses are not included in this paper. The digital processing takes the IBM-PC/AT 12 seconds for one data point and 48 minutes for a 24 (16 x 15) points rectangular grid at present time. A two-time faster processing speed is going to be achieved by using the data symmetry in the 2-D FFT operation in our laboratory. More significant improvement may be achieved by using an updated computer or implementing the AT with an array processor.

6 Acknowledgments

This work was supported by the Army Research Office through Contract DAA20388K0033 (Scientific Officer: Dr. G.L. Anderson) and the Office of Naval Research through Contract N0001482K0566 (Scientific Officer: Dr. Yapa D.S. Rajapakse). These supports are gratefully acknowledged.

7 References

1. F.P.Chiang, "A new family of 2D and 3D experimental stress analysis techniques using laser speckles," SM Archives, Vol.3, pp.27-58 (1978).
2. F.P.Chiang, J.Adachi, R.Anastasi and J.Beatty, "Subjective laser speckle method and its application to solid mechanics problems," Optical Engineering, Vol.21, pp.379-90 (1982).
3. G.H. Kaufmann, A.E. Ennos, B. Gale and D.J. Pugh, "Electro-optical readout system for analysis of speckle photographs," Inst. of Phys., Vol.13, pp.579-84 (1980).
4. B. Ineichen, P. Eglin and R. Dandliker, "Hybrid optical and electronic image processing for strain measurements by speckle photography," Applied Optics, Vol. 19, pp.2191-5 (1980)
5. G.E. Maddux, R.R. Corwin and S.L. Moorman, "Improved automated data reduction device for speckle metrology," Proc. of SESA 1981 Spring Meeting, pp. 248-58 (1981).
6. R. Meynart, "Instantaneous velocity field measurements in unsteady gas flow by speckle velocimetry," Applied Optics, Vol. 22, pp.535-40 (1983).
7. D.W.Robinson, "Automatic fringe analysis with a computer image processing system," Applied Optics, Vol.22, pp.2169-76 (1983).
8. J.M.Huntley, "An image processing system for the analysis of speckle photographs," Journal of Physics, E:19, pp.43-49 (1986).
9. G.T.Reid, "Automatic fringe pattern analysis: a review." Optics and Lasers in Engineering, Vol.7, pp.37-68 (1986/7).
10. D.J.Chen and F.P.Chiang, "Digital processing of Young's fringes in speckle photography," Optical Engineering, Vol.29, pp.1413-20 (1990).
11. C.Wykes, "Use of electronic speckle pattern interferometry (ESPI) in the measurement of static and

- dynamic surface displacements," *Optical Engineering*, Vol.21, pp.400-6 (1982).
12. K.A. Stetson and W.R. Brohinsky, "Electrooptic holography and its application to hologram interferometry." *Applied Optics*, Vol.24, pp.3631-7 (1985).
 13. A.J. Moore and J.R. Tyrer, "An electronic speckle pattern interferometer for complete in-plane displacement measurement," *Measurement Science and Technology*, Vol.1, pp.1024-30 (1990)
 14. C.R. Mercer and G. Beheim, "Fiber optic phase stepping system for interferometry," *Applied Optics*, Vol.30, pp729-34 (1991)
 15. F.M. Santoyo, M.C. Shellabear and J.R. Tyrer, "Whole field in-plane vibration analysis using pulsed phase-stepped ESPI," *Applied Optics*, Vol.30, pp.717-21 (1991).
 16. W.H. Peters and W.F. Ranson, "Digital image techniques in experimental stress analysis," *Optical Engineering*, Vol. 21, pp.427-31 (1982).
 17. T.C.Chu, W.F.Ranson, M.A.Sutton and W.H.Peters, "Application of digital-image-correlation techniques to experimental mechanics," *Experimental Mechanics*, Vol.25, pp.232-44 (1985).
 18. N. Takai and T. Asakura, "Vectorial measurements of speckle displacement by 2-D electronic correlation method." *Applied Optics*, Vol.24, pp.660-5 (1985).
 19. I. Yamaguchi, "Automatic measurement of in-plane translation by speckle correlation using a linear image sensor," *Journal of Physics, E: Scientific Instruments*, Vol.9, pp.944-9 (1986).
 20. M.A. Hamed, "Object-motion measurements using pulse-echo acoustical speckle and two-dimensional correlation," *Experimental Mechanics*, Vol.27, pp.250-4 (1987).
 21. H.A. Bruck, S.R. McNeill, M.A. Sutton and W.H. Peters, " Digital image correlation using Newton-Raphson method of partial differential correlation," *Experimental Mechanics*, Vol. 29, pp. 261-7 (1989).
 22. C. Lee, Y.J. Chao, M.A. Sutton, W.H. Peters and W.F. Rason, "Determination of plastic strains at notches by image-processing methods," *Experimental Mechanics*, Vol.29, pp.214-20 (1989).
 23. D.J.Chen and F.P.Chiang, "Optimal sampling resolution and range of measurement in digital speckle correlation: Patt I. Laser speckle method," *Proceedings of SEM Spring Conference on Experimental Mechanics*, Cambridge, MA, pp. 133-8 (1989).
 24. Z.L. Kahn-Jetter and T.C. Chu, "Three dimensional displacement measurements using digital image correlation and photogrammic analysis," *Experimental Mechanics*, Vol.30, pp.10-16 (1990).
 25. D.J.Chen and F.P.Chiang, "Optimal sampling and range of measurement in displacement-only laser speckle correlation," to appear in *Experimental Mechanics*.
 26. D.J. Chen and F.P. Chiang, "Computer speckle interferometry," *Proc. of Int. Conf. on Hologram Interferometry and Speckle Metrology*, pp.49-58 (1990).

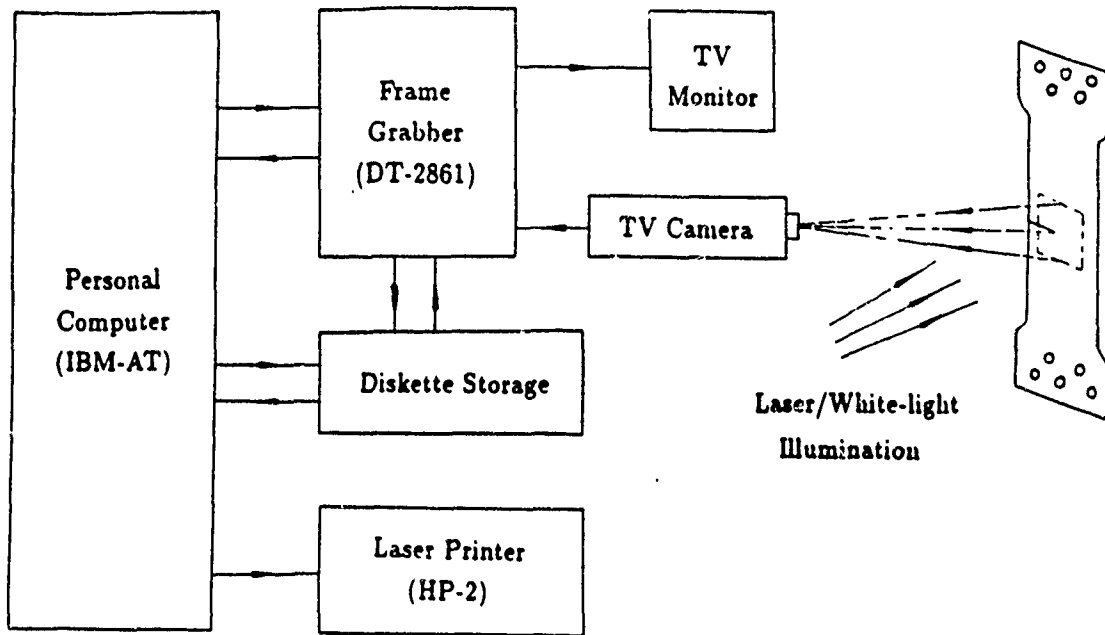


Fig.1 Schematic of data acquisition and image processing system

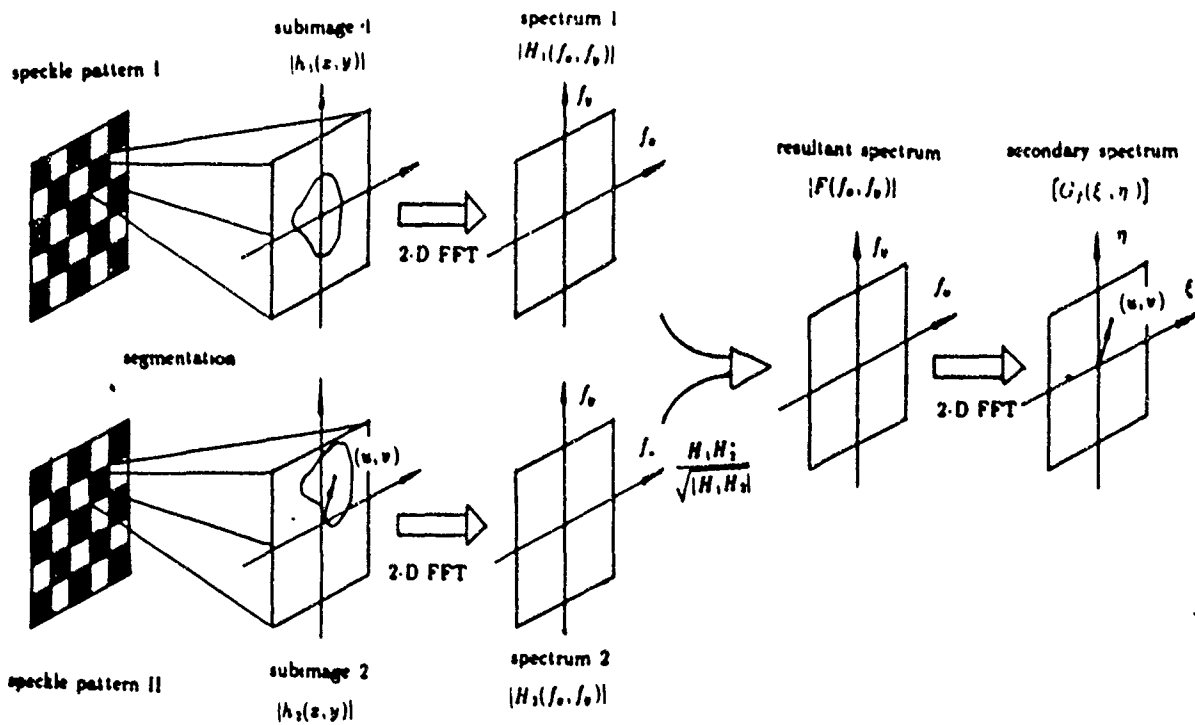


Fig.2 Data processing procedures

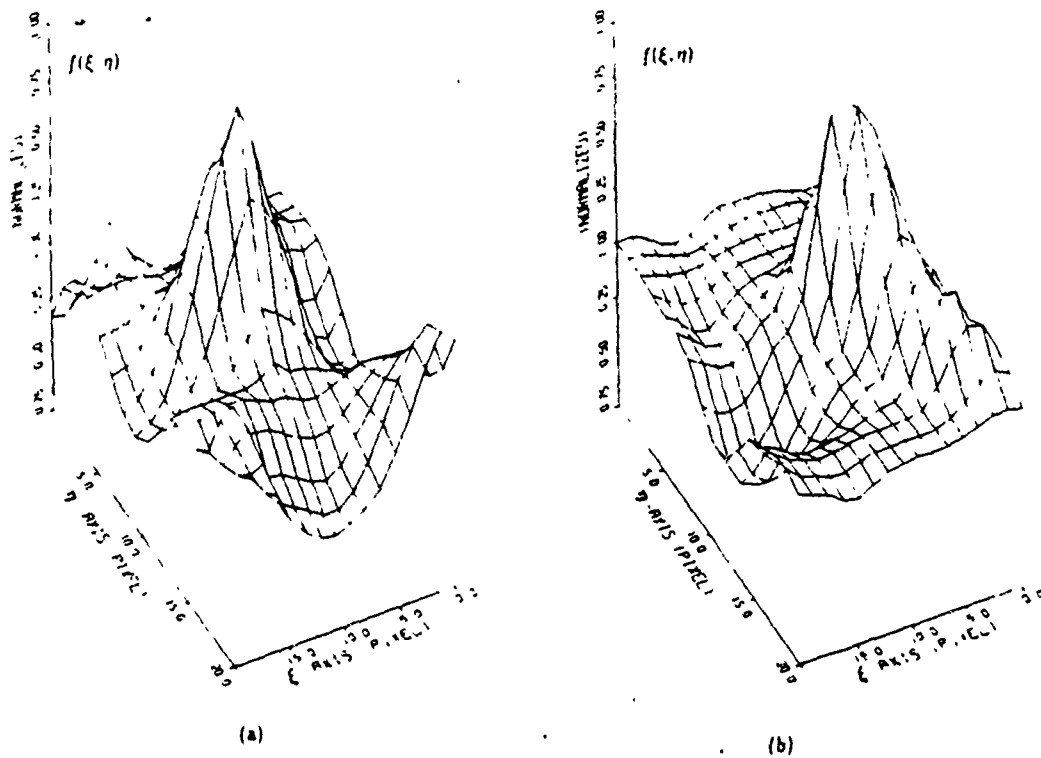


Fig.3 Typical signal hills obtained in secondary spectral domain when local displacements are (a) zero and (b) nonzero (central 16×16 pixels plotted only)

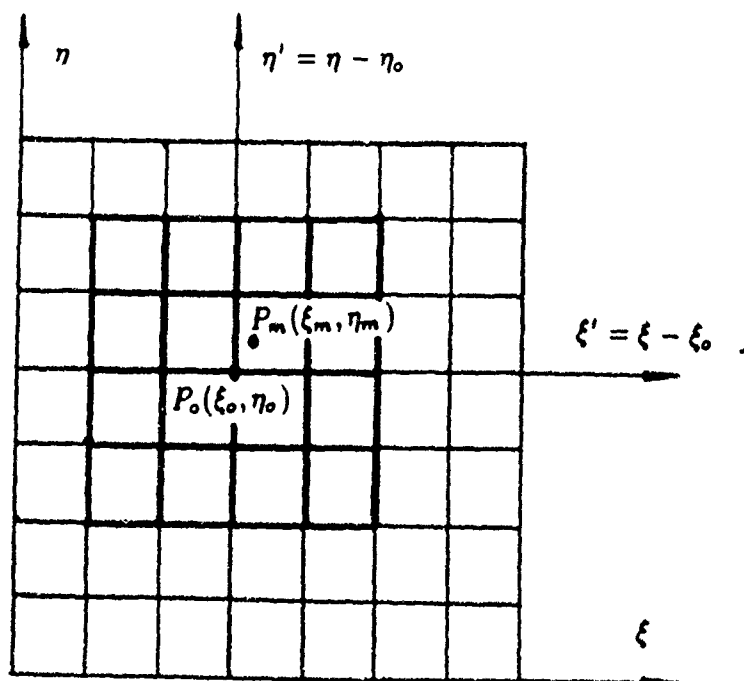
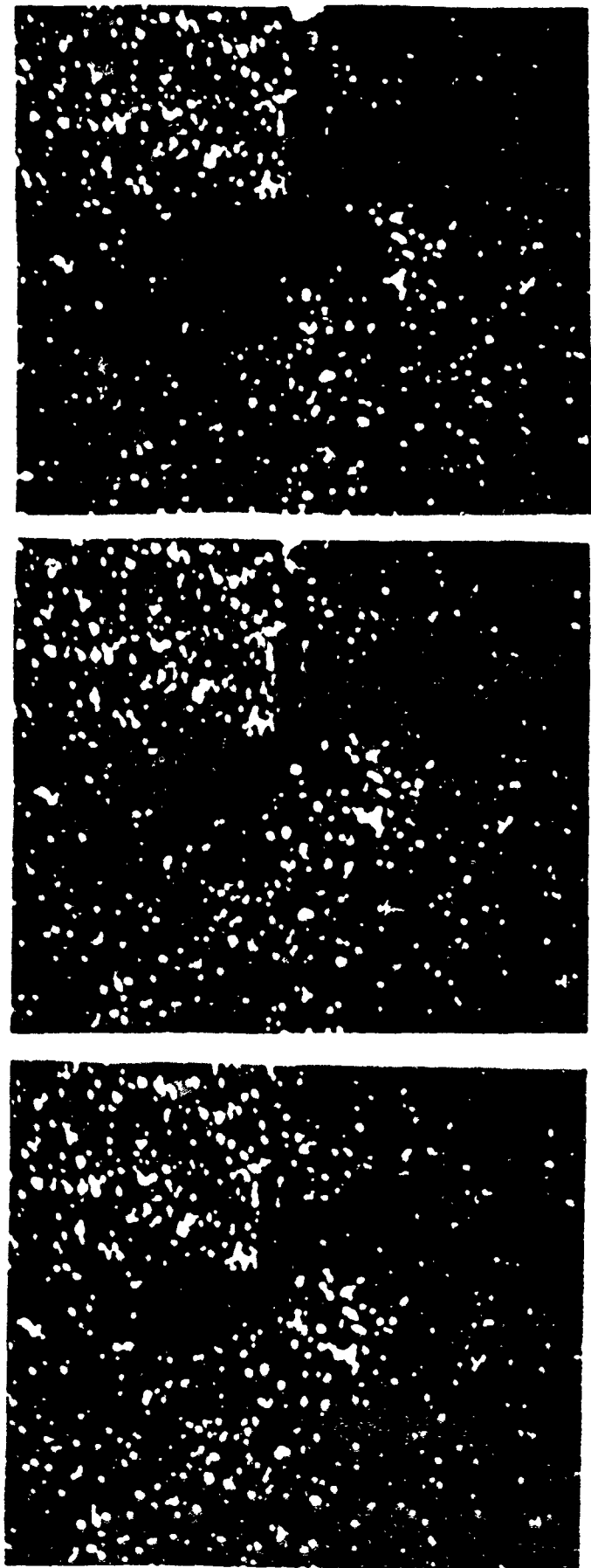


Fig.4 Subpixel maximum detection near a signal hill (highlighted region used in biparabolic fitting)



(a)

(b)

(c)

Fig.5 Digital speckle patterns around a crack tip at
(a) zero load, (b) load level I, and (c) load level II

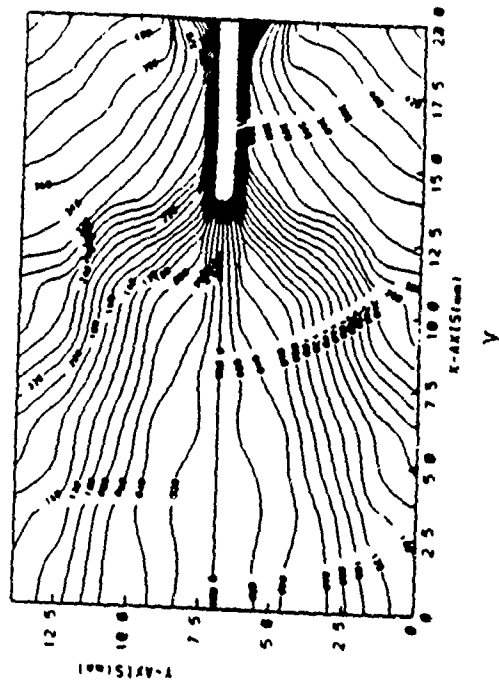
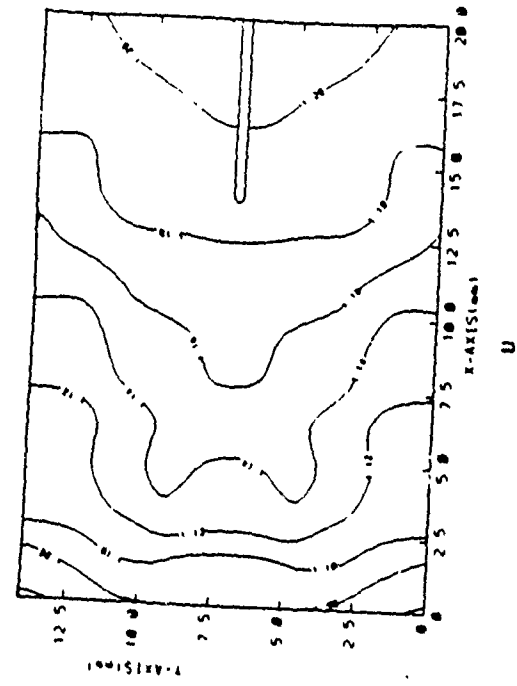


Fig. 7 u- and v-displacement contours of crack tip field at load level II (sensitivity=0.02mm/line)

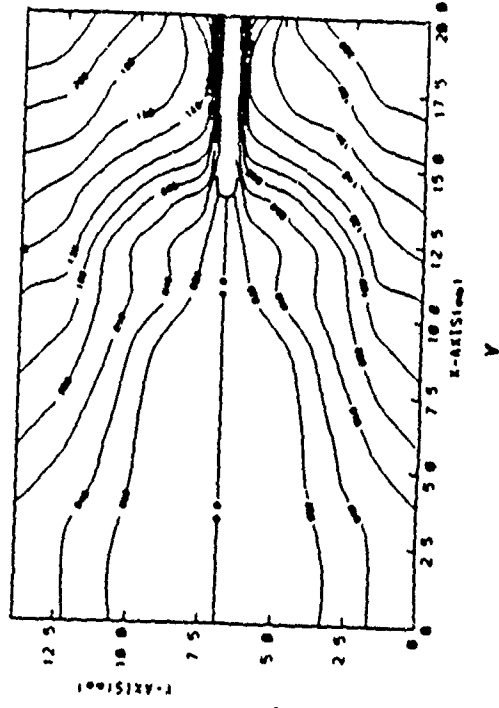
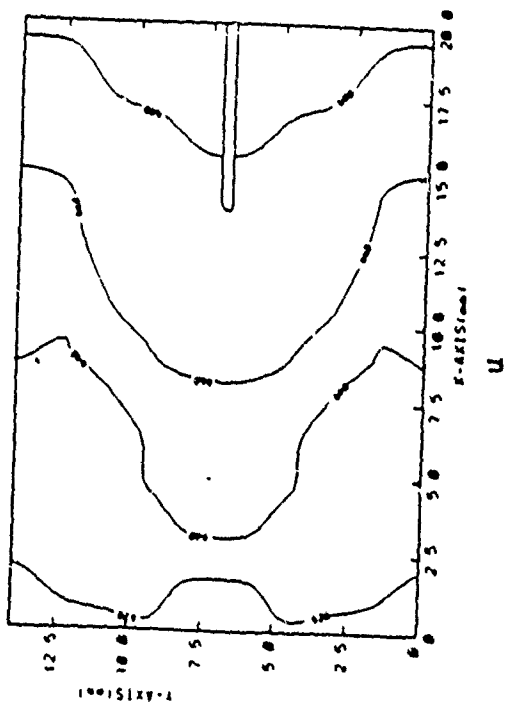


Fig. 6 u- and v-displacement contours of crack tip field at load level I (sensitivity=0.02mm/line)

Range of Measurement of Computer-Aided Speckle Interferometry (CASI)

D.J. Chen, S. Li, T.Y. Hsu and F.P. Chiang
Laboratory for Experimental Mechanics Research
Department of Mechanical Engineering
State University of New York
Stony Brook, NY 11794-2300

ABSTRACT

Measurable area, accuracy, and ranges of displacement, strain and rotation of the CASI are estimated. The measurable area for a given video camera is proportional to the area of each individual speckle. The displacement uncertainty is proportional to speckle size. The measurable displacement ranges from a few percent of speckle size to several centimeters. Strain and rotation up to 3 ~ 5% can be detected. Experimental results from CASI on crack opening displacement (COD) and normal strain distribution ahead of a crack tip of an aluminum specimen under mode I loading are presented.

1 Introduction

In the past two years, two approaches of computer-aided speckle interferometry (CASI) were developed.^{1,2} The fully automated processes, including speckle registration and information extraction, provide a whole field survey of a 2-D displacement in a pointwise fashion. The systems constitute simple and reliable optical as well as electrical setup. The methods retain all nondestructive features of conventional optical speckle interferometry and provide extended range.

In this work, we analyze the range of CASI. Although ranges of the two approaches are comparable, we concentrate our discussion on the second approach.² The system used is shown in Fig.1.

2 Measurable Area and Accuracy

Unlike the conventional speckle photography in which the frame resolution of the recording film is usually abundant, the digital resolution of a video camera is rather limited. It is true that the finer the resolution, the less the distortion resulting from a sampling process, but it reduces the actual area that can be measured with a given capacity of the camera. Therefore, there is a trade-off between the reliability and capability of the measurement.

In the earlier work of Chen and Chiang,³ an optimal sampling resolution of speckles was estimated with experimental verification. By that resolution, the relation between the sampling interval of the sensor array (T) and the subjective speckle size at the image plane (S) is

$$T \approx 0.4S. \quad (1)$$

For the convenience of further analysis, we convert this sampling rate into an expression on specimen, i.e.,

$$T_o \approx 0.4S_o, \quad (2)$$

where $T_o (= T/M)$ is the equivalent sampling interval and $S_o (= S/M)$ is the equivalent speckle size on specimen, and M the image magnification factor. It should be noted that S_o is approximately the objective speckle size when white light speckles are used.

The accuracy of the displacement components from direct detection of signal hill in the discrete spectral domain is limited by the pixel size of the sensor array.² For more accurate determination, a subpixel maximum searching for displacement components is required. This was done by a biparabolic fitting near the signal hill.² To our experiences, at the optimal sampling rate the uncertainty from the subpixel detection varies from 1% to 5% of the speckle size, depending on the degree of decorrelation between the two speckle patterns. Thus the unsolvable displacement on a specimen is

$$\delta_o = 0.01S_o \sim 0.05S_o. \quad (3)$$

Equation (3) indicates that the unsolvable displacement is proportional to the speckle size (see Fig. 2). Thus, for accurate measurement small speckles are preferred.

From Eq.(1) and (2), we see that the image magnification must be selected consistently with the sampling interval of the sensor and speckle size on the object, i.e.,

$$M = 2.5T/S_o. \quad (4)$$

This relation implies that the measurable area on a specimen is limited with a given video camera. For instance, with a camera resolution of 512×512 pixels the maximum measurable area on the specimen is

$$A_o = (512T_o)^2 \approx 42000S_o^2. \quad (5)$$

For speckles from 0 to 100μ , the dependence of the area on speckle size is shown in Fig. 2. It is seen that the measurable area is proportional to the area of each individual speckle on the object. Obviously, to measure large area, large speckles should be used. However, as shown above, larger speckles will also result in lower accuracy. In practice, speckle size must be selected judiciously by considering both effects.

3 Extended Range of Displacement

Local displacement of a specimen always brings about decrement of the correlated area in a subregion between two speckle patterns (Fig.3). We consider a general two-dimensional displacement case in Fig. 3. When both u and v approach $L_s/2$, where L_s is the subimage size, the correlated area in the subimage drops to one-quarter of the total subregion. More critically, when either u or v is larger than $L_s/2$, the signal hill will move out of the 32×32 pixel region in the secondary spectral domain and aliasing will arise.^{2,4} Thus, it seems that the measurable 2-D displacement on the specimen is limited by

$$\text{Max}(|u|, |v|) \leq L_s/(2M) = 16T/M \approx 6.4S_o. \quad (6)$$

However, in practice this displacement limit may be easily lifted by using an image-shifting technique. When there is a large rigid body displacement, an intentional rigid shift towards each other of the two speckle patterns can be introduced before the image segmentation. The real displacement vector at each subregion can be obtained by subtracting the shifted distance vector from the searched displacement vector. Thus, the measurable displacement is unlimited as long as the two speckle patterns remain well correlated. Our observations show that specimen displacement up to 20 ~ 40 mm in white speckle method, and 2 ~ 4 mm in laser speckle method can be measured successfully without dramatic decorrelation between the two speckle patterns.

4 Range of Strain and Rotation

Other decorrelation factors may arise from in-plane strain and rotation of a specimen (Fig.4). Let's first consider the effect of the local normal strains, ϵ_{xx} and ϵ_{yy} , in Fig.4a. In a special case, when the two normal strains are equal, the relation between the reference and deformed subimages is equivalent to a magnification mismatch between the two speckle patterns. Using numerically simulated images, Bailey et al have analyzed the effects of magnification mismatch on cross-correlation performance.⁵ In CASI, the equivalent effects will be the shortening and flattening of the signal hill in the secondary spectral domain.² Obviously, the wider is the signal hill and the smaller is its peak value, the lower are the reliability and the accuracy of the measurement. Therefore, there is a limited range of the measurable normal strain. Bailey et al have suggested a rule of thumb of magnification mismatch using simulated images.⁵ Applying that suggested rule to speckle images and noticing the equivalent magnification mismatch resulting from the normal strains, the following limit is found,

$$N_s |\epsilon| \leq 1, \quad (7)$$

where N_s (=32) is the array size of the subimage and $\epsilon = \epsilon_{xx} = \epsilon_{yy}$. From Eq. (7) it is seen that the relative displacement of a border pixel to the center of the subimage should not exceed one-half pixel width. Recall that the speckle size is two-and-half pixel at the optimal sampling rate, it implies that the relative displacement of the border elements cannot exceed one-fifth of the typical speckle size. In general, the two normal strains are rarely equal. A more conservative rule of thumb for general normal strains may be selected as,

$$\text{Max}(|\epsilon_{xx}|, |\epsilon_{yy}|) \leq 1/N_s. \quad (8)$$

Recall that the subimage size is 32×32 , then the limit of the normal strains becomes,

$$\text{Max}(|\epsilon_{xx}|, |\epsilon_{yy}|) \leq 3.1\%. \quad (9)$$

At this limit, there are nine (3×3) on-pixel positions for which there is partial overlap between the two original 32×32 speckle subregions. And the peak value of the signals drops to 0.53 of that obtained at zero strain.⁵

In the case of in-plane shear strain (ϵ_{xy}) or rotation (ω_{xy}), the effect on the correlation function is similar to that of the normal strains (Fig.4b and 4c). They also spread the correlation hill and decrease its peak

value. By comparing the correlation errors resulting from the shear strain or rotation with that from normal strains, a similar limit can be found as

$$\text{Max}(|\epsilon_{xy}|, |\omega_{xy}|) \leq 1/N_s. \quad (10)$$

For general 2-D deformation, a limit can be obtained by combining Eq.(9) with (10), i.e.,

$$\text{Max}(|\epsilon_{xx}|, |\epsilon_{yy}|, |\epsilon_{xy}|, |\omega_{xy}|) \leq 3.1\%. \quad (11)$$

In reality, many other decorrelation factors between the two speckle subregions may occur. For instance, the deformation of speckle pattern may not follow exactly the same deformation of the specimen. Other decorrelation factors may arise from temporal and spatial variations of the illumination, electrical noise of data acquisition system, and digitizing error in the image sampling and quantization, etc. However, all these decorrelation factors are quite small in small deformation range. The major decorrelation inside a subimage is a result of local displacement, strain and rotation. Indeed, using white light speckles we have achieved successful inspection of strains as large as 5 %. While using laser speckles, strains up to 2 % has been measured. The small range of laser speckle method may be explained by the quick decorrelation of the laser speckles.

5 Experiment

An aluminum specimen with a single edge crack was tested under tensile loading. White light was used in illumination. An area of about $20 \times 20 \text{ mm}^2$ around the crack tip was inspected. The load was applied in three steps. Four speckle patterns, before and after each load level, were registered by the frame grabber (Fig.5). Rigid body displacement between each pair of successive speckle patterns was as large as several millimeters. Such large displacement was first estimated by displaying each speckle pattern on a monitor and tracing a particular speckle at the central region of the pattern by a cursor, and was then eliminated by introducing an intentional rigid shift towards each other of the two speckle patterns. Although displacement distribution at each deformation level could be obtained at 240 (15×16) points in the whole field, we merely made use of two rows of data points in this work, one above and one below the crack. The separation of the two rows, which is referred to as gage length in Fig.6 ~ 9, is 2.5 mm on specimen. Taking the difference (Δv) between the v -displacement components of the the two rows, crack opening displacement (COD) was obtained. Fig.5 shows COD distributions behind crack tip (at $x > 13.13 \text{ mm}$) at load level I and II. It is seen that Δv does not vanish even before crack tip (at $x < 13.13 \text{ mm}$), although it is quite small. This nonzero distribution results from the large normal strain ϵ_{yy} ahead of the tip. In fact, normal strain ($\epsilon_{yy}|_{y=0}$) ahead of crack tip can be simply obtained from the the ratio between Δv and the gage length. Fig.7 shows the so obtained results at load level I and II, from which one can clearly observe the singular strain distributions. For the propagating crack at load level III, the COD and $\epsilon_{yy}|_{y=0}$ were similarly obtained, and are shown in Fig.8 and Fig.9, respectively. From the Δv distribution in Fig.8 one may also easily view the length of crack advance. Unusually, there is some strong fluctuation in the $\epsilon_{yy}|_{y=0}$ distribution. This is because the crack had propagated at about 20° below the horizontal axis (see Fig.5), and the new crack tip extended out of the range of the 2.5 mm gage. Consequently, a compensatory measurement was made by using the v -displacement

difference between two other data rows with a longer gage length of 5 mm (Fig.9). It is seen from Fig.9 that the new estimated $\epsilon_{yy}|_{y=0}$ curve has a more reasonable singular distribution than that from the 2.5 mm gage. It should be noted that such a distribution may under estimate the $\epsilon_{yy}|_{y=0}$ at the near-tip region since it is an average survey over the 5 mm gage. The results in Fig.(7) and (9) also show that plastic strains as large as 5 ~ 8% are effectively monitored.

6 Conclusion

Measurable area, accuracy, and ranges of displacement, strain and rotation by CASI are analyzed. An optimal sampling rate was used in speckle digitalization. Unsolvable displacement varies from 1% to 5% of equivalent speckle size on specimen, depending on degree of decorrelation between the two corresponding speckle patterns. Measurable area is proportional to the area of each individual speckle on object. At subimage size of 32×32 pixels, a theoretical upper limit of 3.1% was found for in-plane strains and rotation. Using white light speckles, displacement up to 20 ~ 40 mm, and strain and rotation up to 5% can be determined. While by laser speckles, displacement up to 2 ~ 4 mm, and strain and rotation up to 2% may be examined.

7 Acknowledgments

We thank the financial supports provided by the Army Research Office through Contract DAA20388K0033 (Scientific Officer: Dr. G.L. Anderson) and the Office of Naval Research through Contract N0001482K0566 (Scientific Officer: Dr. Yapa D.S. Rajapakse).

8 References

1. D.J. Chen and F.P. Chiang, "Computer speckle interferometry," Proc. of Int. Conf. on Hologram Interferometry and Speckle Metrology, pp.49-58 (1990).
2. D.J. Chen, F.P. Chiang, Y.S. Tan and H.S. Don, "Computer-aided speckle interferometry (CASI): II. An alternate approach using spectral amplitude and phase information," to appear in Proc. SPIE, Vol. 1554A.
3. D.J. Chen and F.P. Chiang, "Optimal sampling and range of measurement in displacement-only laser speckle correlation," to appear in Experimental Mechanics.
4. A. Rosenfeld and A.C. Kak, Digital Picture Processing, Vol.1, Academic Press, New York (1982).
5. H.H. Bailey, F.W. Blackwell, C.L. Lowery and J.A. Ratkovic, "Image correlation: Part I. Simulation and analysis," The Rand Corporation, R-2057/1-PR (1976).

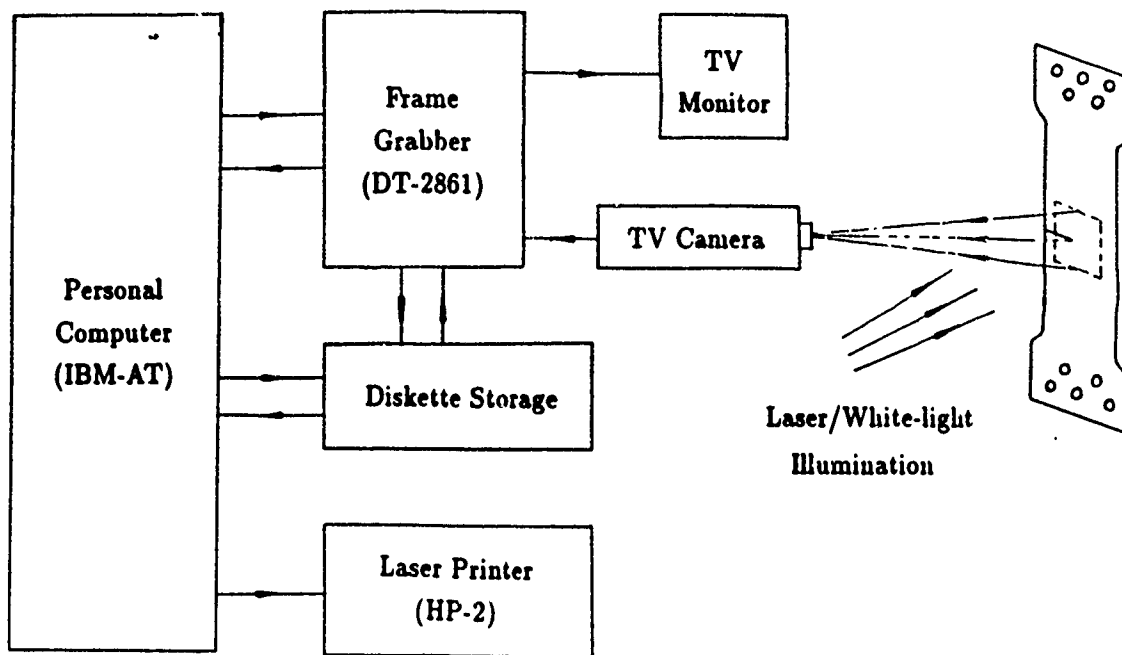


Fig.1 Schematic of data acquisition and image processing system

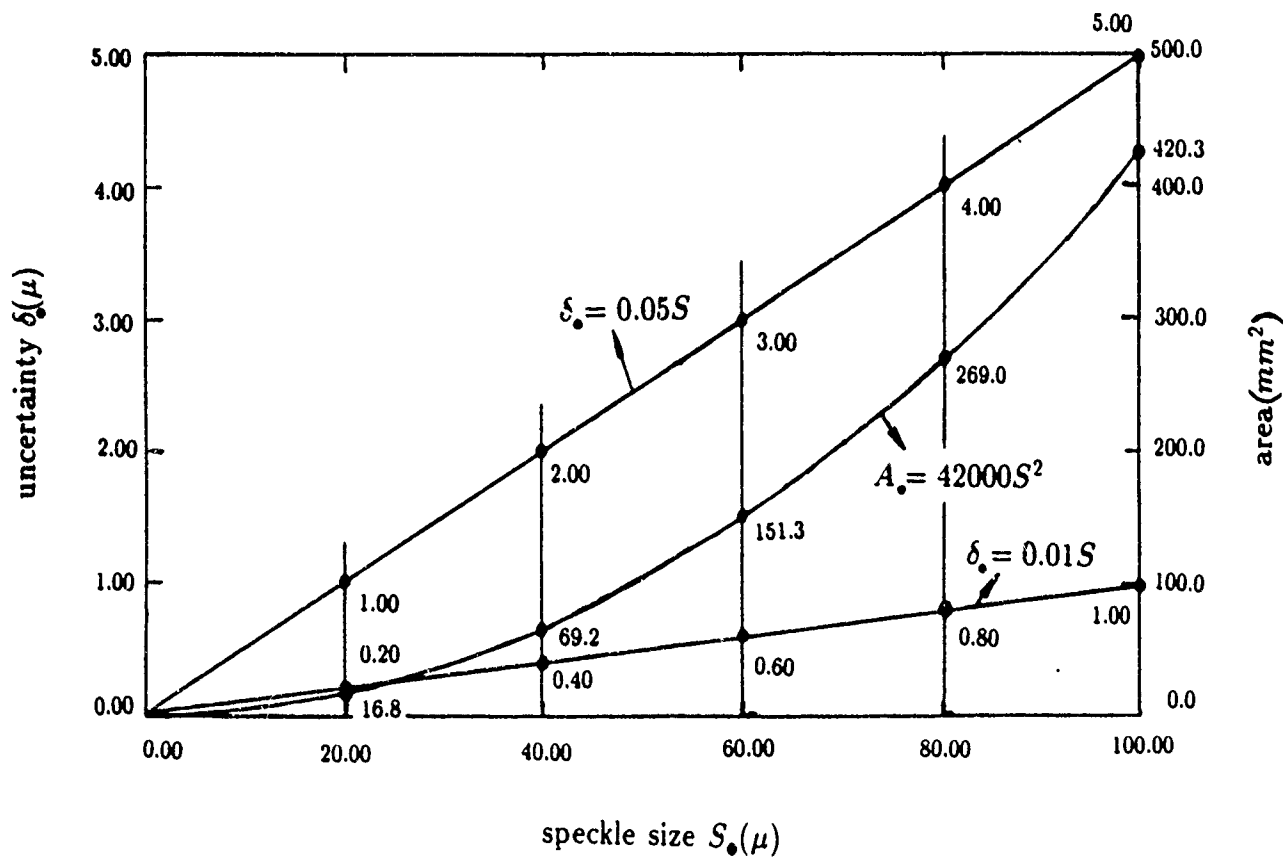


Fig.2 Measurable area and uncertainty vs speckle size

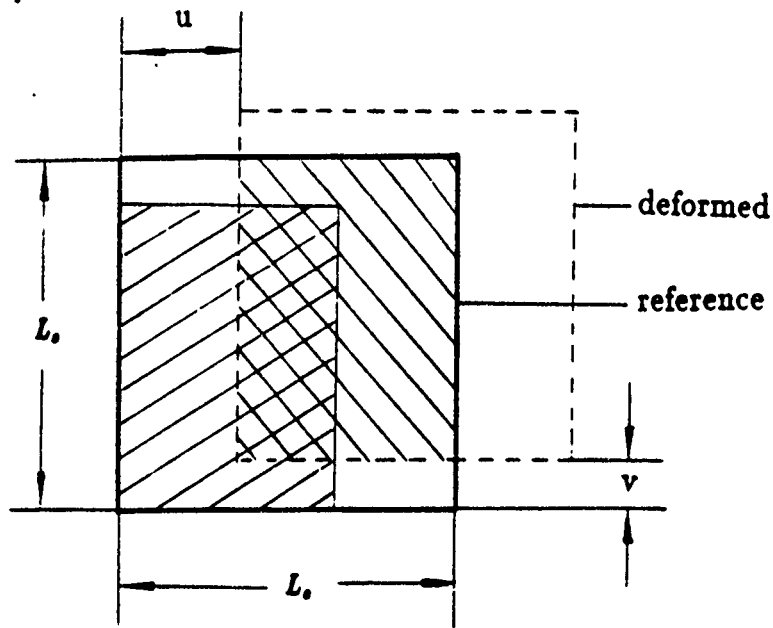




Fig.3 Decrement of correlated area in a subimage resulted from local displacement ( correlated area in reference speckle pattern,  correlated area in deformed speckle pattern)

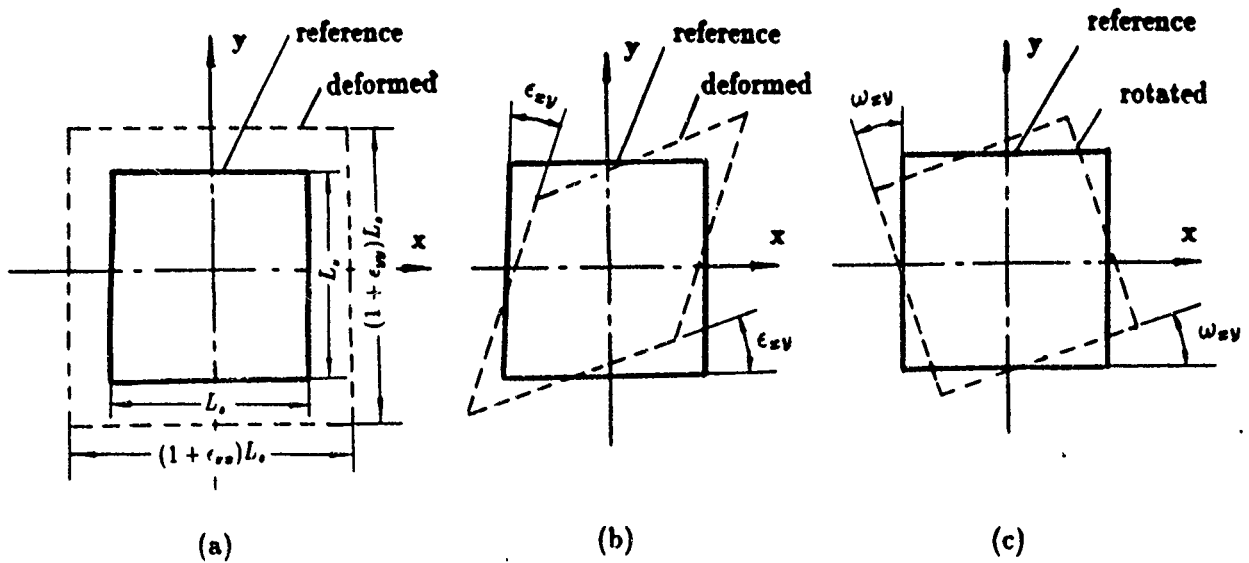
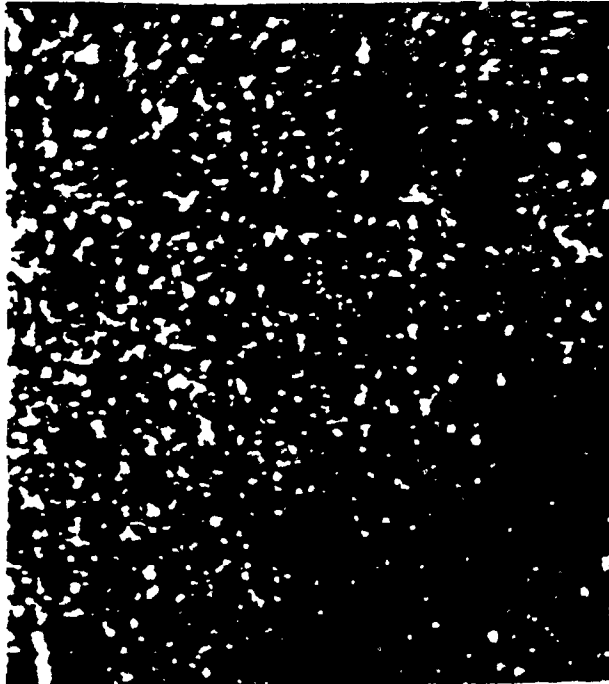
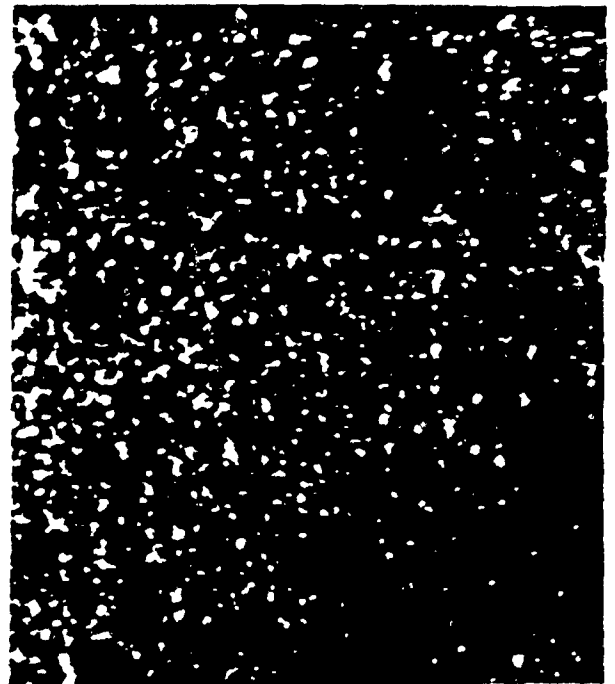


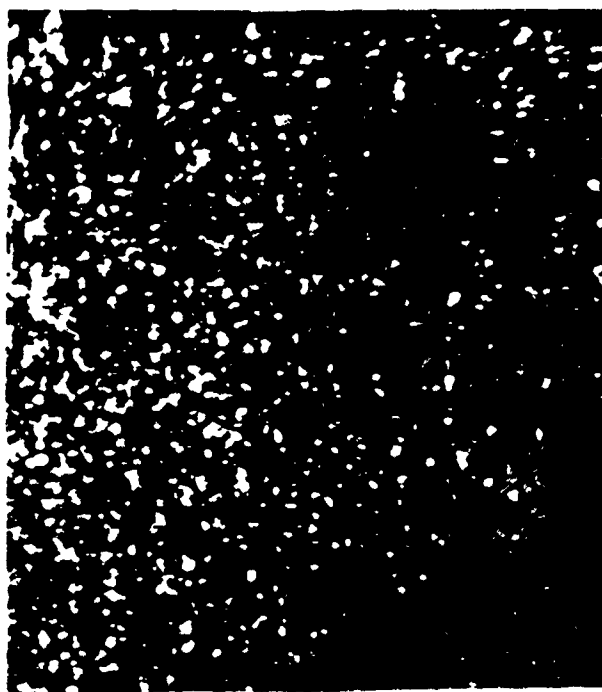
Fig. 4- In-plane strains and rotation; (a) normal strains; (b) shear strain; (c) rotation.



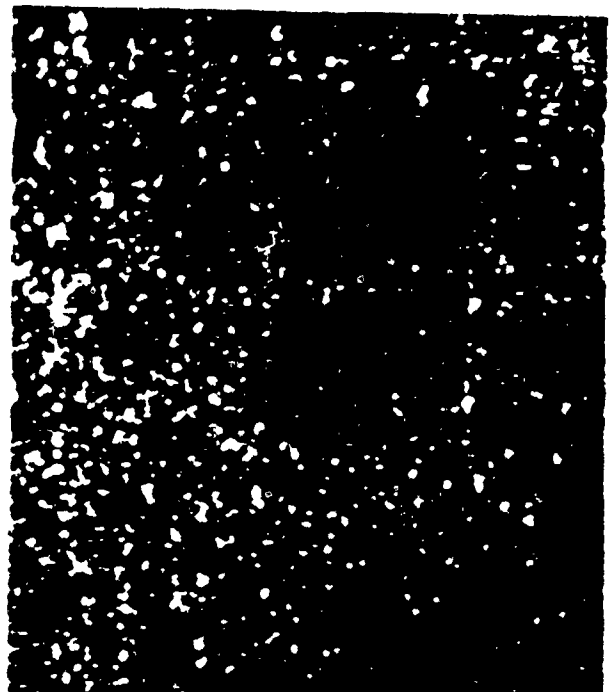
(a)



(b)



(c)



(d)

Fig.5 Digital speckle patterns around a crack tip at (a) zero load, (b) load level I, (c) load level II, and (d) load level III

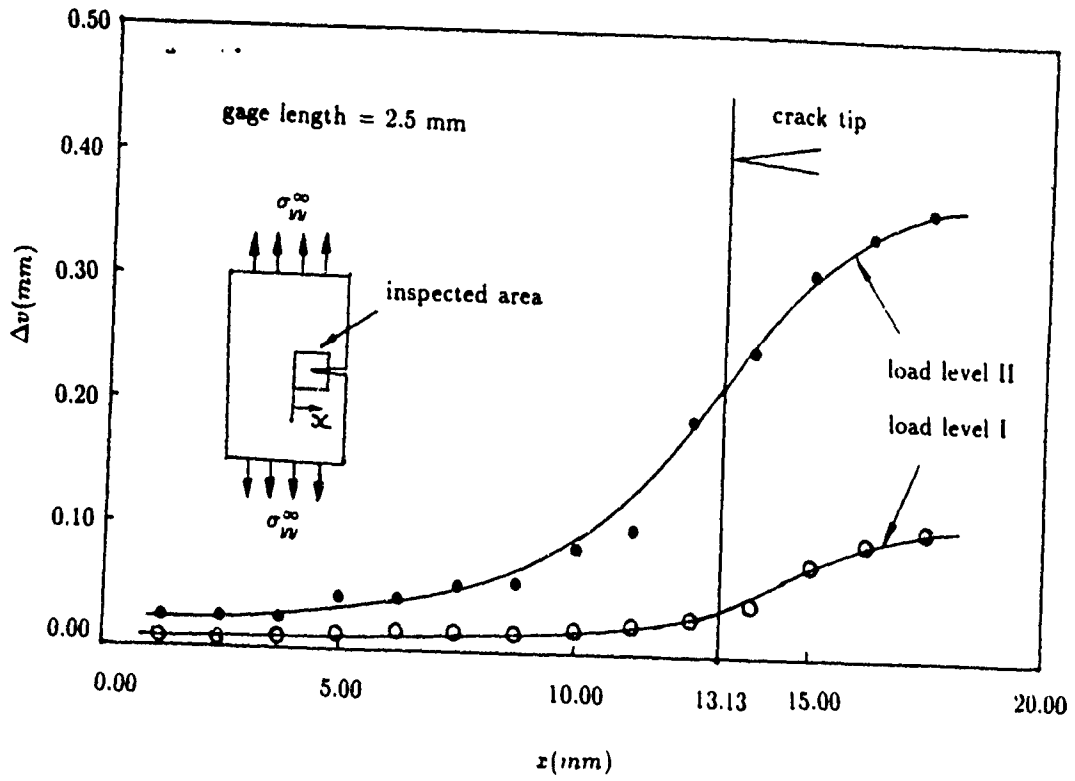


Fig.6 COD at load level I and II

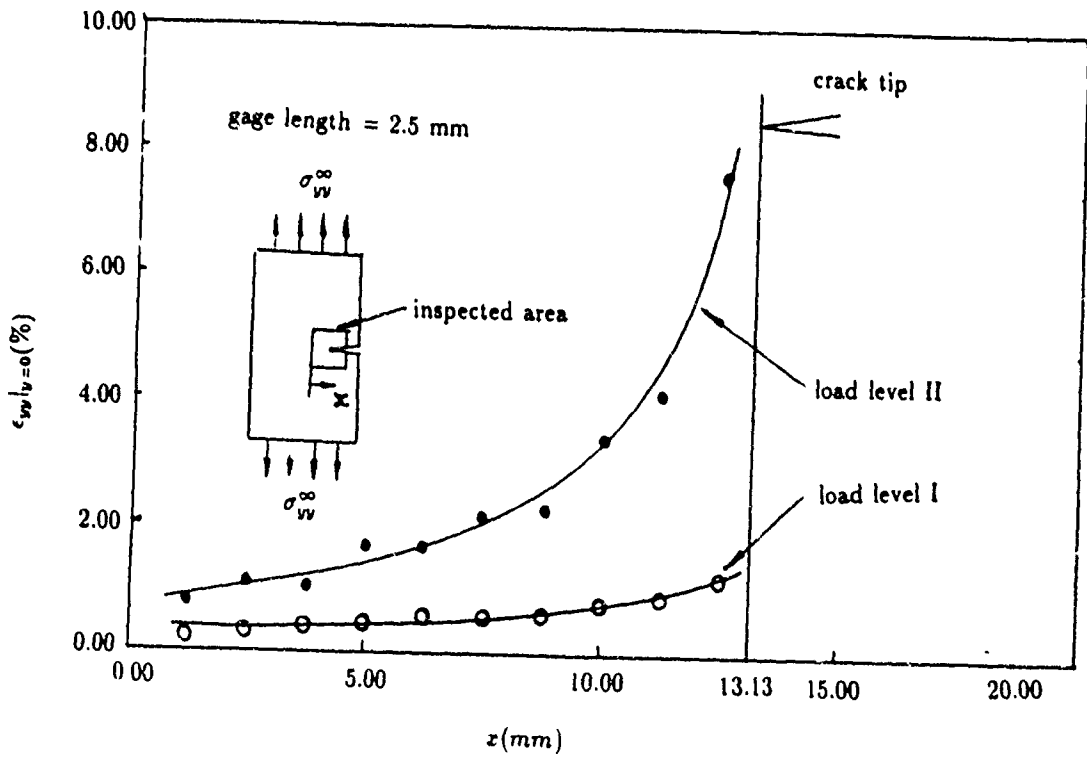


Fig.7 $\epsilon_{yy}|_{y=0}$ in front of crack tip at load level I and II

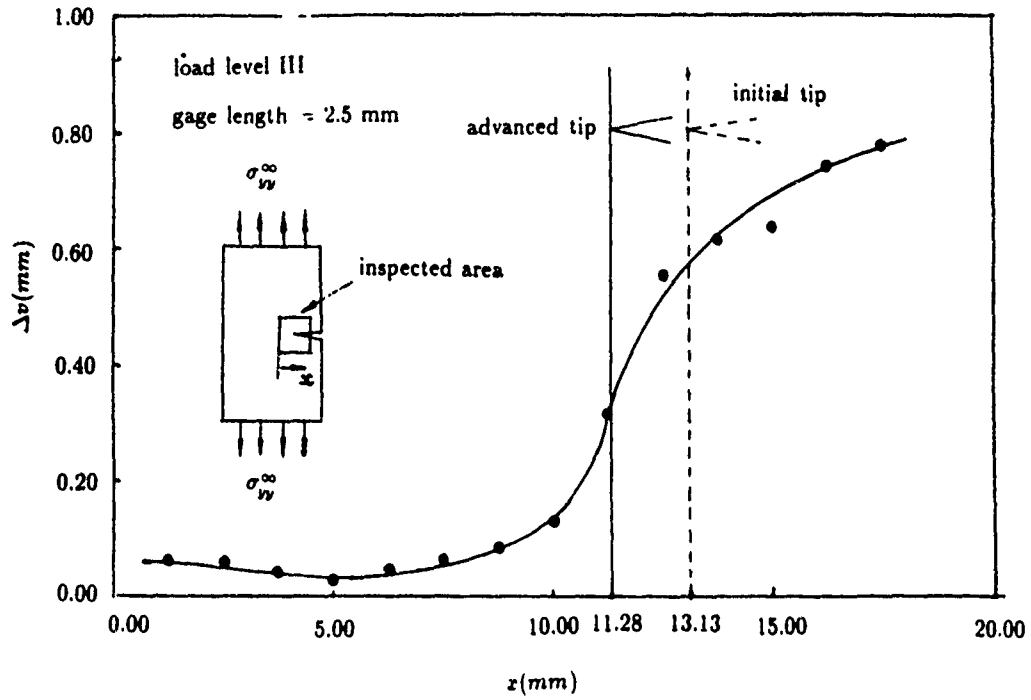


Fig. 8 COD of advanced crack at load level III

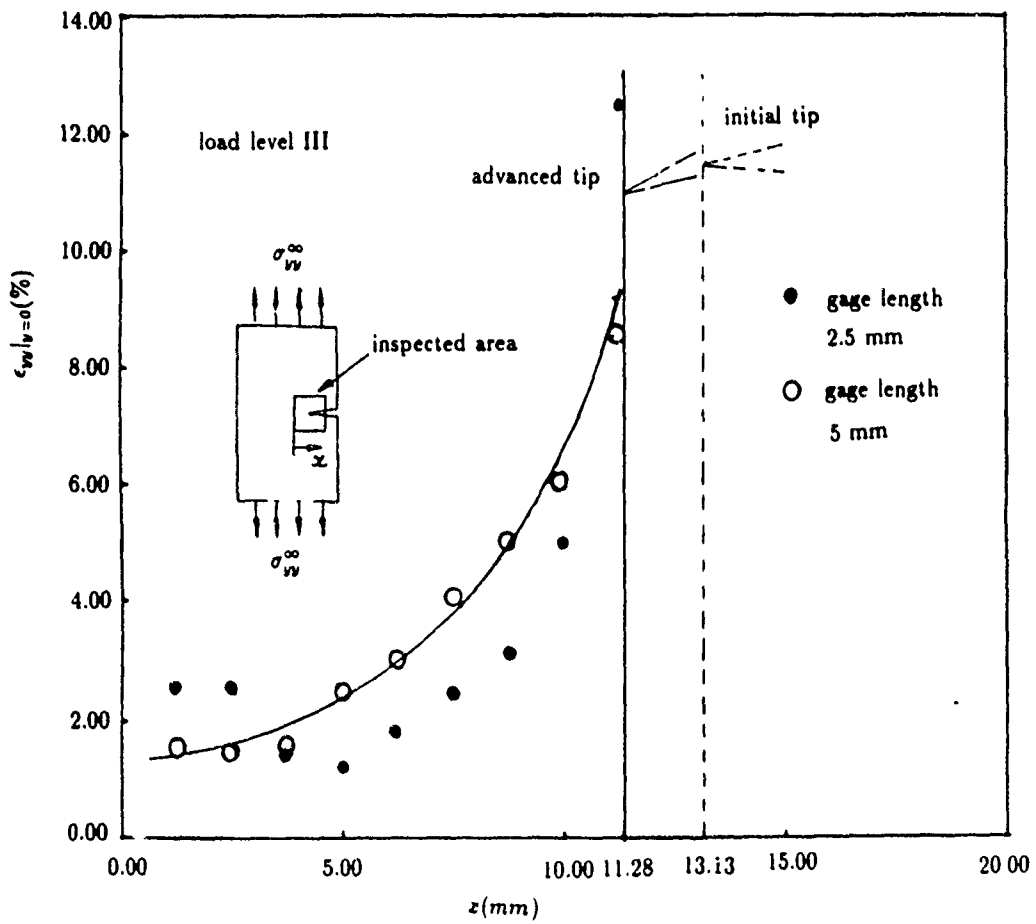


Fig. 9 $\epsilon_{yy}|_{y=0}$ in front of advanced crack tip

DETERMINATION OF PLASTIC ZONE AT CRACK TIP USING LASER SPECKLE DECORRELATION

F.P. CHIANG & Y.Z. DAI

Laboratory for Experimental Mechanics Research
State University of New York at Stony Brook
Stony Brook, N.Y. 11794-2300, USA

ABSTRACT

Speckle patterns from a metallic surface vary as a function of plastic deformation the material experiences and hence may be used to detect the plastic zone size. The speckle patterns are recorded and compared using cross correlation analysis based on an image processing system. This technique, which has the merits of being non-contact, remote sensing, and high sensitivity, is applied to measuring the size of plastic zone around a hole in an aluminum plate with finite width. A comparison with theoretical and finite element results is presented.

INTRODUCTION

The size and shape of plastic zone is of importance to the understanding of ductile fracture of metals. Most plastic zones are calculated numerically. Experimental determination usually requires quantitative values of plastic strains - a time consuming process. A non-contact, remote sensing technique for measuring plastic zone size is presented based on plasticity induced surface roughness and speckle pattern decorrelation.

The surface roughness of a metallic material changes with increase in surface strain [1]. As a result, the speckle pattern from the material surface changes correspondingly. This phenomenon suggests methods to relate plastic strain quantitatively with the variation in speckle pattern. Some techniques based on this principle have been developed [2,3]. Yet, the sensitivity and reliability of these techniques need improvement.

The present non-contact, remote method for determining plastic zone is based on an image processing system. A cross correlation technique is employed to analyze the digitized speckle patterns. In particular the technique is applied to the determination of the plastic zone around a central hole in a plate with finite width. The experimental

result is in good agreement with that by finite element analysis.

SPECKLE PATTERN CORRELATION

The proposed technique makes use of a set-up as schematically shown in Fig.1. The point of interest on a test object was illuminated with a 20mm He-Ne laser. To increase the lateral resolution of the technique, the diameter of the laser beam was reduced through a system consisting of a spatial filter, an aperture, a collimating lens and a converging lens. A piece of ground glass was placed at a distance of 0.8m from the test object. A digital camera was connected to a monitor and supported by a computer which provided correlation analysis capabilities.

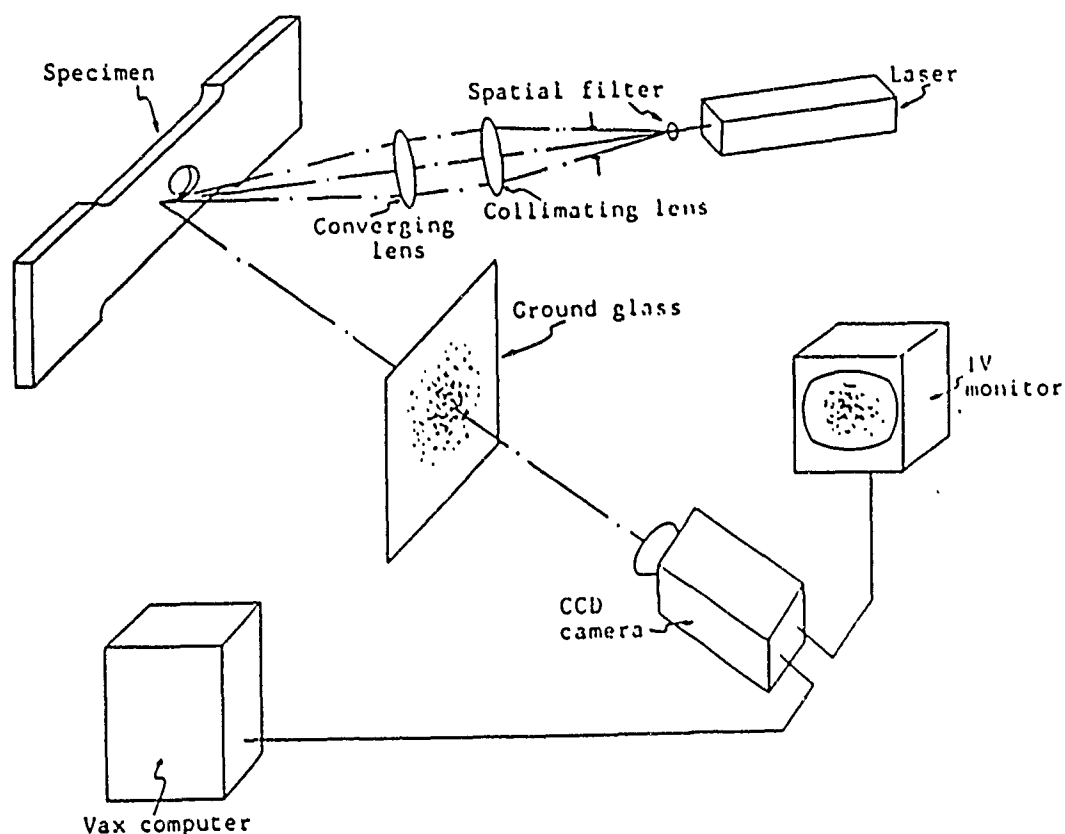


Figure 1 Schematic illustration of the experimental set-up

Cross correlation [3] is used to characterize the digitized speckle patterns such as those shown in Fig.2. For two speckle patterns described by $g(i, j)$ and $f(i, j)$, the normalized zero shift cross correlation $C_c(g, f)$, which is a measure of the resemblance between the two functions, is defined as

$$C_c(g, f) = \frac{\sum_{i=1}^M \sum_{j=1}^N g(i, j) \times f(i, j)}{\sqrt{\sum_{i=1}^M \sum_{j=1}^N g^2(i, j) \times \sum_{i=1}^M \sum_{j=1}^N f^2(i, j)}} \quad (1)$$

where M, N are the dimensions of $g(i, j)$ and $f(i, j)$, respectively. The cross correlation coefficient of two speckle patterns, one from elastic zone and the other from plastic zone, is smaller than that all from elastic zone and hence may be used to detect plastic zone.

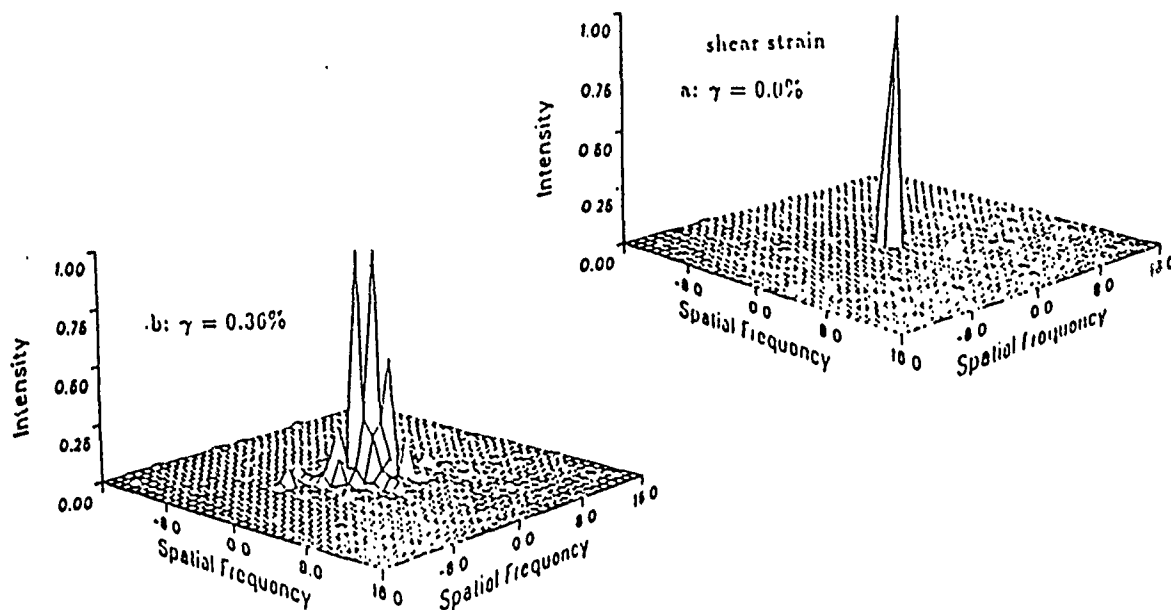


Figure 2 Speckle patterns from elastic and plastic zones

EXPERIMENTAL RESULT

An aluminum plate of 3.2mm thick with a central hole was used in the experiment. The geometry of the specimen is shown in Fig.1 where the width of the specimen is 76mm and the radius of the hole is 6.4mm. One side of the specimen surface was polished until a RMS roughness value of approximately $0.05\mu\text{m}$ was obtained. The specimen was loaded axially at $\sigma = 0.765\sigma_0$, where σ is the remote normal stress along the loading direction, and σ_0 is the yield strength of the material. The specimen was then mounted on a stage which is capable of translating the specimen in two directions in 0.006mm increments. Speckle patterns at different points on the specimen surface was digitized and processed to yield cross correlation coefficient.

Figure 3 shows the plastic zone determined using the method described above. The solid line represents a computed effective strain of 0.01% by a finite element analysis. As can be seen both the experimental and computed results agree reasonably well. The theoretical solution for plastic zone around a circular hole in an infinite flat plate is also included (broken line). As can be seen that it underestimated the size of plastic zone.

It has been demonstrated that speckle pattern cross correlation technique can be applied to the detection of the incipience of plasticity with a sensitivity of 0.01%. With further automation, the technique has the potential of being developed into a practical tool which may be applied to determining plastic zone at crack tip.

The authors wish to thank the Army Research Office, Engineering Science Division for providing the financial support through contract No. DAAL0388K0033.

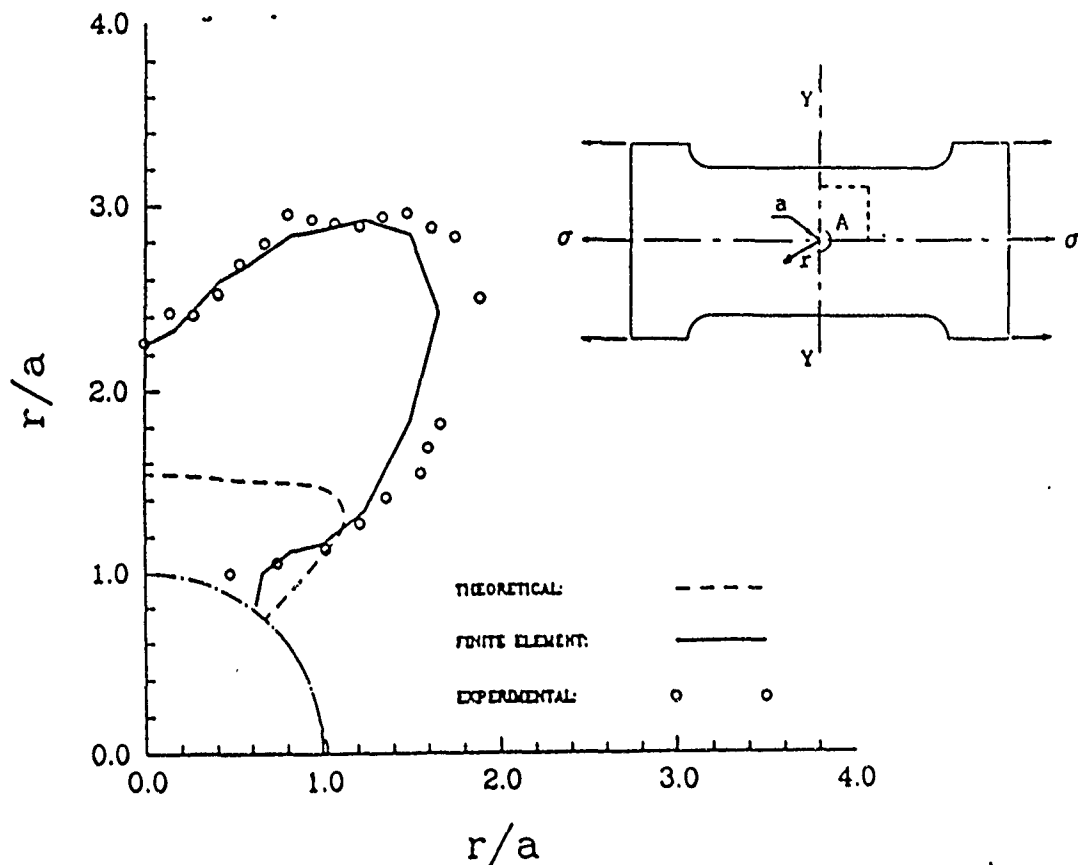


Figure 3 Plastic zone around a hole in a finite width plate ($\sigma = 0.765\sigma_0$)

REFERENCES

1. P.F. Thomson, B.V. Shafer, The roughening of free surface during plastic working, Int. J. of Mach. Tool Des. Res., 22,(4), pp.261-264, 1982
2. M. Miyagawa, A. Azushima, Measurement of plastic deformation by means of a laser beam, Proceedings of the 9th World Conference on Nondestructive Testing, Session 4a-4, pp.1-7, Melbourne, Australia, Nov.,1979
3. C. Lee, Y.J. Chao, M.A. Sutton, et al., Determination of plastic strain at notches by image-processing methods, Experimental Mechanics, 29,(2), pp.214-220, 1989
4. A. Rosenfeld, C.A. Kak, Digital Picture Processing, Academic Press, pp.19, 1982

Damage Monitoring of Metal Materials by Laser Speckle Assisted by Image Processing Technique*

(Basic Experiment)

Akira Kato**, Yu-Zhong Dai*** and Fu-Pen Chiang***

This paper presents a method to detect damage of metals with no contact using a laser speckle technique. This method is based on observations of the change of the laser speckle pattern depending on the surface roughness and texture change of the material caused by slip bands due to plastic deformation. The laser speckle pattern was analyzed automatically and quantitatively using an image processing system in this experiment. A new parameter, c_s , was derived to statistically express the features of the speckle pattern. It was found that this parameter can express the features of the speckle pattern accurately and has excellent reproducibility. Also, it allows an accurate measurement of plastic strain. The experiment was carried out for cases of both static tensile and fatigue loading using aluminum alloy as the material.

Key Words: Nondestructive Inspection, Laser Speckle Method, Image Processing, Plastic Deformation, Fatigue

1. Introduction

Both static plastic deformation and fatigue damage of metal materials are accompanied with plastic strain, producing slip bands on the material surface. Therefore, the surface roughness and surface profile change with the progress of plastic deformation or fatigue damage. The laser speckle pattern changes with the change of surface roughness and profile. If the relationship between the laser speckle pattern and the amount of plastic deformation or fatigue damage is known in advance, we can estimate the damage of metal materials based on the speckle pattern.

Measurement of plastic deformation by the laser or white-light speckle technique has been reported by several researchers⁽¹⁾⁻⁽⁴⁾. Those papers investigated the method to estimate plastic strain based on the speckle pattern change depending on the surface texture change induced by slip bands. However, the data

obtained in those experiments scatter slightly, and the results are not sufficiently reliable. It is essential in this kind of experiment to analyze the laser speckle pattern quantitatively and to find a parameter that can express the features of speckle pattern change accurately. In this investigation, an image processing system was employed to analyze the laser speckle pattern automatically and quantitatively. As the distribution of the light intensity of the laser speckle varies drastically, we considered a new statistical parameter to express the features of the laser speckle pattern. The reproducibility of this parameter was examined and the accuracy of the measurement was investigated.

In the earlier stage of fatigue, slip bands generate on the metal surface, microscopic shear strain is stored in the slip bands and the slip bands become dense. The surface texture changes and surface roughness becomes marked with the progress of fatigue damage^{(5),(6)}. If we use the laser speckle technique, there is a possibility of detecting the fatigue damage of metals in the early stage of fatigue before initiation of cracks. In this study, we observed the change of the speckle pattern during fatigue cycles under operation

* Received 2nd October, 1990

** Chubu University, Kasugai, Aichi, Japan

*** State University of New York at Stony Brook, Stony Brook, New York, U.S.A.

estimated by expression (76). The calculated results shown in the figures quantitatively inform us of the dependency of each parameter. Expression (76) is of closed form; thus, the calculation is very easy. Consequently, the simple expression may be practical to develop the characterization process of flaws.

It should be mentioned here that the characteristic factors affecting the backscattering wave information should be carefully checked and chosen. Inversely determined characteristic features of flaws from measured backscatter intensity data should be verified according to comparison with the real features before a decision process is offered in service. It is necessary to determine if other factors exist, and how these factors affect the backscatter intensity.

References

- (1) Budreck, D. E. and Achenbach, J. D., 3-D Ultrasonic Scattering from Planar Cracks by the Boundary Element Method, In Review of Progress in Quantitative Nondestructive Evaluation, Vol. 7A (eds. Thompson, D. O. and Chimenti, D. E.), Plenum Press, New York (1988), p. 103.
- (2) Nagy, P. B. and Adler, L., The Role of Coherent Backscattering in Quantitative NDE, In Review of progress in Quantitative Nondestructive Evaluation, Vol. 7A (eds. Thompson, D. O. and Chimenti, D. E.), Plenum Press, New York (1988), p. 113.
- (3) Ogilvy, J. A., Wave Scattering by Rough Surfaces, In Review of Progress in Quantitative Nondestructive Evaluation, Vol. 7A (eds. Thompson, D. O. and Chimenti, D. E.), Plenum Press, New York (1988), p. 123.
- (4) Ng, K. W. and Ngoc, T. D. K., Application of Ultrasonic Scattering, In Review of Progress in Quantitative Nondestructive Evaluation, Vol. 7A (eds. Thompson, D. O. and Chimenti, D. E.), Plenum Press, New York (1988), p. 131.
- (5) Bar-Cohen, Y., Ultrasonic NDE of Composites - A Review, (eds. Achenbach, J. D. and Rajapaske, Y.), Martinus Nijhoff Publ (1987), p. 187.
- (6) Auld, B. A., General Electromechanical Reciprocity Relations Applied to the Calculation of Elastic Wave Scattering Coefficients, Wave Motion, Vol. 1 (1979), p. 3.
- (7) Achenbach, J. D., Flaw Characterization by Ultrasonic Scattering Methods, (eds. Achenbach, J. D. and Rajapaske, Y.), Martinus Nijhoff Publ (1987), p. 67.
- (8) Ohyoshi, T. and Achenbach, J. D., Effect of Bottom-Surface Reflections on Backscatter from Porosity in a Composite Layer, In Review of Progress in Quantitative Nondestructive Evaluation, Vol. 7B (eds. Thompson, D. O. and Chimenti, D. E.), Plenum Press, New York (1988), p. 1045.
- (9) Angel, Y. C. and Achenbach, J. D., Reflection and Transmission of Elastic Waves by a Periodic Array of Cracks, Trans. ASME, Jour. Applied Mechanics, Vol. 52, (1985), p. 33.
- (10) Sneddon, I. N., The Use of Integral Transforms, (1972), Chap. 5, McGraw-Hill, New York, p. 298.
- (11) Achenbach, J. D., Wave Propagation in Elastic Solids, (1973), p. 34, North-Holland, Amsterdam.
- (12) Kassir, M. K. and Sih, G. C., Three Dimensional Crack Problems, Noordhoff Publ, Netherlands (1975), p. 2.

not this
page

of a testing machine and the possibility of fatigue monitoring by the laser speckle technique was investigated.

2. Experimental Procedure

2.1 Experimental system

The material used in this experiment is 6061 aluminum alloy. The initial roughness and directional feature of the specimen surface has a large influence on the speckle pattern. If the surface texture is directional, the speckle pattern is directional and not axisymmetrical. Also, if the initial surface roughness is large, speckle pattern change is small in the range of small plastic strain. Therefore, if it is required to detect a small change of speckle pattern, the specimen surface should initially be smooth and nondirectional. In this experiment, the specimens were polished finally with aluminum dioxide powder in random directions so that the surface texture became macroscopically isotropic and uniform. The initial surface roughness was about $0.1 \mu\text{m Ra}$ for every specimen.

Figure 1 shows the layout of the experimental system. He-Ne laser was illuminated on the specimen surface and the laser speckle pattern was formed on the ground glass placed in front of the specimen. The diameter of the laser beam was about 1 mm. The image of the speckle pattern was input into the image processing system (Hamamatsu C1000) through a TV camera and then A/D converted. Resolution of the image memory was 256×256 and the gray scale of one pixel was 256. The brightness of the laser speckle was

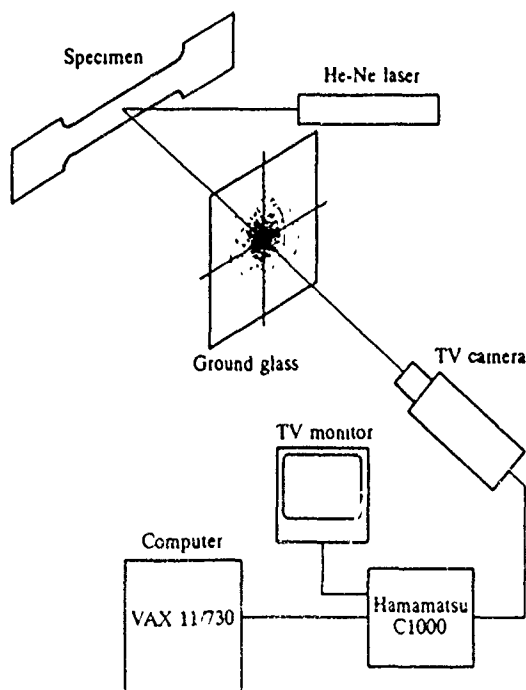


Fig. 1 Configuration of experimental system

expressed by a numerical value from 0 to 255 in the image memory. The digital image data were transferred to the main memory of a minicomputer (VAX 11/730) and then stored in its magnetic tape. Software to analyze the speckle image was developed on the minicomputer.

2.2 Data analysis

Figure 2 shows examples of the gray level distribution of the speckle patterns obtained from the image processing system for different magnitudes of plastic strain caused by uniform tension. These figures show that the speckle pattern changes clearly with the magnitude of plastic strain. The peak gray level at the center of the speckle image decreases and the gray level distribution spreads out with the increase of plastic strain.

Figure 3 shows the relationship between the ratio of M_{\min}/M_{\max} and plastic strain. Here, M_{\min} and M_{\max} are the minimum and maximum values, respectively, among second-order moments of inertia of the gray level distribution around the four different axes shown in the chart of Fig. 3. If the ratio M_{\min}/M_{\max} is close to 1, it means that the second-order moments of inertia around the four axes take almost the same value; that is, the gray level distribution of the speckle pattern is round and axisymmetrical. From the figure, this ratio is smaller than 1 when plastic strain is small, but it approaches 1 with the increase of plastic strain. This means that the gray level distribution is slightly directional initially, but it becomes nondirectional and axisymmetrical around the centroid of the gray level distribution when plastic deformation becomes large. Hence we considered the following parameter to express the spread of the gray level distribution around the centroid of the distribution.

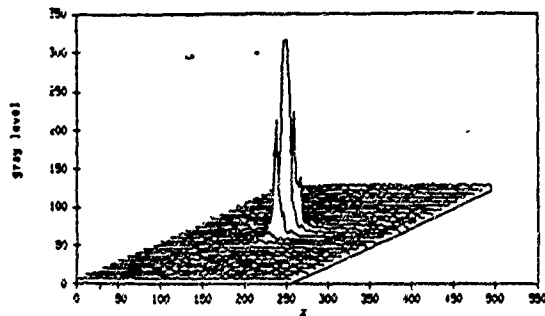
First, the image of the speckle pattern was smoothed to reduce the noise. We used a moving averaging method of 3×3 pixels for smoothing. Then the centroid, $O(x_0, y_0)$, of the gray level distribution was obtained (Fig. 4(a)), and the average gray level, $g(r)$, of pixels on a circle c with the center O and the radius of r was obtained (Fig. 4(b)). The average gray level $g(r)$ is expressed by the following expression,

$$g(r) = \frac{\sum_c g(x, y)}{n_c} \quad (1)$$

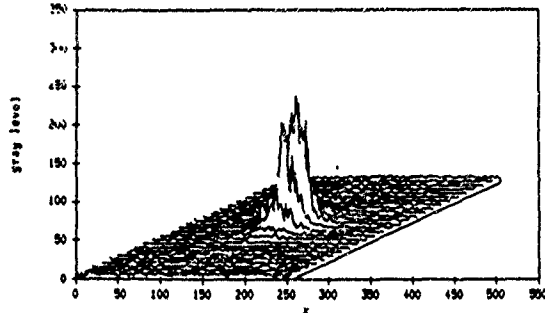
where $g(x, y)$ is the gray level at a point (x, y) on circle c of the speckle image. Then $\sum_c g(x, y)$ is the summation of the gray level of pixels on circle c and n_c is the number of pixels on the same circle. Also,

$$r = \sqrt{(x - x_0)^2 + (y - y_0)^2}$$

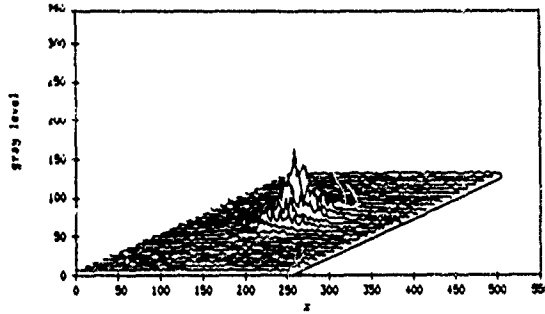
Consequently, the function $g(r)$ expresses the gray level distribution in the radial direction around the



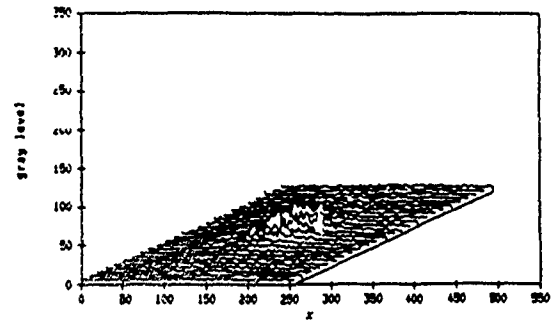
(a) Initial ($\epsilon_p=0\%$)



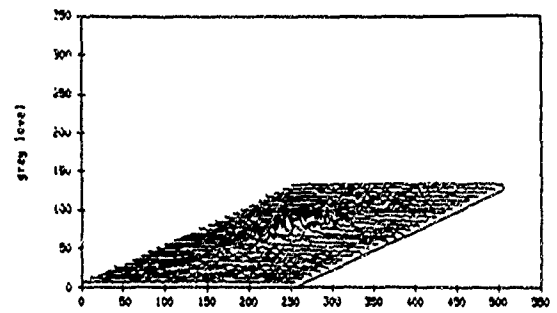
(b) $\epsilon_p=1.55\%$



(c) $\epsilon_p=3.16\%$



(d) $\epsilon_p=5.20\%$



(e) $\epsilon_p=6.60\%$

Fig. 2 Speckle pattern for different plastic strains

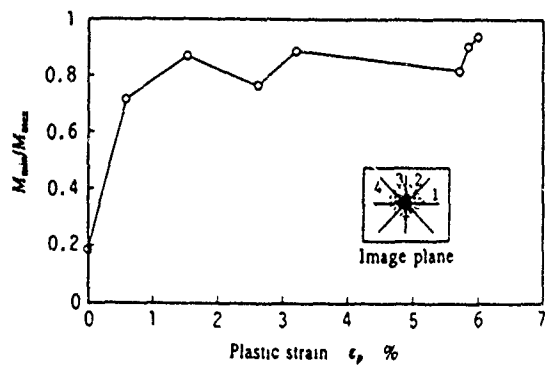
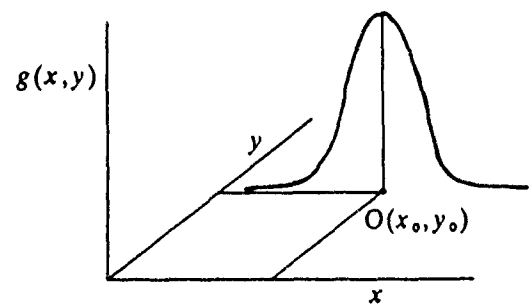
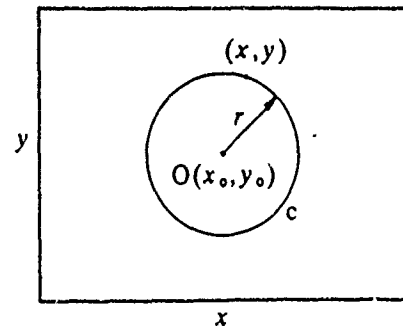


Fig. 3 Relationship between M_{min}/M_{max} and plastic strain



(a) Gray level distribution



(b) Image plane

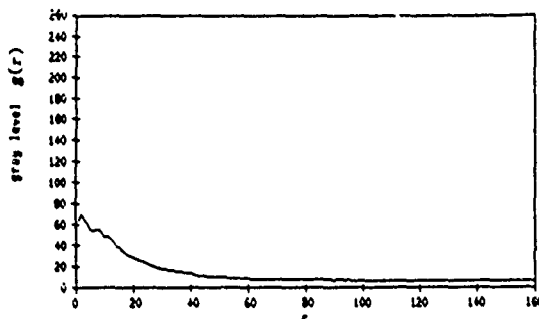
Fig. 4 Procedure for obtaining $g(r)$

centroid O (Fig.5 (a)).

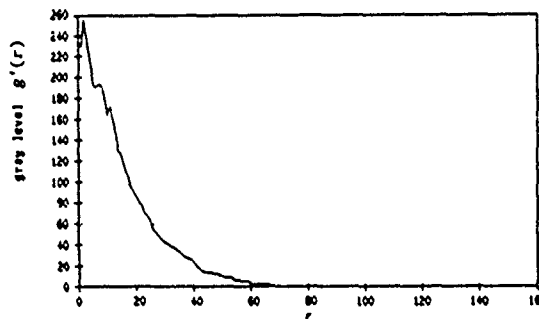
Then $g(r)$ was normalized to reduce the error induced by the difference of speckle brightness. We used the following linear mapping function for normalization:

$$g'(r) = \frac{255(g(r) - g_{min})}{g_{max} - g_{min}} \quad (2)$$

This function makes the maximum gray level in $g(r)$, g_{max} , to be 255 and the minimum value, g_{min} , to be 0. The normalized gray level distribution is shown in Fig.5 (b). Figure 6 shows the difference in the normalized gray level distribution depending on the plastic strain. The gray level distribution, $g'(r)$, is found to spread with the increase of plastic strain. Hence, we considered the following parameter that can express the feature of the gray level distribution:



(a) $g(r)$



(b) $g'(r)$ (Normalized)

Fig. 5 Normalization of $g(r)$

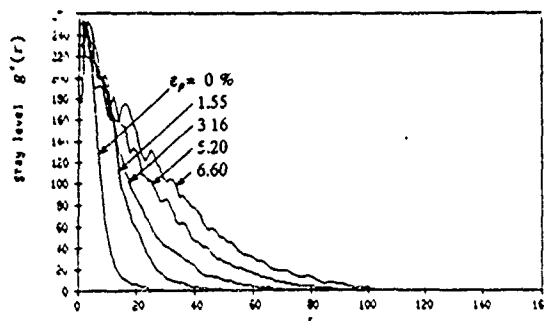


Fig 6 Difference in gray level distribution $g'(r)$

$$c_A = \frac{\sum_{r=0}^{N_r} g'(r)r^2}{\sum_{r=0}^{N_r} g'(r)} \quad (3)$$

where N_r is the limit of r , which was taken as 110 in this analysis.

Next, we examined the influence of brightness of the laser speckle on the c_A parameter. Figure 7 shows the relationship between c_A and g_{max} when the speckle pattern was obtained at the same point of a specimen with different brightness of the laser speckle. The speckle brightness was changed by two methods. One is that the aperture of the lens of the TV camera was changed under a constant intensity of the laser light and the other is that the intensity of the laser light was changed with an ND filter under a constant aperture of the lens. This figure shows that the c_A parameter is almost constant in the wide range of g_{max} (from 20 to 160) and takes almost the same value in both cases. This means that the c_A parameter is not significantly influenced by the brightness of the laser speckle, and both the change of aperture and the change of light intensity have almost the same effect on the c_A parameter.

3. Measurement of Plastic Strain

First we applied this method to the measurement of plastic strain. Here the aluminum specimens were loaded by uniaxial tension. Plastic strain was measured by the moiré method. Figure 8 shows the relationship between the surface roughness (R_a) and the magnitude of plastic strain. It is found that the surface roughness increases with the increase of plastic strain. Figure 9 shows the relationship between the c_A parameter of the laser speckle pattern and the magnitude of plastic strain ϵ_p . The experiment was conducted under four different testing conditions of speckle brightness. The first one was that the aperture of the TV camera was constant but the light intensity of the laser light was adjusted with an ND filter so that the

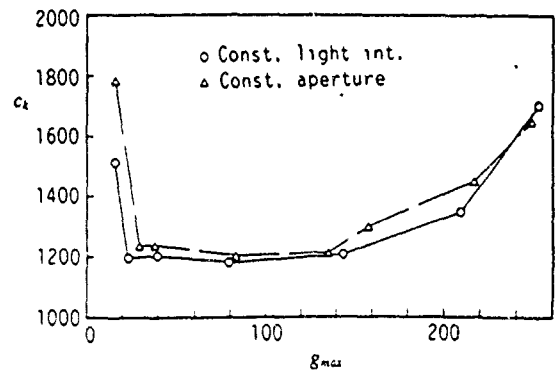


Fig 7 Relationship between c_A and g_{max}

average gray level of the speckle image became approximately constant. The others were that the aperture of the lens of the TV camera was fixed at three different values, i. e., $f/32$, $f/16$ and $f/8$, under a constant light intensity. It is found that the values of c_A obtained under different conditions of speckle brightness coincide very well and the c_A parameter is not influenced by speckle brightness in all cases of plastic strain. The parameter c_A increases with ϵ_p and has a tendency similar to the surface roughness shown in Fig. 8.

Figure 10 shows the relationship between c_A and ϵ_p obtained from two specimens under the same experimental condition. In the figure, the data of specimen A are the averages of values shown in Fig. 9. The values of c_A obtained from the two different specimens coincide very well. This shows that the c_A parameter has a very good reproducibility and that there is a fixed relationship between c_A and plastic strain ϵ_p . Figure 10 shows that there is a linear relationship between $\log(c_A)$ and $\log(\epsilon_p)$. This relation can be expressed by the following equation :

$$\log(c_A) = a + b \log(\epsilon_p) \quad (4)$$

Alternatively, it can be expressed by

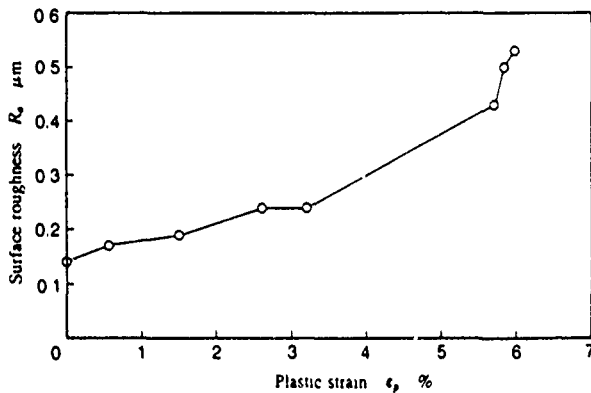


Fig 8 Relationship between surface roughness and ϵ_p

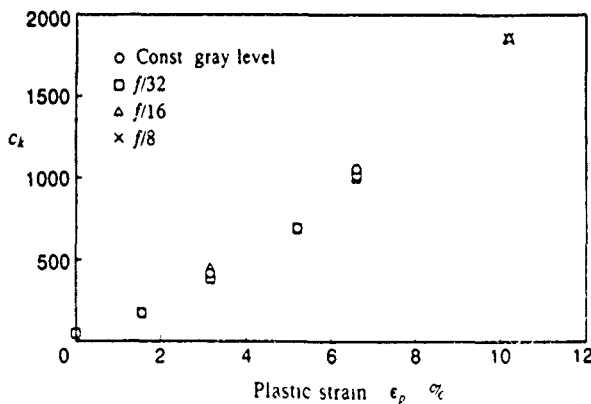


Fig 9 Relationship between c_A and ϵ_p

$$\epsilon_p = c(c_A)^d \quad (5)$$

Here, $\log(c) = -a/b$ and $d = 1/b$. The values of c and d were obtained by the least squares method : $c = 3.99 \times 10^{-4}$ and $d = 0.74$. Therefore, we can estimate the magnitude of plastic strain from the value of the c_A parameter using this equation. It was found that this parameter c_A is not influenced significantly by the brightness of the laser speckle and that it has a very good reproducibility. It can be concluded that this parameter will allow us to make an accurate measurement of plastic strain.

4. Fatigue Test

In the fatigue test, we used the notched specimens shown in Fig. 11. The laser light was illuminated on the edge at the notch root of the specimen, as shown in the figure. The testing machine used was a servo-hydraulic fatigue testing machine (Instron). The specimens were subjected to uniaxial tensile fatigue load. The specimen was set in the testing machine and the laser light was illuminated on the specimen surface. The experimental layout is almost the same as in Fig. 1 except that we used a high-speed camera (SP2000, Kodak) to obtain the speckle image and a personal computer (IBM PS/2) to analyze the image data. The image of the speckle pattern was taken using the high-speed camera during fatigue loading without stopping the testing machine or unloading the specimen. This high-speed camera can obtain images up to 2 000 frames/sec. This equipment performs A/D conversion

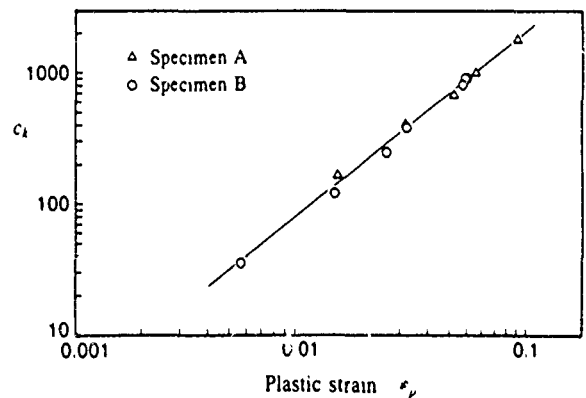


Fig 10 Relationship of c_A - ϵ_p for two specimens

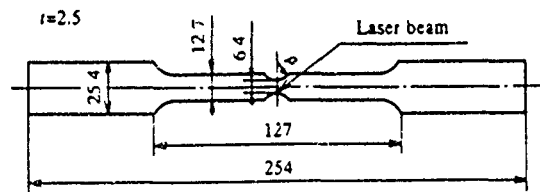


Fig 11 Test specimen for fatigue test

of the image signal and stores the digital image in its memory as well as on the magnetic tape. The resolution of this image memory is 238×192 and the gray scale is 64. As the frequency of the cyclic loading was set at 6 Hz, which was not so high, we used the high-speed camera with the speed of 200 frames/sec. During the fatigue test, an image of the laser speckle pattern was taken at a certain interval of fatigue cycles. The image data of the speckle pattern were transferred to the personal computer, and the same analysis as that in the static tensile test was made.

We observed the speckle pattern change with the progress of fatigue damage under several kinds of stress amplitude. We used specimens with a notch radius of 6.4 mm. As the notch radius is rather large, the stress concentration factor due to the notch can be considered to be 1. The conditions of fatigue stress at the notch root were as follows: the mean stress was fixed to $\sigma_m = 217$ MPa and the stress amplitude was changed as $\sigma_a = 155, 167, 186, 205$ and 217 MPa. Figure 12 shows the relationship between stress amplitude and fatigue life, N_f . Figure 13 shows the relationship between the c_k parameter of the laser speckle pattern and N/N_f . Here, N is the number of stress cycles. It is normalized by fatigue life, N_f , for the corresponding stress amplitude shown in Fig. 12. The value of c_k is very small compared to that in the case of static plastic deformation, especially when σ_a is small. Also, the change of the c_k value is very small when σ_a is smaller than 186 MPa. However, the c_k value increases monotonously with N/N_f . The c_k value becomes large with the increase of σ_a , and the change of the c_k parameter with N/N_f becomes perceptible. It is found that the c_k value changes markedly during the earlier cycles of fatigue and is saturated for larger N .

As the plastic strain induced by fatigue loading is very small compared to that of the static tensile test, the change of the speckle pattern is small and the c_k value does not change so markedly in the fatigue test.

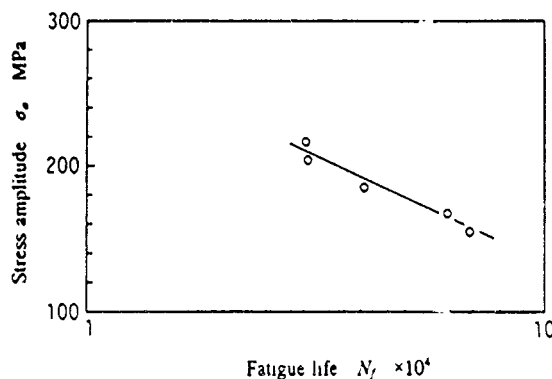


Fig. 12 Relationship between stress amplitude and fatigue life

However, as shown in Fig. 13, the c_k value changes appreciably with the increase of fatigue cycle when stress amplitude is rather large and it may be concluded that damage detection is possible in the early stage of fatigue by observing the change of the c_k parameter.

5. Conclusions

The laser speckle technique was applied to the measurement of plastic strain and the detection of fatigue damage, assisted by an image processing system in this study. The image processing system was found to be very effective in analyzing the speckle pattern quantitatively. The c_k parameter which we derived in this study expresses the features of the laser speckle pattern numerically and has a very good reproducibility. The relation between c_k and plastic strain is expressed by a simple equation in the range of plastic strain ($\epsilon_p < 10\%$) where we made experiments. Then, plastic strain can be estimated easily using this equation. It is concluded that the magnitude of plastic strain can be estimated accurately by the laser speckle technique.

The speckle pattern did not change so dramatically with fatigue loading, and the change of the c_k parameter was very small. Also, the fluctuation of the c_k parameter was not as distinctive as that of static tensile test. Therefore, detection of fatigue damage will not be as accurate. However, the c_k parameter changed appreciably when the stress amplitude was rather large. It can be concluded that there is a possibility of detecting fatigue damage by this method.

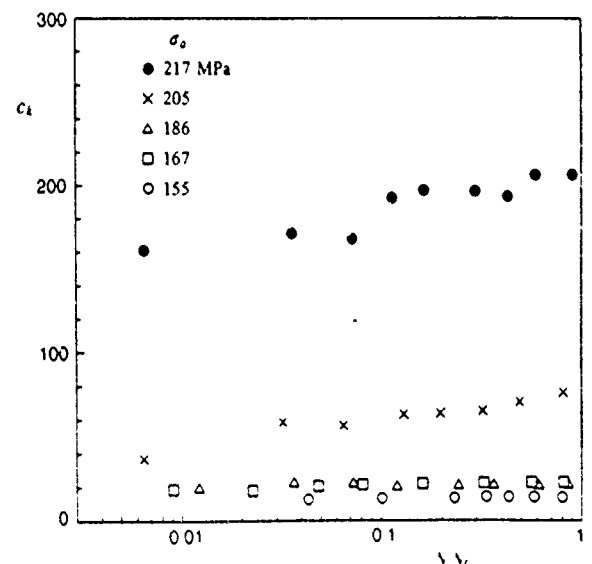


Fig. 13 Relationship between c_k and N/N_f under different stress amplitudes

Acknowledgments

Financial support for the work reported has been provided in part by U.S. Army Research Office, Engineering Science Division, through contract No. DAA0388K033. The major part of the work was done when A. Kato was a visiting scholar at SUNY at Stony Brook.

References

- (1) Miyagawa M. and Azushima A., "Measurement of Plastic Deformation by Means of a Laser Beam", Proc. of Ninth World Conference on Non-Destructive Testing (1979), p. 1.
- (2) Azushima A., Miyagawa M., Ohira T. and Kishi T., "Measurement of the Plastic Zone around a Fatigue Crack of a Stainless Steel Structure Member with a Laser Beam", Proc. of the 1982 SESA-JSME Joint Conference on Experimental Mechanics (1982), p. 709.
- (3) Lee C., Chao Y. J., Peters W. H. and Sutton M. A., "Surface Texture and Plastic Deformation", Proc. of the Southeastern Conference on Theoretical and Applied Mechanics (1986), p. 275.
- (4) Lee C., Chao Y. J., Sutton M. A., Peters W. H. and Ranson W. F., "Determination of Plastic Strains at Notches by Image-Processing Methods", Experimental Mechanics, Vol. 27, No. 2 (1989), p. 214.
- (5) Mughrabi H., Wang R., Differt K. and Essmann U., "Fatigue Crack Initiation by Cyclic Slip Irreversibilities in High-Cycle Fatigue", ASTM STP811 (1983), p. 5.
- (6) Lin T. H., Lin S. R. and Wu X. Q., "Micromechanics of an Extrusion in High-Cycle Fatigue", Proc XVII International Congress on Applied Mechanics, Grenoble, France (1988).

Plastically Induced Surface Roughness: Phenomena and Mechanism

Y.Z. Dai *

Brown and Sharpe Co., North Kingstown, RI 02852

F.P. Chiang

SUNY at Stony Brook, Stony Brook, NY 11794

Abstract

The surface roughness of a metallic material varies as a function of the plastic deformation it has experienced. Investigations on plastically induced surface roughness in literature are incomplete because the roughness parameter used can not define the surface roughness completely. Furthermore the mechanism of plastically induced surface roughness is, to a large extent, unknown. The purpose of this work is therefore two fold: Firstly, a more clear picture of the phenomena is provided through a more detailed description of surface roughness and its relation to material property, stress mode, and strain path. Secondly, the mechanism of the plastically induced surface roughness is experimentally studied in both macroscopic and microscopic scales.

Introduction

Studies on the roughening of free surfaces of metallic materials due to plastic deformation have been reported.¹⁻⁶ These studies discussed the influence of stress mode,^{1,2,4,5,8} grain size,^{1,2,7} material property,^{4,5,8} temperature,⁷ strain rate,⁷ and hydraulic pressure⁷ on surface roughness characteristics. Some models, which describe the relation between some of the above mentioned factors and surface roughness, have also been introduced.^{1,3,6,9} Despite the existence of some discrepancies due to the difference in test con-

*Formerly a graduate student at SUNY, Stony Brook.

ditions, such as the initial surface roughness, some conclusions have been drawn: For room temperature tests, the major conclusion is that the surface vertical roughness increases linearly proportional to the magnitude of plastic deformation and to the average grain size of the test specimen, while it is independent of stress mode, hydraulic pressure, strain rate, and material property.

This conclusion was made based on some experimental observations and the corresponding empirical relations. No investigations, to the authors' knowledge, have been made on the mechanism of plastic deformation induced roughening. Besides, almost all the studies only discussed the surface vertical characteristics, while the behavior of surface horizontal characteristics during plastic deformation has remained, to a large extent, unknown.

This work is designed to investigate the surface roughness response to plastic deformation of different materials and to study its mechanism by using coarse grained specimens such that the deformation of each individual grain due to plastic deformation may be revealed and characterized. It is expected that this approach will provide an explanation to the plastic deformation induced roughening phenomena.

Roughening Phenomena

Plastic deformation roughens free surfaces of metallic material by producing, among other things, slip bands within grains and relative rotation, sliding, among grains as shown in Fig.1. A rough surface may be completely described by two parameters, one vertical (root-mean-square roughness σ) and the other horizontal (correlation length T). These parameters may be approximately evaluated by using the profiles measured by a mechanical profilometer such as that shown in Fig.2. (A measured surface profile is shown in Fig.6a.) For a surface profile represented by discrete values $h(i)$ measured by such a system, these two parameters are defined as

$$\sigma = \sqrt{\frac{1}{N} \sum_{i=1}^N h^2(i)} \quad (1)$$

and

$$\frac{\sum_{i=1}^{N-T} h(i) \times h(i+T)}{\sum_{i=1}^N h^2(i)} = \frac{1}{\epsilon} \quad (2)$$

where N is the number of sampling points, and ϵ is equal to 2.718.

Surface roughness response to plastic deformation was studied on four different materials, namely: copper, aluminum, stainless steel, and hot-rolled steel. The specimens were cut along the rolling direction into a 'dog-bone' shape with dimensions of 2.15mm thick, 20.00mm wide and 125mm long. The surfaces of the specimens were then polished by a cloth buffer wheel and aluminum powder to an initial surface finish of $\sigma < 0.05\mu\text{m}$. Specimens were loaded on a universal testing machine to plastic deformation at incremental steps. At each step, the specimens were unloaded for surface roughness measurements. Surface roughness was measured using a system as systematically shown in Fig.2 with a stylus tip of 5 μm . The plastic deformation was evaluated in terms of true strain which is defined as

$$\epsilon_t = \ln(l/l_0) \quad (3)$$

where ϵ_t is the true plastic strain; l_0 and l are the specimen gage length before and after plastic deformation, respectively.

Figure 3 shows the experimental results of both the vertical and the horizontal surface roughnesses versus true strain ϵ_t for the four materials. As a first approximation the vertical roughness σ is linearly proportional to true plastic strain (Fig.3a); and the horizontal parameter, the correlation length T decreases with the increasement of tue strain(Fig.3b). These relations may be characterized by simple mathematical models as following

$$\sigma = a\epsilon_t + b \quad (4)$$

and

$$T = \frac{c}{\epsilon_t + d} + e \quad (5)$$

where a, b, c, d and e are constants.

As shown in Fig.3 surface vertical roughness seems to be independent of material property(Fig.3a); while the surface horizontal roughness seems to be dependent of material property(Fig.3b).

The influence of stress mode in terms of the ratio of principal stresses, which varied from -0.577 to -3.732 in a study, was investigated previously using specimens made of copper alloy in a disk shape as shown in Fig.4. Experimental results⁸ suggest that the surface roughness is independent of stress mode.

The strain path effect on surface roughness was also investigated using the set-up as systematically shown in Fig.4. Different strain paths in terms of principal strain ratio, which varied from -3.1 to 0.233 in the present work, were obtained by loading the specimen through different angle α . Figure 5 shows some of the strain paths achieved in the test, some of which were achieved by a one-step, while the others by a two-step loading to the final effective strain of 2.3%. Observations made on copper alloy material show that surface roughness is dependent on strain path. However, if no reverse yielding occurs, the process is virtually independent of strain path. (A strain path with reverse yielding is shown in the Fig.5 for $\alpha = 115^\circ \rightarrow -15^\circ$).

A Model for Roughening Mechanism

The plastically induced surface roughening mechanism was experimentally investigated using aluminum specimens with a grain size of 1.4 millimeters. Microscopically, the fundamental cause to surface roughening is due to dislocation movement. Macroscopically, the process manifests itself in a number of ways: Firstly, more and more slip bands come into being within each individual grain. Secondly, grains rotate with respect to each other and the resulting out-of-plane component of the relative rotation contributes to surface roughness. Thirdly neighboring grains may slide relatively to each other forming steps at the grain boundaries as can be seen in Fig.6a. Some other mechanisms, such as the bending of a single grain (section A-B in Fig.6a) and the rotation of multi-grains (section C-D in Fig.6a), may also come into play.

In order to provide an understanding to the roughening phenomenon, it is necessary that the dominant mechanism be isolated among the less significant factors. It was done by decomposing a measured profile into low and high spatial frequency components, and then comparing the surface roughness calculated for each component with that of the original profile. An example showing the decomposed low frequency component (Fig.6b), and the high frequency component (Fig.6c) is also given in Fig.6. Experimental evidence suggests that surface roughness σ, T of the low frequency component are almost the same as that of the measured profile which indicates that low frequency component is the dominant factor.

Specimens with grain sizes in the order of millimeter were used to study the plastically induced relative grain rotation. The grain boundaries in each specimen's surface was marked and the specimen was then polished

to almost mirror finish. The relative grain rotation was measured by using a set-up which consists of a laser, a partial mirror, a translation stage and an observation screen as schematically shown in Fig.7. The laser beam was directed to a grain on the specimen surface and the specular reflection point was marked on the observation plane. The specimen was then translated such that the neighboring grain was now under illumination and the specular reflection point from this grain was then determined. The relative angle between the two grains was determined by the distance between the two marks and the distance between the observation plane and the specimen. The relative grain rotating angle at each plastic deformation level was the average of the rotating angles measured for ten pairs of grains. Figure 8 shows the average grain rotation θ versus true plastic strain for specimens with four different grain sizes. They all appear to increase linearly with plastic strain.

Based on the experimental evidence mentioned above, an attemptive explanation to the plastic induced roughening may be given as follows.

For $0 < \epsilon_t < 3.5\%$: At this range the surface polishing process renders all the grains on a specimen surface be approximately on a plane. As a result, the initial *rms* roughness σ is small and the correlation length T is large. As the plastic deformation increases, slip bands come into being and grains begin to rotate with respect to each other. Because plastic deformation is small, some grains tend to be bonded together to rotate as a group with respect to the other grains or grain groups. With the progression of plastic deformation, less and less grains remain bonded resulting in the correlation length T becoming smaller and smaller. The combined effect of change in correlation length and grain rotation results in a increase of σ with respect to an increase in ϵ_t .

For $3.5\% < \epsilon_t < 18\%$: At this stage, almost all the grains have been involved in the relative rotation with their neighboring grains. This results in a slower generation rate of new low frequency profile components. Therefore the correlation length T approached saturation. The relative grain rotation still increases linearly with respect to plastic stain making the surface vertical roughness σ increase approximately linearly with plastic strain.

Conclusion

Through the experimental study and some analysis for the material stud-

ied under the experiment conditions, we may conclude that: 1. surface vertical roughness increases linearly proportional to plastic strain, that 2. surface horizontal correlation length decreases with plastic strain, that 3. surface vertical roughness is independent of material property, stress mode, and dependent of strain path, that 4. surface horizontal correlation length is dependent of material property, that 5. plastic deformation roughens surface by introducing slip bands within grains and relative rotation and sliding between grains, that 6. the low frequency component of a surface rprofile contributes dominantly to surface roughness parameters σ and T , that 7. the horizontal surface roughness parameter in terms of correlation length is proportional to a specimen's average grain size and it becomes saturated at a certain plastic deformation, that 8. the relative rotation between grains increases linearly to the amount of plastic deformation, and that 9. the surface vertical *rms* roughness is mainly due to grain rotation.

Acknowledgment

We would like to thank the Army Research Office, Engineering Science Division for financial support through contract No.DAA03-88-K-0033.

References

- [1] Osakada, K., Oyane, M. (1971), On the roughening of free surface in deformation process, *Bulletin of the JSME* 14, p.171-177.
- [2] Fukuda, M., Yamaguchi, K., Takakura, N., Sakano, Y. (1974), Roughening phenomenon on free surface of products in sheet metal forming, *J. Jap. Soc. Technol. Plast.*, 15, p.994-1002.
- [3] Parmar, A., Mellor, P.B. (1977), A new model for the prediction of instability and limit strains in thin sheet metal, *Int. J. Mech. Sci.*, 19, p.389-398.
- [4] Thomson, P.F., Nayak, P.U. (1980), The effect of plastic deformation on the roughening of free surfaces of sheet metal, *Int. J. of Mach. Tool Des. Res.*, 20, p.73-86.
- [5] Thomson, P.F., Shafer, B.V. (1982), The roughening of free surface during plastic working, *Int. J. of Mach. Tool Des. Res.*, 22, p.261-264.
- [6] Dautzenberg, J.H., Kals, J.A.G. (1984), Stress state and surface roughness, *Adv. Technol. of Plasticity*, 1, p.186-191.

- [7] Azushima, A., Miyagawa, M. (1986), Effect of working factors and metallurgical factors on roughening phenomenon on free surface of sheets, *J. of the Jap. Soc. for Technol. of Plasticity*, 27, p.1261-1267.
- [8] Chiang, F.P., Dai, Y.Z., Xu, B.Q., Kato, A. (1990), Study of surface roughening under different stress modes by correlation, in: *Hologram Interferometry and Speckle Metrology*, p.36-42, Baltimore.
- [9] Dai, Y.Z., Chiang, F.P. (1990), Application of scattering theory to plastic strain estimation, in: *The 4th Int. Sympo. on Nondestructive Characterization of Materials*, Annapolis.
- [10] Bennett, J.M., Mattsson, L. (1989), *Introduction to Surface Roughness and Scattering*, Optical Society of America, Washington D.C.
- [11] Dieter, G.E. (1968), *Mechanical Metallurgy*, Mc-Graw-Hill, New York.
- [12] Li, X.M., Chiang, F.P. (1990), High temperature creep crack growth in coarse-grained aluminum, in: *ASME 1990 Winter Annual Meeting*, Dalas

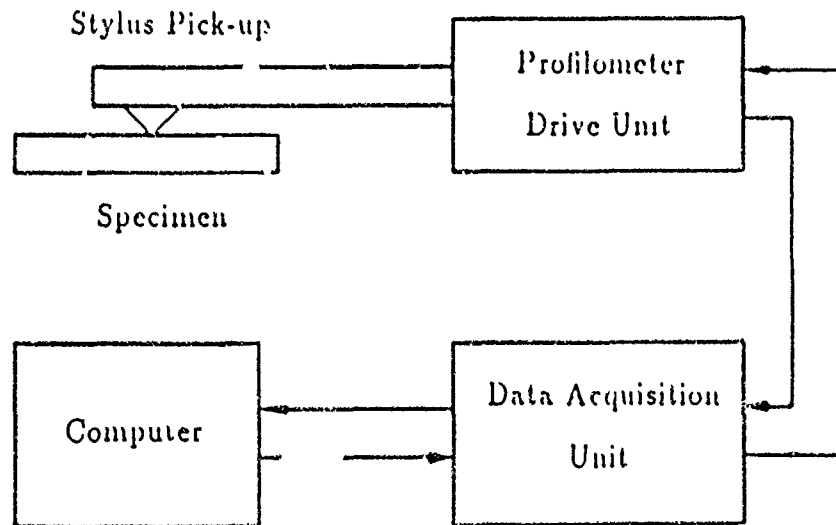


Figure 2 Set-up for surface roughness study

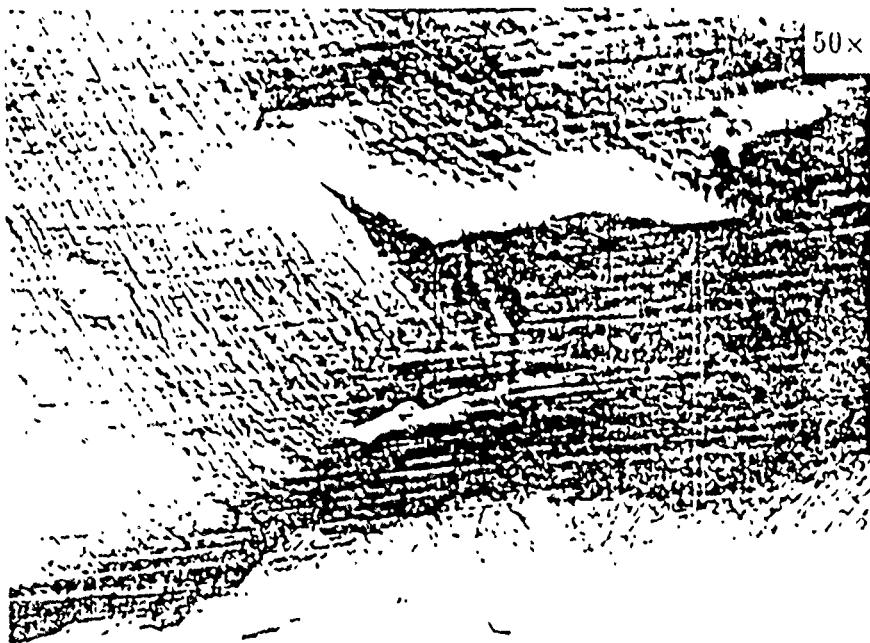
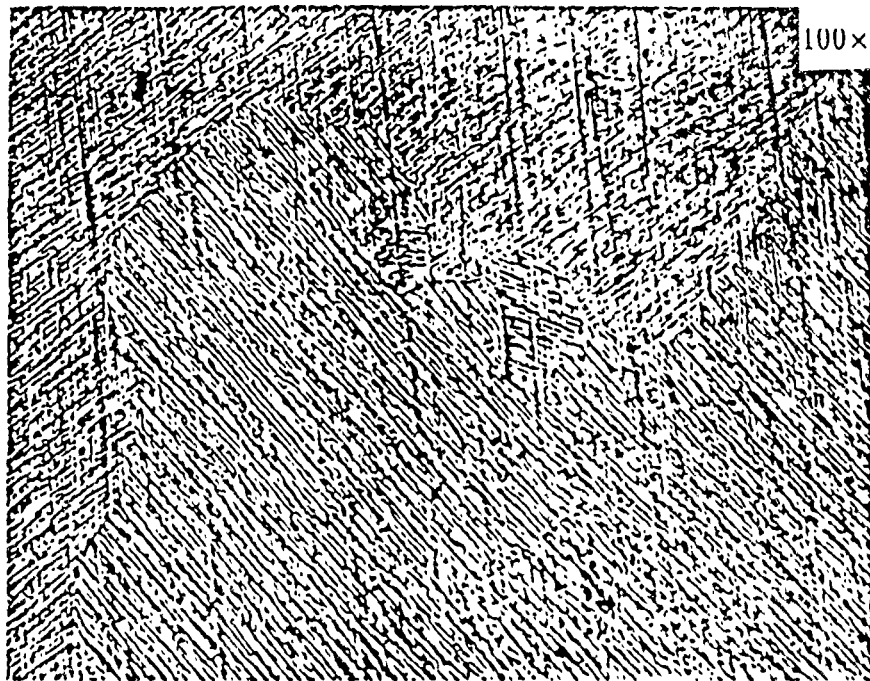


Figure 1 Plastically roughened surfaces revealing (a) slip bands within individual grains (b) grain rotation, grain boundary sliding among grains

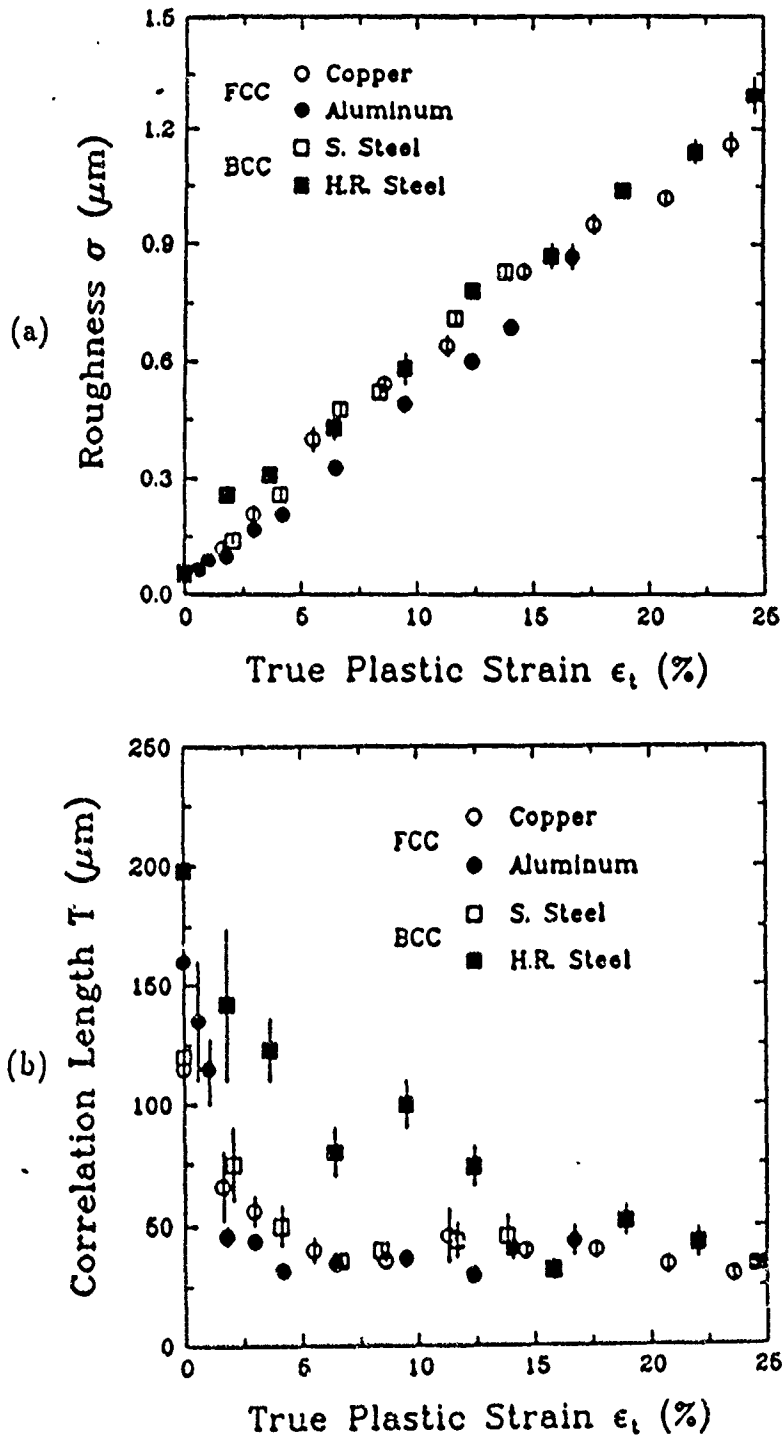


Figure 3 Relation between surface roughness and plastic strain

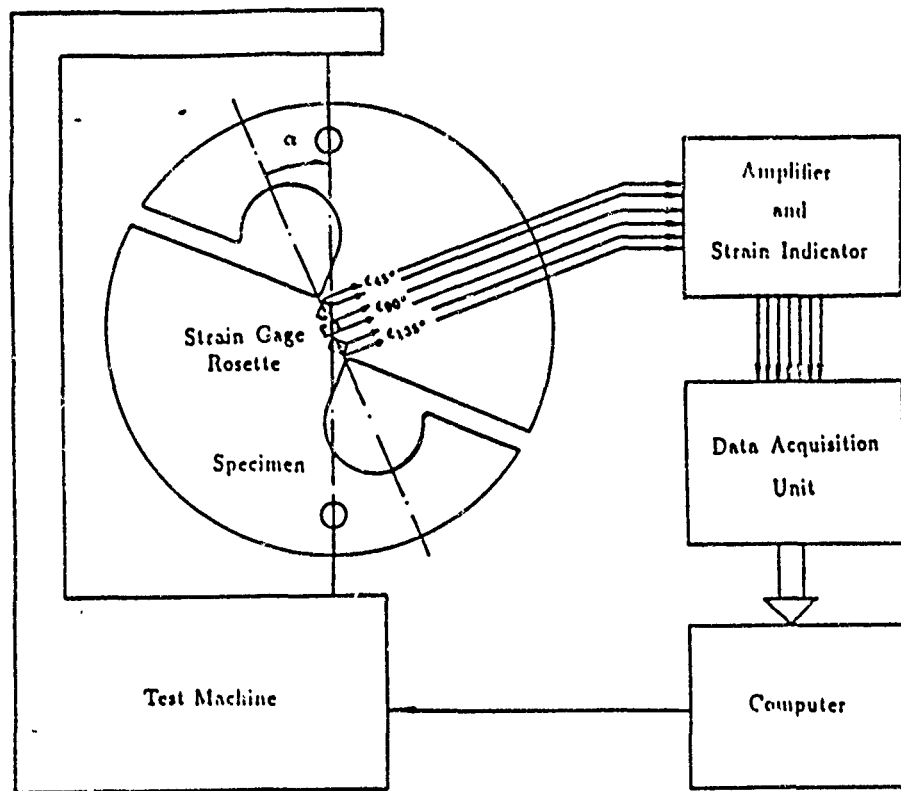


Figure 4 Schematic set-up for strain path effect study

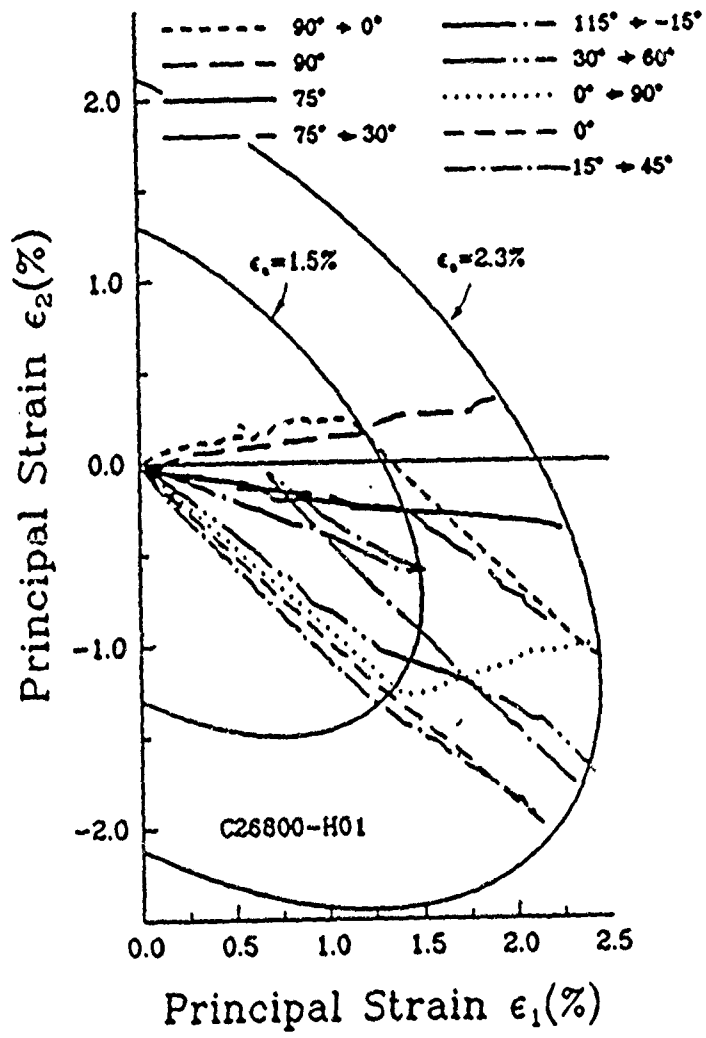


Figure 5 Stain path effect on surface roughness

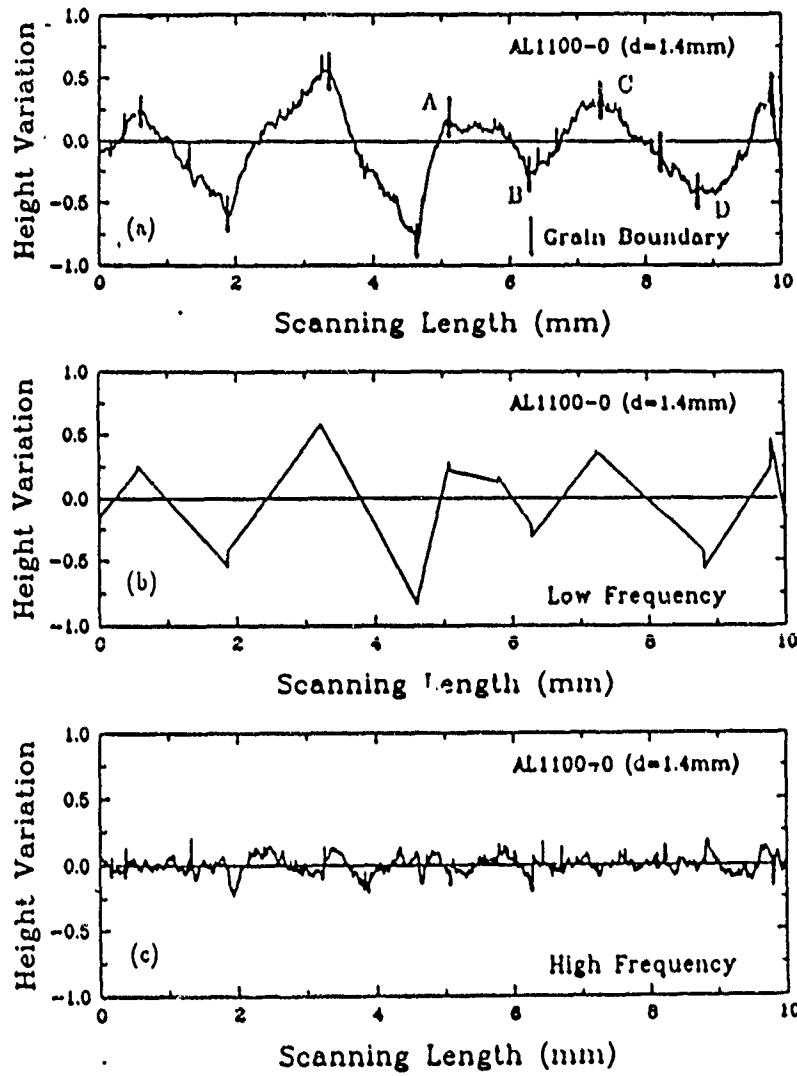


Figure 6 Decomposition of a profile (a) measured surface profile (b) low frequency component (c) high frequency component

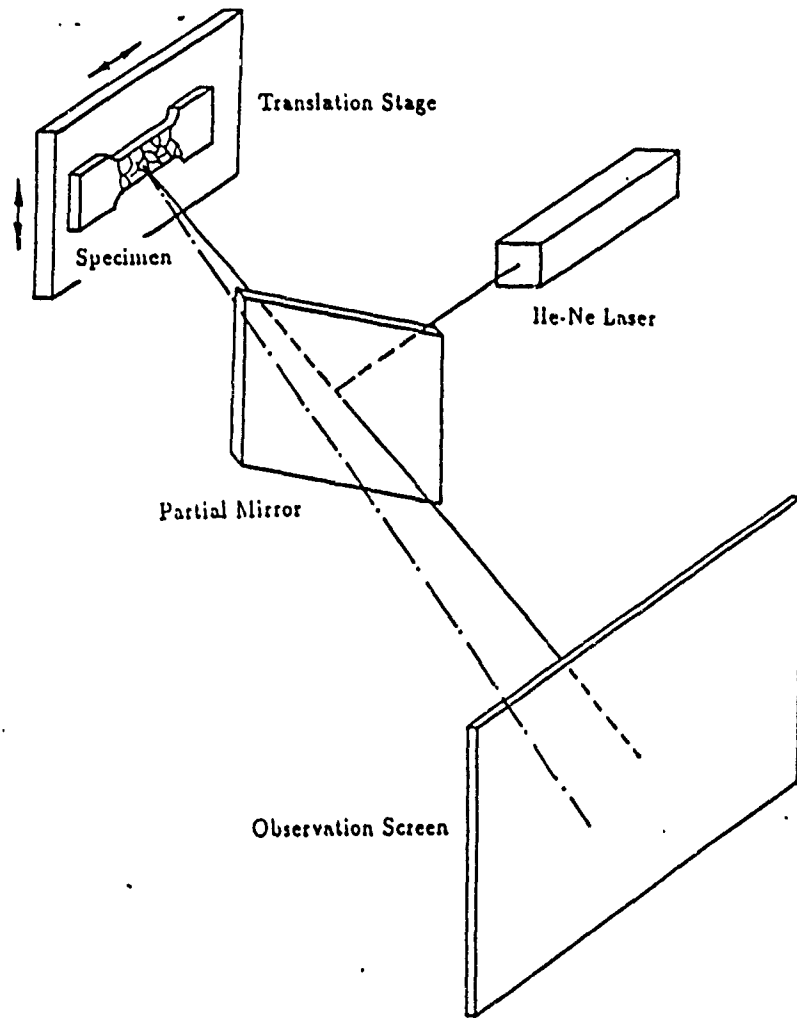


Figure 7 Set-up for relative grain rotation measurement

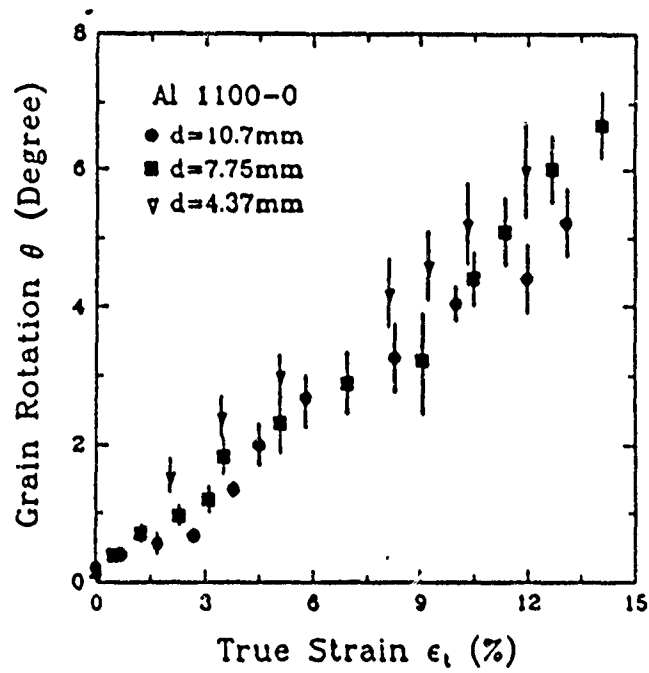


Figure 8 Relation between the relative grain rotation and plastic deformation

F

Experimental Measurement of Crack Tip Strain Field in A Single Crystal

X. M. Li and F. P. Chiang

Laboratory for Experimental Mechanics Research
State University of New York at Stony Brook
Stony Brook, NY 11794-2300, U.S.A.

J. Wu and M. Dudley

Department of Materials Science
State University of New York at Stony Brook
Stony Brook, NY 11794, U.S.A.

(June 1991)

- * Presented at International Seminar on LARGE PLASTIC DEFORMATIONS - Fundamental Aspects and Applications to Metal Forming, Aug. 7 - 9, 1991, Fontainebleau, France.

Abstract

Crack tip deformation fields in an aluminum single crystal are quantitatively obtained by applying three-dimensional moiré interferometry technique. Tensile specimen made of aluminum single crystal has an edge crack on $(\bar{2}23)$ plane with its tip along $[10, 7, 2]$ direction. Specimen with this specific orientation undergoes power law hardening with hardening index of n being 3.1. The experimental arrangement for moiré interferometry results in a displacement sensitivity of $0.47 \mu\text{m}/\text{fringe}$. Displacement fields at the free surface near crack tip are mapped with moiré fringe patterns. Both effective strain and maximum shear strain within a small region of less than 2 mm from the crack tip have been evaluated. In partial agreement with earlier analytical and numerical solutions, the results of this experiment show that the deformation fields near crack tip are divided into several different angular sectors, and concentrated shear are found at these sector boundaries where displacements are continuous. The experimental results also show that the effective strain and maximum shear strain remain constant (independent of angle θ) within certain angular sectors.

1 Introduction

An atomic sharp crack in ductile crystal is often blunted before it can propagate. As the result of complicated slip activities, the crack tip stress and deformation fields in such materials exhibit some unique characteristics. Bilby, Cottrell and Swinden (1963) proposed a model for elastic-plastic crack, known as the BCS model, in which an elastic crack was simulated by a double pile-up of dislocations. By considering the interactions between dislocations and crack, the plastic zone size (caused by dislocation pile-up in the crack tip region) was calculated. Crack tip blunting was first considered by Kelly, Tyson and Cottrell in 1967. They postulated that the localized shear stress in the vicinities of the crack could cause crack tip blunting. Rice and Thomson (1974) obtained a criterion for crack tip blunting by analyzing the condition for dislocation emission from an atomic sharp crack. Ohr and co-workers (Ohr and Narayan, 1980; Kobayashi and Ohr, 1980, 1981, 1984; Chang and Ohr, 1981; Ohr and Chang, 1982; Horton and Ohr, 1982) did a series of *in-situ* experiments on different metal thin foils to investigate the crack tip deformation fields. Their results confirmed the BCS theory. In further improvement they proposed a dislocation free zone model of fracture. Aforementioned theoretical and experimental studies together with the works of other investigators, e.g. Vitek (1976), Vitek and Chell (1980), Majumdar and Burns (1981, 1983), Thomson and Sinclair (1982), Lin and Thomson (1986), Weertman, Lin and Thomson (1983), Weertman (1978), Weertman and Weertman (1989) among others, have formed the framework of the dislocation theory of fracture. However, in these theories the crystals have been considered as linear elastic isotropic materials, and the slips are confined into one or at most two in-plane slips. They can not represent the real situations of single crystal fracture considered here. Furthermore, Ohr and co-worker's experiments were performed on thin crystal foils. As they pointed out in their paper (Chang and Ohr, 1983), the

thickness of the specimen had significant influences on the results.

Analytic solutions for crack tip stress and deformation fields in single crystals were also reported by Rice and co-workers. For elastic-ideally plastic single crystals, assuming crystal flow according to the critical resolved shear stress criterion, Rice and Nikolic (1985) and Rice (1987) have arrived asymptotic solutions for near tip stress field. In the case of ductile single crystals, Rice and Saeedvafa (1988) and Saeedvafa and Rice (1989) considered hardening effects and assumed that crystal obeying Taylor power law hardening, and proposed HRR type solutions for crack tip singular fields. These solutions reveal the formation of angular sectors around crack tip. The solution given for an elastic-ideally plastic crystal near a tensile crack (Rice, 1987) shows that the state of stress remains constant within each sector, and changes discontinuously from sector to sector which dictate that the corresponding displacement component u_r be discontinuous across sector boundaries. This is in contrast with the limiting case (perfect plasticity) of HRR type solution (Saeedvafa and Rice, 1989), in which displacement components are continuous while strain components are discontinuous yet bounded across sector boundaries. Hawk's (a preliminary summary is included in Rice, Hawk and Asaro, 1990) numerical analysis for elastic-ideally plastic crystals showed the discontinuity of displacement component u_r at sector boundaries and corresponding strain field with a Dirac singular form. Analytic solutions of Rice and co-workers were constructed with small-strain formulation, in which finite lattice rotations were ignored. As shown by Asaro (Asaro, 1979, 1983), finite lattice rotation may result, depending on the direction of lattice rotation, in geometrical softening or hardening. This in turn may cause several competing modes of deformation, and plays an important role in the formation of shear band (Asaro and Rice, 1977; Peirce, Asaro and Needleman, 1982, 1983).

Recent numerical analysis by Mohan, Ortiz and Shih (1990, 1991a,b) have considered finite deformation and finite lattice rotation, as well as 3-D crystallographic

geometry. In their finite element calculations, they used an experimentally based self-hardening rule (Chang and Asaro, 1981; Peirce, Asaro and Needleman, 1983; Kocks, 1970) showing an initial stage of rapid hardening followed by a saturation stage. For small deformation, the crystal exhibits a high rate of hardening. When finite strain is approached, crystal behaves like ideally plastic. Their results differed from the analytic solutions of Rice and Nikolic (1985) and Rice (1987) in the region deep inside the plastic zone where plastic saturation was attained. Indeed large lattice rotations were found in this region. It is believed that the large lattice rotations will contribute to the activation of all slip systems and affect significantly the nature of the near tip solutions. As shown by Mohan, Ortiz and Shih (1991a), for example, the development of in-plane stress components of magnitude comparable to that of the out-of-plane shear stresses in the vicinities of anti-plane shear crack tip resulted in the loss of the anti-plane character of the solution. In the region outside the crack tip where plastic strain was small and crystal hardened rapidly, numerical calculations of Mohan, Ortiz and Shih (1991a,b) tended to be consistent with analytical solutions of Rice and Saeedvafa (1988) and Saeedvafa and Rice (1989). Mohan, Ortiz and Shih's (1991b) results were in close agreement with experimental observations of Shield, Kim and Nikolic (1989). Shield, Kim and Nikolic (1989) studied the near tip deformation field in a four-point bend specimen of Iron-Silicon single crystal. In their experiments, they mapped the near tip displacement field of one component using a moiré microscope, and observed narrow bands of fringe kinks representing a jump in displacement component. From parallel and equally spaced fringe patterns, as they stated in their paper, strain component in opening direction changed from one constant value to another across these narrow bands. However, no quantitative measurement of the displacement field near a crack tip was offered.

In this study, we measured displacement fields near a tensile crack tip in an aluminum single crystal by applying a three-dimensional technique of moiré interfer-

ometry. Three components of displacement field at the free surface near crack tip were directly obtained through moiré fringe patterns. The arrangement of moiré interferometry resulted in a displacement sensitivity of $0.47 \mu\text{m}/\text{fringe}$. The specimen used in this experiment had a crystal orientation with crack face on $(\bar{2}23)$ plane and its tip along $[10,7,2]$ direction (as shown in Fig.2a), and underwent power-law hardening with hardening index of n being 3.1 (as shown in Fig.3a). Both effective strain and maximum shear strain within a small region of less than 2 mm around crack tip were evaluated. Evident from the quantitative measurement of crack tip deformation fields are the formations of angular sectors with sector boundaries emanating from crack tip. The strain fields are found to be locally constant (independent of θ) within certain sectors, and concentrated shears are obtained at these sector boundaries. Comparison with earlier analytical and numerical solutions, as well as experimental observations are also discussed.

2 Plane strain solution

2.1 Plane strain yield surface

Aluminum single crystals (they are f.c.c. crystals) have 12 possible slip systems. Plastic flow occurs on a given slip system only when the resolved shear stress on that system reaches a critical value, and this value increases as the material hardening takes place. It is assumed that the crystal obeys Taylor's power-law hardening. Therefore the critical resolved shear stress remains the same for all possible slip systems as hardening develops (equal hardening on all systems). The stress-strain relation at sufficiently large strain is given by

$$\gamma = a\tau^n \quad (1)$$

where a is the material hardening constant, n is the material hardening index, γ is the effective shear strain, i.e. the sum of the absolute shears on all slip systems (as defined by Rice and Saeedvafa, 1989), and τ is the critical resolved shear stress (same for all possible slip systems).

Following Rice's theory (1973), yield surface of such a single crystal, under the assumptions of plane strain deformation and incompressibility, can be represented as a closed curve in a plane whose axes are $(\sigma_{11} - \sigma_{22})/2\tau$ and σ_{12}/τ . This curve is found to be the inner envelope of the planar yield surfaces for individual slip systems or groups of systems, and has the form of a polygon. For our specific problem, the plane of deformation was defined as $x_1 - x_2$ plane with x_1 along $[\bar{1}2\bar{2}]$ direction and x_2 along $[\bar{2}23]$ direction. The yield surface in the space of the ratio of the stresses to the critical resolved shear stress τ was determined and shown in Fig.3b. The active slip systems for this specific orientation are also depicted in this figure.

2.2 Flat sectors and vertex sectors

The yield surface shown in Fig.3b has five flat segments and five vertex points. Each flat segment corresponds to one active slip system, and each vertex point to two active slip systems, resulting in two types of near tip solutions. If we rotate our coordinate system, according to Saeedvafa and Rice (1989), such that for flat sectors the axis of $(\sigma'_{11} - \sigma'_{22})/2\tau$ is parallel to the corresponding flat segment and for vertex sectors the axis of $(\sigma'_{11} - \sigma'_{22})/2\tau$ is passing through the corresponding vertex point (as shown in Fig.4a and Fig.4b). Then the displacement components and stress components for a given sector can be represented in new coordinate system as follows

$$u'_1 = B_1 x'_2 |x'_2|^{-\frac{n}{n+1}} F_1(C_2) \quad (2)$$

$$u'_2 = B_2 x'_1 |x'_1|^{-\frac{n}{n+1}} F_1(C_2) \quad (3)$$

$$\sigma'_{13} = (n+1)^{-\frac{1}{2}} \beta^{\frac{n+1}{2}} (B_1 |x'_2|^{-\frac{n}{n+1}} + B_2 |x'_1|^{-\frac{n}{n+1}})^{\frac{1}{2}} F_2(C_2) \quad (4)$$

$$\sigma'_{11} = -\frac{\sigma'_{12}}{n+1} \left(n \frac{B_2 x'_1 |x'_1|^{-\frac{n}{n+1}}}{B_1 x'_2 |x'_2|^{-\frac{n}{n+1}}} - \frac{x'_1}{x'_2} \right) - \frac{\cos \omega}{B_1 x'_2 |x'_2|^{-\frac{n}{n+1}}} F_2(C_2) \quad (5)$$

$$\sigma'_{22} = -\frac{\sigma'_{12}}{n+1} \left(n \frac{B_1 x'_2 |x'_2|^{-\frac{n}{n+1}}}{B_2 x'_1 |x'_1|^{-\frac{n}{n+1}}} - \frac{x'_2}{x'_1} \right) + \frac{\sin \omega}{B_2 x'_1 |x'_1|^{-\frac{n}{n+1}}} F_2(C_2) \quad (6)$$

for the flat sectors, and

$$u'_1 = \frac{x'_1 (A_1 |x'_2|^{-\frac{1}{n+1}} + A_2 |x'_1|^{-\frac{1}{n+1}})^{n+1} - (2\alpha)^{n+1} \cos \psi}{(2\alpha)^{n+1} A_1 |x'_2|^{-\frac{1}{n+1}}} F_1(C_2) \quad (7)$$

$$u'_2 = -\frac{x'_2 (A_1 |x'_2|^{-\frac{1}{n+1}} + A_2 |x'_1|^{-\frac{1}{n+1}})^{n+1} + (2\alpha)^{n+1} \sin \psi}{(2\alpha)^{n+1} A_2 |x'_1|^{-\frac{1}{n+1}}} F_1(C_2) \quad (8)$$

$$\sigma'_{12} = 0 \quad (9)$$

$$\sigma'_{11} = A_1 |x'_2|^{-\frac{1}{n+1}} F_3(C_2) \quad (10)$$

$$\sigma'_{22} = -A_2 |x'_1|^{-\frac{1}{n+1}} F_3(C_2) \quad (11)$$

for the vertex sectors.

Where

$$x'_1 = \begin{cases} r \cos(\theta - \omega) & \text{for flat sectors.} \\ r \cos(\theta - \psi) & \text{for vertex sectors.} \end{cases}$$

$$x'_2 = \begin{cases} r \sin(\theta - \omega) & \text{for flat sectors.} \\ r \sin(\theta - \psi) & \text{for vertex sectors.} \end{cases}$$

$$F_1(C_2) = a^{\frac{1}{n+1}} [(n+1)C_2]^{\frac{1}{n+1}} \quad (12)$$

$$F_2(C_2) = \left[\frac{(n+1)C_2}{a} \right]^{\frac{1}{n+1}} \quad (13)$$

A_1 and A_2 are constants in each vertex sector, B_1 and B_2 are constants in each flat sector, J_2 is a common constant for all sectors (related to J-integral, as shown in Rice and Saeedvafa, 1988), β is the distance from flat segment to origin as defined in Fig.4a, and α is the distance from vertex point to origin as defined in Fig.4b, r and θ are polar coordinates as defined in Fig.1, and ψ and ω are angles as defined in Fig.4a and Fig.4b.

There are two unknown constants involved in each angular sector, and one unknown common constant C_2 for all sectors. A_1 and A_2 (or, B_1 and B_2) can be determined from its sector boundary conditions. C_2 has the unit of J-integral, and is used for the normalization of the outer field. In the present study, we are interested in the near tip angular sector arrangement to aid the analysis of our experimental results. The constant C_2 is therefore left undetermined.

2.3 Sector arrangement

The flat segment on yield surface is bounded by two vertex points, and the vertex point is adjoined by two flat segments. The validity of previous solutions are therefore limited as follows (Saeedvafa and Rice, 1989)

$$\frac{2}{\tan 2(\omega - \psi^-)} \leq G(x'_1, x'_2) \leq \frac{2}{\tan 2(\omega - \psi^+)} \quad (14)$$

for the flat sectors, and

$$\frac{2(n+1)}{\tan 2(\psi - \omega^-)} \leq Q(x'_1, x'_2) \leq \frac{2(n+1)}{\tan 2(\psi - \omega^+)} \quad (15)$$

for the vertex sectors.

Where

$$G(x'_1, x'_2) = G_1(x'_1, x'_2) + G_2(x'_1, x'_2) + G_3(x'_1, x'_2) + G_4(x'_1, x'_2)$$

$$Q(x'_1, x'_2) = Q_1(x'_1, x'_2) - Q_2(x'_1, x'_2) - Q_3(x'_1, x'_2)$$

$$G_1(x'_1, x'_2) = \frac{n}{n+1} \left(\frac{A_2 x'_1 |x'_1|^{-\frac{n+1}{n}}}{A_1 x'_2 |x'_2|^{-\frac{n+1}{n}}} - \frac{A_1 x'_2 |x'_2|^{-\frac{n+1}{n}}}{A_2 x'_1 |x'_1|^{-\frac{n+1}{n}}} \right)$$

$$G_2(x'_1, x'_2) = \frac{1}{n+1} \left(\frac{x'_2}{x'_1} - \frac{x'_1}{x'_2} \right)$$

$$G_3(x'_1, x'_2) = \frac{(n+1)^{\frac{1}{n}} \cos \omega}{\beta^{\frac{n+1}{n}} B_1 x'_2 |x'_2|^{-\frac{n+1}{n}} (B_1 |x'_2|^{-\frac{n+1}{n}} + B_2 |x'_1|^{-\frac{n+1}{n}})^{\frac{1}{n}}}$$

$$G_4(x'_1, x'_2) = \frac{(n+1)^{\frac{1}{n}} \sin \omega}{\beta^{\frac{n+1}{n}} B_2 x'_1 |x'_1|^{-\frac{n+1}{n}} (B_1 |x'_2|^{-\frac{n+1}{n}} + B_2 |x'_1|^{-\frac{n+1}{n}})^{\frac{1}{n}}}$$

$$Q_1(x'_1, x'_2) = \frac{A_2 x'_1 |x'_1|^{-\frac{n+1}{n}}}{A_1 x'_2 |x'_2|^{-\frac{n+1}{n}}} - \frac{A_1 x'_2 |x'_2|^{-\frac{n+1}{n}}}{A_2 x'_1 |x'_1|^{-\frac{n+1}{n}}} - n \left(\frac{x'_1}{x'_2} - \frac{x'_2}{x'_1} \right)$$

$$Q_2(x'_1, x'_2) = \frac{(2\alpha)^{n+1} \cos \psi}{A_1 x'_2 |x'_2|^{-\frac{n+1}{n}} (A_1 |x'_2|^{-\frac{n+1}{n}} + A_2 |x'_1|^{-\frac{n+1}{n}})^n}$$

$$Q_3(x'_1, x'_2) = \frac{(2\alpha)^{n+1} \sin \psi}{A_2 x'_1 |x'_1|^{-\frac{n+1}{n}} (A_1 |x'_2|^{-\frac{n+1}{n}} + A_2 |x'_1|^{-\frac{n+1}{n}})^n}$$

ψ^- and ψ^+ are angles as defined in Fig.4a, and ω^- and ω^+ are angles as defined in

Fig.4b.

The boundaries of sectors ($\theta = \theta_f$, for flat sectors; and $\theta = \theta_v$, for vertex sectors) are determined from the following equations

$$[G(x'_1, x'_2)]_{\theta_f} = \frac{2}{\tan 2(\omega - \psi \pm)} \quad (16)$$

for flat sectors, and

$$[Q(x'_1, x'_2)]_{\theta_v} = \frac{2(n+1)}{\tan 2(\psi - \omega \pm)} \quad (17)$$

for vertex sectors.

In the previous section, displacement components of both flat and vertex sectors have the form of a function of x'_1 and x'_2 multiplied by $F_1(C_2)$; and stress components of both flat and vertex sectors have the form of a function of x'_1 and x'_2 multiplied by $F_2(C_2)$. With C_2 , and therefore $F_1(C_2)$ and $F_2(C_2)$, left as unknown constants for all sectors, we were able to determine the two unknown constants in each sector from boundary conditions at crack surface together with continuity conditions at each sector boundaries.

The crack surface boundary conditions are

$$\sigma_{13} = 0 \quad \text{at } \theta = \pm\pi \quad (18)$$

$$\sigma_{23} = 0 \quad \text{at } \theta = \pm\pi \quad (19)$$

The sector arrangement corresponding to the yield surface is shown in Fig.5a. Constants B_1 and B_2 in flat sector I were determined from boundary conditions at

sector	$\hat{\phi}$	sector	θ
I	$-180.000^{\circ} \sim -91.031^{\circ}$	VI	$45.580^{\circ} \sim 78.174^{\circ}$
II	$-91.031^{\circ} \sim -64.274^{\circ}$	VII	$78.174^{\circ} \sim 86.473^{\circ}$
III	$-64.274^{\circ} \sim -60.125^{\circ}$	VIII	$86.473^{\circ} \sim 117.821^{\circ}$
IV	$-60.125^{\circ} \sim 34.891^{\circ}$	IX	$117.821^{\circ} \sim 180.000^{\circ}$
V	$34.891^{\circ} \sim 45.580^{\circ}$		

Table 1: Boundaries between angular sectors around crack tip

the crack surface, i.e. Eqs.(18)&(19) at $\theta = -\pi$. The sector boundary between flat sector I and vertex sector II was resolved from Eq.(16). Constants A_1 and A_2 in vertex sector II were calculated from continuity conditions of stress components at the boundary between flat sector I and vertex sector II. The sector boundary between vertex sector II and flat sector III was obtained from Eq.(17). Constants B_1 and B_2 in flat sector III were found from continuity conditions of displacement components at the sector boundary between vertex sector II and flat sector III. Following the same procedure, two unknown constants in each sector, and therefore sector boundaries can be calculated for all sectors. Since Eqs.(2)-(11) and Eqs.(16)&(17) are all nonlinear equations which can only be solved numerically, there exist more than one solutions. The boundary conditions at the crack surface, i.e. Eqs.(18)&(19) at $\theta = +\pi$, were checked from solutions in the last sector (flat sector IX) to ensure that the solution so obtained was correct.

Calculated sector boundaries for this specific crystal orientation are collected and shown in table 1. The sector arrangement in the plane of deformation is also shown in Fig.5b.

3 Experiment

3.1 Specimen

The tensile specimen of aluminum single crystal was cut from one huge single crystal sheet obtained from strain-annealing process for crystal growth. The specimen used in this experiment had the dimension of $40 \times 13 \times 2 \text{ mm}^3$. An edge crack with finite radius of about $76 \mu\text{m}$ was machine cut, and followed by annealing at 150°C for 5 hours to eliminate the possible residual stress induced during machining process. By using X-ray diffraction technique, the crystal orientation was determined such that the crack surface was on $(\bar{2}23)$ plane and its tip along $[10,7,2]$ direction. Specimen surface was then polished and chemically etched. The material was calibrated through a simple tension test, using the same aluminum single crystal, resulting in a stress-strain curve as shown in Fig.3a. Ramberg-Osgoog relation written in a piece-wise formulation was used to fit the curve, i.e.

$$\frac{\epsilon}{\epsilon_0} = \begin{cases} \sigma/\sigma_0 & \epsilon \leq \epsilon_0 \\ a(\sigma/\sigma_0)^n & \epsilon \geq \epsilon_0 \end{cases} \quad (20)$$

where σ_0 is the yield stress; ϵ_0 is the yield strain; n is the material hardening index; and a is the material hardening constant. The calibrated material parameters are indicated in Fig.3a.

3.2 Three-dimensional moiré interferometry

Moiré interferometry was used to measure the near tip three-dimensional displacement component. The detailed analysis of moiré interferometry can be found in many papers (e.g. Post, 1987; Asundi, 1989). Fig.6 shows the principle of moiré interferometry. A high frequency reflective phase grating either replicated onto the specimen surface or onto a glass plate used as a real reference grating is illuminated by two collimated laser beams A and B. These two coherent laser beams which are incident at

symmetrical angles $\pm\alpha$ are diffracted by reflective phase grating. The incident angle α is adjusted such that the direction of diffracted beam of +1 order (resulted from the incident beam of A), and therefore the -1 order (resulted from B) are 0° . If the reflective phase grating remains unchanged, then both diffracted beams of +1 order and -1 order will have plane wavefront, and no interference pattern will be observed. If the grating was replicated onto the specimen surface and distorted as the result of in-plane deformation, then both diffracted beams of +1 order and -1 order would have warped wavefront. The interference of these two coherent laser beams with warped wavefront will result in a moiré fringe pattern which is contours of in-plane displacement components. If a real reference grating was used and remained unchanged, but beam A was reflected back from the specimen with a mirror surface and had a warped wavefront resulted from out-of-plane deformation, the diffracted beam of +1 order would have a warped wavefront while -1 order retain its plane wavefront. Thus a moiré fringe pattern which represents contours of out-of-plane displacement component would be generated from the interference of these two diffracted beams.

Fig.7 shows the experimental set-up for simultaneous measurement of in-plane and out-of-plane displacement. A reflective cross-grating (phase grating) of 1200 line/mm was replicated onto the specimen surface, and a real reflective reference grating (phase grating) of the same frequency was placed perpendicular to the specimen surface. Camera A was focused on the specimen surface to record the in-plane displacement fields, while camera B was focused on the real reference grating to obtain the out-of-plane displacement field. A He-Ne laser with a wavelength of $\lambda = 633$ nm was used as the light source. The laser beam was split into two collimated laser beams. Two opaque screen were used to block these two beams from striking on the specimen surface, the reference grating, and the mirrors. By allowing light from sections c' and d' impinging on reflective grating c and specimen grating d , u_1 -field moiré fringe pattern can be obtained at camera back A. Let light passing through

sections a' and b' and illuminating adjustable mirrors a and b , u_2 -field moiré fringe pattern can be observed at camera back A. To record u_3 -field moiré fringe pattern at camera back B, reflective grating e and specimen surface f are illuminated by coherent laser beams through sections e' and f' . This experimental arrangement resulted in a displacement sensitivity of $0.417 \mu\text{m}/\text{fringe}$. The u_1 -field, u_2 -field, and u_3 -field moiré fringe patterns are the contours of displacement components along x_1 , x_2 , and x_3 -direction respectively, and they are governed by the following equations

$$u_1 = \frac{1}{2} N_1 p \quad (21)$$

$$u_2 = \frac{1}{2} N_2 p \quad (22)$$

$$u_3 = \frac{1}{2} N_3 p \quad (23)$$

where u_1 , u_2 , and u_3 are displacement components along x_1 , x_2 , and x_3 -direction respectively; N_1 , N_2 , and N_3 are fringe orders of u_1 -field, u_2 -field, and u_3 -field moiré fringe patterns respectively ($0, \pm 1, \pm 2, \dots$); p is the pitch of the specimen grating (same as the pitch of the real reference grating).

To evaluate the near tip deformation field we proceeded to calculate the strain distributions from the displacement values. Moiré fringe patterns (such as shown in Fig.8) were digitized and numerically differentiated with respect to the spatial coordinates x_1 and x_2 using a smoothed cubic spline approximations (Berghuaus and Cannon, 1977). The in-plane strain components were obtained from the following relations

$$\epsilon_{11} = \frac{1}{2} p \frac{\partial N_1}{\partial x_1} \quad (24)$$

$$\epsilon_{22} = \frac{1}{2} p \frac{\partial N_2}{\partial x_2} \quad (25)$$

$$\epsilon_{12} = \frac{p}{4} \left(\frac{\partial N_2}{\partial x_1} + \frac{\partial N_1}{\partial x_2} \right) \quad (26)$$

Under the assumption of material incompressibility, we have (neglect the elastic deformation near the crack tip)

$$\epsilon_{33} = -(\epsilon_{11} + \epsilon_{22}) \quad (27)$$

The effective strain ϵ_e and the maximum shear strain γ_{max} are calculated using the following equations

$$\epsilon_e = \frac{\sqrt{2}}{3} [(\epsilon_{11} - \epsilon_{22})^2 + (\epsilon_{22} - \epsilon_{33})^2 + (\epsilon_{33} - \epsilon_{11})^2 + 6\epsilon_{12}^2]^{\frac{1}{2}} \quad (28)$$

$$\gamma_{max} = \left[\left(\frac{\epsilon_{11} - \epsilon_{22}}{2} \right)^2 + \epsilon_{12}^2 \right]^{\frac{1}{2}} \quad (29)$$

The displacement components, the effective strain, and the maximum shear strain were then converted into that in the polar coordinate system, and are shown in Fig.9 and Fig.10, respectively.

3.3 Experimental results

Fig.8 shows the moiré fringe patterns obtained at $\sigma_{\infty} = 0.8\sigma_0$, where σ_{∞} was averaged over the uncracked ligament. These moiré fringe patterns represent the contours of displacement components at the free surface near crack tip. It is evident that the deformation fields are divided into several different angular sectors with sector boundaries emanating from crack tip at $\theta = -89.0^\circ$, $\theta = -64.0^\circ$, and $\theta = 34.0^\circ$. These are also distinguished from the plots of effective strain and maximum shear strain as a function of θ at different radial distance from the crack tip (as shown in Fig.10). Concentrated shear were found at some sector boundaries. Both effective strain and maximum shear strain were locally constant (independent of angle θ) within certain angular sectors. Only three angular sectors were observed at this stage of deformation.

4 Discussion and conclusion

Small-strain solutions of Rice and co-workers did not include the effects of lattice rotation. As pointed out by Mohan, Ortiz and Shih (1991a,b), notable discrepancies will arise if finite deformation and finite lattice rotation are allowed. This is expected since finite lattice rotation will induce geometrical hardening or softening, and in turn may affect yield condition as well as the hardening law. In the numerical analysis of Mohan, Ortiz and Shih (1991a,b), finite deformation in the vicinities of crack tip have exceeded its saturation and thus been treated with perfect plasticity. The results of their calculation did differ from those small-strain solutions of Rice and Nikolic (1985), of Rice (1987), and of Rice, Hawk and Asaro (1990) for elastic-ideally plastic crystals, but agreed well with the experimental observations of Shield, Kim and Nikolic (1989). However, in the region away from crack tip where deformation was small and crystal exhibited high hardening, the results of Mohan, Ortiz and Shih (1991a,b) tended to be consistent with the analytical solutions of Rice and Saeedvafa (1988) and Saeedvafa and Rice (1989) for high hardening crystals.

In this study, the deformation fields measured in the region near crack tip were relatively small (one order of magnitude smaller than those observed in Shield, Kim and Nikolic's experiment (1989)). Crystal exhibited power-law hardening, and therefore perfect plasticity can not be applied. The observations from this experiment provide clear evidence for the formation of angular sectors around crack tip with sector boundaries emanating from crack tip at $\theta = -91.0^\circ$, $\theta = -64.0^\circ$, and $\theta = 34.0^\circ$. Corresponding to these observed sector boundaries, the analytical solution based on Saeedvafa and Rice's analysis (1989) predicts sector boundaries at $\theta = -91.031^\circ$, $\theta = -64.274^\circ$, and $\theta = 34.891^\circ$. An interesting observation can be made in comparison between the experimental result and the analytical solution. Despite the narrow region in sector III (see Fig.5b) which could not be identified from the experimental

observations, all the three distinguished angular sectors (sector I, II, and IV) were the results of active slip on $(11\bar{1})$ plane. In other words, all the three slip systems on $(11\bar{1})$ plane have been activated at this deformation level, and they contribute to the formation of angular sectors. However from Fig.10 it is showed that the effective strain and maximum shear strain remained constant in these sectors. This does not agree with the theoretical predictions (Saeedvafa and Rice, 1989). The locally constant strain was also obtained by Shield, Kim and Nikolic (1989), in which they observed parallel and equal spaced moiré fringes indicating a constant strain within certain sectors. Note that the theoretical solution was obtained under the assumptions of plane strain deformation and material incompressibility, while experimental measurement was taken at the specimen free surface where three-dimensional deformation actually occurred. Near tip lattice rotation may also contribute to the discrepancies.

The formation of angular sectors around crack tip with concentrated shear at sector boundaries have been presented as an unique characteristic of crack tip plasticity field in single crystals. The fact that locally constant strain were found within certain sectors can not be explained by solutions available at present. Within the context of small-strain deformation, it is not clear that to what extent the lattice rotation may influence the near tip stress and deformation fields. Some of the experimentally observed features remain to be further explained.

Acknowledgements

This work is financially supported by the Office of Naval Research, Mechanics Division, through Contract No. N0001482K0566 (Scientific Officer, Dr. Y. Rajapakse). The authors are indebted to Prof. C. F. Shih for providing us their papers in advance. Technical assistance from Ms. Y. Y. Wang during this experiment is also appreciated.

Yap a D.

Reference

- Asaro, R. J., 1979, 'Geometrical effects in the inhomogeneous deformation of ductile single crystals', *Acta Metall.*, Vol.27, pp.445.
- Asaro, R. J., 1983, 'Micromechanics of crystals and polycrystals', *Adv. Appl. Mech.*, Vol.23, pp.1.
- Asaro, R. J., and Rice, J. R., 1977, 'Strain localization in ductile single crystals', *J. Mech. Phys. Solids*, Vol.25, pp.309.
- Asundi, A., Cheung, M. T., and Lee, C. S., 1980, 'Moiré interferometry for simultaneous measurement of u, v, w', *Exp. Mech.*, Vol.29, No.3, pp.258.
- Berghaus, D. G., and Cannon, J. P., 1973, 'Obtaining derivatives from experimental data using smoothed-spline functions', *Exp. Mech.*, Vol.13, pp.19-²³
- Bilby, B. A., Cottrell, A. H., and Swinden, K. H., 1963, 'The spread of plastic yield from a notch', *Proc. R. Soc. Lond.*, Vol.A272, pp.304.
- Chang, Y. W., and Asaro, R. J., 1981, 'An experimental study of shear localization in aluminum-copper single crystals', *Acta Metall.*, Vol.29, pp.241.
- Chang, S. J., and Ohr, S. M., 1981, 'Dislocation-free zone model of fracture', *J. Appl. Phys.*, Vol.52, pp.7174.
- Chang, S. J., and Ohr, S. M., 1983, 'Effect of thickness on plastic zone size in BCS theory of fracture', *Int. J. Fract.*, Vol.21, pp.3.
- Horton, J. A., and Ohr, S. M., 1982, 'TEM observations of dislocation emission at crack tips in aluminum', *J. Mater. Sci.*, Vol.17, pp.3140.
- Kelly, A., Tyson, W. R., and Cottrell, A. H., 1967, 'Ductile and brittle crystals', *Phil. Mag.*, Vol.15, pp.567.
- Kobayashi, S., and Ohr, S. M., 1980, 'In-situ fracture experiments in b.c.c. metals', *Phil. Mag.*, Vol.42, pp763.
- Kobayashi, S., and Ohr, S. M., 1981, 'In-situ observations of the formation of plastic zone ahead of a crack tip in copper', *Scr. Metall.*, Vol.15, pp.343.
- Kobayashi, S., and Ohr, S. M., 1984, 'Dislocation arrangement in the plastic zone of propagating cracks in nickel', *J. Mater. Sci.*, Vol.19, pp.2273.
- Kocks, U. F., 1970, 'The relation between polycrystal deformation and single-crystal

deformation', *Metall. Trans.*, Vol.1, pp.1121.

Lin, I. H., and Thomson, R., 1986, 'Cleavage, dislocation emission, and shielding for cracks under general loading', *Acta Metall.*, Vol.34, pp.187.

Majumdar, B. S., and Burns, S. J., 1981, 'Crack tip shielding: an elastic theory of dislocations and dislocation arrays near a sharp crack', *Acta Metall.*, Vol.29, pp.579.

Majumdar, B. S., and Burns, S. J., 1983, 'A Griffith crack shielded by a dislocation pile-up', *Int. J. Fract.*, Vol.21, pp.229.

Mohan, R., Ortiz, M., and Shih, C. F., 1990, 'Mode mixity effects on crack tip deformation in ductile single crystals', (To be published).

Mohan, R., Ortiz, M., and Shih, C. F., 1991a, 'An analysis of cracks in ductile single crystals. I: anti-plane shear', (To be published).

Mohan, R., Ortiz, M., and Shih, C. F., 1991b, 'An analysis of cracks in ductile single crystals. II: mode I loading', (To be published).

Ohr, S. M., and Chaug, S. J., 1982, 'Dislocation-free zone model of fracture: comparison with experiment', *J. Appl. Phys.*, Vol.53, pp.5645.

Ohr, S. M., and Narayan, J., 1980, 'Electron microscope observations of shear cracks in stainless steel single crystals', *Phil. Mag.*, Vol.A41, pp.81.

Peirce, D., Asaro, R. J., and Needleman, A., 1982, 'An analysis of nonuniform and localized deformation in ductile single crystals', *Acta Metall.*, Vol.30, pp.1087.

Peirce, D., Asaro, R. J., and Needleman, A., 1983, 'Material rate dependence and localized deformation in crystalline solids', *Acta Metall.*, Vol.31, pp.1951.

Post, D., 1987, in *Handbook on experimental mechanics*, (ed. A. S. Kobayashi), Prentice Hall, Englewood Cliffs, New Jersey, Chapter 7.

Rice, J. R., 1973, 'Plane strain slip-line theory for anisotropic rigid/plastic materials', *J. Mech. Phys. Solids*, Vol.21, pp.63.

Rice, J. R., 1987, 'Tensile crack tip fields in elastic-ideally plastic crystals', *Mech. Mater.*, Vol.6, pp.317.

Rice, J. R., Hawk, D. E., and Asaro, R. J., 1990, 'Crack tip fields in ductile crystals', *Int. J. Fract.*, Vol.42, pp.301.

Rice, J. R., and Nikolic, R., 1985, 'Anti-plane shear cracks in ideally plastic crystals',

- J. Mech. Phys. Solids*, Vol.33, pp.595.
- Rice, J. R., and Saeedvafa, M., 1988, 'Crack tip singular fields in ductile single crystals with Taylor power law hardening. I: Anti-plane shear', *J. Mech. Phys. Solids*, Vol.36, pp.189.
- Rice, J. R., and Thomson, R., 1974, 'Ductile versus brittle behavior of crystals', *Phil. Mag.*, Vol.29, pp.73.
- Saeedvafa, M., and Rice, J. R., 1989, 'Crack tip singular fields in ductile single crystals with Taylor power-law hardening. II: plane strain', *J. Mech. Phys. Solids*, Vol.37, pp.673.
- Shield, T. W., Kim, K. S., and Nikolic, R., 1989, in *Micromechanics: Experimental Techniques*, AMD-V.120, (ed. W. N. Sharpe, Jr.), ASME, New York, pp.33.
- Thomson, R., and Sinclair, J. E., 1982, 'Mechanics of cracks screened by dislocations', *Acta Metall.*, Vol.30, pp.1325.
- Vitek, V., 1976, 'Yielding from a crack with finite root radius loaded in uniform tension', *J. Mech. Phys. Solids*, Vol.24, pp.67.
- Vitek, V., and Chell, G. C., 1980, in *Dislocation modeling of physical systems*, (eds. M. F. Ashby et al.), Pergamon Press, New York, pp.92.
- Weertman, J., 1978, 'Fracture mechanics: a unified view for Griffith-Irwin-Orowan cracks', *Acta Metall.*, Vol.26, pp.1731.
- Weertman, J., Lin, I. H., and Thomson, R., 1983, 'Double slip plane crack model', *Acta Metall.*, Vol.31, pp.473.
- Weertman, J., and Weertman, J. R., 1989, *Dislocation based fracture mechanics*, Springer-Verlag, Berlin, (in preparation).

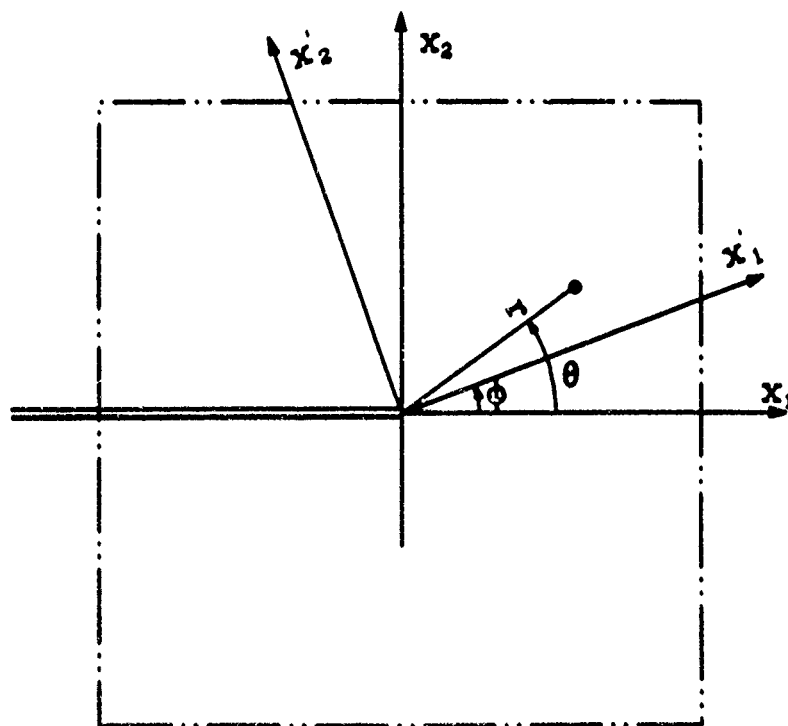


Fig. 1 Coordinate system.

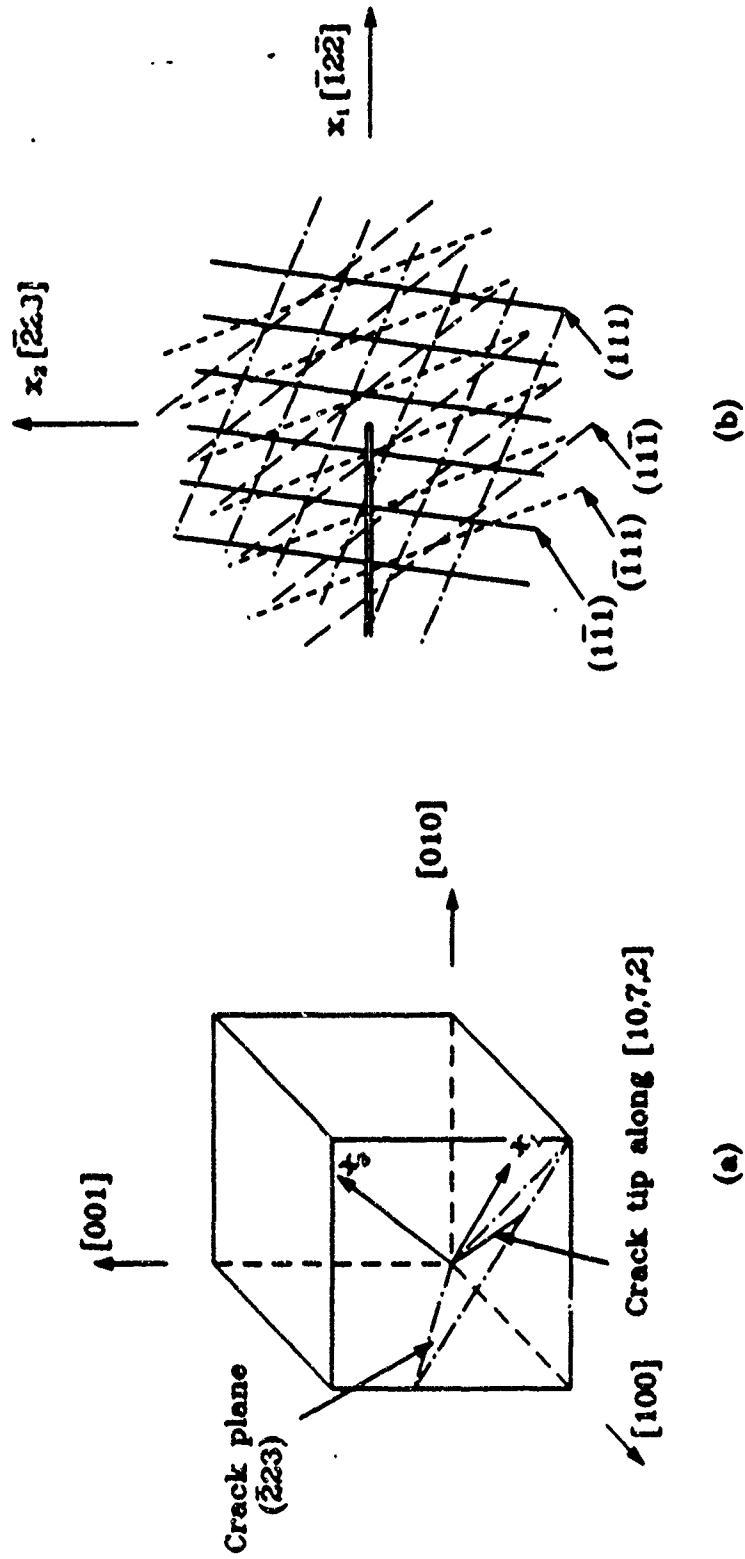


Fig. 2 Crystal orientation for aluminum single crystal under consideration.

- (a) Crack on $(\bar{2}23)$ plane and its tip along $[10,7,2]$ direction.
- (b) Crack in plane of deformation. Families of straight lines are traces of slip plane intersection with $x_3 = \text{constant}$.

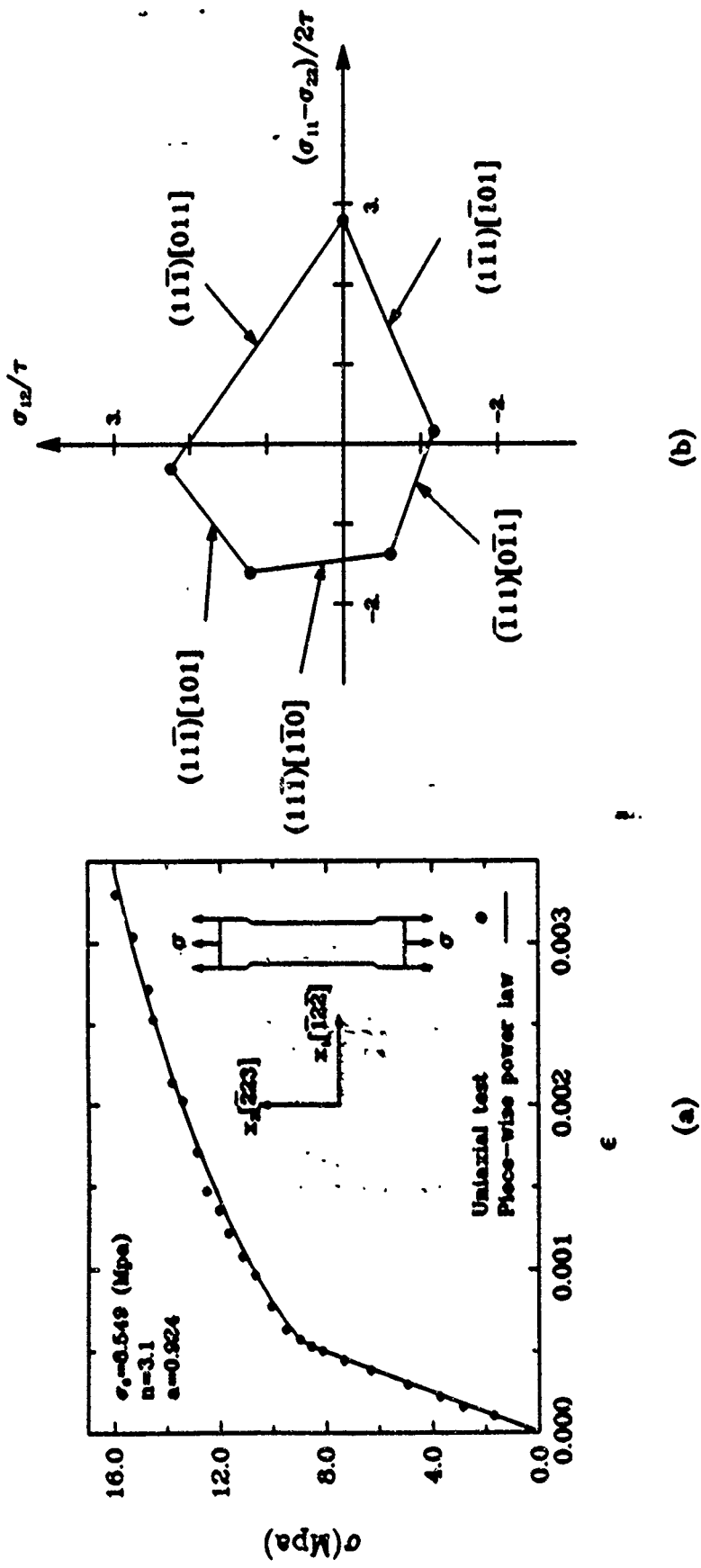


Fig. 3 (a) Stress-Strain relation of aluminum single crystal under consideration.
 (b) Yield surface and active slip systems for this orientation.

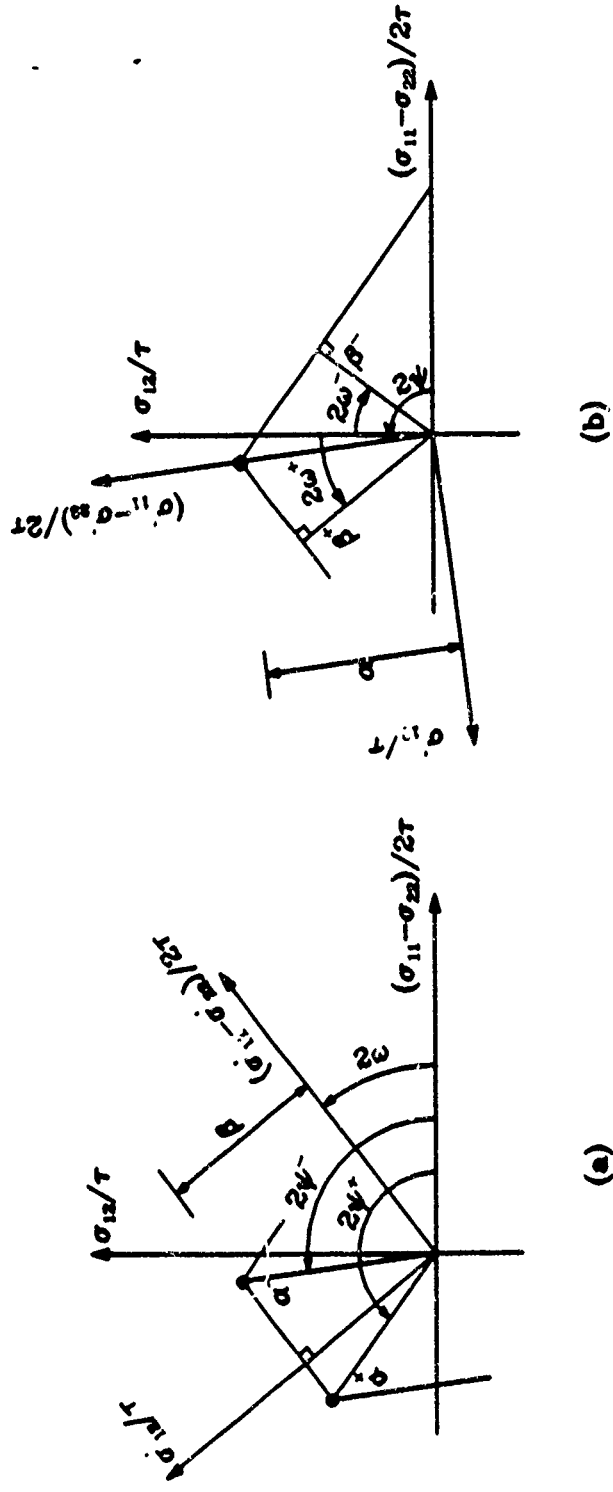
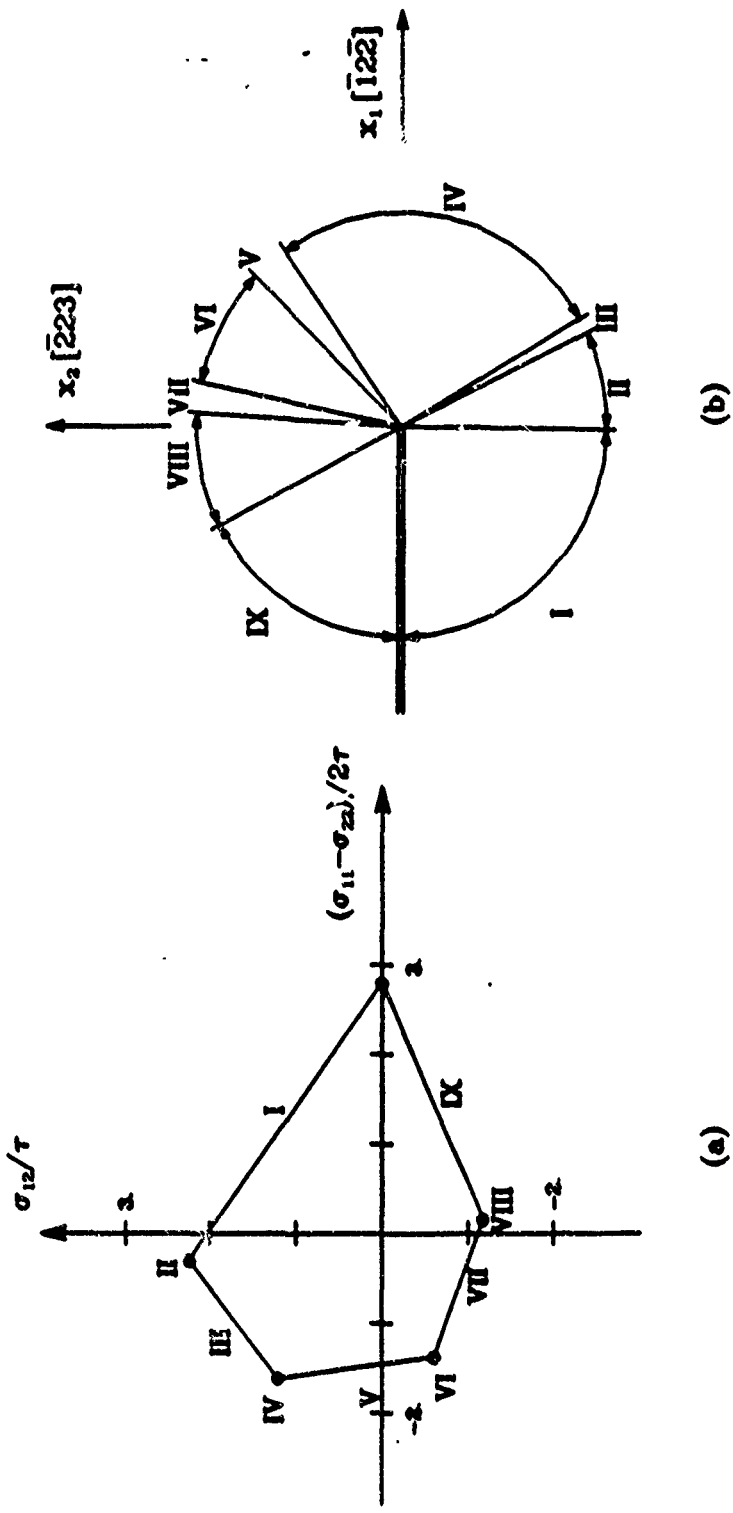


Fig. 4 (a) Coordinate system for flat sector on yield surface.
 (b) Coordinate system for vertex sector on yield surface.



(a)

(b)

Fig. 5 Angular sectors around crack tip in aluminum single crystal under consideration. (a) Sector arrangement corresponding to yield surface. (b) Sector arrangement in $x_1 - x_2$ plane.

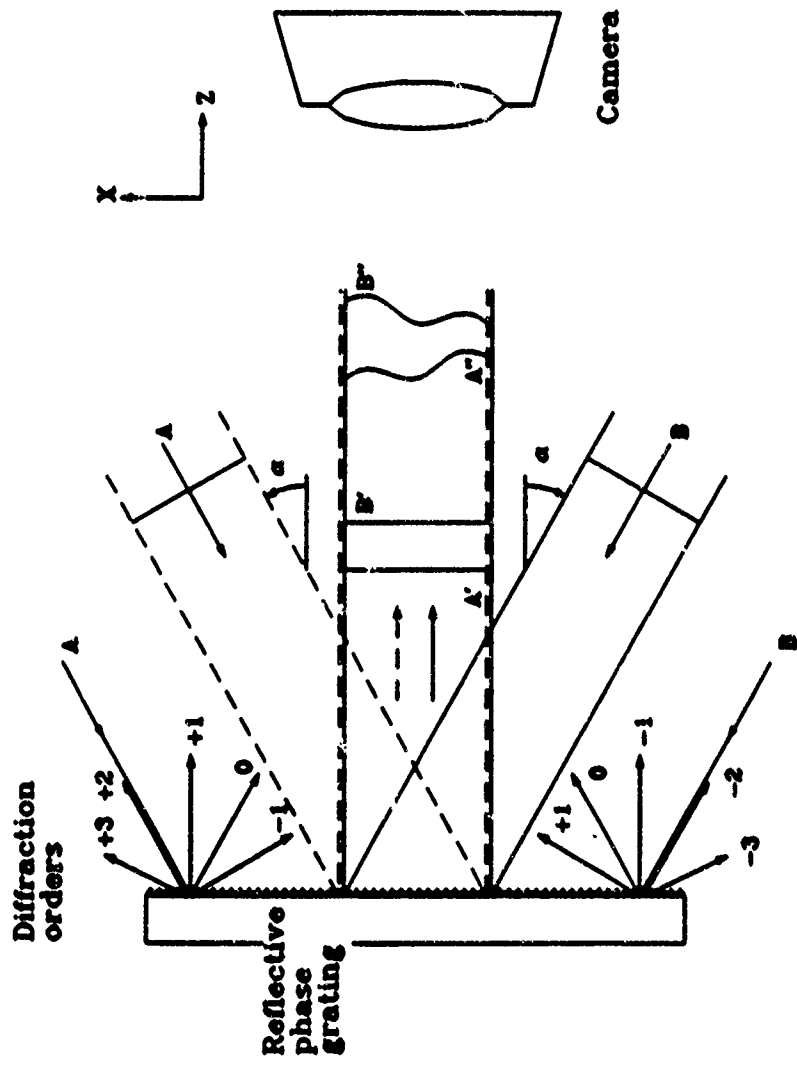


Fig. 6 Principle of moiré interferometry

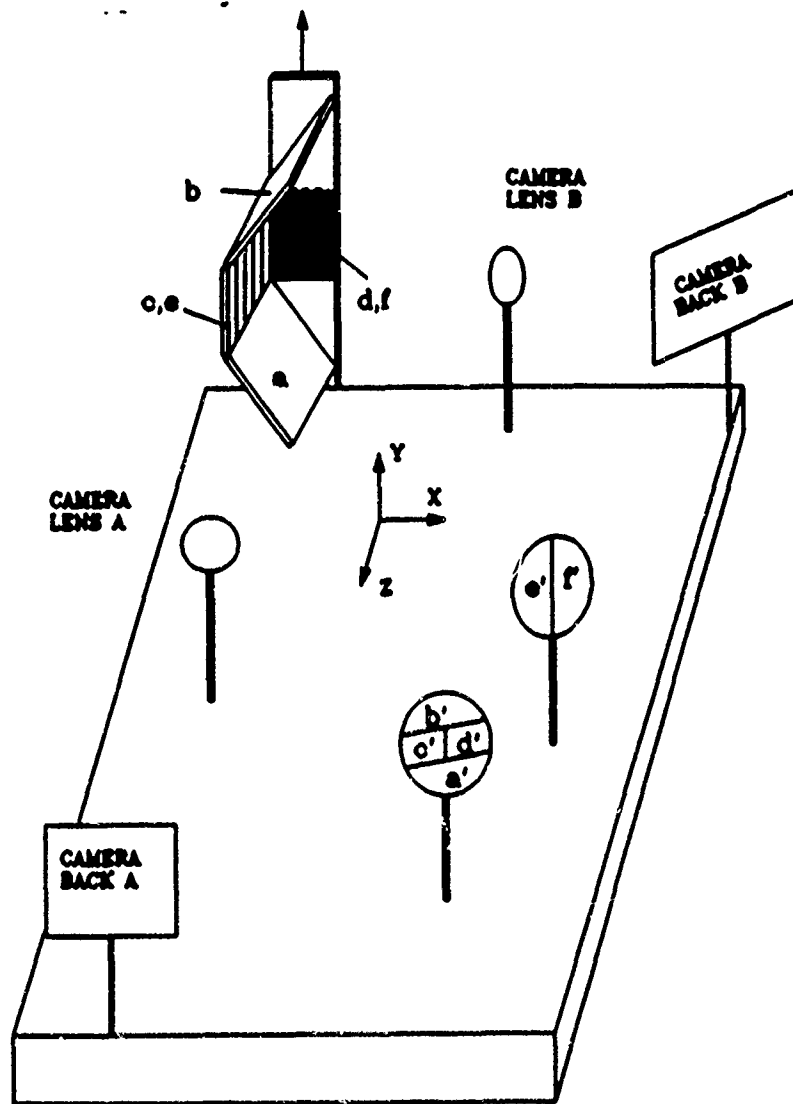
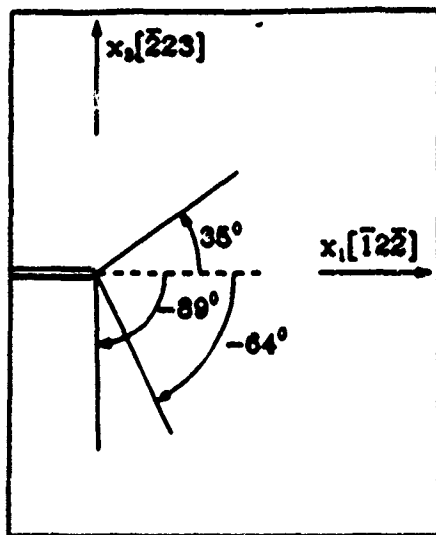


Fig. 7 Experimental setup for simultaneous measurement of u_1 , u_2 , and u_3 -displacement.



(a)



(b)



(c)



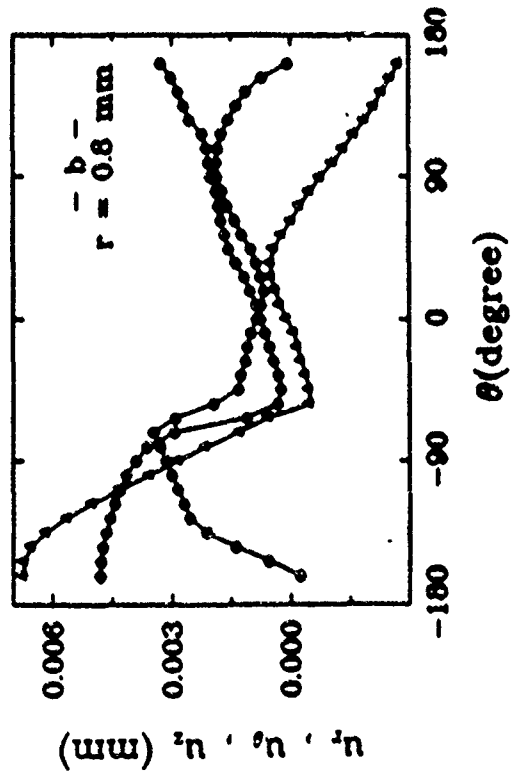
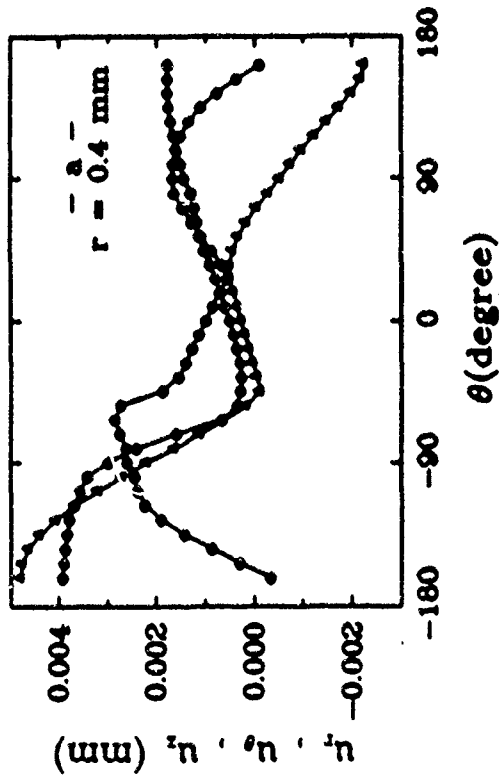
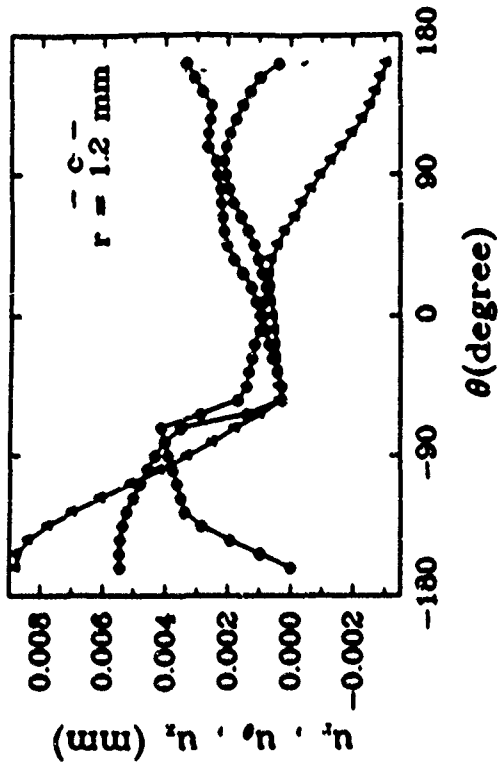
(d)

Fig. 8 Moiré fringe patterns ahead of crack tip

(a) Crack in plane of deformation

(b) u_1 -field, (c) u_2 -field, and (d) u_3 -field

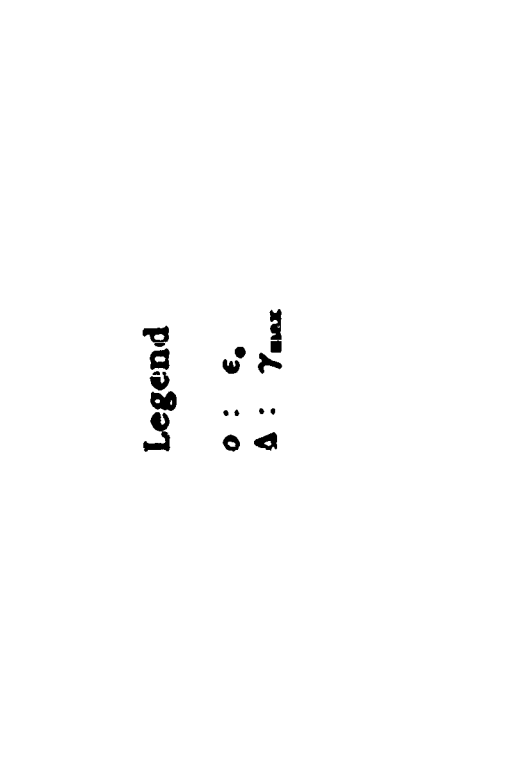
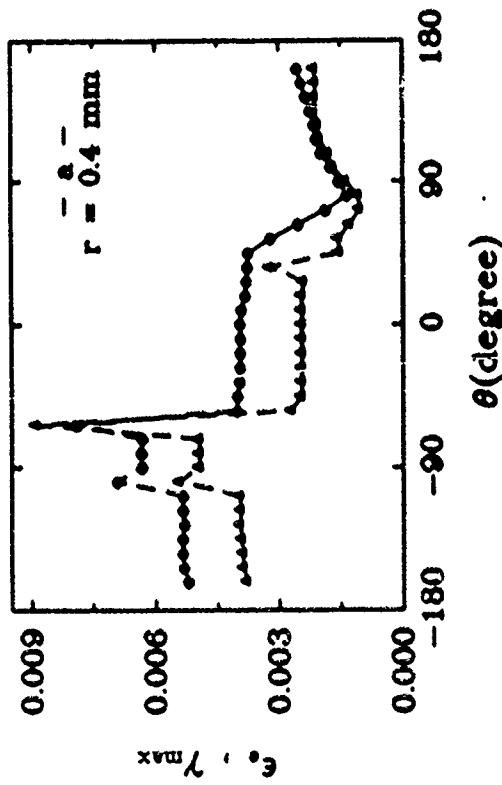
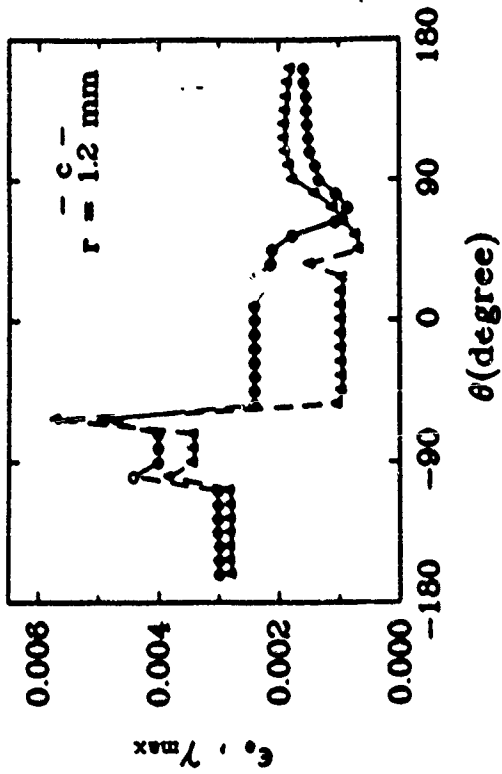
4.17 mm



Legend

- : u_t
- △: u_θ
- : u_n

Fig. 9a,b,c θ -Variation of relative displacements at different radial distance from crack tip



Legend
 o : ϵ_e
 A : γ_{max}

Fig. 10a,b,c --Variation of effective strain and maximum shear strain at different radial distance from crack tip

Material Testing by Computer Aided Speckle Interferometry

Wang, Y. Y., Chen, D. J. and Chiang, F. P.¹

A precise evaluation of material properties is always important in structural analysis and design. A conventional way to examine the stress-strain relation of a material is by means of electrical-resistance strain gage. The sensitivity of such a method is quite high (up to $1\mu\epsilon$), but its applicability is often restricted within a small deformation range ($\leq 1\%$). Although some specific treatment may increase the range substantially, it can not be used to very soft materials; for the gages mounting on the surface of the specimen will strengthen the material drastically. Geometry moire is another method for material testing[1]. With a fine grating printed on one surface of a specimen, the strengthening effect adherent to strain gage method is eliminated. However, the sensitivity of this method is usually low. For instance, if a grating with a frequency of 40 line/mm is used, a $1000\mu\epsilon$ strain can only generate one fringe over a distance of 25 mm. Therefore, a large gage-length is necessary, although the method is effective to finite deformations. Moire interferometry is a much more sensitive approach over geometry moire[2], while it requires complicated techniques and precious optical devices. In addition, both moire methods require either manual or electronic fringe pattern recognition. Manual fringe pattern analysis is tedious and time consuming and the resulting accuracy depends upon the operator's skill. Electronic fringe pattern recognition would either require the presence of many fringes, or involve dedicated phase-shifting devices.

A new technique named computer aided speckle interferometry (CASI) has been developed recently [3,4]. The technique, being non-destructive and remote-sensing, has achieved full automation in both the speckle pattern registration and information extraction. The system constitutes simple and reliable optical as well as electrical set-up. It is effective in a large deformation range, and

¹Wang, Y. Y. and Chen, D. J. (SEM Member) are graduate students, Department of Mechanical Engineering and Chiang, F. P. (SEM Fellow) is Director, Laboratory for Experimental Mechanics Research, SUNY at Stony Brook, NY

affords a complete survey of 2-D displacement fields. In its basic form, two speckle patterns of a specimen, one before and one after the deformation, are captured by a CCD (charged-coupling device) camera and registered in a computer. The U and V displacement distributions between the two speckle patterns are analyzed pointwise at various locations using a two-step fast Fourier transform (FFT); and in-plane strains are obtained by numerical differentiations on the resulting U and V displacement fields. Although the measurable strain is limited within 5% in a single deformation step[5], substantially higher finite strains can be inspected by registering more speckle patterns during a deformation process and analyzing the incremental strain between each pair of successive speckle patterns.

In this work, we investigate the applicability of CASI to material testing. Two types of photoelastic materials are evaluated. The first one, being a rigid material, is tested in its elastic range (up to $8000\mu\epsilon$). While the whole range is also monitored by strain gages, our aim here is to verify the accuracy of the CASI by the strain gage results. The second specimen, being a ductile one, is inspected in a large strain range (up to 20%), where strain gages fail to work. It is the second specimen that demonstrates the advantages of CASI.

Experimental Procedures

1. Treatment of specimen.

The specimen configuration is depicted in Fig.1. The testing area of the specimen is cleaned and sprayed alternately with black and white paints. The thickness of the paint is made negligibly thin and the tiny spots (speckles) uniformly distributed. As an example, one enlarged copy of the registered speckle pattern is demonstrated in Fig.2. A two-element rosette is mounted on the back surface of the first specimen which has a higher modulus. The inspected area by the camera is about $20\text{mm} \times 20\text{mm}$ for both specimens, while the dimension of the actual area is trivial for strain evaluations.

2. Experimental set-up.

Fig.3 illustrates the experimental set-up. The specimen is illuminated by a white light source and loaded in simple tension by a Tinius Olsen testing machine. The speckle patterns of the specimen are captured by a CCD camera and digitized by a DT-2861 frame grabber. The digital images are stored in an IBM-AT personal computer. Each specimen is loaded to several deformation steps. One initial image is recorded before loading and more images are recorded after each successive loading step. The data processing is performed by the computer.

3. Data processing.

The data processing is done by two steps. In the first step, the displacement fields (U and V) are obtained by analyzing the two speckle patterns corresponding to certain deformation levels[3,4]. In the second step, the longitudinal and the lateral strains (ϵ_{yy} and ϵ_{xx}) are deduced from the resulting displacement fields. The U and V fields acquired by CASI are two matrixes with 16×16 elements, in which each element stands for the displacement component at one point. The distance between any two adjacent points of the array is 32 pixels. Since the tested specimens are under uniaxial tension, the U displacement at a vertical cross section within the inspected area can be regarded as uniformly distributed and so as the V displacement at a horizontal cross section. Thus, a resultant U (V) displacement distribution along x (y) axis is obtained by averaging all columns (rows) of the U (V) field matrix. The strains ϵ_{xx} and ϵ_{yy} are finally determined by finding out the slopes of the $U(x)$ and $V(y)$ distributions using a least square fitting. It should be noted that since both displacements (U and V) and distances (x and y) are estimated by pixels, dimensional measurement on the detected area is not needed in the strain evaluation.

Since the ductile specimen undergoes much larger deformation than the rigid one, different procedures of data processing are adopted for the two specimens. For the rigid specimen, the U and V displacement fields at each loading level are obtained respectively by comparing the

speckle pattern at that loading level with the initial one. The deduced strains are, therefore, equivalent to the engineering strains at that loading level. For the ductile specimen, the relative displacement fields between each two successive loading levels are first investigated by processing the two corresponding speckle patterns. Thus the resulting strains are merely the incremental strains during that loading stage. Since the incremental strains at each step are small compared to the total strain, the effective true strains at a certain loading level can be obtained by summing all the incremental strains before that level, i.e.,

$$\epsilon_j^T = \sum_{i=1}^j \Delta \epsilon_i \quad (1)$$

where $\Delta \epsilon_i$ is the incremental strain during the i th loading step, and ϵ_j^T the true strain after the j th loading level. The engineering strain can be simply converted from the true strain by,

$$\epsilon_j = \exp(\epsilon_j^T) - 1. \quad (2)$$

The results reported in the following section are engineering strains.

Results and Conclusions

The estimated stress and strain relations of both specimens are illustrated in Fig.4. For the first specimen, the data obtained by strain gage method is also plotted as a comparison. The standard deviations of the strains evaluated by CASI from those by the strain gages are $100\mu\epsilon$ for normal strain and $59\mu\epsilon$ for lateral strain(Fig.4a). The Young's modulus and the Poisson's ratio measured by both methods are listed in Table 1. It can be seen from the table that the relative differences between the two sets of results are 4.0% for the Young's modulus and 2.5% for the Poisson's ratio. The differences are acceptable in a sense of engineering practice. The second specimen has a non-linear stress-strain relation(Fig.4b). The experimental data are fitted by the least square method using a power law as $\epsilon = \alpha\sigma^n$, where α and n are estimated as 3.96 and 1.39. An approximate proportional stress-strain relation is found under 1% strain. In the approximate linear range, the elastic modulus is 0.039 GPa and the Poisson's ratio is 0.44.

It is worthwhile to mention that the accuracy of the strain evaluation does not depend upon the dimension of the inspected area. The size of the detected area on the specimen is selected based on the size of the speckles. In fact, there is no critical size for the detected area. According to the optimal sampling criteria of the speckle pattern[5], the sampling interval or pixel size is arranged to be about one-third of the speckle size in the image plane. When larger speckles are built on the surface of the specimen, a smaller optical magnification is selected and a larger area is monitored; when smaller speckles, such as laser speckles, are used, a larger optical magnification is employed and a smaller area is monitored. The uncertainty of the measured displacement components is proportional to the speckle size, however, the uncertainty of the evaluated strains is independent of the speckle size. This is because the size of the effective gage-length is also proportional to the speckle size at the optimal sampling rate.

CASI is found to be a good method for measuring finite deformation. The uncertainty of the measured strain range is under $100 \mu\epsilon$. The method retains the non-contact feature of conventional optical speckle interferometry, it is suitable for inspecting stress and strain relations of low modulus materials. The superiority of it over the conventional optical speckle method is that the manual film processing and fringe pattern analyzing are not in requisition.

Acknowledgments

This work was supported by the Army Research Office through Contract DAAL20388K0083 and the office of Naval Research through Contract N0001482K0566. The supports are gratefully acknowledged.

References

- [1] F. P. Chiang, "Moire methods of strain analysis," Chap. VI, *Manual of Experimental Stress Analysis*, 3rd ed., A. S. Kobayashi, Ed., Society for Experimental Stress Analysis, Brookfield Center, CONN 1978
- [2] D. Post, "Moire interferometry," Chap.7, *Handbook of Experimental Mechanics*, A. S. Kobayashi, Ed., Prentice Hall, Englewood Cliffs, NJ 1987
- [3] D. J. Chen and F. P. Chiang, "Computer speckle interferometry," Proc. of Int. Conf. on Hologram Interferometry and Speckle Metrology, pp. 49-58 (1990)
- [4] D. J. Chen, F. P. Chiang, Y. S. Tan and H.S. Don, "Computer aided speckle interferometry (CASI): II. An alternate approach using spectral amplitude and phase information," Proc. SPIE, Vol. 1554A, pp.706-717 (1991)
- [5] D. J. Chen, S. Li, T. Y. Hsu and F. P. Chiang, "Range of measurement of computer aided speckle interferometry," Proc. SPIE, Vol. 1554A, pp.922-931 (1991)

Captions

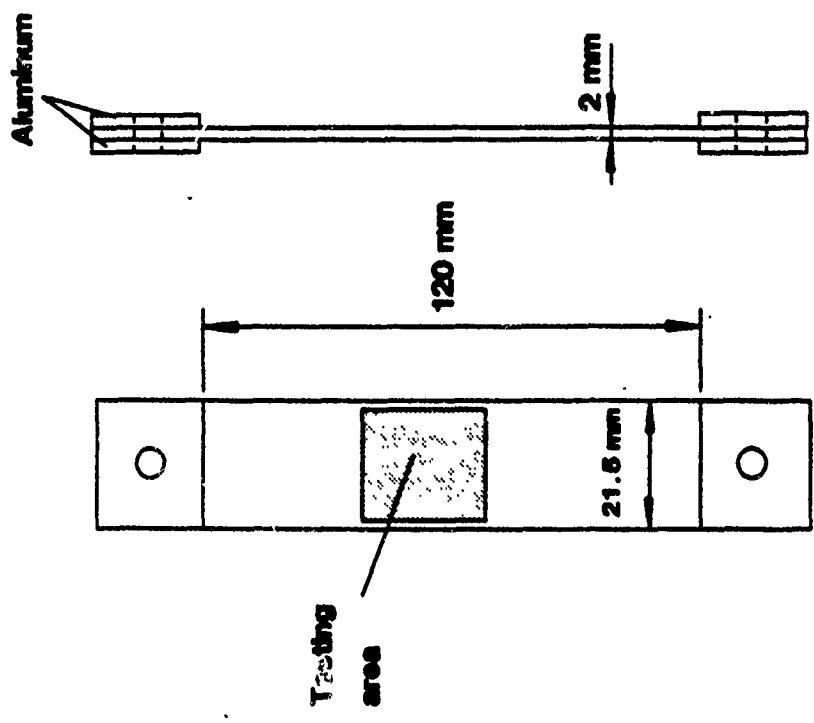
Table 1 Estimated material constants of specimen No.1.

Fig.1 Specimen configuration.

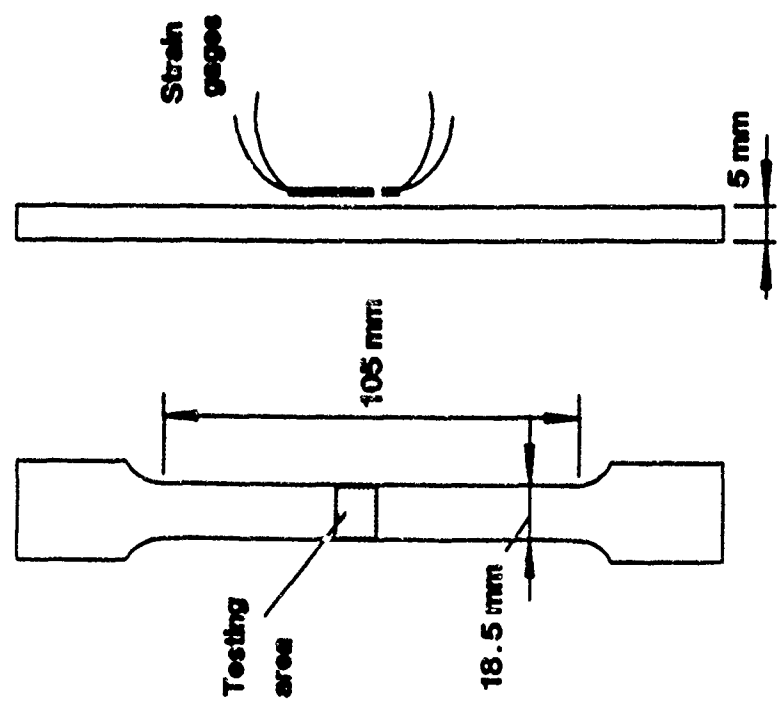
Fig.2 Registered speckle pattern.

Fig.3 Schematic of data acquisition and image processing system.

Fig.4 Material properties measured by CASI and strain gages.



(b) Specimen 2



(a) Specimen 1

Fig. 1 Specimen configuration

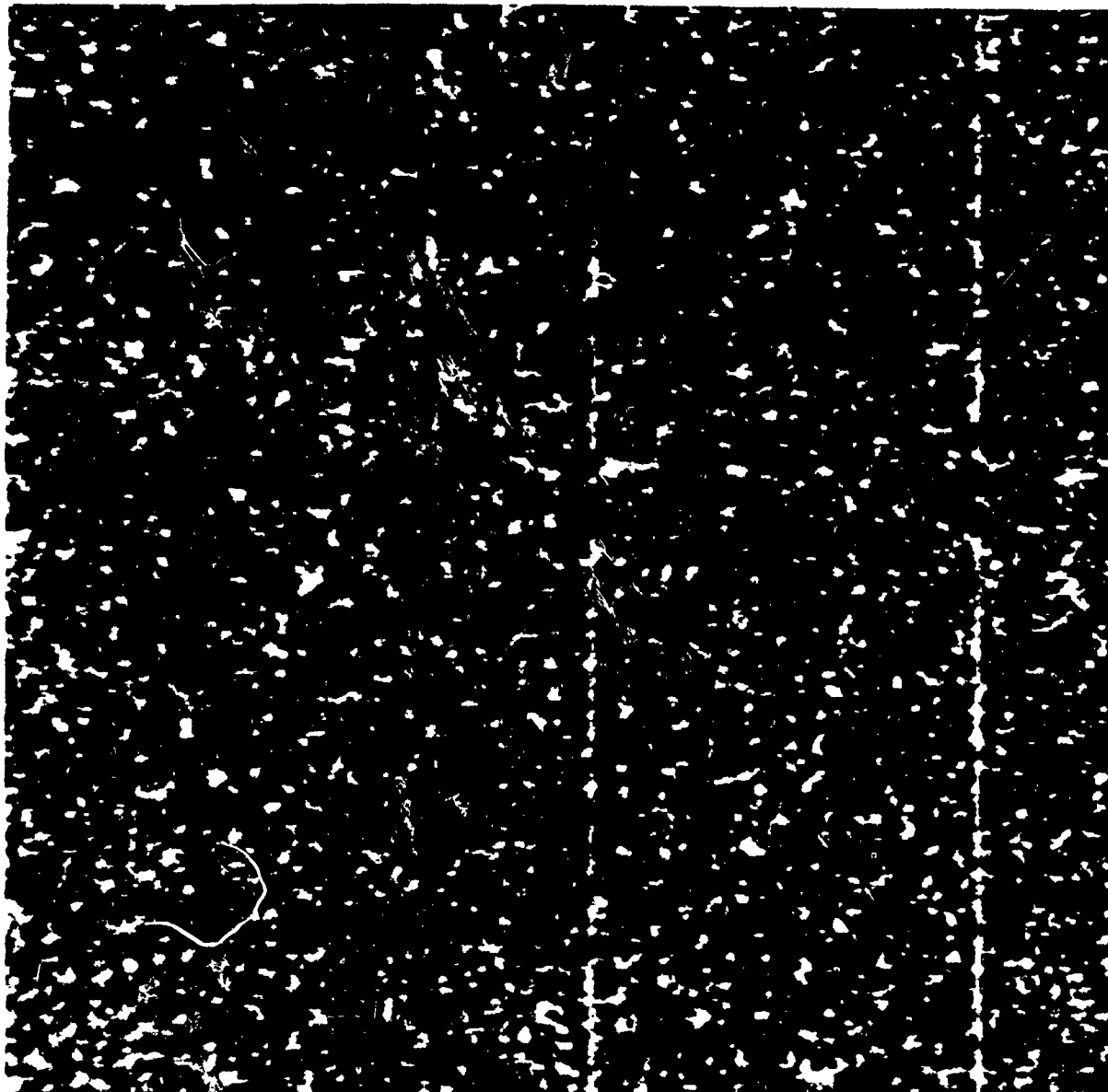


Fig . 2 Registerd speckle pattern

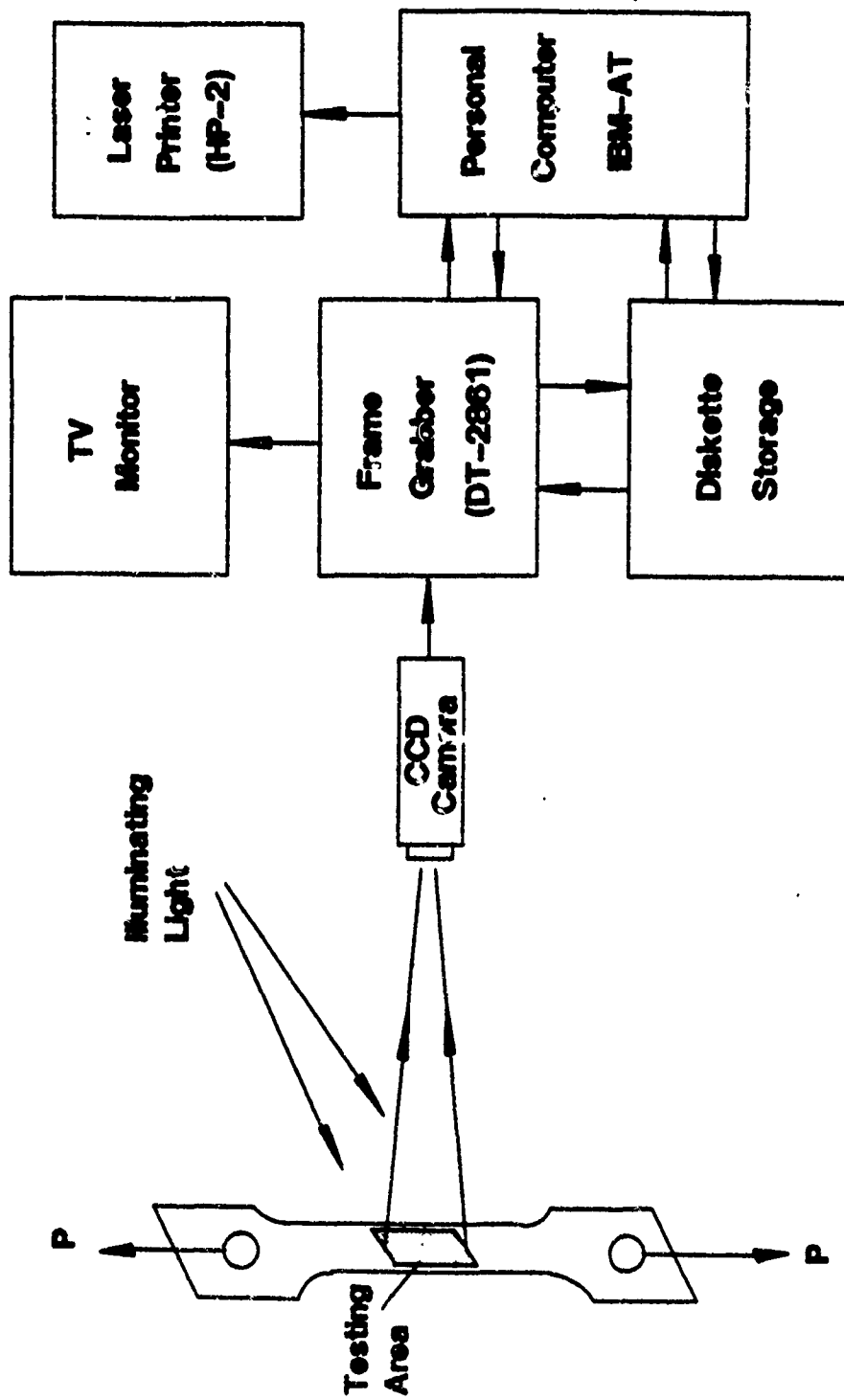
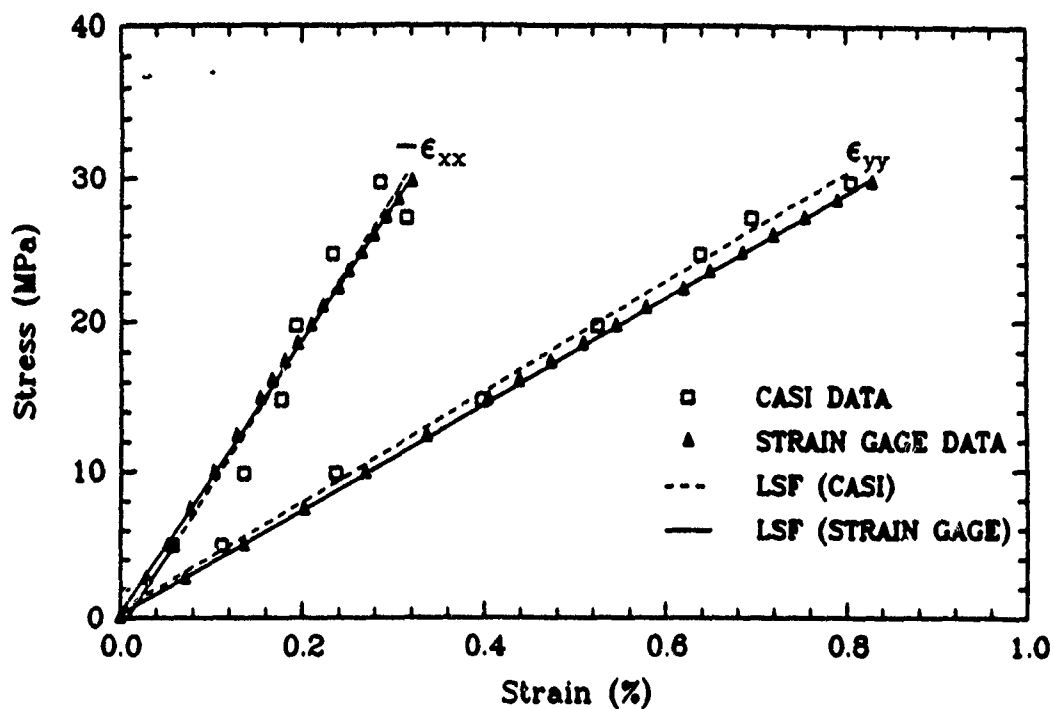
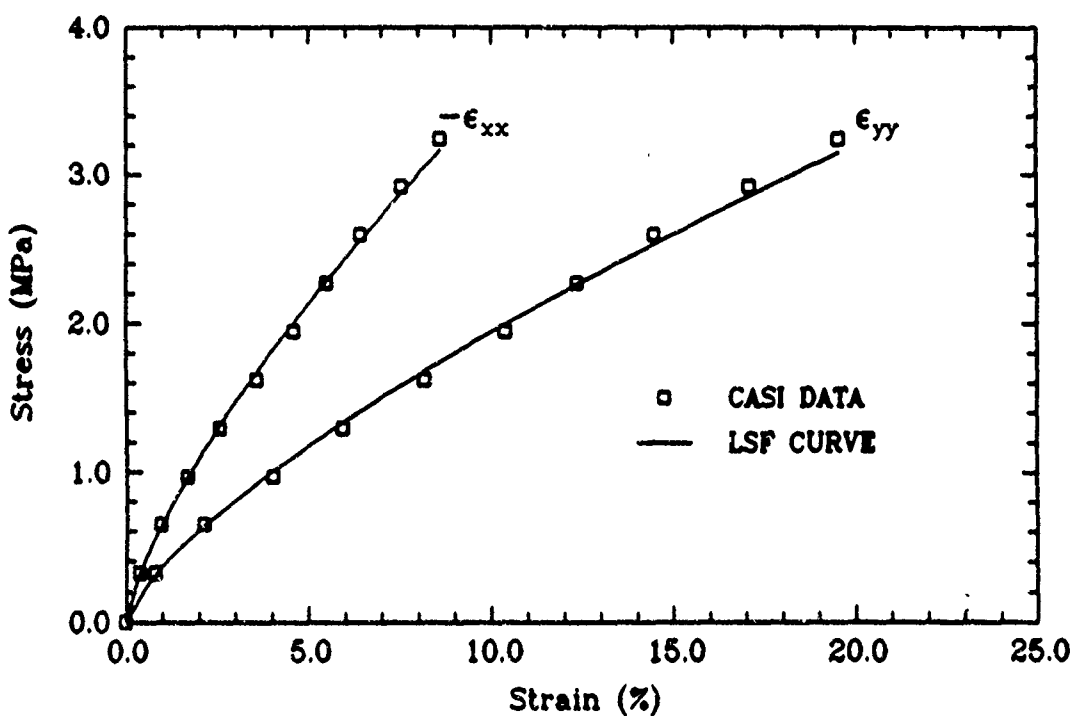


Fig. 3 Schematic of data acquisition and image processing system .



(a) Stress-strain relation of specimen No.1



(b) Stress-strain relation of specimen No.2

Fig.4 Material properties measured by CASI and strain gages.