Center for Research and Education in Optics and Lasers

CREOL



FINAL REPORT TO THE BALLISTIC MISSILE DEFENSE ORGANIZATION AND THE OFFICE OF NAVAL RESEARCH

"CREOL SUPPORT TO ISTEF IV - COHERENT LASER RADAR"

ONR Contract No. N00014-92-C-0123

December 1996

19961223 002

DTIC QUALITY INSPECTED 4

DISTRIBUTION STATEMENT A
Approved for public release
Distribution Unlimited

University of Central Florida
4000 Central Florida Blvd.
Orlando, FL 32816-2700

REPORT DOCUMENTATION PAGE

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 the collection of information, including suggestions for reducing the 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	3. REPORT TYPE AND DATES COVERED
	15 Dec 1996	Final; July 1992 - January 1996

4. TITLE AND SUBTITLE OF REPORT
"CREOL Support to ISTEF IV - Coherent Laser Radar"

5. FUNDING NUMBERS
C:N00014-92-C-0123

6. AUTHOR(S)
C. M. Stickley, A. R. Weeks, P. Gatt, J. Harvey, H. Myler

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Center for Research and Education in Optics and Lasers (CREOL)
Univeristy of Central Florida
P.O. Box 162700
4000 Central Florida Blvd.
Orlando, FL 32816-2700

8. PERFORMING ORGANIZATION REPORT NUMBER:
96-05

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Ballistic Missile Defense Org./ Innovative Science and Technology Directorate
Office of Naval Research
800 N. Quincy Street
Arlington, VA 22217-5660

10. SPONSORING/MONITORING AGENCY REPORT NUMBER:

11. SUPPLEMENTARY NOTES:

12a. DISTRIBUTION AVAILABILITY STATEMENT
Unclassified/Unlimited

12b. DISTRIBUTION CODE

13. ABSTRACT (Maximum 200 words)

This report summarizes achievement in the following areas: • We designed, built, and successfully tested in the field (at BMDO's ISTEF) a coherent laser receiver which demonstrated a M-fold CNR gain and 1/M reduction in signal fading probability, where M is the number of independent apertures/coherent detectors. • We produced a design for such a receiver having eight apertures/detectors. • We developed software, including a graphical user interface, for simulation of lidar in x, y, and t; the software is available from Dr. A. Weeks. • We compared coherent CO_2 and YAG lidars against the same rotating target; the major difference was the presence of glints with the CO_2 system which tended to overshadow rotation rate information. • An eye-safe Doppler lidar was built for vibration sensing. • A technique for tracking thru turbulence was developed in which the resolution is not affected by the turbulence. • Two image reconstruction algorithms were developed which, relative to other algorithms, avoid stagnation, improve global convergence, and dramatically reduce computational complexity. • The effects of reciprocal path scattering on coherent lidars from turbulence and target roughness is assessed.

14. SUBJECT TERMS Coherent lidar; multiple apertures; multiple aperture lidar; lidar simulation; speckle effects; glints; eye-safe lidar; tracking; image reconstruction; reciprocal path scattering; enhanced backscatter

15. NUMBER OF PAGES: 360

16. PRICE CODE

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

Center for Research and Education in Optics and Lasers

CREOL



FINAL REPORT TO THE BALLISTIC MISSILE DEFENSE ORGANIZATION AND THE OFFICE OF NAVAL RESEARCH

"CREOL SUPPORT TO ISTEF IV - COHERENT LASER RADAR"

ONR Contract No. N00014-92-C-0123

December 1996

University of Central Florida
4000 Central Florida Blvd.
Orlando, FL 32816-2700

TABLE OF CONTENTS

1.0 INTRODUCTION AND SUMMARY 1-1

2.0 COHERENT LASER RADAR ARRAY TECHNOLOGY AND SYSTEMS 2-1
- 2.1 Background 2-1
- 2.2 Accomplishments 2-1
- 2.2.1 Analysis of Space Diversity Receivers for Coherent Array Laser Radar Systems 2-6
- 2.2.1.1 Equal Gain, Maximal Ratio and Selection Diversity Receivers 2-7
- 2.2.1.2 Mean CNR Improvement 2-8
- 2.2.1.3 Carrier Fading Reduction Factors 2-9
- 2.2.1.4 Effect of Signal Correlation Between Receiver Apertures 2-10
- 2.2.2 Dual Aperture Coherent Array Receiver Design and Development 2-10
- 2.2.3 The Design and Development of an Eight Aperture Coherent Array Receiver 2-15
- 2.2.4 Low-cost InGaAs Photoreceiver Design 2-33
- 2.2.5 Heterodyne Efficiency: "Where are the Missing dBs?" 2-34
- 2.2.5.1 Theoretical Background 2-34
- 2.2.5.2 Experiment Design 2-35
- 2.2.5.3 Results and Discussion 2-36
- 2.2.6 Field Experimental Test of Two-Aperture System 2-37
- 2.2.6.1 Experiment and Test Design and Layout 2-37
- 2.2.6.2 Results 2-37
- 2.3 Conclusions 2-38
- 2.4 References 2-39
- 2.5 Publications 2-40
- 2.6 Appendix 2.A - P. Gatt, T. Costello, and D. Gantala, "Dual Aperture Measurements of Two km (round trip) Amplitude and Phase Perturbations of One Micron Laser Speckle Propagating Through a Turbulent Atmosphere ", Proc. SPIE 2471, 254-258, (1995) 2-41
- 2.7 Appendix 2.B - P. Gatt, T. P. Costello, D. A. Heimmermann, D. C. Castellanos, A. R. Weeks, and C. M. Stickley, "Coherent Optical Array Receivers for the Mitigation of Atmospheric Turbulence and Speckle Effects", Appl. Opt., 35, 5999-6009, (October 20, 1996). 2-42

3.0 TARGET-INDUCED SPECKLE EFFECTS IN LIDAR APPLICATIONS 3-1
- 3.1 Background 3-1
- 3.1.1 Fully Developed Speckle (FDS) 3-3
- 3.1.2 Partially Developed Speckle (PDS) 3-5
- 3.2 Accomplishments 3-8
- 3.2.1 Target Characteristics 3-8
- 3.2.2 Experimental Measurements 3-11
- 3.2.2.1 BRDF Measurements 3-11
- 3.2.2.2 Speckle Measurements 3-13
- 3.2.3 Effects of Error Sources upon Measured Speckle PDF's 3-17
- 3.2.4 Mean Speckle Size as a Function of Intensity Threshold Level 3-20
- 3.3 Conclusion 3-25
- 3.4 References 3-26
- 3.5 Publications, Presentations and Theses 3-27
- 3.6 Appendix 3A - T. L. Alexander, J. E. Harvey and A. R. Weeks, "Average Speckle Size as a Function of Threshold Level: Comparison of Experimental Measurements with Theory", Appl. Opt. **33**, 8240-8250 (10 Dec 1994) 3-28

3.7 Appendix 3B T. L. Alexander, J. E. Harvey, and D. Hefele, "Experimental Verification of a Theoretical Model for Speckle Intensity Excursion Areas", Proc. SPIE **2222-63** (1994) 3-29

4.0 LIDAR SIMULATION 4-1
4.1 Background 4-1
4.2 Accomplishments 4-6
4.2.1 The Models for the Optical Components used within the Laser Radar Simulator 4-9
4.2.1.1 Free Space Propagation and Lens Theory 4-9
4.2.1.2 Detector 4-12
4.2.1.3 Phase Screen Theory 4-17
4.2.1.4 Laser 4-20
4.2.1.5 Wave Plates 4-23
4.2.1.6 Beamsplitters 4-27
4.2.3 The Graphical User Interface (GUI) for the Laser Radar Simulator 4-28
4.2.3.1 Menu Section 4-29
4.2.3.2 Simulation Menu 4-30
4.2.3.3 Simulation Parameters Window 4-30
4.2.3.4 Aperture Specification 4-31
4.2.3.6 Acousto-optic Modulator 4-33
4.2.3.7 Beamsplitter (polarizing) 4-34
4.2.3.8 Beamsplitter (non-polarizing) 4-35
4.2.3.9 Detector 4-35
4.2.3.10 Lens 4-37
4.2.3.11 Mirror 4-38
4.2.3.12 Phase Screen 4-38
4.2.3.13 Point Detector 4-39
4.2.3.14 Signal Splitter 4-40
4.2.3.15 Target 4-40
4.2.3.16 Wave Plate 4-41
4.2.3.17 Schematic Section 4-42
4.2.3.18 Propagation Line Characteristics 4-43
4.2.3.19 Displaying Data 4-43
4.2.3.20 Example Simulation of a Monostatic Laser Radar 4-47
4.2.4 Experiments Performed with Laser Radar Simulator 4-54
4.2.4.1 Speckle Simulations 4-54
4.2.4.2 Simulation of Partially Developed Speckle 4-58
4.2.4.3 Experiments Demonstrating Evolution of Speckle with Propagation Distance 4-60
4.2.4.4 Gaussian Moments Experiments 4-61
4.2.4.5 Demonstration of Multiple Aperture Laser Radar Simulation 4-67
4.3 Conclusions 4-75
4.4 References 4-75
4.5 Publications, Presentations and Theses 4-76
4.6 Appendix 4A - K. J. Gamble, A. R. Weeks, H. R. Myler, and W. A. Rabadi, "Results of Two-dimensional Time-evolved Phase Screen Computer Simulations", Proc. SPIE, 2471, 170-180, June 1995 4-78

5.0 SINGLE APERTURE COHERENT LIDAR EXPERIMENTS 5-1
5.1 Background 5-1
5.2 Accomplishments 5-1
5.2.1 Rotating Target Experiments with YAG and CO_2 Coherent Lasers 5-1
5.2.2 Vibration Sensing with a CW Coherent Lidar 5-3
5.3 References 5-4

5.4 Publications, Presentations and Theses 5-5
5.5 Appendix 5A - P. Gatt, T. Costello, C. Vogt, and C. M. Stickley, "Laser radar spectrum of a rotating random rough surface", Proc. SPIE 1936, 165-179, (1993). 5-6
5.6 Appendix 5B - M. Acharekar, P. Gatt and L. Mizerka, "Laser Vibration Sensor", Proc. SPIE 2472, 2-11 (1995) 5-7

6.0 RECIPROCAL PATH SCATTERING THROUGH RANDOM MEDIA 6-1
6.1 Background 6-1
6.2 Accomplishments 6-3
6.2.1 Reciprocal Path Imaging (RPI) of Extended Objects 6-4
6.2.2 Interferometric Sensitivity in Pointing and Tracking Accuracy 6-7
6.3 Conclusions 6-11
6.4 References 6-12
6.5 Publications, Presentations and Theses 6-13
6.6 Appendix 6A - J. E. Harvey and A. Kotha, "Sparse Array Configurations Yielding Uniform MTF's in Reciprocal Path Imaging Configurations", Opt. Com. **106**, 178-182 (March 1994 6-15
6.7 Appendix 6B - J. E. Harvey, A. Kotha and R. L. Phillips, "Image Characteristics in Applications Utilizing Dilute Subaperture Arrays , Appl. Opt. **34**, 2983-2992 (1995) 6-16
6.8 Appendix 6C - J. E. Harvey, S. P. Reddy, and R. L. Phillips, "Precision Pointing and Tracking through Random Media by Exploiting the Enhanced Backscatter Phenomenon", Appl. Opt. **35**, 4220-4228 (1996 6-17

7.0 IMAGE RECONSTUCTION 7-1
7.1 Background 7-1
7.2 Accomplishments 7-2
7.2.1 Multiresolution Error Reduction (MRER) Approach to Image Reconstruction 7-2
7.2.2 Wavelet Decomposition Approach to Image Reconstruction 7-11
7.3 Conclusions 7-20
7.4 References 7-22
7.5 Publications, Presentations and Theses 7-23
7.6 Appendix 7A - Rabadi W.A., H. R. Myler and A. R. Weeks, "An iterative multiresolution algorithm for image reconstruction from the magnitude of its Fourier transform," *Optical Engineering*, 35:4, (1996) 7-25
7.7 Appendix 7B - Rabadi W. A., H. R. Myler, A. R. Weeks, K. J. Gamble, "Pyramid framework for image reconstruction from nonimaged laser speckle" Proc. SPIE **2484**, (1995) 7-26

8.0 RECIPROCAL PATH SCATTERING AND LASER RADAR 8-1
8.1 Background - What is RPS? 8-1
8.2 Accomplishments 8-2
8.3 Conclusions 8-3
8.4 References 8-3
8.5 Publications, Presentations and Theses 8-4
8.6 Appendix 8A - R. A. Murphy, P. Gatt, and R. L. Phillips, "Reciprocal Path Scattering from Surfaces and Turbulence in Monostatic Ladar - Theory", Proc. SPIE 1968, 650-661 (1993) 8-5
8.7 Appendix 8B - R. A. Murphy and R. L. Phillips, "Atmospheric Turbulence Reciprocal Path Scattering Effects for a Ground-Based Monostatic Laser Radar Tracking a Space Target", Proc. SPIE 2828, 52, (1996). 8-6

1.0 INTRODUCTION AND SUMMARY

C. Martin Stickley, Ph.D.

This is the final report to the Office of Naval Research on contract N00014-92-C-0123 which covered the period of 1 July 1992 - 30 January 1996. CREOL has conducted research in the area of laser radar since about 1987 when we began such research as a way to provide direct and indirect support to BMDO's Innovative Science and Technology Experimentation Facility (ISTEF) located at Kennedy Space Center. Our efforts to provide such support have been summarized in a number of final reports to the government:

- "Study and Measurements of Turbulence on Optical Systems Operating Through the Atmosphere", Annual Report for 1988.
- "Final Report to SDIO/Office of Naval Research", Contract No. N00014-89-K-0125; July 1991
- "CREOL Support to ISTEF II", final report to BMDO/ONR on Contract No. N00014-90-J-4084; March 1993
- "CREOL Support to ISTEF III", final report to BMDO/ONR on Contract No. N00014-91-C-0227; August 1993.

At the outset of this fourth contract, ISTEF was growing and maturing as a research facility, and therefore, while the title of this program was 'CREOL Support to ISTEF IV', less activity was directed to providing such 'support' to ISTEF than in the past. Instead, the thrust of the work was to develop technology much like any other university contractor, i.e. technology which would be useful to the future mission of BMDO.

Thus, the main goal of our program became one of developing an understanding of a number of basic issues dealing with coherent laser radars: how to use multiple apertures to increase lidar performance and how to design such systems; how to deal with partially and fully developed speckle in coherent laser radar; the effect of reciprocal path scattering on coherent lidar; how to reconstruct images of a target when one has only information on the Fourier transform magnitude of illuminated object; and how can large signal losses, which are often observed in coherent lidar systems, be explained? Addressing the latter issue stimulated us to develop software (including a graphical user interface) for simulation of a lidar system, and to undertake measurements of heterodyne inefficiency. In the course of this program we also built a two aperture coherent lidar to test our theories and design approach for a multiple aperture coherent laser receiver, and we developed a concept for high precision target tracking through atmospheric turbulence.

We trained and graduated a number of students, most of whom followed a thesis-option. The names of these students, where they are employed and their thesis titles are included in Table 1.1.

The following will be a summary of the content and conclusions of the chapters of this report; the chapter topics were areas of prime focus for the program. Each chapter contains reproduced copies from journals of relevant publications produced in the course of this research program.

Chapter 2, Coherent Laser Radar Array Technology and Systems, covers much of the "mainstream" of this research program, namely, the design, development and test of a prototype of the eight aperture system, now being evaluated at ISTEF. Dr. Philip Gatt led and performed much of this work, starting with an analysis of space

Table 1.1: Students Trained and Graduated

Name	Employed At	Thesis Title (if any)
Dr. Philip Gatt	Coherent Technology, Inc., Boulder, CO	"A Simulation and Theoretical Analysis of the Doubly Stochastic Homodyne K Scattering Model"
Ken Jerkatis	Recon Optical, Chicago, IL	
Robert Heilmann		
Dean Heimmermann	Harris Corporation, Melbourne, FL	"Nd:YAG Frequency Stabilization Techniques"
Terri Alexander	EG&G, Kennedy Space Center Titusville, FL	"Target Induced Speckle Effects in Laser Applications"
Dr. Robert Murphy	Southwest Research Institute, Warner-Robbins, GA	"Scattering from Rough Surfaces and Atmospheric Turbulence in Monostatic Laser Radar Systems"
Thomas Costello	Air Net Comm., Melbourne, FL	
Dr. Wissam Rabadi	Texas Instruments, Dallas, TX	"Multiresolution Image Reconstruction"
Stephen Reddy	Broadband Communications, Inc. Melbourne, FL	"Precision Pointing and Tracking Through Atmospheric Turbulence Utilizing the Reciprocal Path Scattering Phenomenon"
Diana Castellanos	Harris Corporation, Melbourne, FL	"Analysis and Design of an Electro-Optics Phase-Locked Loop for use in a Coherent Laser Array Receiver"
David Hefele	Invivio, Orlando, FL	"A Computer Aided Design Package for Multiple Aperture Laser Radar"
Wilson Perez	Software Technology, Inc., Orlando, FL	
Kevin Gamble	(Working on Ph.D. at CREOL)	"A Computer Simulation of a Multiple Aperture Coherent Laser Radar"

diversity receivers in order to determine what their performance might be versus their complexity. His work showed that the optimum approach to building a multiple aperture coherent receiver in which the IF signals from each aperture/detector are combined would be to compare only the phases of each IF signal (and not the phase and the amplitude) to a reference and then use a phase-locked loop to keep that phase difference close to zero. The marginal performance improvement to be obtained by compensating for both phase and amplitude differences was deemed to be not worth the cost in terms of hardware complexity.

After the signal selection/combining approach was settled, a two-aperture system with appropriate electronics was built and tested at ISTEF in October 1994. The system was tested over ISTEF's one kilometer range and performed in a vary satisfactory manner. One of the important observations from that test was that the absolute optical phase difference between two apertures only 25cm apart could be as large as 500 radians. Being able to deal with a change this large became a major issue in the final design of the phase-locked loops in the receiver. The final design approach and construction of the eight aperture system was led by Assoc. Prof. Arthur

Weeks, and is described in Section 2.2.2 and 2.2.3. In addition to coping with large relative phase shifts, the control system was re-designed by Diana Castellanos, including the redesign of the PZT elements by Thomas Costello and Chie Gagge. Improvements were made to the control system to increase its frequency response by an order-of-magnitude for the eight element system relative to the two-element system. Predictions were that the eight element system would be able to maintain a phase lock between all channels without exceeding more than 10° of phase error over the 5000 Hz fluctuation bandwidth of the atmospheric turbulence. This was achieved and will be reported in the final report for the follow-on contract.

This contract funded the project through the point where most of the electronic parts for the eight element system were purchased and the system partially assembled. Assembly of the system was completed on the follow-on contract through NCCOSC/San Diego State University Foundation.

Finally in Chapter Two a brief summary is given of efforts to measure the heterodyne mixing efficiency as part of our larger effort to understand "where are the missing dBs?" when trying to account for signal power in a coherent lidar system. The culprit is usually thought to be (a) differences in size of the two beams on the photosurface, and (b) differences in the structure of the phase fronts of the two beams. To avoid these issues each beam was injected into a sign mode fiber and then combined using a 3dB single-mode fiber optic coupler. Provision was made for adjusting the polarization of each fiber entering the coupler. Upon exiting the fiber coupler, the mixed fields were collimated, polarized, and focused into a photoreceiver. After a number of fine tuning adjustments were made, we were left with an unexplained loss of 3dB (50%). We attribute this to depolarization in the fiber and/or the fiber optic couplers, and concluded that the experiment should be repeated with polarization-preserving fiber which we did not have. Time pressure to use the test set up for the laser radar experiments prevented us from conducting this experiment using polarization-preserving fiber.

In **Chapter 3, Target-Induced Speckle Effects in Lidar Applications**, which was led by Assoc. Prof. J. Harvey, we characterized speckle from a number of different targets and sought to understand speckle effects on coherent lidar systems. In particular we found the following:

- Regarding speckle and the bidirectional reflectance distribution function, speckle is superimposed upon the BRDF. Thus the conventional (incoherent) BRDF determines the local mean intensity of the speckle pattern produced by the given target. Hence, the value of the BRDF (in the direction of the receiver) integrated over the target is necessary, in addition to the local normalized speckle statistics, when calculating the performance of lidar receivers.

- An extensive experimental investigation of the speckle size (or shape) as a function of the intensity threshold level for detection showed agreement with a theoretical model previously developed by F. Kragh, a former UCF student. Agreement was best at high threshold levels (about twice the mean intensity) but significant departures from his theory were found at low intensity levels. An explanation for this departure is offered after examining the initial approximations and assumptions made in the theory. Thus we have shown that at least at high intensity levels, we can predict the intensity of the speckle when combined with knowledge of the BRDF.

- While relatively smooth targets produce glints (narrowly directed reflected beams) at the geometrical angle of reflection from the target , the probability density function of the speckle outside of the glint is the same as

for a rough target. This is so since, in the observation plane, path length differences of greater than a wavelength will occur between the phase fronts scattered from the target surface so long as $d > \lambda L/D$ where d = distance in the observation plane, λ = wavelength, L = distance from the target to the receiver, and D = diameter of the target. Also, the intensity structure of the glint is spatially modulated by the speckle pattern with a strength that depends on the degree of scattering from the surface of the target.

- The scattering properties of the 3M retroreflective tape used to cover the target boards for our field tests of the lidar experiments were stronger by 200 to 500 fold compared with that of a perfectly reflecting Lambertian surface. This made it possible for us to do 2 kilometer round trip lidar experiments with a 35 mW laser.

Chapter 4, Lidar Simulation, summarizes the software, including the graphical user interface, developed primarily by David Hefele, Kevin Gamble, and Assoc. Profs. Arthur Weeks and Harley Myler. This simulation is unique in its ability to simulate a multiple aperture coherent laser radar in x, y, and t. This is accomplished by propagating two-dimensional waveforms between the transmitter/receiver and the target, and in allowing the target and atmosphere to evolve as a function of time. Each time the atmosphere evolves or the target changes its aspect, the perturbations introduced on the optical wavefront incident from the laser transmitter are then propagated back to the receiver and detected coherently, for each of the receiving apertures. Further,

- the individual components of an actual coherent lidar such as wave plates, polarizing beamsplitters, lenses, and detectors are modeled;
- arbitrary light polarizations can be modeled;
- scattering models including depolarization effects can be included;
- light propagation in the Z direction is described by the Rayleigh-Sommerfeld diffraction integral for which the Fresnel approximation ($Z^3 >> \pi a^4/\lambda$) is used;
- lens routines are used having increased accuracy compared to other simulations; and
- spatially and temporally evolved scattering is included in the simulation.

The output from the coherent detection process is obtained as an ac beat signal (the IF signal) resulting from mixing the signal with the local oscillator which is offset by 27.1 MHz from the signal. Noise calculations include detector noise and shot noise; the detector model also calculates the heterodyne efficiency as a function of detector size.

The software is easy to use as a result of a graphical user interface having been developed. The central feature of this interface is a window in which icons, representing optical devices, can be placed and connected to form an optical schematic. Pop-up windows are used to set the physical parameters of the devices in the schematic window. The technique developed for debugging an optical schematic is very similar to that used in debugging an electronic schematic. We plan to use this software to help design the optical link from ground to satellite in the forthcoming Photon Laser Satellite program. A copy of this software can be obtained by writing to Dr. Arthur Weeks.

Chapter 5 summarizes our **Single Aperture Coherent Lidar Experiments**, that were performed by Dr. Philip Gatt and students. These were undertaken to better understand the real world effects of speckle and target roughness. Most interestingly, we compared the coherent lidar performance of a single frequency Nd:YAG laser

with that of a single frequency CO_2 laser (having a frequency 10x lower or a wavelength 10x longer) against the same target - a rotating nail. In planning this experiment we hoped that we would experience the effects of "partially developed speckle" using the CO_2 laser compared to "fully developed speckle" with the Nd:YAG laser.

In examining the data from these two experiments, while the Doppler-width data for the Nd:YAG laser agreed with our theory, we found a 5x disagreement with the CO_2 data in that the frequency width of the scattered light was 5x narrower than we expected. Further probing of this data showed, however, that the width of the data in the 'wings" agreed with our predictions, and that the much narrower in the "main lobe" of the CO_2 data could be explained by a strong glint from a narrow region of the surface of the nail, since the width of the Doppler shift is proportional to the width of the region from which the glint originates; this would be expected to be considerably smaller than the full diameter of the nail. Thus, in coherent lidars, glints can mask the subtle signals which carry rotation rate information expected from this type of lidar. On the other hand, strong glints are, of course, useful in that they aid in detecting the presence of perhaps an otherwise unobservable target.

A second experiment involved working with Schwartz Electro-Optics of Orlando, FL to demonstrate an eye-safe Döppler lidar. This was done by combining SEO's cw holmium laser operating at 2.1 microns with CREOL's receiver employing an InGaAs detector and gold-coated reflecting optics. The system was easily able to distinguish between the vibration spectrum of a Honda and that of an Oldsmobile as these cars idled in the parking lot. This demonstrated that this eye-safe coherent lidar technology can be used for non-cooperative vehicle identification in a real-world environment.

Chapter 6, Reciprocal Path Scattering Through Random Media, describes two innovations developed on this contract which will improve optical systems involved with reciprocal path scattering. The phenomenon of retroreflection from rough surfaces or suspended particles (also called the "opposition effect", "enhanced backscatter", or the "double passage effect") is a naturally occurring conjugate-wave phenomenon and has been shown to be caused by constructive interference between reciprocal multiple scattering paths. We therefore use the phrase 'reciprocal path imaging' to describe attempts to exploit this phenomenon for obtaining diffraction-limited images of extended objects obscured by a random phase screen such as a turbulent atmosphere.

In particular, we have developed a unique and original prescription for constructing both 1D and 2D arrays of subapertures that yield a modulation transfer function exhibiting continuous and uniform (i.e. smooth) spatial frequency response up to an arbitrarily high cutoff frequency. This techniques is complex and is discussed thoroughly in the report and the appendix to Chapter 6, thus no attempt will be made here to explain it.

Secondly, and more importantly, Dr. Harvey and Dr. Phillips have developed a technique which should enable a target to be tracked through turbulence and have the resolution be limited by the diffraction limit of the pointing telescope rather than some characteristics of the turbulent atmosphere. Reciprocal path scattering produces an enhanced backscatter in the direction precisely along the boresight of the pointing telescope. The interference fringes produced by the reciprocal path scattering from the target being tracked are temporally stable in the focal plane even in the presence of time varying turbulence. Choosing the width-to-separation ratio of the dual slits appropriately and utilizing orthogonal polarizations to suppress the time-varying common path scattered radiation allows one to achieve interferometric sensitivity in pointing accuracy through a turbulent atmosphere A laboratory proof-of-concept experiment demonstrated an rms measurement error of approximately 1% of the angular width of

the diffraction-limited point-spread function. We expect that this new precision pointing and tracking technique has applications in ground-to-space laser communications, laser power beaming to satellites, and theater missile defense target tracking scenarios.

Chapter 7, Image Reconstruction, discusses the development of a suite of reconstruction algorithms for the processing of laser speckle images. This work was done by Wissam Rabadi and Assoc. Profs. Harley Myler and Arthur Weeks. Laser speckle images provide a measurement of the Fourier transform magnitude of an illuminated object while the phase information associated with the magnitude is lost. In order to be able to reconstruct the original image of the object, the phase information must be reconstructed.

The reconstruction of a signal from the knowledge of its Fourier transform magnitude remains a very difficult problem and many algorithms have been developed in the past two decades to solve this problem. Currently, the most successful and practical approach in solving the phase retrieval problem utilizes the iterative algorithms that seek numerical solutions that minimize the distance between the measurement and its estimate rather than finding an exact solution. These algorithms suffer from several drawbacks that set a limit to the size and complexity of the images that can be reconstructed. A major disadvantage of the iterative algorithms is the stagnation problem where the algorithm becomes trapped in a local minimum. The second problem is the slow convergence of the algorithm, and the third is its computational cost.

In this contract, two multiresolution adaptations of a basic iterative image reconstruction algorithm were developed that enable the algorithm to avoid stagnation, improve global convergence and dramatically reduce computational complexity.

The first of these approaches is based on Burt's pyramid and the second is based on wavelet decomposition. These two methods have the following advantages: 1) they provide a rough and quick estimate of the solution at a low resolution that may later be refined as the algorithm progresses. This coarse-to-fine strategy enables the algorithm to avoid stagnation by providing a better initial guess and giving the algorithm a higher likelihood of arriving at a global minimum. 2) They can improve the convergence rate by decomposing the search space into orthogonal subspaces that can reduce the low frequency component of the error responsible for the slow convergence of the algorithm. 3) Since the number of independent variables to be processed at each coarser level are less than that at the full resolution grid, the reduction dramatically reduces the computational cost of the algorithm and ensures a faster convergence rate.

The computer simulations that we have processed indicate that the multiresolution approaches developed here are not only capable of avoiding stagnation and reducing the computational complexity of the iterative algorithm, but also produce reconstructions that are superior in quality to the single-grid approaches and are more robust in the presence of noise.

Chapter 8 addresses the subject of **Reciprocal Path Scattering and Laser Radar.** This work was done by Robert Murphy and Prof. Ronald Phillips. A monostatic ladar system is one in which the transmitter and receiver optical paths utilize the same aperture for beam transmission and reception, respectively. The transmitted beam propagates through the turbulent atmosphere to the target, is back-scattered by the target surface, and then propagates downward through the atmosphere in the direction of the receiving aperture. The two distinct scattering processes encountered by the propagating beam are due to rough surface scattering by the target, and scattering

caused by atmospheric refractive index variations over the path due to atmospheric turbulence. Reciprocal time-reversed multiple scattering paths exist for each scattering process due to EM reciprocity; the combination of these processes sequentially produces coherent (in-phase) wave segment pairs at the receiver. The coherent component of the received power contains diffraction-limited information which is exploited by active coherent imaging applications.

At laser wavelengths, the surface irregularities of physical targets are comparable to the wavelength of the illuminating beam. This surface roughness contributes significantly to the scattered cross section of the target. Multiple scattering processes are experienced by wave segments incident on the surface. Reciprocal path coherent wave segment pairs exist for multiply scattered wave segments within the valley of the rough target surface, as well as for surface wave resonant polaritons. Coherent backscattered intensity has been observed as large as 80 times the diffuse backscattered incoherent intensity.

With respect to beam propagation, the Earth's lower atmosphere is a random inhomogenous medium due to local temperature gradients caused by solar heating. The refractive index of the atmosphere is a function of the local temperature. Temperature differences between the ground and the air produce turbulence cells (or eddies), each with slightly different temperature. These eddies are effectively individual scatterers producing both refractive and diffractive scattering events. Coherent reciprocal path wave segment pairs exist for double passage (transmitter to target - target to receiver) propagation through atmospheric turbulence. When the target is in the far-field of the turbulence layer, turbulence double passage produces a backscattered signal which is at maximum 3 dB above the diffuse backscattered incoherent intensity, as confirmed by numerical experiments.

Rough surface and turbulence double passage RPS processes each produce coherent backscatter individually. We hypothesize that in combination, the interrelation of the two processes is multiplicative, with turbulence double passage being the governing scattering process for the relative magnitudes of the coherent and incoherent intensities sensed by the receiver. This because the scattering processes occur sequentially; and the return passage through turbulence is the last scattering process that occurs during the round trip propagation from the transmitter to the receiver.

A model is presented (Appendix 8A) which quantifies the combined effect of rough surface and turbulence double passage reciprocal path scattering (RPS) on the power received by a ground-based monostatic laser radar. This is not a rigorous analysis; instead a phenomenological modification to the traditional laser radar equation (LRE) for power linkage is presented, which includes the first order effects of atmospheric turbulence as well as rough surface and double passage RPS. Within the constraints of this analysis, rough surface RPS in combination with turbulence double passage RPS increases the predicted received power of a monostatic system over that predicted by the traditional LRE. The theoretical formulation is constructed such that the relative magnitudes of the coherent and diffuse received power components are easily separated for RPS exploitation purposes.

Further, a phenomenological model is developed (appendix 8B) for the strength and spatial width of the backscattered coherent intensity peak produced by reciprocal path scattering through atmospheric turbulence. The model is applied to a ground-based monostatic laser radar tracking a space target under optically saturated atmospheric turbulence conditions. The models for the amplitude and width of the RPS peak are based upon the spatial coherence widths of the propagating fields over the up-link and down-link paths within the atmospheric

turbulence as well as the cross-sectional area of the up-link beam. The results indicate that the phenomenological model for RPS presented in this analysis provides surprisingly accurate predictions for a variety of saturated turbulence propagation environments.

2.0 COHERENT LASER RADAR ARRAY TECHNOLOGY & SYSTEMS

Philip Gatt, Ph.D., Arthur R. Weeks, Ph.D., R. L. Phillips, Ph.D., and C. M. Stickley, Ph.D.

2.1. Background

Laser radar (LIDAR) systems are a direct extension of conventional microwave radars to wavelengths from the ultraviolet to the far-infrared regions. LIDAR systems are capable of simultaneous measurement of range, reflectivity, velocity, temperature, azimuth, and elevation angle [1]. These capabilities offer a wide range of applications such as target tracking and acquisition, surface terrain mapping, vibration sensing, atmospheric composition characterization, and missile guidance. Despite the many variations to LIDAR systems, all of them share a single operating principle. Laser radiation is transmitted from the radar system to a target; the radiation interacts with the target and the propagating media, and returns to a receiver. The detected signal is processed in a specific manner so that the desired information is extracted from it. When the receiver uses the same optics as the transmitter the system is called *monostatic* laser radar system; when the receiver is placed in a different location relative to the transmitter, the system is called *bistatic* laser radar system.

Coherent detection offers improved signal-to-noise ratio over incoherent detection but with the added complexity of hardware implementation as shown in Figure 2.1. Incoherent detection is defined as the direct measurement of the received signal intensity which is function of target distance, target characteristics such as shape, size, and reflectance and background noise. Background noise in optical signals includes the effects of atmospheric turbulence, scattering from the atmosphere, the earth surface, the clouds, and any other source that may cause corruption of the return beam [1]. Incoherent detection is used in range finding, image systems, and atmospheric characterization applications [1].

Coherent detection is a mixing process in which the return signal is optically combined with a local oscillator (LO) to produce a signal that is captured by a photodetector. The mixing of these two optical fields produces a current that is proportional to the product of the fields given by [2]:

$$I_o \propto |E_s|^2 + |E_{Lo}|^2 + 2|E_s||E_{Lo}|Cos(\omega_s - \omega_{Lo})\,t \qquad (2.1)$$

where E_S and E_{LO} are the optical signal and the local oscillator fields respectively, and ω_s and ω_{Lo} are their respective frequencies. The detected signal is a function of the local oscillator power, the return signal power, and the background noise. Coherent detection is often referred to as homodyne or heterodyne. In homodyne detection the LO is extracted from the laser source of the system using a beam splitter; if the LO is frequency shifted before it is mixed with the return signal the system is called *offset homodyne* LIDAR. In heterodyne detection the LO is generated using a laser independent of the transmitter source [3]. Homodyne and heterodyne detection have been used to measure the landing speed of commercial aircraft very accurately. However, a heterodyne system cannot tell whether or not the target is approaching, since it senses only the magnitude of the velocity. This problem, called *velocity ambiguity* [3], is solved by offset homodyne detection. With this system, a stationary target will produce a signal at the same frequency as the modulator used to produce the frequency shifted LO; an advancing target will

produce a frequency higher than the modulator frequency, and a receding target will produce a frequency lower than the modulator frequency [3]. Heterodyne systems are useful for very long-range detection.

Coherent and incoherent detection are typically compared by means of the detected signal-to-noise ratio. As derived by Jelalian [1], SNR for incoherent and coherent detection are

$$SNR = \frac{\eta_D P^2_{SIG}}{hf[2B(P_{SIG} + p_{BK})] + K_1 P_{DK} + K_2 P_{TH}} \tag{2.2}$$

$$SNR = \frac{\eta_D P_{SIG} P_{LO}}{hfB[P_{SIG} + p_{BK} + P_{LO}] + K_1 P_{DK} + K_2 P_{TH}}, \tag{2.3}$$

where

SNR= electrical signal power / electrical noise power
η_D = detector quantum efficiency
h = Planck's constant (6.626×10^{-34} J-S)
f = transmission frequency
B = electronic bandwidth
P_{SIG} = received signal power
P_{DK} = equivalent dark current power
P_{TH} = equivalent receiver thermal noise
P_{LO} = local oscillator power
K_1, K_2, : dark current and detector current responsivity constants
P_{BK} = equivalent background power

Equation (2.2) shows that the SNR for incoherent detection decreases as the power of the return signal decreases while for coherent detection, as given by Equation (2.3), the LO power is an additional source of noise that can be increased to a high enough level so the system becomes *background noise limited.* Under this condition, the LO power terms cancel out, and the SNR is then directly proportional to the received signal power rather than proportional to the received signal power squared as in the case of incoherent detection. When the LO power is high enough so that its noise overcomes the background noise, the system is said to be *shot noise-limited,* and Equation (2.3) reduces to:

$$SNR = \frac{\eta_D P_{SIG}}{hfB} \tag{2.4}$$

For background noise-limited incoherent detection Equation (2.2) becomes:

$$SNR = \frac{\eta_D P^2_{SIG}}{2hfBP_{BK}} \tag{2.5}$$

By inspection of Equations (2.4) and (2.5), it can be concluded that for very weak received signals coherent detection provides higher SNR than incoherent detection. For strong signals however, in the limit, incoherent systems approach the performance of coherent systems.

Optical signals propagated through the atmosphere are affected by random fluctuations in the atmosphere as a result of temperature, humidity and pressure variations [4]. Turbulence results in a non-homogeneous index of refraction of the atmosphere over the propagation path. These fluctuations in the index of refraction cause spatial and temporal variations in the optical signal [1]. Since the return signal power is drastically reduced by these

effects, the SNR decreases. It is known from experimental results that most of the return signal phase corruption is caused by atmospheric turbulence, and that the 3 dB bandwidth of this phase disturbance is approximately 5 Hz [6].

From previous studies, the intensity statistics of a laser beam propagating through turbulence have been assumed to be K-distributed, while the phase statistics have been assumed to be Gaussian [6]. These fluctuations in the phase and intensity give rise to two detection problems: phase randomness and signal fading. Figure 2.2(a) depicts an array of detectors, each with a particular speckle cell being detected; amplitude characteristics of the individual speckle cells are shown in Figure 2.2(b).

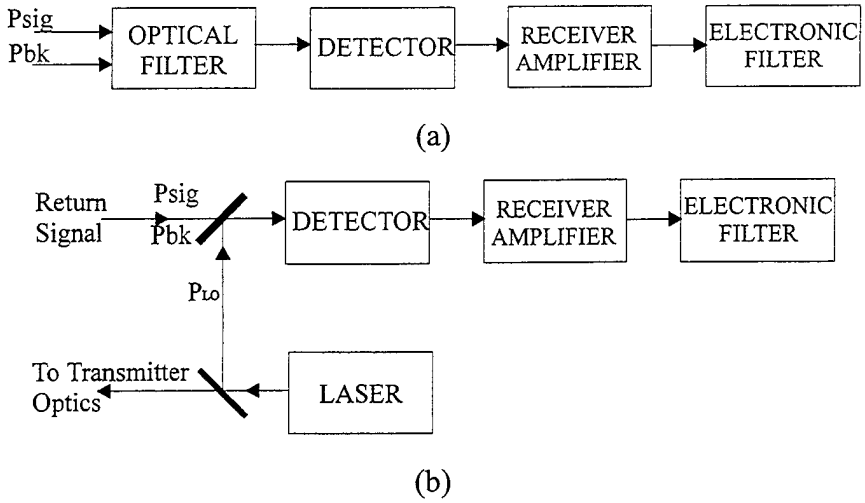


Figure 2.1. (a) Incoherent Detection and (b) Coherent Detection.

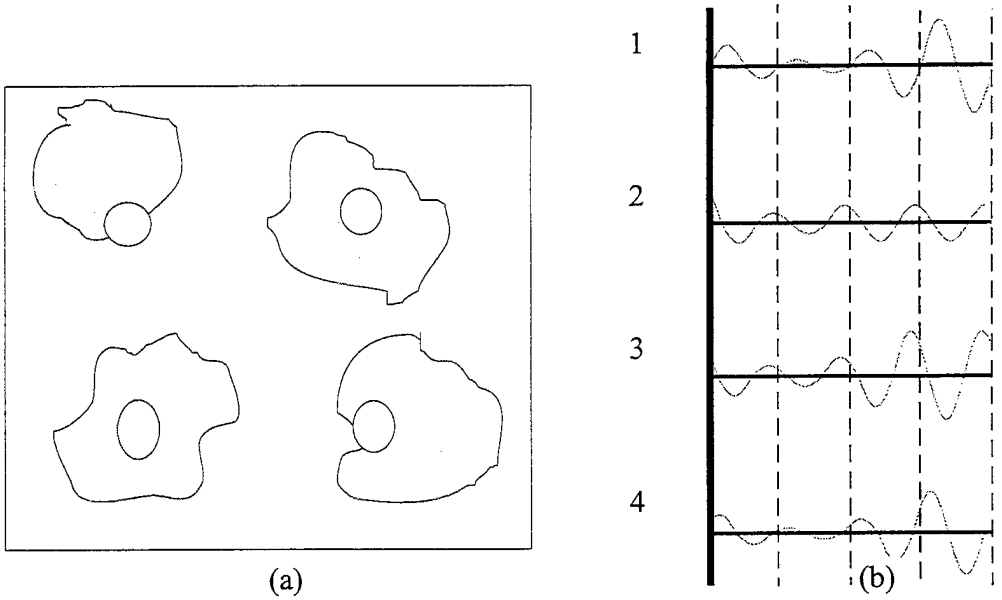


Figure 2.2: Speckle Cells Characteristics (a) Detected Speckle Cells and (b) Individual Amplitude Characteristics.

An array of detectors offers a good solution for both problems. It has been shown [7] that a detector array increases the effective receiver aperture and enhances SNR by the number of array apertures present. Phase randomness is solved by configuring the geometry and location of the array so that only one speckle cell is captured

by each detector. In this manner, each detected signal has constant phase. The size of the detectors and their center-to-center distance are functions of the *atmospheric-turbulence-induced transverse coherent length,* ρ_0, given by [7]:

$$\rho_O = \left[2.91k^2 \int_0^R Cn^2(z)\left(1-\frac{z}{R}\right)^{5/3} dz \right]^{-3/5} \tag{2.6}$$

where k = $2\pi/\lambda$ is the wave number, C_n^2 is the atmospheric strength of turbulence parameter, R is the detection range, and z is the direction of propagation.

However, the signal detected by each array component is a very weak amplitude-varying signal; therefore, it is necessary to add these signals coherently to obtain a composite signal which has an increased SNR and reduced fading. This objective is accomplished only by the in-phase addition of the signals. This process is illustrated in Figure 2.3. The phases of channels 2, 3 and 4 of Figure 2.2(b) are adjusted so that each signal is aligned with the signal from channel 1. This produces an output from the summer which has an increased SNR, and reduced fading compared to the individual inputs. Therefore, to co-phase these channels an electro-optic phase-locked loop (EOPLL) is required. The purpose of the EOPLL is to co-phase two or more optical carriers in such a way that makes the in-phase coherent addition of the two carriers possible. The in-phase addition of the optical carriers will result in a laser radar signal with improved SNR. Figure 2.3 (b) shows the output of the summer, generated by co-phasing the four inputs and then adding them.

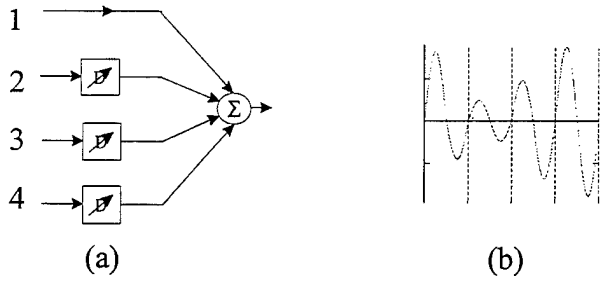


Figure 2.3: In-Phase Addition of Optical Signals (a) Co-Phasing Electronics and (b) Composite Output.

Figure 2.3 shows that one method of reducing fading due to the atmosphere is to use an array of coherent receivers where each of the received signals are co-phased with each other and then coherently added together. Co-phasing of the multiple apertures can be accomplished either optically or electronically after the optical signal has been detected. The use of electrical co-phasing techniques requiring more than 2π radians of phase shift is relatively hard to implement. Larger phase shifts are possible using optical techniques. The idea of co-phasing optical signals in not a new idea but has existed in the literature for a number of years. Unfortunately, due to the hardware complexity of building a coherent receiver most of the co-phasing techniques have been focused on laser cavity resonance improvements and pointing of array antennas. One of the first papers on optical co-phasing is that of L.H. Enloe and J. L. Rodda [8] from Bell Telephone Laboratories in Holmdel, N.J., where in the early 1960's they successfully synchronized two He-Ne lasers with a phase difference of less than a third of a degree. This technique is a direct derivation from the conventional automatic frequency control (AFC) loop [9] developed in the 1930's for synchronous communications. In this experiment the lasers are optically combined and fed to a

photomultiplier. When the frequency of the lasers is identical and their phases differ by (90° – Φ), the output of the photomultiplier consists of a D.C. term that is proportional to SinΦ. For small angles SinΦ = Φ; hence, the output of the photomultiplier is in a sense proportional to the phase difference between the two lasers. The phase error term is extracted and applied to a PZT. One of the mirrors of the controlled laser cavity is mounted on the PZT. The phase error signal applied to the PZT causes changes in the resonator cavity length; therefore, the laser frequency changes and decreases the phase difference between the two lasers. When the loop is locked, the controlled laser tracks the frequency of the uncontrolled laser so that the phase error is always less than ±90°. This experiment was performed to synchronize lasers under acoustical and mechanical disturbances. This phase synchronization technique was capable of maintaining lock for several hours.

The technique developed by Enloe and Rodda has been applied through the years. The report by T. J. Kane and A. P. Cheng [10] from Lightwave Electronics Corporation describes the implementation of a piezo-electrically tunable Nd:YAG laser ring. Unlike the Enloe-Rodda experiment, this scheme uses only one PZT-mirror combination located with the controlled laser, while the other laser operates as a stable reference. In this experiment the tuning range is limited by the PZT sensitivity and its excitation voltage range to ±100 MHz. A paper published by T. Day, A. D. Farinas, and R. L. Byer [11] describes a 1.06 μm optical phase-locked loop (OPLL) designed specifically for a coherent homodyne communication system. Like the experiment conducted by Kane and Cheng [10], this OPLL is implemented with only one PZT-mirror combination. Unlike the other techniques described, this experiment introduces a loop filter to compensate the frequency response of the amplified error signal. The loop filter selected is an ideal integrator combined with high frequency compensation; in this manner the loop has very high gain at low frequencies and a small noise bandwidth. These characteristics make this type of OPLL ideal for the co-phasing of several coherent detected signals. This system achieves tight locking with transmitter power of less than –34 dBm.

Leonid Kazovsky and Bente Jensen [12] added yet another variation to the basic Enloe-Rodda technique. They extended the application of the OPLL to solve the problem of frequency stabilization for multichannel optical communications. In this scheme the reference laser, called *master*, is phase modulated to produce a number of sidebands. Controlled lasers, called *slaves*, are locked to the master laser + sidebands using an array of OPLL's. This master-slave arrangement has successfully phase-locked up to 20 sidebands in addition to the fundamental frequency of the master laser within 3 kHz of each given frequency. In this experiment the required optical power in each sideband is only –50 dBM.

W. M. Newbert, K. H. Kudielka, W. R. Leeb, and A. L Scholtz [13] reported an experimental optical phased array antenna for laser space communications that uses OPLL's to provide a defined phase relation among the individual subantenna beams; this defined phase relation is achieved by controlling the optical length of each subantenna path relative to a reference signal. The purpose of such array antenna is to collect as much optical power as possible without the use of large-diameter optics, as it would be the case of a single aperture system. In this experiment the optical path-length control is implemented using PZT fiber stretchers [14]. This OPLL utilizes a digital phase detector [13], an analog loop filter, and an integrating amplifier. As a result, this control loop

represents a third-order PLL. This OPLL rejects disturbances as fast as 10 kHz, and its overall settling time is approximately 0.7 ms.

As presented in the original proposal, the purpose of this task is to show the feasibility of using multiaperture technology to improve the signal-to-noise ratio of coherent detected signals. The previous optical co-phasing methods provide a starting point for the development of a multi-aperture coherent array receiver. To meet the original proposed goal, a number of tasks had to be addressed in the process of designing and building a multi-aperture system:

1. Analyze and study the several methods of co-phasing signals together
2. Understand the effects that the atmosphere has on signal fading
3. Predict the signal-to-noise improvement as a function of the number of array elements.
4. Investigate techniques that can be used to provide an adequate lock-in range (i.e., a way to track and lock the two receivers even though the phase difference at their inputs becomes many times 2π)
5. Incorporate fiber optic mixers in the two element receiver
6. Incorporate pigtailed detectors and grin lenses as additional receiver apertures
7. Build, design, and test a two-element optically co-phased array receiver
8. Perform laboratory tests using the two-element system to verify its performance
9. Perform tests at ISTEF over the 1-Km range using the two-element system
10. Using the data obtained from the ISTEF experiments modify the design for the eight-element system.

The following sections describe the details of these tasks and how the results were used in the design of the eight-element coherent array receiver.

2.2. Accomplishments

In brief, our major accomplishments in this part of the program are:

- Analysis of three space diversity receivers for coherent array laser radar systems
- Design/test and demonstration of a dual aperture coherent array test bed (preliminary prototype)
- Dual aperture measurements of 2 km (round-trip) intensity and phase pertubations of one micron laser speckle propagating through a turbulent atmosphere
- Design of an eight aperture coherent array test bed (refined design)

These will be discussed in the following sections. The group's technical presentations and publications on this program are listed in Section 2.5. Two of these are reproduced in Appendices 2A and 2B.

2.2.1. Analysis of Space Diversity Receivers for Coherent Array Laser Radar Systems

In the initial stages of this program a detailed analysis of the application of three popular RF space diversity receivers for coherent array laser radars was performed. The details of this analysis are presented in a copy of a journal paper, which is attached in Appendix 2.A. The fundamental advantages a coherent array receiver

provides over a single element coherent receiver are it's inherent ability to mitigate carrier fading and to increase the mean CNR, both of which are limited by either atmospheric turbulence or target-induced speckle.

2.2.1.1. Equal Gain, Maximal Ratio and Selection Diversity Receivers

Optical coherent array receivers belong to a more general class of signal processing architectures called space diversity receivers. Major advances in space diversity receivers occurred in the 1950's where the technology was being applied to mobile radio communications. During this time frame three popular architectures emerged: 1) the selection receiver, 2) the equal-gain (EG) receiver and 3) the maximal-ratio (MR) receiver. In this section we highlight the results of a detailed performance analysis of these three receiver architectures for an optical coherent array receiver as described in Figure 2.4.

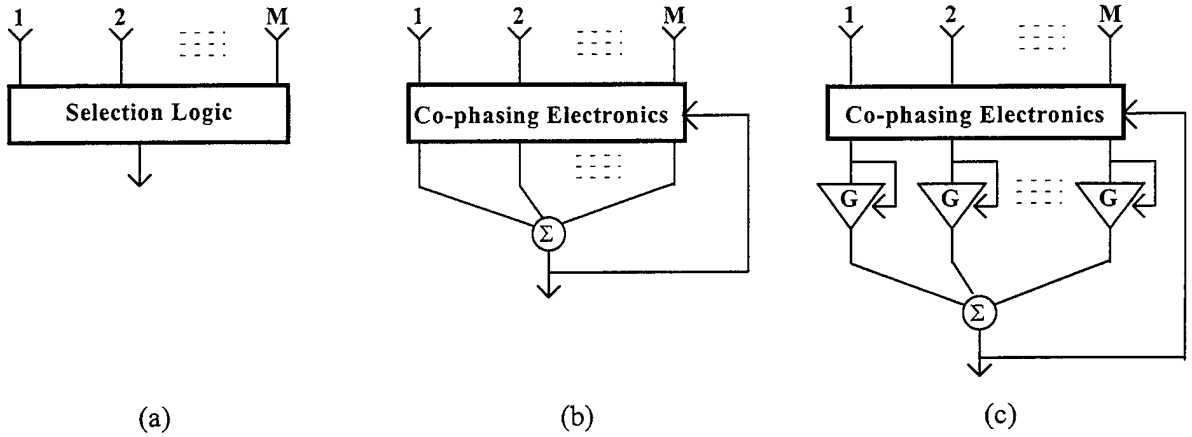


Figure 2.4: Space diversity receiver architectures:(a) selection, (b) equal-gain, and (c) maximal-ratio.

The maximal-ratio receiver, which is sometimes referred to as the ratio-squarer receiver, was first proposed by Kahn for mobile radio communications. In this architecture, the RF signals are co-phased, have their amplitudes separately adjusted, and then summed to generate a composite RF signal with improved qualities. The MR receiver is the optimal receiver, in that it yields a carrier with the highest mean CNR and lowest CNR fading. This type of signal processing technique can be thought of as a method which produces a synthetic local oscillator field that is both amplitude and phase-matched to the distorted received signal. Therefore, it yields the highest possible heterodyne efficiency. Using this conceptual approach, it becomes obvious that the optimum electronic gain for each receiver should be one which synthesizes a local oscillator field amplitude which is proportional to the received signal field amplitude.

The implementation of a MR receiver requires the development of signal processing hardware which adjusts both the delay and the gain elements of the array RF outputs. The rate at which these must be adjusted will be dictated by the rate at which the target speckle and atmospheric turbulence fluctuates. Target speckle fluctuation rates are dependent upon transverse target velocities, while the atmospheric fluctuations typically do not exceed one kHz [5].

An equal-gain receiver differs from the MR receiver in that equal gains are applied to all the RF signals, but no effort is made to match the amplitudes. Thus only the phase of the synthetic local oscillator field is adjusted

to match the signal field. As we will show in the next section, this phase-matching-only technique results in carrier statistics which are very close to those which would be obtained from the optimal MR receiver.

A selection receiver is quite different from the above two coherent summing receivers. In this type of receiver, the single strongest RF signal is switched to the output, while all the other array element signals are discarded. This, the simplest type of coherent array receiver, does little to improve the mean CNR. However, it does provide a reasonable reduction in the signal fading, albeit with less of an improvement than would be achieved using a MR receiver.

2.2.1.2 Mean CNR Improvement

Analytic expressions for the mean CNR of the three types of can be written in terms of the single aperture CNR, $\bar{\delta}$, and the number of receiver apertures M. The derivation of these expressions (Ref. 5) assumed the following: 1) each receiver had equal average noise and carrier powers, 2) the statistics of the carrier amplitude is Rayleigh, which is typical for fully developed dynamic speckle, and 3) that the signal fields in each aperture are spatially independent (i.e., the spacing between the apertures is such that the amplitude and phase statistics of each detected signal is independent).

$$\mathrm{E}[\Delta_{MR}(t)] = M\bar{\delta}. \tag{2.7}$$

$$\mathrm{E}[\Delta_{EG}(t)] = (1 + (M-1)\pi/4)\bar{\delta}. \tag{2.8}$$

$$\mathrm{E}[\Delta_{S}(t)] = M\bar{\delta}\sum_{k=0}^{M-1}\binom{M-1}{k}\frac{(-1)^k}{(k+1)^1}. \tag{2.9}$$

In the limit of large M, the mean EG receiver CNR approaches $\pi/4$ (or -1.05 dB) of the mean MR CNR. The selection receiver's mean CNR turns out to be well approximated by a logarithmic function of the number of receiver apertures. The above expressions are demonstrated graphically in Figure 2.4

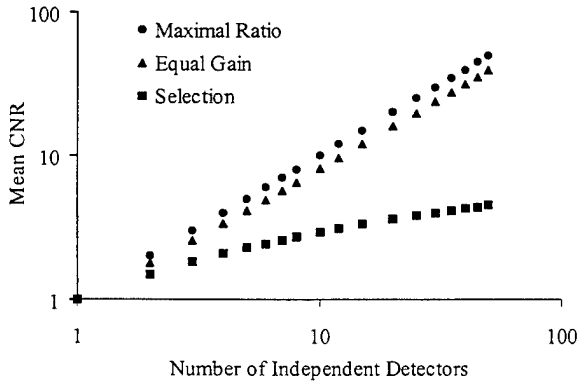


Figure 2.5: Mean CNR for the MR, EG, and selection coherent array receivers.

An alternative way to view these results is to consider the number of independent array elements required to increase the mean CNR by 10 dB. It is straightforward to show that the resulting number of array elements is 10,

13, and 12,000 for the MR, EG, and selection methods respectively. Similarly, to reduce the fading strength (normalized variance) by a factor of ten, 10, 11, and 31 independent array elements would be required for the MR, EG, and selection receivers respectively.

2.2.1.3 Carrier Fading Reduction Factors

The amount of carrier fading is quantified statistically via the normalized CNR variance, which is unity for a single aperture coherent transceiver illuminating a diffuse target. Expressions for the CNR variance, listed below, were derived during this program and were used to predict the fading reduction as a function of the number of receiver apertures.

$$\sigma n^2_{\Delta_{MR}} = 1/M. \tag{2.10}$$

$$\sigma n^2_{\Delta_{EG}} = \frac{2\left(-(2\pi^2 - 8\pi)M^2 + (5\pi^2 - 20\pi + 16)M - (3\pi^2 - 12\pi + 8)\right)}{M(\pi M + 4 - \pi)^2}. \tag{2.11}$$

$$\sigma n^2_{\Delta_S} = \frac{2}{M}\frac{\sum_{k=0}^{M-1}\binom{M-1}{k}\frac{(-1)^k}{(k+1)^3}}{\left[\sum_{k=0}^{M-1}\binom{M-1}{k}\frac{(-1)^k}{(k+1)^2}\right]^2} - 1. \tag{2.12}$$

Thus the normalized variance, for the MR receiver, decreases by a factor proportional to the reciprocal of the number of recovers. The EG receiver's normalized variance can be shown to be at worst only 9.26% larger than the MR normalized CNR variance, whereas, the selection receiver's normalized variance reduces much more slowly. The above expressions are depicted graphically in Figure 2.5

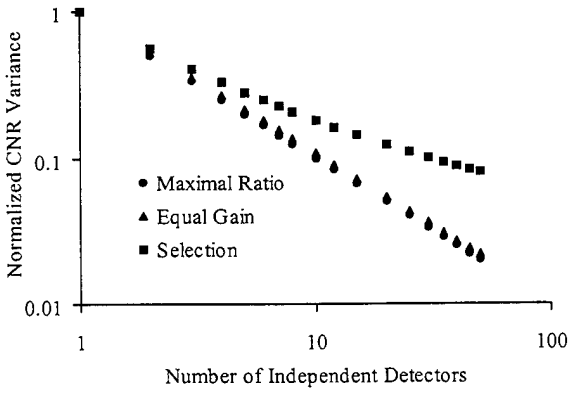


Figure 2.6: Normalized CNR variance for the MR, EG, and selection coherent array receivers.

2.2.1.4 Effect of Signal Correlation Between Receiver Apertures

In some configurations (i.e., dense arrays, or if the array elements are small and closely spaced compared to the speckle correlation length) the RF amplitudes will be correlated due to overlapping speckle. In the case of the MR receiver, if the signal amplitudes are not spatially independent, then the signal fading will be worse than predicted by the theoretical expressions given in the above section, however the mean CNR for an MR receiver will remain unchanged. In the limit, as the speckle size exceeds the overall array size, the normalized variance of the composite RF signal will increase from $1/M$ to that which is obtained using a single detector (i.e., unity). In other words, in this extreme limit, the array behaves just like a single, equal area, large aperture conventional coherent receiver.

2.2.2. Dual Aperture Coherent Array Receiver Design and Development

Even though the maximum ratio architecture for combining several received signals provides the best improvement in the mean CNR and the best reduction in fading, this architecture is by far the most difficult to design and build. In comparing the equal gain method with the maximum ratio method, for array receivers in which the number of apertures is less than 100, the improvement in the mean CNR of using the maximum ratio method over the equal gain method can not be justified due to the increased complexity in hardware implementation. Hence, the approach taken in designing the two-aperture array receiver is to use the equal gain method. This essentially reduces the problem to co-phasing each of the two receiver apertures with each other.

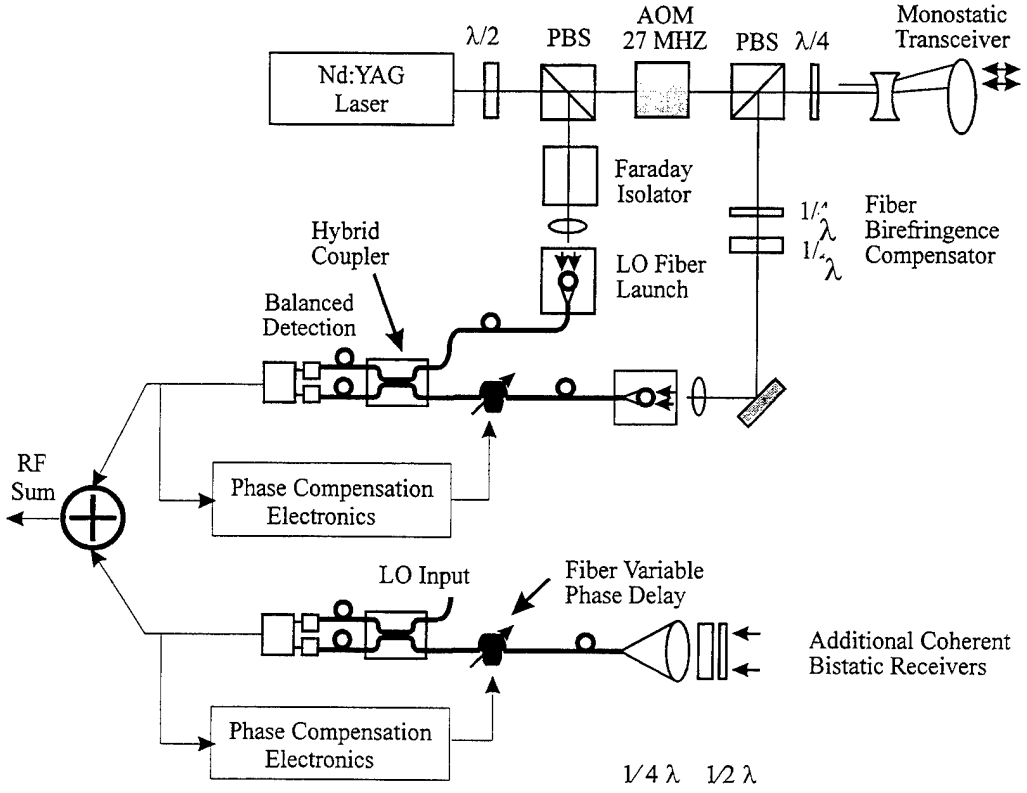


Figure 2.7: The two-element transceiver system.

We have developed a dual aperture offset homodyne coherent laser radar system illustrated in Figure 2.7. The system utilizes a Nd:YAG 50 mW commercially available laser source. A Faraday isolator produces 40 dB of feedback isolation. The focused beam, which is vertically polarized, goes through a half-wave (HW) plate that causes the polarization of the laser beam to have a circular component. The beam is then propagated through a polarizing beam splitter (PBS) that reflects the vertically polarized part of the signal by 90° to become the local oscillator (LO). This HW plate and PBS combination provides a variable adjustment of the laser transmitter power and the LO beam. The LO is then launched into a fiber to be mixed later with the received signal.

The horizontally polarized part, the transmitter field, passes through an acousto-optic modulator (AOM) that modulates the beam by 27.12 MHz, and then through a second PBS that transmits straight ahead only horizontally polarized light. This PBS is followed by a quarter-wave (QW) plate that produces right-hand circularly polarized light. The PBS-QW plate combination is known as an optical switch or polarization duplexer. This optical switch separates the return field from the transmitter field in the monostatic transceiver. The received optical beam then passes through a static fiber birefringence compensator that consists of a QW and a HW plate and then is launched into a fiber to be mixed with the LO beam via a fiber optic coupler.

The bistatic receiver is implemented using a commercially available fiber beam collimator. The optical alignment of the bistatic receiver is accomplished by maximizing the overlap between the back-propagated electric field at the target with the transmitted field at the target. The received signal is then fed into a variable fiber phase delay unit from which its output is optically combined with the LO via a fiber optic coupler. The output of the fiber optic coupler is then detected using a photo-detector-pre-amplifier stage. The output of this is then fed back to the phase compensation electronics shown in Fig. 2.8.

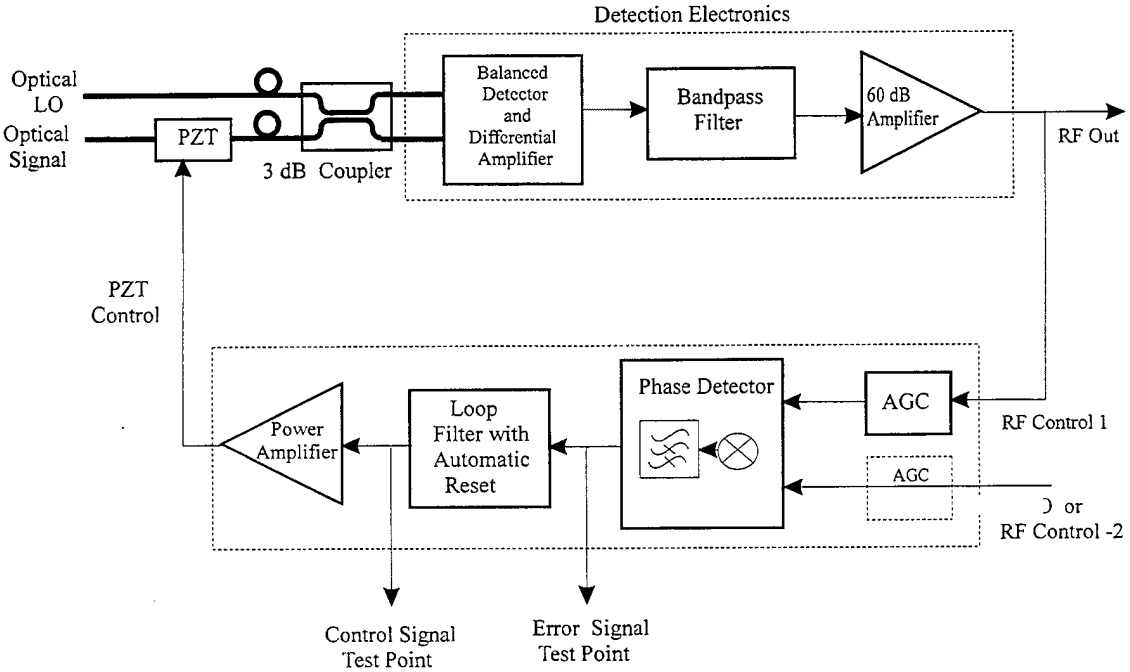


Figure 2.8: Electro-Optic Phase-Locked Loop Block Diagram.

This drawing illustrates all the components of the electro-optic phase-locked loop (EOPLL) and how it is integrated into the transceiver. The phase compensation electronics are part of the electro-optic phase-locked loop. The phase compensation electronics consists of the automatic gain-controlled amplifier (AGC), the phase detector, the loop filter, and the high power amplifier.

There are two major categories of phase detectors: sequential circuits and multiplier circuits [15]. The output of a sequential phase detector is a function of the time interval between a zero crossing of the input signal and a zero crossing of the reference signal. A multiplier circuit generates an output that is proportional to the phase difference between the input signal and the reference signal. In general sequential circuits do not operate as well as multipliers in the presence of noise because their response depends on waveform edges. For this reason a multiplier circuit was selected to implement the phase detector for the CREOL EOPLL.

An input electrical signal to a multiplier circuit can be represented by:

$$\theta_i = A_i Sin\omega_0 t \tag{2.13}$$

and a reference signal that has a phase difference with respect to the other input ϕ it can be written as:

$$\theta_0 = A_0 Sin(\omega_0 t + \phi) \tag{2.14}$$

The multiplier's output signal is then given by:

$$v_{PD} = \theta_i \theta_0 = \frac{A_i A_0}{2} K_m Cos\phi - \frac{A_i A_0}{2} K_m Cos(2\omega_0 t + \phi) \tag{2.15}$$

where K_m is the mixer gain. For the EOPLL application, the double frequency term is eliminated by the lowpass loop filter before it reaches the PZT; therefore, the error function of the loop is:

$$v_e = \frac{A_i A_0}{2} K_m Cos(\phi) \tag{2.16}$$

The error signal is proportional to phase differences about 90°, since $v_e = 0$ when $\phi = \pi/2$. For small changes in phase $\Delta\phi$, the phase difference is:

$$\phi = \frac{\pi}{2} + \Delta\phi \tag{2.17}$$

and the error signal becomes:

$$v_e = \frac{A_i A_O}{2} K_m Cos(\frac{\pi}{2} + \Delta\phi) = \frac{A_i A_O}{2} K_m Sin(\Delta\phi) \cong \frac{A_i A_O}{2} K_m \Delta\phi \tag{2.18}$$

Figure 2.9 gives the measured voltage output of the phase detector as a function of phase error for constant values for the input amplitudes of both signals (A_i, A_o = 0.5 vrms). For small phase errors shown in the center of the plot, the transfer curve in nearly linear validating the approximation made in Equation (2.18). Unfortunately, Equation (2.18) clearly shows that the phase detector output depends on the amplitude of the input signals. Hence fading in the received signal results in variations in the phase detector output. In order to have a constant phase detector gain, the amplitude of its input signals must be kept at a constant level, which is accomplished using an *automatic gain control* circuit (AGC).

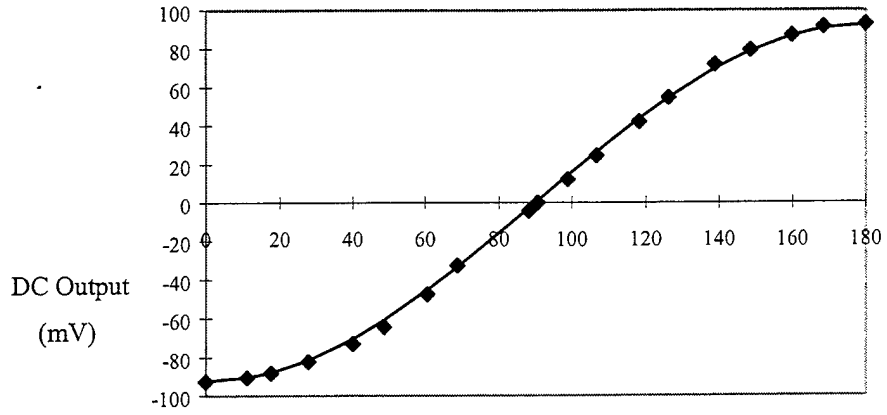


Figure 2.9: Experimentally Measured Phase Detector Performance Curve.

The overall gain of a EOPLL is a function of the amplitude of the input signal [16]; therefore, it is necessary to provide the phase detector with input signals of constant and known amplitude. The addition of an AGC circuit at both phase detector inputs offers a good solution for this problem since an AGC yields a constant amplitude output over a wide dynamic range of input voltages and introduces little or no phase distortion.

The system requirements for the AGC are shown on Table 2.1. The 40 dB dynamic range requirement for the AGC was obtained from amplitude fade data which show that the received signal amplitude variation due to turbulence lies within a 40 dB range. The maximum phase change of 3° or 0.052 radians over the AGC input dynamic range was determined from the 10° phase requirement..

Table 2.1: AGC Design Requirements.

Parameter	Value
Dynamic range	$\geq$ 40 dB
AGC range	1 mV_{RMS} to 100 mV_{RMS}
Phase Change Max.	3 °
Bandwidth	$\approx$ 40 MHz

The plots presented in Figure 2.10 were obtained from actual AGC data. These plots show that the AGC has a dynamic range of 41.5 dB, and less than 2° of phase shift over its range of operation. This last feature, the low phase shift, is what makes this design appropriate for the EOPLL since the goal is to introduce in the system as little phase shift as possible. Thus the response of the loop is determined by the true phase difference between the phase detector inputs alone. Finally, the bandwidth of the AGC was measured to be approximately 35 MHz.

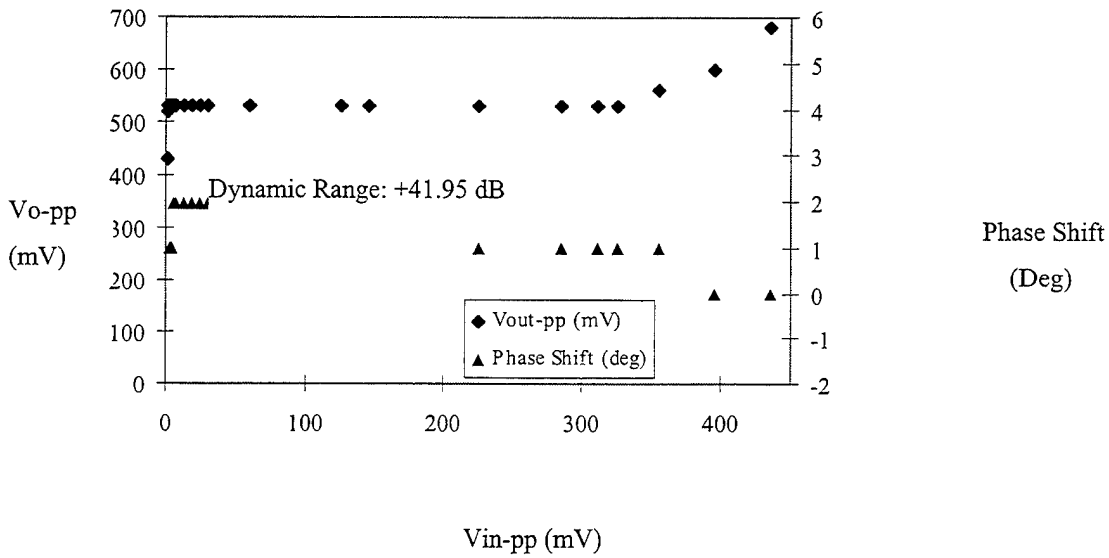


Figure 2.10: Measured AGC Transfer Characteristics

The magnitudes of piezo-electric voltages and forces are very small; therefore, the induced dimensional changes in PZT's are on the order of micro-inches. In particular, the wrapped PZT selected to implement the optical delay of the EOPLL has a gain of only 3.5 radians per volt. However, it has a large linear dynamic range, thus, a voltage swing of 200 V across the electrical terminals of the PZT yields a phase dynamic range of 700 radians peak. A high voltage power amplifier was designed which provides the voltage level necessary to obtain about 500 radians of dynamic range for the control loop of the Laser Radar System of 502.6 radians.

The power amplifier was designed with an APEX PA85A power operational amplifier in an inverting configuration. The PA85A power operational amplifier from APEX is recommended for piezo-transducer excitation since it features ±225V of output swing, 200mA of output current, and can be externally compensated to enhance its stability. The amplifier gain was set to 15 so that an input of ±15V results in the total output swing of ±225V (± 787.5 radians) which the PA85A is capable of delivering. Overvoltage hazards can cause damage to the power operational amplifiers as a sudden voltage increase may be generated by power supplies, inductive loads, and PZT's. Overvoltage can destroy a power operational amplifier through the power supply terminals, the input terminals, or the output terminal. Hence, power supply terminals were protected with unidirectional transient suppresser diodes that withstand high transient stress. Unipolar TranZorb diodes are preferred since they can protect the amplifier from accidental supply reversal or supply loss.

Figure 2.11 illustrates the high voltage amplifier frequency response. The curve was generated from experimental data with an input of 200mV. Circuit testing demonstrates good stability and flat frequency response up to 100 kHz. The response of this high voltage amplifier was satisfactory and it has met the design specifications. An important feature of this design is the protection against the typical overvoltage hazards present in most PZT applications.

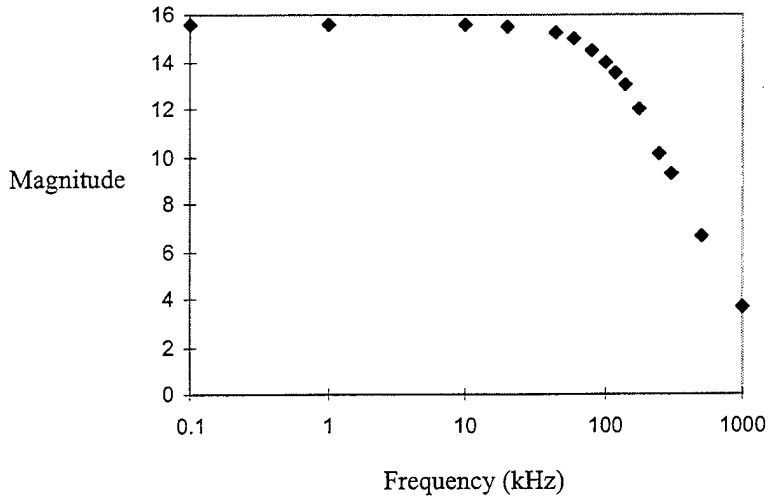


Figure 2.11. Power Amplifier Frequency Response.

The loop filter used in the two-aperture system was a simple integrator containing a adjustable gain constant. This loop filter produced a closed loop response which was flat to 500 Hz. This control system was sufficient to perform initial experiments with the two-element system at ISTEF over the 1km range.

2.2.3. The Design and Development of an Eight Aperture Coherent Array Receiver

Upon the successful demonstration of the coherent array concept using two apertures at ISTEF, CREOL proceeded with the design and development of an eight aperture coherent array. Knowledge obtained during the development and test of the dual-aperture prototype receiver was utilized to design an improved array technology, which has been completed on our follow-on contract. Advances in this technology were made in the following areas:

- Development of a dual detector photoreceiver
- Development of improved PZT devices (electrical models and selection criteria) culminating in the selection of an improved PZT device with a first resonance near 47 kHz
- Development of an improved loop filter design with greater than 5 kHz error bandwidth
- Development of a better understanding of control systems and definition of a detailed performance evaluation procedure.

In this section we describe these advances, with the exception of the dual-detector photoreceiver, since it is described in a later section.

The design of the *hybrid* phase locked loop is based upon a linear model [17]. Although the PLL is nonlinear because the phase detector is a nonlinear device, linearity can be assumed when the loop is locked, since under the locked condition it operates with a very small phase error. Since $Sin(\Delta\theta) \cong \Delta\theta$ for $\Delta\theta < 30°$, the phase detector can be modeled as a linear device when the phase difference is small. As a linear device, the phase detector is modeled as a summing junction and a gain block that is proportional to the level of the two inputs. Using the linear

model, the hybrid phase-locked loop can be represented by the block diagram given in Figure 2.12. The input signal has a phase of $\Theta_i(t)$, and the PZT output has a phase of $\Theta_o(t)$. The phase detector output or error signal is

$$v_e(t) = K_e(\Theta_i - \Theta_o)\,, \tag{2.19}$$

where K_e is the phase detector gain factor and has units of volts per radian.

The error signal is filtered and amplified. The loop filter suppresses noise and high frequency components. The loop filter and PZT transfer functions will be specified in the following sections. For now, they will be represented by $K_F F(s)$ and $K_P PZT(s)$, where K_F, the loop filter gain factor, is unitless, and K_P, the PZT gain factor, has units of radians per volt. In the diagram of Figure 2.11 the gain block K represents the overall gain of the loop, which is given by:

$$K = K_e\, K_F\, K_P\, K_{HVA} \tag{2.20}$$

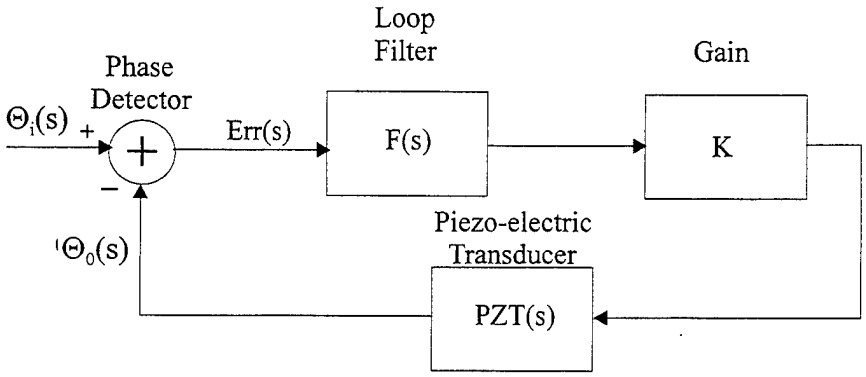


Figure 2.12. Hybrid PLL Linear Model.

All the terms in Equation 2.20 have been defined, except K_{HVA}. This factor represents the gain of the high voltage amplifier that drives the PZT and that is needed in the system to provide appropriate signal levels to meet the requirements of the PZT.

The block diagram of Figure 2.12 can be described by the following equations:

$$\Theta_e(s) = \Theta_i(s) - \Theta_0(s) \tag{2.21}$$

$$\Theta_0(s) = \Theta_i(s)F(s)PZT(s)K \tag{2.22}$$

The characteristic loop transfer functions can be derived from Equations (2.21) and (2.22). Replacing Equation (2.21) in (2.22) yields the following:

$$\Theta_0(s) = \left[\Theta_i(s) - \Theta_0(s)\right]F(s)PZT(s)K \tag{2.23}$$

$$\frac{\Theta_0(s)}{\Theta_i(s)-\Theta_0(s)} = F(s)PZT(s)K \tag{2.24}$$

$$\frac{\Theta_i(s)}{\Theta_0(s)} = \frac{1+F(s)PZT(s)K}{F(s)PZT(s)K} \tag{2.25}$$

$$H(s) = \frac{\Theta_O(s)}{\Theta_I(s)} = \frac{KF(s)PZT(s)}{1+KF(s)PZT(s)} \tag{2.26}$$

Equation 2.26 is known as the Closed-loop Transfer Function of the EOPLL, since it gives the output to input relation when the loop is closed. Another important loop transfer function is the Error Function Err(s), which establishes the phase error to input phase relation. Placing Equation (2.21) in Equation (2.22) yields the following equations:

$$\Theta_e(s) = \Theta_i(s) - \Theta_e(s)F(s)PZT(s)K \tag{2.27}$$

$$\Theta_i(s) = \Theta_e(s)\left[1+F(s)PZT(s)K\right] \tag{.28}$$

$$Err(s) = \frac{\Theta_e(s)}{\Theta_i(s)} = \frac{1}{1+F(s)PZT(s)K} = 1-H(s) \tag{2.29}$$

Finally, the Open Loop Transfer Function, G(s), is given by:

$$G(s) = KF(s)PZT(s) \tag{2.30}$$

Ideally, when the loop is locked, in the limit, H(s) in Equation (2.30) approaches 1 whereas $Err(s) \rightarrow 0$. Equations (2.26), (2.27) and (2.28) describe the control system, and typically are utilized to evaluate the performance of the system. A detailed analysis of the PZT and its transfer function must be given in order to determine the system performance, since the PZT characterization will be the determining factor for the loop filter design. Once K_P PZT(s) and K_FF(s) are defined, the performance of the hybrid phase-locked loop can be evaluated.

As established during the two-element experiments at ISTEF [5], most of the phase corruption of the optical signal is caused by atmospheric turbulence. Therefore, the system requirements are determined by the statistics of the phase fluctuations. Collected measurements show that atmospheric turbulence causes as much as ±40 waves of phase deviation or 502.65 radians peak to peak. From the same set of measurements the phase power spectrum was obtained, and it was determined that at 5 Hz the phase disturbance is around 502 radians, while at 1 kHz it is only about 0.709 radians. These two figures are used to set the specifications of the control system.

From the block diagram of Figure 2.12 with the loop open the error signal can be expressed as:

$$\Theta_e(s) = \frac{\Theta_0(s)}{KF(s)PZT(s)} \tag{2.31}$$

In Equation (2.31) the overall system gain is K, as given by Equation (2.20). To set the gain specifications for the system, F(s) and PZT(s) in Equation (2.31) can be disregarded; these expressions do not involve gain, since

all the gain factors of the system are represented in the overall system gain K. The maximum phase error allowed in the loop is set at 10° or 0.17 radians, as specified by the design requirements of the system. From Equation (2.31) the gain is given by:

$$K = \frac{\Theta_0(s)}{\Theta_e(s)}\Bigg|_{\Theta_e(s)=0.17} \qquad (2.31)$$

Substituting the 5 Hz phase disturbance and the maximum error in Equation (2.13)of 0.17, a required minimum gain at 5 Hz of approximately 3000 or 70 dB is obtained:

$$K_{5Hz} = \frac{502}{0.17} \approx 3000 \qquad (2.33)$$

Similarly, the required minimum gain at 1 kHz is approximately 4.1 or 12.2 dB, as given by:

$$K_{1KHz} = \frac{0.709}{0.17} \approx 4.1 \qquad (2.34)$$

These specifications assure that the electro-optic phase-locked loop (EOPLL) can effectively co-phase signals propagating through atmospheric turbulence while maintaining the system requirement of less than 10° of phase error. With the system gain specifications given, the next step is to define the PZT and loop filter transfer functions. With these functions defined, all the components of the EOPLL will be specified, and the performance of the system can be evaluated.

PZT's are manufactured in different shapes and configurations, and can be used in several different ways as optical delay devices. For the EOPLL a PZT cylinder is most advantageous for ease in which optical fiber can be wrapped around the device, and because of its high frequency response. Piezo-electric ceramics present resonant modes that depend on the size and shape of the device. Thin-walled cylinders exhibit three dominant resonant modes: one axial and two radial. One of the radial modes is determined by the wall thickness and is usually higher than 300 kHz. The other radial mode, called the "hoop" mode, depends on the diameter of the cylinder. This displacement mode produces most of the phase variation in the signal propagating through the fiber by changing the length of the fiber. The axial mode depends on the length of the device and can appear in the same frequency range as the hoop mode [18].

A PZT cylinder of 0.125 in. of wall thickness, a diameter of 1in., and 0.75 in. long was found to be the most appropriate device for the optical delay unit within the CREOL laser radar system. The dimensions of the PZT were selected to produce a hoop resonance at approximately 47 kHz. It was desired to produce an axial resonance significantly higher than the hoop resonance, so that a single resonant mode model can be used to model this device. PZT-4 from Morgan Matroc INC., lead zirconate titanate, was the material selected for this application. The selection of this material was based on the high elastic modulus and low power dissipation of this material [18].

The use of an equivalent electrical circuit has been of common practice to model electrical and mechanical systems. There are several basic equivalent circuits for modeling a PZT. The circuit shown in Figure 2.13 was

found to be the most appropriate for the PZT cylinder chosen to implement the optical delay in the EOPLL, since it represents a PZT with a single resonant mode.

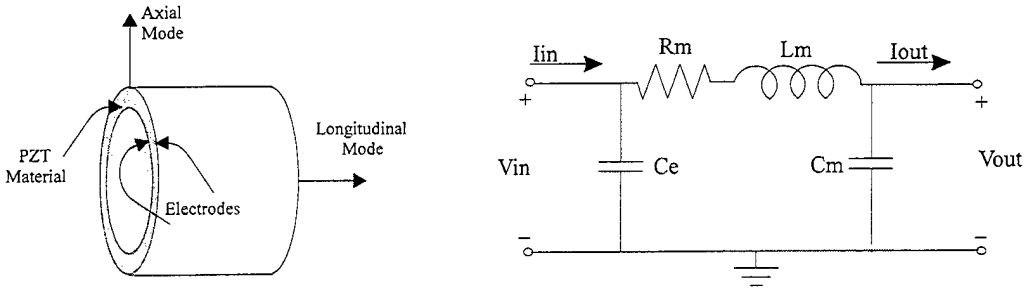


Figure 2.13: PZT and Equivalent Circuit Model.

In the equivalent circuit the input represents the PZT terminals, where the excitation voltage is applied, and the output represents the deformation. The output can be considered as the mechanical terminals, where the mechanical energy transfer to or from the PZT takes place. The inductance represents the effective vibrating mass of the device [18]. Capacitor *Ce* represents the electrical capacitance and capacitor *Cm* represents the mechanical compliance of the PZT. Resistor *Rm* represents the mechanical impedance. The circuit current represents the velocity of the energy transfer. It is important to point out that any equivalent circuit is only an approximation to the actual behavior of the device. This approximation is valid and accurate only up to frequencies slightly higher than the first resonant frequency of the device.

Using parametric circuit analysis techniques the equivalent circuit is described by the following equations:

$$\begin{bmatrix} Vin \\ Iin \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ sCe & 1 \end{bmatrix} \begin{bmatrix} 1 & Rm + sLm \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ sCm & 1 \end{bmatrix} \begin{bmatrix} Vout \\ Iout \end{bmatrix} \tag{2.35}$$

$$\begin{bmatrix} Vin \\ Iin \end{bmatrix} = \begin{bmatrix} 1 + CmRms + LmCms^2 & Rm + Lms \\ (Ce + Cm)s + RmCeCms^2 + LmCeCms^3 & 1 + RmCes + LmCms^2 \end{bmatrix} \begin{bmatrix} Vout \\ Iout \end{bmatrix} \tag{2.36}$$

Since the circuit is not loaded $I_{out} = 0$; hence, Equation (2.36) becomes:

$$\begin{bmatrix} Vin \\ Iin \end{bmatrix} = \begin{bmatrix} 1 + CmRms + LmCms^2 \\ (Ce + Cm)s + RmCeCms^2 + LmCeCms^3 \end{bmatrix} \begin{bmatrix} Vout \\ Vout \end{bmatrix} \tag{2.37}$$

Equation (2.37) yields much insight about the behavior of the PZT. From this equation the transfer function and the characteristic impedance of the device can be derived.

The transfer function is given by:

$$PZT(s) = \frac{Vout}{Vin} = \frac{1}{1 + RmCms + LmCms^2} \tag{2.38}$$

The input current is given by:

$$Iin(s) = \left[(Ce + Cm)s + RmCeCms^2 + LmCeCms^3 \right] \times Vout \tag{2.39}$$

Combining equations (2.38) and (2.39) an expression for the impedance of the PZT can be obtained as follows:

$$Zin(s) = \frac{Vin}{Iin} = \frac{1 + CmRms + LmCms^2}{(Ce + cm)s\left[1 + \frac{CeCmRm}{Ce + Cm}s + \frac{LmCeCm}{Ce + Cm}s^2\right]} \tag{2.40}$$

The impedance of the PZT given by equation (2.40) is a function of the form

$$Z(s) = \frac{1 + \frac{s}{Q\omega_n} + \frac{s^2}{\omega_n^2}}{C_T s\left[1 + \frac{s}{Q_p\omega_p} + \frac{s^2}{\omega_p^2}\right]} \tag{2.41}$$

Comparing equations (2.40) and (2.41) yields:

$$f_n = \frac{1}{2\pi}\left[\frac{Ce + Cm}{LmCeCm}\right]^{1/2} \tag{2.42}$$

$$f_p = \frac{1}{2\pi\,(LmCm)^{1/2}} \tag{2.43}$$

$$Q_n\omega_n = \frac{1}{CmRm} \tag{2.44}$$

$$Q_p\omega_p = \frac{Ce + Cm}{RmCeCm} \tag{2.45}$$

At very low frequencies Equation (2.40) can be approximated with a linear function with slope of

$$m = \frac{1}{(Ce + Cm)2\pi} \tag{2.46}$$

Hence, Equations (2.41) through (2.46) can be used to calculate the values of the model *parameters Ce, Cm, Lm* and *Rm* from the measured performance of the PZT.

Figure 2.14 depicts the experimental set-up used to obtain measurements for f_p, f_n, m, Q_n and Q_p. As it will be shown later, the addition of resistor *Re* in series with the PZT, as shown in Figure 2.14, changes the response of the PZT from a second order to a third order system. Nevertheless, if *Re* is very small, its effect is negligible; therefore, the measurements of the critical frequencies and the slope obtained with this experimental set-up will yield a good approximation to the model parameters.

The input signal is a sine wave of known amplitude. By manually sweeping the frequency of the input signal, f_n and f_p are located at the frequencies of highest and lowest output levels, V_{out}, respectively. The measured frequencies for a 0.125 × 1 × 0.75 in. PZT-4 cylinder with eight layers of fiber wrapped around it were found to be f_n = 47.74 KHz and f_p = 46.36 KHz. Q_n and Q_p are calculated dividing f_n and f_p by their respective measured 3 dB bandwidth. The 3 dB bandwidth is found by measuring the frequencies around f_n and f_p at which the magnitude of the output differs by 3 dB from the magnitude at f_n and f_p. The 3 dB bandwidth around both f_n and f_p was measured to be 200 Hz; therefore, Q_n was calculated as 238.7, while Q_p was found to be 231.8.

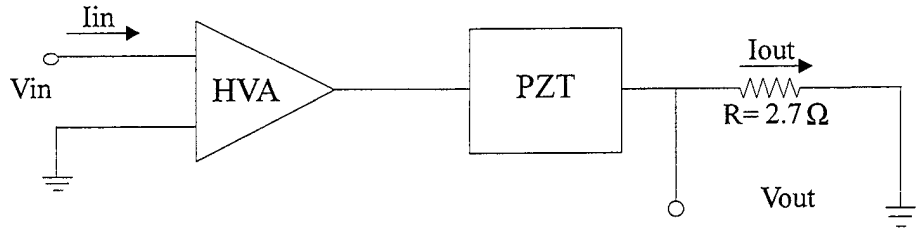


Figure 2.14. Experimental set-up for PZT Model Parameters.

Remembering that Z = Vin / Iin, Y = 1/ Z, I = Vout/ Re, and Y = Vout/ (Vin × Re), a set of input and output voltages measured at low frequencies can be obtained to calculate the slope of the admittance plot. The slope of the admittance plot is the inverse magnitude of the impedance slope. The slope of the admittance plot at low frequencies was found to be 32.075 × 10^{-9} Ω^{-1}. The PZT model parameters shown on Table 2.2 were obtained solving Equations (2.41), (2.42), and (2.45) simultaneously, and averaging the values of *Rm* calculated from Equations (2.44) and (2.46).

Table 2.2: PZT Model Parameters.

Parameter	Value
Ce	4.815×10^{-9} F
Cm	0.284×10^{-9} F
Lm	41.48×10^{-3} H
Rm	59.93 Ω

The measured total capacitance of the PZT, $C_T = Ce + Cm$, is 5.31 nF; calculating C_T from the PZT parameters derived from experimental data yields a value of 5.1 nF, which is very closed to the direct measured value of 5.31 nF. This comparison shows that the model parameters obtained are a very good approximation to the actual PZT parameters. The impedance plot of the PZT model shown in Figure 2.14 was generated from Equation (2.39) with the model parameters; this plot exhibits the expected critical frequencies at around 46.3 KHz and 47.6 KHz.

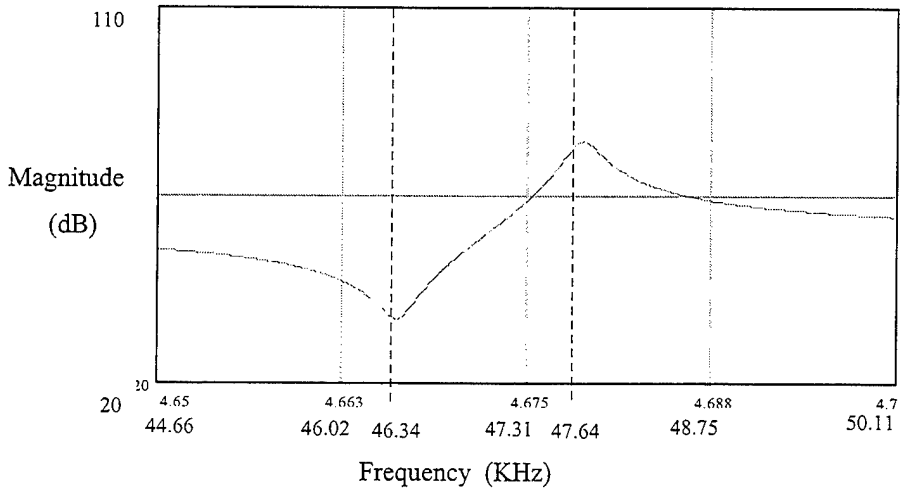


Figure 2.15. Impedance Plot for 0.125×1×0.75 in. PZT-4 Cylinder.

Equation (2.40) was useful to determine the model parameters, but from the control system point of view the area of interest for the PZT is its transfer function, as given by Equation (2.38). The plots shown in Figure 2.16 (a) and (b) were generated using equation (2.38) and the model parameters found from the previously described measurements. The resonant peak shown in the magnitude response, Figure 2.16 (a), at approximately 46 KHz is due to the high mechanical Q of the PZT. A mechanical Q of approximately 230 and a gain of 3.5 rad/volt were evaluated in the laboratory. These results are very close to the Q measurements obtained from experimental data, which gave Q values of 238.7 and 231.8.

The resonant peak in the PZT response represents a problem for the design of the control system. The presence of this resonant peak will reduce the overall gain that the system can tolerate before becoming unstable. Considering this problem, two possible solutions were proposed to reduce the resonant peak: a series resistance, and a notch filter.

The addition of a series resistance reduces the mechanical Q of the PZT, but also introduces some phase loss. The presence of the series resistance transforms the PZT transfer function from the response given by Equation (2.38) to a response of the form:

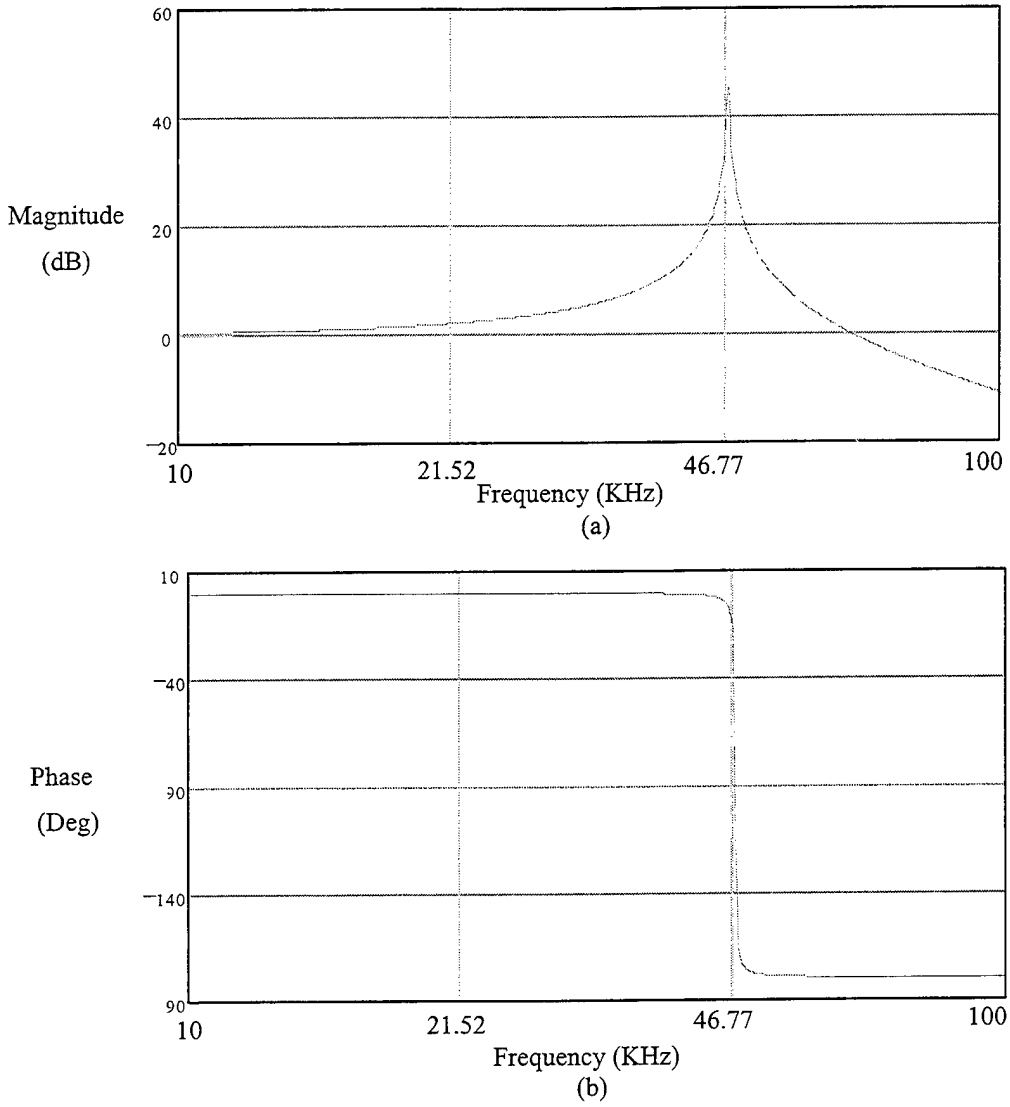


Figure 2.16: (a) Magnitude Response of the PZT Transfer Function and (b) Phase Response of the PZT Transfer Function

$$PZT(s) = \frac{1}{1 + a_1 s + a_2 s^2 + a_3 s^3} \tag{2.47}$$

where $a_1 = Rm\ Cm + Re\ (Ce + Cm)$, $a_2 = LmCm + ReRmCeCm$, and $a_3 = ReLmCeCm$. Figures 2.17 and 2.18 show the PZT response for several values of the added series resistance. Re is varied from 100Ω to 900Ω in increments of 200Ω. These plots show that the addition of the series resistance effectively reduces the resonant peak of the PZT, but it also causes some phase loss. As seen in the magnitude plot, most of the peak attenuation takes place when Re is less than 600Ω; beyond that value, an increase in Re does not yield significant peak attenuation. The phase plot shows that there is a more or less uniform increment in phase loss as Re increases. Hence, beyond 600Ω the amount of phase loss is significantly higher than the degree of attenuation that can be achieved. As a result of this analysis,

a series resistance of 510Ω was selected to reduce the resonant peak of the PZT. This series resistance provides approximately 20 dB of peak attenuation with a phase loss of only 10° at 10 KHz.

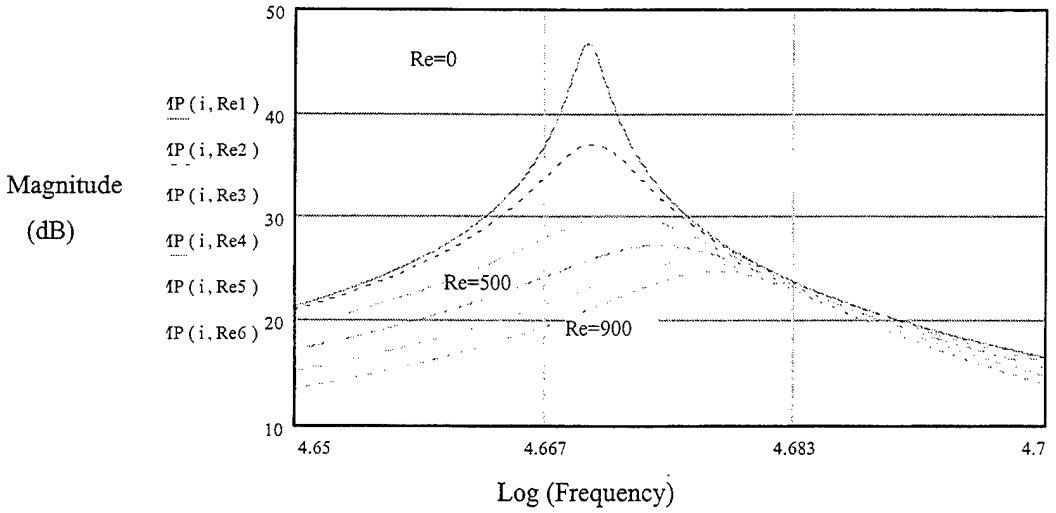


Figure 2.17: Effect of Re on the Magnitude of the PZT Response

The second method considered to address the problem represented by the resonant peak in the PZT's mechanical displacement curve was the addition of a notch filter. A classical design approach is to implement two zeros and two complex poles to produce a notch filter. A detailed analysis of analog notch filters is presented by Valkenburg and analyzed here. The second-order notch characteristic is given by [19]:

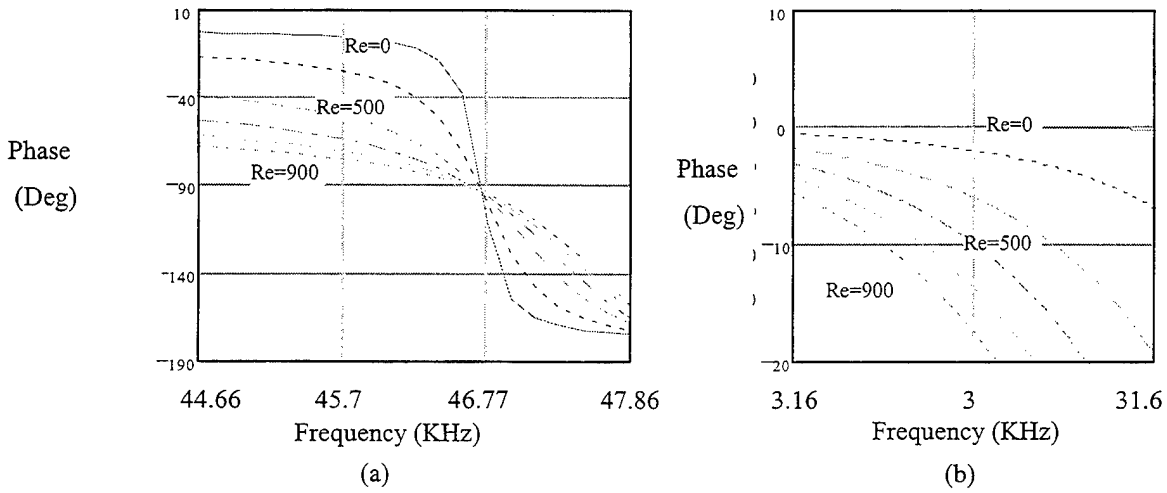


Figure 2.18: Effect of Re on the Phase of the PZT Response (a) Overall Response and (b) Response Around 10 KHz.

$$N(s) = \frac{s^2 + \omega_z^{\,2}}{s^2 + (\frac{\omega_o}{Q})s + \omega_0^{\,2}} \tag{2.48}$$

At the notch frequency ω_z the magnitude of the transfer function has a value of zero. At high and low frequencies the asymptotic values are $|N(0)| = \omega_z^2 / \omega_o^2$ and $|N(\infty)| = 1$.

The relative values of ω_z and ω_o that give rise to the highpass, regular and lowpass notch characteristics shown in Figure 2.19 are given by

$$\frac{\omega_z}{\omega_o} < 1 \tag{2.49}$$

$$\frac{\omega_z}{\omega_o} = 1 \tag{2.50}$$

$$\frac{\omega_z}{\omega_o} > 1 \tag{2.51}$$

Since the purpose of the notch filter in this feedback loop is only to suppress the resonant peak of the PZT, the desired notch filter for this application is the regular notch. Therefore, the transfer function of the desired notch filter is given by:

$$N(s) = \frac{s^2 + \omega_o^{\,2}}{s^2 + (\frac{\omega_o}{Q})s + \omega_0^{\,2}} \tag{2.52}$$

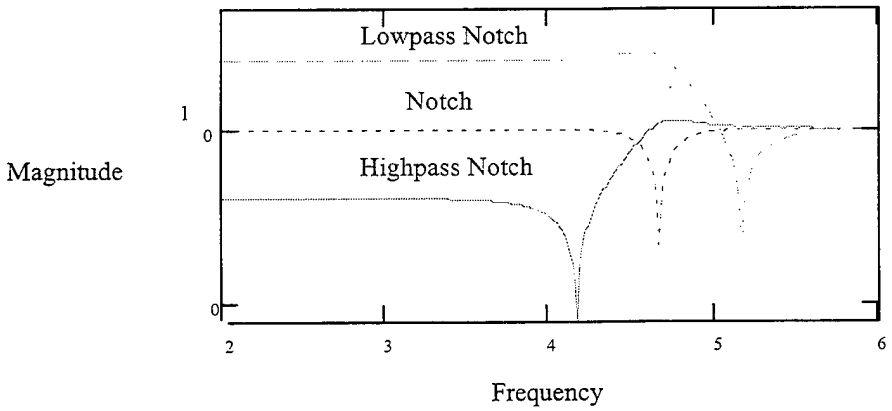


Figure 2.19: Magnitude Plot of Different Notch Filter Characteristics.

From the PZT model presented previously it is known that a PZT of $0.125 \times 1 \times 0.75$ in. connected in series with a 510Ω resistance has a resonance of approximately 28 dB at 46.77 KHz. Hence, the notch frequency should be set at 46.77 KHz. The Q of the notch poles, in terms of the desired attenuation, is defined as [19]:

$$Q = \frac{\omega_o}{BW\sqrt{10^{\wedge 0.1\alpha} - 1}} \tag{2.53}$$

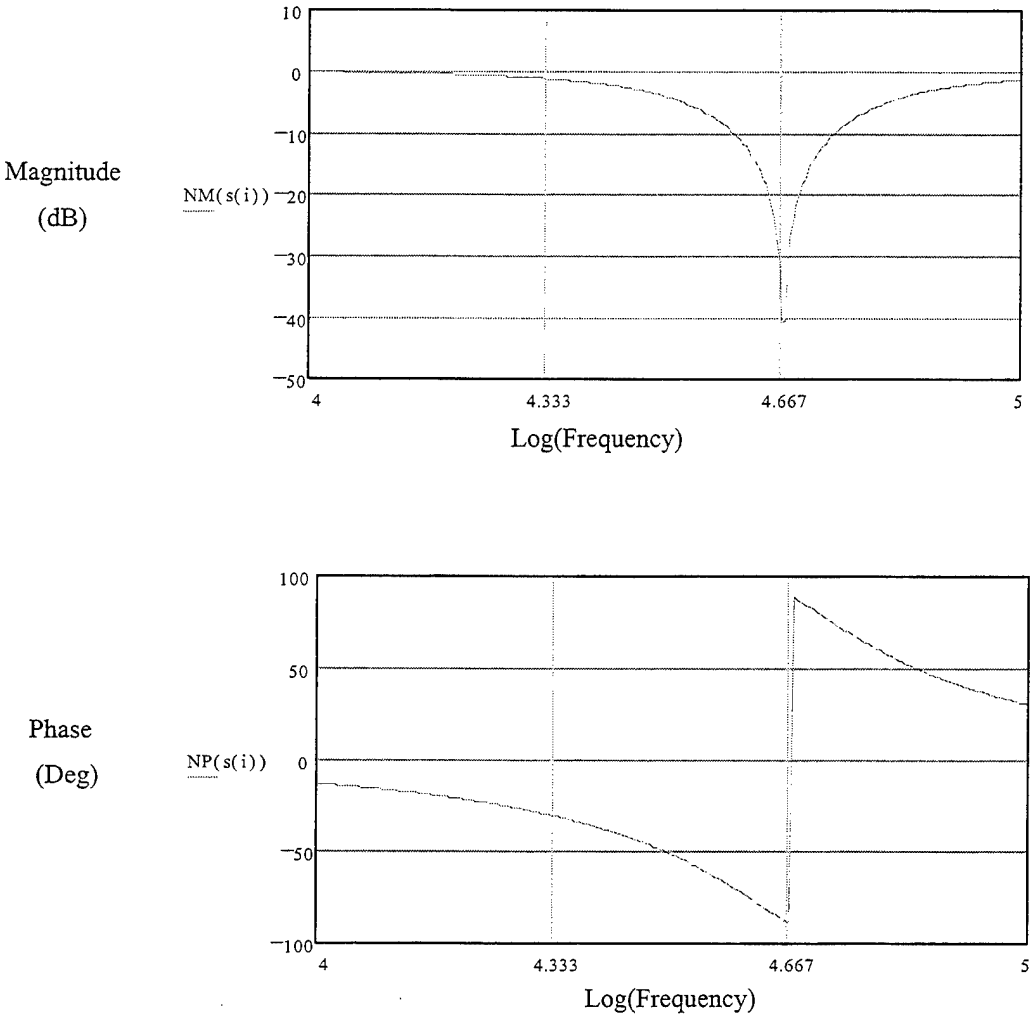


Figure 2.20:. Notch Filter Magnitude and Phase Response.

The frequency response of the Q-compensated PZT given by Figure 2.21 corresponds to the open-loop transfer function without any type of loop compensation; that is, in Equation (2.31) F(s) = 1. This frequency response shows that the system needs compensation to meet the specifications established previously. Therefore, a loop filter is necessary in the system to enhance the gain at low frequencies, and provide sufficient phase and gain margin to make the system stable. Other important function of the compensator is to maximize the overall system gain. Maximum system gain minimizes the system error, as shown by Equation (2.30).

Since a linear phase-locked loop model is a feedback network, classical linear control theory may be used to design the loop filter and analyze the closed-loop system performance. There are three figures of merit that have been used traditionally to measure closed-loop system performance: stability, bandwidth and error.

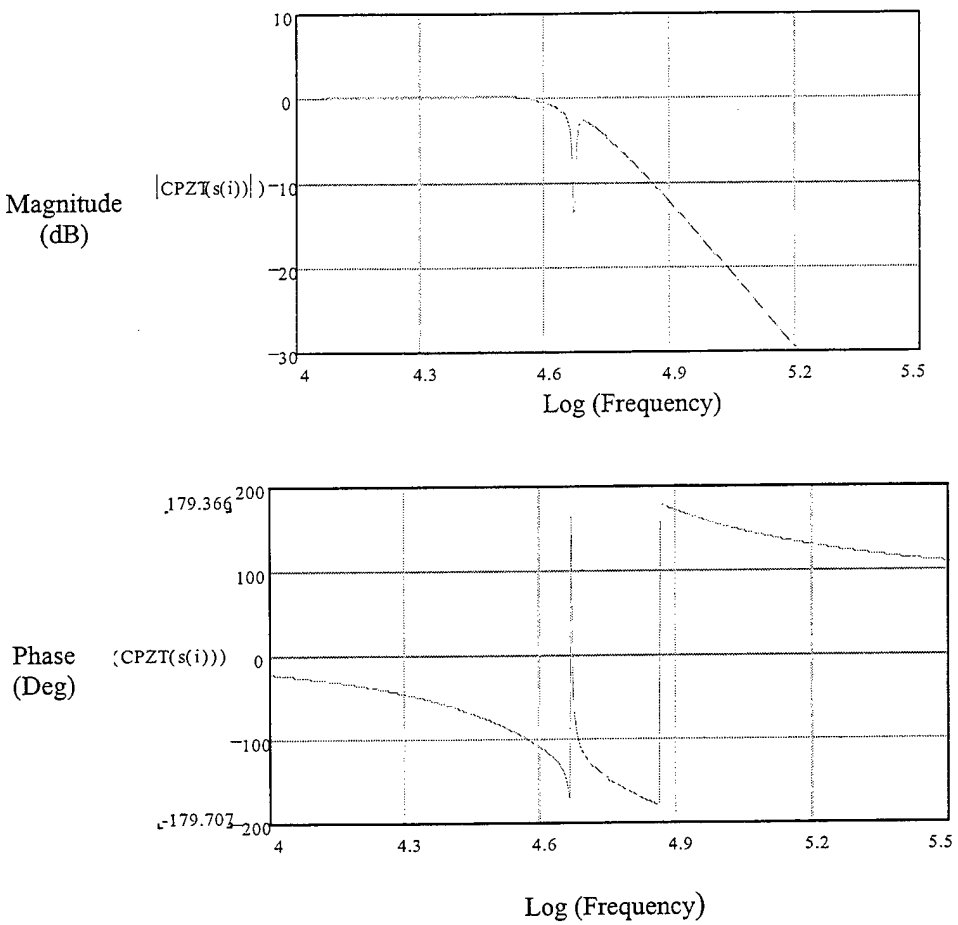


Figure 2.21: PZT-Re-Notch Combination Frequency Response.

A system is stable if all the poles of the closed-loop transfer function are in the left hand side of the s-plane. Establishing the locations of these poles and their migrating path, as the loop gain varies, gives insight into the behavior of the system. The Bode criterion for unconditional stability is that the gain must be less than 1 (0 dB) before the phase reaches −180° [Gardner 1979]. The phase margin is measured as the amount of phase needed to reach −180° when the gain is 1. A phase margin of 35° to 45° is considered to yield an acceptable degree of stability this also assures a gain margin of at least 1.

System error is a very important parameter in the loop analysis. The error is the normalized signal level at the output of the summing junction. The EOPLL can only support $\pm\pi/2$ phase error before it becomes unlocked, while at the same time, a small error is necessary for proper tracking. As discussed earlier, it is required that this control system operates with a phase error less than 10° over a frequency range from DC to 5000 Hz.

For the loop filter design presented here, the Q-compensated PZT, as given by Equation (2.55), is used to model the PZT transfer function. Loop filter design is a science and an art. Computer simulations were generated to verify the response of the loop filter design. In order to evaluate the effectiveness of the loop filter it is necessary to evaluate the performance of the EOPLL with all its elements in place in the open loop configuration. The approach taken for the loop filter design was to use several integrators combined with a single zero; the integrators provide a

great enhancement on the DC gain, whereas the zero helps to recoup some of the phase lost as a result of the two poles represented by the two integrators. Such a loop filter has a transfer function of the form:

$$F(s) = K_F \frac{\frac{s}{a}+1}{s^2} \tag{2.56}$$

where a is the zero location in rad/sec, and K_F is the loop filter gain.

The open-loop function with this type of loop filter is then given by:

$$G(s) = K \times CPZT(s)F(s) = K \times CPZT(s) \times \frac{\frac{s}{a}+1}{s^2} \tag{2.57}$$

where CPZT(s) is given by Equation (2.55), and K is the overall system gain as given by Equation (2.20). In this open-loop response there are two parameters that must be analyzed: K, the overall gain, and a, the zero location. These parameters have to be selected in such a manner that the stability and the gain requirements for the control system are met.

First, the effect of the zero location will be analyzed. Figure 2.22 shows the open loop response plots generated from Equation (2.57) for an overall gain equal to 1, and different zero locations. From the magnitude plots, Figure 2.22(a), it is seen that at frequencies below 3 KHz the zero location does not have any effect on the response, but at higher frequencies the zero location has an impact on the magnitude of the response. For example, at around 10 KHz there is a gain enhancement of about 10 dB when the zero is moved from 6000 rad/sec to 830 rad/sec. Although the gain improvement is desirable, the location of the zero has a more significant impact in the phase of the response. For stability reasons, it is desirable to have as much phase margin as possible; the phase plot, Figure 2.22(b), shows that as the zero moves further out in the frequency axis, the phase margin decreases. Therefore, the best zero location is 830 rad/sec, since this location yields the best phase margin, and at the same time provides some gain enhancement.

Figure 2.23 shows the open loop response plots generated from Equation (2.57) with four different values of the overall system gain K, and with the zero at 830 rad/sec, which is the location that provides the best phase margin. From the magnitude plot, Figure 2.14 (a), it is seen that the maximum loop gain tolerated by the system without becoming unstable is 2×10^8. Notice that the curve with higher gain reaches a phase of -180° before reaching unity gain; therefore, it does not meet the Bode criterion for unconditional stability.

The analysis of Figure 2.23 gave the maximum gain that the system can tolerate without becoming unstable with the zero fixed at 830 rad/sec, but it is important to know how stable the system is. To analyze stability it is necessary to quantize how far the phase is from −180° at the zero dB frequency crossover point of the magnitude response. This measurement is possible from Figure 2.23, but in order to obtain a more accurate value, the plots shown in Figure 2.24 were used. These plots correspond to the open-loop response with an overall gain of 2×10^8. Figure 2.24 (a) shows the magnitude of the response around 34.3 kHz; the frequency at which the phase is -180°.

where α is the required attenuation in dB, and BW is the band of frequencies for which α dB of attenuation is desired. Setting BW to 1.5 kHz and α to 30 dB, Equation (2.35) yields Q = 0.985 ≈ 1. The transfer function for the required notch to suppress the PZT resonance is:

$$N(s) = \frac{s^2 + (46.77 \times 10^3 \times 2\pi)^2}{s^2 + (46.77 \times 10^3 \times 2\pi)s + (46.77 \times 10^3 \times 2\pi)^2} \tag{2.54}$$

Magnitude and phase plots for the notch filter characterized by equation (2.54) are shown in Figure 2.20. It is important to notice that the notch filter only represents a phase penalty of approximately 15° at 10 kHz. Figure 2.21 demonstrates how the notch filter combined with a 510Ω series resistance effectively eliminates the PZT resonant peak with minimum phase loss. The Q-compensated PZT transfer function is then given by:

$$CPZT(s) = \frac{N(s)}{1 + (RmCm + \mathrm{Re}(Ce + Cm))s + (LmCm + \mathrm{Re}\,RmCeCm)s^2 + \mathrm{Re}\,LmCeCms^3} \tag{2.55}$$

where *Cm, Ce, Lm* and *Rm* are the PZT model parameters; *Re* is the 510Ω resistor used in conjunction with the notch filter to suppress the mechanical Q of the PZT. N(s) is the notch filter transfer function as given by Equation (2.54).

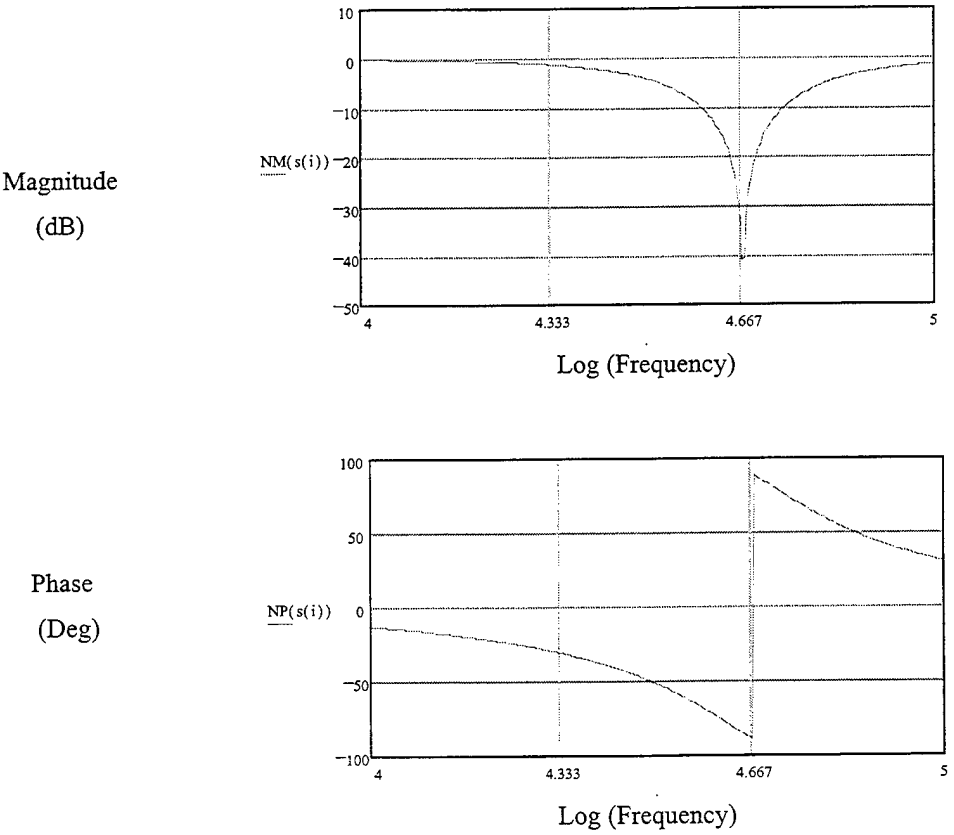


Figure 2.20: Notch Filter Magnitude and Phase Response.

The frequency response of the Q-compensated PZT given by Figure 2.21 corresponds to the open-loop transfer function without any type of loop compensation; that is, in Equation (2.31) F(s) = 1. This frequency response shows that the system needs compensation to meet the specifications established previously. Therefore, a loop filter is necessary in the system to enhance the gain at low frequencies, and provide sufficient phase and gain margin to make the system stable. Other important function of the compensator is to maximize the overall system gain. Maximum system gain minimizes the system error, as shown by Equation (2.30).

Since a linear phase-locked loop model is a feedback network, classical linear control theory may be used to design the loop filter and analyze the closed-loop system performance. There are three figures of merit that have been used traditionally to measure closed-loop system performance: stability, bandwidth and error.

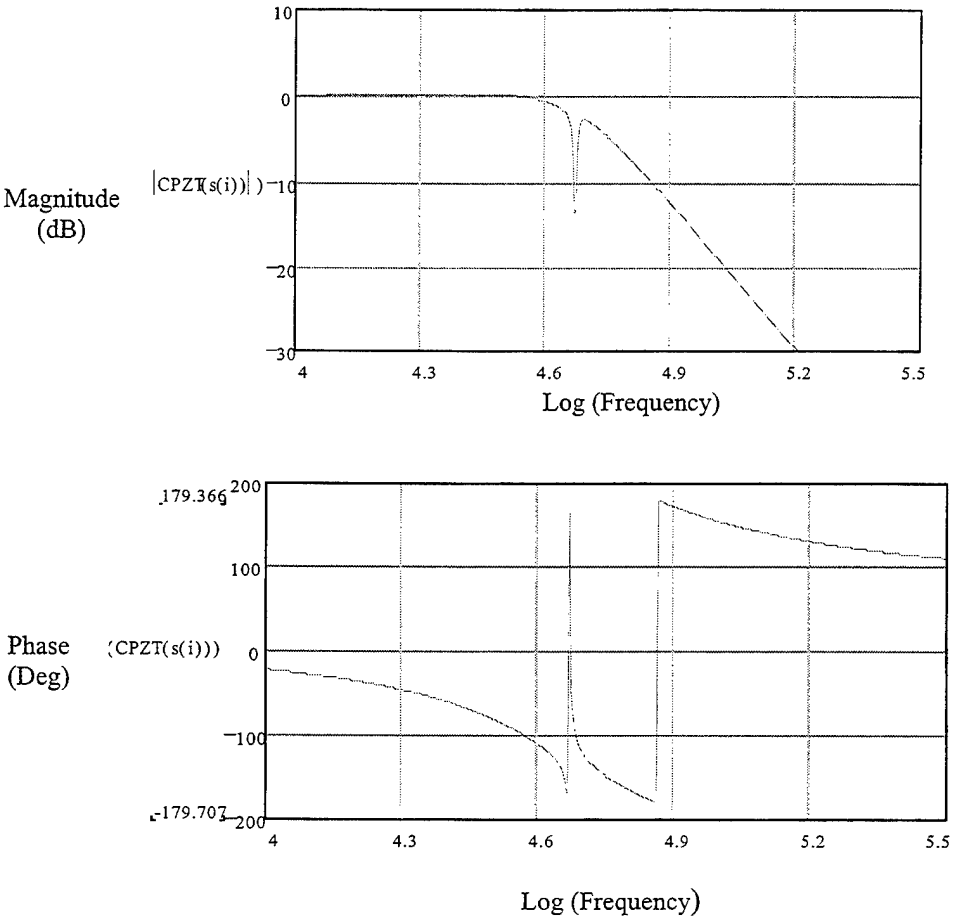


Figure 2.21: PZT-Re-Notch Combination Frequency Response.

A system is stable if all the poles of the closed-loop transfer function are in the left hand side of the s-plane. Establishing the locations of these poles and their migrating path, as the loop gain varies, gives insight into the behavior of the system. The Bode criterion for unconditional stability is that the gain must be less than 1 (0 dB) before the phase reaches −180° [Gardner 1979]. The phase margin is measured as the amount of phase needed to reach −180° when the gain is 1. A phase margin of 35° to 45° is considered to yield an acceptable degree of stability this also assures a gain margin of at least 1.

System error is a very important parameter in the loop analysis. The error is the normalized signal level at the output of the summing junction. The EOPLL can only support $\pm\pi/2$ phase error before it becomes unlocked, while at the same time, a small error is necessary for proper tracking. As discussed earlier, it is required that this control system operates with a phase error less than 10° over a frequency range from DC to 5000 Hz.

For the loop filter design presented here, the Q-compensated PZT, as given by Equation (2.55), is used to model the PZT transfer function. Loop filter design is a science and an art. Computer simulations were generated to verify the response of the loop filter design. In order to evaluate the effectiveness of the loop filter it is necessary to evaluate the performance of the EOPLL with all its elements in place in the open loop configuration. The approach taken for the loop filter design was to use several integrators combined with a single zero; the integrators provide a great enhancement on the DC gain, whereas the zero helps to recoup some of the phase lost as a result of the two poles represented by the two integrators. Such a loop filter has a transfer function of the form:

$$F(s) = K_F \frac{\frac{s}{a} + 1}{s^2} \tag{2.56}$$

where a is the zero location in rad/sec, and K_F is the loop filter gain.

The open-loop function with this type of loop filter is then given by:

$$G(s) = K \times CPZT(s)F(s) = K \times CPZT(s) \times \frac{\frac{s}{a} + 1}{s^2} \tag{2.57}$$

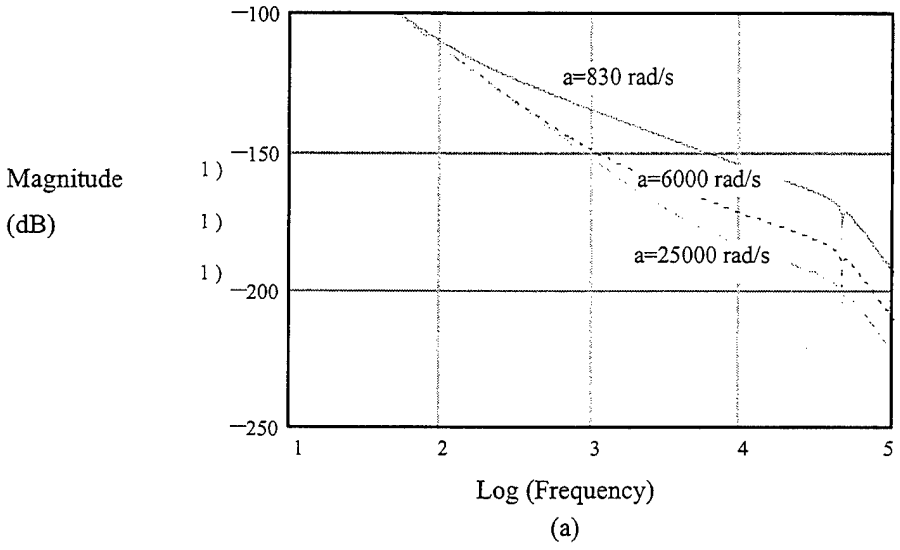
where CPZT(s) is given by Equation (2.55), and K is the overall system gain as given by Equation (2.20). In this open-loop response there are two parameters that must be analyzed: K, the overall gain, and a, the zero location. These parameters have to be selected in such a manner that the stability and the gain requirements for the control system are met.

First, the effect of the zero location will be analyzed. Figure 2.22 shows the open loop response plots generated from Equation (2.57) for an overall gain equal to 1, and different zero locations. From the magnitude plots, Figure 2.22(a), it is seen that at frequencies below 3 KHz the zero location does not have any effect on the response, but at higher frequencies the zero location has an impact on the magnitude of the response. For example, at around 10 KHz there is a gain enhancement of about 10 dB when the zero is moved from 6000 rad/sec to 830 rad/sec. Although the gain improvement is desirable, the location of the zero has a more significant impact in the phase of the response. For stability reasons, it is desirable to have as much phase margin as possible; the phase plot, Figure 2.22(b), shows that as the zero moves further out in the frequency axis, the phase margin decreases. Therefore, the best zero location is 830 rad/sec, since this location yields the best phase margin, and at the same time provides some gain enhancement.

Figure 2.23 shows the open loop response plots generated from Equation (2.57) with four different values of the overall system gain K, and with the zero at 830 rad/sec, which is the location that provides the best phase margin. From the magnitude plot, Figure 2.14 (a), it is seen that the maximum loop gain tolerated by the system

without becoming unstable is 2×10^8. Notice that the curve with higher gain reaches a phase of -180° before reaching unity gain; therefore, it does not meet the Bode criterion for unconditional stability.

The analysis of Figure 2.23 gave the maximum gain that the system can tolerate without becoming unstable with the zero fixed at 830 rad/sec, but it is important to know how stable the system is. To analyze stability it is necessary to quantize how far the phase is from −180° at the zero dB frequency crossover point of the magnitude response. This measurement is possible from Figure 2.23, but in order to obtain a more accurate value, the plots shown in Figure 2.24 were used. These plots correspond to the open-loop response with an overall gain of 2×10^8. Figure 2.24 (a) shows the magnitude of the response around 34.3 kHz; the frequency at which the phase is -180°. Figure 2.24 (b) shows the phase of the open loop response around 19.2 kHz; the frequency at which the magnitude of the response is 1 or 0 dB.



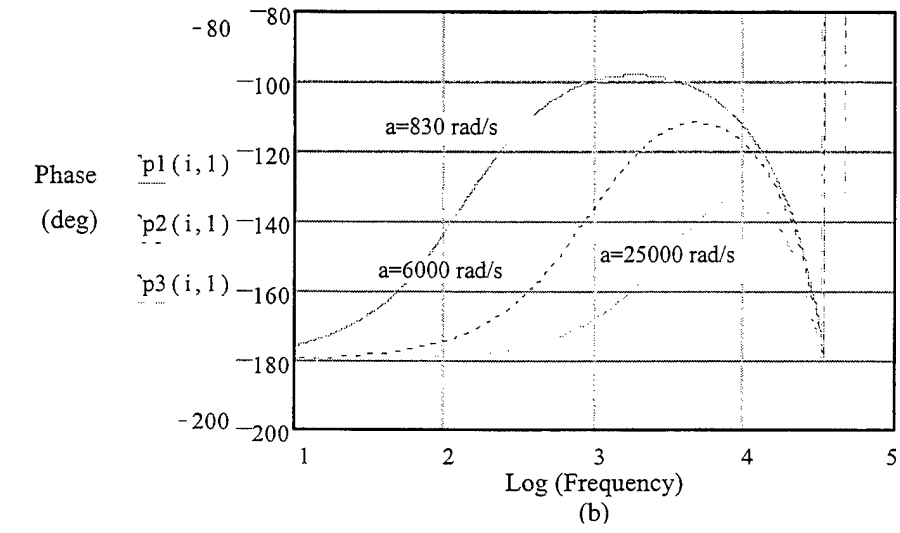
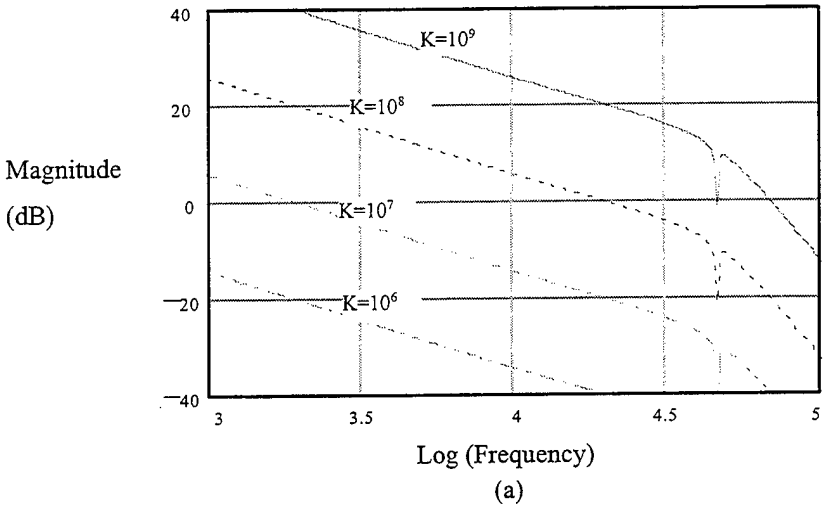


Figure 2.22: Effect of the Zero Location on the Open-Loop Response (a) Magnitude Response and (b) Phase Response.



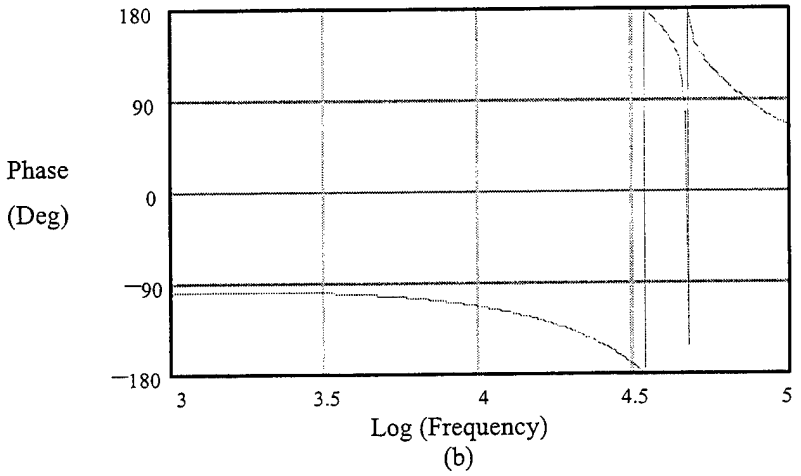


Figure 2.23: Effect of the Total Gain on the Open-Loop Response
(a) Magnitude Response and (b) Phase Response.

The gain margin, as shown in Figure 2.24 (a), is approximately 5 dB; this result means that the overall system gain can still be increased by a factor of 1.7 before the system becomes unstable. The phase plot from Figure 2.24(b) yields a phase margin of approximately 45°. The measurements from Figure 2.24 show that the system meets the requirements for unconditional stability with a maximum overall gain of 2×10^8. It is also important to verify that the control loop meets the design specifications concerning gain factors at given frequencies, as established by Equations (2.33) and (2.34). Figure 2.23 (a) shows a gain of approximately 25 dB at a frequency of 1 KHz. This gain factor is greater than the required gain of 10 dB at 1 KHz given by Equation (2.34).

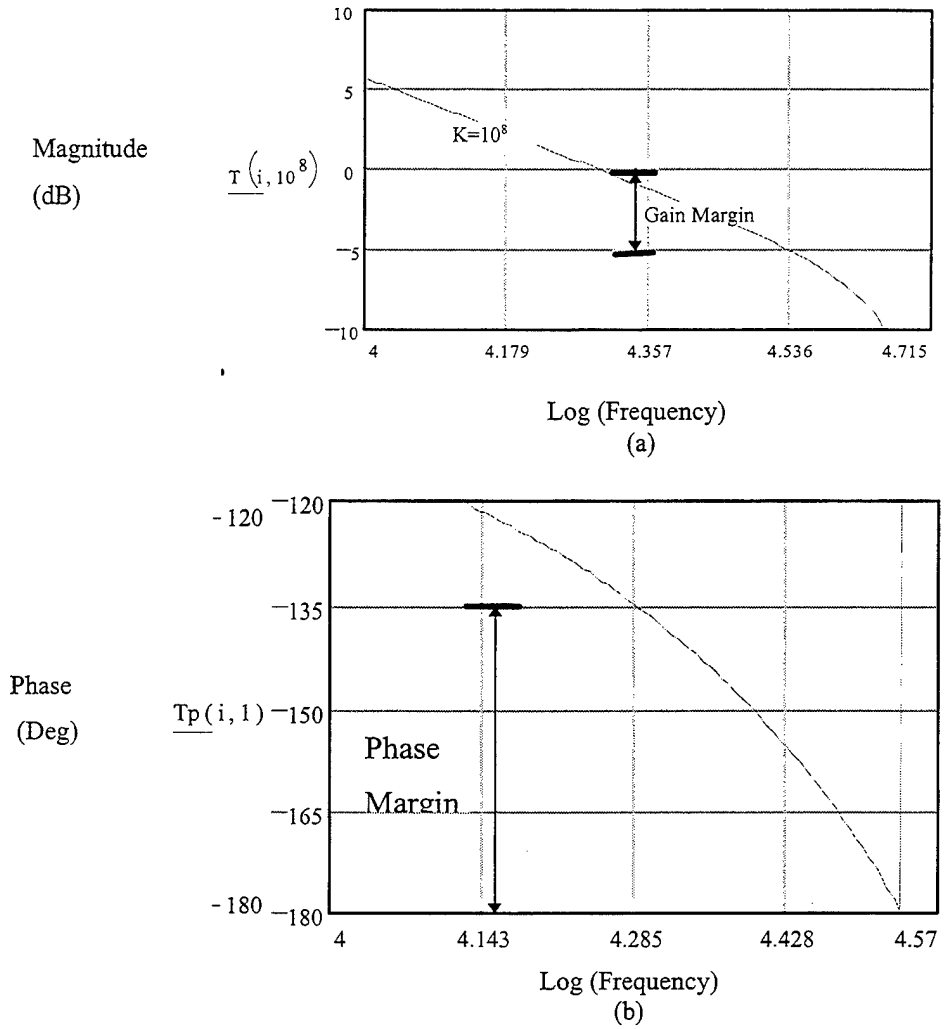


Figure 2.24: Loop Stability Measurement (a) Gain Margin and (b) Phase Margin.

Figures 2.25 and 2.26 give the magnitude of the closed-loop transfer function and the magnitude of the error function as a function of frequency; these functions are given by Equations (2.26) and (2.29), respectively. The closed-loop plots show how the closed-loop response stays constant at 0 dB or 1, and has a 3 dB bandwidth of approximately 30 kHz. The error plots show how as the frequency increases the error also increases. This behavior is expected, since the error function is equal to one minus the closed-loop function, as given by Equation (2.29). The logarithmic plot of the error function illustrates the advantage of using an integrator in the closed loop to reduce the phase error to zero as the frequency approaches zero. It is important to add that the closed-loop function, as well as the error function, do not yield any information about the stability of the control system.

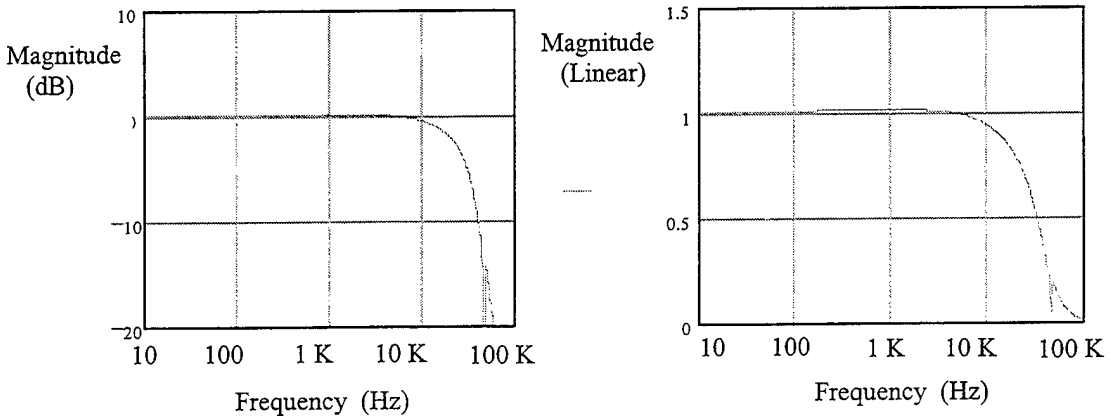


Figure 2.24: Closed-Loop Response Magnitude.

In summary, the simulation shows that the control system designed meets the specifications, as well as the unconditional stability requirements. The compensated loop is characterized by the parameters given in Table 2.3.

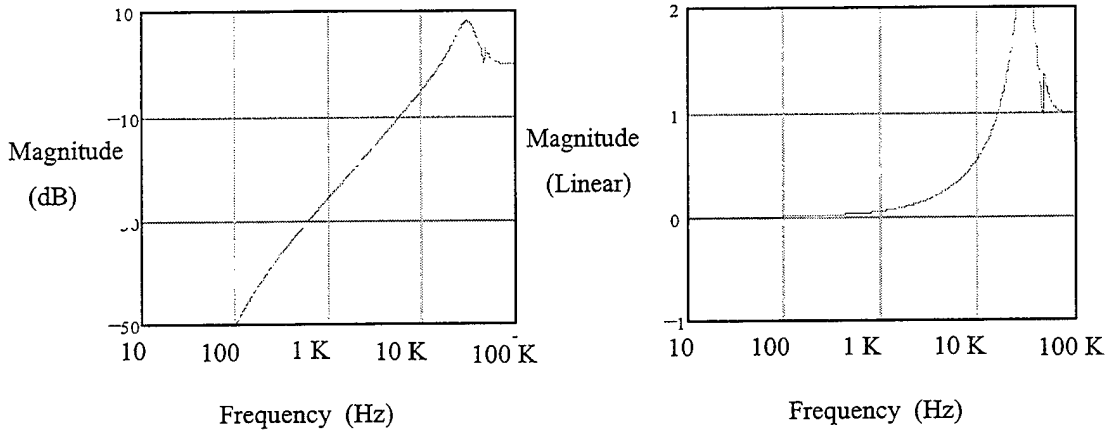


Figure 2.26: Error Function Magnitude.

Table 2.3: Simulated EOPLL Parameters.

Parameter	Value
Maximum overall gain	2×10^8
Gain margin	5 dB
Phase margin	45°
Gain at 1 kHz	30 dB
Loop Filter Zero Location	830 rad / sec

2.2.4. Low-cost InGaAs Photoreceiver Design

The photoreceiver (detector/pre-amplifier) is a critical component of the lidar system. Further, for coherent arrays in excess of 10 apertures, it is important to minimize the cost of each subsystem. For these two reasons CREOL investigated the design and development of a low-cost photoreceiver with the following requirements: high detector quantum efficiency, moderate bandwidth (BW > 150 MHz), moderate transimpedance gain, and moderate-to-low noise.

The primary purpose of the photodetector in the EOPLL is to transform the optical signal from the LIDAR into an RF electronic signal that is then processed by the control system. The current generated in the photodetector by the optical signal is proportional to the product of the LIDAR signal field and the optical LO field, as given by Equation (2.1). In the process of mixing the received signal with the LO, using a $2 \rightarrow 1$ fiber optic coupler, the two mixed optical signals that are produced that are 180° out of phase. Hence a balanced detector provides the differential detection required to combine both signals together removing common mode noise. As a result, a balanced detector increases the signal-to-noise ratio by 3 dB. The system requirements for the balanced photo-detector and differential pre-amplifier are summarized in Table 2.4.

Table 2.4: Photo-Detector and Pre-Amplifier Requirements.

Photodetector	Differential Amplifier
Responsivity:> 0.5 A/W @ 1.06 μm	Transimpedance gain: 1000Ω
Bandwidth: >500 MHz	Swing into 50 Ω: +2V
Dark current: < 10 nA	Bandwidth: 80 MHz

The photodetector selected for the CREOL coherent LIDAR array receiver was the ETX–100T from EPITAX. This photodetector has been designed for high speed optical communications and sensing applications [20]. It features high responsivity between 850 and 1700 nm, low dark current and low capacitance. The selection of this photodetector was based upon the characteristics given on Table 2.5.

Table 2.5: Photodiode ETX–100T Specifications.

Responsivity	Bandwidth	Dark Current	Rise Time
0.9 A/W@ 1300 nm	105 MHz	0.3 nA	180 ps

The wide band operational amplifier CLC425 from Comlinear Corporation was selected for the pre-amplification stage since it was designed for use as a photo-diode transimpedance amplifier [21]. Some features of the CLC425 are given in Table 2.6. Figure 2.27 shows the complete schematics for the balanced photo-detector pre-amplifier. The final board layout and implementation of this circuit required that a good ground plane be formed to prevent any undesired high frequency oscillations.

Table 2.6: CLC425 Comlinear Amplifier Specifications.

Gain-bandwidth product	1.7 GHz
Input voltage noise	1.05 nV /√Hz
Input current noise	1.6 pA /√Hz
Slew rate	350 V /μs
Gain range	±10 - ±1000

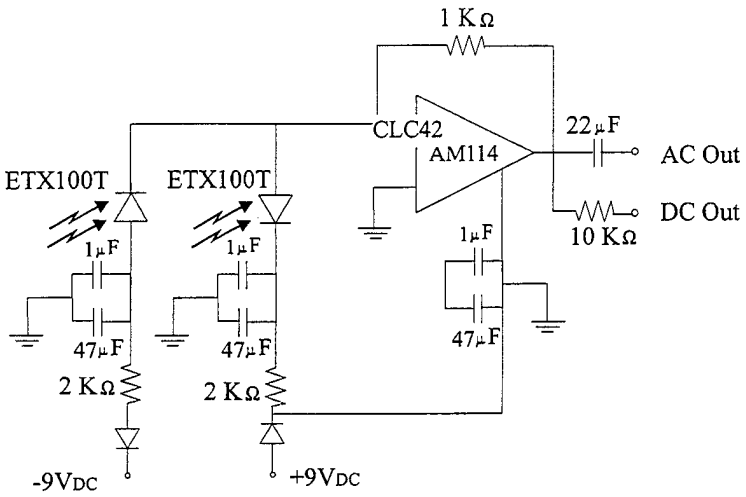


Figure 2.26: Photodetector Pre-Amplifier Schematics.

2.2.5. Heterodyne Efficiency: "Where are the Missing dBs"

Throughout the history of coherent laser radar, there have been discrepancies between predicted and experimentally measured CNR. Typical discrepancies are on the order of three dB to ten dB. There are potentially many contributors to these discrepancies, including unaccounted losses in the transceiver, atmosphere and the target; inaccurate estimates of the heterodyne detection efficiency, γ, (also known simply as the heterodyne efficiency). In this section of the report, the results of our efforts to measure the mixing efficiency, γ_m, which is proportional to the heterodyne efficiency are described and compared to the theoretical limit.

2.2.5.1 Theoretical Background

The heterodyne efficiency is the ratio of the mean power in the AC photocurrent to the mean AC power in an ideal case when (1) the aperture-collected signal power, P_s, is completely received by the detector and (2) the local oscillator and the signal fields (U_{lo} and U_s) are spatially matched. For the case when both the local oscillator and the received signal fields are co-polarized, the heterodyne efficiency is given by [23]

$$\gamma = \frac{\left| \int_{A_d} U_{lo}(A)\, U_s^*(A) dA^2 \right|^2}{P_{lo} P_s}, \tag{2.56}$$

where A_d is the detector area. The heterodyne efficiency can also be rewritten in the following form

$$\gamma = \frac{P_d}{P_s} \frac{\left| \int_{A_d} U_{lo}(A)\, U_s^*(A) dA^2 \right|^2}{P_{lo} P_d} = \gamma_{so} \gamma_m, \tag{2.57}$$

where P_d is the portion of the signal power which is collected by the detector, γ_m, is the mixing efficiency, and γ_{so} is the *spillover efficiency*. Thus the heterodyne efficiency is the product of two quantities: (1) a power loss term which captures the effects of signal spillover and (2) a mixing efficiency term which represents the spatial mismatch of the signal and local oscillator fields. The mixing efficiency, which should be unity for perfectly matched signal and LO fields, is often the source of many lost dBs. This is especially true for transceivers which are subject to thermally and mechanically induced misalignments.

It can be shown that the mixing efficiency is the ratio of half the mean electrical AC power ($V^2rms^{/2}$) and the product of the mean DC local oscillator and signal photovoltages, provided the photoreceiver produces an equal response to both AC and DC signals. That is

$$\gamma_m = V^2_{rms}/2V_{lo}V_s. \tag{2.58}$$

2.2.5.2 Experiment Description

In our earliest attempts to characterize a coherent receiver's performance, a bulk optics testbed was designed to measure the effects of target speckle, misalignment, and defocus on the mixing efficiency. However, for the bulk optics case, even under ideal conditions (a specular target located at the focus of the transceiver), it was found that the best efficiency that could be obtained was on the order of 25%. This was postulated to be due to a beam diameter mismatch, as the two arms of the interferometer were unequal. In an attempt to remove this uncertainty a single-mode fiber testbed was designed and measurements were made of the efficiency under several polarization conditions.

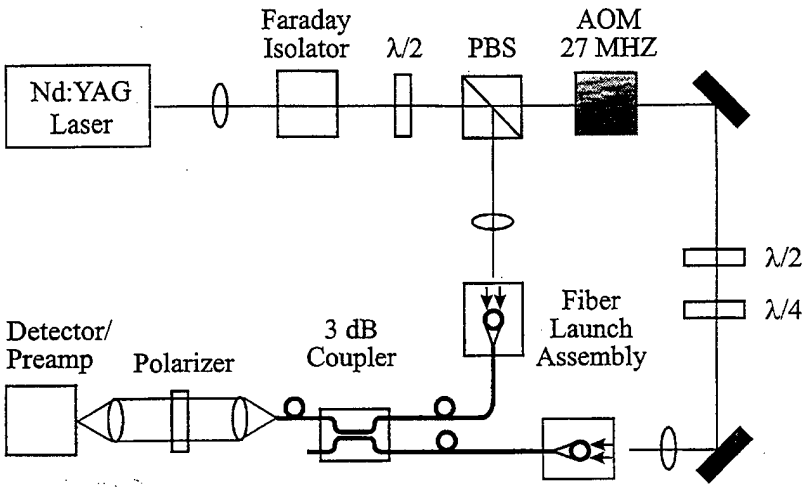


Figure 2.28: Single mode fiber-optic heterodyne-mixing efficiency measurement testbed.

In this experiment, the local oscillator field and a 27 MHz frequency-shifted signal field were each launched into the two input ports of a 3 dB single-mode fiber optic coupler. Half and quarter wave plates were utilized to match the polarization state of the signal field with that of the LO field by maximizing the AC signal amplitude. Upon exiting the fiber, the mixed field was collimated, polarized and focused onto a photoreceiver. In a separate experiment, the 3 dB frequency of the photoreceiver was measured to be approximately 200 MHz with no discernible attenuation at 27 MHz (note, a first order linear circuit model would predict a voltage attenuation factor of 0.991 or -0.08 dB at 27 MHz, which can be considered negligible). This measurement was conducted by frequency shifting one stable laser relative to another to trace out the spectral response using the "peak-hold" function of the RF Spectrum Analyzer (HP8560A). Unfortunately, since both the RF Spectrum Analyzer and the AC output of the photoreceiver inhibit DC signals, the DC response could not be compared to the AC response. However, an analysis of the photoreceiver's circuit (Fig. 2.27), indicates equal gain for the AC and DC signal outputs.

2.2.5.3 Results and Discussion

Several measurements were made with and without the polarizer in place. On average the mixing efficiency measured with the polarizer was about 50% (-3 dB), whereas, without the polarizer, the mixing efficiency was typically much lower (c.f., Table 2.7). The additional reduction in mixing efficiency obtained without the polarizer indicates that the fiber, and/or the fiber-coupler, introduced significant depolarization of the two fields.

Table 2.7: Experimental Measurements of the Heterodyne-mixing Efficiency for the Single-Mode Fiber-Optic Testbed.

	w/o Polarizer	w/ Polarizer
Measured DC LO photo-voltage	213.5 mV	89 mV
Measured DC signal photo-voltage	56 mV	31.5 mV
Measured RMS carrier amplitude	94.5 mV	53 mV
Calculated heterodyne efficiency	61.1%	70.8%

The 3 dB loss, for what was thought to be ideal conditions, is troubling and exemplifies the typical discrepancies between theoretical CNR predictions and experimental measurements. For this experiment, the discrepancy can probably be attributed to a combination of many factors, the most likely of which is spatial depolarization of the beam resulting from random spatial birefringence in the fiber. To further explore such losses, the experiment should be repeated with polarization-preserving single-mode fibers.

2.2.6 Field Experimental Test of Two-Aperture System

2.2.6.1 Experiment and Test Design and Layout

A real-time, dual-aperture, continuous wave, equal-gain coherent array was tested at the BMDO Innovative Science and Technology Experimentation Facility (ISTEF) at the Kennedy Space Center. An optical schematic of the testbed is shown in Fig. 2.7. The offset-homodyne transceiver consisted of two separated apertures coupling the received signals into single mode optical fibers. Phase compensation was accomplished by an electronic servo loop utilizing a fiber-wrapped cylindrical PZT as the active phase modulator. The laser source was a commercially available, narrow linewidth (5 kHz), 50 mW, Nd:YAG non-planar ring laser (Lightwave Electronics 122-1064-50). An acousto-optic modulator in the transmitter leg provided a 27.12 MHz frequency offset from the local oscillator. The two one cm effective diameter receiver apertures were positioned 25 cm apart and aligned to an extended target located one km down a horizontal range. The separation distance ensured that the receivers were sensing spatially-independent samples (speckles) of the return field, since for this particular transceiver/target configuration the speckle diameter should have been equal to the transmitter diameter (i.e., one cm). The target was constructed from a ¼ inch thick aluminum plate covered with 3M reflectance tape (3M #7610; see Chapter III for a discussion of the properties).

2.2.6.2 Results

Ten 900 ms records of the IF signals (channel outputs 1 and 2, and their coherent sum, i.e., electro-optically co-phased and summed) were digitized and subsequently processed to extract the significant statistics. These statistics for the strong signal case, summarized in Table 2.8, demonstrate a near factor of two increase in the mean CNR and a near ½ reduction in fading strength which is commensurate with theoretical predictions. The less-than-excellent agreement for the weak signal case was probably due to an increased sensitivity to noise and signal measurement errors under low CNR conditions. The traces in Fig. 2.29, which are 50 ms segments of the digitized

IF signals, demonstrate this improvement. For an incoherent sum, no improvement is obtained for the mean CNR and normalized variance as theory would suggest.

Table 2.8: CNR Statistics Calculated from Measurements of the Signal plus Noise and the Noise Alone

	CNR Statistics			
	Strong Signal		Weak Signal	
	Mean CNR	Normalized Variance	Mean CNR	Normalized Variance
Channel 1	22.9	1.56	1.07	1.34
Channel 2	33.3	1.39	1.48	1.13
Coherent Sum - Exp.	50.6	0.78	1.70	0.88
Theory	50.2	0.77	2.27	0.65

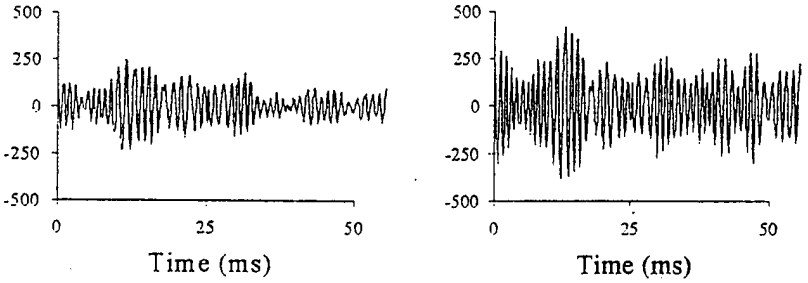


Fig 2.29: Typical traces of the IF signals: left trace is from a single channel; right trace is the coherent sum of the two channels.

For the first time we have demonstrated the operation of an adaptive, coherent-combining laser receiver. We have shown that the mean CNR will increase nearly linearly with the number of array elements M and the noramlized CNR variance will decrease at a rate close to 1/M, in accordance with theory.

2.3. Conclusions

The multiaperture technology developed under this program has lead to the design, implementation and testing of a two-aperture transceiver that was demonstrated in field tests to be capable of removing the atmospheric phase fluctuations induced between the two channels. The analysis has shown that it is not cost effective to use the maximal ratio receiver in comparison to the equal gain receiver. The small CNR improvement of the maximal ratio receiver over the equal gain receiver does not justify the additional hardware complexity required to implement the maximum ratio receiver.

The experience obtained with the two-aperture system over the 1 km range at ISTEF provided valuable insight into the design of the eight-element receiver. This lead to a complete analysis and design of the PZT to be used as the phase modulator for the eight-element system. Improvements were made to the control system to increase its frequency response by an order of magnitude for the eight-element system as compared to the two-element system. Predictions show that the eight-element system will be able to maintain a phase lock between all channels and will not exceed more that 10° of phase error over the 5000 Hz fluctuation bandwidth of the

atmospheric turbulence. It is expected that the eight element system would easily meet the design specifications. While it did, this will be the subject of reports on the follow-on contract.

2.4. References

1. Jelalian, A. V., [1992], *Laser Radar Systems*. Norwood: Artech House.
2. Katzman, M., [1987], *Laser Satellite Communications*. Englewood Cliffs: Prentice Hall.
3. Nordstrom, R., and L. J. Berg, [1990]. "Coherent Laser Radar: Techniques and Applications," *Lasers and Optronics,* June, pp.51-56.
4. Banakh, V., and V. L. Mironov, [1987], *Lidar in Turbulent Atmosphere*. S. Chomet, trans. Norwood: Artech House.
5. Gatt, P., T. P. Costello, D. A. Heimmerman, D. C. Castellanos, A. R. Weeks, and C. M. Stickley, [1996], "Coherent Optical Array Receivers for the Mitigation of Atmospheric Turbulence and Speckle Effects," *Applied Optics*, 5999-6009, (Oct. 20, 1996).
6. Gatt, P., T.P. Costello, and D. H. Gantala, [1995], "Dual Aperture Measurements of Two Km (round-trip) Intensity and Phase Perturbations One Micron Laser Speckle Propagating Through a Turbulent Atmosphere," Proc. SPIE 2471, 254-258.
7. Kin Pui Chan, and D. K. Killinger, [1991], "Enhanced Detection of Atmospheric-Turbulence-Distorted 1-μm Coherent Lidar Return Using a Two-Dimensional Heterodyne Detector Array," *Optics Letters*, Vol. 16, No. 16, August, pp. 1219-1221.
8. Enloe, L. H., and J. L. Rodda, [1965], "Laser Phase-Locked Loop," *Proceedings of the_IEEE,* February, pp. 165-166.
9. Travis, C., [1935], "Automatic Frequency Control," *IRE Proceedings,* Vol. 23, pp. 1125-1141.
10. Kane, T. J., and A. P. Cheng, [1988], "Fast Frequency Tuning and Phase-Locking of Diode-Pumped Nd:YAG Ring Lasers," *Optic Letters,* Vol. 13, No. 11, November, pp. 970-972.
11. Day, T., A. D. Farinas, and R. L. Byer, [1990], "Demonstration of a Low Bandwidth 1.06 μm Optical Phase-Locked Loop for Coherent Homodyne Communication." *IEEE Photonics Technology Letters*, Vol. 2, No. 4, April, pp. 294-296.
12. Kazovky, L. G., and B. Jensen, [1990], "Experimental Relative Frequency Stabilization of a Set of Lasers Using Optical Phase-Locked Loops," *IEEE Photonics Technology Letters,* Vol. 2, No. 7, July, pp. 516-518.
13. Neubwert, W. M., K. H. Kudielka, W. R. Leeb, and A. L. Scholtz, [1994], "Experimental Demonstration of an Optical Antenna for Laser Space Communications," *Applied Optics,* Vol. 33, No. 18, June, pp. 3820-3830.
14. Jackson, D. A., R. Priest, A. Dandridge, and A. B. Tveten, [1980], "Elimination of Drift in a Single-Mode Optical Fiber Interferometer Using a Piezoelectrically Stretched Coiled Fiber," *Applied Optics.*Vol. 19, No. 17, September, pp. 2926-2929.
15. Gardner, F. .M., [1979], *Phase-Lock Techniques*. New York: John Wiley and Sons.
16. Ziemer, R. E., and W. H. Tranter, [1995], *Principles of Communications. Systems, Modulation, and Noise.* Boston: Houghton Mifflin Company.
17. Castellanos, D., "Analysis and Design of an Electro-Optic Phased-Locked Loop for Use in a Coherent Laser Array Receiver," Masters Thesis, University of Central Florida, Fall 1995.
18. Morgan Matroc, INC., [1994], *Guide to Modern Piezo-electric Ceramics*. Bedford: Morgan Matroc, INC.
19. Valkenburg, M. E., [1982], *Analog Filter Design.* New York: CBS College Publishing.
20. Epitax Opto-Electronic Devices, [1993], *Epitax Data Book.* West Trenton: Epitax Opto-Electronics.
21. Comlinear Corporation, [1995], *Solutions with Speed.* Fort Collins: Comlinear Corporation.

22. Robert Mongeon, private communication.

23. Zhoa, Y., M. J. Post, and R. M. Hardesty, "Receiving Efficiency of Monostatic Pulsed Coherent Lidars, 1: Theory", Appl. Opt. **29**, 4111-4119 (1990).

2.5 Publications, Presentations and Theses

1. P. Gatt, T. P. Costello, D. A. Heimmermann, D. C. Castellanos, A. R. Weeks, and C. M. Stickley, "Coherent Optical Array Receivers for the Mitigation of Atmospheric Turbulence and Speckle Effects", to be published in Appl. Opt., 35, 5999-6009, (October 20, 1996).
2. P. Gatt, T. Costello, and D. Gantala, "Dual Aperture Measurements of Two km (round trip) Amplitude and Phase Perturbations of One Micron Laser Speckle Propagating Through a Turbulent Atmosphere", Proc. SPIE 2471, 254-258, (1995).
3. C. M. Stickley, R. Phillips, A. Weeks, and P. Gatt, "Demonstration of an Adaptive Coherent-Combining Laser Radar Receiver", to be published by Springer in the Proceedings of the 18th International Laser Radar Conference, Berlin, Germany, July 1996.
4. C. M. Stickley, A. R. Weeks, P. Gatt, T. Costello, and D. Heimmermann, "Coherent Optical Array Receivers for Mitigation of Atmospheric Turbulence and Speckle Effects", to be published in Proceedings of the 2nd NATO/IRIS Joint Symposium, London, England, June 1996.
5. C. M. Stickley, "Overview of CREOL Programs in Laser Radar", Proc. SPIE 1936, 147-156, (1993).
6. C. M. Stickley, P. Gatt, D. Heimmermann, and W. Perez, "Heterodyne Laser Radar Array Receiver for the Mitigation of Target-Induced Speckle", Proc. SPIE 1936, 157-164, (1993).
7. P. Gatt, T. Costello, and D. Castellanos, "CREOL's Real-Time Dual Aperture CW Nd: YAG Coherent Array", SPIE/AeroSense, Orlando, Florida, April 1995.
8. P. Gatt, A. R. Weeks, T. Costello, D. C. Castellanos, and C. Gagge, "A Design Approach Towards and Prototype for a Multi-aperture Coherent Array Receiver", LASERS '94, Quebec, Quebec, Canada, December 1994.
9. C. M. Stickley, P. Gatt, J. Harvey, A. Weeks, H. Myler and R. Mongeon, "CREOL Laser Radar Activities at BMDO's ISTEF", presented at the 17th Intl. Laser Radar Conference, Sendai, Japan, July 1994.
10. C. M. Stickley, P. Gatt, W. Perez, and D. Heimmermann, "Coherent Laser Radar Array Receivers: Theory & Experiment", presented at the 7th Conference on Coherent Laser Radar, Paris, France, July 1993.
11. C. M. Stickley, P. Gatt and W. Perez, "Application of Space Diversity Receiver Techniques to Coherent Laser Radar", Annual IEEE/LEOS Meeting, Post-Deadline Paper PD-8, Boston, Massachusetts, November 1992.
12. D. C. Castellanos, "Analysis and Design of an Electro-Optics Phase-Locked Loop for Use in a Coherent Laser Array Receiver", M.S. Thesis, University of Central Florida, November 1995.
13. D. Heimmermann, "Nd:YAG Frequency Stabilization Techniques", M.S. Thesis, University of Central Florida, December 1994.

2.6. Appendix 2.A

P. Gatt, T. Costello, and D. Gantala, "Dual Aperture Measurements of Two km (round trip) Amplitude and Phase Perturbations of One Micron Laser Speckle Propagating Through a Turbulent Atmosphere", Proc. SPIE 2471, 254-258, (1995).

Dual Aperture Measurements of Two km (round-trip) Intensity and Phase Perturbations One Micron Laser Speckle Propagating Through a Turbulent Atmosphere

Philip Gatt, Tom P. Costello, and Dhinakarraj H. Gantala

Center for Research and Education in Optics and Lasers (CREOL)
University of Central Florida (UCF),
12424 Research Parkway, Orlando, Florida 32826

ABSTRACT

The results of recent measurements of the intensity and phase fluctuations of one micron speckle propagating through a two kilometer (round-trip) turbulent path are presented. The data was collected at BMDO's Innovative Science and Technology Experimentation Facility (ISTEF) at the Kennedy Space Center in Florida using a dual-aperture Nd:YAG coherent array transceiver developed at CREOL. The dual-aperture measurement technique allowed for the discrimination of atmospheric turbulence induced phase perturbations from relative target/platform motion induced phase modulation, since the target motion produces common phase modulation in both receivers. The phase and intensity were found to be Gaussian and K-distributed respectively.

Keywords: speckle, turbulence, k-distribution, coherent detection, coherent array, phase detection

1. INTRODUCTION

Early theoretical research in the area of speckle propagation through turbulence predicted negative exponential intensity statistics.[1] However, experimental results showed that the negative exponential intensity statistics were only valid for weak and strong turbulence conditions.[1] This prompted researchers to revise the intensity statistics models, which led to the use of the K-distribution as a model for the intensity statistics of laser speckle propagating through turbulence. The phase statistics, on the other hand, have been assumed to be Gaussian.

The development of the K-distribution model for the intensity fluctuations, resulted by considering that turbulence modulates the mean value of the speckle intensity, <I>. The conditional PDF $p_I(I/<I>)$ for fully developed vacuum speckle is know to be negative exponentially distributed. While, the statistics of the mean value of the intensity $p_{<I>}(<I>)$ is log-normal for weak turbulence. Thus, the intensity statistics can be calculated from conditional statistical theory. That is,

$$p_I(I) = \int p_I(I/<I>)p_{<I>}(<I>)d<I>$$

To make the math more tractable, the log-normal model is approximated by a gamma distribution. This leads to the well-known K-distribution model.

$$p_I(I) = \frac{2}{\Gamma(\alpha)}\left(\frac{\alpha}{<I>}\right)^{(\alpha+1)/2} I^{(\alpha-1)/2} K_{\alpha-1}\left(2\sqrt{\frac{I\alpha}{<I>}}\right)$$

The K-distribution has the property that as the normalized variance, σ_n^2, approaches unity the K-distribution asymptotically approaches a negative exponential distribution (straight line on Figure 1)

0-8194-1824-2/95/$6.00

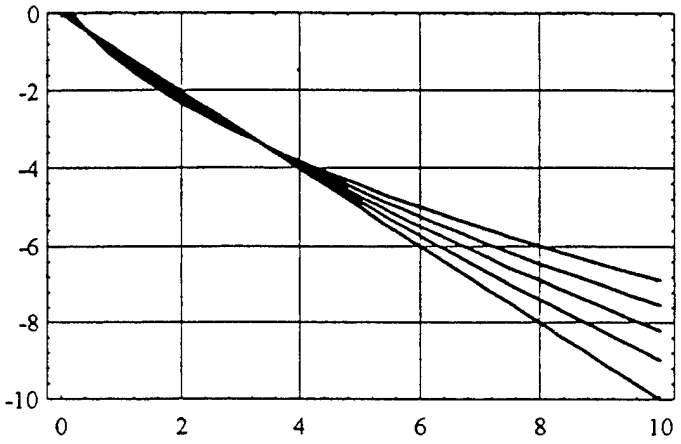


Figure 1, Family of K-distributions,
corresponding to a normalized variance of 2, 1.5, 1.25, 1.1, 1.0 from top to bottom

2. EXPERIMENT DESCRIPTION

While demonstrating a dual aperture coherent laser radar array receiver on BMDO/ISTEF's 1 km horizontal range, speckle propagation through turbulence data was fortuitously collected. The receiver consisted of 1 cm monostatic and bistatic apertures, separated by 25 cm. Single mode optical fiber was employed to mix the signals with a frequency offset local oscillator (27.12 MHz). The target was a diffuse retroreflectance target (gain ≈ 500) located at the end of the 1 km range.

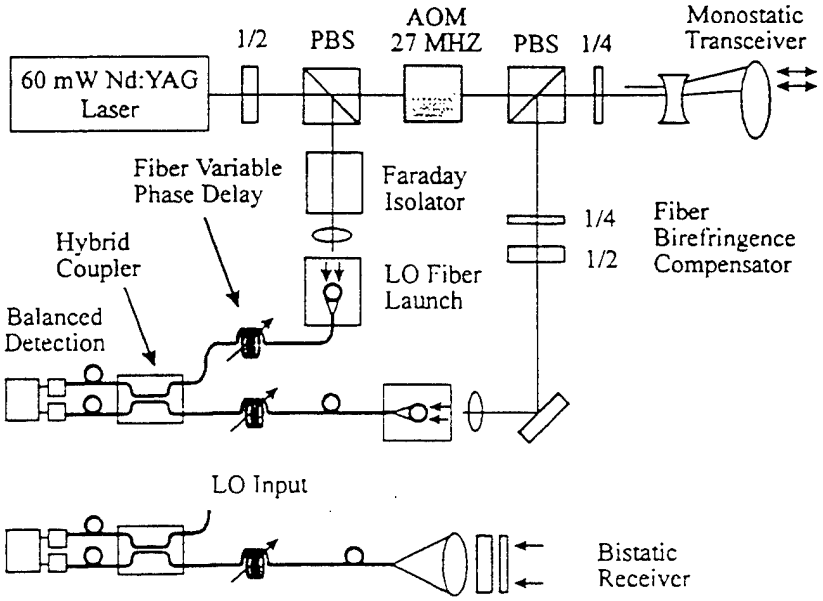


Figure 2. CREOL's Dual Aperture Nd:YAG Coherent Array Receiver

The raw RF signals were digitized at 10 kHz with eight bits. Ten records, each consisting of 8192 points were recorded for both the monostatic and bistatic receivers. The amplitude and phase were demodulated, from the digitized RF signals, using an off-line digital synchronous detection architecture.

2. EXPERIMENTAL RESULTS

2.1 Phase statistics

Target motions could in theory, be canceled by making simultaneous phase measurements from two apertures separated by a distance greater than ρ_o. However, upon analysis of the phase data, we observed that the low frequency component, in some cases showed a high degree of correlation between the two apertures, and in other cases did not. It was therefore assumed that this large scale, low frequency phase was dominated by atmospheric turbulence and not target motion, thus the raw phase data was analyzed rather than the difference of the two channels phase.

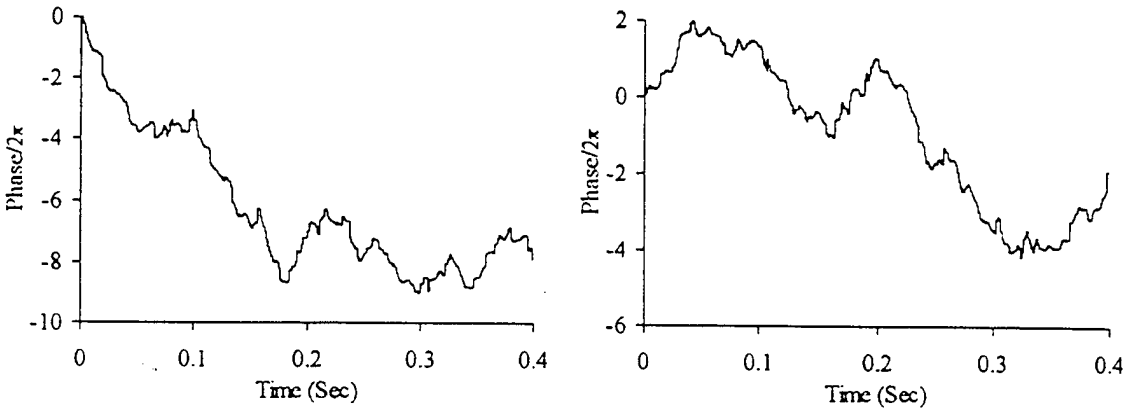


Figure 3. Typical temporal phase traces

a) Monostatic b) Bistatic)

The PDF of the phase were calculated using a standard histogram algorithm, Figure 4. From this data we observed that as much as +/- 40 waves of phase deviation occurred. The expected PDF of the phase data was Gaussian, and one can see that the PDFs shown in Figure 4 are somewhat Gaussian in shape. To obtain better PDFs, one needs more independent samples of the phase data.

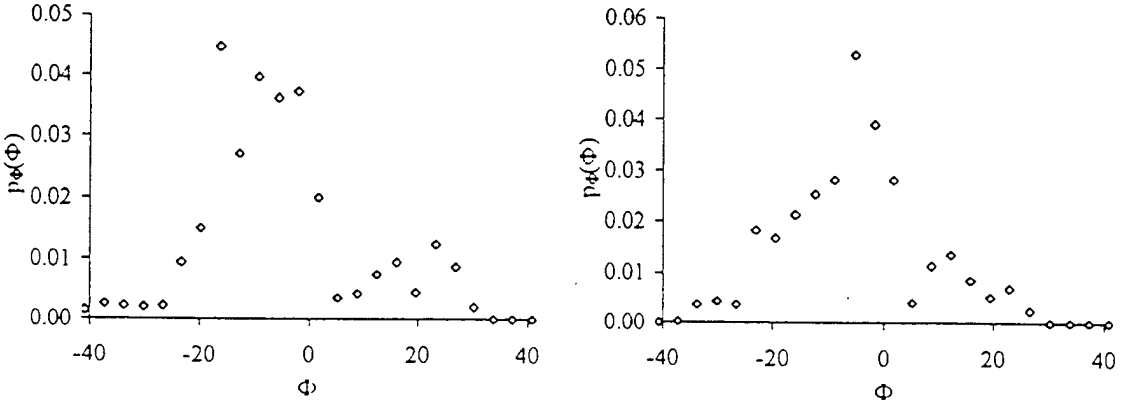


Figure 4. Phase PDFs

a) Monostatic b) Bistatic)

As shown in Figure 5, the 3 dB bandwidth of the phase was only about 5 Hz. Thus, of the 81920 data points(10 records at 8192 points per record), only 41 (81920*5 Hz/10 kHz) of the data points were independent. Hardly enough to obtain reasonable PDFs.

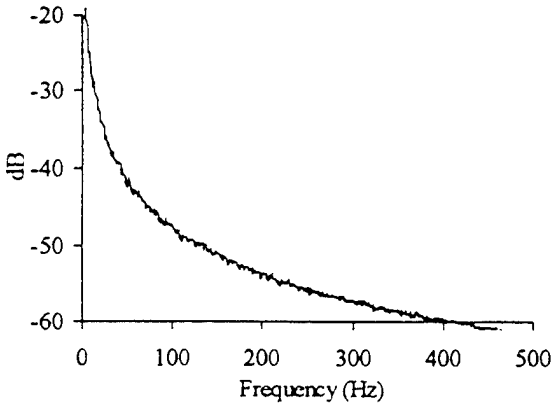


Figure 5 Phase Power Spectrum

2.2 Intensity statistics

Data was collected both in the mid-morning and in the mid-afternoon. A typical time trace of the demodulated intensity is shown in Figure 6 and intensity PDFs are shown in Figures 7 and 8.

Figures 7 and 8, indicate that the intensity statistics are well described by a K-distribution (smooth curves). The K-distribution fit was obtained by setting the k-distribution normalized variance to be equal to the normalized intensity variance.

It was found that the strength of the intensity variations (normalized variance) was higher in the morning than in the afternoon and higher for the bistatic receiver than the monostatic receiver.

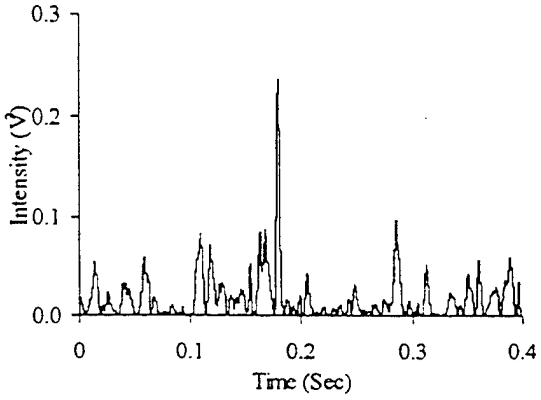


Figure 6 Typical Intensity time trace

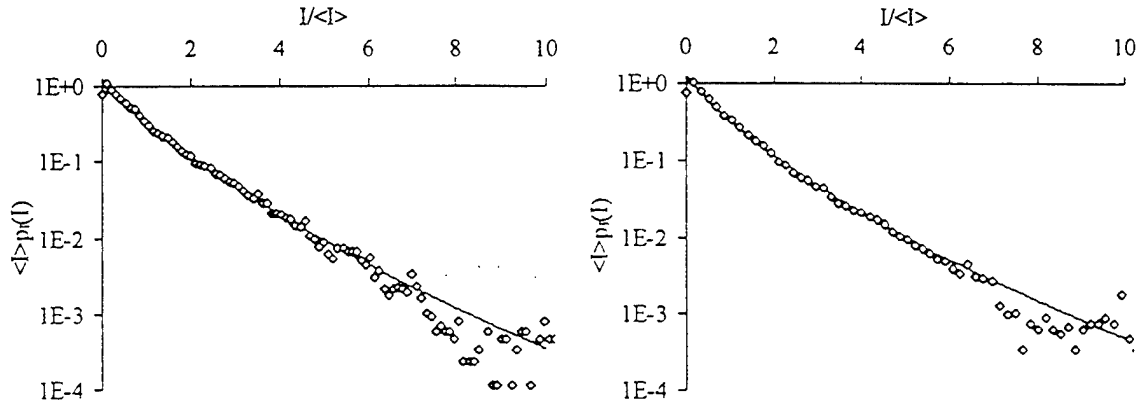


Figure 7 Intensity PDFs, for the mid-morning data collection cycle

a) Monostatic ($\sigma_n^2 = 1.33$)

b) Bistatic ($\sigma_n^2 = 1.47$)

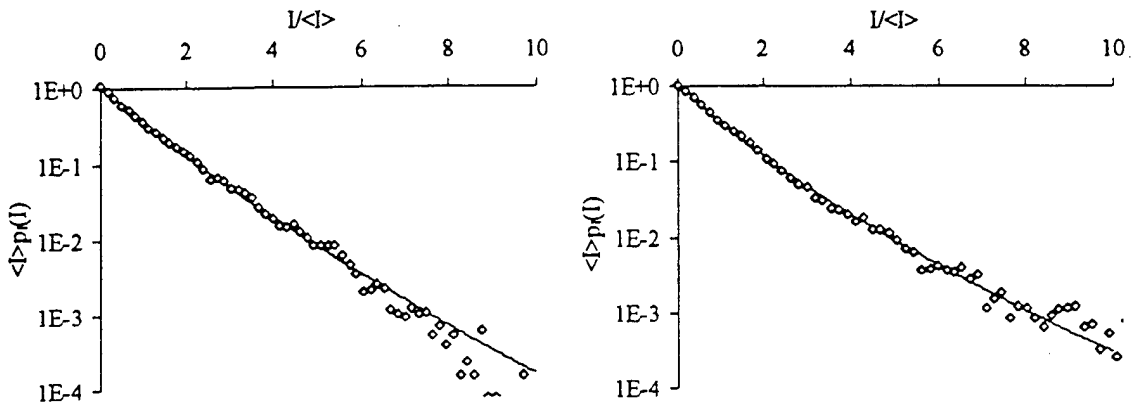


Figure 8 Intensity PDFs, for the mid-afternoon data collection cycle

a) Monostatic ($\sigma_n^2 = 1.14$) b) Bistatic ($\sigma_n^2 = 1.28$)

The intensity fluctuations were measured to be much faster than the phase fluctuations (cf., Figure 9). The measured bandwidth of the intensity fluctuations were seen to be on the order of 75 Hz as opposed to 5 Hz for the phase.

3. SUMMARY

We observed peak phase deviations, on the order of +/- 40 waves across the 1 km horizontal path, of which most are presumed to be due to large, slowly moving, turbulence eddies (outer scale), with high frequency modulations due to smaller scale, faster eddies.

Target motion was thought to have been negligible compared to atmospheric turbulence induced phase fluctuations since no consistent correlation was observed between the two channels' phase signals.

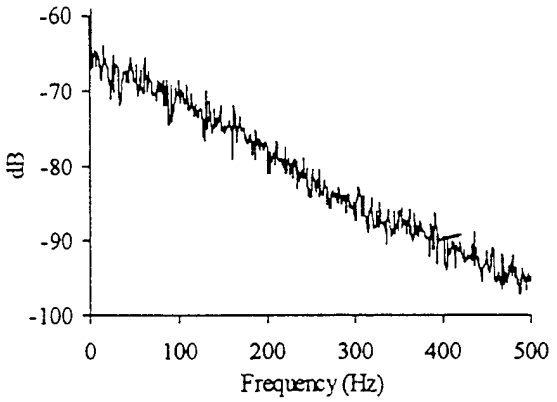


Figure 9. Demodulated Intensity Spectra

The Intensity statistics were shown to be a very well modeled by a K-distribution, for all four sets of intensity data (mid-morning and mid-afternoon; monostatic and bistatic). Stronger scintillation was observed for the bistatic receiver than the monostatic receiver, probably due to the self-compensating effects of the reciprocal path in the monostatic receiver.

4. ACKNOWLEDGMENTS

This work was sponsored by BMDO and ONR under contract N00014-92-C-0123.

5. REFERENCES

1. Pincus, Philip A., Fossey, Michael E., Holmes, J.F., and Kerr, J.R. "Speckle Propagation Through Turbulence: Experimental" JOSA, Vol. 68, No. 6, pp. 760-762, Jun. 1978

2. Gudimetla, V.S.Rao, Holmes, J. Fred, "Probability density function of the intensity for a laser-generated speckle field after propagation through the turbulent atmosphere", JOSA, Vol. 72, No. 9, pp. 1213-1218, Sep., 1982

2.7. Appendix 2.B

P. Gatt, T. P. Costello, D. A. Heimmermann, D. C. Castellanos, A. R. Weeks, and C. M. Stickley, "Coherent Optical Array Receivers for the Mitigation of Atmospheric Turbulence and Speckle Effects", Appl. Opt., 35, 5999-6009, (October 20, 1996).

Coherent optical array receivers for the mitigation of atmospheric turbulence and speckle effects

Philip Gatt, Thomas P. Costello, Dean A. Heimmermann, Diana C. Castellanos, Arthur R. Weeks, and C. Martin Stickley

A description is given of the design, operation, and test over a 2-km path (roundtrip) of a continuous wave, coherent laser array receiver that uses two independent aperture–receivers whose intermediate frequencies are electro-optically co-phased in real time and then added as a proposed way to overcome effective aperture limitations imposed by atmospheric turbulence and to mitigate signal fading associated with atmospheric turbulence and speckle effects. The experiment resulted in a mean carrier-to-noise ratio increase of 1.8, which is within 1% of the theoretical predictions, when the two signals were phase locked, versus no increase without phase locking. Further, the carrier fading strength, or normalized carrier-to-noise ratio variance, was reduced by a factor of 0.53, which is within 2% of the theoretical prediction. The bandwidth of the electro-optic phase-locked loop was measured to be of the order of 600 Hz, which is adequate to compensate for atmospheric refractive turbulence fluctuations. © 1996 Optical Society of America

Key words: Coherent optical array receivers, heterodyne arrays, space diversity receivers, coherent optical communications, atmospheric turbulence compensation, electro-optic phase-locked loop.

1. Introduction

We describe the development of an array technology for coherent laser communication or cw ladar receivers. The benefit of this technology is both an increase in the mean carrier-to-noise ratio (CNR), caused by an increase in heterodyne mixing efficiency, and a reduction in carrier fading, caused by phase coherent averaging effects. The physical layout of these optical array receivers can be diverse, ranging from fully monolithic apertures focusing light onto a detector array to a distributed array of small light-weight apertures coupling light into a coherent fiber-optic receiver. Both implementations provide the same benefit over a single-element coherent receiver.

Coherent array receivers, though developed in the 1950's for mobile radio communications,[1–3] were first suggested in 1965 by Goldstein et al.[4] to overcome the aperture limitations imposed by atmospheric refractive turbulence. The finite transverse coherence length, r_o, of an optical field propagating through refractive turbulence imposes an upper bound on the effective transceiver diameter of a conventional coherent laser radar, which in turn limits the maximum mean CNR for a given laser source.[5] In 1976, Fink and Vodopia[6] published the first detailed treatment of the mean CNR (or mixing efficiency) enhancement provided by a coherent array. They showed that a properly designed coherent array receiver could theoretically yield near-perfect heterodyne mixing efficiency, even in the presence of severe atmospheric turbulence. In principle, this can be accomplished by creating a local oscillator (LO) field, which spatially matches the amplitude and phase of the aberrated signal field. In this paper we show that this approach leads to an increase in the average CNR up to M fold, where M is the number of elements in the receiver array.

A second, and perhaps more important, advantage

When this research was performed the authors were with the Center for Research and Education in Optics and Lasers, University of Central Florida, Orlando, Florida 32816-2700. P. Gatt is now with Coherent Technologies, Inc., Boulder, Colorado 80302; T. Costello is with AirNet, Melbourne, Florida 32901; and D. Heimmermann and D. Castellanos are with the Harris Corporation, Melbourne, Florida 32902.

Received 14 July 1995; revised manuscript received 27 February 1996.

0003-6935/96/305999-11$10.00/0

© 1996 Optical Society of America

of a coherent array is its ability to produce a carrier with much reduced amplitude fluctuations, or fading, caused by dynamic speckle effects. On the one hand, a conventional coherent receiver detecting a fully developed dynamic speckle field yields a normalized signal intensity variance (i.e., fading strength) of unity.[7] On the other hand, as shown in this paper, a well designed M-element coherent array will reduce carrier fading by a factor of $1/M$, in a manner similar to the well-known aperture-averaging technique employed in direct detection receivers.[8]

Others have reported their efforts to develop optical coherent arrays. However, in those efforts the processing of the signals from the coherent array elements was quite different from the approach described here. For example, researchers at MIT Lincoln Laboratory have reported on a 12-element HgCdTe linear array to increase the frame rate of a coherent imaging ladar. This facilitated frame averaging of the speckled intensity images resulting in an image with reduced speckle noise or an increased intensity signal-to-noise ratio.[9–13] At the University of South Florida, researchers recently reported their results of a four-element coherent array utilizing a monolithic receive aperture and a quadrant InGaAs detector. In their first experiments,[14,15] the four independent radio frequency (RF) signals were video detected and averaged, leading to improved signal-to-noise ratio statistics. However, their approach resulted in a loss of Doppler information. In a later paper, they reported their results of non-real-time coherent processing of pulsed lidar returns.[16]

This paper reports on the results of our theoretical analysis, computer simulation, and a field experiment of a real-time fully coherent heterodyne array receiver. Our approach utilizes sparsely distributed array elements, which couple the signal photons into single-mode fiber-optic heterodyne receivers, and employs real-time phase alignment of the RF carriers by using an electro-optical phase-locked loop. Section 2 describes three popular space-diversity reception techniques, developed for mobile radio communications, which are also applicable to coherent laser radar receivers. In Section 3 the performance of these three receiver architectures are compared, on a statistical basis, in terms of their relative improvement factors for the mean CNR and the carrier fading. In Section 4 the results of a computer simulation of an eight-aperture coherent array are presented. The simulation not only demonstrated the performance enhancement of an eight-element coherent array but also provided confidence in the CNR moment estimation equations that were utilized in the analysis of the field experiment data. Section 5 describes the field experiment hardware and a comparison of the experimental results to theoretical predictions. Section 6 discusses some potential applications for this technology, and Section 7 summarizes our results.

2. Space-Diversity Receivers

Optical coherent array receivers belong to a more general class of signal processing architectures, called space-diversity receivers. Major advances in space-diversity receivers occurred in the 1950's when the technology was being applied to mobile radio communications. In this paper the application of three common space-diversity signal processing architectures to optical coherent arrays is investigated. These receiver architectures are depicted in Fig. 1.

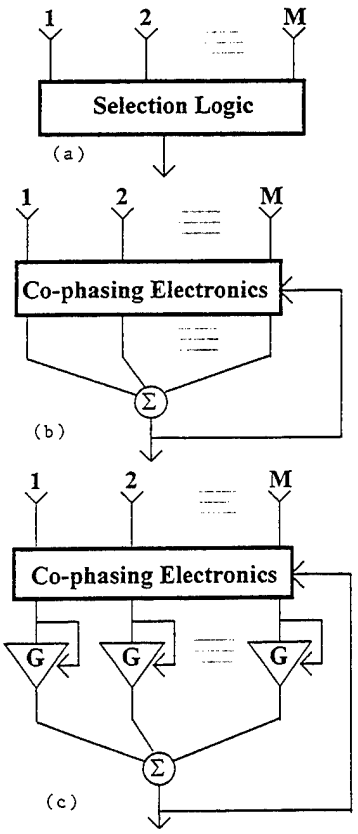


Fig. 1. Space-diversity receiver architectures: (a) selection, (b) EG, (c) MR.

The maximal-ratio (MR) receiver, which is sometimes referred to as the ratio-squarer receiver, was first proposed by Kahn for mobile radio communications.[1] In this architecture, the RF signals are cophased, have their amplitudes separately adjusted, and then summed to generate a composite RF signal with improved qualities. The MR receiver is the optimal receiver in that it yields a carrier with the highest mean CNR and lowest CNR fading. This type of signal processing technique can be thought of as a method that produces a synthetic LO field that is both amplitude and phase matched to the distorted received signal. Therefore, it yields the highest possible heterodyne efficiency. With the use of this conceptual approach, it becomes obvious that the optimum electronic gain for each receiver should be one that synthesizes a LO field amplitude that is proportional to the received signal field amplitude.

The implementation of a MR receiver requires the development of signal processing hardware that adjusts both the delay and the gain elements of the array RF outputs. The rate at which these must be adjusted will be dictated by the rate at which the

target speckle and atmospheric turbulence fluctuates. Target speckle fluctuation rates can be as high as 100 MHz, depending on target range, rotation rate, and transverse velocity, whereas atmospheric fluctuations generally do not exceed 1 kHz.

An equal-gain (EG) receiver differs from the MR receiver in that equal gains are applied to all the RF signals. Thus only the phase of the synthetic LO field is adjusted to match the signal field. As we show in Section 3, this phase-matching-only technique results in carrier statistics that are very close to those that would be obtained from the optimal MR receiver.

A selection receiver is quite different from the above two coherent summing receivers. In this type of receiver, the single strongest RF signal is switched to the output, while all the other array element signals are discarded. This, the simplest type of coherent array receiver, does little to improve the mean CNR. However, it does provide a reasonable reduction in the signal fading, albeit with less of an improvement than would be achieved by using a MR receiver.

3. Carrier-to-Noise Ratio Performance Enhancements for Space-Diversity Receivers

The moments of the CNR (or equivalently the carrier power) for the three space-diversity receivers discussed above provide a relative measure of the improvement these receivers offer over a single-element coherent receiver. The mean CNR is a measure of the average carrier power relative to the noise power, and the normalized variance of the CNR is a measure of carrier fading. Here we present expressions for these two statistical measures.

Throughout this analysis we assume that (a) the statistics of the signal amplitude are Rayleigh distributed, which is typical for fully developed laser speckle; (b) each receiver's average carrier power is equal; (c) the noise power in each receiver is equal and dominated by signal shot noise; and (d) each detector has the same quantum efficiency.

A. Signal Model

Coherently detected signals are modeled as narrow-band RF signals with additive Gaussian noise, i.e.,

$$x(t) = a(t)\cos[\omega_c t + \phi(t)] + n(t), \tag{1}$$

where amplitude $a(t)$ is assumed to be Rayleigh distributed, ω_c is the carrier frequency, phase $\phi(t)$ is assumed to be uniformly distributed over a 2π interval, and noise $n(t)$ is assumed to be zero-mean Gaussian shot noise. The narrow-band assumption implies that the amplitude and phase fluctuate at a rate that is slow compared with the carrier frequency.

To explain the performance increase that an array receiver provides over a single-element receiver, we first investigate the statistics of the CNR, $\delta(t)$, for a conventional single-detector coherent optical receiver. Then we expand this theory to model the CNR, $\Delta(t)$, of the three types of coherent array receivers discussed in Section 2.

B. Conventional Single-Element Coherent Receivers

Usually the CNR is defined as an average quantity. That is, it is defined as the ratio of the average carrier power to the average noise power. However, we take a more general approach here, which facilitates quantification of the carrier fading. In this approach we define the instantaneous CNR as the ratio of the instantaneous signal power to the average noise power. For a quantum or shot-noise-limited coherent optical receiver, the CNR is know to be given by[17]

$$\delta(t) = \frac{1/2a^2(t)}{E[n^2(t)]} = \frac{\eta\gamma(t)P(t)}{h\nu B}, \tag{2}$$

where η is the detector quantum efficiency, γ is the heterodyne mixing efficiency, $P(t)$ is the signal power incident upon the detector, h is Planck's constant, ν is the optical frequency, and B is the receiver's noise-equivalent bandwidth. Amplitude $a(t)$ and noise $n(t)$ are each proportional to many common factors, such as the LO power, the detector responsivity, and the electronic gain, which cancel in the signal-shot-noise-limited CNR equation. Therefore, we define the amplitude and noise as normalized quantities given by

$$a(t) = [2\eta\gamma(t)P(t)]^{1/2}, \tag{3}$$

$$E[n^2(t)] = h\nu B. \tag{4}$$

The single-element-receiver mean CNR is found by taking the expected value of Eq. (2):

$$\bar{\delta} = E[\delta(t)] = \frac{\eta E[\gamma(t)P(t)]}{h\nu B}. \tag{5}$$

For a fully diffuse target and when atmospheric turbulence is negligible, the statistics of the amplitude are known to be Rayleigh distributed; thus the CNR will be negative exponentially distributed, and as such its normalized variance or fading strength is unity ($\hat{\sigma}_\delta^2 = 1$), where $\hat{\sigma}_\delta^2 = \sigma_\delta^2/\bar{\delta}^2$.

C. Coherent Array Receivers

Here we provide an overview of the CNR statistics for the three space-diversity receivers discussed above. In this analysis it is assumed that the signals in each array element are spatially uncorrelated, and that each receiver has equal average carrier and noise powers.

1. *Maximal-Ratio Receiver*

The instantaneous CNR, $\Delta_{sum}(t)$, for a summing coherent array (MR or EG) is the power ratio of the phase-coherent addition of the signal amplitudes from each element of the array to the incoherent addition of the noise. If an arbitrary voltage gain, $G_i(t)$, is assumed for each receiver in the array and if equal noise powers are assumed, the resultant com-

posite mean CNR for an M element array is given by

$$\Delta_{\text{sum}}(t) = \frac{\eta\left\{\sum_{i=1}^{M} G_i(t)[\gamma_i(t) P s_i(t)]^{1/2}\right\}^2}{h\nu B \sum_{i=1}^{M} G_i^2(t)}. \quad (6)$$

The optimum gain settings for a MR receiver are proportional to the amplitude of the signal itself.[3,6] That is,

$$G_i(t) = k a_i(t) = k[2\eta\gamma_i(t) P s_i(t)]^{1/2}, \quad (7)$$

where k is an arbitrary constant of proportionality. This result is consistent with the notion that when the electronic gains and phase delays are backpropagated into the LO, the optimum gain and phase adjustments would result in an amplitude and phase match of this synthetic LO field to the distorted signal field. Applying the optimum gains to Eq. (7) results in the following expression for the MR receiver:

$$\Delta_{\text{MR}}(t) = \frac{\eta}{h\nu B} \sum_{i=1}^{M} \gamma_i(t) P s_i(t) = \sum_{i=1}^{M} \delta_i(t). \quad (8)$$

Therefore, the instantaneous MR receiver CNR is the sum of the component array element CNR's, and the mean CNR is simply M times the single-element mean CNR or $M\bar{\delta}$. Because the CNR statistics of the individual receivers follow a negative exponential distribution and have equal means, the statistics of the MR receiver will be gamma distributed, with a normalized variance equal to $1/M$.[18]

2. Equal-Gain Receiver

For the EG receiver the mean CNR is found by evaluating the expected value of Eq. (6) when all the gains are equal. The result for Rayleigh-distributed amplitudes can be shown to be given by[3]

$$E[\Delta_{\text{EG}}(t)] = [1 + (M - 1)\pi/4]\bar{\delta}. \quad (9)$$

In the limit of large M, the mean EG receiver CNR approaches $\pi/4$ of the mean MR CNR. Therefore, at worst (i.e., large M), an EG receiver will yield a mean CNR enhancement that is 78% (or −1.05 dB) of what would be obtained from a MR receiver. The normalized variance is given by[19]

$$\hat{\sigma}_{\Delta_{\text{EG}}}{}^2 = \frac{2[-(2\pi^2 - 8\pi)M^2 + (5\pi^2 - 20\pi + 16)M - (3\pi^2 - 12\pi + 8)]}{M(\pi M + 4 - \pi)^2}, \quad (10)$$

which, in the limit of large M, is only 9.26% larger than the MR normalized CNR variance.

3. Selection Receiver

For the selection receiver, the moments are given by[3]

$$E[\Delta_S{}^n(t)] = n! M \bar{\delta}^n \sum_{k=0}^{M-1} \begin{bmatrix} M-1 \\ k \end{bmatrix} \frac{(-1)^k}{(k+1)^{n+1}}, \quad (11)$$

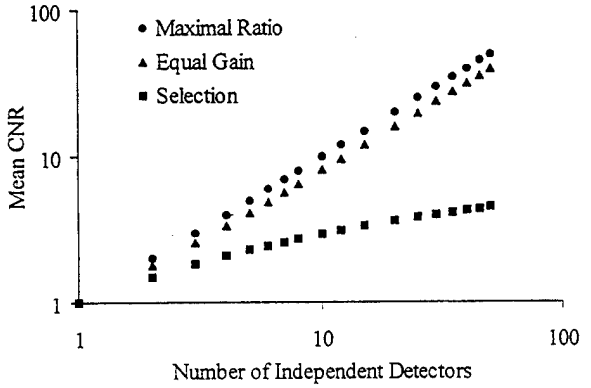


Fig. 2. Mean CNR for the MR, EG, and selection coherent array receivers.

and the normalized CNR variance is given by

$$\hat{\sigma}_{\Delta_S}{}^2 = \frac{2}{M} \frac{\sum_{k=0}^{M-1} \begin{pmatrix} M-1 \\ k \end{pmatrix} \frac{(-1)^k}{(k+1)^3}}{\left[\sum_{k=0}^{M-1} \begin{pmatrix} M-1 \\ k \end{pmatrix} \frac{(-1)^k}{(k+1)^2}\right]^2} - 1. \quad (12)$$

As a way to facilitate a conceptual comparison of the selection receiver's performance, closed-form approximations to the mean CNR and the normalized CNR variance were obtained from a least-squares curve fit. Best fits for the mean CNR and for the normalized CNR variance were obtained for logarithmic and power law functions, respectively. These approximations are given by

$$E[\Delta_S(t)] \approx [0.815 + 0.94 \ln(M)]\bar{\delta} \quad \text{for } 1 \le M \le 50, \quad (13)$$

$$\hat{\sigma}_{\Delta_S}{}^2 \approx 0.8 M^{-0.62} \quad \text{for} \quad 1 \le M \le 50. \quad (14)$$

4. Diversity Receiver Performance Comparison

Figures 2 and 3 provide a performance comparison of these three space-diversity receivers for a range of array elements from 1 to 50. As indicated by these figures, the MR receiver is optimum in terms of its ability to improve the mean CNR as well as reduce the CNR fading. As discussed above, the EG receiver's performance is only slightly worse. The selection receiver, in contrast, exhibits relatively poor mean CNR performance enhancement, but it does provide a reasonable reduction in the CNR fading. The approximate power law behavior of the selection receiver's normalized CNR variance is apparent by the near-linear slope in Fig. 3.

An alternative way to view these results is to con-

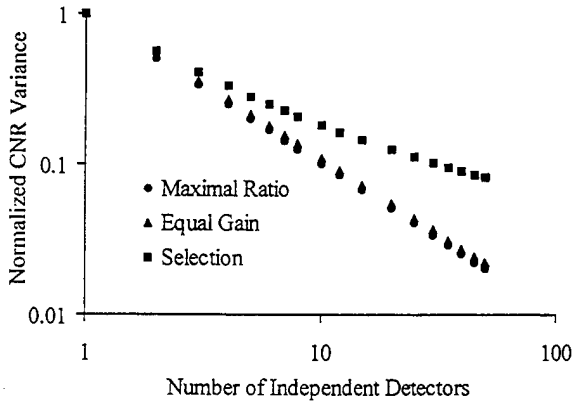


Fig. 3. Normalized CNR variance for the MR, EG, and selection coherent array receivers.

sider the number of independent array elements required to increase the mean CNR by 10 dB. It is straightforward to show that the resulting number of array elements is 10, 13, and 12,000 for the MR, EG, and selection methods, respectively. Similarly, to reduce the fading strength (normalized variance) by a factor of 10 would require 10, 11, and 31 independent array elements for the MR, EG, and selection receivers, respectively.

5. *Effect of Signal Correlation on the Carrier-to-Noise Ratio Moments of a Maximal-Ratio Receiver*

In some configurations (i.e., in dense arrays, or in those in which the array elements are small and closely spaced compared with the speckle correlation length) the RF amplitudes will be correlated as a result of overlapping speckle. In this case, the signal fading will be worse than predicted by the theoretical expressions given in the above section; however the mean CNR, for an MR receiver, will remain unchanged. In the limit, as the speckle size exceeds the overall array size, the normalized variance of the composite RF signal will increase from $1/M$ to that which is obtained by using a single detector (i.e., unity). In other words, in this extreme limit the array behaves just like a single, equal area, large-aperture conventional coherent receiver.

6. *Incoherent Summation of the Radio Frequency Carriers*

The EG receiver discussed above requires the coherent addition of the detected signals, which implies that the signals must be co-phased. Here we present a heuristic argument to show that there is no enhancement in the mean CNR or the CNR variance when the RF signals are added without phase compensation (i.e., when they are added incoherently on a field basis).

The mean CNR will be no better than it is for a single-detector case because both the mean signal power and the noise power increase by the same factor. The noise power increases by a factor of M just as it does for a co-phasing EG receiver. The resultant signal amplitude is the amplitude of an M step random walk, where the length of each step is Rayleigh distributed and the direction is uniformly distributed. For this type of random process, it is well known that the resultant amplitude is again Rayleigh distributed, with a mean value equal to $\sqrt{M}$ times the mean of the M constituent carrier amplitudes.[18] Therefore, the resultant mean carrier power will increase by a factor of M, as opposed to the M^2 increase realized by the MR coherent addition of the RF signals. Additionally, the normalized CNR variance will remain unity because the resultant amplitude is Rayleigh distributed. Thus no reduction in the CNR fading will be achieved.

4. Numerical Simulation of an Eight-Aperture Coherent Array

A Monte Carlo simulation model of an eight-aperture EG array was developed to demonstrate the predicted performance enhancement provided by an EG receiver. The simulation experiments served not only as a precursor to a field experiment but also to generate confidence in the theoretical expressions developed in Section 3. Further, the simulation model facilitated the testing of the fidelity of expressions developed to measure the CNR moments of real heterodyne detected signals.

A direct measurement of the CNR moments from real signal return data is not possible because carrier amplitude $a(t)$ cannot be entirely separated from noise $n(t)$. Instead, one must calculate the CNR moments indirectly. A comparison of the moments of the measured signal, $x(t)$, and separate measurements of the moments of the noise-only signal results in the following simple set of transformation equations:

$$\bar{\delta} = \frac{E[x^2(t)]}{E[n^2(t)]} - 1, \tag{15}$$

$$E[\delta^2(t)] = \frac{2}{E[n^2(t)]}\left\{\frac{E[x^4(t)]}{3E[n^2(t)]} - 2E[x^2(t)] + E[n^2(t)]\right\}, \tag{16}$$

$$\hat{\sigma}_\delta{}^2 = E[\delta^2(t)]/\bar{\delta}^2 - 1. \tag{17}$$

From a series of numerical experiments, it was found that for strong-to-weak specified mean CNR's, the above two statistical estimates converged toward the design values as the record length increased. However, for very weak mean CNR's (less than -10 dB), the second moment estimate, Eq. (16), failed to converge even for long record lengths. The estimating formula for the mean CNR, Eq. (15), had a much greater range of convergence (approximately -40 dB).

The lack of convergence was found to be a result of the fact that the odd-order moment terms, which were assumed to be zero in the mathematical derivation, become nonnegligible for very weak CNR's. These odd-order moments should average out as the

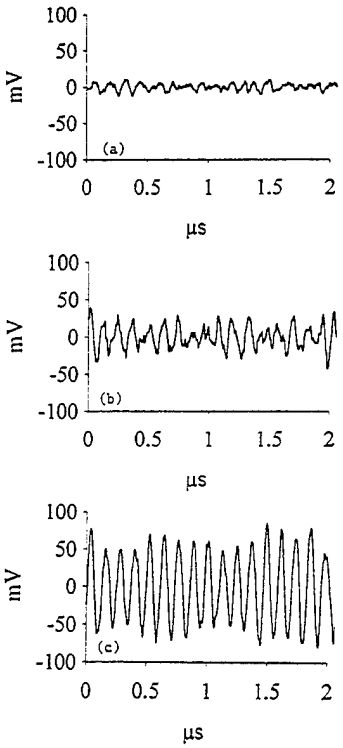


Fig. 4. Portions of the simulated signals: (a) one of eight simulated array element RF signals plus narrow-band noise with a specified mean CNR = 0 dB, (b) incoherent sum of eight signals, (c) coherent sum of eight signals.

record length is increased; however, as the record length increases, round-off errors accumulate, resulting in a lack of convergence. When the odd-order moments were included in the calculations, convergence was obtained for very low mean CNR's. Unfortunately, for real experimental data, one is not able to measure these odd-order terms, nor could one simultaneously monitor both the signal plus noise and the noise only. Therefore, one would not expect to obtain accurate estimates of the CNR second moment under very low CNR conditions.

A computer simulation was conducted with eight independent RF signals and used standard numerical statistical inversion techniques to generate synthetic return signals with the appropriate amplitude (Rayleigh), phase (uniform), and noise (narrow-band Gaussian) statistical distributions and power spectra. Record lengths were such that approximately 1000 independent samples of the Rayleigh-distributed amplitudes were recorded.

Figure 4 shows typical segments of one of the eight simulated RF signals, all with specified mean CNR's equal to 0 dB, and the resultant coherent and incoherent sum. Table 1 contains the values of the CNR moments as estimated by the formulas given in Eqs. (15) and (16). As demonstrated in the Table, the incoherent summation of the intermediate frequency carriers does not improve the CNR statistics whatsoever. However, for the EG coherent sum, the improvement is clear and consistent with the theory presented in Section 3.

Table 1. Summary of Calculated Moments of the CNR for a Simulated Eight-Aperture Coherent Array

Parameter	Addition	
	Incoherent	Coherent
Mean CNR		
Theoretical	1.00	6.74
Simulated	1.04	6.79
Normalized variance		
Theoretical	1.00	0.135
Simulated	1.05	0.139

5. Field Experiment

A real-time, dual-aperture, continuous wave, equal-gain coherent array was demonstrated at the Ballistic Missile Defense Organization's Innovative Science and Technology Experimentation Facility (ISTEF) at the Kennedy Space Center. The offset-homodyne transceiver, which is depicted in Fig. 5, consisted of two separated apertures coupling the received signals into single-mode optical fibers. Phase compensation was accomplished by an electronic servo loop utilizing a fiber-wrapped cylindrical piezoelectric transducer (PZT) as the active phase modulator. Extension to larger array dimensions can be accomplished by adding additional fiber-coupled bistatic receivers and their corresponding servo loops.

A. Hardware Description

The laser source was a commercially available, narrow-linewidth (5 kHz), 50-mW, Nd:YAG nonplaner ring laser (Lightwave Electronics 122-1064-50). The output beam had an elliptical TEM_{00} mode with an ellipticity ratio equal to 1.3:1.0. The combination of a half-wave plate and a polarizing beam splitter (PBS) provided a mechanism to divide the laser power variably into both the transmitter and the LO fields. The Faraday isolator provided 40 dB of feedback isolation.

The local oscillator was launched into a single-mode fiber by using a standard fiber beam launch system incorporating a 4-mm spherical focusing lens and an additional weak lens. The two-lens combi-

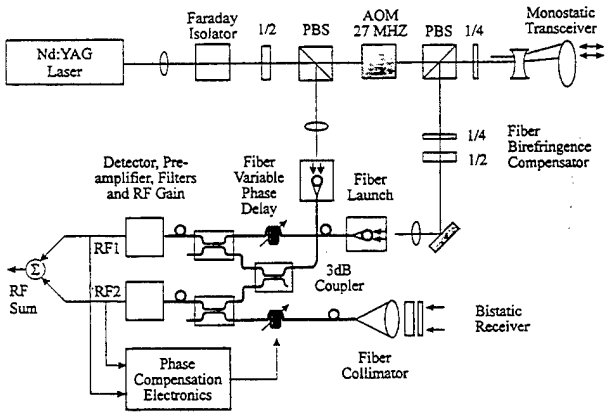


Fig. 5. Dual-aperture offset-homodyne transceiver.

nation facilitated optimization of the fiber launch efficiency. Routinely, fiber launch efficiencies in excess of 80% were obtained. Optical gel was not used, as it was found to be susceptible to contamination from both moisture and dust.

An acousto-optic modulator (AOM) in the transmitter leg provided a 27.12-MHz frequency offset from the LO. Traditionally, two AOM's are employed in an offset-homodyne receiver (i.e., one in the transmitter and one in the LO) to provide frequency isolation of the signal beat note from leakage of the AOM's RF drive. However, by careful shielding of the AOM's RF drive signal, RF leakage was reduced to a level far below the noise floor, and thus the need for a second AOM was circumvented.

In the monostatic transceiver, the return field was separated from the transmitter field by using a polarization duplexer, which consisted of a quarter-wave plate and a PBS. An off-axis 5× beam expander provided a collimated 1-cm output beam diameter and minimized transmitter feedthrough. The target-backscattered light was coupled into a second fiber beam launch system after passing through a static fiber birefringence compensator consisting of a quarter-wave plate and a half-wave plate.

The second receiver, which was bistatic, was constructed from a commercially available fiber beam collimator, whose effective diameter was 1 cm. Coarse optical alignment of the bistatic receiver was accomplished by maximizing the on-target overlap of the bistatic receiver's backpropagated field with the transmitter field.

Optical detection was achieved by using a 1-GHz InGaAs p-i-n photodiode. The photocurrent was amplified with a 200-MHz, 1000-Ω transimpedance amplifier, and the effective noise bandwidth was reduced to 1 MHz by a bandpass RF filter. The combination of a 90-dB RF amplifier and variable RF attenuators provided signal levels commensurate with the requirements for the electro-optic phase-locked loop (EOPLL).

Phase compensation was achieved with an in-house designed EOPLL. This loop consisted of an electronic automatic gain control, which had an input dynamic range in excess of 40 dB, a phase detector, loop filter, a variable loop gain, and a phase modulator. The optical phase was controlled by a fiber-wrapped (100 wraps) 2-in. (~5 cm) cylindrical PZT. The fiber-wrapped PZT's response was measured to be 0.04 (rad/wrap)/V and was flat until the first resonance, which occurred at 10 kHz. This provided a peak-to-peak phase control of 1600 rad [0.04 (rad/wrap)/V × 100 wraps × 400 V (peak to peak)], or ~250 λ.

Only one phase modulator and control system was required for the dual-aperture experiment. The second phase modulator (monostatic path in Fig. 5) facilitated servo-loop evaluation and optimization. It provided a mechanism for measuring the loop's response to deterministic forcing functions. The closed-loop's step response was observed to be dominated by a second-order pole whose damped natural

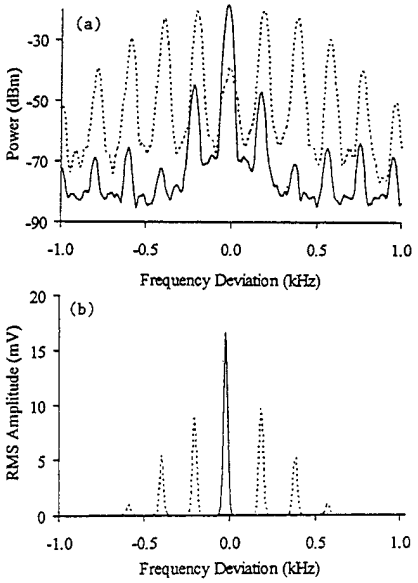


Fig. 6. Closed loop (solid curve) and open loop (dotted curve) of the sinusoidally modulated optical carrier: (a) power spectra, (b) amplitude spectra.

frequency was proportional to the loop gain. The loop gain was adjusted to provide a reasonable trade-off between the loop's bandwidth (i.e., damped natural frequency) and the loop stability. This was accomplished by adjusting the loop gain until the loop damping ratio, ζ, was at or near 0.25, for which the dampened natural frequency was measured to be 3 kHz. For this loop gain, the bandwidth of the servo loop, as defined by the residual phase error, was measured to be approximately 600 Hz, which is adequate to compensate for atmospheric turbulence.

Figure 6 illustrates the results of a test of the sensor's phase-locking performance at 200 Hz. In this test, a deterministic phase error was introduced in the local oscillator by sinusoidally phase modulating the test PZT phase modulator at 200 Hz with a modulation index of 2.405 rad. The open loop spectrum (dotted curve) is close to the theoretical prediction for a sinusoidally phase modulated carrier with the carrier containing less than 0.5% of the total power. The closed-loop power spectrum (solid curve) was generated by locking the heterodyne-detected signal to the AOM's drive signal. This curve demonstrates that most of the signal power, measured to be 99.6%, was transferred from the sidebands into the carrier.

B. Experiment Setup

The two 1-cm effective diameter receiver apertures were positioned 25 cm apart and aligned to an extended target located 1 km down a horizontal range. This separation distance ensured that the receivers were sensing spatially independent samples (speckles) of the return field, because for this particular transceiver–target configuration the speckle diameter should have been equal to the transmitter diameter (i.e., 1 cm).

The target was constructed from a 0.25-in. (~0.6

cm) thick aluminum plate covered with 3M reflectance tape (3M 7610). This tape, which did not appreciably depolarize the scattered radiation, provided a measured target antenna gain of 28 dB. The speckle, though not measured quantitatively, was anticipated to be fully developed, because the target consisted of many small retroreflecting spheres randomly positioned along the z axis. Laboratory observations of the speckle, generated by laser scatter from the 3M tape, showed no unique symmetry or character other than the enhanced reflectance.

The experiment at the ISTEF was conducted in mid-October 1994 under moderate turbulence conditions. Atmospheric turbulence provided the temporal variations in the return signal's amplitude and phase by a combination of turbulence and speckle effects. The turbulence effects were both scintillation and phase distortions. In addition to these first-order effects, the turbulence also causes dynamic speckle effects. As the turbulence modulates the beam position and beam structure on the target, an exchange of effective scatterers takes place, causing the speckle to evolve dynamically in a manner called speckle boiling.[20,21]

The measured spot size at 1 km was 30 cm, whereas the theoretical free-space beam diameter should have been approximately 20 cm. Because the far-field diffraction angle is known to be the root-mean-square (RMS) sum of the free-space diffraction angle and the turbulence-induced diffraction angle,[22] the turbulence-induced transverse coherence length, r_o, was estimated to be of the order of the transmitter diameter (i.e., 1 cm). Unfortunately, C_n^2 meters were not available during the experiment and thus r_o could not be verified.

In the experiment the bistatic receiver's RF signal was co-phased to the monostatic receiver's signal. A simpler alternative approach would have been to co-phase both signals to a stable local RF reference. However, in the latter approach, the message information (phase modulation) is removed from both carriers, whereas in the former approach the message information is preserved.

Throughout the experiment two types of data were collected. First, long-term (60 s) video averages of the signal power spectra were recorded on a Hewlett-Packard HP8560A RF spectrum analyzer. The average power spectra provided a direct measure of the mean CNR. However, because only one spectrum analyzer was available for the experiment, the video-averaged spectra had to be recorded sequentially.

Second, simultaneous measurements of the RF signals were obtained with a ten-channel, high-speed, eight-bit digitizer (Tektronix DSA602). The 27.12-MHz RF signals were digitized at a 10-kHz sample rate, which was far less than the Nyquist rate for the carrier but well above the Nyquist rate for the signal (turbulence-induced fluctuations) bandwidth, which was measured to be 200 Hz. The effect of this subsampling was (a) to beat the carrier frequency down to an aliased frequency of approximately 800 Hz, (b) to fold the narrow-band (1 MHz) noise into the half-

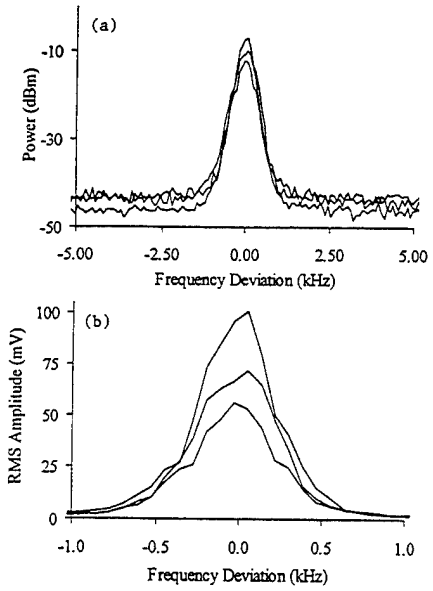


Fig. 7. Signal spectra for the 2-km (round-trip) field experiment, with coherent addition (top trace), incoherent addition (middle trace), and typical single channel spectra (lower trace): (a) power spectra, (b) amplitude spectra.

sample rate frequency space (i.e., 5 kHz), which makes the noise look white but does not change the mean CNR, and (c) to preserve both the amplitude and phase of the carrier.

The digitized RF signals were collected in ten disjoint 8192 point records. Because the bandwidth of the amplitude spectra was approximately 200 Hz and the sample rate was 10 kHz, the number of independent samples contained in the ten 8192 blocks was of the order of 1600 (10 blocks $\times$ 8192 points/block $\times$ 200 Hz/10,000 Hz), which is sufficiently large for reasonable estimates of the mean and normalized variance.

The data were collected under both moderate and weak mean CNR conditions. The weak CNR data were obtained by attenuating the transmit power by 20 dB. As a way to facilitate the analysis of the system performance, the electronic gains of the two receiver channels were adjusted so as to equalize their respective noise powers. Then the receiver wave plates were adjusted to equalize the average signal power in each receiver.

C. Experiment Results

Figure 7 displays typical samples of the sequentially collected, video-averaged power and amplitude spectra for the weak CNR data. The bottom trace in Fig. 7 is a representative sample of one of the two channel's signal spectra. The middle and top traces are the spectra of the incoherent and coherent sum of the two channel's signals, respectively. The improvement in the mean CNR is discerned from the figures by subtracting the increase in the noise power (3 dB measured in the wings of the data) from the increase in the peak signal power (6 dB for coherent addition and 3 dB for incoherent addition). Therefore, con-

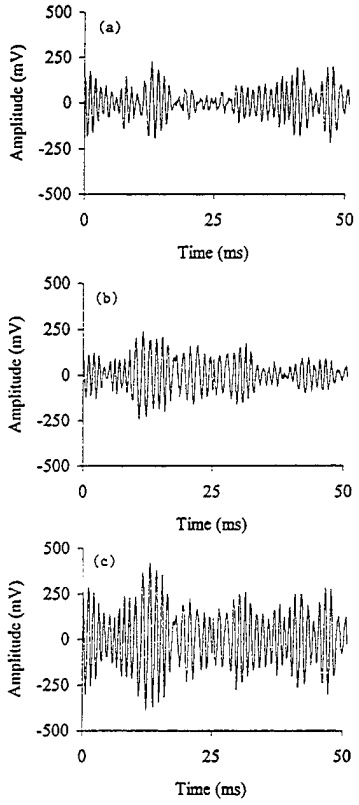


Fig. 8. Portions of simultaneously collected subsampled RF signals: (a) monostatic, (b) bistatic, (c) EG coherent sum.

sistent with theory, a nearly 3-dB increase in the mean CNR occurred for coherent addition, while no improvement was obtained for incoherent addition.

It should be mentioned that the receiver noise bandwidth was 1 MHz, whereas the spectrum analyzer's resolution bandwidth was just 300 Hz. Therefore, the actual electronic noise level was 35 dB higher than indicated in the spectrum analyzer traces. Thus, the true mean CNR, presented to the phase-compensation servo loop, was approximately only 0 dB [−11 dBm (signal), −46 dBm (noise) + 35 dB (noise bandwidth reduction factor)], yet stable phase locking was still obtained as exhibited by the 3-dB increase in the mean CNR. This was possible because the control loop's bandwidth was much less than the receiver noise bandwidth.

Figure 8 displays typical simultaneous segments of the subsampled monostatic signal, the co-phased bistatic signal, and the resultant equal-gain RF signal for the moderate mean CNR experiment. These figures demonstrate in the time domain both the increase in the mean CNR and a reduction in the CNR fading.

Estimates of the CNR moments are provided in Table 2 for both the moderate mean CNR and low mean CNR experiments. We calculated these statistical estimates using the formulas given in Eqs. (15) and (16) from measurements of the signal (carrier plus noise) and measurements of the noise alone.

In the development of the theory several, simplifying assumptions were made to produce a tractable result. These assumptions were (a) that the mean CNR of all the receivers were equal, and (b) that the amplitude statistics were Rayleigh distributed (or the intensity statistics were negative exponentially distributed, which is characterized by a normalized CNR variance equal to unity). However, these two assumptions were not fully realized by the experimental data. Therefore, two modifications to the theory were developed to facilitate comparison of the measured results with theoretical predictions. The first modification was to interpret the mean CNR of each receiver $\bar{\delta}$ in Eq. (9), to represent the average of the mean CNR's of the receivers; the second was to assume the normalized CNR variance $\hat{\sigma}_{\Delta_{EG}}{}^2$ in Eq. (10), to represent a reduction factor for the average of the receiver's normalized CNR variance.

For the coherent sum of the moderate CNR data, we observe very good agreement between the approximate theory and the measured results. In this case the resultant mean CNR increased by a factor of 1.8 over the average of the receiver's mean CNR. This result is within 1% of the approximate theory. Likewise, the reduction in the normalized CNR variance was measured to be 0.53 compared with an approximate theoretical reduction of 0.52. However, for the low CNR data, the mean CNR of the coherent sum was 25% below the approximate theoretical prediction, and the normalized CNR variance of the coherent sum was 36% higher than the prediction. This lack of agreement with theory was most likely caused by an increased sensitivity to noise and signal measurement errors under low CNR conditions. For the

Table 2. Dual-Aperture Coherent Array CNR Statistics[a]

	Coherent Sum				Incoherent Sum		
	Monostatic	Bistatic	Experiment Sum	Theory Sum	Monostatic	Bistatic	Sum
Moderate CNR							
Mean	33.3	22.9	50.6	50.2	26.5	28.0	27.6
Normal Variance	1.39	1.56	0.78	0.77	1.19	1.42	1.11
Low CNR							
Mean	1.48	1.07	1.7	2.27	1.29	1.37	1.31
Normal Variance	1.13	1.34	0.88	0.65	1.16	1.43	1.14

[a]A round trip of 2 km was made through atmospheric turbulence.

incoherent sum, no improvement is observed for the mean CNR and the normalized CNR variance, as theory would suggest.

It is interesting to note that the normalized variance for both receivers is higher than would be predicted for a fully diffuse speckle field (i.e., unity), and that the bistatic signal's normalized variance is higher than that of the monostatic signal. The deviation from negative exponential CNR statistics is due to atmospheric turbulence effects. It is well known that the intensity statistics of speckle propagating through atmospheric turbulence are K distributed,[23] and the resulting normalized variance is typically larger than unity. The observed lower normalized variance in the monostatic signal's CNR is to be expected, because for the monostatic receiver the beam propagates through the same turbulence in both directions.

6. Technology Applications and Discussion

A. Continuous Wave Applications

In a communications application such as in a satellite-to-ground system, the turbulence-induced phase noise can be mitigated by co-phasing each of the M receiver carriers to a stable RF local oscillator, rather than to one of the other received signals, provided the message modulation format (i.e., modulation index and message frequency) is carefully designed to facilitate frequency separation of the phase noise from the message signal. Furthermore, any high-frequency phase information (message) must be common to all the receivers.

In a cw radar application, whereby an EOPLL is employed to compensate the phase noise, the application must be one for which the phase noise (caused by turbulence or target-induced speckle) fluctuation rates are within the bandwidth of the phase-locked loop. If not, then incoherent signal processing of the demodulated array outputs would be an attractive approach. One application for the cw sensor is long-range vibration sensing of relatively stationary targets.

A byproduct of this coherent array research was the development of the EOPLL technology. While our first-generation system had only 600 Hz of bandwidth, the second-generation system (for which the electronics and the EOPLL have been constructed) has a bandwidth of at least 5 kHz (it has only a 6° tracking error at 5 kHz). The EOPLL technology can be applied to many engineering problems whereby precision optical path-length compensation is required. Some additional potential applications include laser frequency stability, holography, and precision displacement compensation.

B. Pulsed Applications

In hard-target pulsed laser radar applications the duty cycle of the return pulse train is typically very low. For these applications, it would probably be less costly and simpler to compensate the phases with a real-time digital signal processor rather than the analog EOPLL. Phase compensation would be relatively straightforward, because for short pulse laser radars one would not expect intrapulse phase noise. Thus, only one phase adjustment would be required per pulse per channel. Virtually any low-duty cycle pulsed hard-target laser radar could take advantage of this technology.

C. Incoherent Signal Processing of the Demodulated Array Outputs

If the phase noise is too fast for the phase compensator, or if one simply wants to avoid the complex EOPLL hardware, incoherent averaging of the demodulated array outputs (i.e., velocity, range) still results in signals with reduced fading. However, the mean CNR is not improved. The array, in this case, supplants the need for temporal averaging, thus reducing the laser pulsed repetition frequency requirement. A specific application of this technology is long-range hard-target imaging, in which a coherent receiver would be required to meet the range requirements and a spatial array would be employed to mitigate target speckle effects.

D. Transmitter Compensation

It should also be pointed out that additional gains can be achieved if the transmitter, not just the receiver, is compensated.[24–26] In a properly designed adaptive transmitter, refractive turbulence beam spreading can be reduced, resulting in more on-target photons and thus an improved CNR. In our study, however, we transmitted through just one channel and received in all the others.

7. Summary

This paper has shown that coherent laser radar array receivers can mitigate the adverse effects of target speckle and atmospheric turbulence. We showed that the EG receiver, which is much simpler to implement, performs within 1 dB of the optimal MR receiver. Specifically, for an EG receiver, the mean CNR will increase nearly linearly with the number of array elements M, and the normalized CNR variance will decrease at a rate close to $1/M$. In addition, the effect of correlation between neighboring apertures will increase carrier fading and should be kept to a minimum by employing an aperture separation distance larger than the mean speckle diameter.

The results of an eight-aperture coherent array numerical simulation were presented, which provided confidence in the statistical models of the CNR enhancement and also facilitated the testing of algorithms for processing the measured field data.

Results of a 2-km round-trip field demonstration of a dual-aperture real-time equal-gain coherent array showed it to perform extremely close to theoretical predictions under moderate mean CNR conditions. For low CNR levels, however, the measured results deviated from the theory by approximately 25%. We also observed less carrier fading, of the order of 10–20%, for the monostatic receiver than for the bistatic receiver.

Currently we are developing an eight-aperture array. This second-generation array includes an enhanced EOPLL. The principal modifications to the loop include smaller PZT's, which exhibit higher frequency response and an optimized loop filter. Multiple fiber wraps around the PZT cylinders have been employed to compensate for the reduced peak-to-peak phase deviation associated with the smaller PZT's. We found that eight layers of fiber provide approximately four times the response of a single layer.

Upon completion of this eight-aperture–receiver system, we expect that the mean CNR will improve 6.8 fold, the fading will be reduced by a factor of 1/7.4, and the system will be able to compensate for phase fluctuation rates, caused by either atmospheric turbulence or target movement, as much as 5–10 kHz.

Funding for this research was provided by the Ballistic Missile Defense Organization's Innovative Science and Technology Directorate and was administered by Matthew White and William Stachnik of the Office of Naval Research under contract N00014-92-C-0123. The development of the coherent array technology described in the paper is the culmination of several years of research and development by many contributors. Robert Mongeon provided invaluable technical guidance in the developing phases of the project. Philip Gatt developed the theory, the design approach, and led the experimental work at ISTEF. A. Weeks, Dean A. Heimmermann, and Diana C. Castellanos developed the real-time receiver servo electronics. Thomas P. Costello and Chie L. Gagge developed the fiber-optic receivers, and Wilson P. Perez and Dhinakaraj H. Gantala designed and executed the data-processing and computer simulations. The authors gratefully acknowledge the efforts of the ISTEF site contractor (Nichols Research Corporation) staff for their support during the field experiment.

References

1. L. R. Kahn, "Ratio squarer," Proc. IRE **42,** 1704–●●● (1954).
2. D. G. Brennan, "Linear diversity combining techniques," Proc. IRE **47,** 1075–1102 (1959).
3. J. D. Parsons, "Diversity techniques in communications receivers," in *Advanced Signal Processing*, D. A. Creasey, ed. (Peregrinus, London, 1985), Chap. 6.
4. I. Goldstein, P. A. Miles, and A. Chabot, "Heterodyne measurements of light propagation through atmospheric turbulence," Proc. IEEE **53,** 1172–1180 (1965).
5. D. L. Fried, "Optical heterodyne detection of an atmospherically distorted signal wave front," Proc. IEEE **55,** 57–67 (1967).
6. D. Fink and S. N. Vodopia, "Coherent detection SNR of an array of detectors," Appl. Opt. **15,** 453–454 (1976).
7. J. H. Shapiro, B. A. Capron, and R. C. Harney, "Imaging and target detection with a heterodyne-reception optical radar," Appl. Opt. **20,** 3292–3313 (1981).
8. H. T. Yura and W. G. McKinley, "Aperture averaging of scintillation for space-to-ground optical communication applications," Appl. Opt. **22,** 1608–1609 (1983).
9. W. B. Veldkamp and E. J. Van Allen, "Binary holographic LO beam multiplexer for IR imaging detector arrays," Appl. Opt. **22,** 1497–1507 (1983).
10. W. B. Veldkamp, "Holographic local-oscillator beam multiplexing for array heterodyne detection," Appl. Opt. **22,** 891–900 (1983).
11. W. B. Veldkamp and C. J. Kastner, "Beam profile shaping for laser radars that use detector arrays," Appl. Opt. **21,** 345–356 (1982).
12. J. H. Shapiro, "Heterodyne mixing efficiency for detector arrays," Appl. Opt. **26,** 3600–3606 (1987).
13. B. E. Edwards, "Design aspects of an infrared laser radar," Laser Appl. **1,** 47–50 (1982).
14. K. P. Chan and D. K. Killinger, "Enhanced detection of atmospheric-turbulence-distorted 1-μm coherent lidar returns using a two-dimensional heterodyne array," Opt. Lett. **17,** 1237–1239 (1992).
15. N. Sugimoto, K. P. Chan, and D. K. Killinger, "Optimal heterodyne detector array size for 1-μm coherent lidar propagation through atmospheric turbulence," Appl. Opt. **30,** 2609–2616 (1991).
16. K. P. Chan and D. K. Killinger, "Coherent summation of spatially distorted laser Doppler signals by using a two-dimensional heterodyne detector array," Opt. Lett. **17,** 1237–1239 (1992).
17. C. G. Bachman, *Laser Radar Systems and Techniques* (Artech, Wayland, Mass., 1979), Chap. 2.
18. J. W. Goodman, *Statistical Optics* (Wiley, New York, 1985), Chap. 2.
19. This equation results by following an analysis similar to Parsons' analysis for the mean CNR.[3] In this case up to the fourth moment of the field amplitude is required to calculate the second moment of the CNR, which is then used to derive the normalized CNR variance.
20. J. Y. Wang and A. P. Pruitt, "Effects of speckle on the range precision of a scanning lidar," Appl. Opt. **31,** 801–808 (1992).
21. J. F. Holmes, J. S. Peacock, and D. C. Draper, "Optical remote sensing of surface roughness through optical turbulence," Appl. Opt. **33,** 7770–7776 (1994).
22. R. K. Tyson, *Principles of Adaptive Optics* (Academic, New York, 1991), Chap. 2.
23. V. S. Gudimetla and J. F. Holmes, "Probability density function of the intensity for a laser-generated speckle field after propagation through the turbulent atmosphere," J. Opt. Soc. Am. **72,** 1213–1218 (1982).
24. N. E. Zirkind and J. H. Shapiro, "Adaptive optics for large aperture coherent laser radars," in *Laser Radar III*, R. J. Becherer, ed., Proc. SPIE **999,** 117–135 (1989).
25. J. E. Pearson, "Compensation of propagation distortions using coherent optical adaptive techniques (COAT)," in *Optical Design Problems in Laser Systems*, W. R. Sooy, ed., Proc. SPIE **69,** 21–22 (1975).
26. S. A. Kokorowski, T. R. O'Meara, R. C. Lind, and T. Calderone, "Automatic speckle cancellation techniques for multidither adaptive optics," Appl. Opt. **19,** 371–381 (1980).

3.0 TARGET-INDUCED SPECKLE EFFECTS IN LIDAR APPLICATIONS

James E. Harvey, Ph.D.

Coherent lidar return signatures consist of amplitude and phase variations at the receiver. A typical coherent lidar array receiver (CLAR) scenario is illustrated schematically in Figure 3.1. The resulting speckle pattern, whether induced by target surface characteristics or atmospheric turbulence, can significantly decrease the heterodyne efficiency of the CLAR receiver. The spatial and temporal characteristics of the speckle pattern will depend upon the size, shape, material, texture, and motion of the target as well as the wavelength, spectral bandwidth, and state of polarization of the illuminating laser beam. In general the return signature is accompanied by "glints" which significantly complicates the analysis of speckle effects upon CLAR performance. However, because speckle phenomena is inherent in laser radar systems, there exists the possibility of developing techniques for target discrimination using laser speckle.[1]

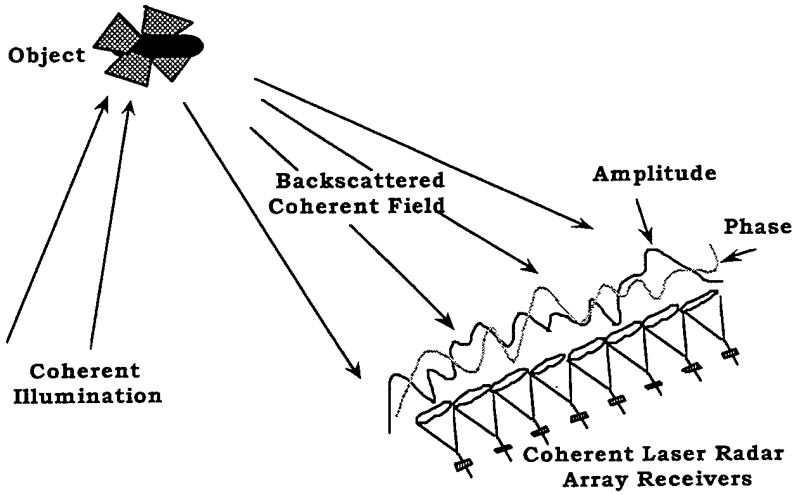


Figure 3.1: Typical CLAR scenario illustrating random phase and intensity variations at the receiver.

Speckle characteristics affecting CLAR experiments include: mean speckle size, mean speckle intensity, mean speckle separation, temporal effects, phase statistics of the received electric field, and polychromatic effects.

For distributed array receivers, the mean speckle size as a function of threshold level is of particular interest. This specific speckle characteristic has been investigated both theoretically and experimentally under this contract. Extensive experimental measurements of average speckle size as a function of intensity threshold level were made for several different targets and illumination conditions. Excellent agreement with theory was obtained for intensity threshold levels greater than approximately twice the mean intensity.

3.1 Background

Speckle is inherently an interference phenomenon produced when a rough object or turbulent medium introduces some degree of randomness to a reflected or transmitted electromagnetic field. Because the vast majority of surfaces are rough on the scale of an optical wavelength, large spatial variations occur in the phase of the

reflected electromagnetic field. As this randomly perturbed wavefront (phase variation) propagates, a three-dimensional speckle pattern is produced.

Figure 3.2 shows the random intensity distribution, or speckle pattern, resulting from the reflection of coherent illumination from a rough object. The speckle pattern consists of a multitude of bright and dark spots of unusual shapes with no preferred orientation and appears unordered and chaotic in nature. Hence, the methods of probability and statistics are used to quantitatively describe the extremely complex patterns of laser speckle, and to relate them to target characteristics.[2]

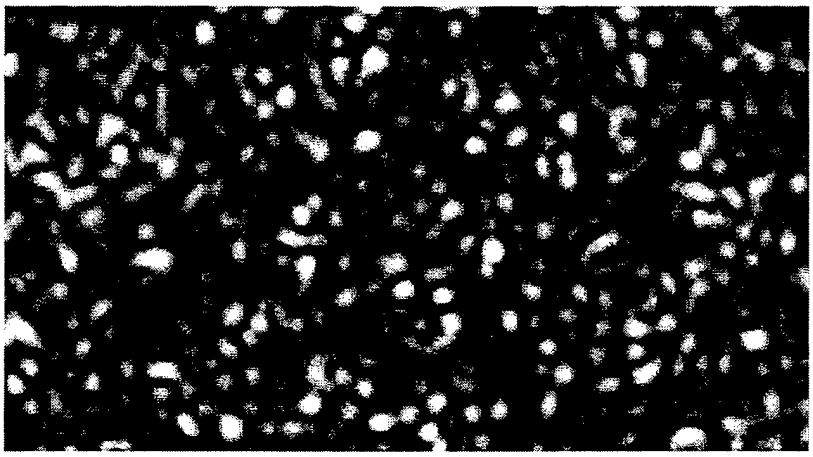


Figure 3.2: Typical laser speckle pattern.

The average size and shape of each speckle lobe, depends on the size and shape of the illuminated portion of the surface or target. A well known relation for the average speckle size can be expressed as

$$d \geq \frac{\lambda L}{D} \tag{3.1}$$

where d is the average speckle diameter (at zero intensity level) due to free-space propagation, λ is the wavelength of the illumination, L is the distance from the target to the receiver, and D is the width of an illuminated square region on the target.[3,4]

The structure of a speckle pattern, including the average size and shape of the individual speckle lobes, depends mainly on the shape of the illuminated portion of the surface or the shape of the target. Since the speckle pattern is an intensity variation over space, any definition of "the average size of a speckle" must be based on the correlation properties of this intensity. Following Goodman [4], the autocorrelation function of the speckle intensity assumes the form

$$\Re_I(\Delta x, \Delta y) = \bar{I}^2 \left[1 + \left| \frac{\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} |P(\xi,\eta)|^2 \exp\left[j\frac{2\pi}{\lambda z}(\xi\Delta x + \eta\Delta y)\right] d\xi d\eta}{\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} |P(\xi,\eta)|^2 \, d\xi d\eta} \right|^2 \right] \tag{3.2}$$

where the function $P(\xi,\eta)$ represents the amplitude of the field incident on the scattering spot. Thus, the intensity autocorrelation function is equal to the magnitude squared of the Fourier transform of the scattering field plus a bias level equal to the squared value of the mean intensity in the speckle pattern.

The *power spectral density* (PSD) function (or power spectrum) of the speckle intensity distribution $I(x,y)$ can be obtained by Fourier transforming the autocorrelation function $\Re_I(\Delta x, \Delta y)$. This is known as the Wiener-Kintchine theorem [5]. Applying a Fourier transformation to Equation (3.6) yields

$$S_I(\nu_x,\nu_y)=\bar{I}^2\left\{\delta(\nu_x,\nu_y)+\frac{\int_{-\infty}^{\infty}\int|P(\xi,\eta)|^2\left|P(\xi-\lambda z\nu_x,\eta-\lambda z\nu_y)\right|^2 d\xi d\eta}{\left[\int_{-\infty}^{\infty}\int|P(\xi,\eta)|^2 d\xi d\eta\right]^2}\right\}. \tag{3.3}$$

The PSD of the speckle pattern consists of a δ-function component at zero frequency, plus a component having the same shape and frequency range of the normalized autocorrelation function of the intensity distribution incident on the scattering area.

Using a square scattering area as an example will help illustrate the relationships described above. Representing the illuminated area on the scattering surface as

$$|P(\xi,0)|^2 = rect(\xi) \tag{3.4}$$

where rect(x)=1 for $|x| \le \frac{1}{2}$, zero otherwise (dimensions of the scattering area are 1 x 1). The autocorrelation function of intensity is

$$\Re_I(\Delta x,0)=\bar{I}^2\left[1+sinc^2\left(\frac{\Delta x}{\lambda z}\right)\right]. \tag{3.5}$$

Now the power spectrum (PSD) of the speckle intensity distribution is obtained by Fourier transforming the autocorrelation function $\Re_I(\Delta x,\Delta y)$ in Equation (3.5) resulting in

$$S_I(\nu_x,0)=\bar{I}^2\left[\delta(\nu_x,0)+(\lambda z)^2 tri(\lambda z\nu_x)\right]. \tag{3.6}$$

These expressions and their relationships to each other and to the speckle intensity distribution $I(x,y)$ are shown in Figure 3-3 for the one dimensional case.

3.1.1 Fully Developed Speckle (FDS)

If the reflecting surface, which is rough on the scale of a wavelength, is so rough that the reflected light produces uniformly distributed random phase changes comparable to 2π radians, then the speckle is called *fully-developed*. In other words, any phase change between $-\pi$ and $+\pi$ is equally likely upon reflection from any point on the surface chosen at random. The speckle effect decreases as the ratio of the surface roughness to the wavelength of illumination is decreased. For example, very little speckle effect could be produced from reflection from a smooth mirror. The relationship is not linear though, because a type of saturation exists in fully-developed speckle.

Once the phase fluctuations exceed 2π radians, any further increases in roughness will produce no perceptible changes in the irradiance statistics.[3]

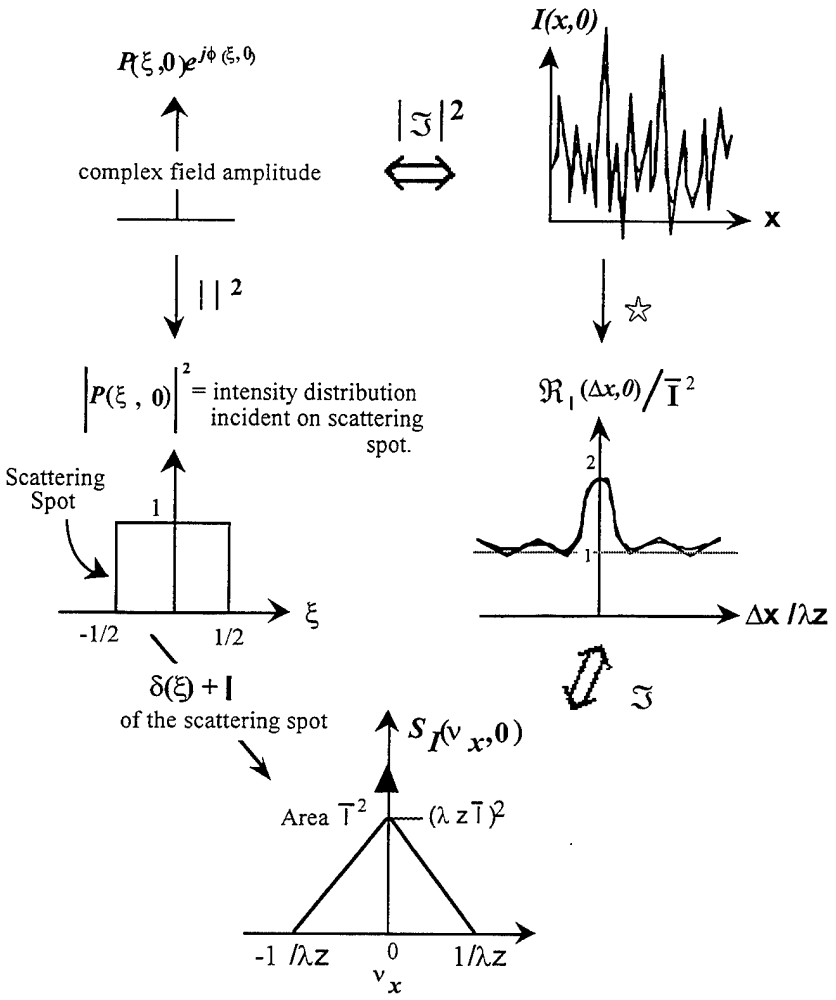


Figure 3.3: Relationships between the pupil function, the intensity distribution, the autocorrelation function of intensity and the power spectral density of intensity. The I denotes autocorrelation and $\Im$ denotes the Fourier transform.

The structure of the surface (roughness) does not affect the size of the speckles, although the actual spatial distribution of the speckle pattern does depend on the properties of the target surface. Two different target surfaces (one may be much rougher than the other), will produce speckles with the same statistics, while the spatial distribution of the two speckle patterns will be completely different.

The most important statistical property of a speckle pattern is the probability density function (PDF) of the intensity *I*. The primary interest in the PDF is the probability that the intensity exceeds a threshold which can be determined from the cumulative distribution function (CDF) by integration over the PDF. What is the probability of observing a bright peak of intensity or a dark null at a given point? Goodman answers this question by treating speckle as a classical random walk phenomenon.

If the number N of elementary contributions is large, then (a) the real and imaginary parts of the complex field at (x,y,z) are independent, zero mean, identically distributed Gaussian random variables, and (b) the intensity I obeys negative exponential statistics, i.e., its probability density function is of the form

$$p_I(I) = \begin{cases} \left(1/\bar{I}\right)\exp\left(-I/\bar{I}\right) & I \geq 0 \\ 0, & \text{otherwise} \end{cases} \tag{3.7}$$

where $\bar{I}$ is the mean or expected intensity.

The above PDF is shown in Figure 3.4 where the intensity has been normalized by the mean intensity $\bar{I}$ and $\bar{I}\,p_I(I)$ is the probability $P(I)$ that the intensity exceeds a threshold I.

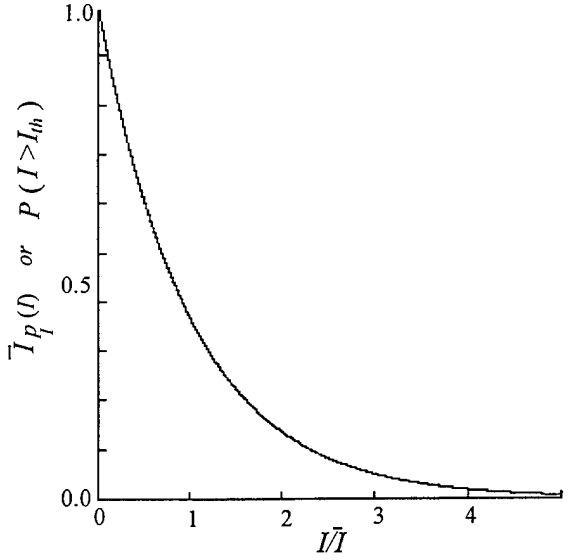


Figure 3.4: Probability density function (PDF) of a speckle pattern produced by linearly polarized light.

Thus, the normalized probability density function of a fully-developed, polarized speckle pattern follows the negative exponential curve shown above. Also, a fundamentally important characteristic of the negative exponential distribution is that its standard deviation precisely equals its mean. Thus, the contrast of a fully-developed, linearly polarized speckle pattern, as defined by

$$C = \sigma_I / \bar{I}\,, \tag{3.8}$$

is always unity.

3.1.2 Partially Developed Speckle (PDS)

Partially developed speckle patterns are produced when a weak diffuser, such as ground glass with a root-mean square (rms) surface roughness smaller than the optical wavelength, is illuminated by coherent light. The weak diffuse object means that it would distribute the phase over less than 2π. The far-field diffraction pattern

generally consists of a central specular peak surrounded by a speckle pattern which is present even at a considerable distance from the peak and follows usual speckle statistics produced by a strong diffuse object.[6] Partially developed speckle patterns have been extensively studied by Dainty,[7] Goodman,[8] Ohtsubo and Asakura [9] and others.

No attempt will be made to present a rigorous theory of partially developed speckle, however, it will be shown that, in many situations, partially developed speckle patterns can be adequately described by combining conventional BRDF data with the usual fully developed speckle statistics.

The speckle patterns shown in Figure 3.5 were computer generated to illustrate the reduction in the speckle effect as the ratio of the surface roughness to the wavelength of illumination is decreased for a rectangular scattering area. The algorithm is based on an amplitude and phase description of the field reflected from the target surface given by

$$E(x,y) = A(x,y)e^{j\phi(x,y)},$$

where (3.9)

$$\phi(x,y) = coef\,\pi\ \ rand(n,n)$$

represents the random phase induced by the n by n target surface. The coefficient (coef) controls the magnitude of uncorrelated random phases (rand(n,n)) varying between zero and one.

A far-field intensity speckle pattern is produced by calculating the squared modulus of the Fourier transform of the field described by Equation 3.5. A coefficient of 2 resulted in the fully developed speckle pattern shown in Figure 3.5(a). Any further increases in the coefficient produce no changes in the nature of the speckle pattern. However, while a coefficient of 0 results in the far-field diffraction pattern for a rectangular aperture shown in Figure 3.5(d), Figures 3.5(b) and (c) are called partially developed speckle patterns and were produced by coefficients (coef) of 1 and 0.35 respectively.

Figure 3.6 verifies the claim that in a partially developed speckle pattern, the region out of the specular beam follows negative exponential statistics. Figure 3.6(a) shows the normalized probability density function for the fully-developed speckle pattern in Figure 3.5(a) and a negative exponential curve for comparison. In Figure 3.6(b) the normalized probability density function for a 200 x 200 pixel area taken from the upper left corner of Figure 3.5(c) just out of the glint region. Again a negative exponential curve is shown for comparison.

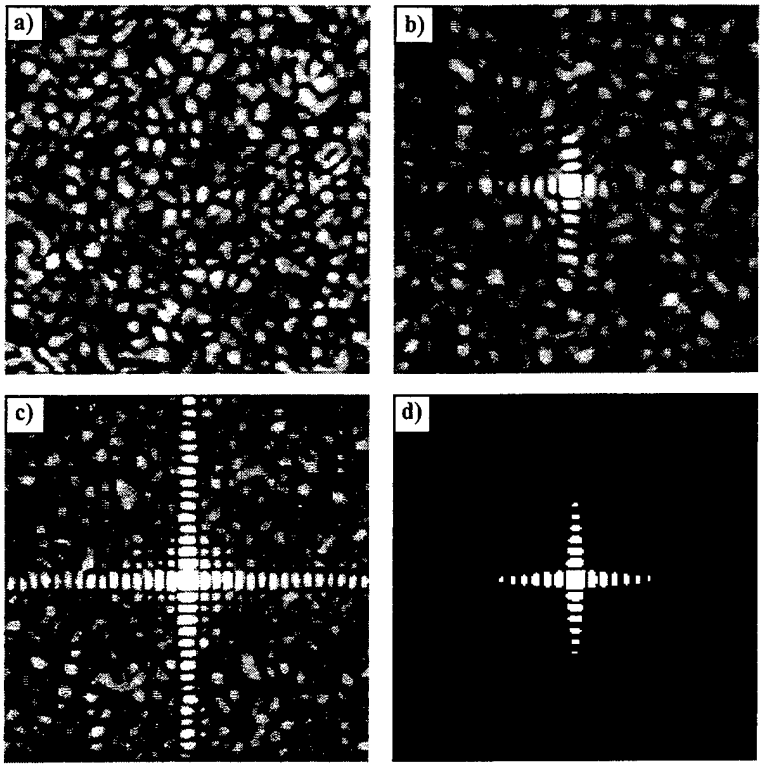


Figure 3.5: Speckle patterns for an increasingly smoother reflecting surface: (a) fully developed speckle, (b) and (c) partially developed speckle, (d) far- field diffraction pattern for a rectangular aperture (no scatter).

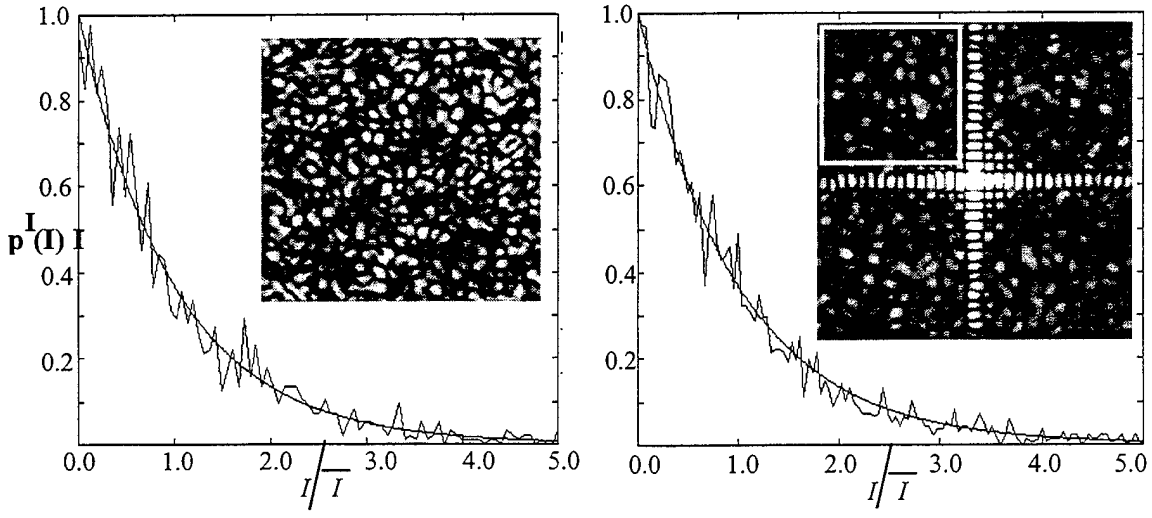


Figure 3.6: a) The normalized PDF for fully developed speckle follows a negative exponential curve, b) for the 'out of the glint' region of a partially developed speckle pattern the PDF also follows negative exponential statistics.

Experimental measurements of partially developed speckle patterns produced by a variety of different target types will be presented in the following sections of this chapter along with the relevant surface characteristics for selected targets.

3.1.3 The Bidirectional Reflectance Distribution Function

The bidirectional reflectance distribution function (BRDF) is the standard radiometric measurement for specifying the angular-scattering distribution of a surface.[1,10] The BRDF describes the scattering of light from a flat surface element in terms of both the illumination direction and the observation direction.

The BRDF is defined as the reflected radiance $dL_r(\theta_r,\phi_r;\theta_i,\phi_i)$ of a surface element dA into a differential solid angle $d\omega_r$ in the direction (θ_r, ϕ_r) divided by the incident irradiance $dE_i(\theta_i, \phi_i)$ producing it.[10] The BRDF is expressed in units of inverse steradians (sr^{-1}).

$$BRDF(\theta_r,\phi_r;\theta_i,\phi_i)=\frac{dL_r(\theta_r,\phi_r;\theta_i,\phi_i)}{dE_i(\theta_i,\phi_i)} \tag{3.10}$$

The angles used in this definition are illustrated in Figure 3.7.

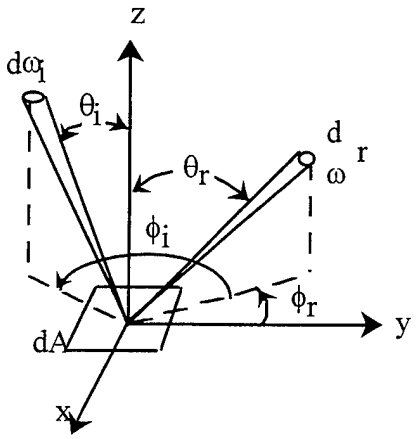


Figure 3.7: Geometry of incident and reflected elementary beams used to define the bidirectional reflectance distribution function (BRDF).

The numerical value of the BRDF for a given pair of incident and reflected directions may vary from zero to infinity. In particular, consider two ideal cases. Theoretically, the BRDF is a constant value for all reflected directions for a perfectly diffuse (Lambertian) surface; and it becomes infinite (as a Dirac delta function) in the specular direction for a perfect mirror surface.

We will show that this conventional (incoherent) BRDF determines the *local* mean intensity of the speckle pattern produced by a given target. Hence, the value of the BRDF (in the direction of the receiver) integrated over the target is necessary in addition to the local normalized speckle statistics when calculating the performance of laser radar receivers.

3.2 Accomplishments

3.2.1 Target Characteristics

In order to produce the desired fully-developed and partially-developed speckle patterns, three different materials were used to develop a wide range of target sizes, shapes and textures. Both planar and three-dimensional

targets such as cylinders were used (see Figure 3.8). A total of eighteen samples of aluminum, steel and glass were subjected to varying degrees and combinations of polishing, sandblasting and acid etching to achieve different surface characteristics. Dr. Jean M. Bennett of the Naval Air Warfare Center, China Lake, California deserves special thanks and acknowledgment for providing surface characterization measurements on many of these samples.

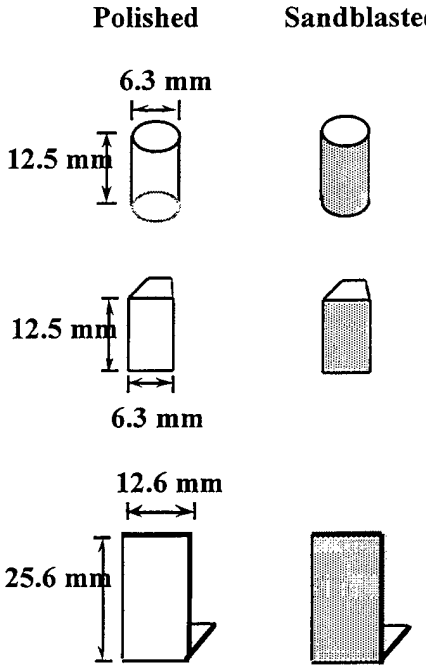


Figure 3.8: Illustration of target sizes and shapes

We illustrate only an example of target surface characterizations in this report. A complete catalog of the target characterizations can be found in Appendix A of Reference.[11]

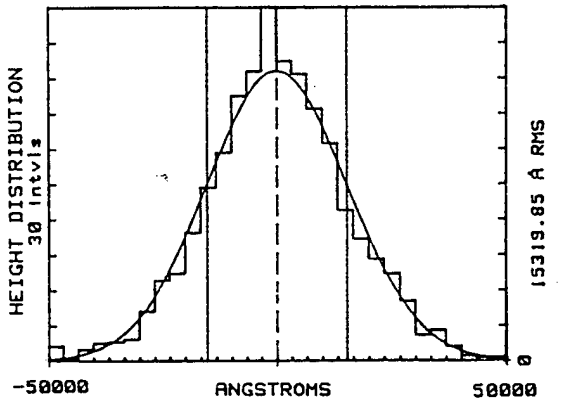


Figure 3.9: Height distribution function for the sandblasted steel (Courtesy of Dr. Jean Bennett).

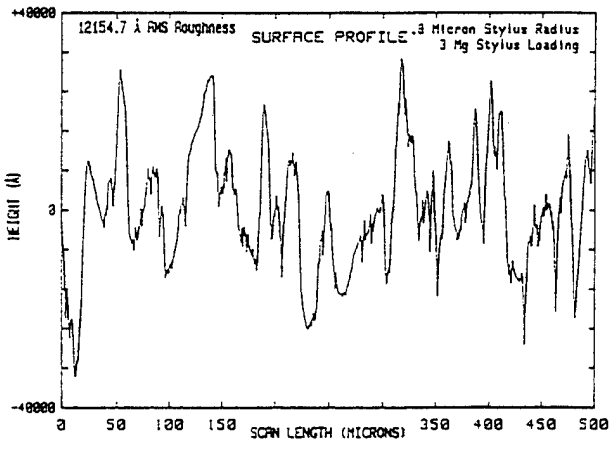


Figure 3.10: Surface profile of the sandblasted steel surface (Courtesy of Dr. Jean Bennett).

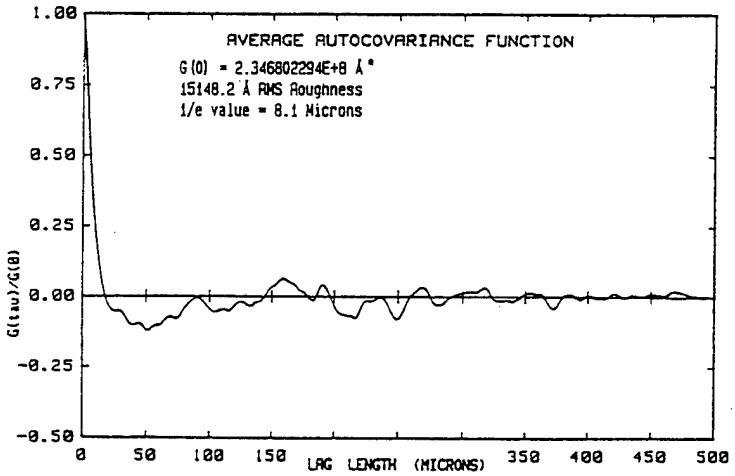


Figure 3.11: Average autocovariance function for the sandblasted steel sample (Courtesy of Dr. Jean Bennett).

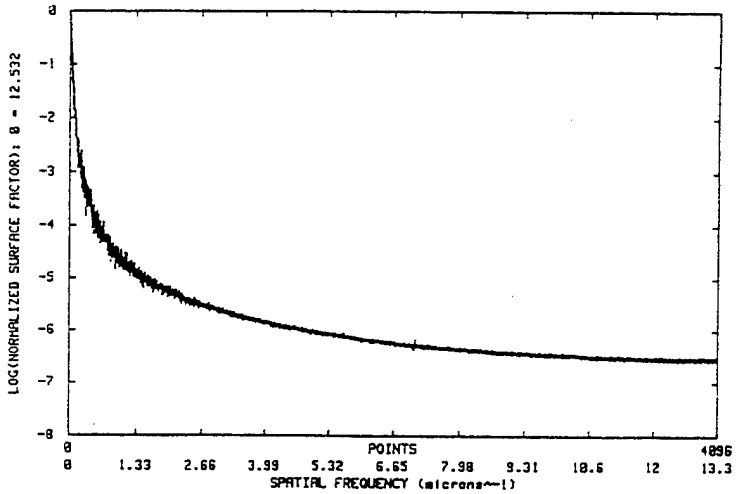


Figure 3.12: Power spectral density function for the sandblasted steel sample (Courtesy of Dr. Jean Bennett).

3.2.2 Experimental Measurements

This section gives a description of the instrumentation and the experimental configurations used for the measurement and collection of BRDF and speckle data.

3.2.2.1 BRDF Measurements

A precision numerically-controlled goniometric radiometer was used for the BRDF surface scatter measurements. The system is fully automated using an A/D converter and a stepping motor controller. A 386 PC with an IEEE-488 bus drives the scanning detector arm, records the scattered intensity measurements, processes the data, and displays the results. The entire instrument (except the computer) is mounted on an optical table which has been enclosed to minimize dust, ambient light and unwanted scattered light within the system. A block diagram of the instrument is shown in Figure 3.13.

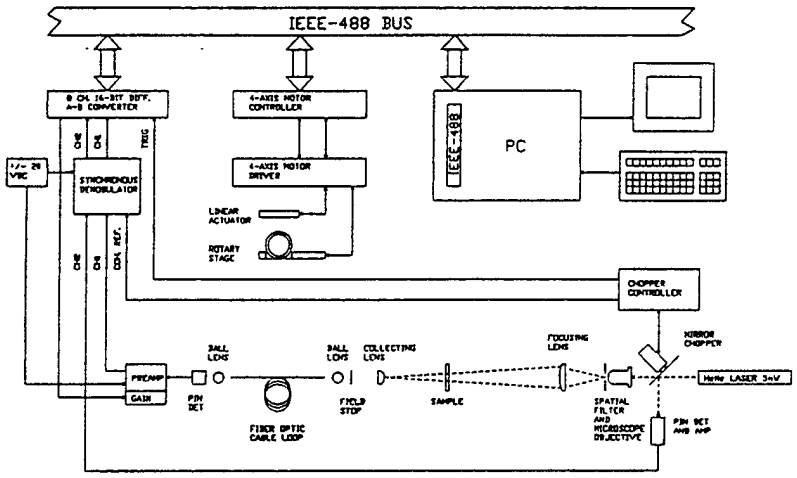


Figure 3.13: Block diagram of the BRDF instrument.

Light incident on the sample is specularly reflected at twice the angle of incidence onto the observation hemisphere. The detector scans a semicircular path in the plane of incidence to measure the angular distribution of scattered light. An optical layout is illustrated schematically in Figure 3.14.

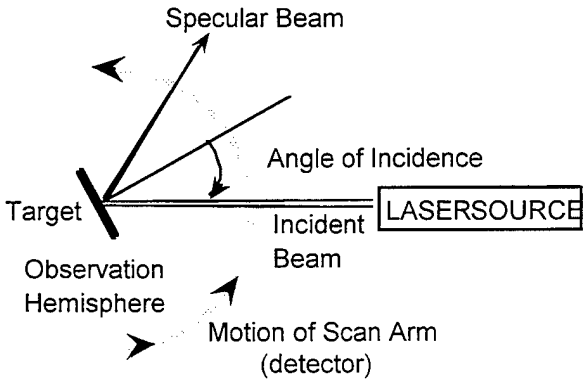


Figure 3.14: Schematic illustration of BRDF instrument.

A typical BRDF curve for a highly reflective (mirror) surface is shown in Figure 3.15. Note that the specular beam is at 10 degrees and the incident beam (-10 degrees) is blocked by the scanning detector when measurements are made in the plane of incidence.

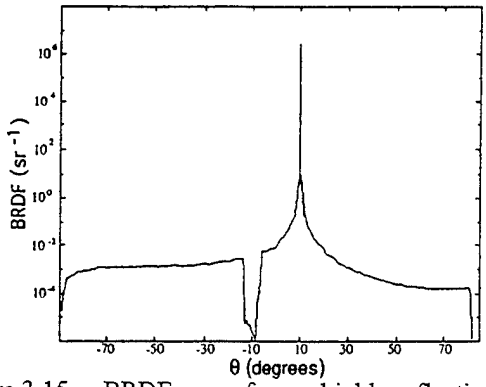


Figure 3.15: BRDF curve from a highly reflective surface.

Of particular interest is the scattering properties of the 3-M retro-reflective tape used to cover the target boards for the field tests of our laser radar experiments. Three different types of retro-reflective tape were examined. Electron micrographs indicated that the tape consisted of a reflective surface covered with micro-spheres which enhance the reflectance in the "retro" direction (see Figure 3.16). The BRDF curves in Figure 3.17 indicate that the particular tape that we chose for our target boards has a reflectance in the retro-direction that is a factor of 200 to 500 greater than a perfectly reflecting Lambertian

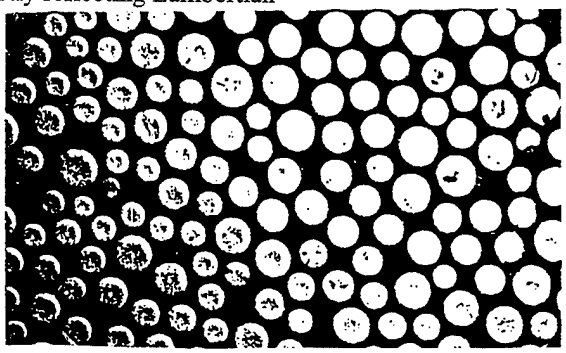


Figure 3.16: Electron micrograph of 3M retro-reflective tape.

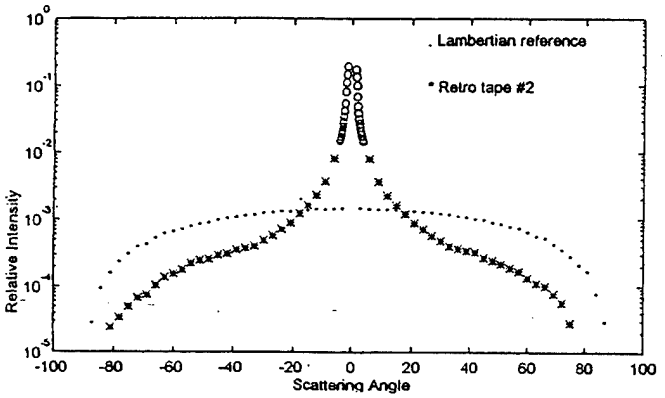


Figure 3.17: BRDF curve of 3M retro-reflective tape.

3.2.2.2 Speckle Measurements

Figure 3.18 shows a schematic diagram of the optical setup used for the collection of speckle data with visible illumination on the target. For these experiments a Particle Measuring Systems Helium-Neon Tunable Laser was used as the light source. Red, orange, yellow and green wavelengths were available with this laser, although most of the experiments were conducted with the red (632.8 nm) wavelength. After the beam exits the laser, it is directed into a 40X microscope objective for beam expansion and a 5 mm pinhole for optical spatial filtering. The collimating lens was mounted on a translation stage making it possible to work with a collimated beam or to use converging or diverging illumination. The beam was then reflected from a mirror which allowed path length adjustability and directional control. When it was necessary to control the size of the beam striking the target an iris was placed between the mirror and the target. The target stage was designed to adapt to the different types of targets allowing for precision positioning and translation or rotation as appropriate. Fully polarized speckle is customarily treated as a special case [8], and a linear polarizer is placed directly in front of the camera to assure that any depolarization effects caused by scattering from the target are eliminated from the speckle data.

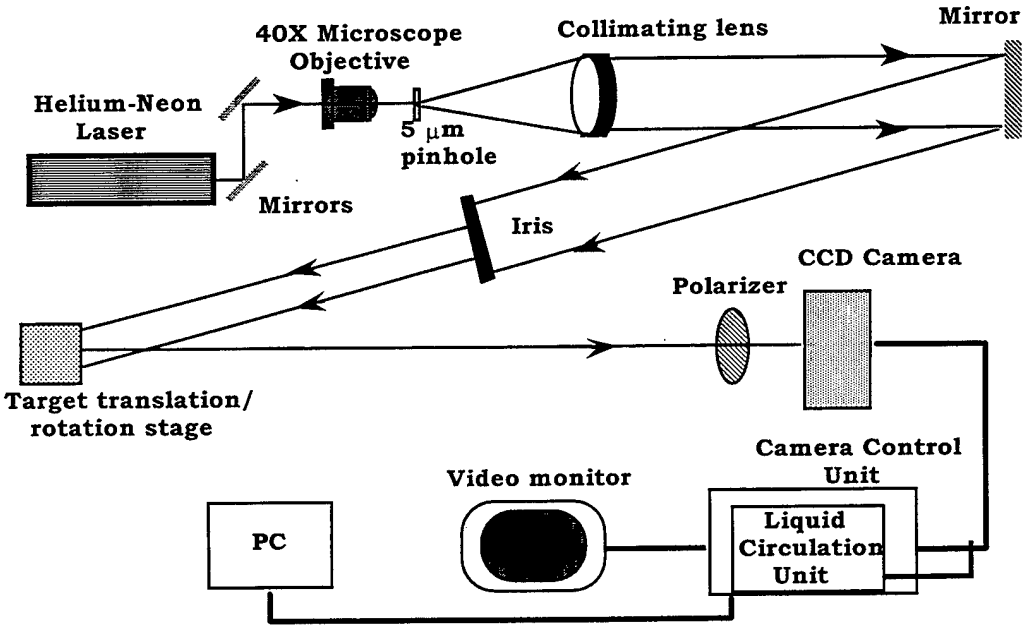


Figure 3.18: Experimental configuration for laser speckle data collection.

The Star I Camera System, a Photometrics Ltd. scientific-grade charge-coupled device (CCD) camera, was selected because of its low noise, high linearity, and high dynamic range. The CCD imager contains 384 x 576 pixels that are 23 mm x 23 mm in size. When a newly acquired image is digitized, the analog-to-digital (A/D) converter assigns a value in A/D units (ADUs) to each pixel. The ADU value, or *pixel intensity value*, is proportional to the number of electrons stored in each pixel and can range from 0 to 4096 (12-bit precision). The camera controller operates the camera head and contains the power supplies, slow scan electronics, and image random access memory used to digitize and temporarily store an image. The liquid circulation unit thermoelectrically cools the camera head which reduces the dark current. The video monitor displays each image

using 256 gray-scale levels. The Star I Camera System is operated through an IEEE-488 interface and imaging processing software using a PC as the host computer.

A few simple preliminary experiments were performed to illustrate the properties of speckle. To illustrate the first-order statistical properties of polarized laser speckle three separate speckle patterns were obtained holding all variables constant (target, illumination size, distance to detector, etc.) except the illumination intensity. The intensity was varied using a neutral density filter and the intensity histograms are shown in Figure 3.19 with the mean and maximum values for each. While the histograms vary significantly with intensity, the normalized probability density functions for the three distributions are virtually identical. Thus, the reason for using a normalized PDF is its insensitivity to the illumination intensity.

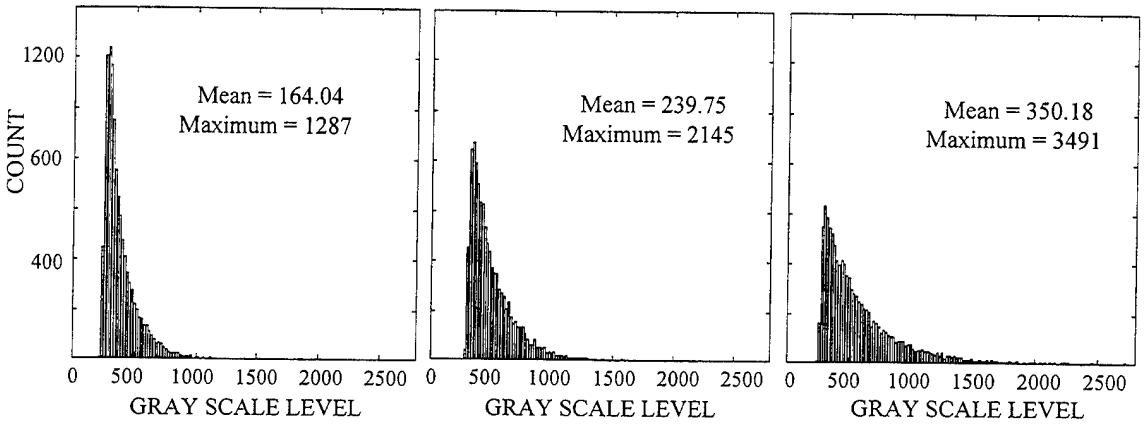


Figure 3.19: Intensity histograms for an experimental configuration where the only changed variable was the illumination intensity.

The normalized PDF for one of the above speckle patterns described above is shown in Figure 3.20 along with a negative exponential curve for comparison and a truncated portion of the corresponding speckle pattern. The nature of this target is such that the resulting speckle pattern is expected to be fully-developed producing a negative exponential PDF distribution. However, real measurements of speckle data can rarely be perfect even in the most controlled circumstances. The departure from the negative exponential curve is undoubtedly due to a variety of measurement error sources as discussed in section 3.2.3.

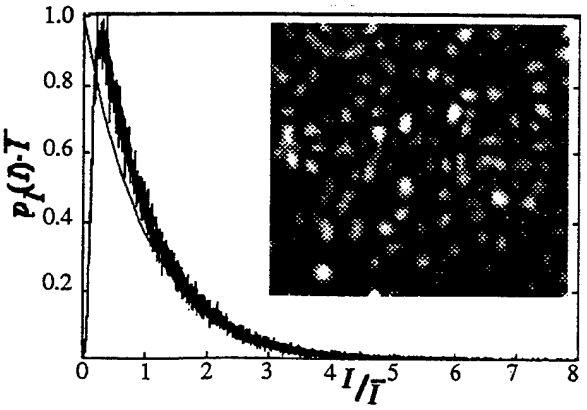


Figure 3.20: Normalized PDF for actual measured speckle data and a negative exponential curve for comparison.

A normalized autocovariance function calculated from measured speckle data from a fully developed speckle pattern is illustrated in Figure 3.21.

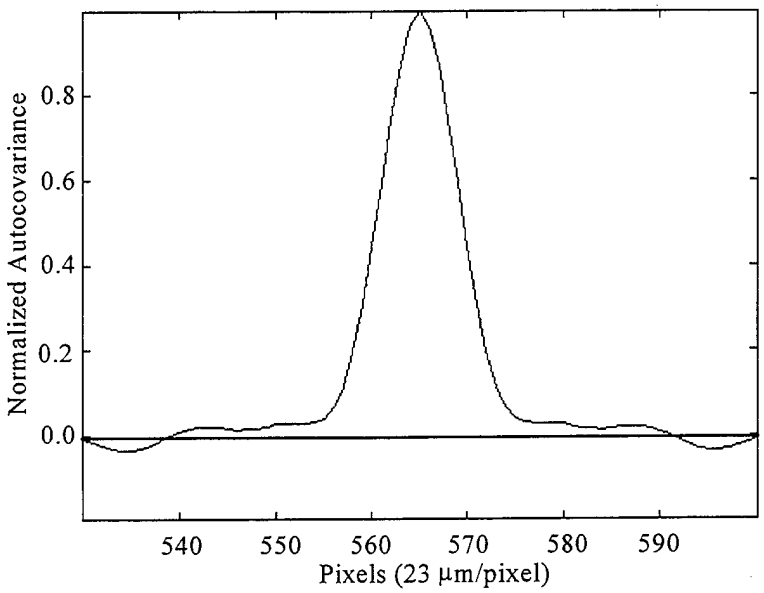


Figure 3.21: Normalized autocovariance function calculated from measured speckle data.

If the reflecting surface is "quasi-smooth" (i.e., rms roughness less than a wavelength), then the far-field partially developed speckle pattern consists of a specular peak or "glint" surrounded by a speckle pattern which exhibits the usual speckle statistics. Figure 3.22 illustrates a portion of the partially developed speckle pattern in the region of the "glint" produced by the polished steel target.

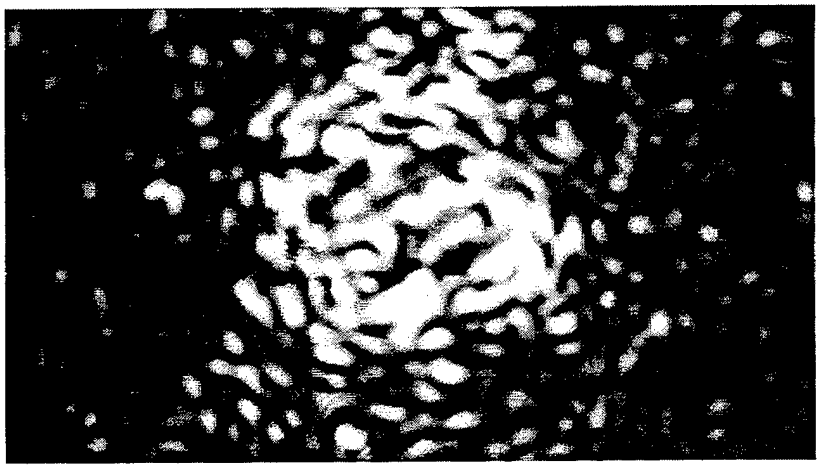


Figure 3.22: Partially-developed speckle pattern (in the glint) produced by the polished steel target.

If we observe and record the intensity speckle pattern far from the region of the glint as illustrated in Figure 3.23, the intensity is obviously much lower; however, the normalized PDF and the normalized autocovariance function from this measured data seems to obey the statistics of fully developed speckle patterns. These quantities are illustrated in Figures 3.24 and 3.25 respectively.

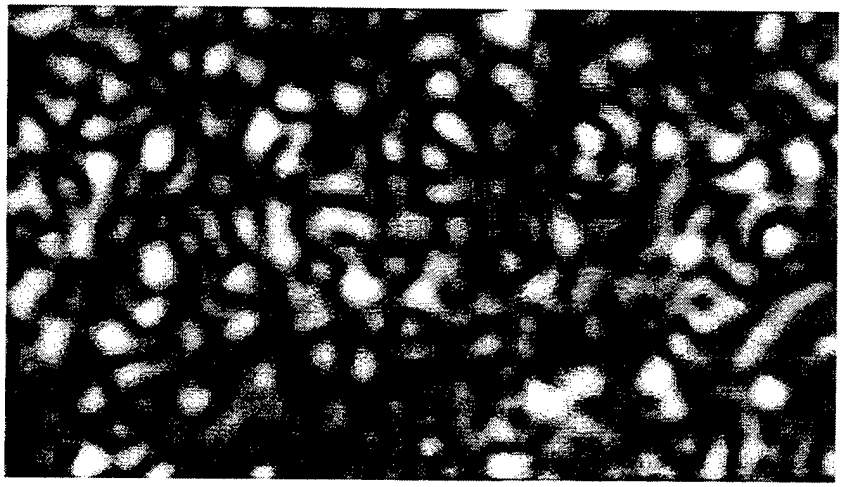


Figure 3.23: Partially-developed speckle pattern (out of the glint) produced by the polished steel out of the glint.

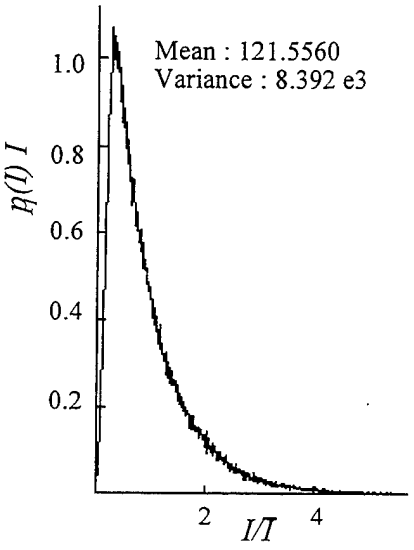


Figure 3.24: Normalized probability distribution function for the speckle pattern shown in Figure 3.22.

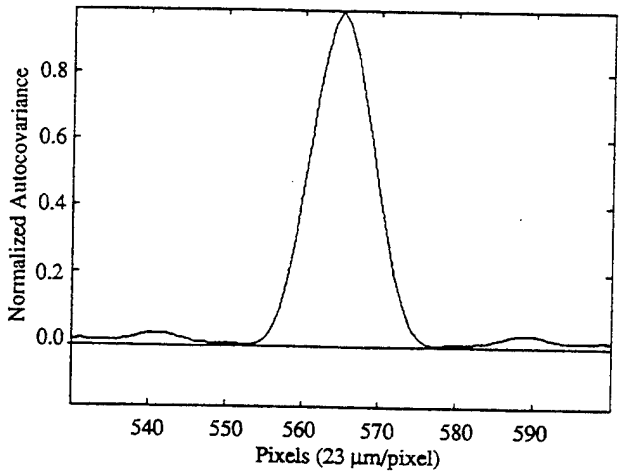


Figure 3.25: Normalized autocovariance function for the speckle pattern shown in Figure 3.22.

One final example of interest is the partially developed speckle pattern produced by a quasi-rough cylindrical target that exhibits a strong "glint", or specularly reflected component. When the CCD camera is in the glint region (Figure 3.26a), extremely elongated speckles are observed. This is to be expected since the specularly reflected cylindrical wavefront appears to arise from a line source at the axis of rotation of the cylinder. Outside of the specularly reflected beam (Figure 3.26b), the speckle pattern is very similar to that observed from a rough planar rectangular target with the same length-to-width ratio as the cylinder.

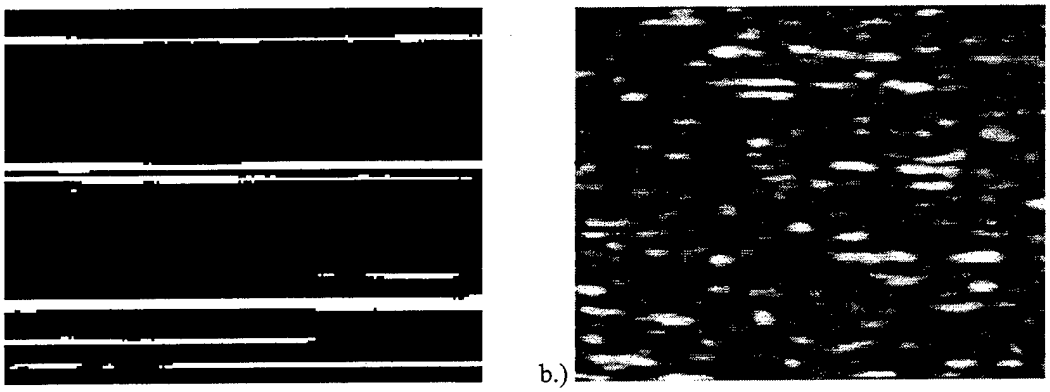


Figure 3.26: Speckle patterns produced by a quasi-smooth cylinder provides information about the physical characteristics of the target: a) in the glint, or specularly reflected beam, and b) out of the glint, in the diffusely scattered region.

Considerable information about this or any other target, including the size, shape and surface texture, can thus be determined by the investigation of the first- and second-order statistical properties of the speckle pattern and the BRDF of the surface. A complete catalog of speckle data from the various targets can be found in Appendices B and C of Reference 11.

3.2.3 Effects of Error Sources upon Measured Speckle PDF's

Experimental measurements of laser speckle are the sum of the actual data and any outside noise contributors that affect the measurement. Factors in the experiment which could degrade the resulting measurement include such things as camera detector dark current and amplifier noise, the finite pixel size of the CCD camera, background radiation, polarization effects, spatial averaging of the speckle pattern due to vibration during the exposure time, turbulence effects in the beam path during the exposure time, finite line-width of the laser beam, etc. Extreme care was taken to control the experimental environment; however, noise still corrupts the negative exponential PDF. It should be remembered that from signal and systems analysis theory, the PDF of the sum of two statistically independent random variables, is equal to the convolution of the two individual PDFs [12]. In other words, for additive noise, the measured speckle intensity $I_m(x,y)$ is

$$I_{measured}(x,y) = I_{speckle}(x,y) + I_{noise}(x,y), \tag{3.11}$$

and the resulting PDF is

$$P_{I_{measured}}(I) = P_{I_{speckle}}(I) * P_{I_{noise}}(I) \ . \tag{3.12}$$

Since there are many independent contributing error sources, the PDF of the noise can be thought of as the convolution of the PDF's of the individual error sources. The central limit theorem thus dictates that the composite noise PDF will approximate a Gaussian function [13] as is illustrated in Figure 3.27.

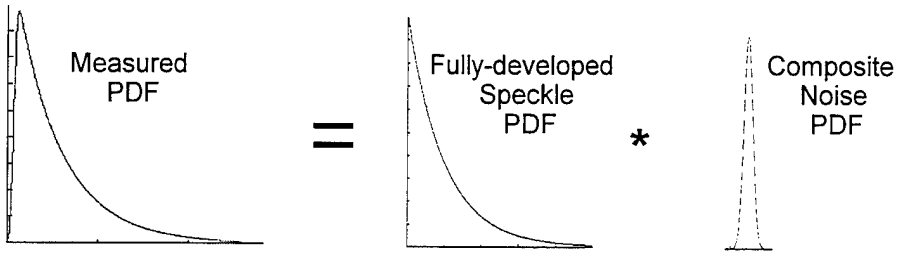


Figure 3.27: For additive noise, the measured PDF is given by the convolution of the ideal fully-developed speckle PDF and the gaussian composite noise PDF.

The specific Gaussian parameters characterizing the composite noise PDF were determined by empirically fitting the convolution of a Gaussian function with the ideal negative exponential PDF to the PDF of the experimentally measured speckle data obtained from a reference surface whose characteristics are such that it should yield a fully-developed speckle pattern:

$$P_{I_{noise}}(I) = \frac{1}{b}\exp\left[-\frac{\pi(x-x_o)}{b}\right]^2, \qquad b = 0.41\,\bar{I}, \quad x_o = b. \tag{3.13}$$

The PDFs from three distinctly different speckle patterns were compared to a theoretical prediction of the PDF based on the composite noise PDF determined above. The first speckle pattern was produced by illuminating a circular region on a sandblasted aluminum target with coherent radiation. The second speckle pattern was produced by flood illuminating a rectangular aluminum target whose surface had also been sandblasted. Both were linearly polarized. The results are shown below in Figure 3.28.

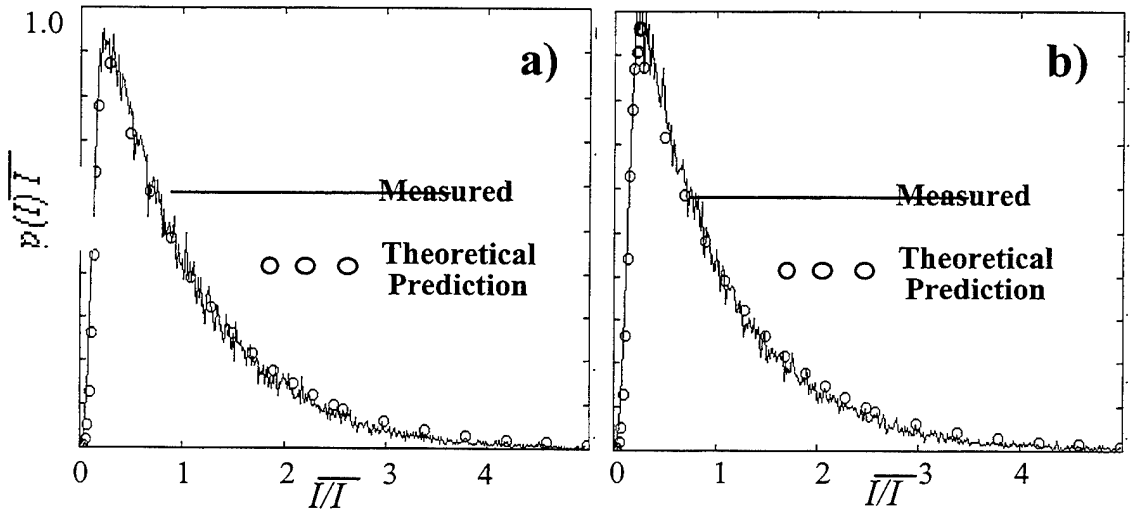


Figure 3.28: The actual PDF and the prediction based on the composite noise PDF for (a) a circular target, and (b) a rectangular target.

Polarization effects upon the speckle PDF are quite pronounced. For circularly polarized light the resulting speckle pattern is the incoherent sum of two orthogonal linearly polarized speckle patterns and, as for the case of unpolarized light, the expected PDF is a Gamma distribution with a shaping constant of M = 2 as shown in Figure 3.29.[14] This particular Gamma distribution is just the convolution of two negative exponential PDFs.

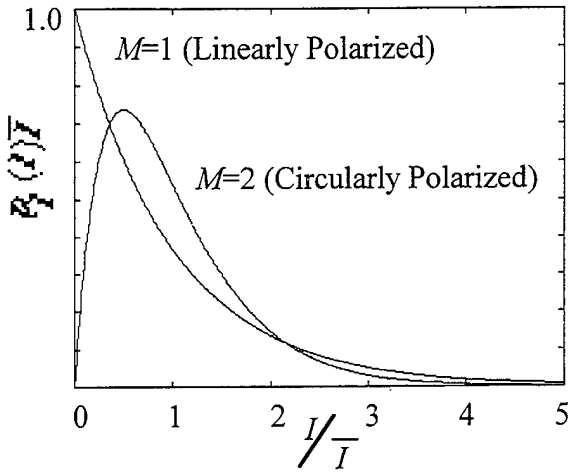


Figure 3.29: The theoretical PDF for speckle patterns produced by linearly polarized and circularly polarized laser light.

The PDF calculated from measured speckle data produced by illuminating a circular region on the sandblasted aluminum target with circularly polarized laser light departs very significantly from the above Gamma distribution. Furthermore, it is not accurately characterized by a Gamma distribution of any shaping constant. However, since both orthogonally polarized speckle patterns are separately corrupted by the composite noise PDF as illustrated in Figure 3.30, the resulting measured speckle PDF is equal to the ideal speckle PDF convolved with a noise PDF resulting from the self-convolution of the composite noise PDF for linearly polarized light.

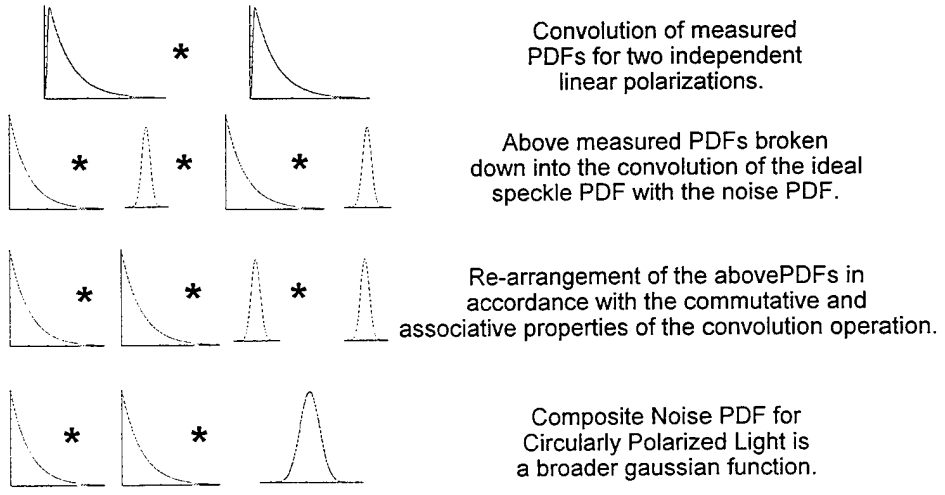


Figure 3.30: Schematic diagram illustrating that the composite noise PDF for circularly polarized light is merely given by the self-convolution of the composite noise PDF for linearly polarized light.

This composite noise PDF for circularly polarized light is thus a Gaussian function which is $\sqrt{2}$ times wider than the composite noise PDF for linearly polarized light.

$$P_{I_{noise}}(I) = \frac{1}{b_2} \exp\left[-\frac{\pi (x - x_o)}{b_2} \right]^2 , \qquad b_2 = b\sqrt{2} \tag{3.14}$$

Figure 3.31 shows the PDF calculated from a measured speckle pattern produced from a circular region on the sandblasted aluminum target compared to the theoretical curve convolved by the composite noise PDF for circularly polarized light described in Eq 3.14. The agreement is again excellent. We have thus experimentally verified our empirically determined composite noise PDF which characterizes our experimental set-up, including all of the potential error sources.

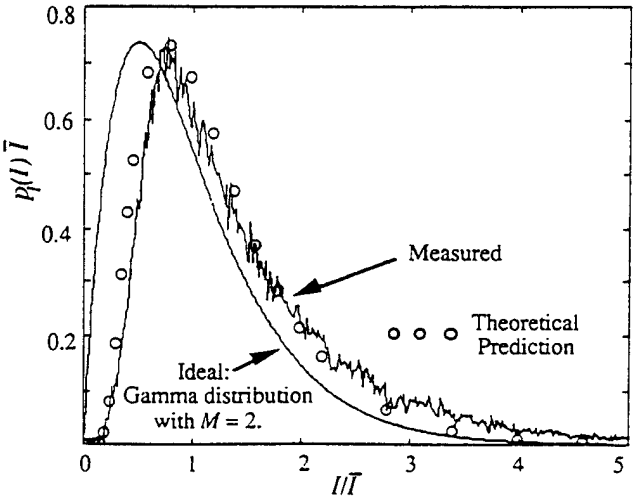


Figure 3.31: The expected Gamma distribution for this speckle pattern and the actual measured PDF with predicted PDF based on the composite noise PDF.

3.2.4 Mean Speckle Size as a Function of Intensity Threshold Level

For laser radar distributed array receivers, the mean speckle size as a function of threshold level is of particular interest. Figure 3.32 shows a three-dimensional isometric plot of speckle pattern with the area at the specific threshold level indicated.

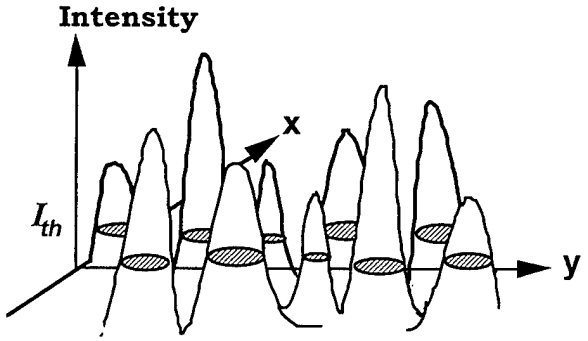


Figure 3.32: Three-dimensional isometric plot of speckle pattern with the area at a particular threshold level indicated.

This average speckle size has been investigated theoretically and reported by Frank E. Kragh in an M.S. Thesis entitled '*Excursion Areas of Intensity Due to Random Optical Waves*' [15]. Kragh's theoretical model is based upon four major assumptions: (i) the intensity PDF of the speckle pattern is a gamma distribution, (ii) the process is spatially ergodic, (iii) the threshold level is high enough to allow only simply connected excursion areas (no doughnut holes), and (iv) the expectation of the ratio of the total speckle area to the total number of speckles is equal to the ratio of the expectation of the total speckle area to the expectation of the total number of speckles, i.e.,

$$\left\langle \frac{\text{total speckle area}}{\text{total number of speckles}} \right\rangle \approx \frac{\langle \text{total speckle area} \rangle}{\langle \text{total number of speckles} \rangle}. \tag{3.15}$$

The above assumptions led to the following theoretical model for the expected speckle size as a function of the normalized intensity threshold level

$$\bar{A} = \frac{A_{tot}}{N} = \frac{\pi \exp\left(M \frac{I_{thres}}{\bar{I}}\right) \Gamma\left(M,\, M \frac{I_{thres}}{\bar{I}}\right)}{\sqrt{\ell_{xx}\ell_{yy}} \left(M \frac{I_{thres}}{\bar{I}}\right)^{M-1} \left(2M \frac{I_{thres}}{\bar{I}} - 2M + 1\right)}, \tag{3.16}$$

where

$\frac{I_{thres}}{\bar{I}}$ = the normalized intensity threshold

$\Gamma\left(M,\, M \frac{I_{thres}}{\bar{I}}\right) = \int_{M \frac{I_{thres}}{\bar{I}}}^{\infty} \exp(-t)\, t^{M-1}\, dt$ = the incomplete gamma function

$M = \bar{I}^2 / \sigma_I^{\,2}$ the shaping constant of that gamma distribution

and

ℓ_{xx} and ℓ_{yy} are the second moments of the normalized spectral density function of the electric field.

A detailed experimental characterization of the average speckle size as a function of intensity threshold level has been performed and is compared with the above theoretical model. For this investigation, speckle data was produced in the laboratory using different types of targets, the variable parameters of Eq. (3.36) were determined for the experimental data, and an image processing code was used to determine the actual area and number of speckles at each intensity threshold level.[16]

The parameters contained in the theoretical model were first determined from experimentally measured intensity speckle data. The gamma distribution shaping constant, M, can be shown to equal the mean intensity squared divided by the variance of the intensity as indicated above. This quantity is readily calculated from the measured speckle data. However, ℓ_{xx} and ℓ_{yy} depend upon the normalized PSD, $S(\xi,\eta)$, of the *field* producing the speckle. Since detectors of optical radiation respond not to field strength, but rather to optical power or intensity, we have to deduce *field* statistics from *intensity* statistics. Fortunately, Siegert's relation

$$\Re_E(x,y) = \sqrt{\Re_I(x,y) - \bar{I}^2} \tag{3.17}$$

allows us to do just that.[17] Figure 3.33 illustrates the relationships allowing us to obtain the normalized *field* PSD from the intensity measurements.

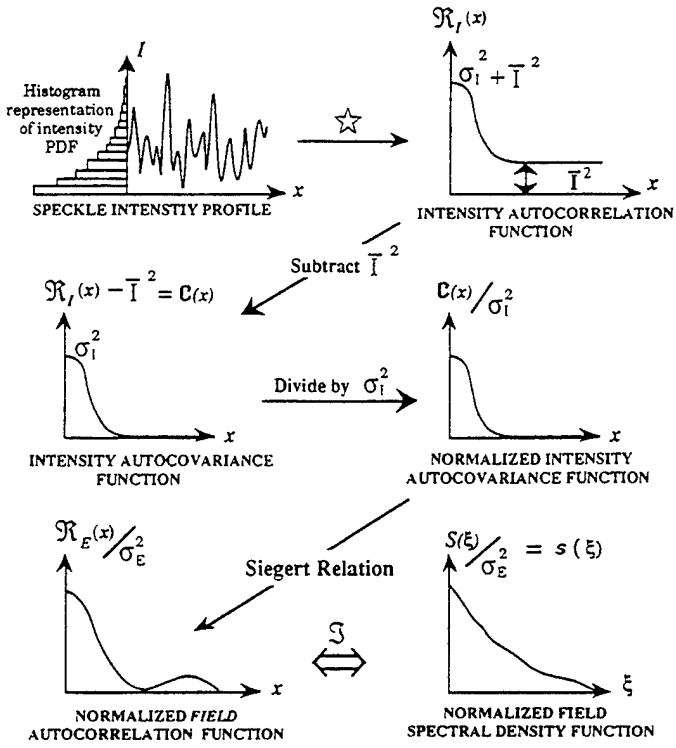


Figure 3.33: Illustration of relationships allowing us to obtain the normalized *field* PSD from the intensity measurements.

The final parameters, l_{xx} and l_{yy}, were then determined by applying the 2nd moment theorem of Fourier transform theory [18] by calculating the peak curvature of the normalized autocorreleation of the field using a parabolic fit to discrete data points.

An image processing computer program was developed to calculate the total area and number of speckles at each threshold level from the experimentally measured and digitized speckle data recorded by the CCD camera. The digitized data was first scanned to find the highest and lowest intensity values recorded. The threshold level was then sequentially increased from the lowest to the highest recorded intensity values. At each threshold level all speckles were identified, counted, and the total speckle area calculated. The average speckle size and the associated threshold level were then stored. The threshold level was then increased and the procedure repeated.

The speckle identification process involves determining the connectivity between pixels to establish boundaries. The concept of *4-connectivity* was used to minimize ambiguities and accurately identify U-shaped and doughnut-shaped speckles [19]. A block diagram of the algorithm used in the code is shown in Figure 3.34.

Comparison of the experimentally measured and theoretically predicted average speckle size vs. intensity threshold level is illustrated in Figures 3.35 and 3.36 for a rotationally symmetric target ($l_{xx} = l_{yy}$) and a non-symmetric (l_{xx} _ l_{yy}) target. Excellent agreement is indicated for both cases for threshold levels above approximately 1.5 times the mean intensity level. Significant departures occur in the comparison at the low intensity levels.

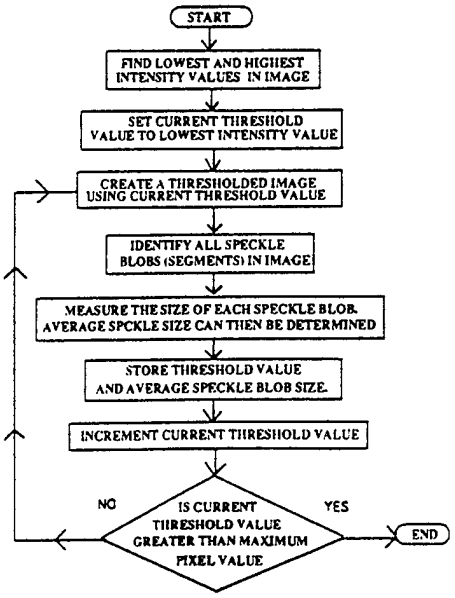


Figure 3.34: Block diagram of the algorithm used for calculating average speckle size from measured speckle data.

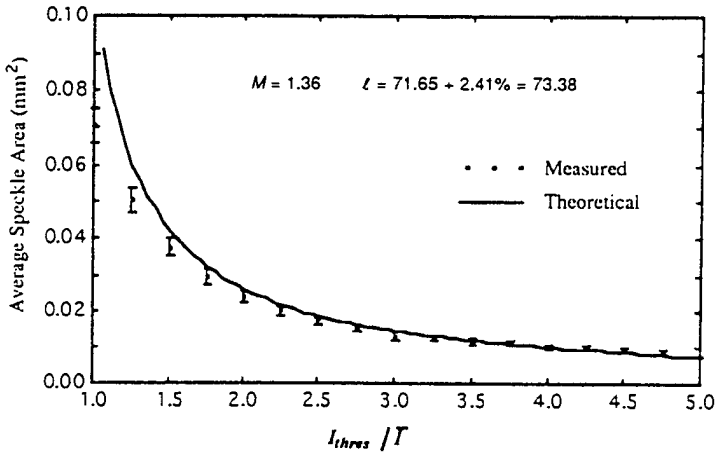


Figure 3.35: Comparison of theoretical predictions of average speckle size as a function of threshold level with experimental measurements for an isotropic speckle pattern.

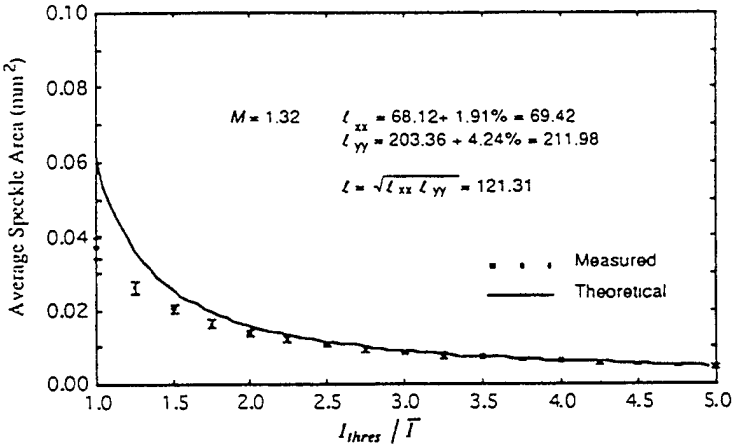


Figure 3.36: Comparison of theoretical predictions of average speckle size as a function of threshold level with experimental measurements for a non-isotropic speckle pattern.

The error bars associated with each measured data point indicate the expected range of experimental values. To determine the error bars, five frames of uncorrelated speckle data was collected from one sample target and the speckle area curve was plotted for each frame. Since no trend was observed in the percent departure from the mean over the range of $I_{thres}/\bar{I}$, the average was used for the error bars.

In both cases for which detailed comparisons are shown, the theoretical curve asymptotically approached the experimental curve at high threshold levels; however, significant departures occur in the comparison at the low intensity threshold levels. Excellent agreement between theory and experiment is observed for intensity threshold levels greater than approximately 2.0 times the mean intensity.

An examination of the initial approximations and assumptions which were made in the development of the theoretical model, and an investigation of the nature of the speckle pattern at low threshold levels helps to explain the departure between the theoretical and experimental curves. The assumption of a Gamma distributed intensity PDF and spatially ergodic speckle data is probably quite accurate; and although we have already seen that the assumption of a Gamma distributed intensity PDF is not strictly valid due to a variety of measurement error sources, it is not clear that the effect of this assumption varies significantly with threshold level. However, we will see that assumptions #3 and #4 discussed earlier are not valid at the lower threshold levels.

Figure 3.37 illustrates speckle data at a threshold level equal to the mean intensity for the 1st speckle pattern investigated. This image clearly exhibits speckles that are not simply-connected (i.e., some doughnut holes are observed). Assumption #3 is therefore not valid at low threshold levels; however, these occurrences do not seem prevalent enough to explain the departure between the two curves at the lower threshold levels.

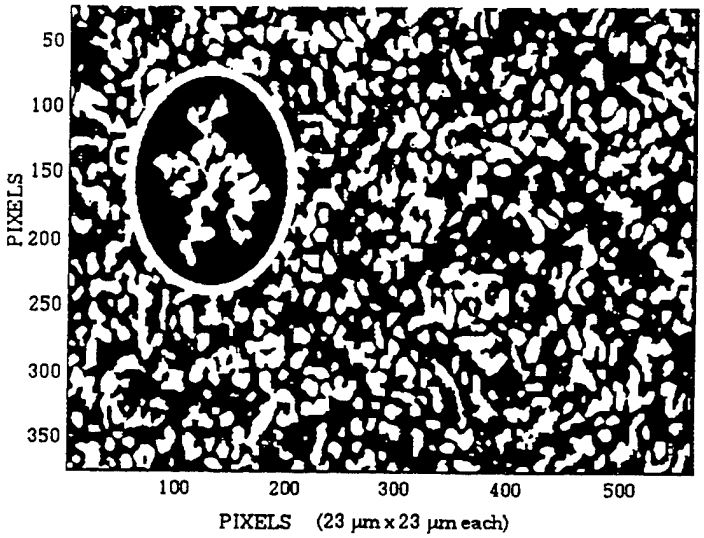


Figure 3.37: Speckle pattern at a threshold level equal to the mean intensity.

More significant than the doughnut holes is the nature of the highlighted speckle in Figure 3.37. Figure 3.38 illustrates the same speckle pattern at four different threshold levels. As the threshold level is reduced the speckles begin to coagulate. A small decrease in threshold may therefore result in a small increase in total speckle area but a drastic change the number of speckles present. For example, at a slightly higher threshold, the total speckle area may change only slightly, but the highlighted speckle may become 5, 6, or even more speckles instead

of one. Thus, assumption #4 is clearly not valid at the lower intensity threshold levels. This effect is undoubtedly the main contributor for the departure between the theoretical curve and the experimental data at the low threshold levels.

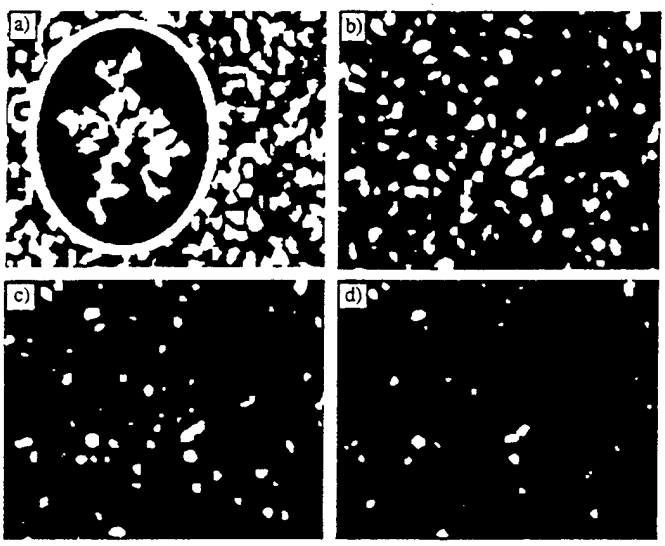


Figure 3.38: Speckle pattern illustrated at four different threshold levels: (a) $I_{th} = \bar{I}$, (b) $I_{th} = 2\ \bar{I}$, (c) $I_{th} = 3\ \bar{I}$, and (d) $I_{th} = 4\ \bar{I}$

As a final check on the theoretical predictions, a computer generated speckle pattern with a nearly perfect negative exponential PDF was treated as actual data and a comparison of the directly measured average speckle size vs. threshold level with the theoretical prediction was made. The results of this comparison exhibited the same behavior as the above laboratory generated speckle samples.

3.3 Conclusions

The interference phenomenon known as laser speckle has been reviewed. Speckle effects, whether induced by target surface characteristics or atmospheric turbulence, routinely cause signal fading and a reduction in the signal-to-noise ratio in laser radar systems. The spatial and temporal characteristics of the speckle pattern will depend on the size, shape, material, texture, and motion of the target as well as the wavelength, spectral bandwidth, and state of polarization of the illuminating laser beam. Additionally, the return signature may be accompanied by 'glints' which complicate the analysis of speckle effects.

A variety of targets were fabricated and characterized. Speckle data was acquired using the different targets and changing the experimental configuration as necessary to study the effects of target size, shape, and texture, as well as illumination wavelength and state of polarization. The speckle patterns and corresponding statistics were discussed.

It was shown that a surface which is smooth on the scale of a wavelength, and accordingly produces a partially-developed speckle pattern, will have fully-developed local speckle statistical properties when observed out of the glint region. A smooth surface and a rough surface of the same size and shape will appear the same when

observed away from the specular component. A demonstration of the normalized probability distribution function's insensitivity to illumination intensity was shown.

It was found that the composite noise PDF (due to the sum of an arbitrary number of error sources producing additive noise) can be experimentally determined and backed out of measured speckle data. This can be done without analyzing the error sources individually.

An extensive experimental investigation of the speckle size as a function of intensity threshold level was conducted. The results were compared to the mathematical model developed by Frank E. Kragh. In the two cases for which detailed comparisons were presented in this report, the theoretical curve asymptotically approached the experimental curve at high threshold levels; however, significant departures occur in the comparison at the low intensity threshold levels. Excellent agreement between theory and experiment is observed for intensity threshold levels greater than approximately 2.0 times the mean intensity. An explanation of the departure between the theoretical and experimental curves was proposed by an examination of the initial approximations and assumptions which were made in the development of the theoretical model.

3.4 References

1. L. G. Shirley, E. D. Ariel, G. R. Hallerman, H. C. Payson, J. R. Vivilecchia, "Advanced Techniques for Target Discrimination Using Laser Speckle", *Massachusetts Institute of Technology Lincoln Laboratory Journal*, Vol. 5 No. 3 (Fall 1992).
2. J. W. Goodman, "Some fundamental properties of speckle", JOSA **66**, 1145-1150 (November 1976).
3. N. George, *Speckle*, a tutorial, The Institute of Optics, University of Rochester, Rochester, New York (November 1979).
4. J. W. Goodman, "Statistical Properties of Laser Speckle Patterns", in *Laser Speckle and Related Phenomena*, edited by J. C. Dainty,Chapter 2 (Springer-Verlag, New York, 1984).
5. J. W. Goodman, *Statistical Optics* pg. 74 (John Wiley & Sons, New York, 1985).
6. J. Ohtsubo, T. Asakura, "Statistical properties of laser speckle produced in the diffraction field", *Applied Optics*, Vol. 16 No. 6 1742 (June 1977).
7. J. C. Dainty, "Coherent Addition of a Uniform Beam to a Speckle Pattern", *Journal of the Optical Society of America*, Vol. 62, 595 (April 1972).
8. J. W. Goodman, "Statistical Properties of Laser Speckle Patterns", in *Laser Speckle and Related Phenomena*, edited by J. C. Dainty,Chapter 2 (Springer-Verlag, New York, 1984).
9. J. Ohtsubo, T. Asakura, "Statistical properties of the sum of partially developed speckle patterns", *Optics Letters*, Vol. 1 No. 3 98 (September 1977).
10. F. E. Nicodemus, "Reflectance Nomenclature and Directional Reflectance and Emissivity," *Applied Optics*, Vol. 9 1474 (1970).
11. T. L. Alexander, *Target Induced Speckle Effects in Laser Radar Applications*, M. S. Thesis, University of Central Florida, (1994).
12. G. Cooper, C. McGillem, "*Probabilistic Methods of Signal and System Analysis*', pg. 126-129 (Harcourt Brace Jovanovich College Publishers, 1986).
13. Gaskill, "*Linear systems, Fourier Transforms, and Optics*", pg. 164 (John Wiley & Sons, Inc. New York 1978).
14. J. W. Goodman, *Statistical Optics* pg. 124-127 (John Wiley & Sons, New York, 1985).
15. F. E. Kragh, *Excursion Areas Of Intensity Due To Random Optical Waves*, M.S. Thesis, University of Central Florida, Electrical Engineering Dept. (1990).

16. T. L. Alexander, J. E. Harvey, and A. R. Weeks, "Average Speckle Size as a Function of Threshold Level: Comparison of Experimental Measurements with Theory", Appl. Opt. **33**, 8240-8250 (1994).

17. E. Jakeman, "Photon Correlation", in *Photon Correlation and Light Beating Spectroscopy*, edited by H. Z. Cummins and E. R. Pike pg. 75-149 (Plenum Press, New York, 1974).

18. R. Bracewell, *The Fourier Transform and its Applications,* pg. 141 (McGraw-Hill Book Company, New York, 1965).

19. R. C. Gonzalez and R. E. Woods, *Digital Image Processing*, pg. 41 (Addison Wesley, Reading , MA, 1992).

3.5 Publications, Presentations, and Theses

1. T. L. Alexander, J. E. Harvey, and D. Hefele, "Experimental Verification of a Theoretical Model for Speckle Intensity Excursion Areas", Proc. SPIE **2222-63** (1994).

2. Terri L. Alexander, *Target Induced Speckle Effects in Laser Radar Applications*, M. S. Thesis, University of Central Florida (May 1994).

2. T. L. Alexander, J. E. Harvey and A. R. Weeks, "Average Speckle Size as a Function of Threshold Level: Comparison of Experimental Measurements with Theory", Appl. Opt. **33**, 8240-8250 (10 Dec1994).

3.6 Appendix 3A

T. L. Alexander, J. E. Harvey and A. R. Weeks, "Average Speckle Size as a Function of Threshold Level: Comparison of Experimental Measurements with Theory", Appl. Opt. **33**, 8240-8250 (10 Dec1994).

a reprint from Applied Optics

Average speckle size as a function of intensity threshold level: comparison of experimental measurements with theory

Terri L. Alexander, James E. Harvey, and Arthur R. Weeks

Speckle phenomena result whenever spatially coherent radiation is reflected from a rough surface or propagated through a random medium such as atmospheric turbulence. Speckle characteristics are therefore a major concern in many laser-imaging or wave-propagation applications. We present the results of experimental measurements of target-induced speckle patterns produced in the laboratory from a variety of targets and illumination conditions. We then compare these experimental measurements with a theoretical model for the average speckle size as a function of intensity threshold level. Excellent agreement is obtained for intensity threshold levels greater than approximately twice the mean intensity level.

Key words: Speckle phenomena, random surface scattering, binary speckle imaging.

Introduction

Speckle is inherently an interference phenomenon produced when a rough object or a turbulent medium introduces some degree of randomness to a reflected or a transmitted electromagnetic field. Speckle is therefore present in laser and microwave echoes and transmissions. It is also seen in stellar images, microscope images, and microfilm viewers. Speckle constitutes a basic noise phenomenon limiting the working resolution of coherent systems in acoustics, microwave and infrared technology, and visible optics.[1-4] The spatial and temporal characteristics of the speckle pattern depend on the size, shape, material, texture, and motion of the illuminated object as well as the wavelength, spectral bandwidth, and state of polarization of the illuminating radiation.

A typical application in which the effects of speckle is important is a coherent laser radar system. The return signature at the receiver consists of a random speckle pattern. Speckle effects, whether induced by target surface characteristics or by atmospheric turbulence, routinely cause signal fading and a reduction in the signal-to-noise ratio in laser radar systems. Coherent laser radar array receivers (CLAR's) are being explored in an effort to mitigate these effects.[5-7] A typical CLAR setup is illustrated in Fig. 1, showing amplitude and phase variations at the array receiver. Speckle characteristics affecting CLAR experiments include mean speckle size, mean speckle strength or intensity, mean speckle separation, temporal effects, phase statistics of the received electric field, and polychromatic effects.

A common expression for the minimum speckle diameter (d) that is due to free-space propagation is determined by the illuminating laser wavelength (λ), the distance of the target from the receiver (L), and the size of the illuminated region on the target (D) (Ref. 1):

$$d \geq \lambda L/D. \tag{1}$$

However, for many applications, a more detailed description of speckle size as a function of intensity threshold level is desirable. The correlation properties of thresholded (or clipped) time-varying laser speckle have been studied in detail.[8-11] Figure 2(a) shows a typical speckle pattern produced when laser light is reflected from a rough surface. A three-dimensional isometric plot of the intensity speckle pattern illustrating the speckle size at a particular intensity threshold level is shown schematically in Fig. 2(b).

Theoretical Development

A theoretical investigation entitled "Excursion areas of intensity due to random optical waves" was con-

The authors are with the Center for Research and Education in Optics and Lasers, University of Central Florida, 12424 Research Parkway, Orlando, Florida 32826.

Received 7 March 1994; revised manuscript received 5 July 1994.
0003-6935/94/358240-11$06.00/0.

© 1994 Optical Society of America.

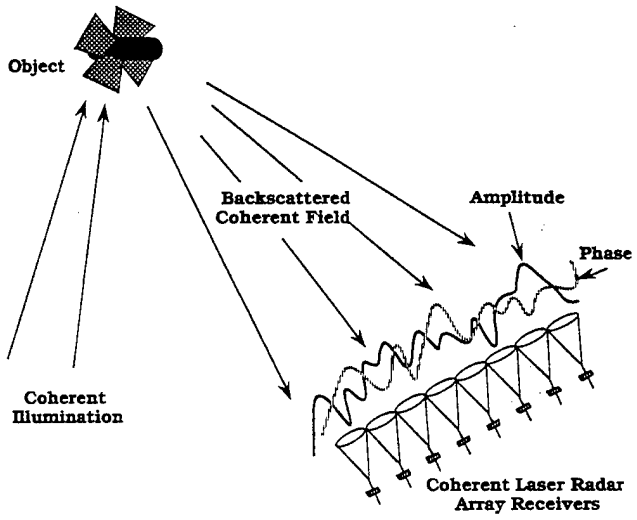


Fig. 1. Typical CLAR setup illustrating random phase and intensity variations at the receiver.

ducted by Kragh in a master's thesis at the University of Central Florida in 1990.[12] The theoretical model developed by Kragh approximates the expected area of excursion regions due to a random intensity. The analysis is general and applies to any two-dimensional gamma-distributed process, including laser speckle. This is an appropriate assumption for laser light that has propagated through the atmosphere because the gamma probability density function (PDF) is applicable for the intensity of a sum of independent Gaussian fields and may be used to describe the image-intensity statistics in coherent imaging.[3]

Kragh defines the average size of a single speckle excursion area as

$$\overline{A} = \left\langle \frac{\text{Speckle area above threshold}}{\text{Total number of speckles}} \right\rangle. \quad (2)$$

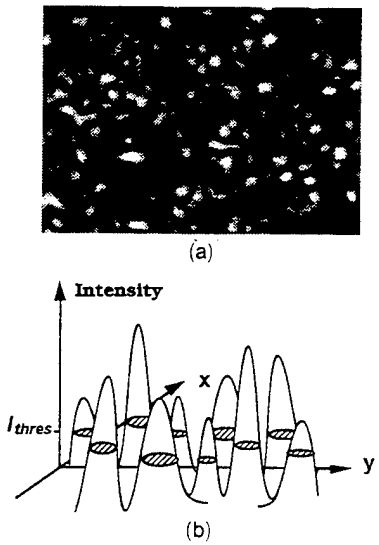


Fig. 2. (a) Typical speckle pattern produced by illumination of a rough surface with laser light; (b) a schematic illustration of speckle size at a specific intensity threshold level.

He then makes the assumption that the expected value of the ratio is equal to the ratio of the expected values, i.e.,

$$\overline{A} = \frac{\langle \text{Speckle area above threshold} \rangle}{\langle \text{Total number of speckles} \rangle} = \frac{A_{tot}}{\overline{N}}, \quad (3)$$

where A_{tot} is the fraction of the speckle area above a given intensity threshold level (I_{thres}) and $\overline{N}$ is the average number of speckles present at I_{thres}. Equation (3) is not strictly true because the numerator and the denominator are not independent. However, in spite of their mutual dependence, this approximation is reasonable at high-intensity threshold levels because both the number of speckles and the total speckle area decrease at a similar rate as the intensity threshold increases. Figure 2(b) gives a visual representation of the average speckle size for an intensity threshold of I_{thres} with the shaded areas representing the excursion areas of interest.

The fractional speckle area above a given threshold intensity I_{thres} in terms of the intensity PDF $p_I(I)$ is given by

$$A_{tot} = S \int_{I_{thres}}^{\infty} p_I(I) \mathrm{d}I, \quad (4)$$

where S is the total illuminated area. As presented by Goodman,[3] the intensity statistics for coherent light illuminating a rough surface is predicted by the gamma PDF:

$$p_I(I) = (M/\bar{I})^M I^{M-1} \frac{\exp[-M(I/\bar{I})]}{\Gamma(M)} \quad \text{for } I \geq 0, \quad (5)$$

where M is the shaping constant, $\bar{I}$ is the average intensity, and $\Gamma(M)$ is the gamma function. Figure 3 shows a plot of $p_I(I)$ as a function of I for various values of M. The horizontal axis is normalized with respect to the average intensity to remove its dependence on the curves. For the special case of $M = 1$ the gamma PDF reduces to the negative exponential PDF. This behavior is characteristic of fully devel-

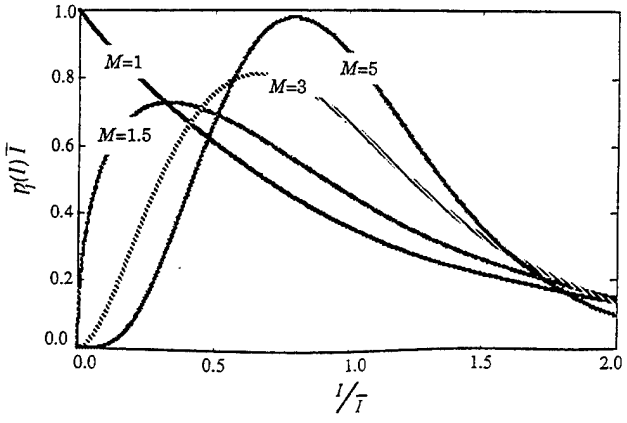


Fig. 3. Gamma probability-distribution function plotted for several values of M.

oped speckle, which results when an electromagnetic wave is reflected from a surface whose roughness is large compared with the wavelength of the incident radiation, and the extent of the incident wave spans many correlation widths of the surface irregularities.[13]

If the surface of the scattering object is rough on the scale of an optical wavelength and the scattered phases are uniformly distributed over $(-\pi, \pi)$, there are a great many independent phase contributions, and the individual linear polarization components of the speckle field follow circular complex Gaussian statistics. Hence the intensity distribution for each polarization obeys negative exponential statistics.[3] This implies an M value of one in Fig. 3. The total detected intensity distribution for the unpolarized condition yields the sum of two negative exponentially distributed intensities, which results in a gamma PDF with a shaping constant (M) of two.[3]

Placing Eq. (5) into Eq. (4) and performing the integration yields

$$A_{\text{tot}} = S\,\frac{\Gamma(M, MI_{\text{thres}}/\bar{I})}{\Gamma(M)}, \tag{6}$$

where

$$\Gamma(M, MI_{\text{thres}}/\bar{I}) = \int_{MI_{\text{thres}}/\bar{I}}^{\infty} \exp(-t)t^{M-1}\mathrm{d}t \tag{7}$$

is the incomplete gamma function.[14]

The derivation of the average number of speckles is not as straightforward as determining the fractional speckle area (A_{tot}) above the intensity threshold I_{thres}. To reduce the complexity of the problem, Kragh made two additional assumptions. The first states that all excursion areas are simply connected. In other words there are no doughnut-shaped speckles (refer to Fig. 4, last speckle on the right). This assumption is obviously valid for high-intensity threshold levels but is not valid for threshold levels much lower than the mean intensity. The second assumption is that the variations in intensity are spatially ergodic, permitting the joint probability distributions of the random field to be determined from a single realization.

Kragh's derivation of the expected number of excursion areas (speckles) at a particular intensity threshold level is a two-dimensional analog of a method for counting excursion intervals of a one-dimensional function developed by Middleton.[15] Consider the excursion areas in Fig. 4. If an individual excursion area is approached from the left, in the x direction, the front edge of the area will have a combination of concave and convex sections. Traveling in the x direction, the intensity is increasing on one side of the area and decreasing on the other side. This is described as I_x (the subscript denotes the first partial derivative with respect to x) being positive on one side and negative on the other side, while between these two extremes, I_x must pass through zero an odd number of times to ensure that the sign changes from one side to the other.

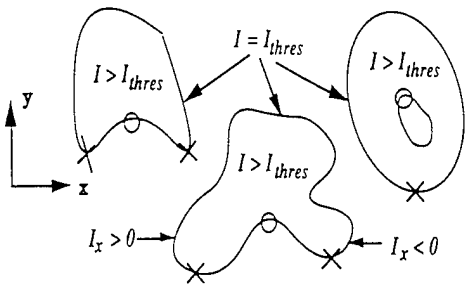


Fig. 4. Contour plots of $I = I_{\text{thres}}$. Points marked with an X satisfy Eq. (8), and points marked with an O satisfy Eq. (9).

On the outer edge of the contour the points at which I_x changes from positive to negative occur once for every convex section, and at this point $I_x = 0$ and $I_{xx} < 0$, which indicates that the intensity reaches a local maximum in the x direction. This situation is indicated with X's in Fig. 4 and shows that, at these local maximums, forward or backward movement in the x direction results in being outside the excursion area.

The concave sections are similarly identified by the points at which $I_x = 0$ and $I_{xx} > 0$ (I_x changes from negative to positive as x increases). At these points (identified by circles in Fig. 4) a local minimum is reached in the x direction, and forward or backward movement in the x direction results in a location inside the area.

Within each of the thresholded excursion areas, the intensity may increase and decrease many times while staying above the threshold level. However, the net change in I_x across an excursion area is from positive to negative; therefore I_x must change from positive to negative one time more than from negative to positive. As a result, there will always be one convex section more than there are concave sections.

Each convex section has exactly one point satisfying the conditions

$$I = I_{\text{thres}}, \quad I_x = 0, \quad I_y > 0, \quad I_{xx} < 0, \tag{8}$$

and each concave section has exactly one point satisfying the conditions

$$I = I_{\text{thres}}, \quad I_x = 0, \quad I_y > 0, \quad I_{xx} > 0. \tag{9}$$

For each excursion area the difference between the number of points satisfying conditions (8) (N_a) and the number of points satisfying conditions (9) (N_b) is one, and the total number of excursion areas is the difference in the total number of points satisfying each condition:

$$\overline{N} = \overline{N_a} - \overline{N_b}. \tag{10}$$

In deriving Eq. (10), we assumed that the boundaries were smooth, and excursion areas with

$$I = I_{\text{thres}}, \quad I_x = 0, \quad I_y > 0, \quad I_{xx} = 0$$

were not included (i.e., speckles of zero area were not counted).

The number of points satisfying the condition of Eq. (8) is

$$N_a = \iint_S \delta[I(x,y) - I_{\text{thres}}, I_x(x,y)]\text{step}[I_y(x,y)] \times \text{step}[-I_{xx}(x,y)]J(x,y)\text{d}x\text{d}y, \quad (11)$$

where $\delta(t_1, t_2)$ is the two-dimensional delta function, step(x) is the unit step function, and $J(x, y)$ is the Jacobian for the mapping of $[I(x, y), I_x(x, y)]$ to (x, y).

The delta function in Eq. (11) ensures that only points that satisfy $I = I_{\text{thres}}$ and $I_x = 0$ are included in the calculation of N_a. The first step function ensures that only points with $I_y > 0$ are counted; the second step function ensures that only points with $I_{xx} < 0$ are counted. The Jacobian is included to permit integration over variables other than the parameters of the delta function, i.e., integration over x and y with I and I_x as the parameters of the delta function.

By analogy the number of points satisfying conditions (9) is

$$N_b = \iint_S \delta[I(x,y) - I_{\text{thres}}, I_x(x,y)]\text{step}[I_y(x,y)] \times \text{step}[I_{xx}(x,y)]J(x,y)\text{d}x\text{d}y. \quad (12)$$

Using Eqs. (11) and (12) and substituting them into Eq. (10) yields

$$\overline{N} = -S \int_{I_y>0} \int_{\text{all } I_{xx}} I_y I_{xx} p_{II_xI_yI_{xx}}(I_{\text{thres}}, 0, I_y, I_{xx})\text{d}I_y\text{d}I_{xx}, \quad (13)$$

where $p_{II_xI_yI_{xx}}(I_{\text{thres}}, 0, I_y, I_{xx})$ is the joint PDF between I, I_x, I_y, and I_{xx}. Because I and I_x are independent of I_y and I_{xx}, Eq. (13) can be rewritten as

$$\overline{N} = -SE\{I_y^+ I_{xx}|_{I=I_{\text{thres}} \text{ and } I_x=0}\}p_{II_x}(I_{\text{thres},0}), \quad (14)$$

where

$E\{\alpha\}$ is the expected value of α and $I_y^+ = \max(0, I_y)$.

The approach taken by Kragh was to take the expectation of Eq. (14) with the chi-squared PDF for the joint PDF and then to perform a transformation of variables to yield the solution for a gamma PDF:

$$\overline{N} = S\frac{(\lambda_{xx}\lambda_{yy})^{1/2} I_{\text{thres}}^{M-1}(M/\bar{I})^M}{\pi\Gamma(M)} \times (2MI_{\text{thres}}/\bar{I} - 2M + 1)\exp[-MI_{\text{thres}}/\bar{I}]. \quad (15)$$

The parameters λ_{xx} and λ_{yy} can be expressed in terms of the second moments, ℓ_{xx} and ℓ_{yy}, of the power spectral density of the random electric field normalized by the variance of the electric field, $s(\omega_x, \omega_y)$:

$$\lambda_{xx} = \ell_{xx}(\bar{I}/2M), \qquad \lambda_{yy} = \ell_{yy}(\bar{I}/2M), \quad (16)$$

where

$$\ell_{xx} = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} \omega_x^2 s(\omega_x, \omega_y)\text{d}\omega_x\text{d}\omega_y,$$

$$\ell_{yy} = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} \omega_y^2 s(\omega_x, \omega_y)\text{d}\omega_x\text{d}\omega_y. \quad (17)$$

Using Eqs. (3), (6), (15), and (16), one obtains the following theoretical model for the average speckle size as a function of intensity threshold level:

$$\overline{A} = \frac{\pi \exp[M(I_{\text{thres}}/\bar{I})]\Gamma[M, M(I_{\text{thres}}/\bar{I})]}{(\ell_{xx}\ell_{yy})^{1/2}[M(I_{\text{thres}}/\bar{I})]^{M-1}[2M(I_{\text{thres}}/\bar{I}) - 2M + 1]}. \quad (18)$$

The average speckle size is thus a funtion of four parameters: $I_{\text{thres}}/\bar{I}$; the shaping constant of the gamma PDF, M; the second moments of the normalized power spectral density of the field, ℓ_{xx} and ℓ_{yy}.

Experimental Measurements

Exhaustive experimental measurements of speckle data were recorded from a variety of target shapes, sizes, and textures with different illumination configurations (converging, diverging, or collimated), wavelengths, and states of polarization.

Laboratory Speckle Data Acquisition

Figure 5 shows a schematic diagram of the optical setup used for the collection of speckle data with visible illumination on the target. For these experiments a helium–neon tunable laser was used as the light source. Red, orange, yellow, and green wavelengths were available with this laser, although most of the experiments were conducted with the red (632.8-nm) wavelength. After the polarized beam exits the laser, it is directed into a 40× microscope objective for beam expansion and a 5-μm pinhole for

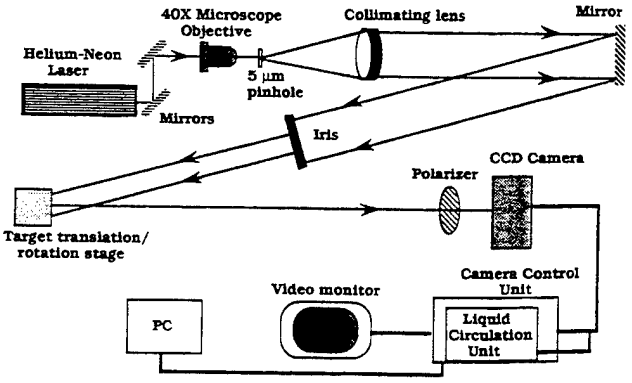


Fig. 5. Experimental configuration for laser speckle data collection.

optical spatial filtering. The collimating lens is mounted on a translation stage, making it possible to work with a collimated beam or to use converging or diverging illumination. The beam is reflected from a mirror and directed onto the target. The size of the beam striking the target is controlled by an iris diaphragm placed between the mirror and the target. A linear polarizer is placed in front of the camera to ensure that any depolarization effects caused by scattering from the target are eliminated from the speckle data.

A scientific-grade charge-coupled-device (CCD) camera was used because of its low noise, high linearity, and high dynamic range. The CCD imager contains 384 × 576 pixels that are 23 μm × 23 μm in size. When a newly acquired image is digitized, the analog-to-digital converter assigns an intensity value (gray level) proportional to the number of electrons stored in each pixel in the range of 0 to 4096 (12-bit precision). The controller operates the camera head and contains the power supplies, slow scan electronics, and image random-access memory, which is used to digitize and temporarily store an image. The liquid circulation unit thermoelectrically cools the camera head, which reduces the dark current. The video monitor displays each image using 256 gray levels. The CCD camera system is operated through an IEEE-488 interface with a dedicated personal computer, and appropriate image-processing software is used to manipulate the data. Figure 2(a) illustrates a typical digitized speckle image, which constitutes the raw data.

Extreme care was taken to isolate the camera and the laser path from the possibility of stray light noise contributions to the speckle image. Obviously any motion of the speckle pattern caused by mechanical vibrations or air turbulence during an observation period would degrade the integrity of the measurements. Accordingly the entire optical assembly is mounted on a Newport vibration-isolated optical table, and the speckle path is protected by an antiturbulence tunnel.

Another important consideration in recording the speckle data is the speckle size on the CCD array. The speckles must be large compared with the pixel size so that the speckle intensity variations can be resolved. Approximately 7 to 10 pixels spanning the diameter of a speckle (50–100 pixels/speckle) provides sufficient sampling and still permits well over 1500 speckles to be recorded in a single frame with the CCD camera. A large number of speckles in each image sample is necessary for meaningful statistical evaluation. We accomplished these sampling conditions by choosing a distance to the target (L) of ~1.5 m and by choosing the diameter (D) of the illuminated region on the target to be ~5.5 mm.

Measurement of Statistical Parameters for Theoretical Model

The parameters contained in the theoretical model expressed by Eq. (18) were determined from experimentally measured intensity speckle data. The mean ($\bar{I}$) and the variance (σ_I^2) of the intensity variations are readily calculated from the raw speckle data. We can find the value of the shaping constant M by taking the expected value of I^2 with the PDF defined by Eq. (5) (Ref. 3):

$$M = (\bar{I})^2/\sigma_I^2, \qquad \sigma_I^2 = \overline{(I^2)} - (\bar{I})^2. \tag{19}$$

The parameters ℓ_{xx} and ℓ_{yy} used in Eq. (18) are the second moments of the normalized PSD of the random field producing the speckle. Because detectors of optical radiation respond not to field strength but rather to optical power or intensity, the measured speckle data is in terms of intensity. Siegert's relation makes it possible to determine the autocorrelation of the electric field from that of the intensity[16]:

$$\Re_E(x, y) = [\Re_I(x, y) - (\bar{I})^2]^{1/2}. \tag{20}$$

Siegert's relation is a direct consequence of the many statistically independent contributions to the field at a point in space due to scattering from a rough surface. Equation (20) relates the autocorrelation of the electric field to the autocovariance function, $c(x, y) = \Re_I(x, y) - (\bar{I})^2$, of the intensity speckle pattern.[17] Hence we can obtain the autocorrelation function of the field by multiplying the square root of the normalized autocovariance function by the mean intensity and dividing by the square root of the PDF shaping constant:

$$\Re_E(x, y) = \bar{I}[c_I(x, y)/\sigma_I^2]^{1/2}/(M)^{1/2}. \tag{21}$$

Because the variance of the field is equal to the mean intensity,

$$\sigma_E^2 = \overline{(E^2)} - (\bar{E})^2 = \bar{I}, \quad \bar{E} = 0, \tag{22}$$

we can write

$$\Re_E(x, y)/\sigma_E^2 = [c_I(x, y)/\sigma_I^2]^{1/2}/(M)^{1/2}. \tag{23}$$

The Wiener–Khintchine theorem[3] (also known as the autocorrelation theorem[18]) states that the PSD of a signal and its autocorrelation function $\Re_E(x, y)$ constitute a Fourier-transform pair:

$$\Re_E(x, y) \Leftrightarrow S(\omega_x, \omega_y). \tag{24}$$

Hence

$$\Re_E(x, y)/\sigma_E^2 \Leftrightarrow s(\omega_x, \omega_y). \tag{25}$$

Figure 6 schematically illustrates the relationship between the measurable speckle intensity data and the normalized PSD of the random electromagnetic field producing the speckle.

We obtain the parameters of interest, ℓ_{xx} and ℓ_{yy}, by calculating the second moments of the normalized PSD of the field, $s(\omega_x, \omega_y)$. However, the second-moment theorem of Fourier-transform theory[19] states

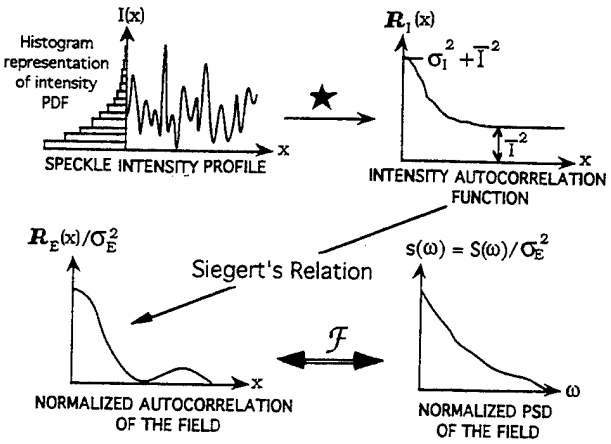


Fig. 6. Illustration of the relationship between the measurable speckle intensity data and the normalized PSD function of the field.

that if $f(x, y)$ and $F(\omega_x, \omega_y)$ are a Fourier-transform pair, then

$$-\frac{1}{4\pi^2} f_{xx}(0, 0) = \int_{-\infty}^{\infty} \omega_x^2 F(\omega_x, \omega_y) d\omega_x d\omega_y. \quad (26)$$

Hence the second moment of a function $F(\omega_x, \omega_y)$ can be obtained from the curvature (second derivative) of its Fourier transform evaluated at zero. This permits us to calculate ℓ_{xx} and ℓ_{yy} by taking the second derivative of $\Re_E(x, y)$ evaluated at zero.

We obtained this peak curvature by using a parabolic fit to the peak data point and its nearest neighbor, as indicated in the inset of Fig. 7. The accuracy of this technique clearly depends on the sampling density.

The percent error plotted against density illustrated in Fig. 7 was prepared by comparison of the result of the above technique to the exact analytic solution for a Gaussian autocorrelation function. A sampling density of 12 samples/speckle results in an error of ~1%, 8 samples/speckle yields a value for the curvature that is accurate to ~2%, and a 10% error results if only 3 samples/speckle are used.

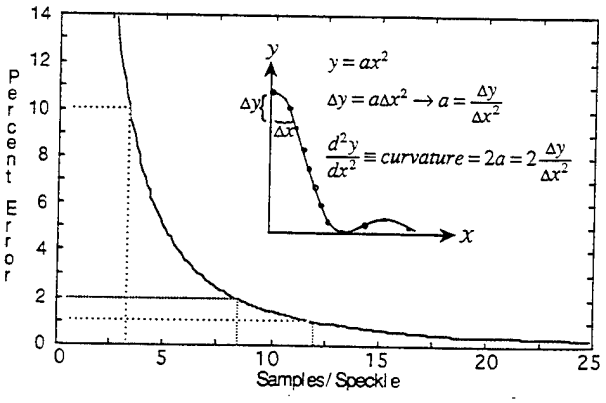


Fig. 7. Calibration curve for curvature calculation; better than 2% accuracy is indicated for sampling densities exceeding 8 samples/speckle.

The above graph can be used not only as an indication of the accuracy of the calculation of the desired second moments, ℓ_{xx} and ℓ_{yy}, but also as a calibration curve from which a correction factor can be applied to the calculated values.

Direct Measurement of Average Speckle Size

An image-processing computer program was developed to calculate the total area and number of speckles at each threshold level from the experimentally measured and digitized speckle data recorded by the CCD camera. The digitized data are first scanned to find the highest and the lowest intensity values recorded. The threshold level is then sequentially increased from the lowest to the highest recorded intensity values. At each threshold level all speckles are identified and counted, and the total speckle area is calculated. The average speckle size and the associated threshold level are then stored. The threshold level is then increased, and the procedure is repeated.

The speckle identification process involves determination of the connectivity between pixels to establish boundaries. The concept of 4-connectivity was used to minimize ambiguities and to identify U-shaped and doughnut-shaped speckles[20] accurately.

As a test case to check this algorithm for accuracy, a 512 × 512 pixel field of four deterministic, pyramid-shaped speckles whose area decreased quadratically with increasing threshold level was numerically formed with intensity values in the range of zero to 255. The average speckle size as a function of threshold level was then calculated and compared with the analytically determined values. Excellent agreement was obtained.

Comparison of Experimental Measurements with Theory

In order to determine the degree and the range of validity of the theoretical model developed in Ref. 12 and summarized by Eq. (18), we present a direct comparison of theoretical predictions and experimental measurements for speckle patterns produced under three distinctly different conditions.

The first speckle pattern investigated was produced by illumination of a circular region on a sandblasted aluminum target with coherent radiation from a helium–neon laser. The size of the illuminated area (D) and the distance (L) to the observation plane was adjusted to produce a 7–10 samples/speckle diameter according to Eq. (1). Figure 8 illustrates a profile of the normalized autocovariance function of the speckle pattern in both the horizontal and the vertical directions. They are virtually identical, indicating an isotropic speckle pattern, as expected from a rotationally symmetric target. The peak curvature of this normalized autocovariance function results in calculated second moments of $\ell_{xx} = \ell_{yy} = 71.65$. If we take the full width of the autocovariance function to be twice the average speckle diameter, our measured

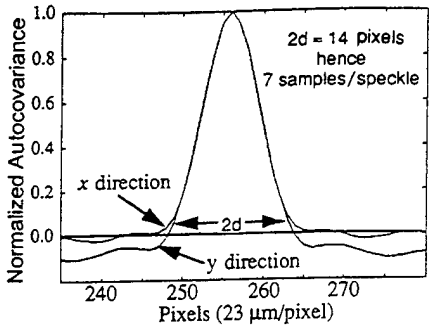


Fig. 8. Horizontal and vertical profiles of the normalized autocovariance function of the intensity speckle pattern produced by illumination of a circular region on a sandblasted region of an aluminum target.

speckle size is consistent with Eq. (1), and the sampling density of 7 samples/speckle diameter results in a correction factor of 2.8% from Fig. 7. This yields a corrected value of $(\ell_{xx}\ell_{yy})^{1/2} = 74.45$, used in Eq. (18).

The normalized PDF was also calculated from the digitized intensity data and is illustrated in Fig. 9 along with a truncated section of the recorded speckle pattern. The departure of this measured PDF from the negative exponential curve ($M = 1$) was caused by a variety of measurement error sources because the conditions of the target were such that a negative exponential PDF should have resulted.

Specific error sources that may have contributed to the departure of the measured PDF from the ideal negative exponential curve include the following: the finite pixel size of the CCD camera, the camera detector dark current and amplifier noise, background radiation, polarization effects, spatial averaging of the speckle pattern caused by vibration during the exposure time, turbulence effects in the beam path during the exposure time, and finite linewidth of the laser beam. If these effects are assumed to constitute additive noise

$$I_{\text{measured}}(x, y) = I_{\text{speckle}}(x, y) + I_{\text{noise}}(x, y), \qquad (27)$$

we can use the fact that the PDF of the sum of two statistically independent random variables is equal to the convolution of the two individual PDF's[21]:

$$P_{I_{\text{measured}}}(I) = P_{I_{\text{speckle}}}(I) * P_{I_{\text{noise}}}(I). \qquad (28)$$

Because there are many independent contributing error sources, the PDF of the noise can be thought of as the convolution of the PDF's of the individual error sources. The central-limit theorem thus dictates that the composite noise PDF will approximate a Gaussian function.[19,21] We determined the specific Gaussian parameters characterizing the composite noise PDF by empirically fitting the convolution of a Gaussian function with the ideal negative exponential PDF to the PDF of the experimentally measured speckle intensity data, as illustrated in Fig. 9:

$$P_{I_{\text{noise}}}(I) = (1/b)\exp[-\pi(x - x_0)/b]^2,$$
$$b = 0.14\bar{I}, \quad x_0 = b. \qquad (29)$$

This set of measured intensity data yielded a value of 1.36 for $M = (\bar{I})^2/\sigma_I^2$, as calculated from Eq. (19). However, this quantity should be thought of as only the ratio of the mean-square intensity to the variance of the intensity data, not as a PDF shaping constant (i.e., because the PDF is no longer a gamma distribution, a given value for M does not determine a unique PDF shape).

Figure 10 shows a comparison of the average speckle area as a function of intensity threshold level, as theoretically predicted from Eq. (18), compared with the experimental measurements for this isotropic speckle pattern. There is excellent agreement for intensity threshold levels greater than roughly twice the mean intensity, with increasing departure between theory and experiment as the threshold level is decreased.

The error bars associated with the experimentally measured data points indicate the range of values obtained from five independent realizations of the speckle pattern produced from this target. We measured these independent realizations by illuminating different uncorrelated areas on the target.

We produced a second (nonisotropic) speckle pat-

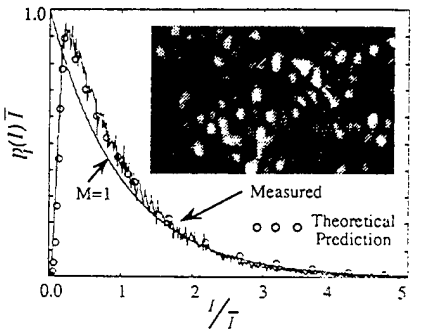


Fig. 9. Truncated section of the recorded speckle pattern illustrated along with the measured PDF compared with an ideal negative exponential and the predicted PDF in the presence of error sources characterized by Eq. (29).

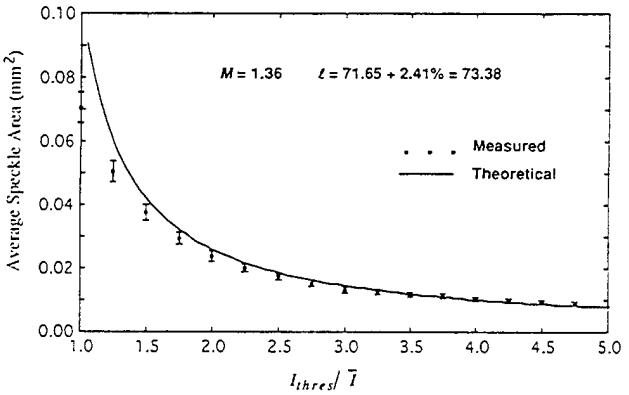


Fig. 10. Comparison of theoretical predictions of average speckle size as a function of threshold level with experimental measurements for an isotropic speckle pattern.

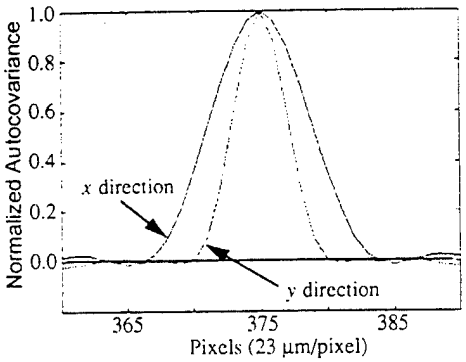


Fig. 11. Vertical and horizontal autocovariance function profiles for a speckle pattern produced by illumination of a rectangular target.

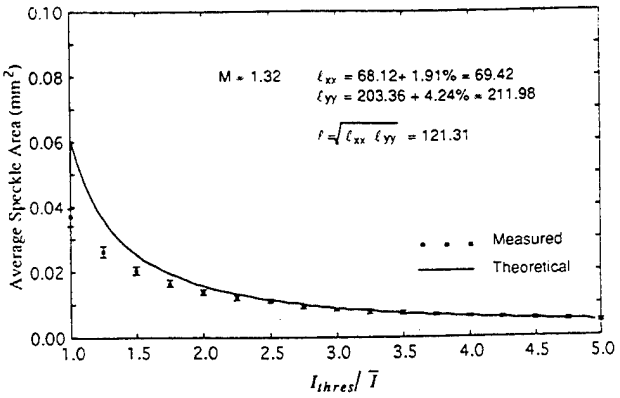


Fig. 13. Comparison of theoretical predictions of average speckle size as a function of threshold level with experimental measurements for a nonisotropic speckle pattern.

tern by illuminating a sandblasted aluminum rectangular target with a 2-to-1 aspect ratio. The ratio of the widths of the autocovariance-function profiles illustrated in Fig. 11 are approximately equal to the reciprocal of the aspect ratio of the rectangular target, as expected. For this target the second moments as determined by the peak curvature of the autocovariance function in the x and y directions are given by

$$\ell_{xx} = 68.12 + 1.91\% = 69.42,$$
$$\ell_{yy} = 203.36 + 4.24\% = 211.98. \quad (30)$$

The correction factor obtained from the calibration curve in Fig. 7 is substantially greater in the y direction, in which the speckles are elongated. This is obviously because of the lower sampling density.

The asymmetrical speckles that resulted from this rectangular target are evident in Fig. 12, along with the normalized PDF, which is again in excellent agreement with the convolution of a negative exponential speckle PDF with the Gaussian function described by Eq. (29) as the composite PDF of all measurement errors associated with the experiment.

This set of measured intensity data yielded a value of 1.32 for $M = (\bar{I})^2/\sigma_I^2$, as calculated from Eq. (19). Again, this quantity should not be thought of as a PDF shaping constant, but only as the ratio of the mean-square intensity to the variance of the intensity speckle data.

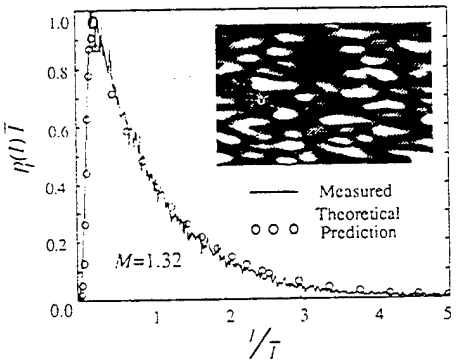


Fig. 12. Truncated section of the recorded (nonisotropic) speckle pattern illustrated along with the normalized PDF compared with the PDF predicted in the presence of error sources characterized by Eq. (19).

Figure 13 shows the comparison of the average speckle area as a function of the intensity threshold level as theoretically predicted from Eq. (18) compared with the experimental measurements for this nonisotropic speckle pattern. Once again, there is excellent agreement for intensity threshold levels greater than roughly twice the mean intensity, with increasing departure between the theory and experiment as the threshold level is decreased.

The previous speckle patterns were produced with a linearly polarized laser beam illuminating the target and a linear polarizer placed in front of the CCD camera to correct for any depolarization caused by the scattering process. Speckle patterns characterized by a PDF shaping constant near unity were thus produced. Speckle patterns exhibiting larger values of M were then produced by placing a $\lambda/4$ wave plate in the incident beam to create elliptically polarized light with varying degrees of ellipticity.

The third test case presented here was a speckle pattern produced by illumination of a circular region on a sandblasted aluminum target with circularly polarized laser light. The resulting speckle pattern is the incoherent sum of two orthogonally polarized speckle patterns, and, as for the case of unpolarized light, a shaping constant of $M \approx 2$ is thus expected.[3]

Figure 14 illustrates a profile of the normalized autocovariance function of the speckle pattern. The peak curvature of this normalized autocovariance function results in calculated second moments of $\ell_{xx} = \ell_{yy} = 53.7$. The width of the normalized autocovariance function is ~20 pixels, from which a correction factor of 1.55% is obtained from the calibration curve in Fig. 7. This yields a corrected total value of $(\ell_{xx}\ell_{yy})^{1/2} = 54.53$, used in Eq. (18) for this target.

The normalized PDF was calculated from the digitized intensity data and is presented in Fig. 15 along with the ideal gamma distribution with a PDF shaping constant of $M = 2$ and the predicted curve based on the previously determined measurement errors.

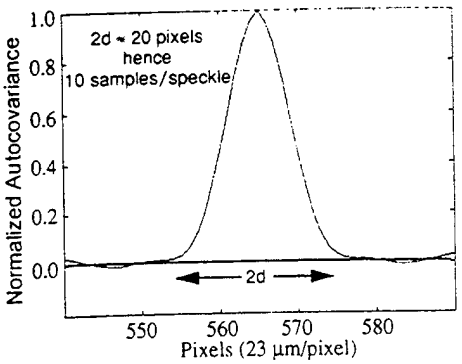


Fig. 14. Normalized autocovariance function from which the second moment is determined.

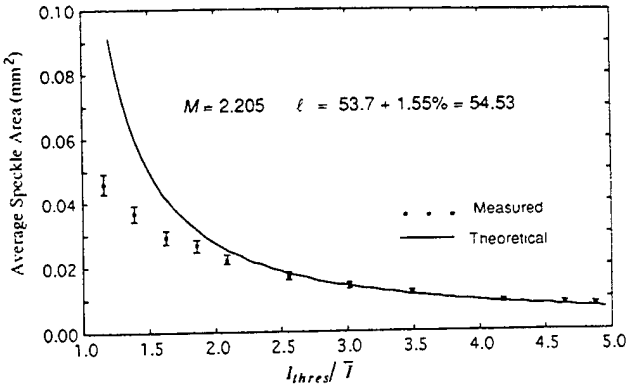


Fig. 16. Comparison of the theoretical prediction of average speckle size as a function of the threshold level with experimental measurements for this speckle pattern.

Because both orthogonally polarized components of the speckle pattern are corrupted independently by the noise PDF, this predicted PDF is the convolution of the ideal PDF with a Gaussian function $(2)^{1/2}$ times wider than that expressed by Eq. (29).

Figure 16 shows the average speckle area as a function of intensity threshold level as theoretically predicted from Eq. (18) compared with the experimental measurements for this speckle pattern produced with circularly polarized light. The mean-square-to-variance ratio of the experimentally measured intensity speckle data was $M = 2.2$. Again we have excellent agreement for intensity threshold levels greater than roughly twice the mean intensity; however, for this circularly polarized illumination the departure between the theoretical predictions and the experimental results increases much more rapidly as the threshold level is decreased.

Discussion of Results and Conclusions

In all three cases for which detailed comparisons were made the theoretical curve asymptotically approached the experimental curve at high threshold levels; however, significant departures occurred in the comparison at the low-intensity threshold levels. Excellent agreement between theory and experiment is observed for intensity threshold levels greater than ~2.0 times the mean intensity.

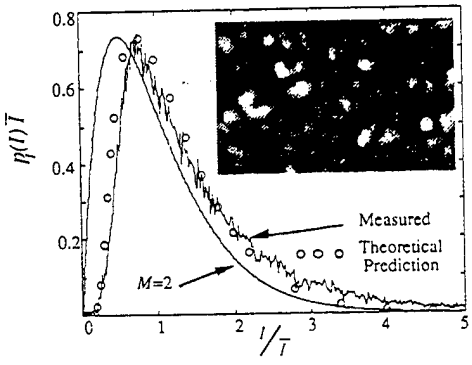


Fig. 15. Truncated section of the recorded (circularly polarized) speckle pattern illustrated along with the normalized PDF compared with a shaping constant of $M = 2$ and the prediction based on the known measurement noise PDF.

An examination of the initial approximations and assumptions that were made in the development of the theoretical model and an investigation of the nature of the speckle pattern at low threshold levels helps to explain the departure between the theoretical and experimental curves. The assumption of spatially ergodic speckle data is probably quite accurate; although we have seen that the assumption of a gamma-distributed intensity PDF is not strictly valid because of a variety of measurement error sources. It is not clear that the effect of this assumption varies significantly with threshold level. However, we will see that the assumption of simply connected speckles and the approximation of Eq. (2) as given in Eq. (3) are not valid at the lower threshold levels.

Figure 17 illustrates speckle data at four different threshold levels for the first speckle pattern investigated. Figure 17(a) clearly exhibits speckles that are not simply connected (i.e., some doughnut holes are observed). The simply connected assumption is therefor not valid at low threshold levels; however, these occurrences do not seem prevalent enough to explain the departure between the two curves. Of more significance is the nature of the highlighted speckle in Fig. 17(a). As the threshold level is re-

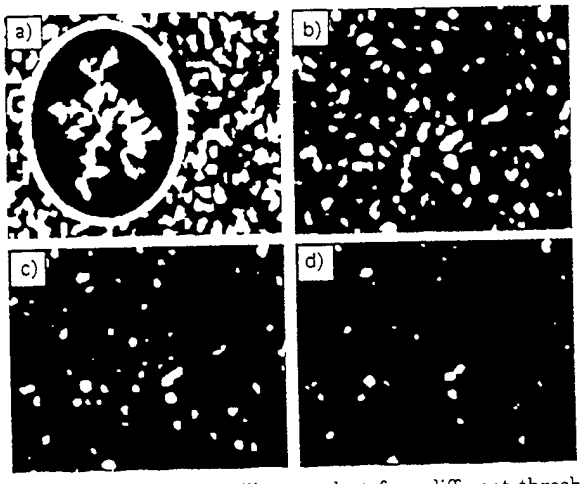


Fig. 17. Speckle pattern illustrated at four different threshold levels: (a) $I_{th} = \bar{I}$, (b) $I_{th} = 2\bar{I}$, (c) $I_{th} = 3\bar{I}$, (d) $I_{th} = 4\bar{I}$.

duced, the speckles begin to coagulate. A small decrease in the threshold may therefore result in a small increase in the total speckle area but a drastic change the number of speckles present. For example, at a slightly higher threshold the total speckle area may change only slightly, but the highlighted speckle may become five, six, or even more speckles instead of one. Thus the approximation of Eq. (2) given by Eq. (3) is clearly not valid at the lower intensity threshold levels. This effect is undoubtedly the main contributor for the departure between the theoretical curve and the experimental data at the low threshold levels.

As a verification that invalid assumptions in the theoretical model rather than experimental measurement errors are responsible for the departure in the data at low threshold levels, we produced an ideal computer-generated speckle pattern by permitting uncorrelated random phase errors (uniformly distributed from $-\pi$ to π) to propagate to the far field. This ideal speckle pattern was subjected to the same image-processing computer program developed to calculate the total area and number of speckles at each threshold level from the experimentally measured speckle patterns. Figure 18 illustrates the normalized autocovariance function and the normalized PDF for this ideal computer-generated speckle pattern. Note that the PDF is a nearly perfect negative exponential curve and does not exhibit the roll-over at the low-intensity levels that was observed for all of the experimentally measured speckle patterns.

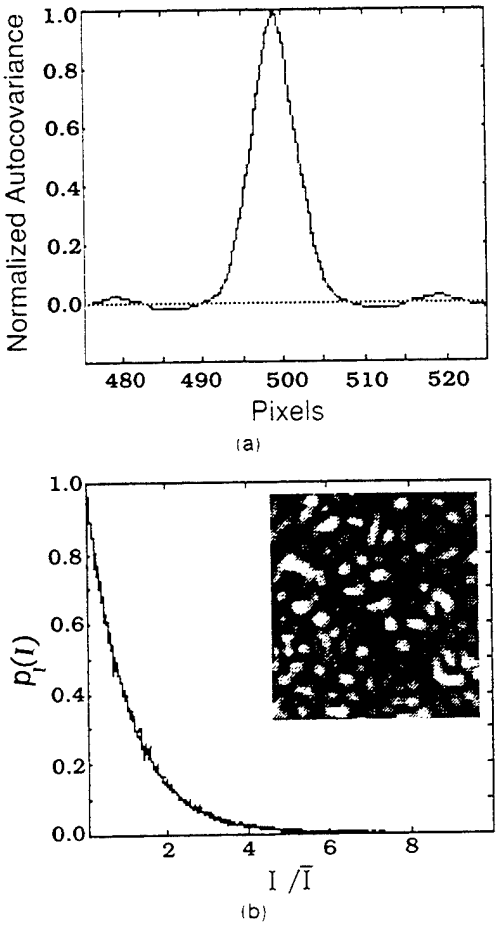


Fig. 18. (a) Speckle-pattern normalized autocovariance function, (b) normalized PDF for computer-generated speckle.

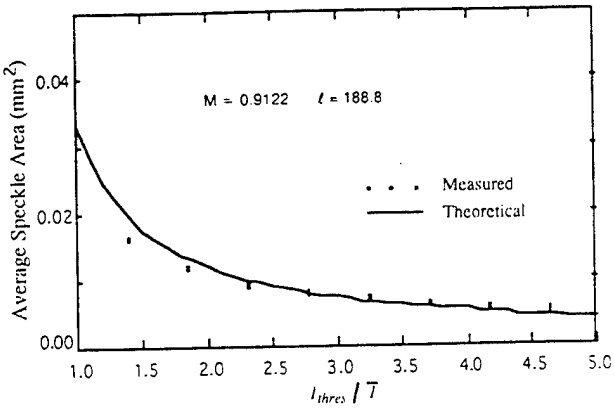


Fig. 19. Comparison of the theoretical prediction with actual data on computer-generated speckle.

In spite of the ideal statistics of the computer-generated speckle pattern, the measured average speckle size versus intensity threshold level still departs significantly from predictions of the theoretical model at the low-intensity threshold levels. Figure 19 illustrates this familiar behavior and verifies that the assumptions in the theoretical model are indeed invalid at low threshold levels. Generalizing the theoretical model to expand its range of validity will be the subject of later study.

This research has been supported by the Ballistic Missile Defense Organization, Innovative Science and Technology Directorate, through the Office of Naval Research under contract N00014-92-C-0123.

References

1. N. George, *Speckle* (The Institute of Optics, University of Rochester, Rochester, New York, 1979).
2. J. C. Dainty, *Laser Speckle and Related Phenomena*, 2nd ed. (Springer-Verlag, New York, 1984), pp. 1–7.
3. J. W. Goodman, *Statistical Optics* (Wiley, New York, 1985), Chap. 4, pp. 124–127; Chap. 7, 340–350.
4. M. Francon, *Laser Speckle and Applications in Optics* (Academic, New York, 1979), pp. i–xi.
5. D. Fink and S. N. Vodopia, "Coherent detection signal-to-noise ratio of an array of detectors," Appl. Opt. **15**, 453–454 (1976).
6. P. Gatt, W. P. Perez, D. A. Heimmermann, and C. M. Stickley, "Hererodyne laser radar array receiver for the mitigation of target-induced speckle," in *Laser Radar Applications*, G. W. Kamerman, ed., Proc. Soc. Photo-Opt. Instrum. Eng. **1936**, 157–164 (1993).
7. P. Gatt, W. P. Perez, D. A. Heimmermann, and C. M. Stickley, "Coherent laser radar array receivers: theory and experiment," presented at the Optical Society of America Seventh Annual Conference on Coherent Laser Radar, Paris, France, 19–23 July 1993.
8. J. Marron and G. M. Morris, "Correlation measurements using clipped laser speckle," Appl. Opt. **25**, 789–793 (1985).
9. J. Marron, "Correlation properties of clipped laser speckle," J. Opt. Soc. Am. A **2**, 1403–1410 (1986).

10. R. Barakat, "Clipped correlation functions of aperture integrated laser speckle," Appl. Opt. **25**, 3885–3888 (1986).
11. A. D. Ducharme, G. D. Boreman, and D. R. Snyder, "Effects of intensity thresholding on the power spectrum of laser speckle," Appl. Opt. **33**, 2715–2720 (1994).
12. F. E. Kragh, "Excursion areas of intensity due to random optical waves," M. S. thesis (University of Central Florida, Orlando, Fla., 1990).
13. E. Jakeman, "Enhanced backscattering through a deep random phase screen," J. Opt. Soc. Am. A **5**, 1638–1648 (1988).
14. I. S. Gradshteyn and I. M. Ryzhik, *Table of Integrals, Series, and Products*, (Academic, New York, 1980), p. 940.
15. D. Middleton, *Introduction to Statistical Communication Theory* (McGraw-Hill, New York, 1960), Chap. 9, pp. 426–427.
16. E. Jakeman, "Photon correlation," in *Photon Correlation and Light Beating Spectroscopy* H. Z. Cummins and E. R. Pike, eds. (Plenum, New York, 1974), Chap. 2, pp. 75–149.
17. A. Papoulis, *Probability, Random Variables, and Stochastic Processes* (McGraw-Hill, New York, 1965), Chap. 9, p. 282.
18. R. N. Bracewell, *The Fourier Transform and its Applications* (McGraw-Hill, New York, 1965), Chap. 6, p. 115.
19. J. D. Gaskill, *Linear Systems, Fourier Transforms, and Optics* (Wiley, New York, 1978), Chap. 7, p. 213.
20. R. C. Gonzalez and R. E. Woods, *Digital Image Processing* (Addison-Wesley, Reading, Mass., 1992), Chap. 2, p. 41.
21. G. Cooper and C. McGillem, *Probabilistic Methods of Signal and System Analysis* (Harcourt Brace Jovanovich, San Diego, Calif., 1986), Chap. 5, pp. 126–129.

3.7 Appendix 3B

T. L. Alexander, J. E. Harvey, and D. Hefele, "Experimental Verification of a Theoretical Model for Speckle Intensity Excursion Areas", Proc. SPIE **2222-63** (1994).

Experimental verification of theoretical model for speckle intensity excursion areas

Terri L. Alexander, James E. Harvey, and David Hefele

The Center for Research in Electro-optics and Lasers (CREOL)
The University of Central Florida
12424 Research Parkway
Orlando, Florida 32826
(407) 658-6818

ABSTRACT

Speckle is inherently an interference phenomenon produced when a rough object or turbulent medium introduces some degree of randomness to a reflected or transmitted electromagnetic field. Speckle characteristics are therefore a major concern in many laser imaging or wave propagation applications. For many applications, a detailed description of speckle size as a function of intensity threshold level is desirable. Extensive experimental measurements of average speckle size as a function of intensity threshold level were therefore made for several different targets and illumination conditions. We then compare these measurements with a theoretical model for excursion areas of speckle intensity. Excellent agreement is obtained for intensity threshold levels greater than approximately twice the mean intensity level.

1. INTRODUCTION

Speckle arises when the field associated with a monochromatic electromagnetic wave can be regarded as resulting from the sum of many independent and randomly phased contributions. Interference between the various contributions produces large amplitude and phase variations resulting in a speckled intensity distribution. Essentially a basic noise phenomenon, speckle limits the working resolution of coherent systems in acoustics, microwave and infrared technology, and visible optics.[1-4] The spatial and temporal characteristics of the speckle pattern will depend upon the size, shape, material, texture, and motion of the illuminated target as well as the wavelength, spectral bandwidth, and state of polarization of the illuminating radiation.

Coherent laser radar return signatures consist of a random speckle pattern at the receiver. Speckle effects, whether induced by target surface characteristics or atmospheric turbulence, routinely cause signal fading and a reduction in the signal-to-noise ratio in laser radar systems. Coherent laser radar *array* receivers (CLAR) are being explored in an effort to mitigate these effects.[5-7] Speckle characteristics affecting CLAR experiments include: mean speckle size, mean speckle strength or intensity, mean speckle separation, temporal effects, phase statistics of the received electric field, and polychromatic effects.

A common expression for the minimum speckle diameter (d) due to free-space propagation is determined by the illuminating laser wavelength (λ), the distance of the target from the receiver (L), and the size of the illuminated region on the target (D)[1]

$$d \geq \lambda L / D \tag{1}$$

However, for many applications, a more detailed description of speckle size as a function of intensity threshold level is desirable. The correlation properties of thresholded (or clipped) time-varying laser speckle has been studied in detail.[8-10] Figure 1(a) shows a typical speckle pattern

0-8194-1526-X/94/$6.00

produced when reflecting laser light from a rough surface. A three-dimensional isometric plot of the intensity speckle pattern illustrating the speckle size at a particular intensity threshold level is shown schematically in Figure 1(b).

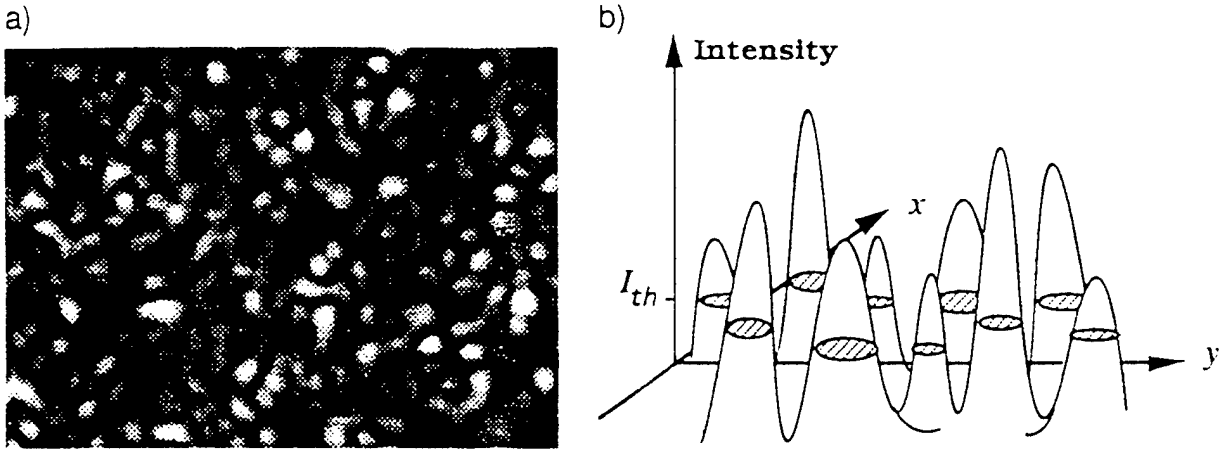


Figure 1. (a) Typical speckle pattern produced by illuminating a rough surface with laser light (b) a schematic illustration of speckle size at a specific intensity threshold level.

2. REVIEW OF THEORETICAL MODEL

A theoretical investigation entitled *Excursion Areas of Intensity due to Random Optical Waves* was conducted by Frank Kragh in a Master's thesis at the University of Central Florida in 1990[11]. He defines the expected size of a single excursion area as

$$\overline{A} = \frac{A_{tot}}{\overline{N}} \tag{2}$$

where A_{tot} is the total sampled area times the fraction of that area that is expected to have an intensity greater than the threshold intensity level, and $\overline{N}$ is the total number of speckles at that threshold level. Figure 2(b) is a schematic representation of the regions where the intensity has exceeded a given threshold level, and the shaded areas are the excursion areas of interest.

Kragh's derivation of the expected number of excursion areas or speckles ($\overline{N}$) at a particular intensity threshold level is a two-dimensional analog of a method developed by Middleton for counting excursions intervals of a one-dimensional function.[12] The resulting theoretical model for the expected area of a single excursion area for a Gamma distributed intensity is given by

$$\overline{A} = \frac{A_{tot}}{\overline{N}} \approx \frac{\pi \exp\left(M\frac{I_{th}}{\bar{I}}\right)\Gamma\left(M, M\frac{I_{th}}{\bar{I}}\right)}{\sqrt{\ell_{xx}\,\ell_{yy}}\left(M\frac{I_{th}}{\bar{I}}\right)^{M-1}\left(2M\frac{I_{th}}{\bar{I}} - 2M + 1\right)} \tag{3}$$

where

$$\Gamma\left(M, M\frac{I_{th}}{\bar{I}}\right) = \int_{M\frac{I_{th}}{\bar{I}}}^{+\infty} \exp(-t)\, t^{M-1} dt \qquad (4)$$

is an incomplete Gamma function[13], I_{th} is the intensity threshold level, $\bar{I}$ is the mean intensity, M is the intensity probability density function (PDF) shaping constant[3], and ℓ_{xx} and ℓ_{yy} are the second moments of the normalized power spectral density (PSD) function of the random field producing the speckle.

The derivation is general and applies to any two-dimensional gamma distributed process. The assumption of a gamma distributed process is appropriate for laser light which has reflected from a rough surface or propagated through the atmosphere since a gamma probability density function (PDF) results for the sum of two independent circular Gaussian random fields and may therefore be used to describe the intensity in coherent optical applications exhibiting speckle phenomena.[3]

In order to accurately count the excursion areas in a speckle pattern, Kragh's second assumption was that all excursion areas are simply connected. In other words, he assumed that there were no doughnut shaped speckles. This assumption is obviously valid for high threshold levels, but will not be valid for very low threshold levels.

A third assumption was that the expectation of the ratio of the total speckle area to the number of speckles can be approximated by the ratio of the expectation of the total speckle area to the expectation of the total number of speckles:

$$\bar{A} = \left\langle \frac{\textit{total speckle area}}{\textit{total number of speckles}} \right\rangle \approx \frac{\langle \textit{total speckle area} \rangle}{\langle \textit{total number of speckles} \rangle} = \frac{A_{tot}}{\bar{N}} \qquad (5)$$

The above relation is not strictly true since the numerator and denominator are not independent. However, in spite of their mutual dependence, this approximation is reasonable at high threshold levels since both the number of speckles and the total speckle area decrease at a similar rate as the threshold level increases. This dependency is therefore partially cancelled by the ratio.

Finally, the variations in intensity were assumed to be spatially ergodic. A random field is *ergodic* if all information about its joint probability distributions (and their statistics) can be obtained from a single realization of the random field.

3. EXPERIMENTAL MEASUREMENTS

Exhaustive experimental measurements of speckle data has been recorded from a variety of target shapes, sizes, and textures with different illumination configurations (converging, diverging, or collimated), wavelengths, and states of polarization.

3.1 Laboratory Speckle Data Acquisition

For these experiments a Helium-Neon Tunable Laser was used as the light source. Red, orange, yellow and green wavelengths were available with this laser, although most of the experiments were conducted with the red (632.8 nm) wavelength. A scientific-grade charge-coupled device (CCD) camera, was used because of its low noise, high linearity, and high dynamic range. The CCD imager contains 384 x 576 pixels that are 23 μm x 23 μm in size. Figure 2(a) illustrates a typical digitized speckle image which constitutes the raw data.

Extreme care was taken to isolate the camera and laser path from the possibility of stray light noise contributions to the speckle image. Obviously any motion of the speckle pattern, caused by mechanical vibrations or air turbulence, during an observation period would degrade the integrity of the measurements. Accordingly, the entire optical assembly is mounted on a Newport vibration-isolated optical table, and the speckle path is protected by an anti-turbulence tunnel.

Another important consideration in recording the speckle data is the speckle size on the CCD array. The speckles must be large compared to the pixel size so that the speckle intensity variations can be resolved. Approximately 7 to 10 pixels spanning a speckle provides sufficient sampling, and still allows well over 1500 speckles to be recorded in a single frame with the CCD camera. Having a large number of speckles in each image sample is necessary for meaningful statistical evaluation. These sampling conditions were accomplished by choosing a distance to target (L) of about 1.5 meters, and the diameter (D) of the illuminated region on the target to be approximately 5.5 millimeters.

3.2 Measurement of Statistical Parameters for Theoretical Model

The parameters contained in the theoretical model expressed by Eq. (3) were determined from experimentally measured intensity speckle data. The mean ($\bar{I}$) and the variance (σ_I^2) of the intensity variations are readily calculated from the raw speckle data.

The equation for the shaping constant M can be found from the following expression for the PDF of a gamma distribution[3]

$$p_I(I) = \left(\frac{M}{\bar{I}}\right)^M I^{M-1} \frac{\exp\left(-M\frac{I}{\bar{I}}\right)}{\Gamma(M)}, \quad \text{for } I \geq 0. \tag{6}$$

where

$$\Gamma(M) = \int_0^\infty \exp(-t)\, t^{M-1} dt \tag{7}$$

is a gamma function of order M. This function is illustrated graphically for several different values of M in Figure 3.

If Eq. (6) is multiplied by I^2, then integrated from zero to infinity, the result is the mean square intensity, $\overline{I^2}$ expressed as a function of M. This equation can then be inverted to yield

$$M = \frac{\bar{I}^2}{\sigma_I^2}, \quad \text{where } \sigma_I^2 = \overline{I^2} - \bar{I}^2 \tag{8}$$

Note that for M =1, the variance is equal to the mean intensity, and the PDF is a negative exponential function as illustrated in Figure 2. This behavior is characteristic of *fully developed speckle* which results when an electromagnetic wave is reflected from a surface whose roughness is large compared to the wavelength of the incident radiation (and the extent of the incident wave spans many correlation widths of the surface irregularities).[14]

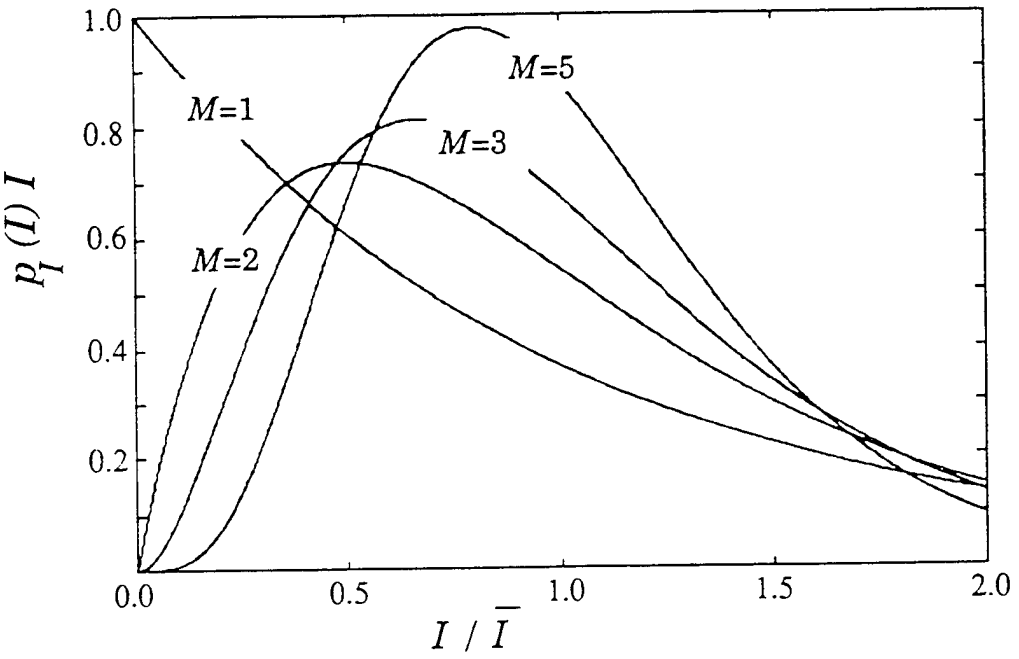


Figure 2. The gamma probability distribution function plotted for several values of M.

The parameters ℓ_{xx} and ℓ_{yy} used in Eq.(3) are the second moments of the *normalized* PSD of the random *field* producing the speckle and are defined as

$$\ell_{xx} = \int_{-\infty}^{\infty} \xi^2 S(\xi,\eta)\,d\xi\,d\eta, \qquad \ell_{yy} = \int_{-\infty}^{\infty} \eta^2 S(\xi,\eta)\,d\xi\,d\eta, \tag{9}$$

where

$$S(\xi,\eta) = S(\xi,\eta)/\sigma_E{}^2. \tag{10}$$

$S(\xi,\eta)$ is the PSD of the random field producing the speckle and $\sigma_E{}^2$ is the variance of that random field.

Since detectors of optical radiation respond not to field strength, but rather to optical power or intensity, the speckle data measured in the laboratory is measured in terms of intensity. Fortunately, the following equation, known as Siegert's relation, makes it possible to determine certain properties of the field from intensity information[15]

$$\Re_E(x,y) = \sqrt{\Re_I(x,y) - \bar{I}^2}. \tag{11}$$

$\Re_E(x,y)$ is the autocorrelation function of the random field and $\Re_I(x,y)$ is the autocorrelation function of the intensity speckle pattern. Siegert's relation is a direct consequence of the many statistically independent contributions to the field at a point in space due to scattering from a rough surface.

But $\Re_I(x,y)-\bar{I}^2$ is just the autocovariance function, C(x,y), of the intensity speckle pattern.[16] Hence, the autocorrelation function of the field is obtained by multiplying the square root of the normalized autocovariance function by the mean intensity and dividing by the square root of the PDF shaping constant

$$\Re_E(x,y)=\bar{I}\sqrt{C_I(x,y)/\sigma_I^2}/\sqrt{M}, \tag{12}$$

or, since the variance of the field is equal to the mean intensity,

$$\sigma_E^2=\overline{E^2}-\bar{E}^2=\bar{I},\ \ \bar{E}=0 \tag{13}$$

we can write

$$\Re_E(x,y)/\sigma_E^2=\sqrt{C_I(x,y)/\sigma_I^2}/\sqrt{M}. \tag{14}$$

The *Wiener-Khintchine* theorem[3] (also known as the autocorrelation theorem[17]) states that the PSD $S(\xi,\eta)$ of a signal and its autocorrelation function $\Re_E(x,y)$ constitute a Fourier transform pair

$$\Re_E(x,y)\Leftrightarrow S(\xi,\eta). \tag{15}$$

Figure 3 schematically illustrates the relationship between the measurable speckle intensity data and the normalized PSD of the random electromagnetic field producing the speckle.

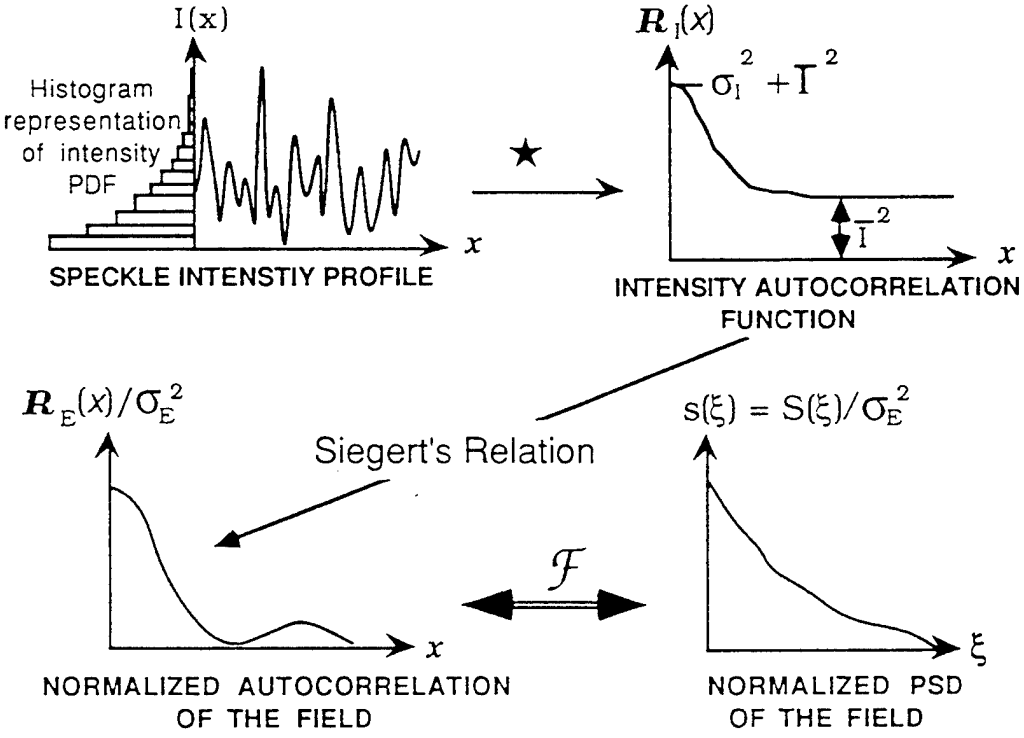


Figure 3. Illustration of relationship between the measurable speckle intensity data and the normalized PSD function of the field.

Finally, the parameters of interest, ℓ_{xx} and ℓ_{yy}, can be obtained by calculating the second moments of the normalized PSD of the field, $S(\xi,\eta)$. However, the second moment theorem of Fourier transform theory[18] states that, if $f(x,y)$ and $F(\xi,\eta)$ are a Fourier transform pair, then

$$-\frac{1}{4\pi^2} f''(0,0) = \int_{-\infty}^{\infty} \xi^2 F(\xi,\eta)\, d\xi\, d\eta. \tag{16}$$

Hence, the second moment of a function F(ξ,η), can be obtained from the curvature (second derivative) of its Fourier transform evaluated at zero. This allows us to calculate ℓ_{xx} and ℓ_{yy} by taking the second derivative of $\Re_E(x,y)$ evaluated at zero.

This peak curvature was obtained by using a parabolic fit to the peak data point and its nearest neighbor as indicated in the inset of Figure 5. The accuracy of this technique clearly depends upon the sampling density.

The percent error plotted against sampling density illustrated in Figure 4 was prepared by comparing the results of the above technique to the exact analytic solution for a Gaussian autocorrelation function. A sampling density of 12 samples/speckle results in an error of approximately 1%, 8 samples/speckle yields a value for the curvature that is accurate to about 2%, and a 10% error results if only 3 samples/speckle are used.

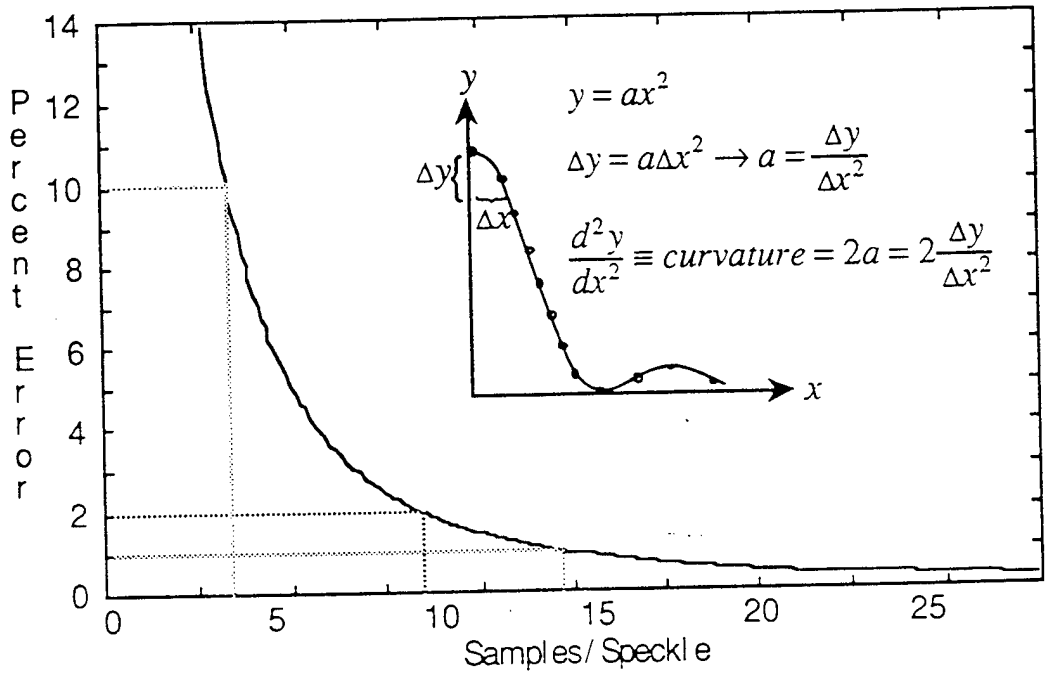


Figure 4. Calibration curve for curvature calculation; better than 2% accuracy is indicated for sampling densities exceeding 8 samples/speckle.

The above graph can be used not only as an indication of the accuracy of the calculation of the desired second moments, ℓ_{xx} and ℓ_{yy}; but as a calibration curve from which a correction factor can be applied to the calculated values.

3.3 Direct Measurement of Average Speckle Size

An image processing computer program was developed to calculate the total area and number of speckles at each threshold level from the experimentally measured and digitized speckle data recorded by the CCD camera. The digitized data is first scanned to find the highest and lowest intensity values recorded. The threshold level is then sequentially increased from the lowest to the highest recorded intensity values. At each threshold level all speckles are identified, counted, and the total speckle area calculated. The average speckle size and the associated threshold level is then stored. The threshold level is then increased and the procedure repeated.

The speckle identification process involves determining the connectivity between pixels to establish boundaries. The concept of *4-connectivity* was used to minimize ambiguities and accurately identify U-shaped and doughnut-shaped speckles.[19]

As a test case to check this algorithm for accuracy, a 512 X 512 pixel field of four deterministic, pyramid-shaped speckles whose area decreases quadratically with increasing threshold level was numerically formed with intensity values ranging from zero to 256. The average speckle size as a function of threshold level was then calculated and compared to the analytically determined values. Excellent agreement was obtained.

4. COMPARISON OF EXPERIMENTAL MEASUREMENTS WITH THEORY

In order to determine the degree and range of validity of the theoretical model developed in Ref. 11 and summarized by Eq. (3), a direct comparison of theoretical predictions and experimental measurements will be presented for speckle patterns produced under three distinctly different conditions.

The first speckle pattern investigated was produced by illuminating a circular region on a sandblasted aluminum target with coherent radiation from a He-Ne laser. The size of the illuminated area (D) and the distance (L) to the observation plane was adjusted to produce 6-10 samples/speckle according to Eq. (1). The intensity speckle pattern was recorded with the CCD camera.

The profiles of the normalized autocovariance function of the speckle pattern in both the horizontal and vertical directions are virtually identical, indicating an isotropic speckle pattern as expected from a rotationally symmetric target. The peak curvature of this normalized autocovariance function results in calculated 2nd moments of $\ell_{xx} = \ell_{yy} = 71.65$. If we take the full-width of the autocovariance function to be twice the average speckle diameter, our measured speckle size is consistent with Eq.(1), and the sampling density of 7 samples/speckle results in a correction factor of 2.8% from Figure 5. This yields a total value of $\ell = \sqrt{\ell_{xx}\,\ell_{yy}} = 74.45$. to be input into Eq.(3).

The normalized PDF was also calculated from the digitized intensity data and is illustrated in Figure 6 along with a truncated section of the recorded speckle pattern. The departure of this measured PDF from the negative exponential curve (M=1) is undoubtedly due to a variety of measurement error sources since the conditions of the target were such that a negative exponential PDF should have resulted.

Specific error sources which may have contributed to the departure of the measured PDF from the ideal negative exponential curve include: the finite pixel size of the CCD camera, the camera detector dark current and amplifier noise, background radiation, polarization effects, spatial averaging of the speckle pattern due to vibration during the exposure time, turbulence effects in the beam path during

the exposure time, finite line-width of the laser beam, etc. If these effects are assumed to constitute additive noise

$$I_{measured}(x,y) = I_{speckle}(x,y) + I_{noise}(x,y) \tag{17}$$

we can use the fact that the PDF of the sum of two statistically independent random variables, is equal to the convolution of the two individual PDFs.[20]

$$P_{I_{measured}}(I) = P_{I_{speckle}}(I) * P_{I_{noise}}(I) \tag{18}$$

Since there are many independent contributing error sources, the PDF of the noise can be thought of as the convolution of the PDF's of the individual error sources. The central limit theorem thus dictates that the composite noise PDF will approximate a Gaussian function.[18] The specific Gaussian parameters characterizing the composite noise PDF were determined by empirically fitting the convolution of a Gaussian function with the ideal negative exponential PDF to the PDF of the experimentally measured speckle intensity data as illustrated in Figure 5.

$$P_{I_{noise}}(I) = (1/b)\exp[-\pi(x - x_o)/b]^2, \quad b = 0.14\bar{I}, \quad x_o = b \tag{19}$$

This set of measured intensity data yielded a value of 1.36 for $M = \bar{I}^2/\sigma_I^2$ as calculated from Eq. (8). However, this quantity should be thought of only as the ratio of the mean square intensity to the variance of the intensity data and not as a PDF shaping constant (i.e., since the PDF is no longer a Gamma distribution, a given value for M does not determine a unique PDF shape).

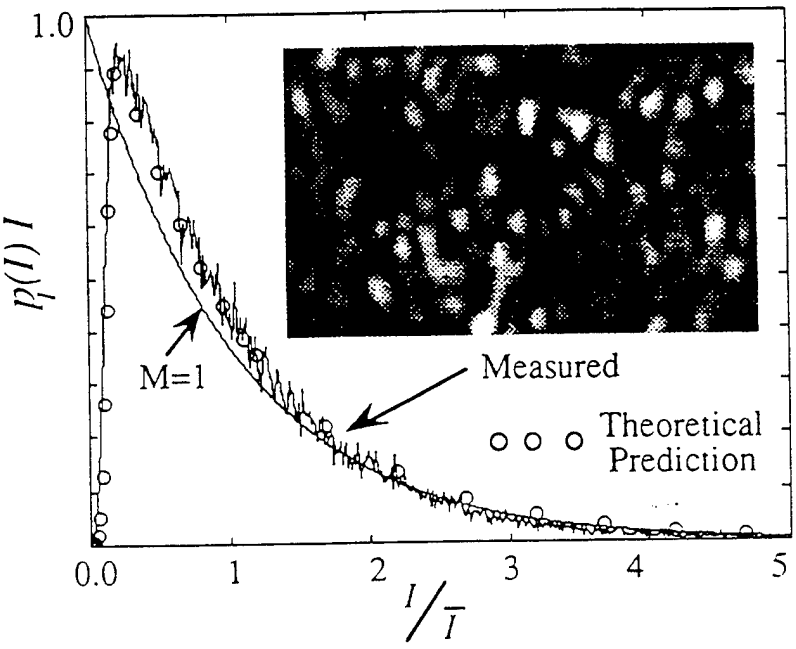


Figure 5. A truncated section of the recorded speckle pattern is illustrated along with the measured PDF compared to an ideal negative exponential and the predicted PDF in the presence of error sources characterized by Eq(19).

Figure 6 shows a comparison of the average speckle area as a function of intensity threshold level as theoretically predicted from Eq. (3) compared to the experimental measurements for this isotropic

speckle pattern. Note that there is excellent agreement for intensity threshold levels greater than about twice the mean intensity, with increasing departure between the theory and experiment as the threshold level is decreased.

The error bars associated with the experimentally measured data points indicate the range of values obtained from five (5) independent realizations of the speckle pattern produced from this target. These independent realizations were measured by illuminating different uncorrelated areas on the target.

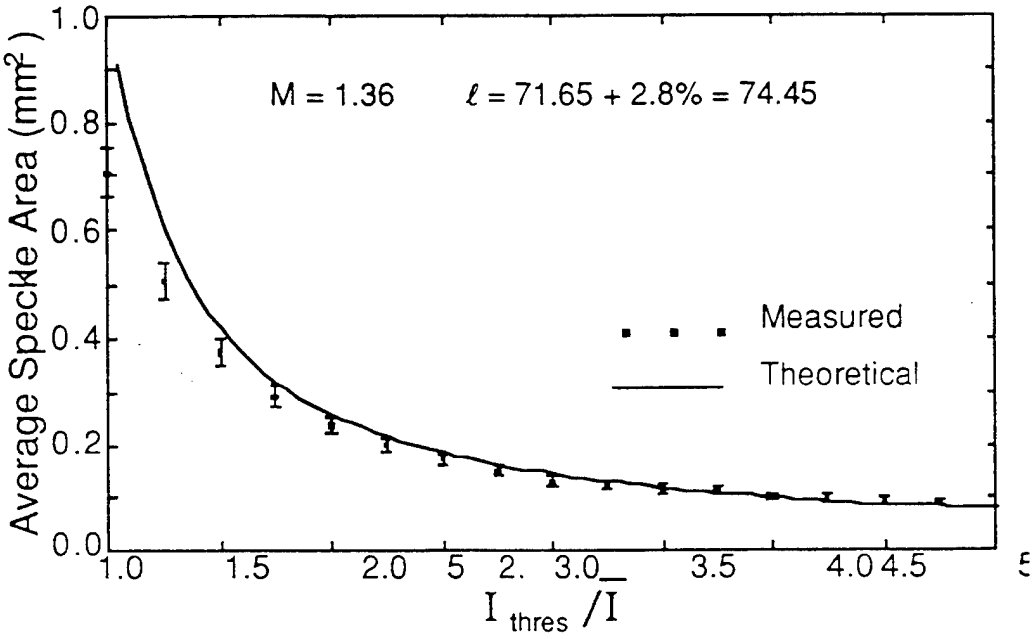


Figure 6. Comparison of theoretical predictions of average speckle size as a function of threshold level with experimental measurements for an isotropic speckle pattern.

A 2nd (non-isotropic) speckle pattern was produced by illuminating a sandblasted aluminum rectangular target with a 2 to 1 aspect ratio. The ratio of the widths of the autocovariance function profiles are approximately equal to the reciprocal of the aspect ratio of the rectangular target as expected. For this target the second moments as determined by the peak curvature of the autocovariance function in the x-direction and the y-direction are given by

$$\ell_{xx} = 68.12 + 1.91\% = 69.42$$

and

$$\ell_{yy} = 203.36 + 4.24\% = 211.98 \tag{19}$$

or

$$\ell = \sqrt{\ell_{xx}\,\ell_{yy}} = 121.31$$

Note that the correction factor obtained from the calibration curve in Figure 5 is substantially greater in the y-direction in which the speckles are elongated. This is obviously due to the lower sampling density.

This set of measured intensity data yielded a value of 1.32 for $M = \bar{I}^2/\sigma_I^2$ as calculated from Eq. (8). Again, this quantity should not be thought of as a PDF shaping constant, but only as the ratio of the mean square intensity to the variance of the intensity speckle data.

Figure 7 shows the comparison of the average speckle area as a function of intensity threshold level as theoretically predicted from Eq. (3) compared to the experimental measurements for this non-isotropic speckle pattern.

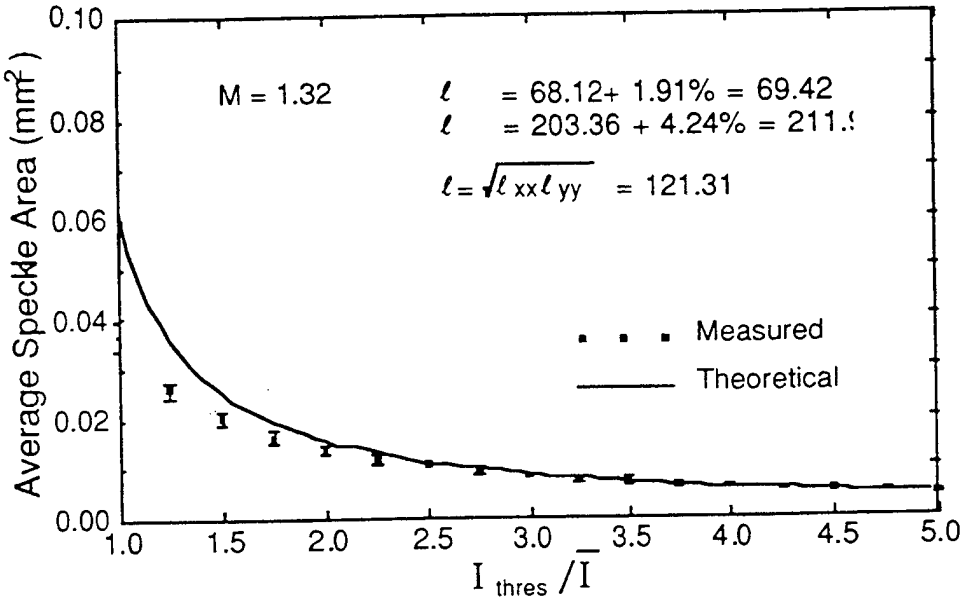


Figure 7. Comparison of theoretical predictions of average speckle size as a function of threshold level with experimental measurements for a non-isotropic speckle pattern.

Once again, there is excellent agreement for intensity threshold levels greater than about twice the mean intensity, with increasing departure between the theory and experiment as the threshold level is decreased.

The previous speckle patterns were produced with a linearly polarized laser beam illuminating the target, and a linear polarizer placed in front of the CCD camera to correct for any depolarization caused by the scattering process. Speckle patterns characterized by a PDF shaping constant near unity were thus produced. Speckle patterns exhibiting larger values of M were then produced by placing a $\lambda/4$ wave plate in the incident beam to create elliptically polarized light with varying degrees of ellipticity.

The 3rd test case presented here was a speckle pattern produced by illuminating a circular region on a sandblasted aluminum target with circularly polarized laser light. The resulting speckle pattern is the incoherent sum of two orthogonally polarized speckle patterns and, as for the case of unpolarized light, a shaping constant of $M = 2$ is thus expected.[3]

The peak curvature of this normalized autocovariance function results in calculated 2nd moments of $\ell_{xx} = \ell_{yy} = 53.7$. The width of the normalized autocovariance function is approximately 20 pixels, from which a correction factor of 1.55% is obtained from the calibration curve in Figure 4. This yields a total value of $\ell = \sqrt{\ell_{xx}\,\ell_{yy}} = 54.53$ to be input into Eq.(3) for this target.

Figure 8 shows the comparison of the average speckle area as a function of intensity threshold level as theoretically predicted from Eq. (3) compared to the experimental measurements for this speckle pattern produced with circularly polarized light. The mean square to variance ratio of the experimentally measured intensity speckle data was M=2.2. Again we have excellent agreement for intensity threshold levels greater than about twice the mean intensity; however, for this circularly polarized illumination, the departure between the theoretical predictions and experimental results increases much more rapidly as the threshold level is decreased.

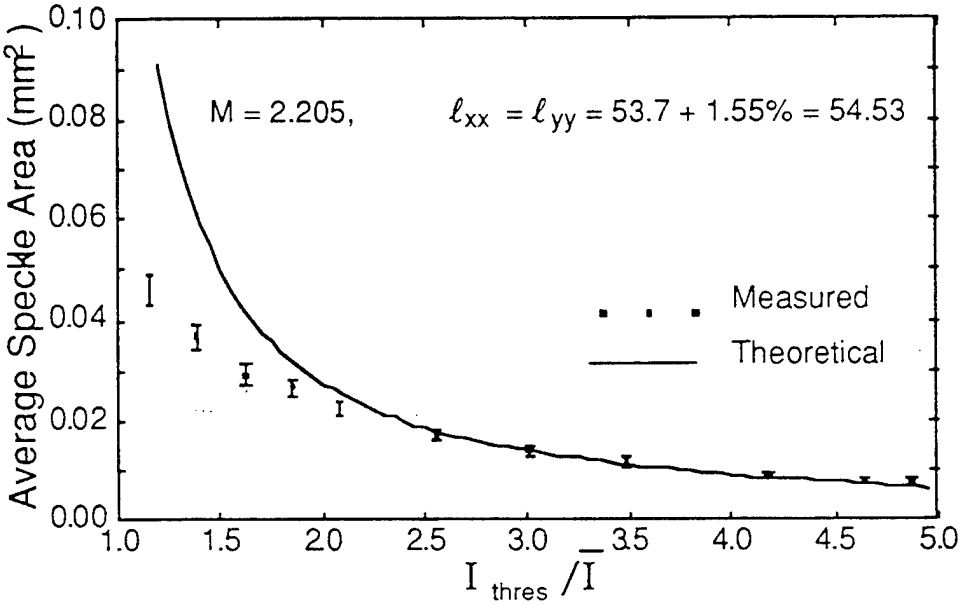


Figure 8 Comparison of theoretical predictions of average speckle size as a function of threshold level with experimental measurements for this speckle pattern.

5. DISCUSSION OF RESULTS AND CONCLUSIONS

In all three cases for which detailed comparisons were made, the theoretical curve asymptotically approached the experimental curve at high threshold levels; however, significant departures occur in the comparison at the low intensity threshold levels. Excellent agreement between theory and experiment is observed for intensity threshold levels greater than approximately 2.0 times the mean intensity.

An examination of the initial approximations and assumptions which were made in the development of the theoretical model, and an investigation of the nature of the speckle pattern at low threshold levels helps to explain the departure between the theoretical and experimental curves. The assumption of spatially ergodic speckle data is probably quite accurate; and although we have already seen that the assumption of a Gamma distributed intensity PDF is not strictly valid due to a variety of measurement error sources, it is not clear that the effect of this assumption varies significantly with threshold level. However, we will see that assumptions #2 and #3 discussed earlier are not valid at the lower threshold levels and that the effects of these assumptions does become increasingly severe as the threshold level is reduced.

Figure 9 illustrates speckle data at four different threshold levels for the 1st speckle pattern investigated. Figure 9(a) clearly exhibits speckles that are not simply-connected (i.e., some doughnut holes are observed). Assumption #2 is therefor not valid at low threshold levels; however, these occurrences do not seem prevalent enough to explain the departure between the two curves. Of more significance, is the nature of the highlighted speckle in Figure 9(a). As the threshold level is reduced

the speckles begin to coagulate. A small decrease in threshold may therefore result in a small increase in total speckle area but a drastic change the number of speckles present. For example, at a slightly higher threshold, the total speckle area may change only slightly, but the highlighted speckle may become 5, 6, or even more speckles instead of one. Thus, assumption #3 is clearly not valid at the lower intensity threshold levels. This effect is undoubtedly the main contributor for the departure between the theoretical curve and the experimental data at the low threshold levels.

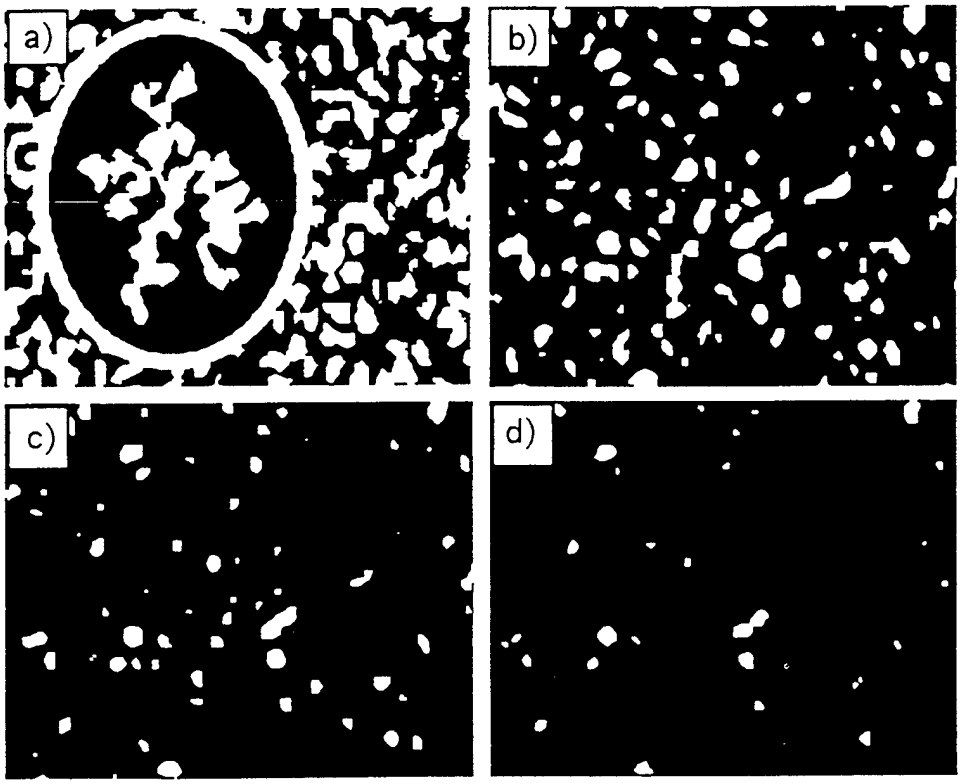


Figure 10. Speckle pattern illustrated at four different threshold levels: (a) $I_{th} = \bar{I}$, (b) $I_{th} = 2\bar{I}$, (c) $I_{th} = 3\bar{I}$, and (d) $I_{th} = 4\bar{I}$

As a final check on the theoretical predictions, a computer generated speckle pattern with a nearly perfect negative exponential PDF was treated as actual data and a comparison of the directly measured average speckle size vs. threshold level with the theoretical prediction was made. The results of this comparison showed the same behavior as the previous three laboratory generated speckle samples.

6. ACKNOWLEDGMENTS

This research has been supported by the Ballistic Missile Defense Organization, Innovative Science and Technology Directorate through the Office of Naval Research under contract N00014-92-C-0123.

7. REFERENCES

1. Nicholas George, "Speckle", The Institute of Optics, University of Rochester, N.Y. (1979).

2. J. C. Dainty, *Laser Speckle and Related Phenomena*, Springer-Verlag, N.Y. (1984).

3. J. W. Goodman, *Statistical Optics* (John Wiley & Sons, New York, 1985).

4. M. Francon, *Laser Speckle and Applications in Optics*, Academic Press, N.Y. (1979).

5. D. Fink and S. N. Vodopia, "Coherent detection SNR of an array of detectors", Appl. Opt. , Vol.15, No. 2, 453-454 (Feb. 1976).

6. P. Gatt, W. P. Perez, D. A. Heimmermann, and C. M. Stickley, "Hererodyne Laser Radar Array Receiver for the Mitigation of Target-induced Speckle", Proc. SPIE 1936-17 (1993).

7. P. Gatt, W. P. Perez, D. A. Heimmermann, and C. M. Stickley, "Coherent Laser Radar Array Receivers: Theory and Experiment", Presented at the OSA 7th Annual Conference on Coherent Laser Radar, Paris, France (July 1993).

8. J. Marron and G. M. Morris, "Correlation measurements using clipped laser speckle", J. Opt. Soc. Am. A, Vol. 2, 1403 (1985).

9. J. Marron, "Correlation properties of clipped laser speckle", J. Opt. Soc. Am. A, Vol. 25, 789 (1986).

10. R. Barakat, "Clipped correlation functions of aperture integrated laser speckle", Appl. Opt. Vol. 25, 3885 (1986).

11. F. E. Kragh, *Excursion Areas Of Intensity Due To Random Optical Waves*, M.S. Thesis, University of Central Florida, Electrical Engineering Dept. (1990).

12. D. Middleton, *Introduction to Statistical Communication Theory* , pg. 426-427 (McGraw-Hill, New York, 1960).

13. I. S. Gradshteyn and I. M. Ryzhik, *Table of Integrals, Series, and Products* , pg 940 (Academic Press, San Diego, 1980).

14. E. Jakeman, "Enhanced backscattering through a deep random phase screen", J. Opt. Soc. Am. A **5**, 1638-1648 (1988).

15. E. Jakeman,"Photon Correlation", in *Photon Correlation and Light Beating Spectroscopy*, edited by H. Z. Cummins and E. R. Pike (Plenum Press, New York, 1974), 75-149.

16. A. Papoulis, *Probability, Random Variables, and Stochastic Processes*, McGraw-Hill, N.Y. (1965).

17 R. N. Bracewell, *The Fourier Transform and its Applications*, pg. 115 (McGraw-Hill, New York, 1965).

18. J. D. Gaskill, *Linear Systems, Fourier Transforms, and Optics*, pg. 213 (John Wiley & Sons, New York, 1978).

19. R. C. Gonzalez and R. E. Woods, Digital Image Processing, pg. 41 (Addison Wesley, Reading , MA, 1992).

20 G. Cooper, C. McGillem, *Probabilistic Methods of Signal and System Analysis*', pg. 126-129 (Harcourt Brace Jovanovich College Publishers, (1986).

4.0. LIDAR SIMULATION

Arthur R. Weeks Ph.D. and Harley R. Myler, Ph.D.

4.1. Background

Modern laser radar systems possess a wide range of uses. There are a variety of detection schemes and modes of operation that can provide a range of information for any given application. For example, it is possible to configure a laser radar system to give range and velocity information about a given object relative to a radar receiver. Also, because laser wavelengths are short compared to microwave wavelengths and the output radiation is spatially coherent, laser radars possess high spatial resolution. This high spatial resolution allows the use of laser radar to effectively track moving objects, or to provide detailed images of objects through range or reflectivity data. This information is very useful in many military applications such as target acquisition and tracking, and in providing maps of surface terrain. The ability to track moving objects also finds use in the civilian world. Air traffic controllers use laser radars to identify and track incoming and outgoing aircraft [1] Commercial airline pilots use laser radar systems for precision aircraft navigation and guidance [2]. Meteorologists use laser radar to directly measure properties of the atmosphere such as wind direction and water vapor content [2].

A simplified diagram showing the operation of a laser radar system utilizing incoherent detection is illustrated in Figure 4.1.

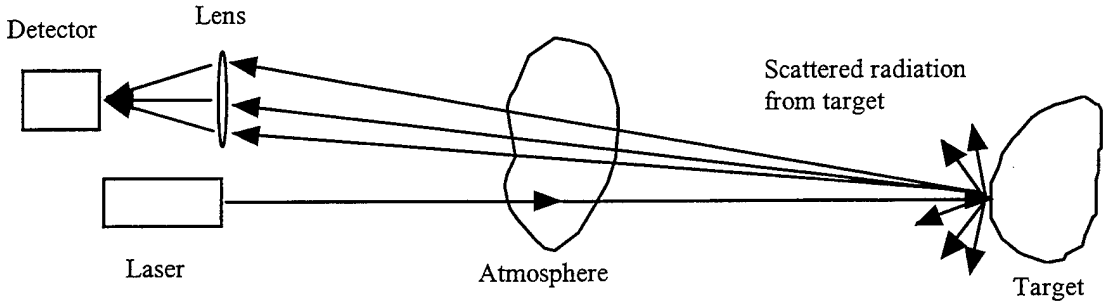


Figure 4.1: Simplified diagram of a laser radar system utilizing incoherent detection.

The operation is straightforward in that the detector measures the received intensity which is a function of target shape, size, reflectance, range to target, and atmospheric turbulence. Generally, a laser radar system using incoherent detection uses a pulsed laser in such a way that the range to a target can be determined by the time it takes for the pulse to travel to the target and back to the detector. This allows simple laser rangefinders to be constructed. Another use for incoherent laser radars is in the production of laser radar images. This is done by scanning the laser across a target and gathering range or intensity data at each scanning position.

The disadvantage of incoherent detection lies in the signal to noise ratio (SNR). An expression for the signal to noise ratio of a laser radar utilizing incoherent detection is [2]

$$SNR = \frac{\eta_D P_{sig}^{\ 2}}{2h\upsilon B\left[P_{sig} + P_{BK}\right] + \frac{\eta_D}{R^2}\left[P_{DK} + P_{TH}\right]}, \tag{4.1}$$

where η_D is the detector quantum efficiency, h = Planck's constant $(6.626 \cdot 10^{-34}\ J \cdot S)$, υ is the carrier frequency, B is the electrical bandwidth of the detector, P_{sig} is the received signal power, P_{BK} is the background power, P_{DK} is the equivalent dark current power of the detector, P_{TH} is the equivalent receiver thermal noise, q is the charge on an electron $(1.6 \cdot 10^{-19}\ C)$, and R is the detector current responsivity [2]. From this expression, it is evident that as P_{sig} decreases, the other noise sources in the denominator of Equation (4.1) will become dominant and the SNR decreases in value as P_{sig} decreases. The solution to this problem lies in being able to differentiate between the temporally and spatially coherent signal light, and incoherent background radiation P_{BK}.

A second method of detection used in laser radar systems is coherent detection. This method can drastically improve the SNR by being able to discriminate between the spatially and temporally incoherent background radiation and the coherent radiation from the laser reflected off the target. A block diagram showing a laser radar that uses coherent detection is shown in Figure 4.2. In Figure 4.2, HWP represents a half-wave plate, QWP represents a quarter-wave plate, PBS represents a polarizing beamsplitter, AOM represents an accousto-optic modulator, and NPBS represents a non-polarizing beamsplitter.

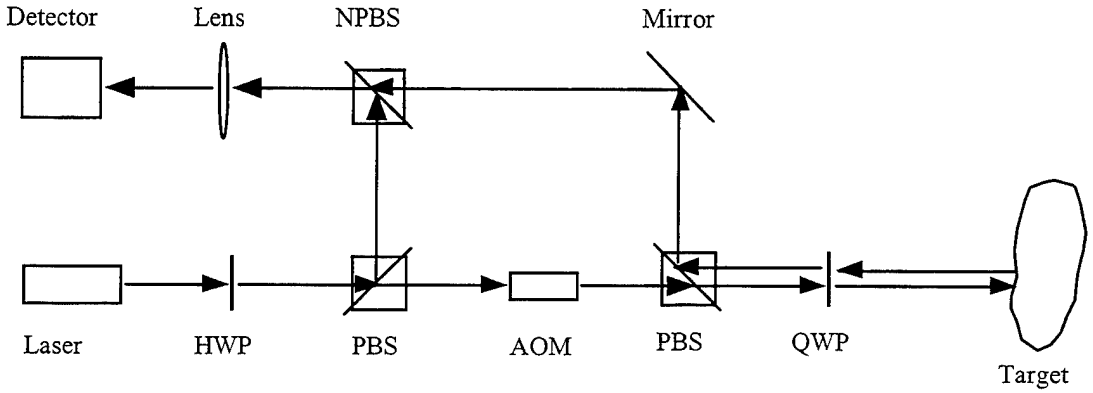


Figure 4.2: A bulk-optic monostatic laser radar system utilizing coherent detection.

As illustrated in Figure 4.2, a lens serves to focus the mixture of the return radar signal and the local oscillator onto the detector. The detector outputs a current that is composed of two components. Let $E_1(x,y)e^{-j\omega_1 t}$ and $E_2(x,y)e^{-j\omega_2 t}$ represent the signal wave and local oscillator, respectively in the detector plane then the current output of the detector will be proportional to

$$I_{out} \propto |E_1|^2 + |E_2|^2 + 2|E_1||E_2| \cdot \cos[(\omega_1 - \omega_2)t]. \tag{4.2}$$

As Equation (4.2) illustrates, the detector current output will consist of a DC component that represents the average power for both the local oscillator and signal waves falling on the detector, and a time varying component that oscillates at the difference frequency between the local oscillator and the return signal.

By making E_1 equal to E_2, full modulation can occur - i.e. the detector signal will vary between 0 and $4|E_1|^2$. In practical applications though, the modulation will be far less than 100%. The primary reasons for this are two-fold. First, the local oscillator power will usually be far greater than the signal power. Second, there will always be phase errors between the local oscillator and signal. The phase errors arise from imperfect transmitter/receiver optics, atmospheric turbulence, diffuse targets, diffraction effects, etc.

An expression for the signal-to-noise ratio, SNR, of a laser radar system using coherent detection is [2]

$$SNR = \frac{\eta_D P_{l.o.} P_{sig}}{h\upsilon B\left[P_{l.o.} + P_{sig} + P_{BK}\right] + \frac{h\upsilon}{2 \cdot q \cdot R}\left[P_{DK} + P_{TH}\right]}, \tag{4.3}$$

where $P_{l.o.}$ is the local oscillator power. If the local oscillator power is made very large so that it provides shot-noise-limited operation of the detector, Equation (4.3) reduces to

$$SNR = \frac{\eta_D P_{sig}}{h\upsilon B}. \tag{4.4}$$

In the incoherent case, if the signal power is much greater than any of the noise sources, then Equation (4.1) reduces to :

$$SNR = \frac{\eta_D P_{sig}}{2h\upsilon B}. \tag{4.5}$$

From Equation (4.4), the advantage of the local oscillator power being much larger than the return signal power can be readily seen. The advantage is that for weak signal power, it provides a higher SNR as compared to incoherent detection.

To maximize the SNR for a laser radar system using coherent detection, it makes sense to gather as much of the signal power scattered by the target as possible. A simple way to do this would be to make the receiver aperture as large as possible. However for coherent detection, this is often counterproductive because the transverse spatial coherence of the signal wave is typically reduced by atmospheric turbulence and irregular target surfaces. The random phase fluctuations across the signal wavefront will seriously reduce the heterodyne efficiency of the system when the corrupted signal wave is mixed with the uncorrupted local oscillator. This in turn limits the peak signal to noise ratio. If the receiver aperture is made smaller so that it has roughly the same diameter as the transverse coherence length of the return signal, the heterodyne efficiency of the detection will increase but the net amount of energy collected by the receiving aperture will decrease. This appears to be a lose-lose situation.

A way to get around the above difficulties is to design a receiver that has several small apertures about the size of the transverse coherence length of the return signal. Each separate aperture has its own local oscillator, so a separate beat signal is calculated for each individual aperture. The individual beat signals will have higher SNRs because the heterodyne efficiency will be higher. If the beat signals are co-phased together and then added coherently, this has the effect of increasing the net signal amplitude, while averaging the noise superimposed on the individual signals. Hence, the SNR has been increased. For N apertures that are co-phased and added together, the SNR will be increased by approximately N. A system of this type has the ability to provide a high SNR in the

presence of strong atmospheric turbulence and target scattering. For this reason it is of great interest to understand the implementation of such a system and all possible configurations. For a more detailed discussion of the implementation and design of a mulitple aperture coherent laser radar system the interested reader is referred to Section 2 of this report

As stated in the original proposal, it is the purpose of this task is to develop a computer simulation that accurately models a multiple aperture coherent laser system. The approach taken in the development of this computer simulator was to divide the simulation program into three components. The first component is a graphical user interface (GUI) that allows the easy visualization and modification of the system under design. Second, are programs that simulate the optical system. The third component is a set of functions that act as a bridge between the GUI and the simulation software. Such bridge functions are one of the defining characteristics of CAD systems. These functions are used to audit the design of the optical system being created. Some of the auditing functions act to prevent errors in design, such as having two lasers independently providing the system with two input beams. Other functions are used to check the placement of components in the system. CAD systems can and do exist without the aid of a GUI. Usually in such systems the user enters his or her data into a text file which is then later checked by an auditing routine. User interfaces such as this are more difficult to work with and are more prone to error than a graphical user interface.

Computer simulation of laser radar systems is not new. Many references can be found in the literature discussing this subject. Wang *et. al.* [3] wrote a simulation that uses radiometric methods to simulate short range incoherent laser radars. In this simulation, the target is taken to be diffuse. The simulation takes the shape of the target and the irradiance function illuminating it and predicts the return signal power function. Along the lines of this, Estep *et. al.* [4] wrote a simulation that takes experimentally determined monostatic reflectivity data on a material and combines this with information on target shape. This information is in turn combined with the irradiance pattern on the target to generate a laser radar cross section. Both of these simulations don't preserve phase information and model incoherent systems. In addition, multiple aperture operation isn't supported.

A simulation by Letalick and Oestberg [1] modeled an imaging coherent laser radar, but unlike the simulation presented here, they did not explicitly calculate the heterodyned signal. In Letalick and Oestberg's simulation, by specifying an atmospheric index of refraction structure constant (C_n^2), the program simulates the effect of turbulence induced noise. This is done by first simulating a noise free intensity image, and then corrupting the original intensity image through a series of transformation equations [4]. The heterodyne efficiency is a fixed parameter specified for the entire simulation This prevents the simulation from being able to model the time evolving change of a heterodyned signal. The purpose of their simulation was to simulate range and intensity images of objects placed in front of different backgrounds. These images were then analyzed to determine ways to improve the contrast between an arbitrary target and background. The simulation presented here is different from the one by Letalick and Oestberg in that both time-evolved and spatial components of the return signal are simulated. Their simulation only provides a single spatial measurement.

The technical report by Peters and Nomiyama [5] simulated a system for analyzing the effect that a rotating diffuse target has on a quadrant detector. This simulation used Fraunhofer diffraction between the receiver and

target, and also used a Fourier transform to model the lens. Because the actual diffraction between optical elements is modeled in this simulation, the actual heterodyne detection process is modeled. This is done by using the lens to focus both the return signal and the local oscillator onto the detector. Next, the beat frequency is calculated by integrating the square of the sum of the local oscillator and signal over the surface of the detector. Because the diffuse target is rotating, this system can model the time evolved heterodyned output signal where each time realization corresponds to an incremental rotation of the target. The main limitation with this simulation is that it does not simulate atmospheric turbulence and it can't simulate effects such as lens defocus.

The report by J. C. Leader [6] does not actually model a coherent laser radar system, but obtains analytical expressions for heterodyne efficiency and then numerically evaluates the results. The procedure Leader uses to generate his numerical data was to specify the return signal and the local oscillator field profiles, and then introduce a coherence function between the two that modeled the effects of target roughness and atmospheric turbulence [6]. Next, a computer was used to find the time averaged heterodyne efficiency by numerically evaluating his results for different target surfaces, atmospheric conditions, receiver parameters, etc. This approach to analyzing a laser radar system does not allow the time evolution of the heterodyne efficiency to be examined.

R. G. Frehlich [7] published a report describing a laser radar simulator similar to Leader's except Frehlich does not include atmospheric turbulence. Frehlich's simulation allows for the target to be located in the near or far field, and explicitly models these propagations using the Huygens Fresnel integral. Frehlich's analysis uses targets that possess diffuse surfaces with delta function autocorrelation functions [7]. The lens that focuses the return signal and the local oscillator onto the detector is modeled by a quadratic phase transformation and the focusing operation is performed by a Fourier transform. Values of the heterodyne efficiency are calculated a single realization at a time by providing statistically independent target surfaces. Ensemble averages are then easily computed by averaging all realizations. However this technique does not allow one to study the real-life time evolution of these effects. In addition, this simulation does not allow for a finite size detector - i.e. the detector in this simulation collects all of the signal and local oscillator power.

The report by Brown *et. al.* [8] entitled 'Simulation of Ground to Space Optical Propagation' presented a computer simulation that modeled the operation of a ground to space adaptive optics system. In their simulation, the ground to space propagation is modeled included including the effects of the atmosphere. The atmosphere was simulated using Kolmogorov phase screens. The simulation also modeled the phase detectors and associated hardware that comprises an adaptive optic system that corrects for atmospheric effects. While this is not a laser radar system, it does share some similarities with the simulation presented here in that both simulations use time correlated phase screens. Though the adaptive optics simulation by Brown *et. al.* did use time correlated phase screens, their approach was to create a single large phase screen and shift it one column at a time across the incident beam [8]. In this manner, the maximum number of time realizations possible was limited by the ratio of the size of the phase screen to the size of the incident laser beam.

A paper by A. L. Kachelmyer [9] determined the heterodyne efficiency of a return signal that was reflected from a mirror possessing aberrations (approximated by Zernike polynomials), and a plane wave local oscillator. By focusing the two beams onto a detector, Kachelmyer studied effects that various optical imperfections would have on the heterodyne efficiency. K. M. Iftekharuddin *et. al.* [10] used computer simulations to examine the effects of

tilt and offset of the local oscillator on the heterodyne efficiency using diffraction-free Bessel beams. K. Tanaka *et. al.* [11] use the same general method as Iftekharuddin except they examined Gaussian beams.

The laser radar computer simulated presented here uses the C programming language and was developed on a Sparc 5 Sun workstation running under X-11 Windows. This simulation will model a multiple aperture laser radar transmitter/receiver system utilizing coherent detection. With this simulation, the user can design a laser radar system by specifying component parameters such as receiving aperture placement and size, and then investigate the effect that factors such as target range, atmospheric turbulence, and target roughness have on the SNR of the output. To meet the original proposed goal, several tasks had to be implemented in the process.

1. Define a Graphical User Interface for the entry of the optical components used in a typical laser radar.
2. Model each one of the optical components used in the laser radar system
3. Combine the GUI interface with the models of each component.
4. Verify the final performance of the laser radar simulator

The remaining sections describe in detail each one of the above tasks and how they were used in the laser radar computer simulator.

4.2 Accomplishments

In brief, our major accomplishments in the part of the program are:

- The individual components of an actual laser radar such as wave plates, polarizing beamsplitters, lenses and detectors are modelled.
- More than one receiver aperture is allowed
- The presence of components like the wave plates and polarizing beamsplitters allows this simulation to handle arbitrary light polarizations.
- A scattering model that includes depolarization effects can easily be modeled.
- Fresnel diffraction is included in the propagation model
- Increased accuracy of the lens routine is achieved compared with previous simulations.
- Spatially and temporally evolved scattering is included in the simulation

To understand the approach taken in the developing the laser radar simulator, a summary of the components used in a single aperture laser radar is first presented followed by the bolck diagram describing a two - element laser radar system. In any multiple aperture laser radar system, there will be at least one transmit aperture that illuminates the target with laser light. If this aperture also acts as a receiver, then it is called the monostatic aperture. The operation of the monostatic aperture as modeled in the computer simulation is illustrated below in Figure 4.3. Figure 4.3 illustrates every component present in the monostatic aperture as well as the light polarization at each point within the system.

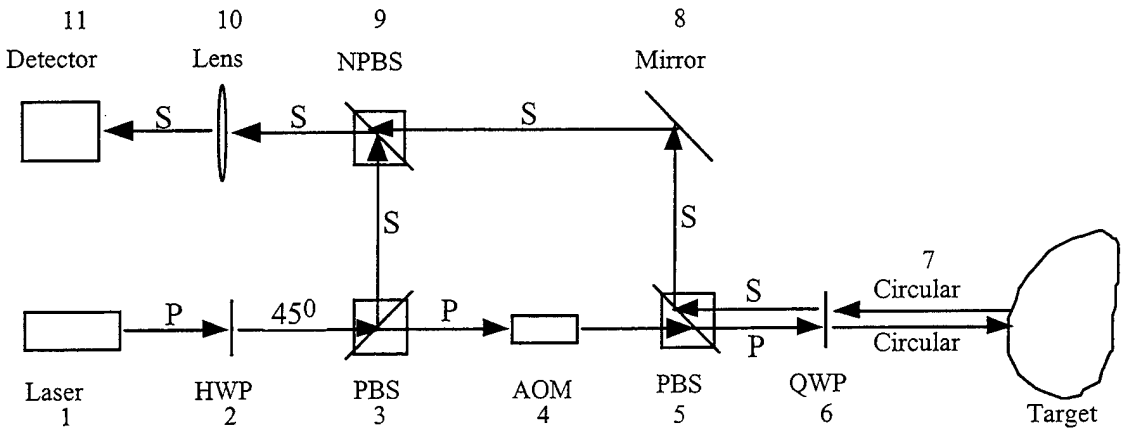


Figure 4.3: Monostatic aperture of the laser radar simulator.

The P represents light polarized along the P polarization reference axis which is in the plane of the paper and perpendicular to the optic axis. The S represents light polarized along the S polarization reference axis which is perpendicular to the plane of the paper and the optic axis (pointing into the paper). In this computer simulation, every component shown in Figure 4.3 is explicitly modeled with the exception of the acousto-optic modulator (component 4). In this simulation, it is just assumed that the return signal has been frequency shifted by the same frequency as the AO modulator's frequency. A summary of each component as it is modeled in the computer simulator is given as follows:

1. The first component is the laser. The laser simulator produces a Gaussian beam with a specified power, polarization, radius of curvature, and beam radius. This beam is propagated throughout the system using ABCD matrices where possible.
2. The second component is the half-wave plate. The wave plate simulator allows either half-wave or quarter-wave operation. In addition, the wave plate simulator allows the angular orientation of the slow axis to be specified. The slow axis of the half-wave plate makes an angle of 22.5^0 with the S axis so that the light exiting the wave plate is linearly polarized at 45^0 with respect to the S and P axes.
3. The third component is a polarizing beamsplitter. The polarizing beamsplitter allows incident light polarized along the P polarization axis to pass straight through the beamsplitter. If the incident light is polarized along the S polarization reference, the light is deflected at a 90^0 angle to the incident light. In addition, the polarizing beamsplitters allow for finite extinction ratios in both the forward and side directions. The S polarization that is deflected by polarizing beamsplitter number 3 is the local oscillator for the coherent detection process.
4. The component following the polarizing beamsplitter is the acousto-optic modulator. The purpose of this component of the laser radar system is to shift the transmit signal by a given frequency.
5. Immediately following the acousto-optic modulator is another polarizing beamsplitter. This beamsplitter works in conjunction with quarter-wave plate number 6 to provide an 'optical transmit/receive switch'.

6. Quarter-wave plate number 6 is part of the optical transmit/receive switch. The function of the optical switch is to allow the same aperture to be used as a transmitter and receiver. The optical switch works by passing the P polarized light emerging from the acousto-optic modulator straight through the polarizing beamsplitter (5) to illuminate the quarter-wave plate (6). This quarter-wave plate has its slow axis rotated by 45^0 with the S reference axis, so that the light passing through it is circularly polarized in the clockwise direction. The light then propagates to the target, where it is scattered back to the receiver. If the scattering is weak, the light will still be circularly polarized when it arrives back at the quarter-wave plate. When the circularly polarized light passes back through the wave plate, it will be linearly polarized along the S axis. This allows the polarizing beamsplitter to deflect the light 90^0 so that it may be mixed with the local oscillator at the non-polarizing beamsplitter, and then focused onto the detector by the lens.
7. The free space propagation between the transmitter/receiver and target is modeled by Fresnel diffraction. For many of the simulated experiments implemented under this research project, there will be phase screens placed within the propagation path to simulate atmospheric turbulence.
8. After the light has returned from the target and gets deflected by the polarizing beam-splitter (component 5), the mirror serves to guide the light so that it may be mixed with the local oscillator. (Note: Diffraction is neglected between wave plate 6 and the lens for the waves scattered off the target. The local oscillator is propagated to the lens with ABCD matrices.)
9. This non-polarizing beamsplitter mixes the return signal and the local oscillator beams.
10. The lens focuses the signal/local oscillator mixture onto the detector. The lens simulator allows the following parameters to be specified : the lens diameter, the lens focal length, distance from the lens plane to the detector plane, and pan and tilt angles to simulate lens misalignment. The lens is modeled by a quadratic phase transformation and a uniform transmittance. The diffraction between the lens and detector is valid for lenses with an F# as low as 4. This is accomplished by reclaiming some of the higher order terms that are normally removed when making the Fresnel approximation.
11. The detector calculates the DC components of the local oscillator and the return signal as well as the AC portion (beat signal). The DC components are fed into a Gaussian noise generator which adds simulated shot noise to the AC output. The detector simulator allows the size, shape, responsivity, sampling rate, and the 3dB point of the detector to be specified. The Gaussian noise generator also allows the dark current and the current noise of the detector to be specified.

In order to simulate the operation of a multiple aperture system, bistatic apertures are also modeled. Figure 4.5 shows a bistatic aperture as it is modeled in the computer simulator. The simulation allows the user to arrange the monostatic aperture and numerous bistatic apertures so that any synthetic aperture configuration can be modeled. One merely has to specify the location of each aperture, the aperture diameter, the focal length of the lens behind it, the detector size and shape, etc. Each receiving aperture is independent of the others in that it has its own set of parameters that govern the output when the signal is mixed with the local oscillator and focused onto the

detector. Section 4.2.1 presents the models for each one of the optical components implemented in the laser radar system, while Section 4.2.2 gives the graphical user interface description used by the laser radar simulator. Finally, Section 4.2.3 gives some results generated by the laser radar simulator.

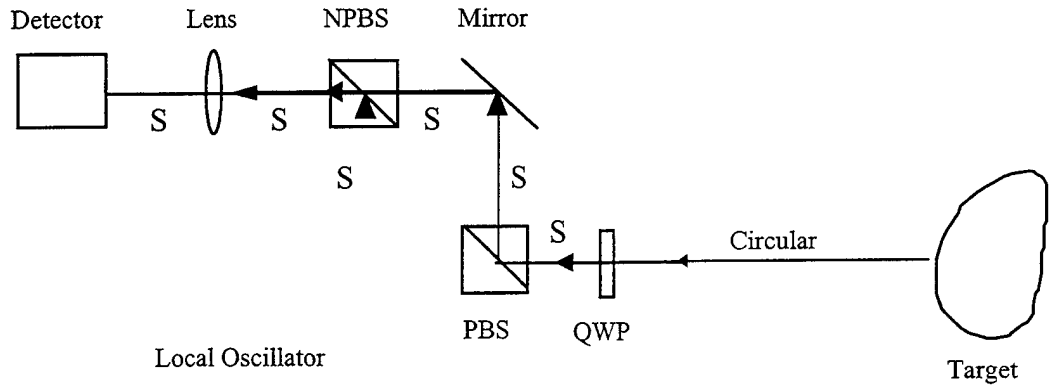


Figure 4.4: A bistatic laser radar receiver.

4.2.1 The models for the optical components used within the laser radar simulator

The four most complex parts of the laser radar simulation are the free space propagation model, the lens model, the detector model, and the phase screen model. This section gives the theoretical background for each of these components plus the other optical components implemented in this computer simulation. The free space propagation and the lens theory will be developed in continuous space, while the phase screen theory and the detector will be presented in discrete space.

4.2.1.1 Free Space Propagation and Lens Theory

A very important part of the simulation is the method by which light waves are propagated between the target and receiver/transmitter and within the receiver/transmitter itself. To accomplish this task, the Rayleigh-Sommerfeld diffraction integral [12] is used.

$$U(\alpha,\beta) = \frac{1}{j\lambda} \iint_{(x,y)\in\Sigma} U(x,y) \frac{e^{jkr}}{r} \cos(\bar{n}\cdot\bar{r})\, dxdy, \tag{4.6}$$

where $r = \{(x-\alpha)^2 + (y-\beta)^2 + z^2\}^{1/2}$,

(x,y) are object plane coordinates,

(α,β) are observation plane coordinates,

$(x,y) \in \Sigma$ denotes integration over object plane aperture,

$\bar{n}$ is the unit normal vector in the object plane,

$\bar{r}$ is the unit vector in the direction of r,

$\cos(\bar{n}\cdot\bar{r})$ is the cosine of the angle between $\bar{n}$ and $\bar{r}$, and

$A(x,y)$ is the complex amplitude distribution in the object plane.

A diagram showing the geometry associated with Equation (4.6) is shown in Figure 4.5.

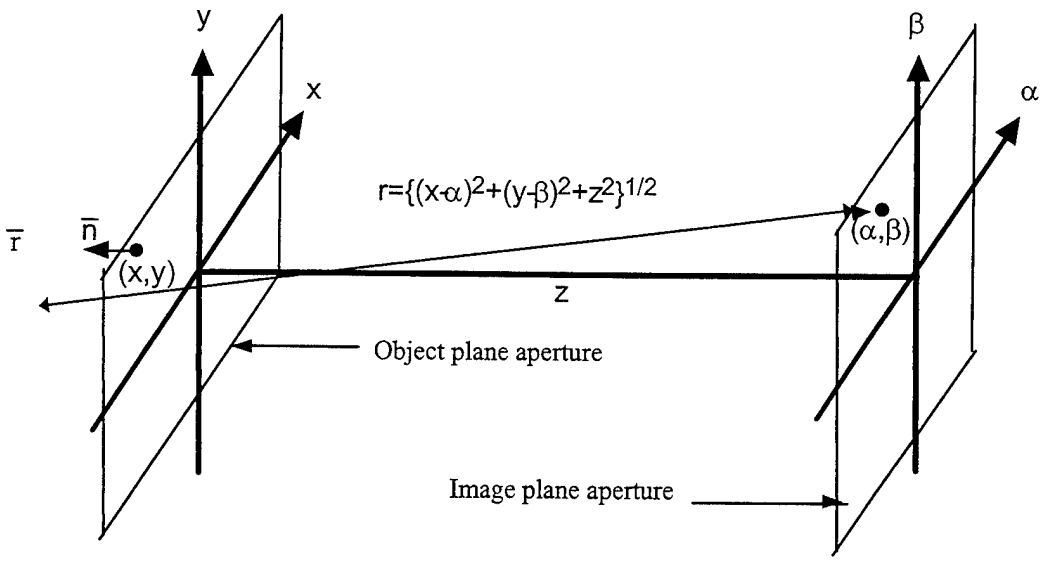


Figure 4.5: Geometry describing Rayleigh-Sommerfeld diffraction.

The actual form of the Rayleigh-Sommerfeld diffraction equation formula used for the free-space propagation within this simulation is derived by first making approximations on r.

Without any approximations,

$$r = \{(x-\alpha)^2 + (y-\beta)^2 + z^2\}^{1/2} = z\{1+\frac{(x-\alpha)^2+(y-\beta)^2}{z^2}\}^{1/2}. \tag{4.7}$$

Now, expanding Equation (4.7) using the Binomial Theorem,

$$z\{1 + \frac{(x-\alpha)^2+(y-\beta)^2}{2z^2} - \frac{[(x-\alpha)^2+(y-\beta)^2]^2}{8z^4} + \ldots\} \tag{4.8}$$

is obtained. In Equation (4.8), all of the higher order terms beyond

$$\frac{(x-\alpha)^2+(y-\beta)^2}{2z^2} \tag{4.9}$$

are negligible if the condition

$$z^3 >> \frac{\pi}{4\lambda}[(x-\alpha)^2+(y-\beta)^2]^2_{max} \tag{4.10}$$

is met. This is known as the Fresnel approximation [12]. Thus

$$U(\alpha,\beta) = \frac{e^{jkz}}{j\lambda z}\iint_{(x,y)\in\Sigma} U(x,y)e^{j\frac{\pi}{\lambda z}[(x-\alpha)^2+(y-\beta)^2]}dxdy. \tag{4.11}$$

Hence, Equation (4.11) can be written as

$$U(\alpha,\beta) = \frac{e^{jkz}}{j\lambda z}\Im^{-1}\left\{\Im\left\{\iint_{(x,y)\in\Sigma} U(x,y)e^{j\frac{\pi}{\lambda z}[(x-\alpha)^2+(y-\beta)^2]}dxdy\right\}\right\}$$

$$= \frac{e^{jkz}}{j\lambda z}\Im^{-1}\left\{\Im\{U(x,y)\}\Im\left\{e^{j\frac{\pi}{\lambda z}(x^2+y^2)}\right\}\right\}$$

$$= e^{jkz}\Im^{-1}\left\{\Im\{U(x,y)\}e^{-j\pi\lambda(f_x^2+f_y^2)}\right\} \tag{4.12}$$

where f_X and f_Y are the spatial frequencies in the x and y directions respectively, and $\Im\{A(x,y)\}$ is the two-dimensional Fourier transform of A(x, y) :

$$\Im\{A(x,y)\} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} A(x,y)e^{-j2\pi(xf_x+yf_y)}dxdy$$

$$\Im^{-1}\{A(x,y)\} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} A(x,y)e^{j2\pi(xf_x+yf_y)}df_xdf_y \tag{4.13}$$

Equation (4.13) uses the two-dimensional Fourier transform to calculate the electric field in the (a, b) observation plane. Equation (4.13) will yield a valid propagation result for any combination of the object aperture, image aperture, and propagation length that satisfies the Fresnel condition.

As stated earlier, a model is needed that predicts the electric field at the focal plane of the lens. The foundation for this operation is once again Equation (4.6). However in this case the approximations used to arrive at Equation (4.11) are not adequate. This is because typical lenses that must be modeled have a F# of around 5.

The key to modifying Equation (4.11) so that it will work for the lens system is in the Binomial expansion for r (Equation (4.8)).

$$r = z + \frac{(x-\alpha)^2+(y-\beta)^2}{2z} - \frac{[(x-\alpha)^2+(y-\beta)^2]^2}{8z^3} + \ldots \tag{4.14}$$

To arrive at Equation (4.11), only the first two terms in Equation (4.8) were kept. In order to allow Equation (4.11) to work with smaller values of z, some of the higher order terms that were ignored must also be included.

The goal of the remainder of this section is to examine the terms that were ignored in deriving Equation (4.11) and re-incorporate them so that lenses with small F numbers can be modeled. In addition, an expression analogous to Equation (4.10) will derived for the new expression that gives the relative lens sizes, detector sizes, and propagation distances for which the new expression is valid.

Expanding the third term in Equation (4.14) yields :

$$j\frac{2\pi}{\lambda}\frac{[(x-\alpha)^2+(y-\beta)^2]^2}{8z^3} = j\frac{\pi}{4\lambda z^3}(x^4-4x^3\alpha-4x\alpha^3+6x^2\alpha^2+\alpha^4+y^4$$
$$-4y^3\beta-4y\beta^3+6y^2\alpha^2+\beta^4+2x^2y^2$$
$$-4x^2y\beta+2x^2\beta^2-4x\alpha y^2+8xy\alpha\beta \qquad (4.15)$$
$$-4x\alpha\beta^2+2\alpha^2y^2-4y\alpha^2\beta+2\alpha^2\beta^2).$$

For most radar applications of interest to the multiple aperture program, the input aperture lenses will have a diameter of approximately 1 centimeter, and the detectors will have a diameter of around 60 microns. Substituting these numbers into Equation (4.15), the terms

$$j\frac{\pi}{4\lambda z^3}(x^4+2x^2y^2+y^4) \qquad (4.16)$$

will be the most significant, and the terms

$$-j\frac{\pi}{4\lambda z^3}(4x^3\alpha+4y^3\beta+4xy^2\alpha+4x^2y\beta). \qquad (4.17)$$

will be the least significant. Hence, including only the terms given in Equation (4.16), the equation used to model lenses with F# as low as 3 is

$$U(\alpha,\beta)=\frac{e^{jkz}}{j\lambda z}\iint_{(x,\,y)\in\Sigma}U(x,y)e^{j\frac{\pi}{\lambda z}[(x-\alpha)^2+(y-\beta)^2]}e^{-j\frac{\pi}{4\lambda z^3}(x^4+2x^2y^2+y^4)}dxdy. \qquad (4.18)$$

Rearranging Equation (4.18) further, we obtain

$$U(\alpha,\beta)=\frac{e^{jkz}}{j\lambda z}e^{j\frac{\pi}{\lambda z}(\alpha^2+\beta^2)}\iint_{(x,\,y)\in\Sigma}\left\{U(x,y)e^{j\frac{\pi}{\lambda z}(x^2+y^2)}e^{-j\frac{\pi}{4\lambda z^3}(x^4+2x^2y^2+y^4)}e^{-j\frac{2\pi}{\lambda z}(x\alpha+y\beta)}\right\}dxdy. \qquad (4.19)$$

Equation (4.19) is the equation that is used to simulate the lens propagation. It is valid for any combination of the lens aperture, detector aperture, and lens/detector distance that satisfies

$$\frac{\pi}{\lambda z^3}r^3r'\le 0.1 \text{ radian}. \qquad (4.20)$$

4.2.1.2 Detector

This section discusses the theory behind the operation of the detector. First, the operation of integrating the intensity of the sum of the signal and L.O. waves over the surface of the detector will be discussed. Second, the simulation of the detector noise will be outlined. Key to the simulation of the detector noise is a lowpass filtering operation used to simulate the finite bandwidth of the detector. The last part of this section will outline the process by which the final detector AC current output is obtained through the collective process of calculating the

heterodyned signal, generating the random detector noise, and then lowpass filtering the detector noise added to the detected AC signal.

A heterodyned signal consists of mixing a return signal wave carrying information with a reference wave (local oscillator). Once the signal wave and local oscillator wave are mixed and focused onto the detector by a lens, the detector must calculate the time varying intensity incident on its surface. This process is illustrated in Figure 4.6. In reference to Figure 4.6, the following equations will describe how the heterodyned signal is obtained from the detector. The process will be illustrated using the S polarization component of the local oscillator and signal wave. Let the S polarization of signal wave in the detector plane be

$$A_S(x,y)e^{j(\varphi_{A_S}(x,y)+\omega_A t)}, \tag{4.21}$$

where $A_S(x, y)$ represents the signal amplitude as a function of position across the detector, $\varphi_{A_S}(x,y)$ represents the phase as a function of position across the detector, and ω_A is the frequency shifted signal wave.

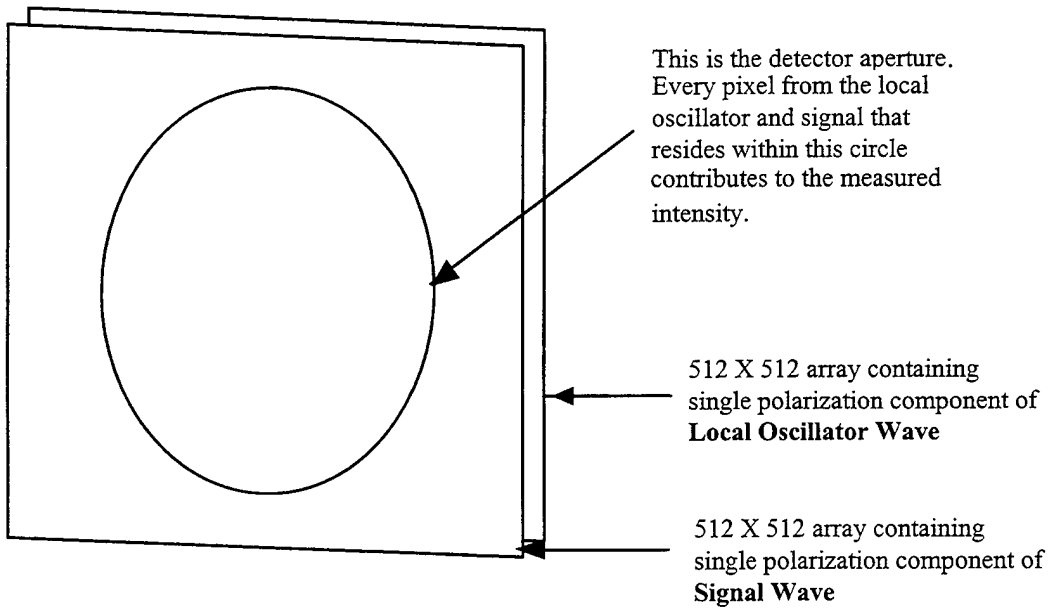


Figure 4.6 : Figure showing simplified detector operation.

Let the S polarization of the local oscillator (L.O.) wave in the detector plane be

$$B_S(x,y)e^{j(\varphi_{B_S}(x,y)+\omega_B t)}, \tag{4.22}$$

where $B_S(x, y)$ is the L.O. amplitude as a function of position across the detector, $\varphi_{B_S}(x,y)$ is the phase of the L.O. as a function of position across the detector, and ω_B is the frequency of the local oscillator (not frequency shifted).

Given Equations (4.21) and (4.22), the detector output current as a function of time for the S component of the radiation incident on the detector is

$$R\frac{\varepsilon_0 c}{2}\sum_{(m,n)\in\alpha}\left|A_S(m,n)e^{j(\varphi_{A_S}(x,t)+\omega_A t)}+B_S(m,n)e^{j(\varphi_{B_S}(x,y)+\omega_B t)}\right|^2\Delta x\Delta y, \tag{4.23}$$

where a represents all points that lie within the detector aperture, R represents the responsivity of the detector (Amperes/Watt), e_0 is the permitivity of free space, c is the speed of light in free space, Dx is the spatial sampling interval in the detector plane in the x-direction, and Dy is the spatial sampling interval in the detector plane in the y-direction. Expanding the arguments within the summation in Equation (4.23) yields

$$\begin{aligned}&|A_S(m,n)|^2+|B_S(m,n)|^2\\&\quad+2|A_S(m,n)||B_S(m,n)|\cos\left\{(\varphi_{A_s}(x,t)-\varphi_{B_s}(x,y))+(\omega_A-\omega_B)t\right\}.\end{aligned} \tag{4.24}$$

Using $\cos(\beta+\gamma)=\cos\beta\cos\gamma-\sin\beta\sin\gamma$ in Equation (4.24) and substituting the result back into Equation (4.24), the final expression for the detector current due to the incident S polarized light is

$$\begin{aligned}R\frac{\varepsilon_0 c}{2}\Delta x\Delta y\Bigg[&\sum_{(m,n)\in\alpha}\left\{|A_S(m,n)|^2+|B_S(m,n)|^2\right\}\\&+\cos(\omega_A-\omega_B)t\sum_{(m,n)\in\alpha}2|A_S(m,n)||B_S(m,n)|\cos(\varphi_{A_s}(m,n)-\varphi_{B_s}(m,n))\\&-\sin(\omega_A-\omega_B)t\sum_{(m,n)\in\alpha}2|A_S(m,n)||B_S(m,n)|\sin(\varphi_{A_s}(m,n)-\varphi_{B_s}(m,n))\Bigg].\end{aligned} \tag{4.25}$$

Equation (4.25) gives the output beat signal of the heterodyned waves as a function of time. To obtain the total current output from the detector, Equation (4.25) must be repeated for the P polarized light incident on the detector (subscripts in Equation (4.25) change to P subscripts). In most laser radar applications the first term of Equation (4.25) (DC term) is not of interest. Letting the last two terms of Equation (4.25) define the AC portion of the detected current , the AC current for a single polarization component will be

$$\begin{aligned}\Omega_\gamma(t)=\ &R\frac{\varepsilon_0 c}{2}\Delta x\Delta y\cdot\\&\Bigg[\cos(\omega_A-\omega_B)t\sum_{(m,n)\in\alpha}2|A_\gamma(m,n)||B_\gamma(m,n)|\cos(\varphi_{A_\gamma}(m,n)-\varphi_{B_\gamma}(m,n))\\&-\sin(\omega_A-\omega_B)t\sum_{(m,n)\in\alpha}2|A_\gamma(m,n)||B_\gamma(m,n)|\sin(\varphi_{A_\gamma}(m,n)-\varphi_{B_\gamma}(m,n))\Bigg],\end{aligned} \tag{4.26}$$

where g is either S or P depending on the polarization. The final AC current detector output without noise will be

$$\Omega_{total}(t)=\Omega_S(t)+\Omega_P(t). \tag{4.27}$$

To make the output more realistic, noise must be added to the AC signal in Equation (4.27). The noise energy for the detector depends on the dark current, the detector current noise, the bandwidth of the detector, and the shot noise due to the DC signal present at the detector. The time varying noise produced by the detection process is simulated by adding uncorrelated zero mean Gaussian distributed random numbers to the AC current output. Zero mean Gaussian distributed random numbers are generated with the following equation [13] :

$$\text{Gaussian R.V.} = \left[-2 \cdot \sigma^2 \cdot \ln X\right]^{1/2} \cdot \cos\theta \,, \tag{4.28}$$

where X is a uniformly distributed random number between 0 and 1, θ is a uniformly distributed random number between 0 and 2p (independent of X), and s^2 is the desired variance. The variance of the noise produced by the detection process is given by [14] :

$$\sigma^2_{\text{noise}} = \left[2q(I_D + I_{DC_S} + I_{DC_P}) + I_N{}^2\right] \cdot \frac{\pi}{2} f_0 \,, \tag{4.29}$$

where q is the charge on an electron ($1.602 \cdot 10^{-19}$ C), I_D is the dark current on the detector, in is the current noise of the detector, f_0 is the 3dB point of the detector,

$$I_{DC_S} = R\frac{\varepsilon_0 c}{2} \sum_{(m,n)\in\alpha} \left\{ \left|A_S(m,n)\right|^2 + \left|B_S(m,n)\right|^2 \right\}, \tag{4.30}$$

and

$$I_{DC_P} = R\frac{\varepsilon_0 c}{2} \sum_{(m,n)\in\alpha} \left\{ \left|A_P(m,n)\right|^2 + \left|B_P(m,n)\right|^2 \right\}, \tag{4.31}$$

which are the DC currents on the detector from the S and P polarizations respectively.

To produce the noise added to the AC detector current, a time sequence of uncorrelated Gaussian random numbers is generated using Equations (4.28) and (4.29). The next step in the process is to filter the noise to match the passband of the detector. A block diagram describing the noise generation process is given in Figure 4.7.

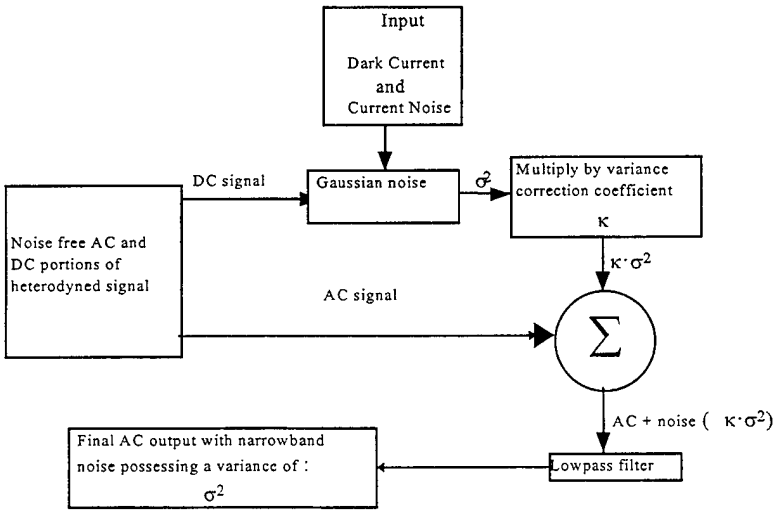


Figure 4.7 : Process of adding narrowband noise to the AC current output of a detector.

A first order lowpass filter is used model the filter for the detector. In the Laplace domain, this filter is defined as [15] :

$$H(s) = \frac{\omega_0}{s+\omega_0}. \tag{4.32}$$

where $w_0 = 2pf_0$ is the 3dB point of the filter in radians/second. To find the discrete version of Equation (4.32) H(z) as required by the simulation, the bilinear transform [15] is used resulting in

$$H(z) = \frac{\omega_0}{\frac{2}{T}\left(\frac{1-z^{-1}}{1+z^{-1}}\right)+\omega_0}, \tag{4.33}$$

where T is the time sampling period of the laser radar simulator. Rearranging Equation (4.33), we obtain

$$H(z) = \frac{Y(z)}{X(z)} = \frac{\frac{\omega_0 T}{2}(1+z^{-1})}{(1-z^{-1})+\frac{\omega_0 T}{2}(1+z^{-1})}. \tag{4.34}$$

where X(z) is the Z-transform of the time sequence of the uncorrelated Gaussian random variables and Y(z) is the Z-transform of the filter output. Rewriting Equation (4.34) yields

$$(1-z^{-1})Y(z)+\frac{\omega_0 T}{2}(1+z^{-1})Y(z) = \frac{\omega_0 T}{2}(1+z^{-1})X(z) \xrightarrow{\text{Rearranging further,}}$$

$$Y(z)\left(1+\frac{\omega_0 T}{2}\right) - \frac{Y(z)}{z}\left(\frac{\omega_0 T}{2}-1\right) = \frac{\omega_0 T}{2}\left(X(z)+\frac{X(z)}{z}\right). \tag{4.35}$$

Applying the inverse Z-transform to Equation (4.35) yields

$$Y(nT)\left(1+\frac{\omega_0 T}{2}\right) - Y[(n-1)T]\left(\frac{\omega_0 T}{2}-1\right) = \frac{\omega_0 T}{2}\left(X(nT)+X[(n-1)T]\right). \tag{4.36}$$

Finally, solving Equation (4.36) for Y(nT) in terms of X[nT], X[(n-1)T], and Y[(n-1)T] yields the difference equation of the lowpass filter given in Equation (4.32) :

$$Y(nT) = \frac{\frac{\omega_0 T}{2}\left(X(nT)+X[(n-1)T]\right) - \left(\frac{\omega_0 T}{2}-1\right)Y[(n-1)T]}{\left(1+\frac{\omega_0 T}{2}\right)}. \tag{4.37}$$

As shown in Figure 4.7, in order to maintain the noise power during the filtering operation, a correction factor is used to scale the variance before implementation of the lowpass filter. The correction factor is [16]

$$\kappa = 1+\frac{1}{\tan\left[\frac{\omega_0 T}{2}\right]}. \tag{4.38}$$

Using the correction factor in conjunction with Equation (4.29) establishes that the variance added to the AC signal before low pass filtering will be

$$\sigma^2_{noise} = \left[2q(I_D + I_{DC_S} + I_{DC_P}) + in^2\right] \cdot \frac{\pi}{2} f_0 \cdot \left\{1 + \frac{1}{\tan\left[\frac{\omega_0 T}{2}\right]}\right\}. \tag{4.39}$$

Applying the lowpass filter to the uncorrelated Gaussian noise with the variance given by Equation (4.39) results in the correct noise variance given by Equation (4.29).

The detector noise generated as shown in Figure 4.7 is finally added to the AC detected signal as defined by Equations (4.26) and (4.27). The result is a noisy detected laser radar signal.

4.2.1.3 Phase Screen Theory

The different subroutines discussed for the laser radar simulator can be linked together to form a variety of different simulated experiments. In several instances, the experiments to be simulated will involve atmospheric turbulence. The approach taken to model atmospheric turbulence in the laser radar simulator is to use phase screens constructed with a Kolmogorov power spectrum. The Kolmogorov power spectrum is defined as [17]

$$\Phi_n(\kappa) = .033 \cdot C_n^2 \kappa^{-\frac{11}{3}}, \tag{4.40}$$

where

$$\kappa = 2\pi \cdot \left[\frac{\alpha^2}{(N\Delta x)^2} + \frac{\beta^2}{(N\Delta y)^2}\right]^{1/2}, \tag{4.41}$$

C_n^2 is the strength of turbulence, κ is the two-dimensional spatial frequency, and dz is the thickness of the phase screen. The terms Dx and Dy are the spatial sampling intervals of the screen in the x and y directions, the terms a/NDx and b/NDy are the spatial frequencies in the x and y directions, and N is the total number of points used to describe the phase screen. Using this model for the Kolmogorov spectrum, the basic procedure for generating a single screen is presented in Figure 4.8.

The first step in Figure 4.8 is to fill an array with uncorrelated Gaussian random numbers with a variance of 1. This is done via a standard independent uniform random number generator used to fill two two-dimensional arrays with statistically independent random numbers between 0 and 1 and 0 and 2p. These arrays are then used to compute a

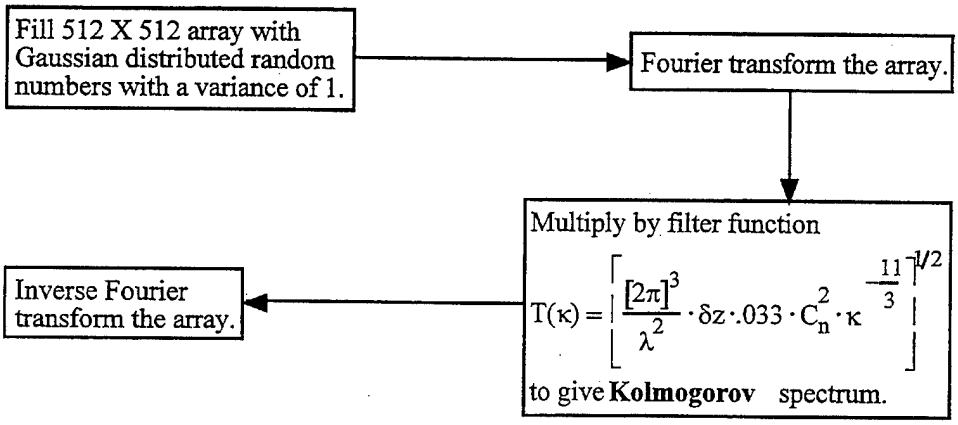


Figure 4.8: Flowchart showing the procedure for generating a single phase screen.

new array of statistically independent Gaussian numbers with a zero mean and an variance s^2 using the equation [16]

$$\text{Gaussian R.V.} = \left[-2 \cdot \sigma^2 \cdot \ln X\right]^{1/2} \cdot \cos\theta \,, \tag{4.42}$$

where q is an uniformly distributed number between 0 and 2p, and X is a uniformly distributed number between 0 and 1 (independent of q). The uncorrelated random numbers generated from Equation (4.42) produce random phase screens that are Gaussian distributed and that contain zero spatial correlation. The second step in the flowchart is to Fourier transform the array of Gaussian random numbers using the discrete Fourier transform

$$\Im\{A(m,n)\} = \Delta x \Delta y \sum_{m=0}^{N-1} \sum_{n=0}^{N-1} A(m,n) e^{-j\frac{2\pi}{N}(mm'+nn')} \,. \tag{4.43}$$

The spatial correlation is applied to each phase screen realization (time sample) by filtering the real and imaginary Fourier components by the filter function [18],

$$T(\kappa) = \left[\frac{[2\pi]^3}{\lambda^2} \cdot \delta z \cdot .033 \cdot C_n^2 \cdot \kappa^{-\frac{11}{3}}\right]^{1/2} \,, \tag{4.44}$$

as indicated in the third step in Figure 4.8. Finally, the last step is to inverse transform the filtered arrays to yield a single time realization of a phase screen possessing the necessary spatial correlation. This is done using the inverse discrete Fourier transform

$$\Im^{-1}\{A(m',n')\} = \frac{1}{N^2 \Delta x \Delta y} \sum_{m'=0}^{N-1} \sum_{n'=0}^{N-1} A(m,n) e^{j\frac{2\pi}{N}(mm'+nn')} \,. \tag{4.45}$$

In instances where an ensemble of temporally independent screens is needed, the process outlined in Figure 4.8 is repeated using a new initial array of random numbers for each screen. However if a series of successive phase screens need to possess temporal correlation, they must be created in a different manner. In order to create the

temporal correlation, two basic operations are performed between successive phase screens. One is to introduce translation or 'drift' from one screen to another. The second is to introduce boiling from one to the other.

To introduce translation from one screen to another, the initial N X N array of random numbers that created the first screen is used for the second screen, except that its rows (or columns) are shifted. The rows or columns left blank from the shifting are then replaced by a new set of random numbers. This process is illustrated in Figure 4.8.

Introducing boiling between successive screens is as straightforward as introducing drift. Figure 4.9 shows boiling being added to a phase screen. As Figure 4.9 illustrates, boiling is introduced by replacing every n^{th} row or column in the phase screen with a new set of random numbers before applying the Kolmogorov filter. In an effort to make the boiling affect more random, instead of replacing the n^{th} row for every successive realization, the rows are replaced in a small uniformly random region about the n^{th} row. The phenomenon of boiling corresponds to a continuing evolution in the profile of the phase screen perhaps caused by thermal gradients [8].

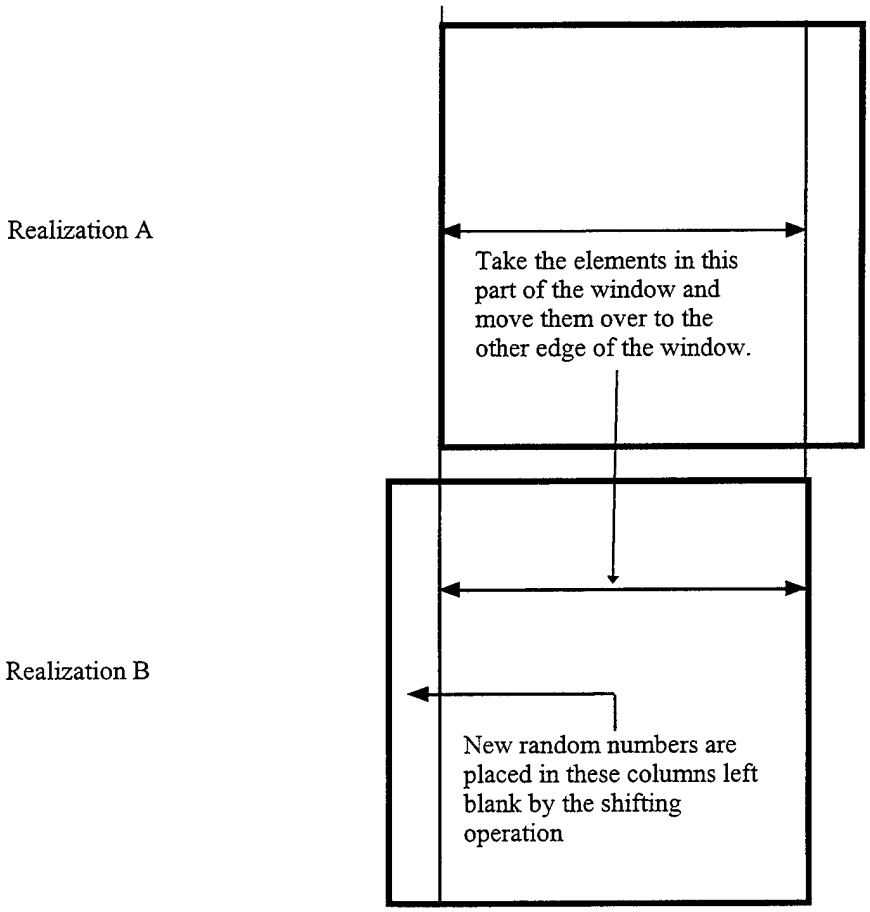
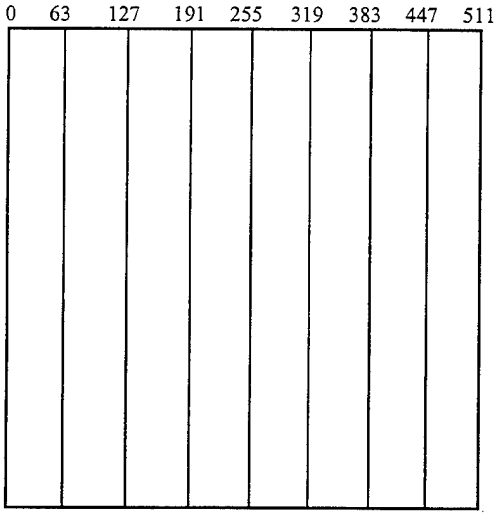


Figure 4.9: Process by which successive phase screens are made to 'translate'.

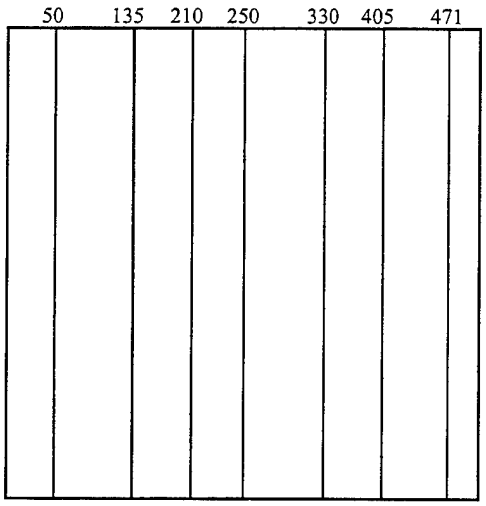
4.2.1.4 Laser

The laser model creates a Gaussian beam that propagates through the system and illuminates the target. This section briefly describes the theory of Gaussian beam propagation and gives all the relevant equations.

A simplified diagram of a laser is shown in Figure 4.10. The equations governing the size and shape of the laser beam exiting the laser are [19]



This picture illustrates every 64th column in a phase screen realization



This picture illustrates how boiling in a successive phase screen realization is made more random. Instead of replacing every 64th column with a new set of random numbers, columns are replaced in a randomized range around every 64th column.

Figure 4.10: Demonstration of boil being added between successive phase screens.

$$E(r,z) = E_0 \frac{\omega_0}{\omega(z)} e^{-jkz} e^{j\tan^{-1}\left(\frac{z}{z_0}\right)} e^{\frac{-r^2}{\omega^2(z)}} e^{-j\frac{kr^2}{2R(z)}}, \qquad (4.46)$$

where

$$\omega^2(z) = \text{square of beam radius} = \omega_0^2[1+(\frac{z}{z_0})^2], \tag{4.47}$$

$$R(z) = \text{radius of curvature} = z[1+(\frac{z_0}{z})^2], \tag{4.48}$$

$$\omega_0 = \text{beam waist} = \sqrt{\frac{\lambda z_0}{\pi}}, \tag{4.49}$$

z_0 = confocal parameter, z = distance from the beam waist, r = radial distance from the optic axis, k is the wavenumber ($2\pi/\lambda$), and E(r,z) is the electric field in the z plane at a radial distance r. Finally, the electric field from the laser is

$$E_0 = \frac{2}{\omega_0}\left[\frac{P_{laser}}{\varepsilon_0 \cdot c \cdot \pi}\right]^{1/2}, \tag{4.50}$$

where P_{laser} is the laser output power. In Figure 4.11, the beam waist is located at the output coupler, so this becomes the z = 0 plane (z-axis coincides with optic axis). For z = 0 Equation (4.47) reduces to $w(0) = w_0$ which is the beam waist, and the radius of curvature of the Gaussian beam becomes infinite. This results in the laser beam curvature matching that of the planar output coupler. In fact, all that is needed to completely describe the Gaussian beam produced by the laser is the beam waist, the beam waist location, the wavelength of the light, and the laser power. By specifying the power in the laser, the value of E_0 can be found with Equation (4.50). By specifying the beam waist and the wavelength, the confocal parameter can be found using Equation (4.49). Knowing the beam waist and the distance z from the beam waist location allows R(z) and w(z) to be determined using Equations (4.47) and (4.48). Substituting all of this information into Equation (4.46) gives the electric field at any value of z. Using Equations (4.46) - (4.50), the electric field at any value of z can be found.

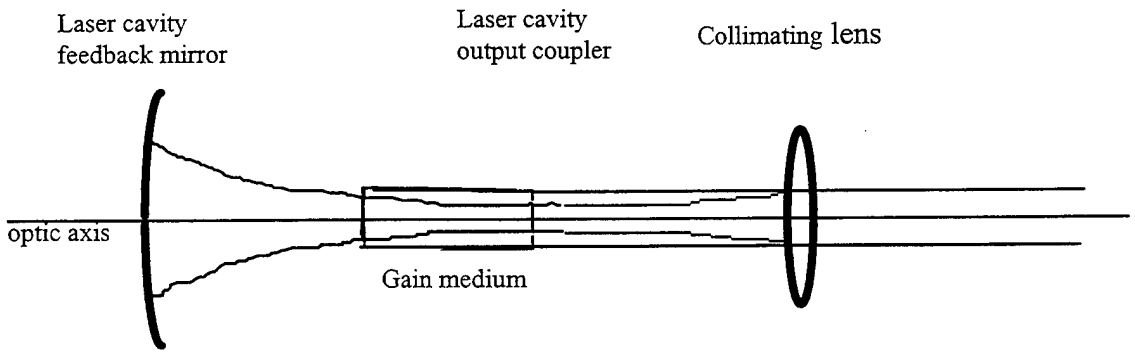


Figure 4.11: Simple diagram of a laser cavity.

The above discussion describes how to propagate a Gaussian beam through free space. But, typically the output of the laser is incident on a lens to collimate the beam. The effect of a lens on a Gaussian beam is easily modeled using ABCD matrices. The key to this analysis is the fact that given the wavelength of the light, a Gaussian beam is completely described by its q parameter [19],

$$q = z + jz_0, \tag{4.51}$$

where z is distance from the beam waist to the lens, and z_0 is the confocal parameter.

The equations for w(z) and R(z) are easily derived from the q parameter in Equation (4.51). So by calculating the effect of a thin lens and a subsequent free space propagation on the q parameter of a Gaussian beam, w(z) and R(z) can be found at any distance past the lens. The effect that ABCD matrices have on the q parameter of a Gaussian beam is given by the equation [19]

$$q' = \frac{Aq + B}{Cq + D}. \tag{4.52}$$

The ABCD matrix for a lens of focal length f is [19]

$$\begin{bmatrix} 1 & 0 \\ -\frac{1}{f} & 1 \end{bmatrix}, \tag{4.53}$$

where A=1, B=0, C=-1/f, and D=1. The ABCD matrix for a free space propagation of distance a is [19]

$$\begin{bmatrix} 1 & \alpha \\ 0 & 1 \end{bmatrix}, \tag{4..54}$$

where A = 1, B = a, C = 0, and D = 1.

Substituting the ABCD values in Equation (4.53) and the q parameter in Equation (4.51) into Equation (4.52) gives:

$$q' = z' + j \cdot z_0', \tag{4.55}$$

where

$$z' = \frac{zf^2 - z^2 f - z_0{}^2 f}{(z-f)^2 + z_0{}^2}, \tag{4.56}$$

and

$$z_0' = \frac{z_0 f^2}{(z-f)^2 + z_0{}^2}. \tag{4.57}$$

Equation (4.57) gives the new q parameter of the Gaussian beam, in which Equation (4.56) represents the location of the new beam waist relative to the lens plane, and Equation (4.55) represents the new confocal parameter. The next step in obtaining w(z) and R(z) is to propagate the q parameter in Equation (4.55) a given distance past the lens. This is done by substituting the ABCD values in Equation (4.54) and the q parameter in Equation (4.55) into Equation (4.52). The result of this calculation is

$$q'' = q' + \alpha = (z' + \alpha) + j \cdot z_0'. \tag{4.58}$$

The last step in finding w(z) and R(z) after focusing by a lens is to derive the actual equations themselves from the q parameter in Equation (4.58). This is done by substituting q'' into the equation

$$e^{-j\frac{k}{2\cdot q''}\cdot r^2} \tag{4.59}$$

Multiplying out the exponent produces the new equations for the beam radius and phase curvature as a function of propagation distance [19]

$$\frac{-jK}{2\left[(2'+2)+jz_0'\right]} = \frac{-kr^2}{2(z'+\alpha)\left\{1+\left(\frac{z_0'}{(z'+\alpha)}\right)^2\right\}} + \frac{-r^2}{\frac{\lambda z_0'}{\pi}\left\{1+\left(\frac{(z'+\alpha)}{z_0'}\right)^2\right\}} \tag{4.60}$$

Upon inspection of Equation (4.60), the denominator of the real part is of the same form as Equation (4.47) and the denominator of the imaginary part is the same as Equation (4.4`in lens are

$$\omega^2(\alpha) = \frac{\lambda z_0'}{\pi}\left\{1+\left[\frac{(z'+\alpha)}{z_0'}\right]^2\right\}, \tag{4.61}$$

and

$$R(\alpha) = (z'+\alpha)\left\{1+\left[\frac{z_0'}{(z'+\alpha)}\right]^2\right\}, \tag{4.62}$$

where a is the distance propagated past the lens. Equations (4.61) and (4.62) completely describe the Gaussian wave as it propagates through a thin lens.

4.2.1.5 Wave Plates

This section discusses the theory behind 'optical wave plates' as implemented in the laser radar simulator. Figure 4.12 shows an electromagnetic wave traveling in the +z direction with the electric field polarized in the y direction. As the light wave in Figure 4.12 propagates along the z-axis, the electric field vector can be expressed as

$$E(z,t) = \mathbf{y}\, E_0 \sin(k\cdot z - \omega t), \tag{4.63}$$

where y is the unit vector in the positive y direction and E_0 is the electric field magnitude, w is the frequency, and k is the wavenumber. The magnetic field vector can be expressed as

$$M(z,t) = -\mathbf{x}\,\frac{E_0}{c}\sin(k\cdot z - \omega t), \tag{4.64}$$

where x is the unit vector in the positive x direction and c is the speed of light in a vacuum. A standard convention in optics is to refer only to the electric field of an electromagnetic wave when specifying polarization information, and this is the convention that will be followed here. In addition, S polarization is represented by the x axis while P polarization is represented by the y axis as shown in Figure 4.12. The x unit vector will be replaced by the S unit

vector, and the y unit vector will be replaced by the P unit vector. The general expression for the polarization of light is [20]

$$\vec{P}(z,t) = \hat{S}\,E_S \sin(k\cdot z - \omega t + \theta) + \hat{P}\,E_P \sin(k\cdot z - \omega t), \tag{4.65}$$

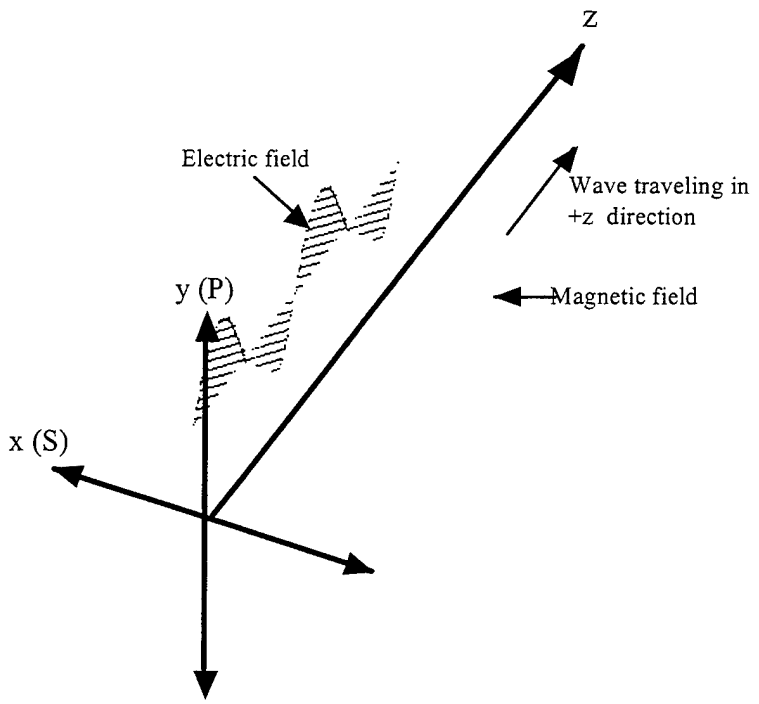


Figure 4.12: Electromagnetic wave traveling in the +z direction with the E-field polarized along the y-axis.

where E_S is the electric field amplitude along the S axis, E_P is the electric field amplitude along the P axis, q is the relative phase difference between the S and P axes, and w is the frequency of the light.

Examining Equation (4.65), it is seen that there are three different types of polarization that are possible. The first is linearly polarized light. This can occur only when q = 0 or q = p. The relative magnitudes of E_S and E_P determine the angle that the polarization vector makes with the S axis. This angle is given by

$$\Phi = \tan^{-1}\left[\frac{E_P}{E_S}\right]. \tag{4.66}$$

The second type of polarization is circularly polarized light. This is only possible when $E_S = E_P$ and $q = \pm\pi/2$. The final and most common type of polarization is elliptically polarized light. It occurs when $E_S = E_P$ and $\theta \neq 0$ or $\pi/2$ or π, but most generally when $E_S \neq E_P$ and q lies anywhere in the range $(0,\pi)\cup(\pi,2\pi)$.

An expression that gives the angle of the major axis relative to the S axis is [20]

$$\Theta = \frac{1}{2}\tan^{-1}\left(\frac{2 \cdot E_S \cdot E_P \cdot \cos\theta}{E_P^2 - E_S^2}\right). \tag{4.67}$$

Figure 4.13 shows elliptically polarized light with the major axis located along the P-axis and the minor axis located along the S axis. From the perspective of someone looking in the direction of the light propagation, the polarization vector will rotate in the counter-clockwise direction.

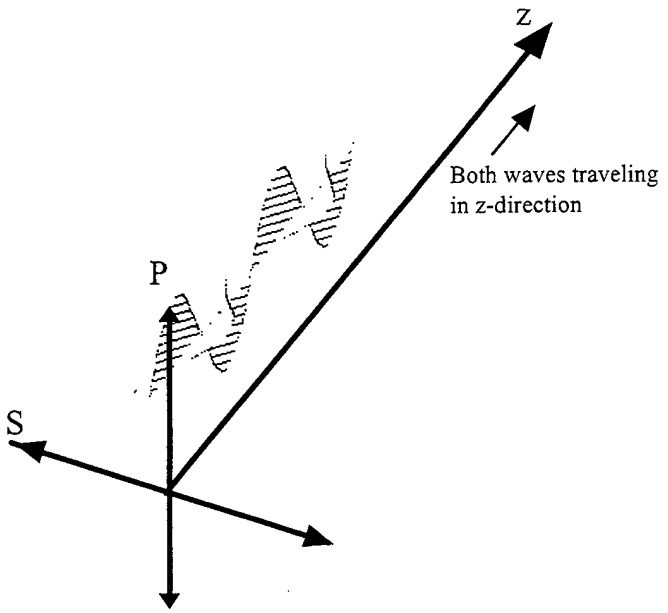


Figure 4.13: Elliptically polarized light.

The difference between linear polarization and elliptical polarization is relative phase between the two orthogonal components. This forms the basis for the wave plate operation. Wave plates are made with birefringent crystals which possess different indices of refraction for different light polarizations traveling through the material. For use as wave plates, these crystals are chosen so that only two orthogonal light polarizations are allowed to propagate through the crystal at normal incidence. By utilizing the fact that the indices of refraction are different for the two polarizations, a relative phase difference can be introduced between the two polarizations upon exiting the crystal. The crystal thickness needed to introduce a given phase difference for a given wavelength of light depends on the relative magnitudes of the orthogonal indices of refraction. The axis in the crystal that possesses the higher index of refraction is known as the **slow axis**, because the speed of light is slower in a higher index material. Conversely, the axis with the smaller index of refraction is known as the **fast axis**.

The two types of wave plates implemented in the laser radar simulator are half-wave and quarter-wave. In modeling both of types of wave plates an arbitrary angular orientation of the slow axis of the wave plate with the S polarization reference axis of the system is required. By allowing an arbitrary angular orientation of the slow axis with the S axis, linearly polarized light can be transformed into circularly polarized light by rotating the slow axis so

that it makes an angle of 45^0 with the S axis. By rotating the slow axis 45^0, the P polarized light will possess equal components along both the fast and slow axes of the wave plate. For a wave plate with a fast axis to slow axis phase difference of p / 2, this operation yields circularly polarized light.

For both half-wave and quarter-wave plate operation, the first step is to rotate the S and P coordinate system so that it coincides with the fast and slow axes of the wave plate. This is illustrated in Figure 4.14. In Figure 4.14, the wave is initially polarized equally along both the S and P axes. To transform the amplitude and phase to the coordinate system of the wave plate, the following equations are used :

$$M_{S'} = [M_S{}^2 \cdot \cos^2\theta + M_P{}^2 \cdot \cos^2(90-\theta) + 2\cos\theta \cdot \cos(90-\theta) \cdot M_S M_P \cdot \cos(\phi_s - \phi_P)]^{1/2}, \quad (4.68)$$

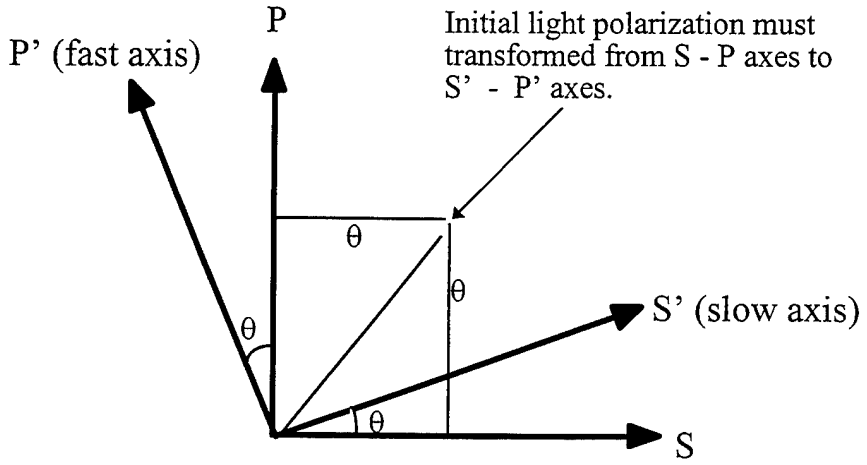


Figure 4.14: A diagram showing a wave plate rotated at an angle q relative to the S and P axes.

$$\phi_{S'} = \tan^{-1}\left[\frac{\cos\theta \cdot \sin\phi_S + \cos(90-\theta) \cdot \sin\phi_P}{\cos\theta \cdot \cos\phi_S + \cos(90-\theta) \cdot \cos\phi_P}\right], \quad (4.69)$$

$$M_{P'} = [M_S{}^2 \cdot \cos^2(90+\theta) + M_P{}^2 \cdot \cos^2(\theta) + 2\cos\theta \cdot \cos(90+\theta) \cdot M_S M_P \cdot \cos(\phi_s - \phi_P)]^{1/2}, \quad (4.70)$$

$$\phi_{P'} = \tan^{-1}\left[\frac{\cos(90+\theta) \cdot \sin\phi_S + \cos(\theta) \cdot \sin\phi_P}{\cos(90+\theta) \cdot \cos\phi_S + \cos(\theta) \cdot \cos\phi_P}\right], \quad (4.71)$$

where M_a is the magnitude of the a^{th} polarization component, f_a is the phase of the a^{th} polarization component, and q is the angle between the slow axis of the wave plate and the S polarization axis. Once these transformations have been performed, depending on whether quarter-wave plate or half-wave plate operation is desired, either $\pi/2$ or π

is added to $\phi_{S'}$ (phase of slow axis). The final step is to transform the light emerging from the wave plate (along the fast and slow axes) back to the original S and P coordinate system.

4.2.1.6 Beamsplitters

This section discusses implementation of polarizing and non-polarizing beamsplitters. The non-polarizing beamsplitter divides the incident radiation based on a power splitting coefficient, and the polarizing beamsplitter divides the incident radiation based on its polarization. The operation of a non-polarizing beamsplitter is illustrated in Figure 4.15. The non-polarizing beamsplitter is described completely by its power transmittance T used to model losses within the beamsplitter and its power splitting ratio R. The transmittance represents the fraction of the incident power that is not lost due to scattering, absorption, etc. The splitting ratio determines the portion of the incident power that passes straight through the non-polarizing beamsplitter and consequently the portion that is deflected. The fraction of the incident electric field that passes straight through the beamsplitter is

$$E_{T_P} = \sqrt{T} \cdot \sqrt{R} \cdot E_{0_P}, \tag{4.72}$$

$$E_{T_S} = \sqrt{T} \cdot \sqrt{R} \cdot E_{0_S}, \tag{4.73}$$

and the fraction that is deflected at a 90^0 angle is

$$E_{D_P} = \sqrt{T} \cdot \sqrt{1-R} \cdot E_{0_P}, \tag{4.74}$$

$$E_{D_S} = \sqrt{T} \cdot \sqrt{1-R} \cdot E_{0_S}, \tag{4.75}$$

where E_0 is the incident radiation, E_T is the transmitted radiation, and E_D is the deflected radiation. The phase information for the transmitted and deflected components remains unchanged from the incident components.

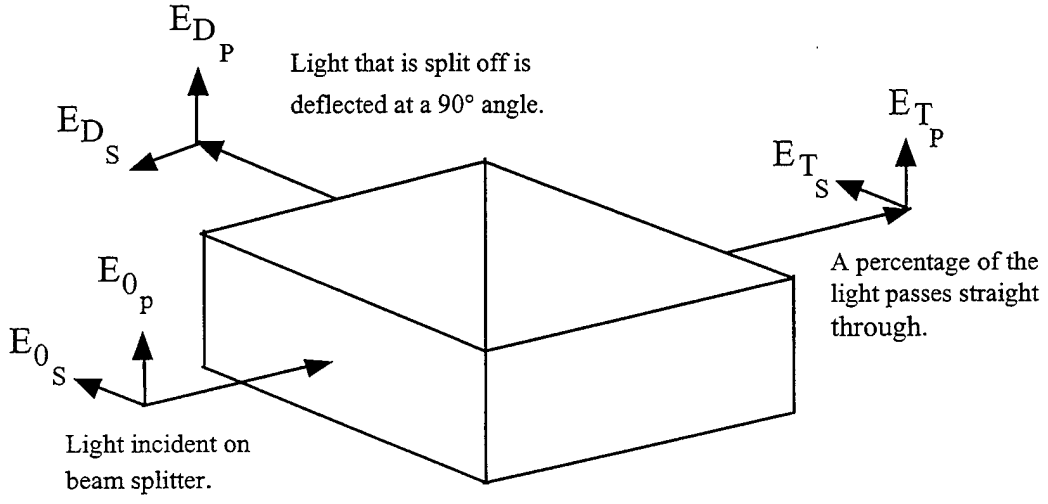


Figure 4.15: Operation of a non-polarizing beamsplitter as utilized in this laser radar simulator.

The operation of the polarizing beamsplitter is illustrated in Figure 4.16. The incident light in Figure 4.16 is illustrated with equal amplitudes of electric field along the S and P polarizations. The transmitted light shows that most of the P polarization component of the incident radiation is allowed to pass through the beamsplitter, while only a small fraction of the S component is transmitted. The deflected light shows that most of the S component of the incident radiation is deflected, while only a small percentage of the P polarization is deflected. The net effect is that a polarizing beamsplitter deflects the S component while transmitting the P component.

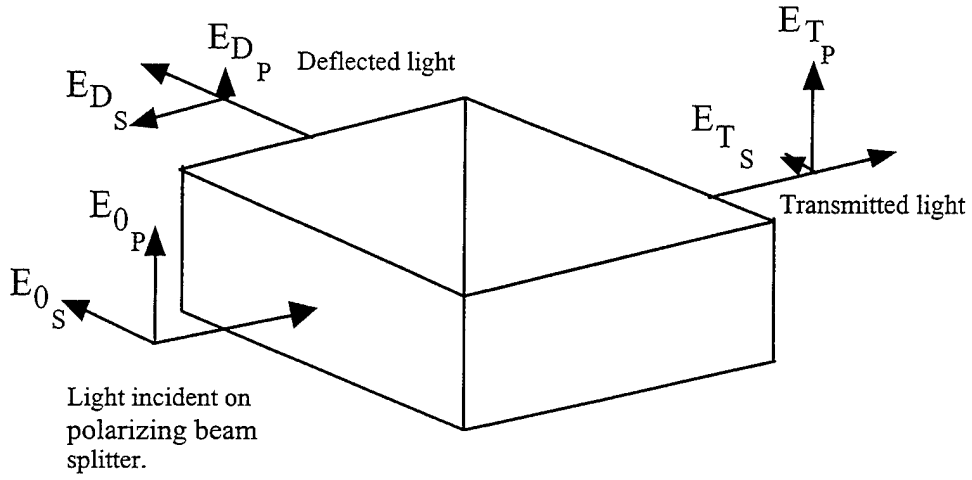


Figure 4.16: Operation of a polarizing beamsplitter.

The fact that a small percentage of the S polarization passes straight through and a small percentage of the P polarization is deflected is a consequence of the finite extinction ratios in real world polarizing beamsplitters. Allowing the transmittance of the beamsplitter to be T, the forward power extinction ratio to be F, and the side extinction ratio to be S, the electric field amplitudes of the transmitted radiation will be

$$E_{T_P} = \sqrt{T} \cdot \sqrt{1-S} \cdot E_{0_P}, \tag{4.76}$$

$$E_{T_S} = \sqrt{T} \cdot \sqrt{F} \cdot E_{0_S}, \tag{4.77}$$

and the electric fields of the deflected radiation will be

$$E_{D_P} = \sqrt{T} \cdot \sqrt{S} \cdot E_{0_P}, \tag{4.78}$$

$$E_{D_S} = \sqrt{T} \cdot \sqrt{1-F} \cdot E_{0_S}. \tag{4.79}$$

As obtained from data sheets for typical polarizing beamsplitters, a sample value of F is 1/200, and a sample value of S is 1/20. The phase information of all the transmitted and deflected radiation is left unchanged.

4.2.3. The graphical user interface (GUI) for the laser radar simulator

The laser radar system developed under this tasks models the effects that various optical devices have on the electric field of a laser beam as it travels through an x by y by z region of space, where y is the height, x is the width, and z is the length of the region of space being simulated. The height and width of this region of space is not

continuous but is modeled as an array of 512 by 512 elements. Each one of these elements contains information about the intensity and phase of the laser beam's electric field for the position of simulation space it represents. Each element has a length of z. The values of x and y are determined by multiplying the spatial sampling interval of the simulation by the number of samples in one row or column of the simulation array.

The main GUI window of the laser radar simulator contains three sections, as shown in Figure 4.17: first, the pull down menu section at the top of the window; second, an area in which optical device icons will be placed so that they form an optical schematic; and third, a box containing the current schematic file name associated with the currently displayed schematic. The pull down menu has three sections: FILE, DEVICES, and SIMULATION. The FILE pull down menu contains options to LOAD a simulation, SAVE a simulation, or QUIT the GUI. The DEVICES pull down menu consists of a list of optical devices, such as laser, detector, lens, etc. The DEVICES menu is used to place a device in the schematic area. The SIMULATION menu contains options to RUN a simulation, change a simulation's parameters, or specify an aperture that can be used to specify the location of an optical device.

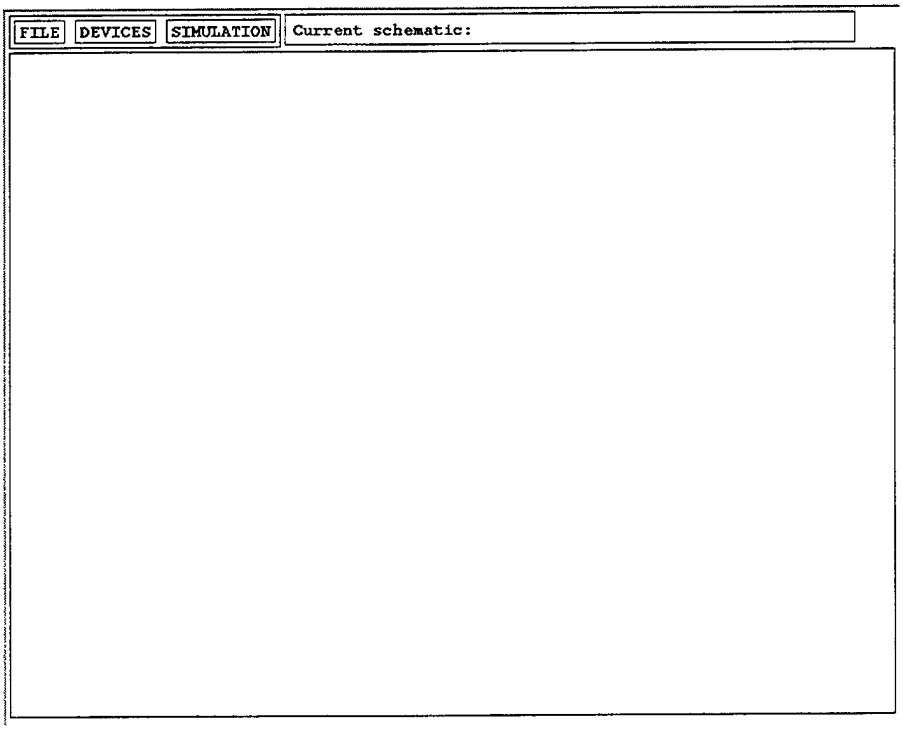


Figure 4.17: Main window for CAD system.

4.2.3.1 Menu Section

As stated above, the menu section consists of three pull down menus: FILE, DEVICES, and SIMULATION. The FILE menu has four entries: LOAD, SAVE, SAVE AS, and QUIT. When LOAD is selected a file selection window (Figure 4.18) will appear that will allow the user to select the simulation schematic file to be loaded. If a schematic is currently in the bench area, a warning will be given to the user with the choice to overwrite the current schematic with the contents of the selected file or to cancel the load request.

The SAVE option will behave in a manner similar to the LOAD option. A file selection box will appear so that the user can select the directory and file in which he or she can save the current schematic. If the user has entered a filename that currently exists then she/he is given a choice to overwrite the contents of the selected file or cancel the save operation. SAVE AS saves a file to the last file name entered by the user. If no file name has been selected then SAVE AS behaves exactly like the SAVE option.

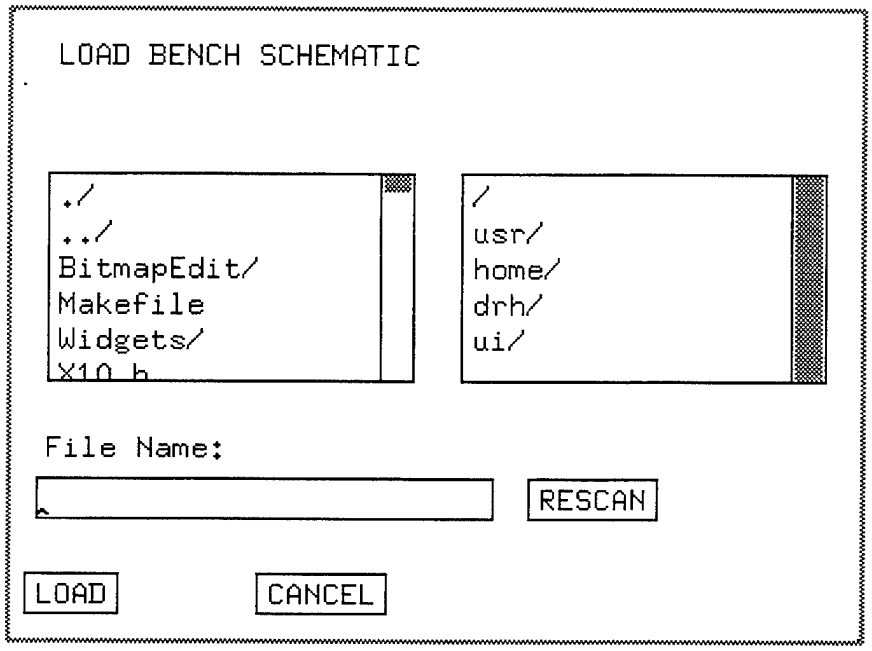


Figure 4.18: File selection window.

QUIT is used to exit the GUI. This command causes all temporary work files to be deleted and allocated memory to be returned to the system. No warning is given to the user to save an existing schematic.

4.2.3.2 Simulation Menu

The SIMULATION menu is used to setup and run a simulation. The RUN option is used to initiate a simulation. The CLEAR option is used to clear the schematic that is currently displayed. The SIMULATION PARAMETERS option produces a pop-up window (Figure 4.3) that contains the simulation's spatial sampling rate, the name of the directory used to store results, and the name of the directory that is to be used to store the intermediate results of a simulation. The APERTURES option produces a pop-up window that contains a scrolling region of aperture locations and sizes. If the user desires, a diagram of the aperture locations and sizes may be displayed using this option.

4.2.3.3 Simulation Parameters Window

The window shown in Figure 4.19 has five parameters that may be modified. The sampling rate is the spatial sampling rate, in samples per meter, of the current simulation. The number of realizations is the number of realizations that are to be performed when a simulation is run. This value is usually used only when a time varying device, such as a phase screen, is used in a simulation. "A. O. M. Frequency Shift" is the amount of shift in

frequency caused by all of the acousto-optic modulators in a simulation. The results directory is the directory that contains the results of a simulation. Such results may include the values a detector recorded or errors that occurred during a simulation. The temporary scratch pad directory is a directory that the simulation programs use for holding intermediate results. This directory is created automatically, usually in a section of the file system devoted to temporary files, such as "/tmp" or "/scratch."

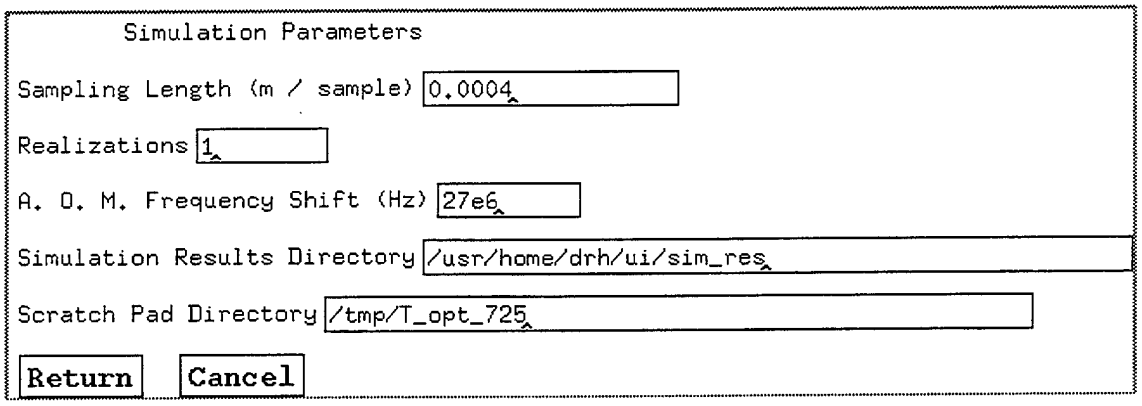


Figure 4.19: Simulation parameters window.

4.2.3.4 Aperture Specification

Many devices, such as beamsplitters and wave plates, can have apertures associated with them. Apertures are used to allow only a particular portion of the input beam to be processed by the device. An aperture is also used to specify the location of a device within the area of space undergoing simulation. This is done by creating an aperture specification that is associated with an aperture number. If no aperture is desired, the device's aperture number is set to zero. An aperture value of zero will cause the entire 512 x 512 input window to be processed by a device and will locate the device in the middle of the area of space that is being simulated.

The aperture menu item produces a window that may be used to add, modify, and delete apertures. To modify aperture information, select an aperture item from the list in the aperture information window, shown in Figure 4.20, by clicking on the appropriate item with the mouse. The selected aperture's information will appear in a window under the aperture information list. The user can then modify this information and press the update button to enter the new information into the aperture list. An aperture may be added by specifying an aperture diameter that is not zero. Likewise, it may be deleted by setting an aperture's diameter to zero. The location and size of the apertures can be viewed by pressing the "Display Apertures" button in the aperture window. This will cause a window to appear that will display the current aperture settings, as shown in Figure 4.21.

The DEVICES menu is used to add devices to the optical schematic area. When the user selects a device, an outline of a device icon appears in the bench area. The user can then move the icon outline to the desired position on the bench. When the user clicks the left mouse button an icon representing the selected device will then appear in the schematic area. The device may be moved to any other area on the bench by placing the mouse pointer over the device, depressing the left mouse button, and dragging the device icon to the desired location.

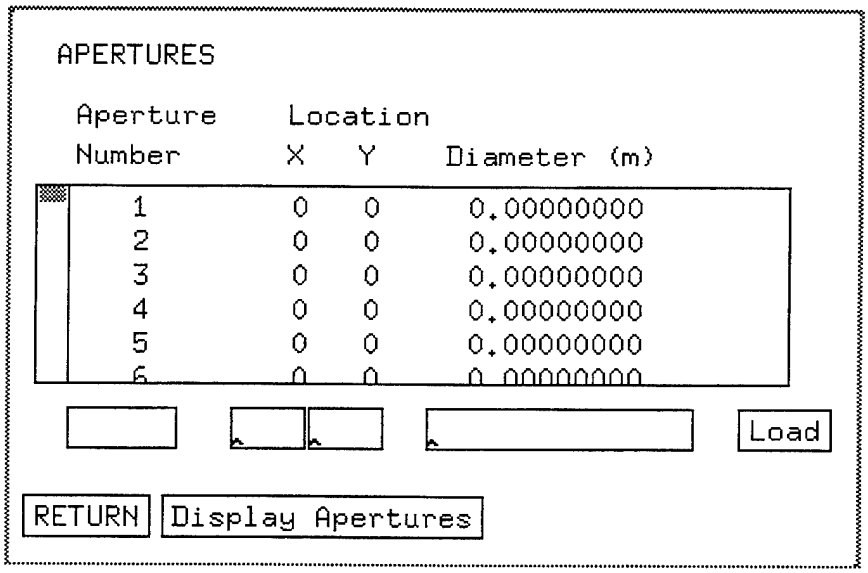


Figure 4.20: Aperture maintenance window.

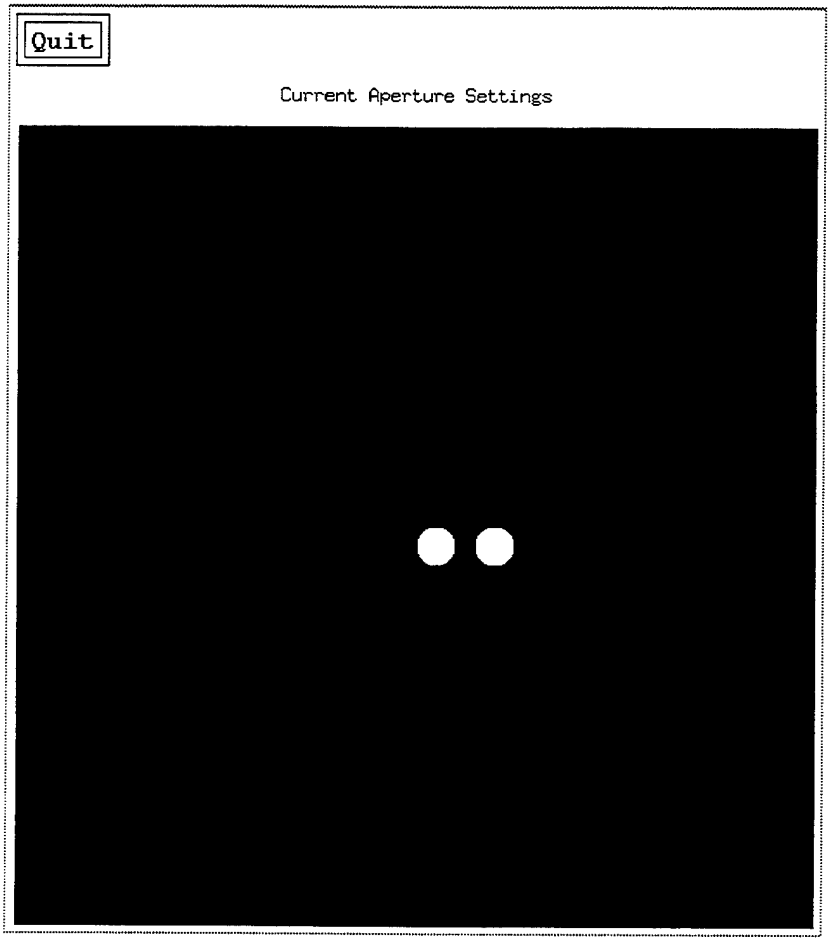
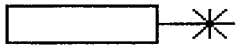


Figure 4.21: Aperture display window.

A device can be deleted by positioning the mouse cursor over the device icon and pressing the right mouse button twice while holding down the shift key. Modifications to the devices characteristics can be performed by pressing the right mouse button twice while the mouse cursor is positioned over the device icon. This will cause a window to be displayed that contains the device's current characteristics. Once the user has modified the device's parameters to the desired values they can be stored by pressing the "Return" button. If the original values are desired pressing the "Cancel" button will set the device's characteristics back to their original values.

4.2.3.5 Laser

The icon for the laser simulation function is shown in Figure 4.22. Figure 4.23 shows the parameter window that is associated with the laser icon. The laser simulation function emits a beam with a Gaussian intensity profile that is linearly polarized and of constant power (i.e., a CW laser). No more than one laser is allowed to provide input to a system. Four characteristics can be set with the laser: wave length in micrometers, power in watts, the angle of polarization with respect to the P axis in degrees, and the laser beam spot size at the beam waist in micrometers located at the laser's output.



Laser

Figure 4.22: Laser function icon.

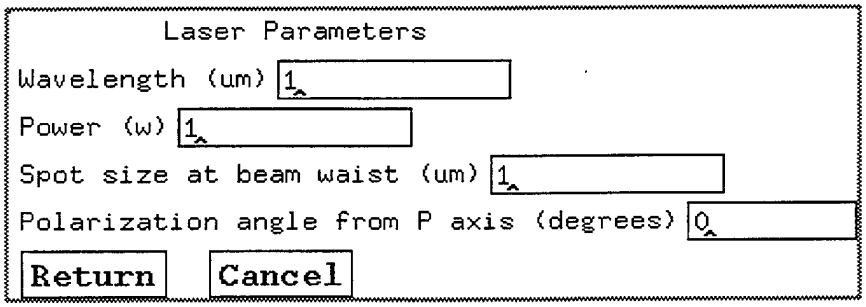


Figure 4.23: Laser parameters window.

4.2.3.6 Acousto-optic Modulator

An acousto-optic modulator sets a flag during a simulation to indicate that the frequency of the beam that passes through the acousto-optic modulator has been shifted. This frequency shift is only used in the detector model. The amount of the shift in frequency is set in the simulation parameters pop up window and applies to every acousto-optic modulator in a simulation. More than one acousto-optic modulator is allowed in a path between a laser and a detector; however, this will not shift the frequency of the beam twice since all the A. O. M. block does is set a flag indicating that a beam has passed through an acousto-optic modulator. The icon that represent an acousto-optic modulator is shown in Figure 4.24.



Figure 4.24: Acousto-optic modulator icon.

4.2.3.7 Beamsplitter (polarizing)

The icon representing a polarizing beamsplitter is shown in Figure 4.25. Polarizing beamsplitters allow only one polarization component of the incident beam to be transmitted while the orthogonal polarization component is deflected. These beamsplitters have four parameters that must be set, as shown in Figure 4.26. Percent transmittance is the percent of incident power that is not lost to scattering caused by defects and impurities in the beamsplitter. The aperture number is the aperture associated with the beamsplitter as defined in the aperture specification window.

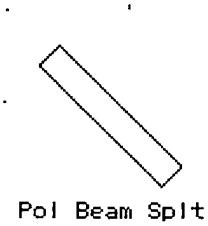


Figure 4.25: Polarizing beamsplitter icon.

The extinction ratio of a polarizer is defined as follows. Two polarizers are aligned so that the maximum amount of incident light is transmitted. The percentage of transmitted light is then measured and recorded. Next, one of the polarizers is rotated 90_ so that the minimum amount of incident light is transmitted. The percentage of transmitted light is then recorded.

The forward extinction ratio is this value for the transmitted portion of the S polarization component of the beam. The side extinction ratio is the extinction ratio associated with the P polarization component of the deflected beam. In other words a small forward extinction ratio will cause almost all of the S component of the incident beam to be deflected and a very small portion of the S component to be transmitted. A large forward extinction ratio will cause the apposite condition to occur.

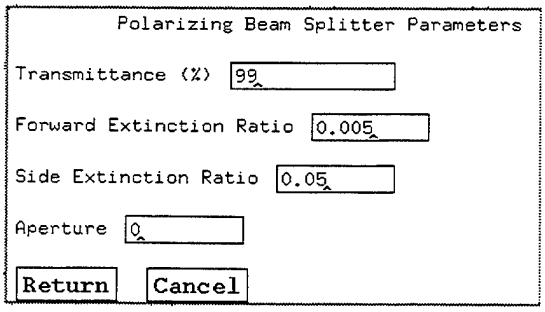
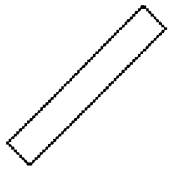


Figure 4.26: Polarizing beamsplitter parameter window.

4.2.3.8 Beamsplitter (non-polarizing)

The beamsplitter function has three parameters (Figure 4.27): percent of power deflected from the incident beam, percent transmittance, and an aperture number. If no aperture is to be used an aperture number of zero should be entered. The "Percent of power deflected from incident beam" parameter indicates the percent of power that is deflected from the incident beam. If the intensity of a beam entering a beamsplitter is 1 Watt / cm^2 and the percent of power deflected from the incident beam is 60% then the intensity of the transmitted beam will be 0.4 Watts / cm^2 and the intensity of the deflected beam would be 0.6 Watts / cm^2. Percent transmittance is the percent of incident power that is not lost to refraction and scattering caused by defects and impurities in the beamsplitter. The beamsplitter icon is shown in Figure 4.28.



Beam Splitter

Figure 4.27: Beamsplitter icon.

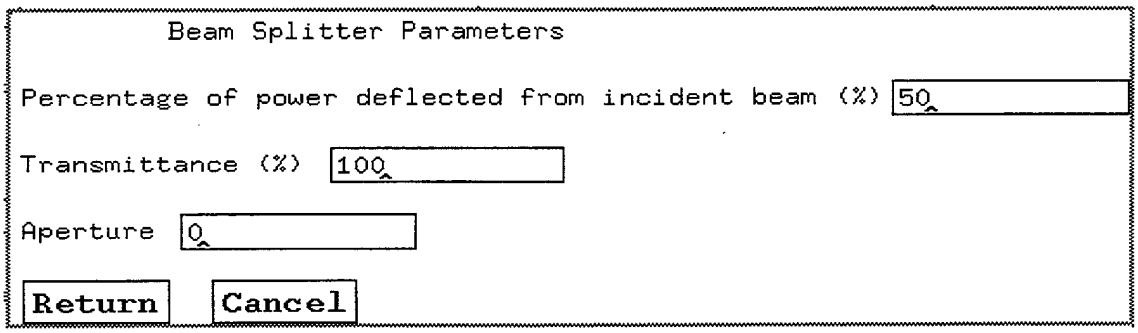


Figure 4.28: Beamsplitter parameters window.

4.2.3.9 Detector

A detector simulation function consists of two devices, as illustrated by the detector icon in Figure 4.29, a focusing lens and the actual detector (a photodiode). The parameters used for this model are shown in Figure 4.30.



Detector

Figure 4.29: Detector icon.

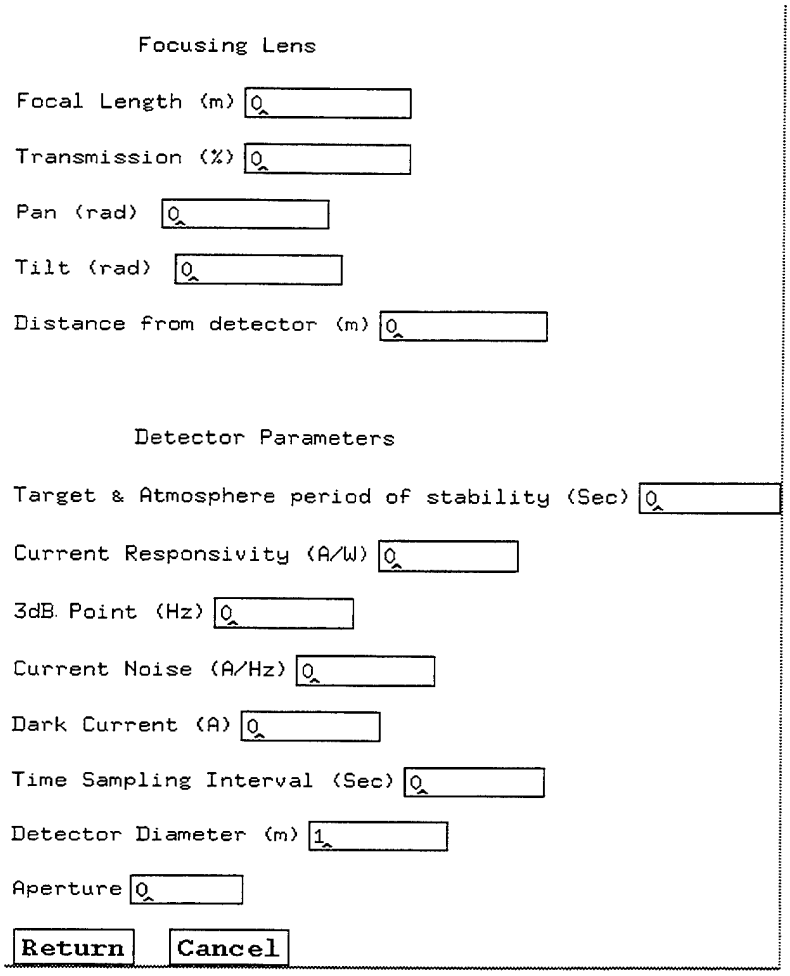


Figure 4.30: Detector parameters.

The "Target and atmosphere period of stability" is the period of time, during one realization, that the phase screen's characteristics and target's characteristics remain constant. The "Time sampling interval" is the temporal sampling interval of the detector. If the "Target and atmosphere period of stability" is one second and the "Time sampling interval" is 0.1 seconds / sample then ten output values will be recorded in the detector's output value file before the characteristics of a phase screen or a target will change. Unless two different frequencies of light are incident upon the detector the detector's output will not vary during a realization of a simulation.

The "Current responsivity" is the ratio of the RMS current out of the detector to the RMS power incident upon the detector [21]. Because the detector has a finite frequency response, a low pass filter is used to simulate the detector's finite bandwidth. The 3 dB point is the half power point of this low pass filter. "Detector Diameter" is simply the diameter of the detector's surface. The "Dark current" is the DC current present on the output of the detector with no illumination present [22]. The "Current noise" is the Shot noise present in the detector, as shown in Figure 4.30.

4.2.3.10 Lens

Figure 4.31 shows the icon for the lens simulation function. The lens function has five parameters that may be set, as shown in the lens parameters window in Figure 4.32. The focal length is the distance from the lens at which a collimated beam will converge to a point. Transmission is the fraction of incident power that is transmitted by the lens. The pan and tilt parameters only apply to a lens that is processing a non-Gaussian beam (i. e., a beam that has passed through a phase screen or has been reflected from a target).

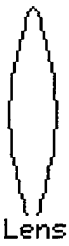


Figure 4.31: Lens icon.

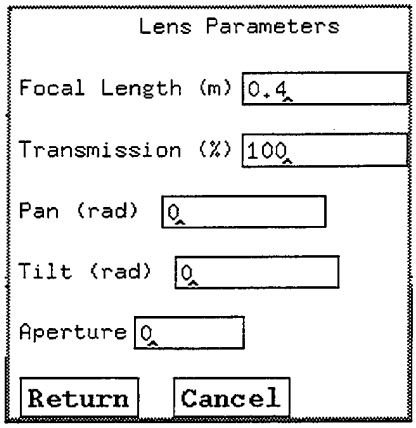


Figure 4.32: Lens parameters.

As shown in Figure 4.33, the pan parameter is the angle between the plane of the lens and the x axis. The tilt angle is the angle between the plane of the lens and the y axis. The aperture is the aperture associated with the lens as defined in the aperture specification window. If no aperture is to be used an aperture number of zero should be entered.

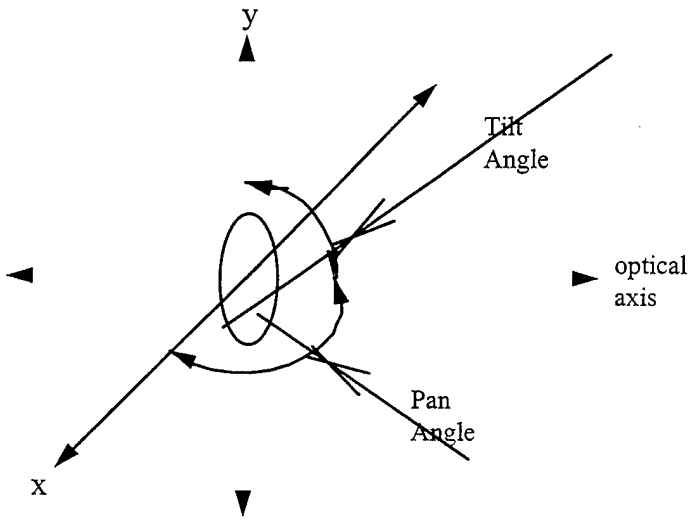


Figure 4.33: Illustration of pan and tilt parameters of lens.

4.2.3.11 Mirror

No parameters have to be entered for the mirror. This routine simulates a mirror with perfect reflectance. The icon for the mirror is shown in Figure 4.34.

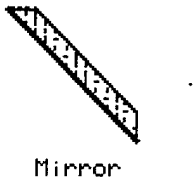


Figure 4.34: Mirror icon.

4.2.3.12 Phase Screen

The phase screen function simulates a phase screen with a Kolmogorov power spectrum as described in chapter two. Figure 4.35 shows the icon used to represent a phase screen. The "Average spacing of columns replaced for boiling" parameter shown in Figure 4.36 specifies the average number of columns in a phase screen that are to be replaced to simulate boiling for each temporal realization. Boiling is the term used to describe the phases perturbations in a beam of light passing through a portion of the atmosphere that contains areas of random density due to different temperatures. Therefore, the columns that are replaced are randomly chosen about a small neighborhood around a multiple of the value entered by the user.

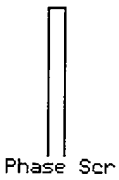


Figure 4.35: Phase screen icon.

"Number of columns shifted for translation" specifies how many columns are to be shifted in a phase screen, for each temporal realization, to simulate the effects of wind . If a translation value of S is specified by the user then S columns are removed from the right side of the phase screen before each temporal realization. The remaining phase screen values are then shifted to the right and S new columns are placed on the left side of the phase screen.

The "Structure constant" is the index of refraction structure constant and can be thought of as the strength of the phase screen's turbulence. The "Spectral index" is the rate of decrease in attenuation, from low spatial frequencies to high spatial frequencies, in the Kolmogorov power spectrum that the phase screen models. "Thickness" is the phase screen's thickness in meters.

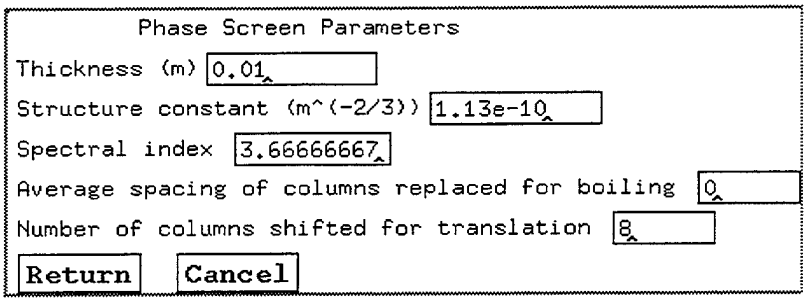


Figure 4.36: Phase screen parameters.

4.2.3.13 Point Detector

This detector, whose icon is shown in Figure 4.37, will record the intensity of the electric field incident upon each aperture that is less than or equal to the spatial sampling interval. The only parameter for this device is the current responsivity as shown in Figure 4.38.



Figure 4.37: Point detector icon.

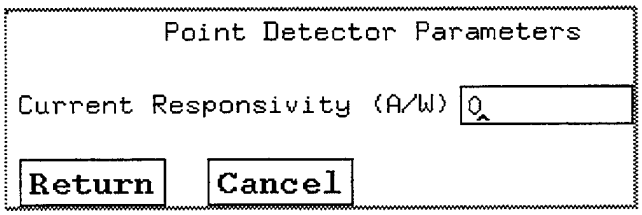


Figure 4.38: Point detector parameter window.

4.2.3.14 Signal Splitter

This icon's name is somewhat misleading. No signal is actually split. A more appropriate name for this icon would be "Simulation Space Splitter". A signal splitter icon, shown in Figure 4.39, is used when a simulation needs to be traced through two different paths within the same space being simulated. An example of this is shown in Figure 4.40. Light from a laser is reflected from a target, and the reflected light intensity is measured by two detectors at different locations. The signal splitter creates two separate propagation lines from the target icon to the detector icons. Without a signal splitter only one propagation line could be drawn from the target to a detector. If only one line could be drawn to a detector then the user would have to run a simulation with a detector in one position, save their results, specify a new detector position and then run another simulation. With a signal splitter only one simulation would have to be run. If more than two paths are required one signal splitter can be used to split the output of another signal splitter.

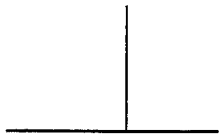


Figure 4.39: Signal splitter icon.

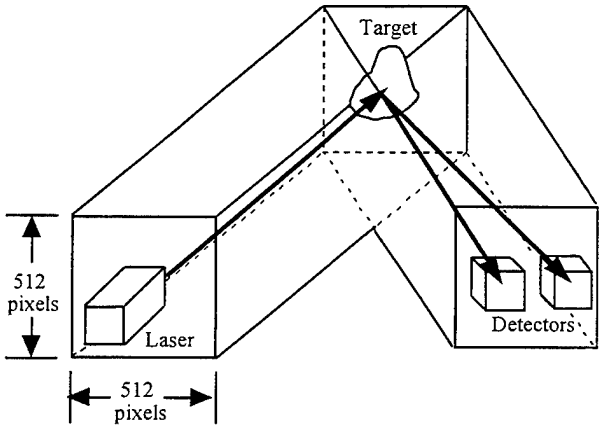


Figure 4.40: Diagram of signal splitter function.

4.2.3.15 Target

The target function adds an uncorrelated, uniformly distributed, random phase, between x and $-x$, where x is the maximum phase change specified by the user, to the optical signal it receives. The graphical icon for a taget is Given in Figure 4.41is The maximum phase change is specified by the "Maximum phase change parameter" as shown in Figure 4.42. The target's size is specified by the "Target height and width parameter." If an invalid target size is entered for the current spatial sampling interval an error message will be displayed when the user attempts to close the target parameters window by pressing the "Return" button.



Figure 4.41: Icon for target function.

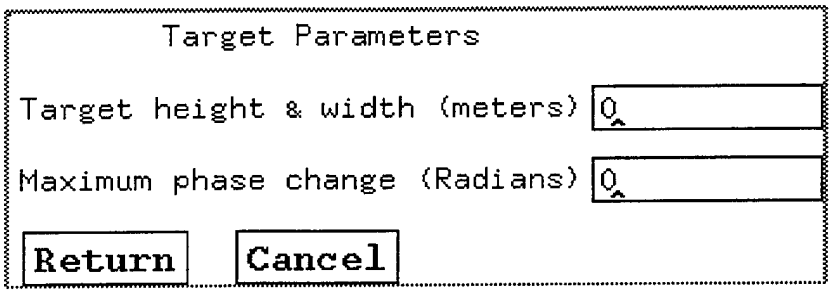


Figure 4.42: Target parameter window.

4.2.3.16 Wave Plate

The icons for the quarter and half-wave plates are shown in Figure 4.43. A wave plate separates the electric field of an incident beam of light into two orthogonal components. One of these components undergoes a phase retardation of 1/4 of a wavelength for a quarter-wave plate or 1/2 a wavelength for a half-wave plate relative to the other orthogonal component. This phase change is due to the fact that the index of refraction along one orthogonal component of a wave plate (the slow axis) is higher than the index of refection along the other orthogonal component.

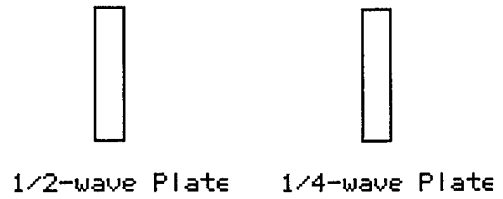


Figure 4.43: Icons for half-wave plate and quarter-wave plate.

Three parameters are required in the wave plate parameter window in Figure 4.44. The rotational angle parameter of this function specifies the angle between the wave plate's slow axis and the S polarization axis. Transmission is the percent of incident power that is transmitted through the wave plate. Aperture number is the reference number of the aperture the wave plate is to be associated with. If no aperture is to be used, a value of zero should be entered.

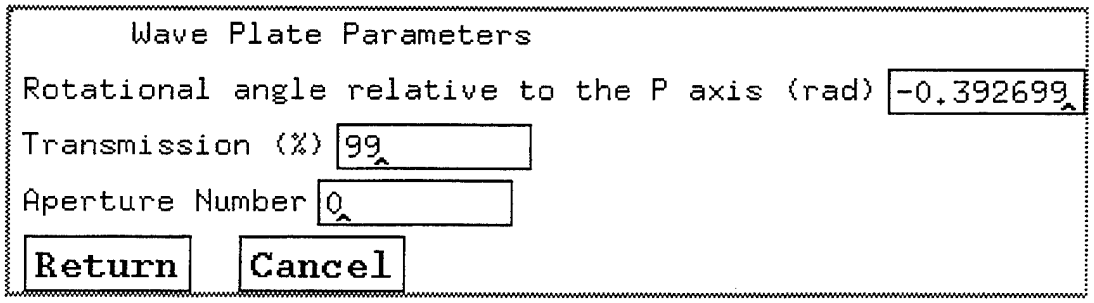


Figure 4.44: Wave plate parameter window.

4.2.3.17 Schematic Section

The optical schematic section of the GUI contains the schematic for the current simulation. Once a device is placed on the bench by using the DEVICES menu it can be connected to another device, or the device's characteristics can be established. To connect one device to another the user clicks on the source device and then on the destination device with the right mouse button. If the user specified an allowable connection an arrow will appear on the screen in the direction of specified propagation. If a connection can not be established, an error message is displayed in a pop up window. An example of how devices are connected is given in Figures 4.45 and 4.46. Figure 4.45 shows a laser connected to a lens. However, it is desired that the laser beam be focused on the target. By clicking on the lens icon and then clicking on the target icon with the right mouse button a propagation line is drawn from the lens to the target, as shown in Figure 4.46. Once the user has placed a device icon in the desired location, the device may be rotated by pressing the center mouse button or by pressing the "r" key on the keyboard.

To alter a device's characteristics, the user double clicks on the device's icon with the right mouse button. A window containing the device's current characteristics will appear. After the user changes the device's parameters, he/she can save them by pressing the "Return" button or cancel the changes made by pressing the "Cancel" button. If invalid parameters have been entered, an error message will appear informing the user of his/her error after the "Return" button has been pressed.

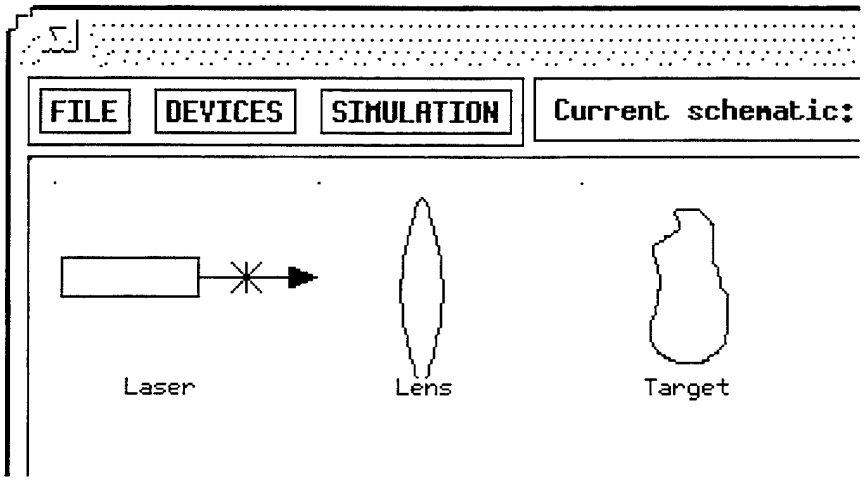


Figure 4.45: Lens and target not connected.

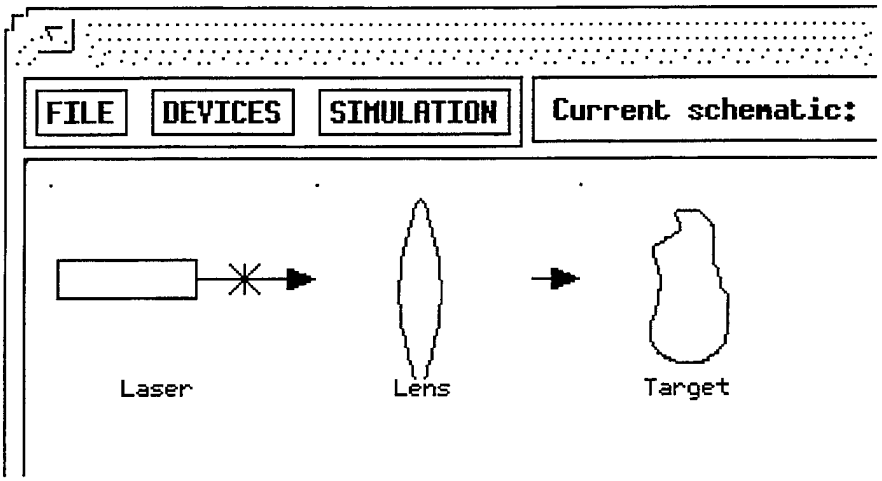


Figure 4.46: Lens and target connected.

4.2.3.18 Propagation Line Characteristics

To modify the propagation distance between two devices, the user should double click on the propagation line with the right mouse button. The propagation distance window, shown in Figure 4.47, will appear. The new propagation distance can then be entered in the "Propagation Distance" value area in meters.

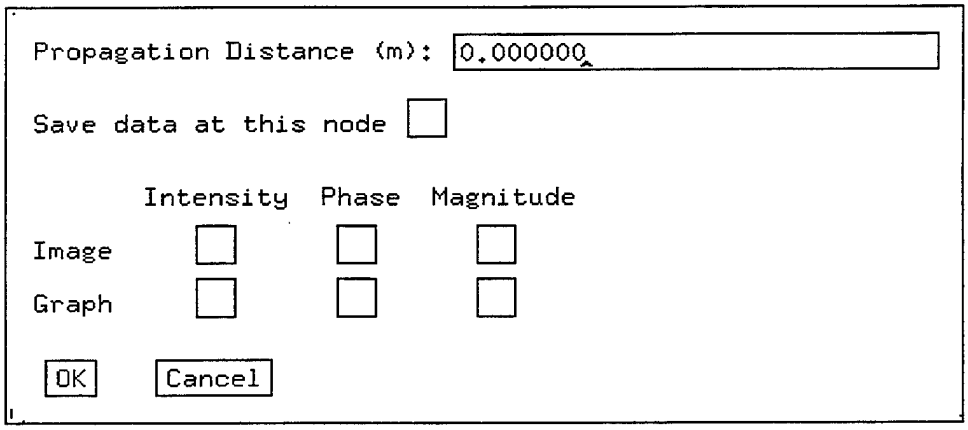


Figure 4.47: Propagation line pop up window.

If the user wishes to examine the data produced by a device, he/she should set the "Save data at this node" toggle switch. This will cause the electric field information of the propagated beam to be saved in a temporary file when a simulation is run. After a simulation is performed, the user can then produce images and graphs of the propagated electric field's phase or intensity.

4.2.3.19 Displaying Data

To display or save image information the user should press the image button under the Intensity, Phase or Magnitude prompts in the propagation line characteristics window shown in Figure 4.47. An image information window (Figure 4.48) will appear that will allow either an image to be displayed or raw image data to be saved.

The "Original S" and "Original P" prompts are used to indicate that the user wishes to examine data about the S and/or P components of a signal that has not passed through an acousto-optic modulator (i.e., a signal at the laser's original frequency). The "Frequency shifted S" and "Frequency shifted P" prompts are used to select the S and P components of a signal whose frequency has been shifted by an acousto-optic modulator. Once the user has selected the desired electric field components the image of these components is displayed by pressing the "Display Image" button. If the user wishes to save raw data the "Save Image Data" button should be pressed. Selecting this function will cause a file selection box to appear, so that the user can specify the file name that should be used to save his or her information. The data is stored in binary format using complex number format with the first 262144 values being real numbers and the next 262,144 values being imaginary. All values are four byte floating point.

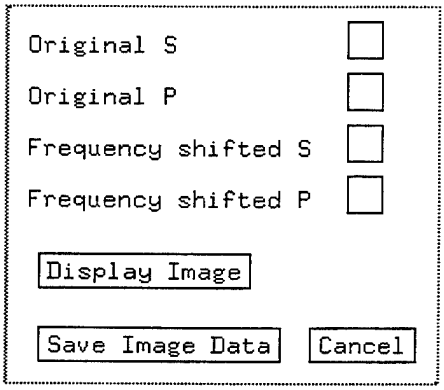


Figure 4.48: Image information pop up window.

If the user wishes to plot intensity, phase, or magnitude information for a row or column, the appropriate button should be pressed by the "Graph" prompt of the propagation line pop up window. A window will then appear (Figure 4.49) in which the user can specify the type of graph he / she would like. The row / column number can be any value from 0 to 511 inclusive. A row or column can be selected by pressing the row / column button. Pressing this button when it is labeled "Column" will cause the label to change to "Row", pressing again will change to label back to "Column." Once the user has specified the appropriate information, the "Plot" button can be pressed to display a graph of the requested data.

For example, if a user wanted to see an intensity plot of the S and P components of the laser beam at a point just before the detector, as shown in Figure 4.50, the user would double click on the propagation line. This will cause a propagation line information window to be displayed, as shown in Figure 4.51. Next, the user would set the toggle switch to save the data at this node, as shown in Figure 4.51. Once a simulation has been run, the user would again double click on the propagation line to display the propagation line information window, but this time the user would press the graph button under the intensity prompt. This will cause a graph information window to appear, such as in Figure 4.52. When the user has specified the information that they want to see, the "Plot" button can be pressed causing a graph of the requested information to be displayed.

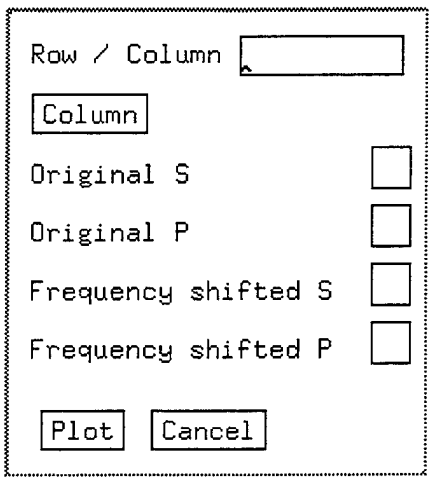


Figure 4.49: Plot data pop up window.

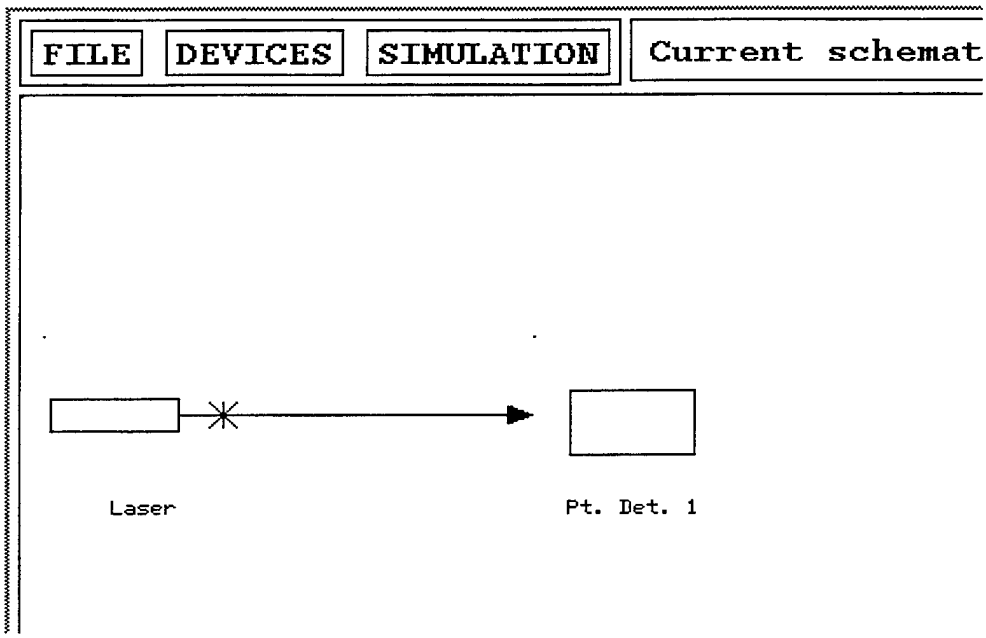


Figure 4.50: Laser incident upon a detector.

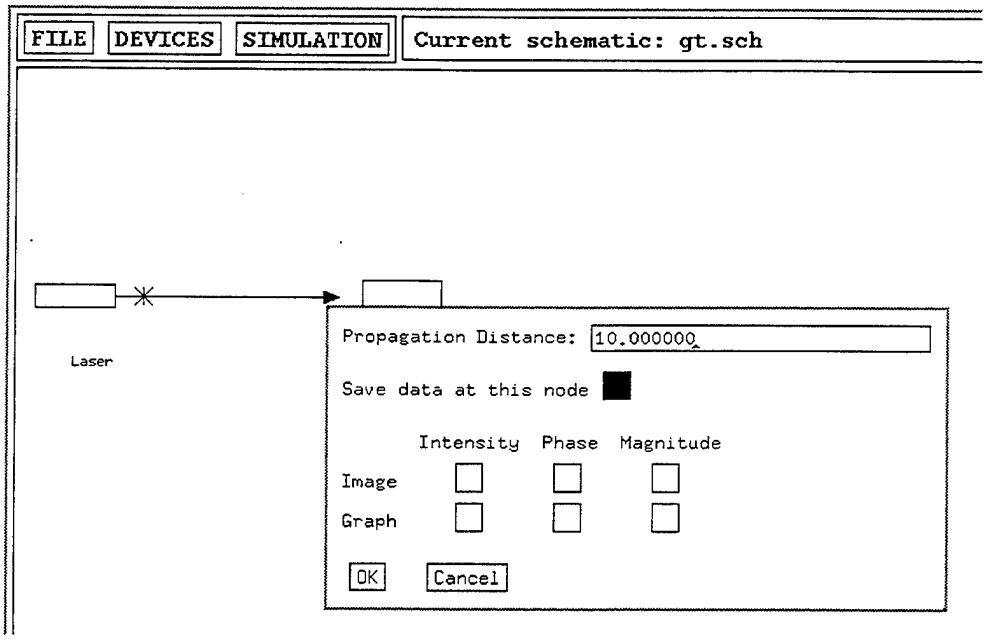


Figure 4.51: Propagation line information window for line between the laser and detector.

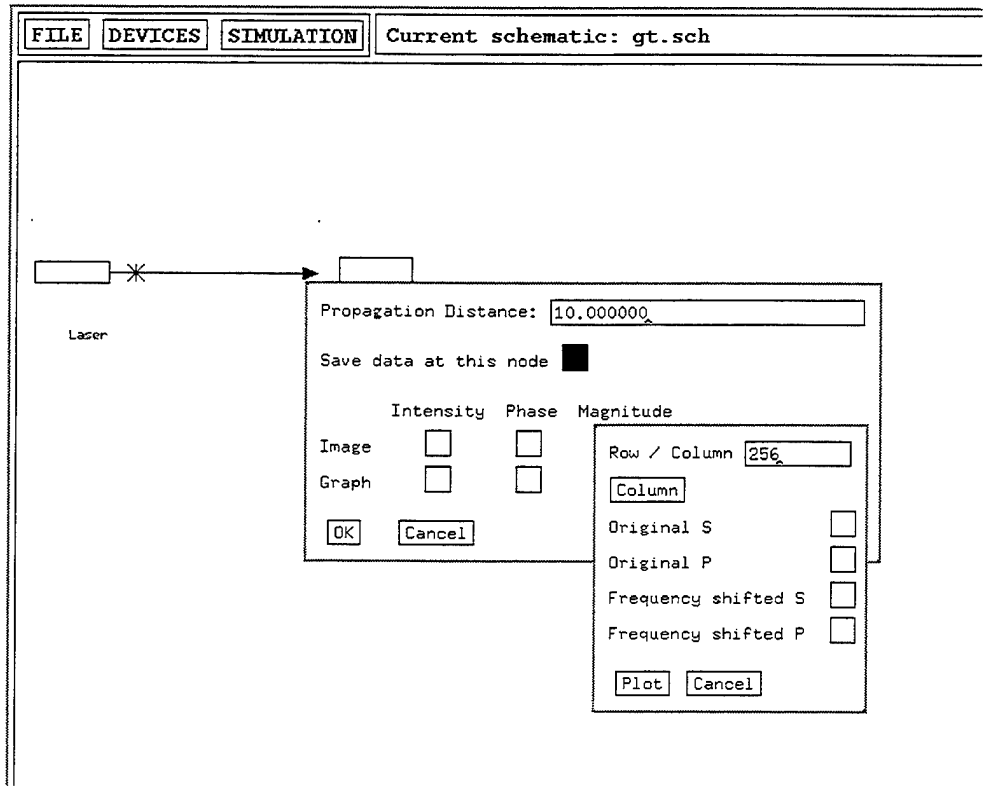


Figure 4.52: Graph information window.

4.2.3.20 Example Simulation of a Monostatic Laser Radar

The following example illustrates how to simulate a monostatic laser radar system. The first step is to clear the schematic window of the GUI. This is done by choosing the "clear schematic" option under the SIMULATION menu. A laser icon is to be placed in the schematic area. To do this the user presses the right mouse button while the mouse pointer is positioned over the DEVICES menu button. Next, while still keeping the mouse button pressed, the user moves the pointer to the "Laser" text area and releases the mouse button. An outline of a icon box will appear as shown in Figure 4.53. The user then moves the icon outline to the desired position and presses the mouse button. This will cause the icon outline to be replaced by the laser icon as shown in Figure 4.54. By pressing the center mouse button twice, the laser icon is rotated 180^{o} so that it points in the correct direction, as shown in Figure 4.55.

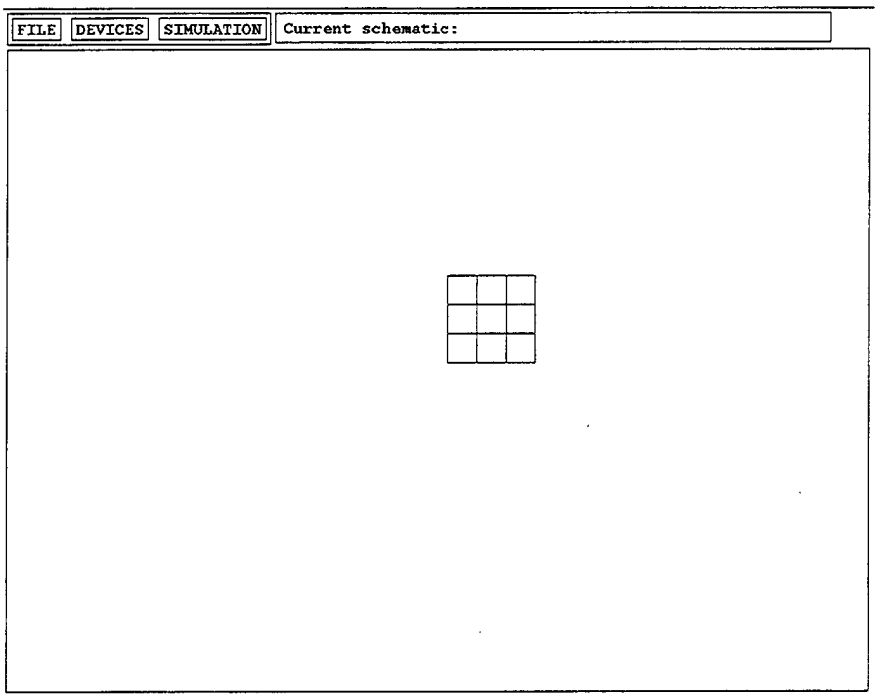


Figure 4.53: Icon outline displayed in schematic area prior to placement of device.

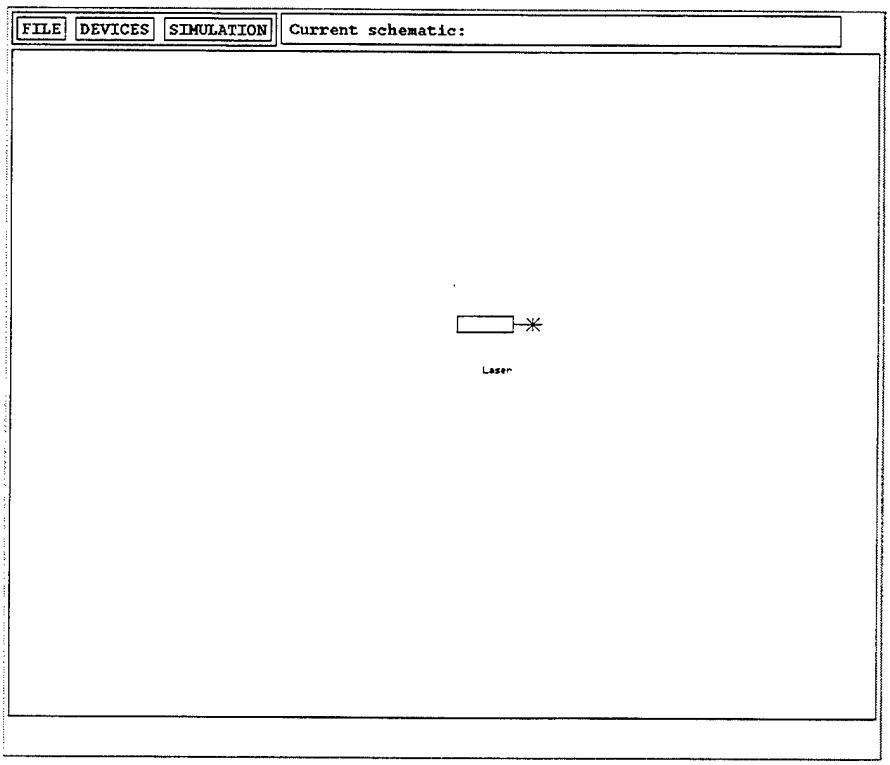


Figure 4.54: Laser icon displayed in place if icon outline.

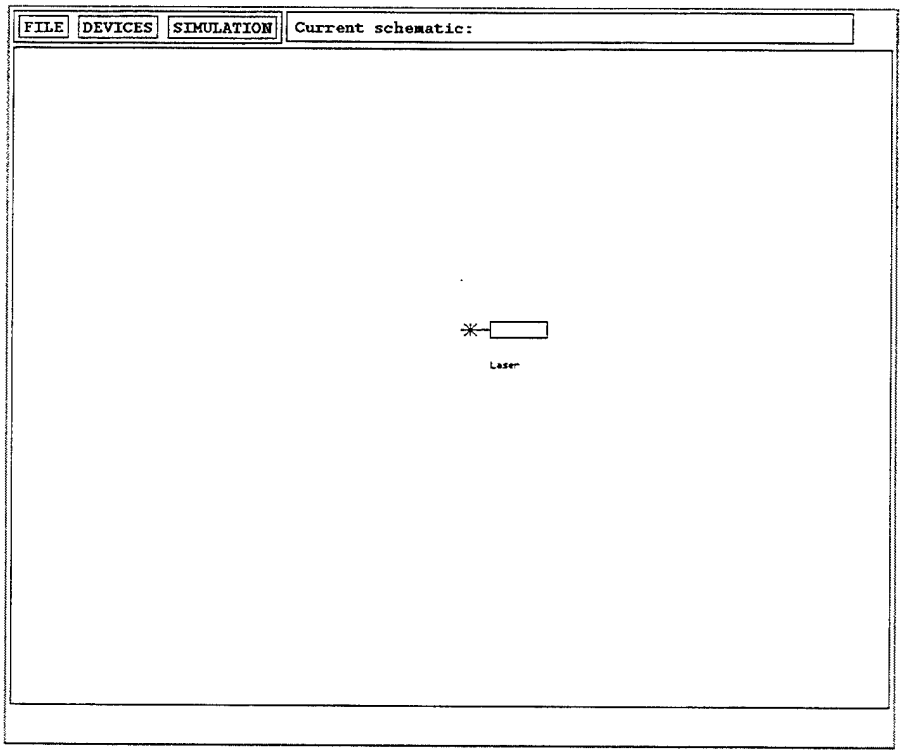


Figure 4.55: Laser icon after being rotated 180°.

Table 4.1: Monostatic laser radar simulation parameters.

Simulation Parameters	
Spatial sampling interval	0.0004 meters / sample
Aperture 1	Location : x=256, y=256. Diameter = 0.01 m
Laser Parameters	
Laser power	0.1 mW
Wavelength	1 um
Waist size	67.5 um
Polarization angle from P axis	0_
Half-wave Plate Parameters	
Rotational angle	$-\pi/8$ radians
Percent transmission	99 %
Polarizing Beamsplitter Parameters	
Forward extinction ratio	1/200
Side extinction ratio	1/20
Percent transmission	99 %
Quarter-wave Plate Parameters	
Rotational angle	$-\pi/4$ radians
Percent transmission	99 %
Phase Screen Parameters	
Structure constant	$1.13 \cdot 10^{-10}$ $m^{-2/3}$
Thickness	0.01 meters
Spectral index	-11/3
Target Parameters	
Maximum phase change	0.98 radians
Height and width	0.01 meters
Detector Focusing Lens Parameters	
Focal length	0.07 meters
Distance from detector	0.07 meters
Pan angle	0 radians
Tilt angle	0 radians
Percent transmission	99 %
Detector Photodiode Parameters	
Current responsivity	0.7 A/W
3dB point	27 MHz
Current noise	0.2 pA / $\sqrt{Hz}$
Dark current	0.5 nA
Diameter	150 um

Once the laser icon is placed in the desired position, its parameters have to be set. Placing the mouse pointer over the laser icon and double clicking will cause the laser parameters pop up window to appear. The laser's wavelength is set to 1 micrometer, its power equals 0.5 milliwatts, the beam waist spot size is set to 67.5 micrometers, and the angle of polarization from the P axis is set to zero degrees. In addition, it should be noted that the laser beam starts at the laser's beam waist. Next, collimating lenses are added as shown in Figure 4.56. The propagation distances are given in Figure 4.62. All parameters for the devices given in this example are shown in Table 4.1. This table contains the parameters for each device used in this simulation. To set the device parameters, the user should position the mouse cursor over the desired device icon and press the right mouse button twice. Pressing the "Return" button in the device parameter window will cause the new parameters to be saved. Pressing "Cancel" will set the parameters back to their original values.

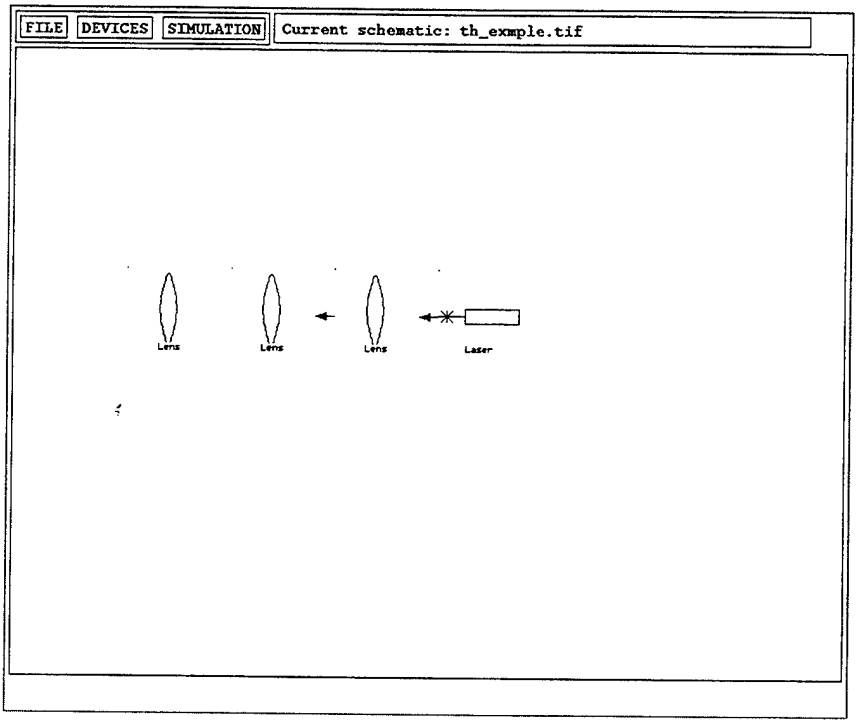


Figure 4.56: Collimating lenses added after laser.

Because the schematic area is too small to accommodate all icons two mirrors are added after the collimating lenses. A half-wave plate is then added after the mirrors as shown in Figure 4.57. The half-wave plate's slow axis is at an angle of 22.5_ with the S axis so that the laser light is linearly polarized at a 45_ angle with respect to the S and P axes. A polarizing beamsplitter is then inserted after the half-wave plate so that part of the laser light is diverted to act as a local oscillator signal. Following this an acousto-optic modulator is added in Figure 4.58 so that the returned signal's frequency can be used to determine if the target is moving away or toward the ladar.

Next, a polarizing beamsplitter - quarter-wave plate combination, shown in Figure 4.59, is added to act as an optical switch that will transmit the target signal while deflecting the returned signal toward the ladar's detector. The quarter-wave plate then converts the linearly polarized light to circularly polarized light. The returned signal is then passed through the quarter-wave plate again, where the polarization is change from circular to linear along the

S axis. Because the returned signal is now polarized along the S axis the polarizing beamsplitter will the deflect the returned signal toward the detector.

Beside having to set the beamsplitter and quarter-wave plate characteristics, an aperture must be assigned to each of these devices. To do this the user selects the "Apertures" option under the "Simulation" menu. This will cause an aperture specification window to appear. Because this is a simulation of a monostatic laser radar only one aperture needs to be specified. This aperture is to be located in the center of the area of space being simulated and has a diameter of 0.01 meters. To set the aperture the user clicks on the element in the aperture specification that is associated with the desired aperture number using the mouse. For this simulation, aperture number one will be used so the user highlights the first aperture entry in the aperture specification list. This will cause the associated aperture information to appear in an area under the aperture list that may be modified by the user. Once the user has entered the desired location and size of the aperture the "Load" button is pressed and the new aperture information is entered into the aperture information list. If the user wishes to see a graphical view of the location and size of the currently active apertures they can press the "Display Aperture" button in the aperture specification window to display a diagram of the current apertures. A phase screen is then added to simulate the turbulence of the atmosphere. After traveling though the phase screen the signal is deflected from the target back through the phase screen to the ladar as illustrated in Figure 4.60. This returned signal is deflected by the polarizing beamsplitter and a mirror toward a beamsplitter that is used to combine the local oscillator signal with the returned signal. Finally, this signal is focused and heterodyned in the detector icon as shown in the final layout of the monostatic laser radar system in Figure 4.61.

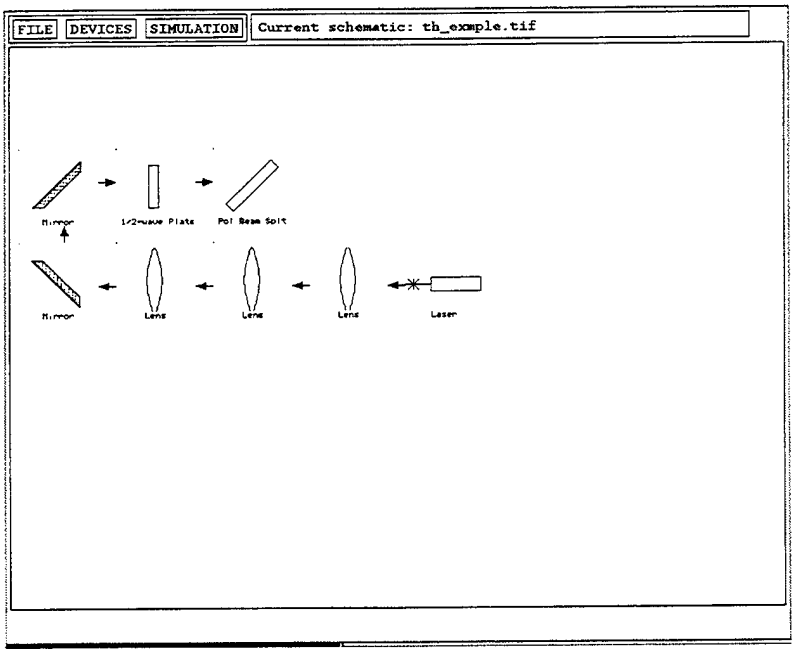


Figure 4.57: Polarizing beamsplitter inserted to create a target and local oscillator signal.

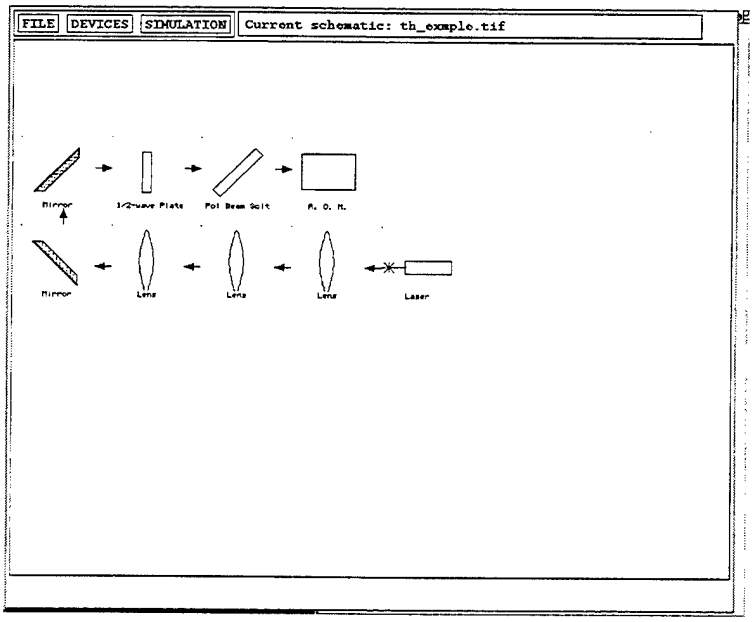


Figure 4.58: Acousto-optic modulator used to shift the frequency of the target signal so that the target's direction of travel can be established.

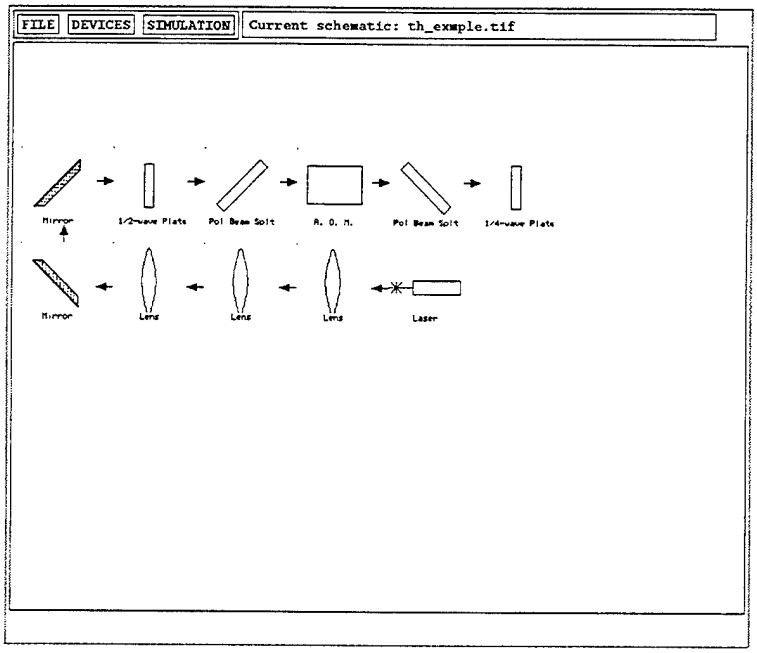


Figure 4.59: Ladar schematic after the polarizing beamsplitter - quarter-wave plate combination placed.

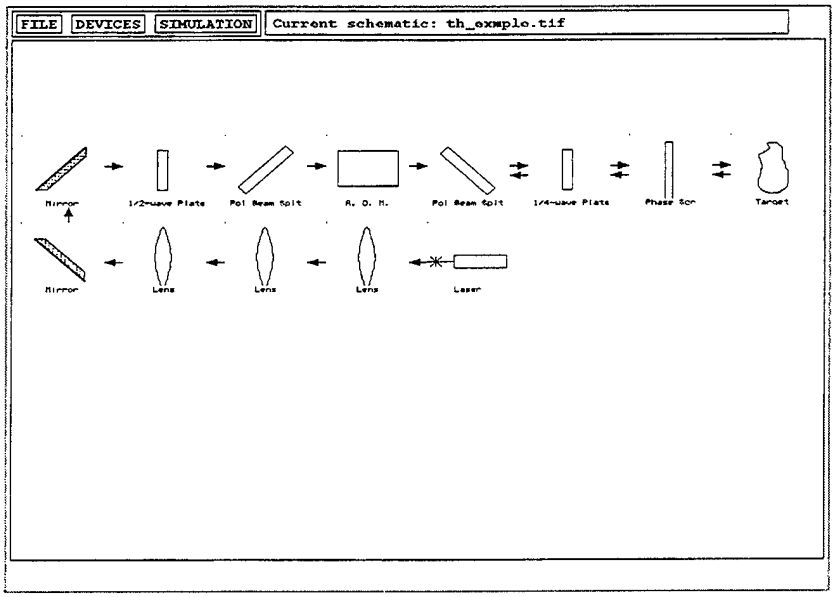


Figure 4.60: Returned ladar signal passing back through a phase screen to the ladar.

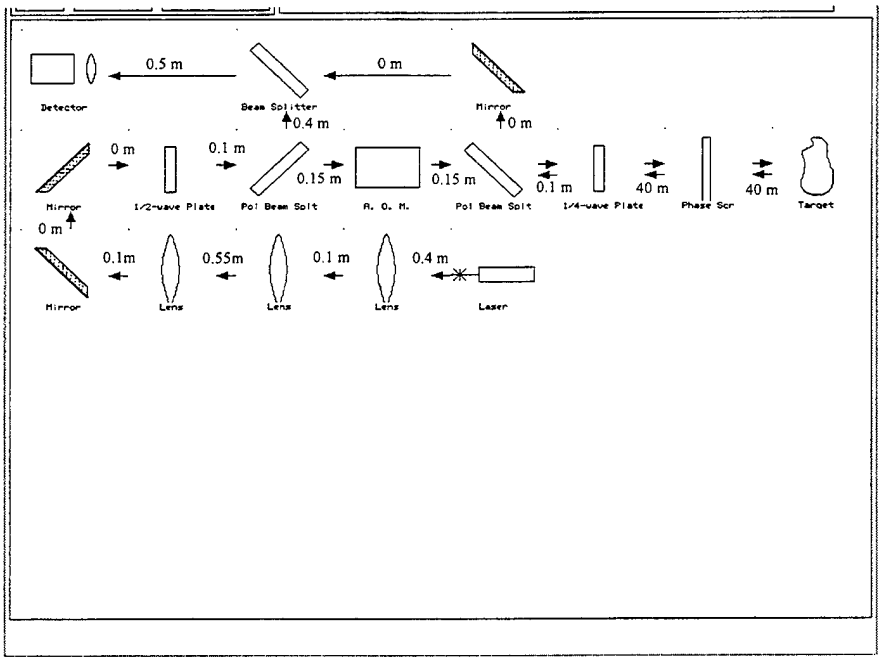


Figure 4.61: Schematic diagram of a monostatic laser radar.

Before a simulation can be run the spatial sampling interval and the frequency shift caused by the acousto-optic modulator have to be specified. This is done in the simulation parameters window (Figure 4.62) which appears when the user selects the "Simulation Parameters" option in the "SIMULATION" pull down menu. The simulation can now be run by selecting the "Run" option under the "SIMULATION" menu. The "Run" option of the simulation pull down menu will stay highlighted as long as a simulation is running. When a simulation is

finished the "SIMULATION" pull down menu will disappear. In the next section, several different examples of using the laser radar simulator will be presented.

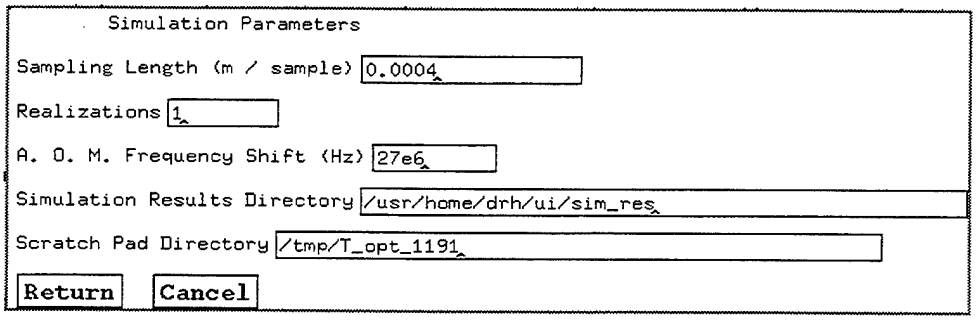


Figure 4.62: Simulation parameters window.

4.2.4 Experiments Performed With Laser Radar Simulator

This section presents three sets of simulation experiments performed with the laser radar simulator discussed in the previous sections [23][24]. The first simulation produces both fully and partially developed speckle and examines the speckle statistics of each. The statistics for fully developed speckle are then compared with theory from Goodman [23]. The second simulation calculates the radial intensity fluctuations of a Gaussian beam passing through a Kolmogorov phase screen and compares the results with theory from Miller *et. al.* [17]. The third simulation verifies the operation of a multiple aperture laser radar system. Intensity images are obtained at several locations throughout the laser beam propagation path.

4.2.4.1 Speckle Simulations

Figure 4.63 gives the geometrical layout used to generate both fully and partially developed speckle. For all of these simulations, the first step in producing speckle is to create an object with a diffuse surface. The object shape used for all speckle simulations is a square. The diffuse reflectance of the square is modeled by allowing its surface to possess spatially uncorrelated, uniformly distributed random height variations. These height variations introduce a random phase between -q and q to the propagating wave that is incident upon the square aperture. The value of q chosen depends on the amount of scattering desired. The square has a reflectance of 1, so when an incident wave strikes it, the portion of the wave that lies within the boundaries of the square has its phase randomized with no change introduced onto the electric field magnitude.

The second step in producing speckle is to propagate the diffuse object into the observation (a, b) plane. If the far field diffraction pattern is desired, the Fourier transform of the square aperture is computed. A list parameters used to produce fully developed speckle in the far field is given in Table 4.2.

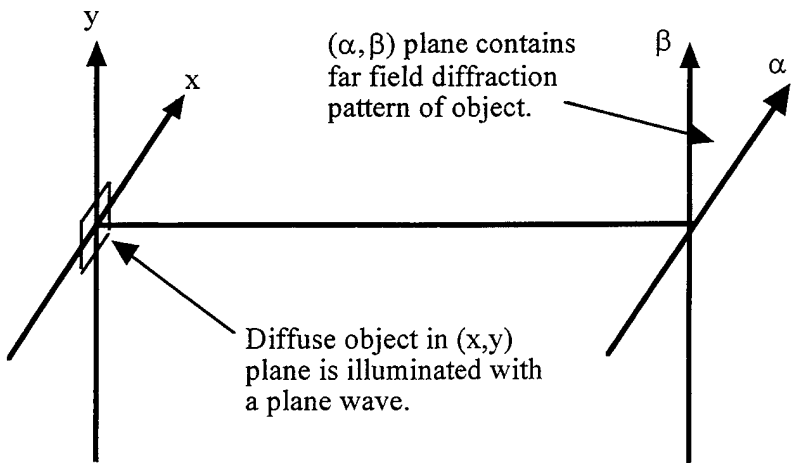


Figure 4.63: Experimental setup for the generation of speckle.

Table 4.2: Parameters used in probability density function verification.

Sampling interval in object plane	.0004 m
Wavelength of incident light	1 μm
Propagation distance	4800 m
Sampling interval in image plane	.0234 m
Object shape	square
Object side length	1.8 cm
Peak to peak phase of random object	2π
Object reflectivity	1
Electric field strength of incident light	1 V/m

The fully developed speckle is produced using the setup in Figure 4.63 in conjunction with the parameters in Table 4.2. Once the simulated pdf's are found from the speckle, they are compared to theory. The intensity distribution of one example speckle pattern generated using the laser radar simulator is shown in Figure 4.64. Figure 4.64 was produced by scaling the original intensity image located in the (a, b) plane to gray levels between 0 (black) and 255 (white). When scaling the original intensity image, it was noted that the minimum intensity value was

$$I_{min} = 3.270 \cdot 10^{-15} \text{ W/m}^2, \tag{4.80}$$

the maximum intensity was

$$I_{max} = 2.932 \cdot 10^{-8} \text{ W/m}^2, \tag{4.81}$$

and the mean intensity value was

$$I_{mean} = 2.992 \cdot 10^{-9} \text{ W/m}^2. \tag{4.82}$$

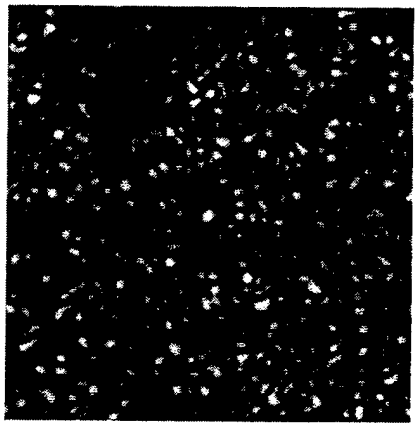


Figure 4.64: 11.98 m by 11.98 m far field intensity pattern produced from the 1.8 cm by 1.8 cm diffuse square.

According to Goodman [23], the pdf of the intensity of fully developed speckle should be negative exponential.

$$p(I) = \frac{1}{\langle I \rangle} \cdot e^{\frac{-I}{\langle I \rangle}}, \tag{4.83}$$

where $\langle I \rangle$ denotes the mean of the intensity. Figure 4.65 shows the resultant pdf. Also shown is the theoretical pdf predicted by Goodman [23]. The theoretical pdf is generated by calculating the mean value of the gray level image in Figure 4.64 and using this value for <I> in Equation (4.83). Equation (4.83) is then evaluated at all gray level values from 0 to 255. The final step in generating the theoretical pdf is to scale the domain so that it ranges from $I_{min}/\langle I \rangle$ to $I_{max}/\langle I \rangle$. The excellent agreement shows that the speckle intensity possesses the statistics expected for fully developed speckle verifying the propagation and the diffuse models used in the laser radar simulator.

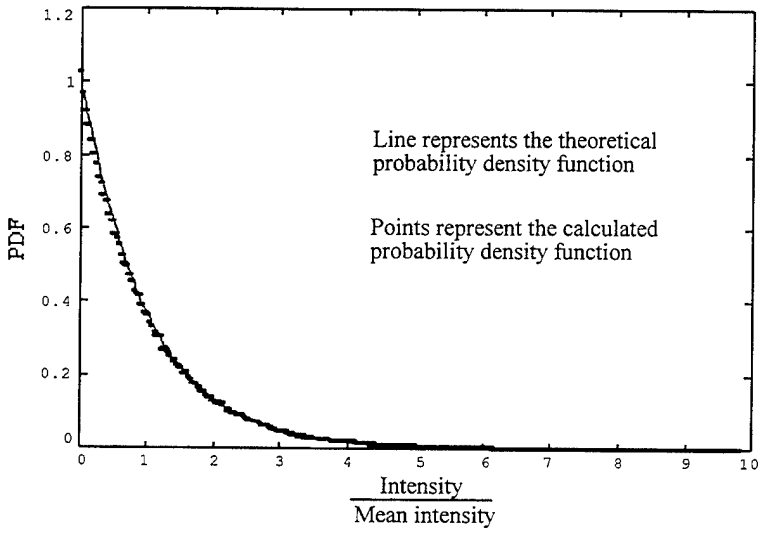


Figure 4.65: Comparison of simulated and theoretical pdf's for the intensity of fully developed speckle.

According to Goodman [23], the electric field magnitude of fully developed speckle should follow a Rayleigh distribution:

$$p(E) = \frac{E}{\frac{2}{\pi}\langle E\rangle^2} e^{\frac{E^2}{\frac{4}{\pi}\langle E\rangle^2}}, \tag{4.84}$$

where $\langle E\rangle$ is the mean value of the electric field amplitude. As with the intensity pdf, the simulated electric field pdf is constructed by first scaling the electric field values between 0 and 255. The pdf for these electric field values is then computed and scaled so that it has a domain of the E /<E> analogous to the intensity pdf in Figure 4.65. To generate the theoretical pdf from Equation (4.84), the value of <E> is replaced with the mean electric field and Equation (4.84) is then evaluated for all gray level values from 0 to 255. The pdf is then scaled so that the domain ranges from E_{min}/<E> to E_{max}/<E>. The simulated and theoretical plots are both shown in Figure 4.66. The excellent agreement between the simulation and theory shows that the simulation is correctly predicting the statistics of the electric field for fully developed speckle.

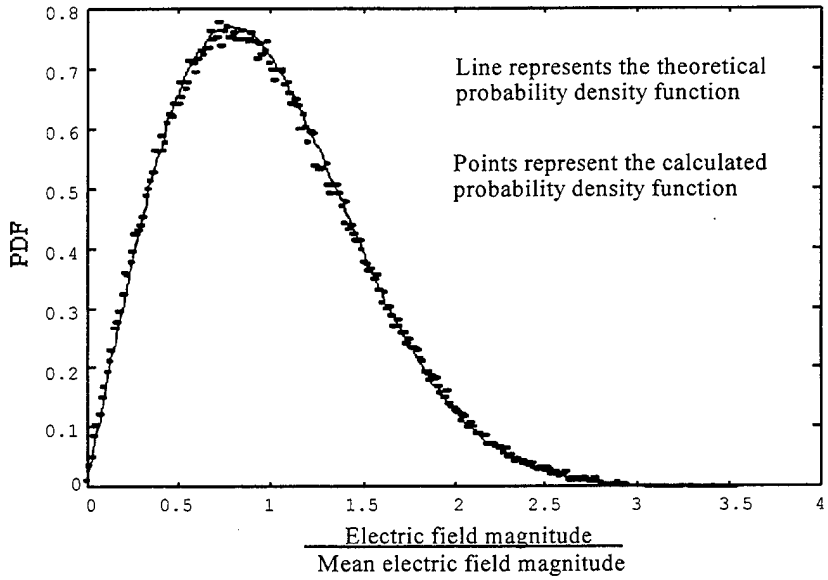


Figure 4.66: Simulated and theoretical pdf's for the electric field magnitude of fully developed speckle.

The final statistics that are compared to theory is for the phase of fully developed speckle. Goodman [23] states that the phase of fully developed speckle should follow a uniform distribution law:

$$p(\theta) = \begin{cases} \frac{1}{2\pi} & 0 \le \theta < 2\pi, \\ 0 & \theta < 0 \text{ or } \theta \ge 2\pi \end{cases} \tag{4.85}$$

The simulated pdf for the phase is calculated in the same basic manner as the simulated pdf's for the intensity and electric field magnitude. The difference is that before the phase is scaled between 0 and 255, extra multiples of

2p must be added or subtracted as needed to assure that all phase values are in the range $[0,2\pi)$. Figure 4.67 shows the simulated and theoretical pdf's for the phase of fully developed speckle. This shows that the statistics of the phase in the simulated speckle agrees with theory.

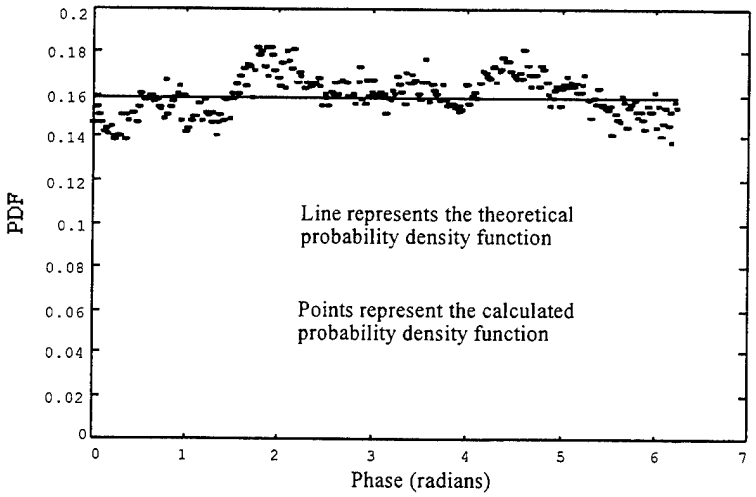


Figure 4.67: Simulated and theoretical pdf's for the phase of fully developed speckle.

4.2.4.2 Simulation of partially developed speckle

The previous section compared simulated statistics of fully developed speckle with the theoretical predictions. In this section, the intensity pdf of partially developed speckle will be presented along with a discussion of the differences from that of fully developed speckle. To accomplish this task, two examples of partially developed speckle are produced using the setup illustrated in Figure 4.63. This simulation uses the parameters listed in Table 4.2 except the for peak-to-peak phase of the object is changed from 2p to 1.7p. This corresponds to spatially uncorrelated random phases between -.85p and .85p residing on the surface of the square object. Reducing the peak-to-peak phase fluctuations of a diffuse object effective make the object smoother. The second simulation uses the same parameters as the first simulation, except the size of the square aperture is reduced to 1.24 cm by 1.24 cm, and the peak-to-peak value of the phase fluctuations is reduced to 1.25p radians.

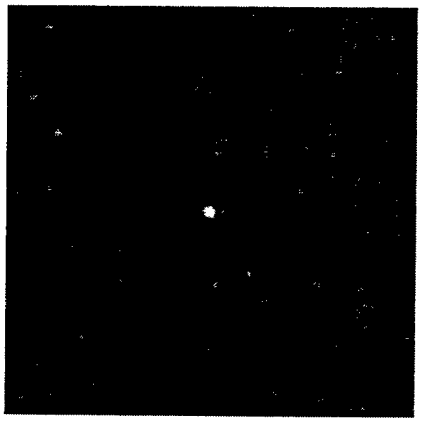


Figure 4.68: 11.98 m by 11.98 m intensity image of partially developed speckle produced by propagating a 1.8 cm by 1.8 cm diffuse square into the far field.

The intensity image resulting from propagating the 1.8 cm by 1.8 cm square aperture into the (a, b) plane is shown in Figure 4.68. As can be seen, the specular part of the image is located in the center of the window. This corresponds to the central lobe in the diffraction pattern produced by the square shape of the object. This bright central lobe is surrounded by dim speckles. The lowest intensity was found to be

$$I_{min} = 1.780 \cdot 10^{-15} \text{ W/m}^2. \tag{4.86}$$

while highest intensity was

$$I_{max} = 1.389 \cdot 10^{-7} \text{ W/m}^2. \tag{4.87}$$

and the mean intensity was

$$I_{mean} = 2.987 \cdot 10^{-9} \text{ W/m}^2. \tag{4.88}$$

Using the values in Equations (4.86) - (4.88), the intensity pdf vs. I/I_{mean} is then calculated from the intensity image in Figure 4.68. The result is shown in Figure 4.69.

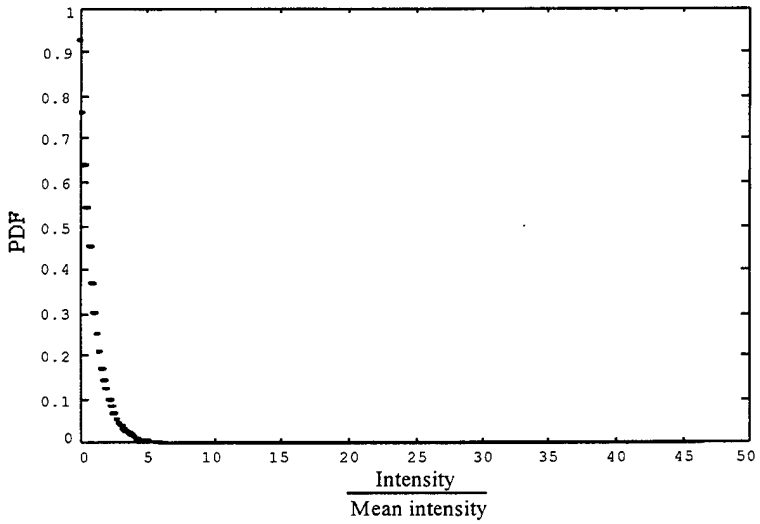


Figure 4.69: The pdf for partially developed speckle produced by a 1.8 cm by 1.8 cm square aperture.

The pdf in Figure 4.69 does appear to be negative exponential for very small values of I/I_{mean}, however the pdf extends out much further than the one for the fully developed speckle in Figure 4.65 due to the specular component that is present. Notice that the domain in Figure 4.65 extends out to 10 while the domain in Figure 4.69 extends out to 50.

The gray-level image produced from the propagation of the 1.24 cm by 1.24 cm square aperture into the (a, b) plane is shown in Figure 4.70. By making the object smaller in the second simulation and propagating the same distance, the specular image is made larger. As can be seen in Figure 4.70, the specular part occupies a larger fraction of the total image when compared to Figure 4.68.

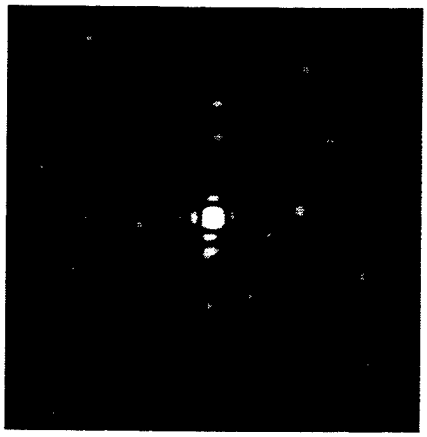


Figure 4.70: 11.98 m by 11.98 m intensity image of partially developed speckle produced by propagating a 1.24 cm by 1.24 cm diffuse square into the far field.

Figure 4.71 show the pdf derived from the intensity image given in Figure 4.70. Its shape is similar in appearance to the pdf in Figure 4.69 except for a shift towards the higher intensities due to larger spatial region being occupied by the specular component. This is most notable in Figure 4.71 for the range of values within of 5 and 200. As can be seen, the effect of decreasing the size of the semi-smooth object along with reducing the peak-to-peak phase fluctuations has made the specular image more predominant in the resultant pdf.

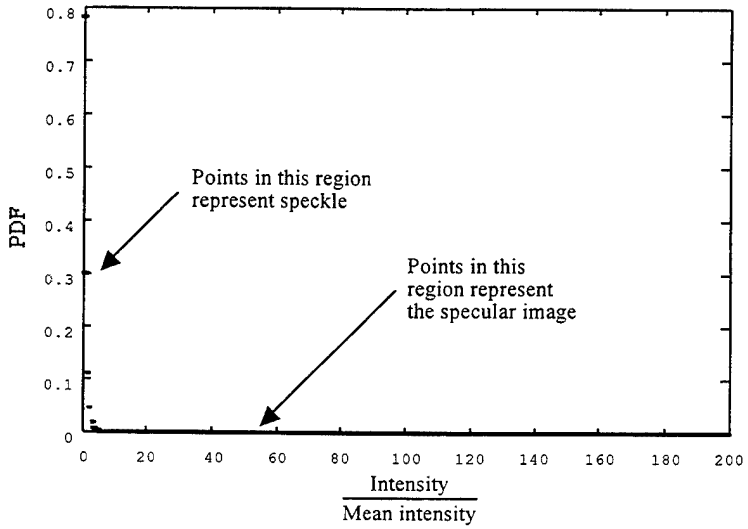


Figure 4.71: 11.98 m by 11.98 m gray level intensity image of partially developed speckle produced by propagating a 1.24 cm by 1.24 cm diffuse square into the far field.

4.2.4.3 Experiments demonstrating evolution of speckle with propagation distance

This sub-section demonstrates how speckle patterns evolve with increasing propagation distance. With reference to Figure 4.2, several intensity images will be obtained by varying the distance between the diffuse object and the observation plane. The simulation parameters used to produce the intensity data are taken directly from Table 4.2, with the exception of the propagation distance. The propagation distance was varied z = 10 m, z =

20 m, z = 30 m, z = 60 m, z = 70 m, and finally, z = 80 m. An image of the initial object will also be presented. Figure 4.72(a) shows a 512 X 512 image of the initial object. The remaining images are presented in order from 10m to 80 m.

When comparing the intensity images, it is seen that certain features remain constant as the diffraction pattern grows with increased propagation distance. This is especially true for the larger propagation distances. This is because for larger propagation distances, the diffraction pattern approaches the far field result. It is interesting to note that for any propagation distance that satisfies the Fraunhofer condition, the resultant diffraction pattern will never change in shape. Only the scaling will change.

4.2.4.4 Gaussian Moments Experiments

The second group of simulated experiments presented here are the Gaussian moments simulations. The goal of these simulations is to obtain the radial fluctuation statistics of a Gaussian beam passing through a time correlated Kolomogorov phase screen. Once the radial fluctuation statistics are computed from the raw intensity data, they are compared with theory from Miller *et. al.* [17]. The first step in constructing the simulation given in Figure 4.73 is to place a laser in the simulation layout area to create a Gaussian beam of a given power, wavelength, and beam waist. The second step is to propagate the Gaussian beam to the phase screen. Once the beam is propagated to the phase screen plane, the laser radar simulator creates a single phase screen realization of a given thickness and turbulence strength. The corrupted Gaussian beam is then propagated to the detector plane. Finally, the detector takes intensity readings starting at the optic axis of the system and extending out in a radial direction. This is illustrated in Figure 4.74.

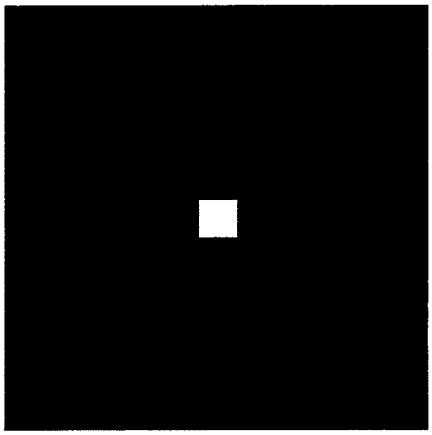


Figure 4.72(a): Intensity image of a diffuse object before propagation.

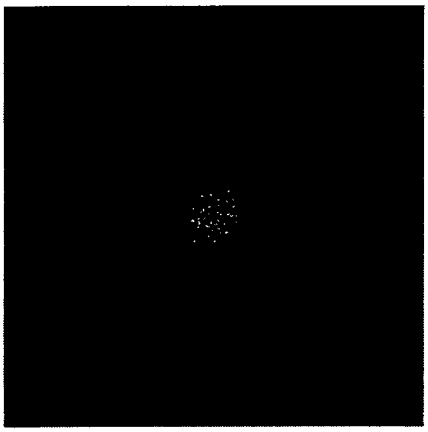


Figure 4.72(b): Intensity image of a diffuse object propagated 10 m.

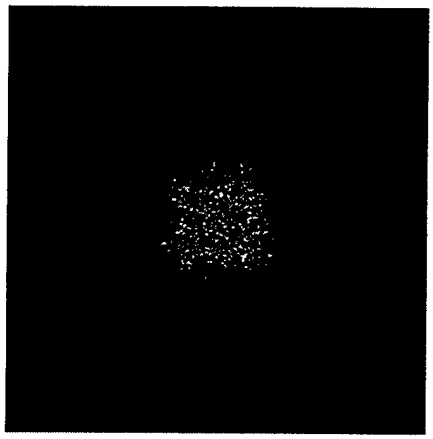


Figure 4.72(c): Intensity image of a diffuse object propagated 20 m.

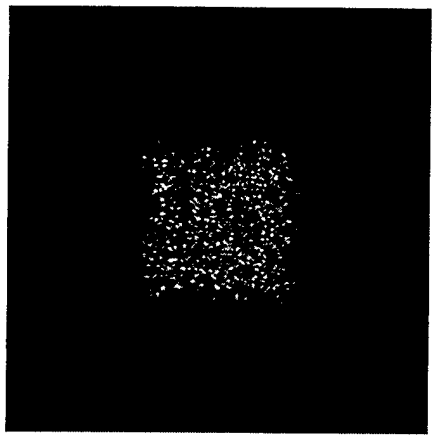


Figure 4.72(d): Intensity image of a diffuse object propagated 30 m.

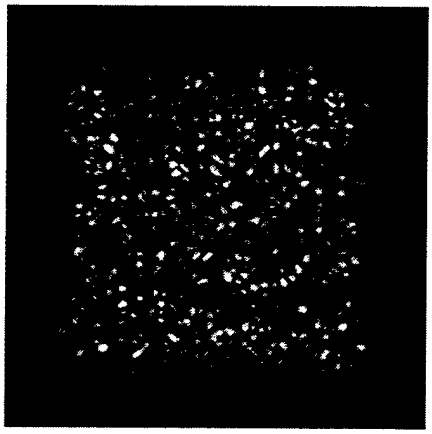


Figure 4.72(e): Intensity image of a diffuse object propagated 60 m.

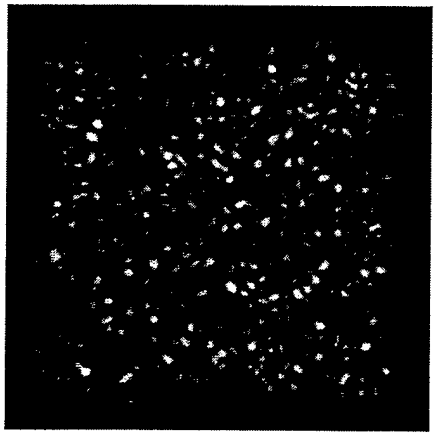


Figure 4.72(f): Intensity image of a diffuse object propagated 70 m.

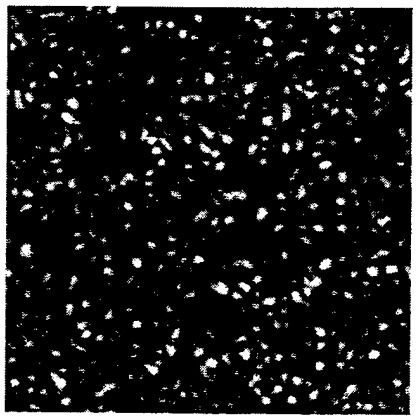


Figure 4.72(g): Intensity image of a diffuse object propagated 80 m.

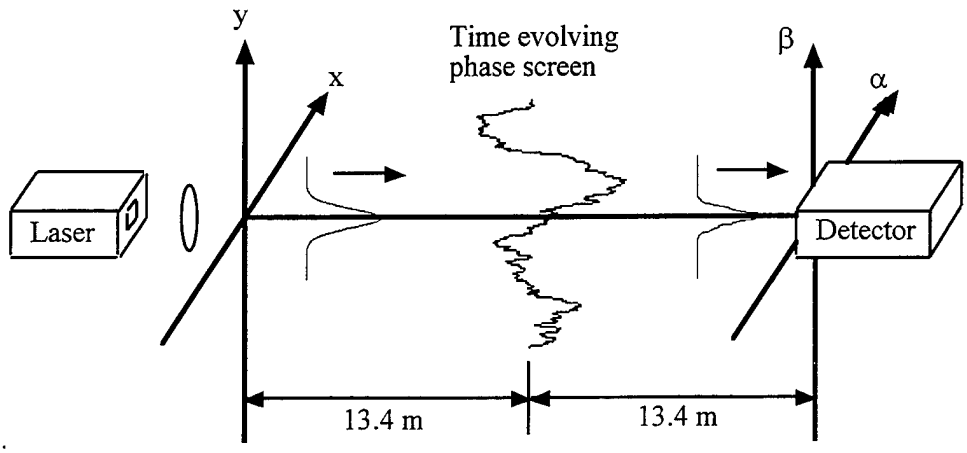


Figure 4.73: Simulated experimental setup for the Gaussian moments simulations.

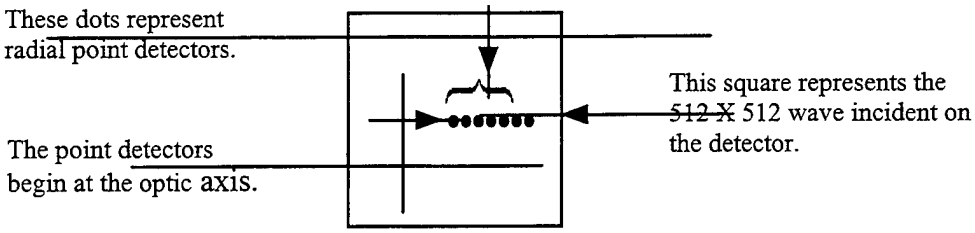


Figure 4.74: Illustration of a point detector arrangement in the detector plane.

The radial intensity measurements are taken in the same positions for each successive phase screen realization. Figure 4.75 shows a flowchart that illustrates the process of gathering intensity measurements for multiple phase screen realizations. Using this framework, the simulated intensity fluctuations can now be compared with theory. Each intensity realization is generated by propagating a Gaussian beam through a translating phase screen, and then computing the resultant intensity patterns in the detector plane. This is done to illustrate the effect that a translating phase screen has on the incident Gaussian beam. Successive phase screen realizations are created by introducing a certain amount of translation and/or boil between the successive phase screen realizations. Referring to Figure 4.75, a Gaussian beam of radius .00943 m and phase curvature of 24398.3 m is incident on the phase screen. The phase screen has a thickness of 1 cm and an index of refraction structure constant of $C_n^2 = 1.0 \cdot 10^{-10}\ m^{-2/3}$. Once the Gaussian beam passes through the phase screen, it is propagated 13.4 m to the detector plane. Figure 4.76 shows the degraded beam as it appears in the detector plane. To demonstrate the effect that a translating phase screen has on the intensity pattern in the detector plane, a new phase screen realization is created by shifting 5.

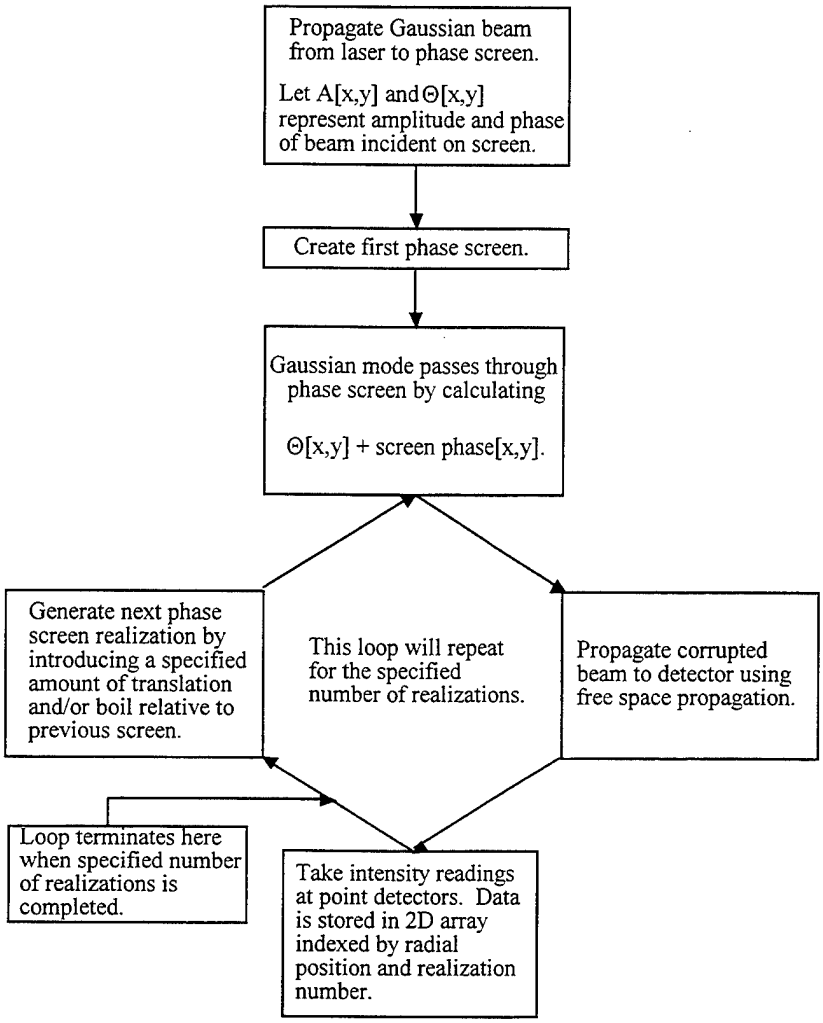


Figure 4.75: Flowchart showing the process of gathering intensity data for multiple phase screen realizations.

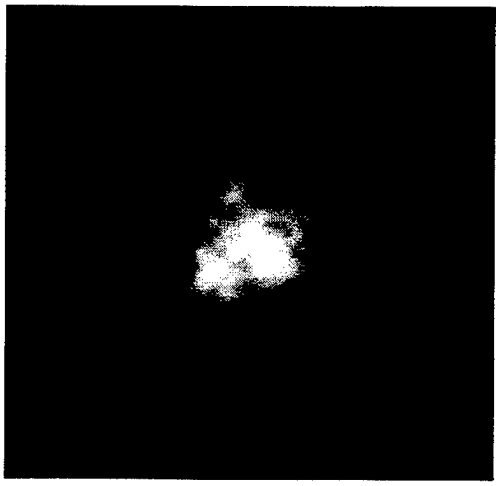


Figure 4.76: Intensity image for a single phase screen realization in the detector plane.

As previously stated, the goal of this simulation is to compare the simulation results with the theory predicted by Miller [17]. From Miller the equation describes the normalized off-axis intensity variance divided by the normalized on-axis intensity variance is defined as

$$\frac{\sigma_N{}^2(r)}{\sigma_N{}^2(0)} = 1 + \frac{4.42 \cdot \Lambda^{5/6} \cdot \frac{r^2}{W^2} \cdot [1 + \alpha]}{3.86 \cdot \left\{ .40 \cdot \left[(1+2\Theta)^2 + 4\Lambda^2 \right]^{5/12} \cdot \cos\left[\frac{5}{6} \tan^{-1}\left(\frac{1+2\Theta}{2\Lambda} \right) \right] - \frac{11}{16}\Lambda^{5/6} \right\}} \tag{4.89}$$

where

$$\Theta = 1 + \frac{L}{R}, \tag{4.90}$$

$$\Lambda = \frac{2L}{kW^2}, \tag{4.91}$$

$$\alpha = 1.0 + .083 \cdot \left[\frac{R}{W} \right]^2, \tag{4.92}$$

W and R are the waist and radius of curvature of the laser beam at the detector plane, k is the wavenumber, and L is the distance from the phase screen to the detector plane. This equation was derived with the assumption that the intensity statistics are in the weak turbulence regime. The reason that the simulated data is compared with the theoretical values of $\sigma_N^2(r)/\sigma_N^2(0)$ is that forming this ratio normalizes out the C_n^2 parameter. This makes the theoretical quantity independent of C_n^2 as along as the weak scattering approximation is satisfied. With a = 0 in Equation (4.89), the theoretical equation is valid only for radial distances very close to the optic axis. The a parameter (defined in Equation (4.89)) is included so that the simulated and theoretical fluctuation statistics can be compared out to distances close to W (defined above). The parameter a is an additional fourth order correction which was not originally part of Equation (4.89).

To compare simulation results with the theoretical equation given in Equation (4.89), the normalized variance as a function of radial distance is computed

$$\sigma_N^2(r) = \frac{\frac{1}{\beta}\sum_{n=1}^{\beta}\left[I(n,r) - \text{mean}(r) \right]^2}{\left[\text{mean}(r) \right]^2}, \tag{4.93}$$

where

$$\text{mean}(r) = \frac{1}{\beta}\sum_{n=1}^{\beta} I(n,r), \tag{4.94}$$

b = number of time realizations, r = distance from optic axis, and I(n, r) represents the raw intensity data collected from the laser radar simulator. Next, the normalized off axis variance is divided by the normalized on axis variance

$$\Omega(r)^2 = \frac{\sigma_N^2(r)}{\sigma_N^2(0)}. \tag{4.95}$$

The parameters used by the laser radar simulator for this simulation is listed in Table 4.3. Figure 4.77 shows $\sigma_N^2(r)/\sigma_N^2(0)$ for the simulated data and compares it to the theoretical curve given by Equation (4.89). The domain in Figure 4.77 is the radial distance from the optic axis in the detector plane normalized by W. When the domain value is equal to 1, the radial distance from the optic axis is equal to the beam waist of the uncorrupted Gaussian beam incident on the detector.

Table 4.3: List of simulation parameters with sample values

Laser power	50 mW
Wavelength	1 μm
Gaussian beam radius at detector (W)	1.509 cm
Gaussian beam phase curvature at detector (R)	$5.872 \cdot 10^{5}$ m
Gaussian beam confocal parameter	715.400 m
Simulation sampling interval	$1.614 \cdot 10^{-4}$ m
Laser to phase screen distance	13.4 m
Spectral index	-11/3
Phase screen thickness	1 cm
Number of columns shifted for translation (multip)	2
Number of realizations (b)	900
Frequency of columns replaced for boiling (boilskip)	0
Index of refraction structure constant (C_n^2)	$1.0 \cdot 10^{-13}$ m$^{-2/3}$
Phase screen to detector distance (L)	13.4 m
Area of point detector	$2.605 \cdot 10^{-8}$ m^2

Examining Figure 4.77, there is good agreement between the simulated result and both of the theoretical curves. However at radial distances greater than W, the simulated curve grows much faster than either theoretical curve. This is in part because more higher order terms are needed in Equation (4.89) for large radial distances. The real difficulty in Figure 4.77 is that $\sigma_N^2(r)/\sigma_N^2(0)$ is seen to drop below 1 for distances very close to the optic axis. According to Equation (4.89), this should never happen.

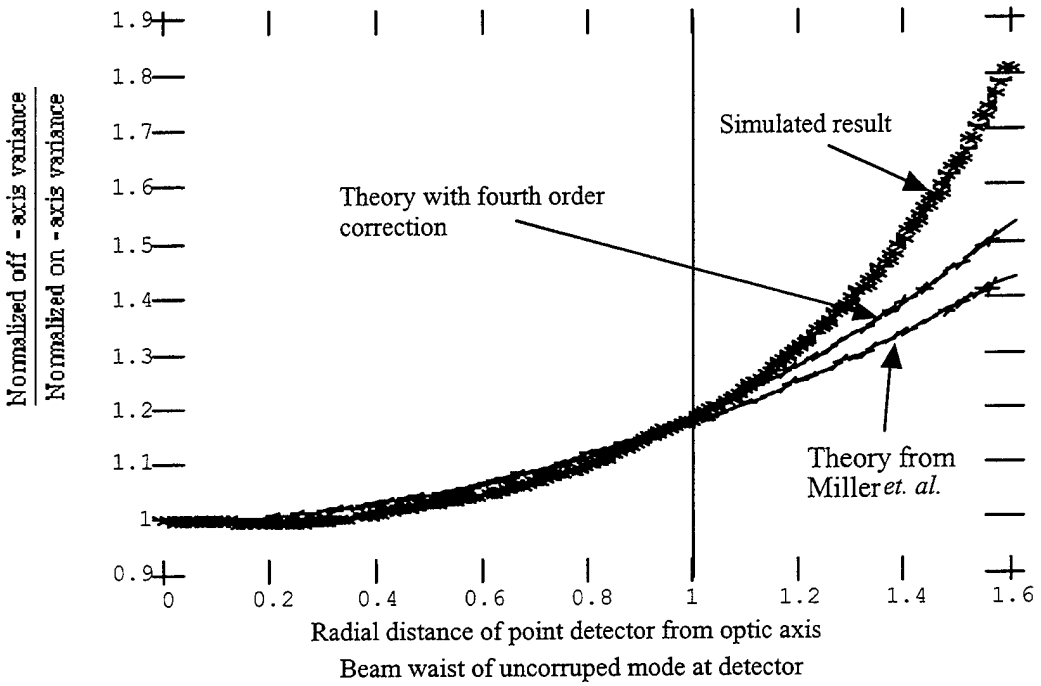


Figure 4.77: Simulated and theoretical results for the simulation parameters given in Table 4.2.

4.2.4.5 Demonstration of Multiple Aperture Laser Radar Simulation

The laser radar that is simulated is a two aperture system which illuminates a rough target through a single phase screen realization. The primary results of this simulation will be the AC beat signal for each aperture, and the intensity images at various points in the laser beam propagation path. These images will illustrate such effects as phase screen perturbations and diffraction effects. Along with the intensity images, one-dimensional slices of the electric field will be given.

The simulated experimental setup for the multiple aperture radar is illustrated in Figure 4.78. This figure includes the arrangement of all components, in addition to the component spacing where required. Diffraction is neglected for the signal wave from the point it enters either receiving aperture up until it is focused onto the detector. Once the laser beam leaves the transmitting aperture, it propagates 40 m to a phase screen. The beam then passes through a phase screen, and then illuminates a rough circular target. The beam reflects off the target, and is then propagated back to the receiving apertures through the phase screen. As indicated in Figure 4.78, the local oscillator beam that is used by the monostatic aperture is also used by the bistatic aperture.

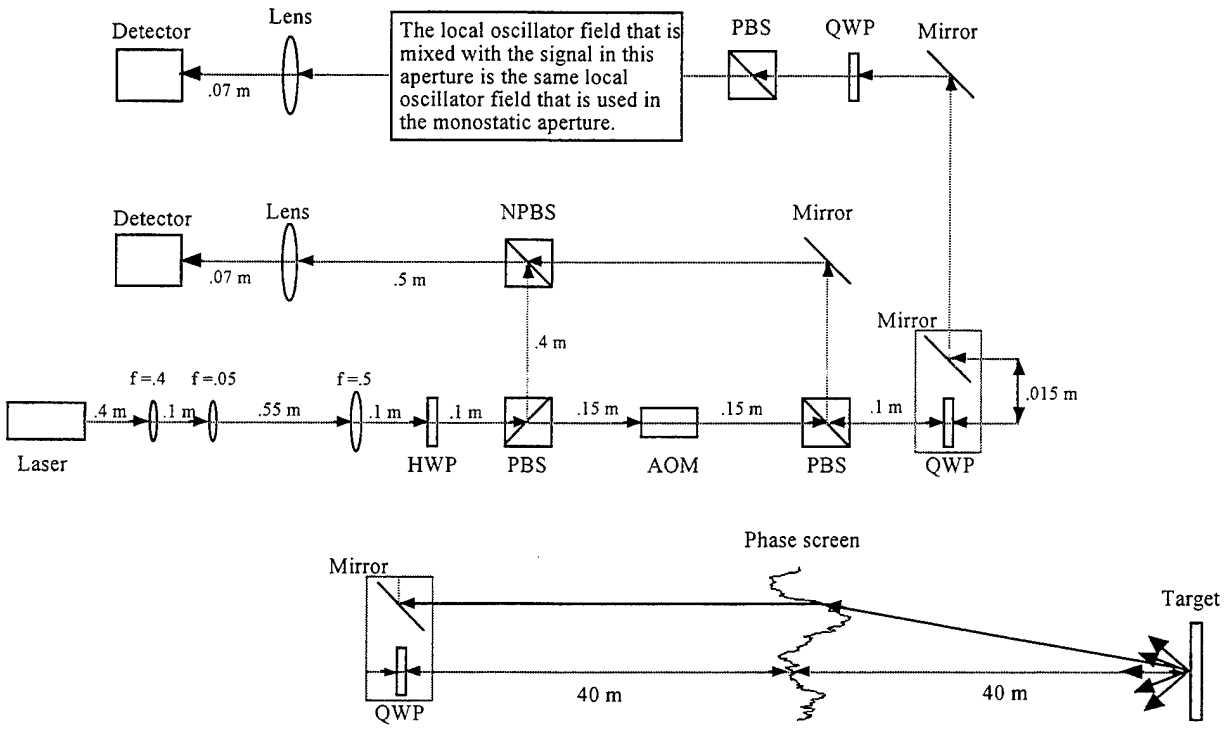


Figure 4.78: Simulated experimental setup for the demonstration of the multiple aperture laser radar simulation.

The target used in this simulation is a disk of 4 cm in diameter. The disk surface is composed of spatially uncorrelated, uniformly distributed phase variations between -.98 radians and .98 radians. The peak-to-peak phase variations of 1.96 radians will not produce fully developed speckle because the scattering is too weak. Figure 4.79 shows a two-dimensional picture of the target. This picture was constructed by scaling the phase variations to integer values between 0 (black) and 255 (white). The parameters used to define each optical component is given in Table 4.4.

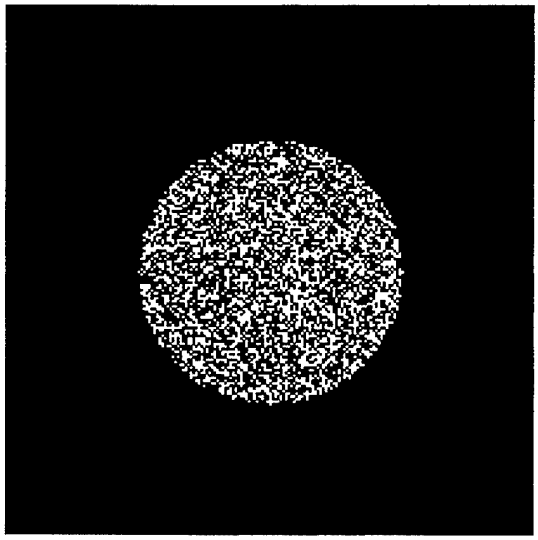


Figure 4.79: Picture showing the target used in the simulation. The target is 4 cm in diameter.

Table 4.4(a): Laser specifications.

Laser power	5 mW
Laser wavelength	1 mm
Laser waistsize	67.5 mm
Laser polarization	Linear along P axis
Spatial sampling interval	400 mm

Table 4.4(b): Half-wave plate (HWP) specifications.

Angle between slow axis and S axis	-p/8
Power transmission	.99

Table 4.4(c): Specifications for all polarizing beamsplitters (PBS) in simulation.

Forward extinction ratio	1/200
Side extinction ratio	1/20
Power transmission	.99

Table 4.4(d): Specifications for all quarter-wave plates (QWP) in simulation.

Angle between slow and S axis	-p/4
Power transmission	.99

Table 4.4(e): Specifications for non-polarizing beamsplitter (NPBS).

Power splitting ratio	.5
Power transmission	.99

Table 4.4(f): Specifications for phase screen.

C_n^2	$1.13 \cdot 10^{-10}$ $m^{-2/3}$
Spectral index	-11/3
Screen thickness	.01 m
Focal length	.07 m
Lens to detector distance	.07 m
Diameter	.01 m
Pan angle	0
Tilt angle	0
Power transmission	.99

Table 4.4(g): Specifications for monostatic and bistatic detectors.

Detector shape	circular
Detector diameter	150 mm
Sampling frequency	864 MHz
3dB point	27 MHz
Dark current	.5 nA
Current noise	.2 $pA/\sqrt{Hz}$
Responsivity	.7 A/W

Now that the specifications for the entire simulation have been defined, the intensity images and electric field slices will be given at different points in the simulation. In order to help organize the material flow better, all of the intensity images will be presented first, followed by the electric field slices. Every two-dimensional intensity image presented here is the S polarization component scaled between 0 and 255. The first intensity image is shown in Figure 4.80. This is the Gaussian beam as it exits the monostatic aperture through the quarter-wave plate. In order to give a visualization of what the phase screen looks like in two dimensions, Figure 4.81 shows the entire 512 X 512 phase screen with the spatial phase perturbations scaled 0 to π between 0 and 255. The Gaussian beam then propagates through the phase screen where it is distorted by an amount depending on the index of refraction structure constant and the thickness of the screen. Figure 4.82 shows the two-dimensional intensity image of the Gaussian beam that is incident on the target. As is evident, it no longer resembles the Gaussian beam in Figure 4.80. Figure 4.83 shows the two dimensional intensity image as it scatters off the target, and propagates back through the phase screen on to the detector.

After the scattered light illuminates the receiver, different parts of the scattered light enter the receiving apertures. The simulation now mixes the signal light and the local oscillator for each aperture and focuses the result onto a detector. This generally results in a beat signal between the frequency shifted local oscillator and signal wave. Figure 4.84 shows an intensity image of the local oscillator beam that has been focused onto the monostatic detector. The slight ringing is caused by the truncation of the Gaussian local oscillator beam by the lens aperture. Figure 4.85 shows an intensity image of the portion of the return signal wave that entered the monostatic aperture. Figure 4.86 shows the intensity image of the portion of the signal wave that entered the bistatic aperture. The intensity images incident on the detector should give a good indication as to the nature of the electric fields that are heterodyned by the detector.

To illustrate the effect of the free space propagations, phase screens, lenses, etc. on the electric field magnitude of the signal wave and local oscillator beam, slices of the magnitude of the electric field will now be presented. Each slice is taken down the center of the 512 X 512 S polarized wave that created each of the intensity images discussed above. The first slice is shown in Figure 4.87. It represents the electric field magnitude of the Gaussian beam exiting the monostatic aperture (transmitter). Figure 4.88 shows an electric field slice taken from the magnitude of the electric field incident on the target. Finally, Figure 4.89 shows a slice taken from the magnitude of the electric field incident on the receiver. The wave at the receiver in is much broader than

the wave at the target. This is caused by the target scattering and an additional pass through the phase screen. Although the semi-smooth target does not produce fully developed speckle, it does cause additional spreading in the signal wave. Figures 4.90 and 4.91 shows the distribution of the electric field and both the monostatic and bistatic detectors.

Now that the electric fields have been mixed and focused onto the monostatic and bistatic detectors, the noise-corrupted AC beat signals will be presented for each aperture. Figure 4.92 shows the AC beat signal for the monostatic aperture. Figure 4.93 shows the AC beat signal for the bistatic aperture. The beat signal for the bistatic aperture has a smaller RMS value and a slightly different phase. This is believable because referring to Figure 4.90 and 4.91, the monostatic detector has substantially more signal energy focused onto it. (Note the difference in vertical scales for Figs. 4.90 and 4.91, and Figs. 4.92 and 4.93.) One would expect a difference in phase since the optical lengths from the transmitter to the monostatic detector and from the transmitter to the bistatic detector are not the same.

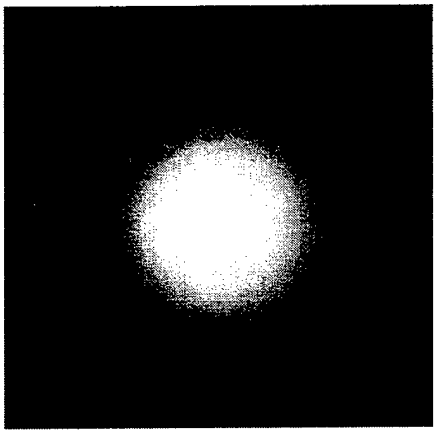


Figure 4.80: Intensity at transmitter.



Figure 4.81: Phase screen scaled between between 0 (0) and 256 (2p).



Figure 4.82 : Intensity at target.

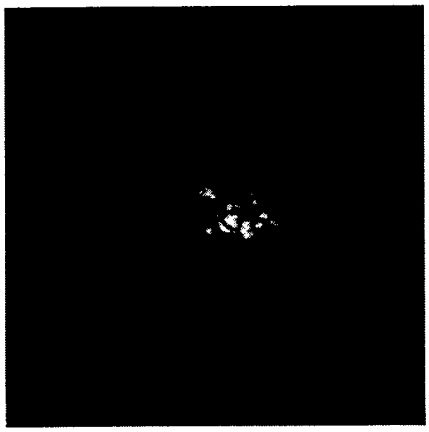


Figure 4.83: Intensity at receiver.

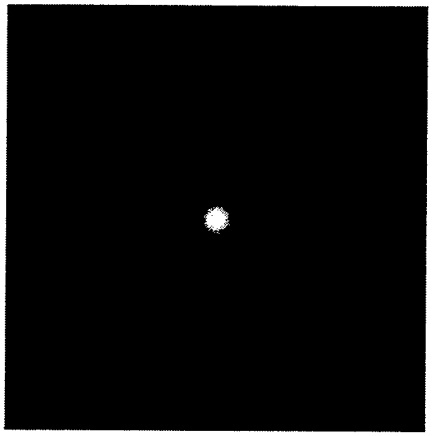


Figure 4.84: Local oscillator on detector.

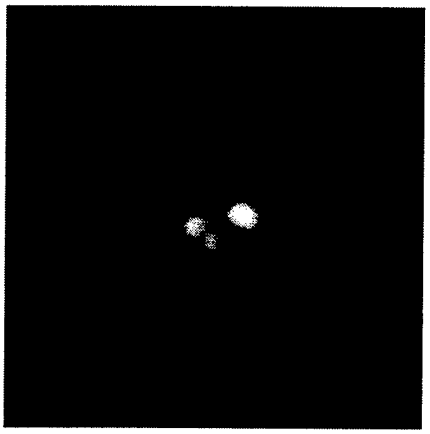


Figure 4.85: Monostatic aperture signal on detector.

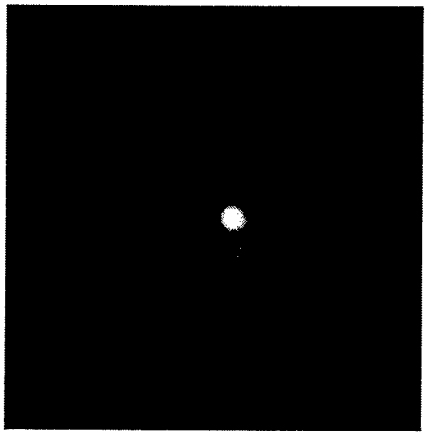


Figure 4.86: Bistatic aperture signal on detector.

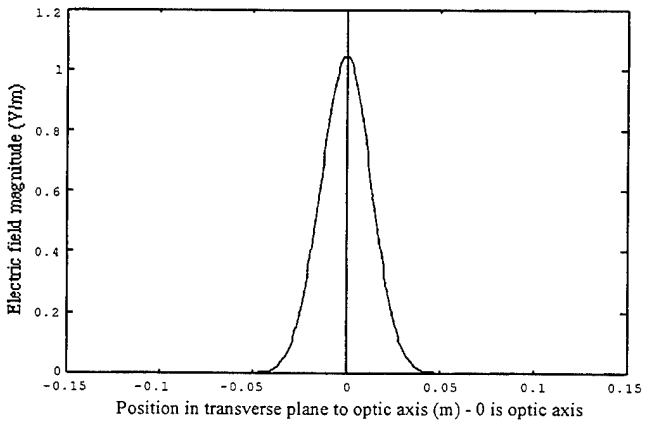


Figure 4.87: Slice of the electric field magnitude of the Gaussian beam leaving the monostatic aperture.

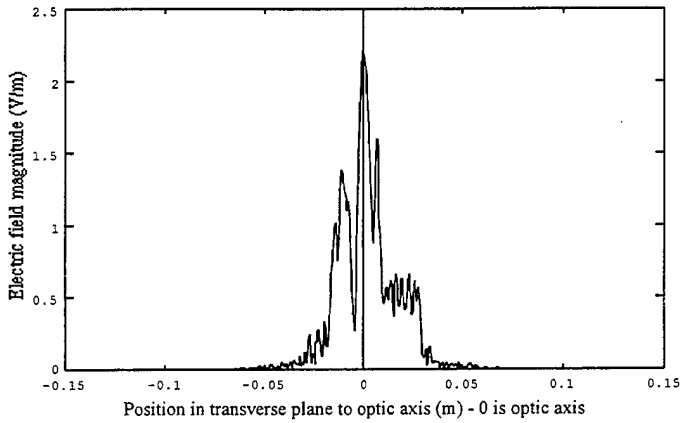


Figure 4.88: Slice from the magnitude of the electric field at the target.

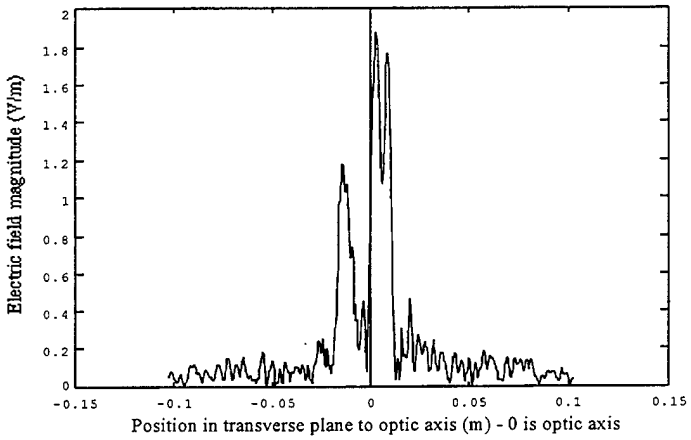


Figure 4.89: Slice from the magnitude of the electric field on the receiver.

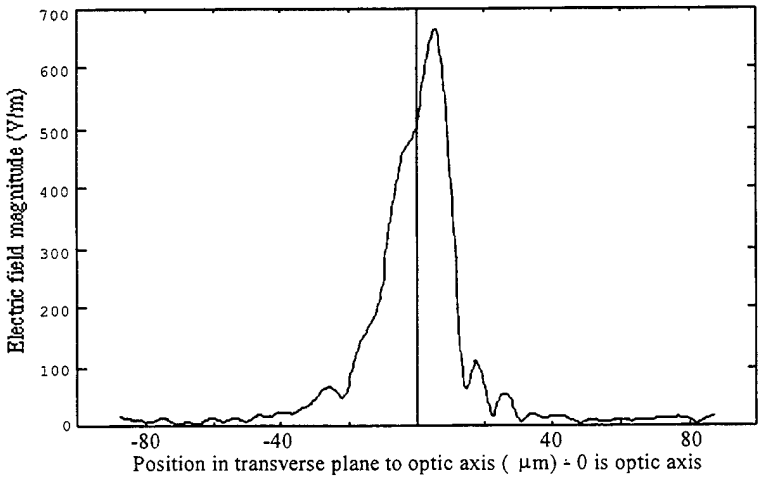


Figure 4.90: Slice taken from the magnitude of the electric field focused onto the monostatic (on - axis) detector.

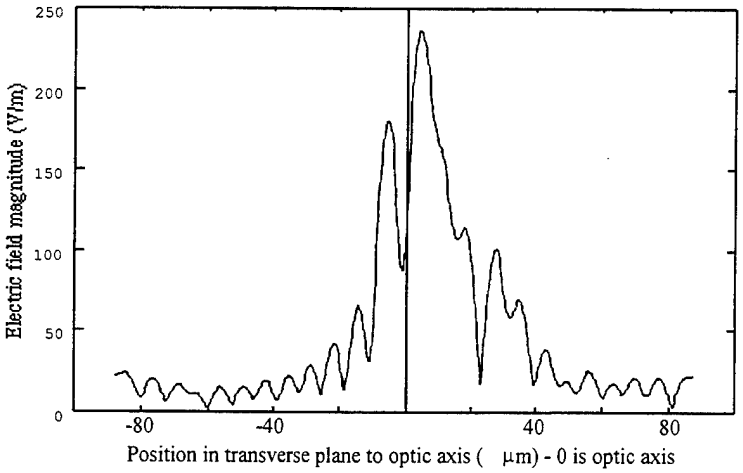


Figure 4.91: Slice taken from the magnitude of the electric field focused onto the bistatic (off - axis) detector.

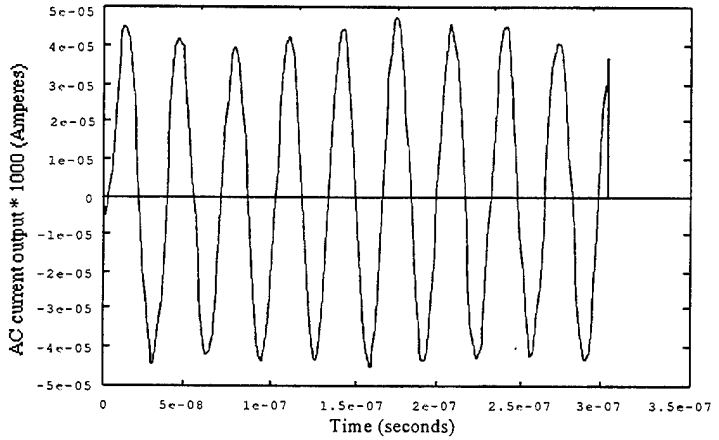


Figure 4.92: AC current output of the monostatic (on - axis) detector. The radius of the detector is 75 mm.

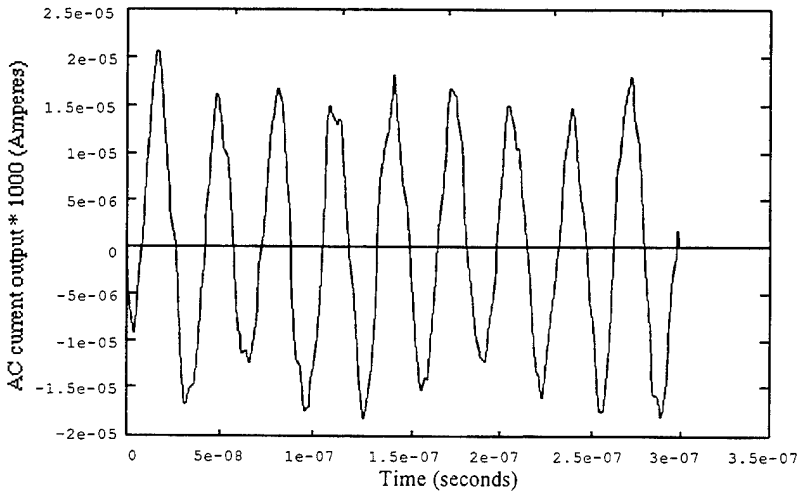


Figure 4.93: AC current output of the bistatic (off - axis) detector. The radius of the detector is 75 mm.

4.3 Conclusions

The laser radar simulation presented here is unique in its ability to simulate a multiple aperture laser radar system in 3 dimensions. This is accomplished by propagating two-dimensional waveforms between the transmitter/receiver and target, and in allowing the target and atmosphere to evolve as a function of time. Each time the atmosphere evolves or the target changes aspect, the perturbations introduced on the light incident from the laser radar transmitter are then propagated back to the receiver and detected coherently, for each of the receiving apertures in the receiver. The output from the coherent detection process can be interpreted in several ways. First, the laser radar output can be recorded as an AC beat signal resulting from the heterodyne detection. The AC output is made more realistic through the modeling of detector noise and shot noise. The detector model also calculates the heterodyne efficiency as a function of detector size. The phase and the peak value of the AC current signal are other possible detector outputs.

In addition to the unique approach used to simulate both space and time the laser radar simulator also provides a convenient means for the modeling of a multiple aperture laser radar system. The ease of use of this software is primarily due to its graphical user interface. The central feature of this interface is a window in which icons, representing optical devices, can be placed and connected to form an optical schematic. Pop-up windows are used to set the physical parameters of the devices in the schematic window. The technique for debugging an optical schematic was very similar to that used in debugging an electronic schematic.

A copy of the software developed under this contract can be obtained by writing to:

Dr. Arthur R. Weeks
University of Central Florida
Electrical and Computer Engineering
4000 Central Florida Blvd.
Orlando, FL 32816-2450
(407) 823-5762

4.4 References

1. Letalick, D., and A. Oestberg, [1984], "Computer Model for an Imaging Coherent Laser Radar," *GRA*, Vol. 84, No. 4, Feb., 48 pages.

2. Jelalian, A. V., [1992], *Laser Radar Systems*. Norwood MA : Artech House.

3. Wang, J., and J. Kostamovaara, [1994], "Radiometric analysis and simulation of signal power function in a short-range laser radar," *Applied Optics*, Vol. 33, No. 18, June, pp. 4069 - 4076.

4. Estep, J., and Z. Gu, [1992], "Ladar Signature Simulation," *SPIE Proceedings on Automatic Object Recognition II*, Vol. 1700, April, pp. 119 - 125.

5. Peters, W. N., and N. T. Nomiyama, [1976], "Laser Quadrant Tracker Simulation," *General Research Corp. Final Technical Report*, RADC-TR-76-204, Aug., 106 pages.

6. Leader, J. C., [1983], "Laser radar Analysis," *McDonnel Douglas Research Laboratories Final Technical Report*, MDC Q1203, July, 43 pages.

7. Frehlich, R. G., [1993], "Coherent Lidar Design and Performance Verification," NASA Annual Report, 1 Dec. 1992 to 1 Dec. 1993 (Colorado University), 97 pages.

8. Brown, W. P., C. Fry, and G. C. Valley, [1982], "Simulation of Ground to Space Optical Propagation," *Hughs Research Laboratories Final Technical Report*, RADC-TR-82-91, May.

9. Kachelmyer, A. L., [1989], "Mirror Surface Aberations and Mixing Efficiency," *SPIE Vol. 1045 Modeling and Simulation of Laser Systems*, pp. 89 - 98.

10. Iftekharuddin, K. M., and M. A. Karim, [1992], "Heterodyne detection by using a diffraction-free beam: tilt and offset effects," *Applied Optics*, Vol. 31, No. 23, Aug., pp. 4853 - 4856.

11. Tanaka, K., and N. Ohta, [1987], "Effects of tilt and offset of signal field on heterodyne efficiency," *Applied Optics*, Vol. 26, No. 4, Feb., pp. 627 - 632.

12. Goodman, J. W., [1988], *Introduction to Fourier Optics*. New York : McGraw-Hill.

13. Weeks, A.R., H.R. Myler, and H. Wenaas, [1993], "Computer generated noise images for the evaluation of image processing algorithms," *Optical Engineering*, Vol. 32, No. 5, May, pp. 982 - 992.

14. Bhattacharya, P., [1994], *Semiconductor Optoelectronic Devices*. Englewood Cliffs, NJ; Prentice-Hall.

15. Proakis, J. G., and D. G. Manolakis, [1992], *Digital Signal Processing Priciples, Algorithms, and Applications*. New York : Macmillan Publishing Company.

16. Gatt, P., [1989], "A Simulation and Theoretical Analysis of the Doubly Stochastic Homodyne K Scattering Model," Ph. D. Discertation, University of Central Florida, August, 162 pages.

17. Miller, W. B., J. C. Ricklin, and L. C. Andrews, [1994], "Effects of the refractive index spectral model on the irradiance variance of a Gaussian beam," *Journal of Optical Society ofAmerica* A / Vol. 11, No. 10, October, pp. 2719 - 2726.

18. Turner, J. L., [1989], "Simulation of Optical Propagation Through Atmospheric Turbulence Using Two-Dimensional Fourier Transform Techniques," Masters Thesis, Naval Postgraduate School, June, 54 pages.

19. Siegman, Anthony E., [1986], *Lasers*. Mill Valley CA : University Science Books.

20. Fowles, G. R., [1989], *Introduction to Modern Optics*. New York : Dover Publications.

21. Jelalian, A. V., [1992], *Laser Radar Systems*. Norwood MA : Artech House Fink, D.,

22. Verdeyen, Joseph T., [1989], *Laser Electronics*, Englewood Cliffs, NJ : Prentice Hall.

23. Goodman, J. W., [1984], "Statistical Properties of Laser Speckle," in *Laser Speckle and Related Phenomena*, ed. J. C. Dainty. New York : Springer-Verlag.

24. Gamble, K. J., [1995], *A Computer Simulation of a Multiple Aperture Coherent Laser Radar*, Masters Thesis, University of Central Florida, August, 269 pages.

24. Hefele, D. R., [1995], *A Computer Aided Design Package for Multiple Aperture Laser Radar*, Masters Thesis, University of Central Florida, December, 566 pages.

4.5. Publications, Presentations and Theses

1. K. J. Gamble, A. R. Weeks, H. R. Myler, and W. A. Rabadi, "Results of Two-dimensional Time-evolved Phase Screen Computer Simulations", Proc. SPIE, 2471, 170-180, June 1995.

2. H. R. Myler, A. R. Weeks, and W. A. Rabadi, "Speckle Simulation Movies for Analysis and Evaluation of Laser Systems", Proc SPIE, 2222, 823-826, June 1994.

3. K. J. Gamble, "A Computer Simulation of a Multiple Aperture Coherent Laser Radar", M.S. Thesis, University of Central Florida, August 1995.

4. D. Hefele, "A Computer Aided Design Package for Multiple Aperture Laser Radar", M.S. Thesis, University of Central Florida, December 1995.

4.6 Appendix 4A

K. J. Gamble, A. R. Weeks, H. R. Myler, and W. A. Rabadi, "Results of Two-dimensional Time-evolved Phase Screen Computer Simulations", Proc. SPIE, 2471, 170-180, June 1995.

Results of two-dimensional time-evolved phase screen computer simulations

K. J. Gamble, A. R. Weeks, H. R. Myler, and W. A. Rabadi

Department of Electrical and Computer Engineering
University of Central Florida
Orlando, Fl. 32816

ABSTRACT

This paper presents a two-dimensional computer simulation of observed intensity and phase behind a time evolved phase screen. Both spatial and temporal statistics of the observed intensity is compared to theoretical predictions. In particular, the intensity statistics as a function of detector position within the propagated laser beam are investigated. The computer simulation program was written using the C-programming language running on a SUN SPARC-5 workstation.

KEYWORDS: Scattering, Laser Propagation, Phase Screen Simulation

2. INTRODUCTION

Computer simulations of random phase screens have been used for many years to study the effect of laser beam propagation through a random medium.[1,2,3,4,5] In each of these works, independent random phase screens were used for each time/space realization. Within each of the independent random phase screens, spatial correlation was introduced using a Kolmogorov power law spectrum. For example, Welch and Phillips[1] used approximately 100 independent random phase screens to simulate the affects of enhance backscattering. Mavroidis[2] *et. al.* used 2000 independent random phase screens to simulate the reconstruction of a coherent illuminated object after a double passage through a random phase screen. Solomon and Dainty[3,4] also used independent phase screens to simulate the reconstruction of an image after double passage through a random phase screen. In their work they used up to 100,000 time realizations.

None of the above mentioned simulations use both time and space correlated random phase screens. The exclusion of time correlated phase screens does not properly simulate the observed intensity fluctuations as a function of time and spatial position. Independent phase screen realizations can only simulate the phenomenon known as "speckle boiling" ignoring the affect of "speckle translation". The observed intensity fluctuations of a laser beam after propagating through a random media contains both effects. Brown[6] *et. al.* simulated time correlated phase screens by first generating a random phase screen that was larger than the optical beam size. The optical beam was then translated in position through this same phase screen simulating the affect of "speckle translation". Unfortunately, the number of time realizations using this approach is limited to the ratio of the phase screen size to the size of the beam incident on the phase screen.

This paper presents a simple method of incorporating temporally correlated screens that removes the limitations associated with Brown's work. Both space and time correlated phase screens are generated that include the effects of both speckle boiling and translation. Next, these phase screens are used to study the changes in laser beam intensity statistics as a function of the radial distance from the optic axis. The results of this simulation is then compared against the theoretical predictions presented by Miller[7] *et. al.* In Section 3, we describe the theoretical models used to generate both time and space correlated phase screens. We also include a short summary of the work presented by Miller[7] *et. al.* In Section 4, we show several results of our simulations.

0-8194-1824-2/95/$6.00

3 THEORETICAL MODELS

Figure 1 shows the simulated experimental setup used to propagate a laser beam through a random phase screen to a detector. The laser is modeled as a single mode Gaussian beam which is specified by its initial beam waist w_0, beam waist location (usually located at the exit of the laser), and output power. The Electric field exiting the laser is:

$$E(r,z) = E_0 \frac{\omega_0}{\omega(z)} e^{-\frac{r^2}{\omega^2(z)}} e^{-jkz} e^{-j\frac{kr^2}{2R(z)}} \tag{1}$$

where

$$\omega^2(z) = \omega_0^2\left[1+\left(\frac{z}{z_0}\right)^2\right] \quad : \quad R(z) = z\left[1+\left(\frac{z_0}{z}\right)^2\right] \quad : \quad z_0 = \text{Confocal parameter} \tag{2}$$

Propagation from the laser to the lens and then to the phase screen is accomplished using standard ABCD matrices. The ABCD matrices simply specifies a new radius of curvature and a new beam size at the phase screen.

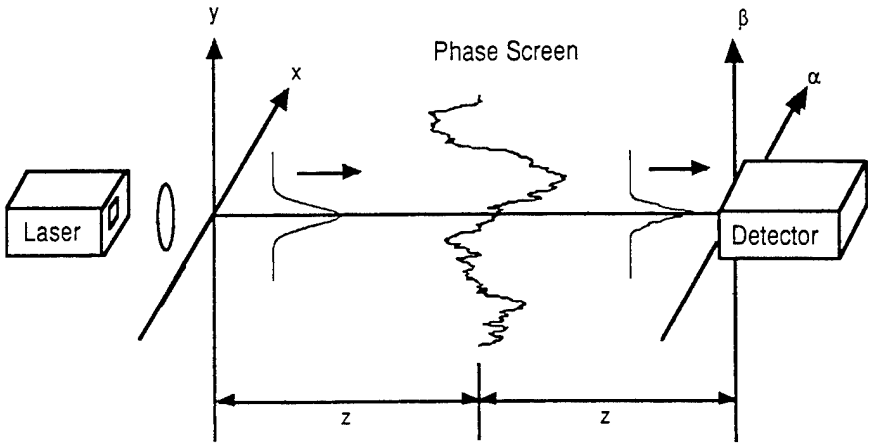


Figure 1. A block diagram describing the phase screen simulation.

Next, a random phase perturbation represented by a phase screen is applied to the incident laser beam. The phase screen is generated via a standard independent uniform random number generator used to fill two arrays with statistically independent random numbers between 0 and 1 and 0 and 2π. These arrays are is then used to compute a new array of statistically independent Gaussian numbers[8] with a zero mean and an variance σ^2

$$\text{Gaussian R.V.} = \left[-2\cdot\sigma^2\cdot\ln X\right]^{1/2}\cdot\cos\theta\ , \tag{3}$$

where θ is an uniformly distributed number between 0 and 2π. X is also an uniformly distributed number between 0 and 1 which is independent of θ. The uncorrelated random numbers generated from equation (3) produce random phase screens that are Gaussian distributed and that contain zero spatial correlation. Spatial correlation is applied to each phase screen realization (time sample) by using Fourier transform methods.[8] The Fourier transform is applied to the uncorrelated phase

screen data generated from equation (3) producing its Fourier components which are then filtered using the Kolmogorov power law :

$$\Phi_n(\kappa) = .033 C_n^2 \kappa^{-\frac{11}{3}} \tag{4}$$

$$\text{Where } \kappa = 2\pi \cdot \left[\frac{\alpha^2}{(N\Delta x)^2} + \frac{\beta^2}{(N\Delta y)^2} \right]^{1/2}, \tag{5}$$

The filter function used which depends on the thickness of the phase screen is:

$$T(\kappa) = \left[\frac{[2\pi]^3}{\lambda^2} \cdot \delta z \cdot .033 \cdot C_n^2 \cdot \kappa^{-\frac{11}{3}} \right]^{1/2}, \tag{6}$$

where C_n^2 is the strength of turbulence, κ is the two-dimensional spatial frequency, and δz is the thickness of the phase screen. The terms Δx and Δy are the spatial sampling intervals of the screen in the x and y directions, $\alpha/N\Delta x$ and $\beta/N\Delta y$ are the spatial frequencies in the x and y directions, and N is the total number of points used to describe the phase screen. The inverse Fourier transform is then applied to the newly filtered components yielding a spatially correlated phase screen.

Generating correlated random phases screens in the above manner makes it straightforward to create successive phase screens that are also correlated in time. First, time varying "speckle translation" is introduced by simply generating new phase screens as described in Equations (3)-(6). Instead of generating a completely new set of random numbers that are uniformly distributed and then applying Equations (3)-(6), the left most N-τ columns of the array of uniformly distributed numbers are shifted to the right by τ columns. The columns left blank by the shifting operation are then filled with a new set of uniformly distributed numbers that also vary between 0 and 1. Equations (3)-(6) are then used to produce a new spatially correlated random phase screen. This new phase screen will be a translated version of the previous phase screen realization. It also will contain some boiling due to the new columns of uniformly distributed numbers add at the left. The net effect is that the phase screen drifts across the incident laser beam for each time realization. The effect of "speckle boiling" is accomplished by replacing ρ columns which are randomly spaced throughout the array of uniformly distributed random numbers with new uniformly distributed numbers also between 0 and 1. This provides additional decorrelation of the phase screen for each time realization.

The third part of the simulation provides the free space propagation model used to propagate the random phase fluctuations introduced onto the laser beam by the phase screen. The free space propagation as shown in Figure 2 is based on the Fresnel diffraction integral and is given in transfer function form as[9]

$$U(\alpha,\beta) = e^{jkz}\, \Im^{-1}\left\{ \Im\{U(x,y)\} e^{-j\pi\lambda z\left[\frac{U^2}{(N\Delta x)^2} + \frac{V^2}{(N\Delta y)^2}\right]} \right\} \tag{7}$$

$$\text{Where} \quad \Im\{A(x,y)\} = \iint A(x,y) e^{-j2\pi(xU+yV)} dxdy \tag{8}$$

$$\Im^{-1}\{A(U,V)\} = \iint A(U,V) e^{j2\pi(xU+yV)} dUdV$$

This equation is valid for propagation distances that satisfy:[9]

$$z^3 >> \frac{\pi}{4\lambda}[(x-\alpha)^2+(y-\beta)^2]^2_{max} \tag{9}$$

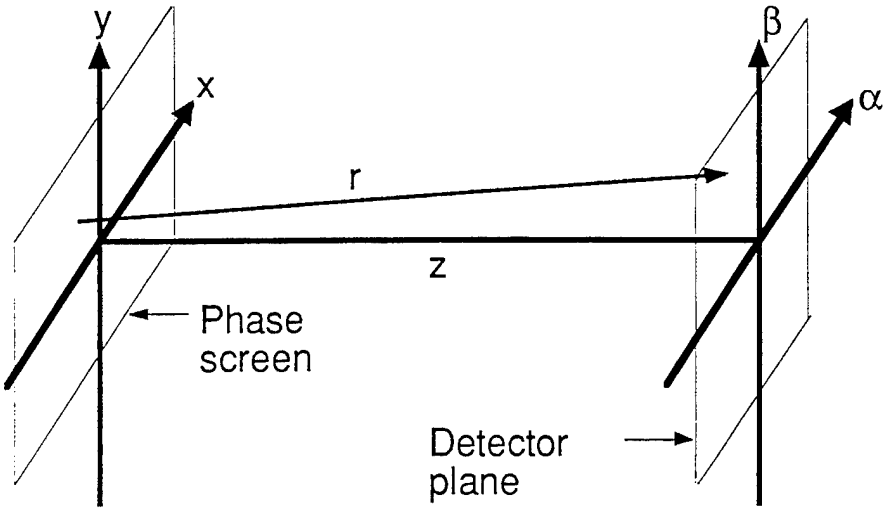


Figure 2. The geometry used to describe the propagation between two points.

The approach taken in this paper used the transfer function method over other techniques in that it provides the same sampling interval in the observation plane as in the z = 0 plane.

The last part of the simulation model simply uses a point detector with an aperture size equal to size of the sampling interval to obtain the observed intensity fluctuations. For the purpose of this, paper the placement of the detector was varied from the optical axis to a chosen radial distance. The work developed by Miller[7] *et. al.* gives the normalized variance of the intensity fluctuations as a function of radial distance r from the optic axis. In order to remove dependence of the refractive index structure constant C_n^2, the normalized off axis variance is divided by the normalized on axis variance. The theoretical equation for normalized off axis variance divided by normalized on axis variance is:[7]

$$\frac{\sigma^2(r)}{\sigma^2(0)} = 1 + \frac{4.42\cdot\Lambda^{5/6}\cdot\frac{r^2}{W^2}}{3.86\cdot\left\{.40\cdot\left[(1+2\Theta)^2+4\Lambda^2\right]^{5/12}\cdot\cos\left[\frac{5}{6}\tan^{-1}\left(\frac{1+2\Theta}{2\Lambda}\right)\right]-\frac{11}{16}\Lambda^{5/6}\right\}} \tag{10}$$

Where $\Theta = 1+\frac{L}{R}$, $\Lambda = \frac{2L}{kW^2}$, W and R are the waist and radius of curvature of the laser beam at the detector plane without a phase screen, k is the wavenumber, and L is the distance from the phase screen to the detector plane. This equation was generated with the assumption that the intensity statistics are in the weak turbulence regime.

4. RESULTS

Figure 1 gives the overall system layout used in the simulation experiments. Table 1 gives the corresponding parameters that were held constant throughout the simulations presented in this

paper. A collimating lens was placed in front of the exiting aperture of the laser to ensure that the laser beam incident on the phase screen had essentially a planer phase curvature. The value of $1.0 \cdot 10^{-13}$ $m^{-5/3}$ for the strength of turbulence was chosen so that the observed intensity statistics were in the weak turbulence regime. A distance of 13.4 m from the phase screen to the detector plane was the maximum allowable distance in which the simulation would be valid. For the given sampling interval, larger propagation distances produce aliasing during the computation of the inverse Fourier transform yielding invalid results.

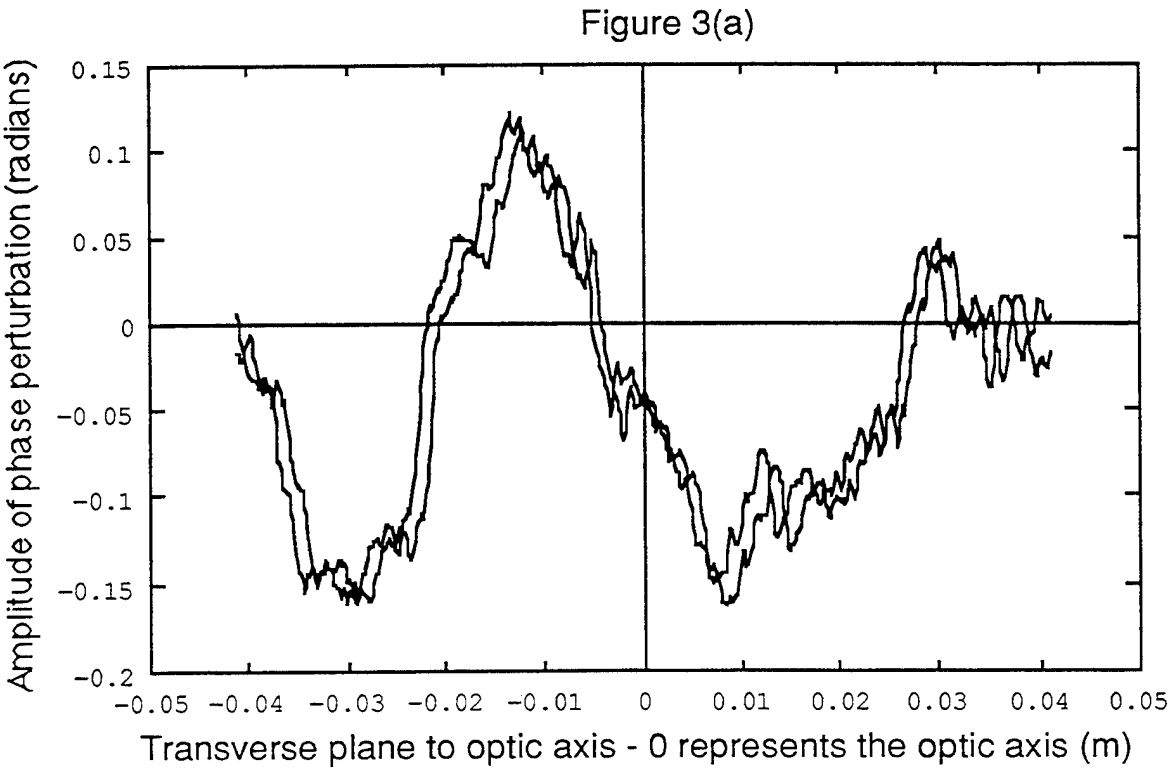
Parameters	Value
Wavelength λ =	$1.0 \cdot 10^{-6}$ m
Sampling interval $\Delta x = \Delta y = \Delta\alpha = \Delta\beta$ =	$1.61 \cdot 10^{-4}$ m
Beam waist at the exit of the lens	.00944 m
Radius of curvature at the exit of the lens	-7644.3 m
Laser output power	50 mW
C_n^2	$1.0 \cdot 10^{-13}$ $m^{-5/3}$
Thickness of the phase screen	.01 m
Distance from the laser to the phase screen	13.4 m
Distance from the phase screen to the detector	13.4 m

Table 1. A listing of the parameters used in the simulation experiments

Figures 3(a) and 3(b) show two different time correlated phase screens with the same Kolmogorov spatial power law spectrums. The two curves in each of Figures 3(a) and 3(b) show the spatial distribution of the phase fluctuations across the phase screen for two different time realizations. The phase screen in Figure 3(a) was generated by translating the arrays of uniformly distributed random numbers to the right by eight columns and then replacing the 8 left most columns in each array with a new set of uniformly distributed random numbers. The phase screen of Figure 3(b), which illustrates "speckle boiling", was generated by replacing 73 columns in each uniformly distributed array with a new set of uniformly distributed numbers. To further improve the "speckle boiling" effect, the actual location of each column replaced was also randomized.

Figure 4 shows a two dimensional intensity image of a Gaussian mode propagated through a phase screen with $C_n^2 = 1.0 \cdot 10^{-10}$ $m^{-5/3}$ and $\delta z = .01$ m. The image was generated by propagating the laser mode 13.4 m to the phase screen, and then propagating the corrupted laser beam 13.4 m to the detector plane. This figure shows the spatial intensity fluctuations that are present in the beam as a function of the phase perturbations induced by the phase screen. Figure 5(a) is a one-dimensional plot of the intensity image in Figure 4 taken as a function of horizontal position and centered in the vertical direction. This plot shows the deviation from a pure Gaussian mode. Figure 5(b) shows another one-dimensional-plot of intensity as a function of horizontal position for a later time realization of a translating phase screen with a translation to the right by 5 columns. Note, the translation of the intensity fluctuations have moved to the right by $5\Delta x$ in Figure 5(b) as compare to Figure 5(a).

The goal of the this paper is to compare the intensity statistics (normalized intensity variance) as a function of radial distance from the optical axis as described by Equation (10). Six hundred successive time realizations were generated using independent uncorrelated time and translation only phase screens. For each simulation, a sequence of intensity fluctuations as a function of time was acquired at a given spatial detector position r. The normalized variance was then computed from the time sequence using:



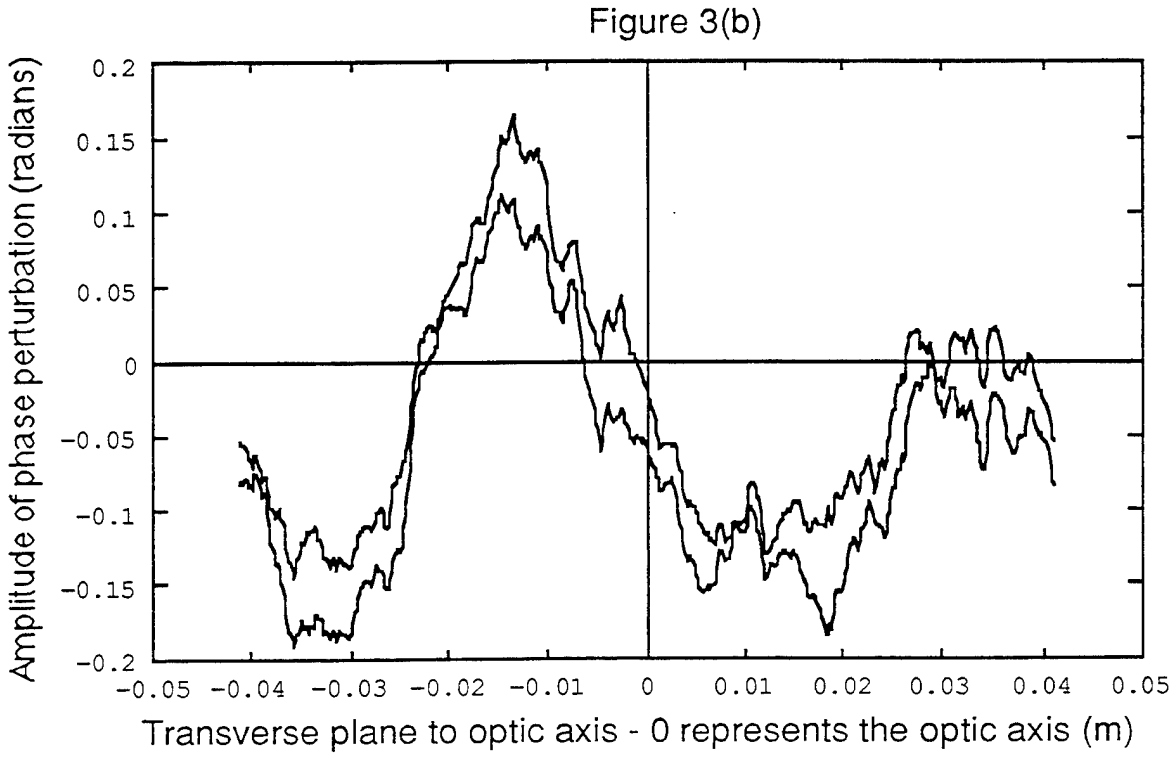


Figure 3. Two different phase screen realizations showing (a) translation and (b) boiling.

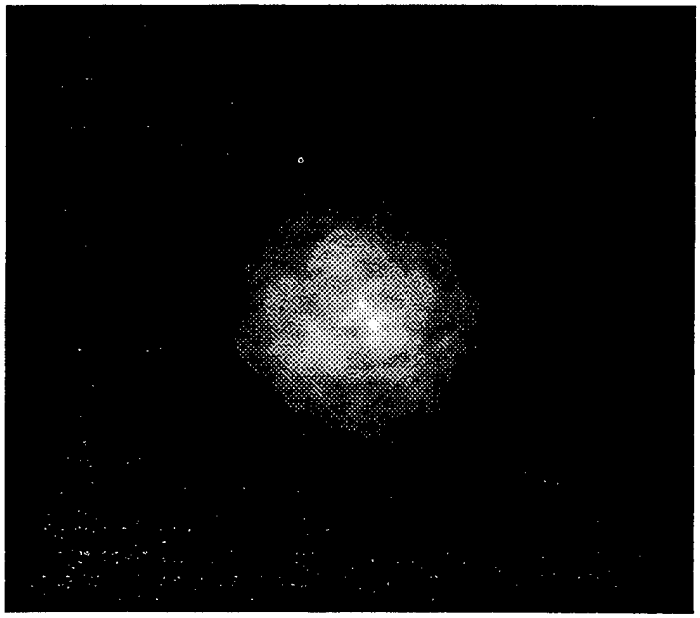


Figure 4. One time realization of the observed intensity at the detector plane.

$$\sigma^2(r) = \frac{\frac{1}{\beta}\sum_{n=1}^{\beta}[I(n,r) - \text{mean}(r)]^2}{[\text{mean}(r)]^2} \tag{11}$$

and

$$\text{mean}(r) = \frac{1}{\beta}\sum_{n=1}^{\beta} I(n,r)$$

where

β is the total number of time realizations (600),

r is the distance from optic axis of the detector, and

n is the nth time realization.

Next, the normalized off axis variance is divided by the normalized on axis variance for comparison with Equation (10).

$$\Omega(r)^2 = \frac{\sigma(r)^2}{\sigma(0)^2} \tag{12}$$

Figures 6(a) and 6(b) gives the simulation results for the normalized variance and the theoretical results computed from Equation (10) as a function of radial position from the optical axis. The only difference between the two figures was the correlation of the phase screens between each time realization. The simulation results of Figure 6(a) was generated using only translation phase screens, with a translation to the right by 3 columns. Figure 6(b) was generated using independent uncorrelated time phase screens. The parameters used in equation (10) were calculated as L = 13.4m, R=4776.4 m and W=.00945 m from the parameter values given in Table 1 and with a confocal parameter z_0 of 279.3 m.

Figure 5(a)

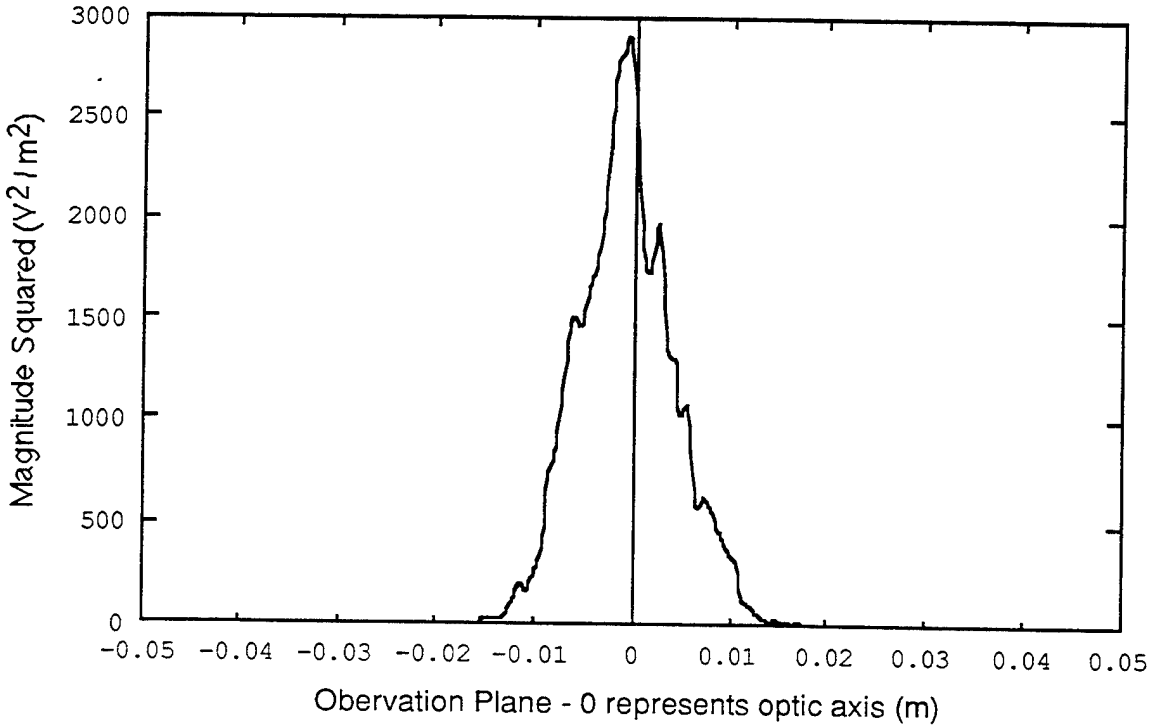


Figure 5(b)

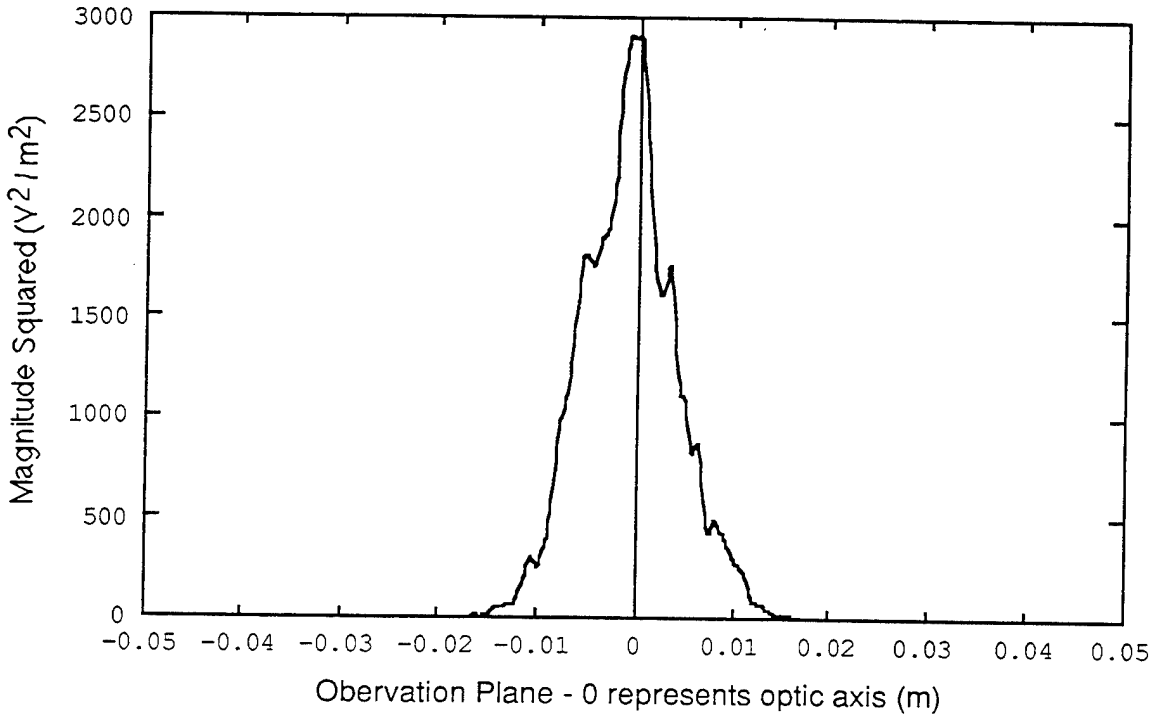


Figure 5. Two successive time realizations of horizontal intensity profiles.

Figure 6(a)

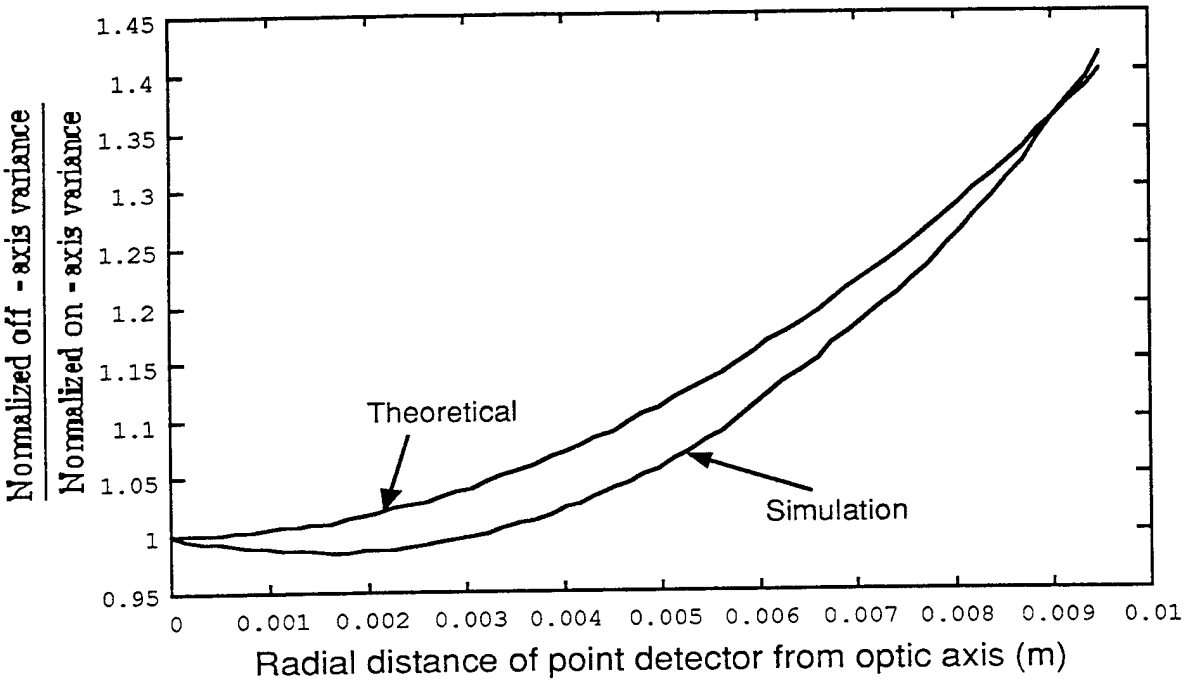


Figure 6(b)

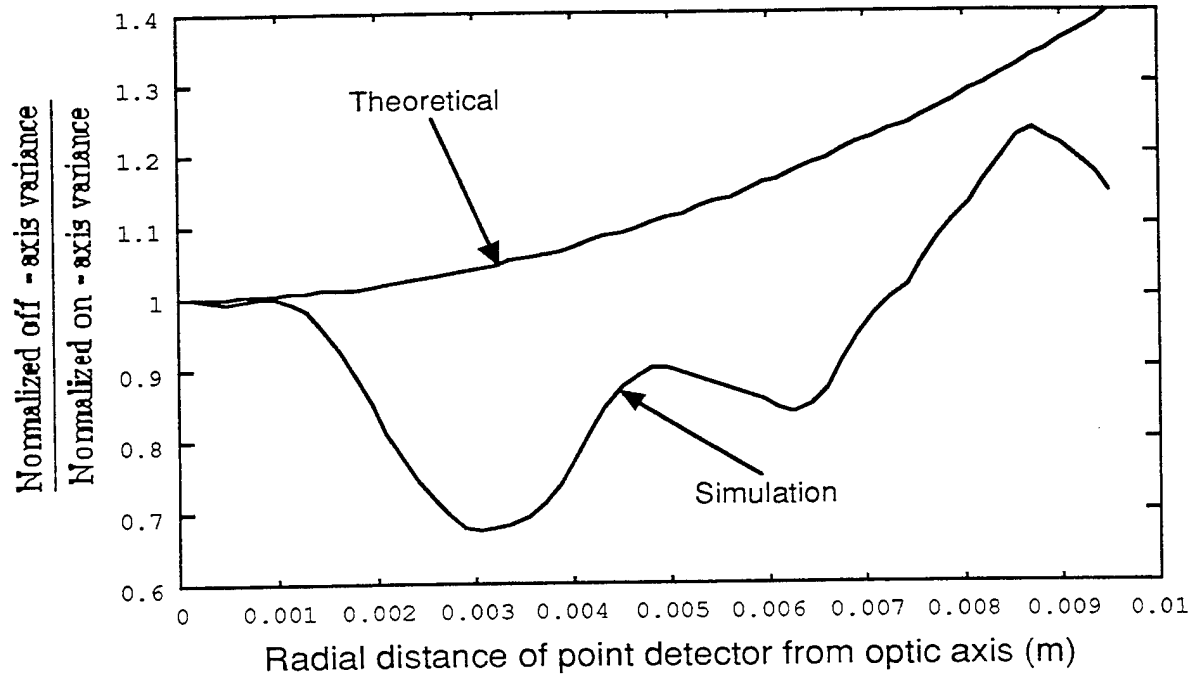


Figure 6. Theoretical and simulated results for the ratio of the off axis normalized variance to the on axis normalized variance for (a) time translated phase screens and (b) temporally uncorrelated phase screens.

The comparison of the simulated and experimental curves in Figure 6(a) shows a close match of the simulation with the theoretical prediction given by Equation (10). For a radial distance of approximately .00915 m, the two curves intersect. This is very close to the diffractive edge of the laser beam at that point, W=.00945 m. The results of Figure 6(b) shows a quite a difference between the theoretical prediction and the simulated results. For radial distances beyond approximately .001 m, the simulated curve lies completely below the theoretical curve.

5. CONCLUSIONS

Based on the data presented in Figures 6(a) and 6(b), the presence of 'speckle translation' is required to better predict the theoretical calculated values for normalized variance as described by Miller *et. al.*[7] The use of independent uncorrelated phase screens does not properly predict the intensity statistics for successive time realizations. During the process of performing these simulations, several trends were noticed. For C_n^2 sufficiently small, the simulated results were independent of C_n^2. As C_n^2 was increased, so that the weak scattering approximation was violated, the ratio of the off-axis normalized variance divided by normalized on axis variance for the temporally correlated phase screens began to dip lower and lower below the minimum theoretical value of 1. As the radial distance increased, this ratio did increase to values greater than 1, but grew more slowly than the curves with smaller values of C_n^2.

6. ACKNOWLEDGMENTS

This research has been supported through the Center for Research and Education in Optics and Lasers (CREOL) under support by the Strategic Defense Initiative Organization, Innovative Science and Technology Office through the Office of Naval Research under contract N00014-92-C-0123.

7. REFERENCES

1. Gisele Welch and Ronald Phillips, "Simulation of enhanced backscatter by a phase screen," Journal of Optical Society of America-A, Vol. 7, Number 4, pp. 578-583 , April,1990.

2 T. Mavroidis, C. J. Solomon, and J. C. Dainty, "Imaging a coherently illuminated object after double passage through a random screen," Journal of Optical Society of America-A, Vol. 8 Number 7 pp. 1385-1390, July, 1991.

3 C. J. Solomon and J. C. Dainty, "Imaging a coherently illuminated object through a random screen using a dilute aperture," Journal of Optical Society of America-A, Vol. 9, Number 8, pp. 1385-1390, August , 1992.

4. C. J. Solomon and J. C. Dainty, "Use of polarisation in double passage imaging through a random screen," Optics Communications, Vol. 87, pp. 207-211, February, 1992.

5. Jeffrey L. Turner, "Simulation of Optical Propagation Through Atmospheric Turbulence Using Two-Dimensional Fourier Transform Techniques," Masters Thesis Naval Postgraduate School pp. 24-26, June 1989.

6. W. P. Brown, C. Fry, G. C. Valley, "Simulation of Ground-To-Space Optical Propagation," Hughes Research Laboratories Final Technical Report RADC-TR-82-91, pp 1-24, May 1982.

7. W. B. Miller, J. C. Ricklin, L. C. Andrews, "Effects of the refractive index spectral model on the irradiance variance of a Gaussian beam," Journal of Optical Society of America A/ Volume 11, Number 10, pp. 2719-2726, October 1994.

8. A. R. Weeks, H. R. Myler, and H. Wenaas, "Computer generated noise images for the evaluation of image processing algorithms," Optical Engineering, Vol. 32, Number 5, pp.982-992, May 1993.

9. J. W. Goodman, Introduction to Fourier Optics, pp 30-70, McGraw-Hill, 1988.

5.0 SINGLE APERTURE COHERENT LIDAR EXPERIMENTS

C. Martin Stickley, Ph.D. and P. Gatt, Ph.D.

5.1 Background

In addition to understanding and developing multiple aperture coherent lidars, it was also an objective of the program to better understand the effect of speckle on lidar performance and to test lidars in real-world situations (i.e., outside of a university laboratory).

The motivation to understand speckle effects stems from the fact that lidars can be built having a 10x difference in frequency - 1.06 micron YAG systems and 10.6 micron CO_2 systems - and that they differed significantly in performance [1] which could be related to whether the speckle was partially or fully developed. "Partially" developed speckle can originate from a target whose surface irregularities are smaller than a wavelength; one would expect "fully" developed speckle in cases where target surface roughness is larger than a wavelength. Thus YAG lasers and CO_2 lasers might perform quite differently.

The motivation to test coherent lidar in a non-laboratory environment was to expose our students to real-world outdoor test problems and thereby to check our receiver system in other than a pristine laboratory.

5.2 Accomplishments

5.2.1 Rotating Target Experiments with YAG and CO_2 Coherent Lasers

During the course of this contract we carried out cw coherent lidar experiments [2] at 1.06 microns with a single frequency Nd:YAG laser and at 10.6 microns with a single frequency CO_2 laser using the same target - a common nail - rotating at 10 or 20 rps, and located 2 meters away from the exit aperture of the table top lidar. Both lidars upshifted the transmitted beam by 27.1 MHz relative to the local oscillator; thus the IF signals were at 27.1 MHz.

Figures 5.1 and 5.2 give the power spectrum of the IF signal for the 1.06 micron system and the 10.6 micron signal, respectively.

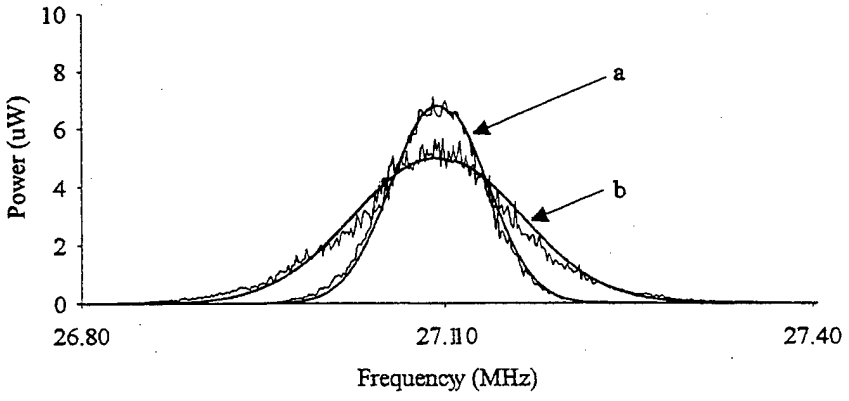


Fig. 5.1: 1.06 micron power spectrum a) 10 Hz; b) 20 Hz, with superimposed Gaussian fits, for 2 meter range to target.

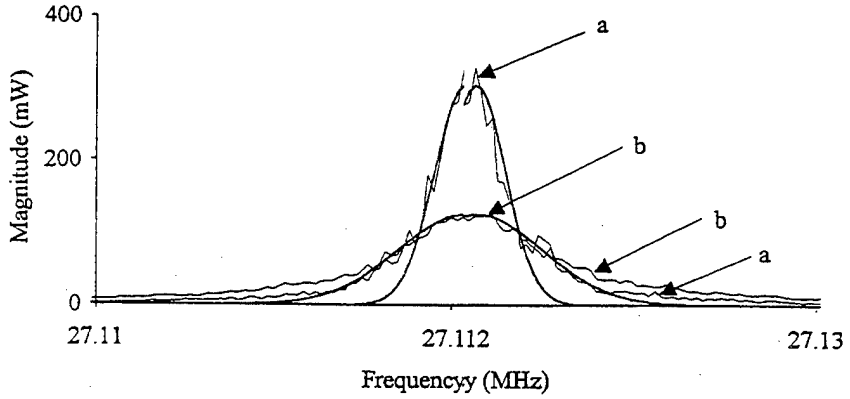


Fig 5.2: 10.6 micron power spectrum a) 10Hz; b) 20 Hz, for 2 meter range to target with superimposed Gaussian fits for the narrow band components.

First, the power spectra in Fig. 5.1 are entirely normal, that is, they are Gaussian in shape, as our theory says they should be [2]; the peak amplitude of the signal drops appropriately at 20 rps due to conservation of energy; and the widths of the spectra are in approximately agreement with our predicted widths.

The situation is different for the CO_2 laser, however: note that the wings of the data do not follow the expected Gaussian curve. This is amplified in Fig. 5.3.

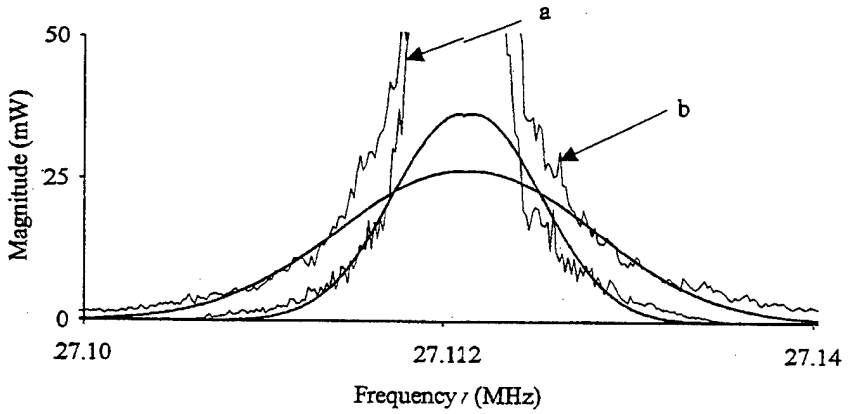


Fig. 5.3: Enlarged 10.6 micron power spectrum a) 10 Hz; b) 20 Hz, for a 2 meter target with superimposed Gaussian fits for the broadband components.

A close study of Fig 5.3 shows that the "wings" of the data can fit Gaussian curves; note these "wings" are broader in frequency than the data in the central lobes.

Further study of the data showed that the spectral width of the "wings" agreed very well with the simple theory which explained the 1.06 micron data so well, whereas the narrow band, large amplitude component was 5x narrower than expected from the theory.

This can be explained by noting that the target is more specular at 10 microns, and what we are observing in the central lobes is the "glint" or specular reflection from the target.

Such a glint is expected to have a narrow frequency width since it originates from a small surface of the target, and this frequency width will be smaller since the Doppler shift from it is proportional to the width of the

glint area. This spectral width should be rotation-rate dependent, and, indeed, this is oberved: 2 kHz at 10 rps and 4.5 kHz at 20 rps.

In conclusion, the "partially" developed speckle showed up in our experiments as a glint from the target's surface. While this was unexpected, it is entirely rational that it should happen that way. Thus, the lesson from this work is that coherent CO_2 laser radars will be affected by glints which can mask the more subtle signal which carries the rotation rate information expected with this type of lidar (i.e., a Doppler width which is based on the diameter of the target). On the other hand, strong glints can be useful in that they aid in detecting the presence of perhaps an otherwise unobservable target.

5.2.2 Vibration Sensing with a CW Coherent Lidar

CREOL's development of a receiver for coherent lidar prompted a collaboration with Schwartz Electro-Optics, Inc. to combine CREOL's receiver with SEO's eye-safe laser and make a cw Doppler lidar for sensing vehicle vibrations [3]. Fig. 5.4 is a schematic drawing of the system. The eye-safe laser is a diode-pumped holmium laser developed by SEO and producing 100 mW. While an intra-cavity etalon was used to alter the laser's frequency, no strong effort was made to make it oscillate in a single mode. Like the other CREOL lidars, the transmitted beam was upshifted by 27.1 MHz. The receiver used an InGaAs, 100 MHz detector packaged by CREOL.

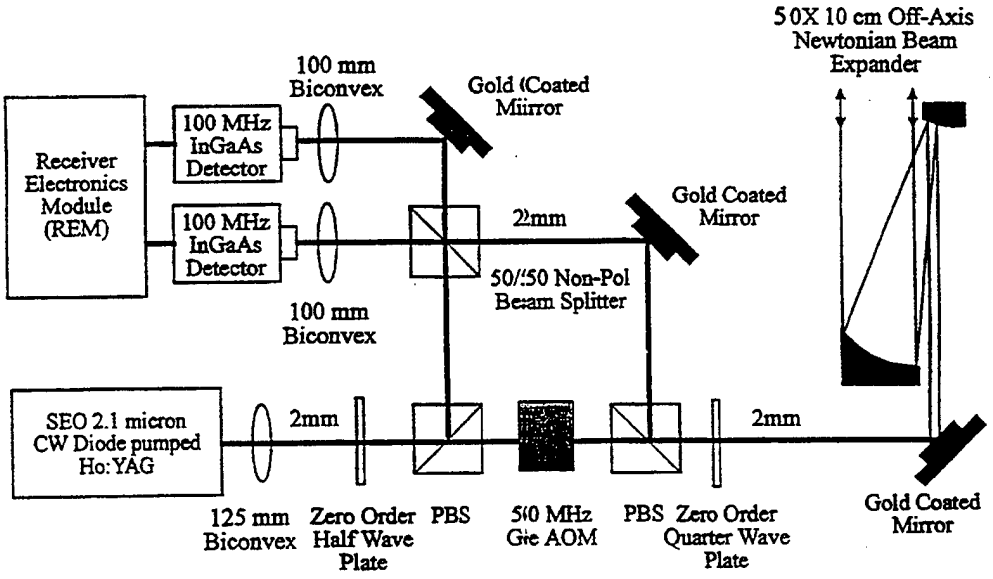


Fig. 5.4: 2 Micron Laser Doppler Radar Experimental Set-up

Calibration of the system was achieved in the laboratory by using a piezoelectric loudspeaker as a target. Results of the calibration indicated that vibration amplitudes as small as 10 nanometers could be sensed.

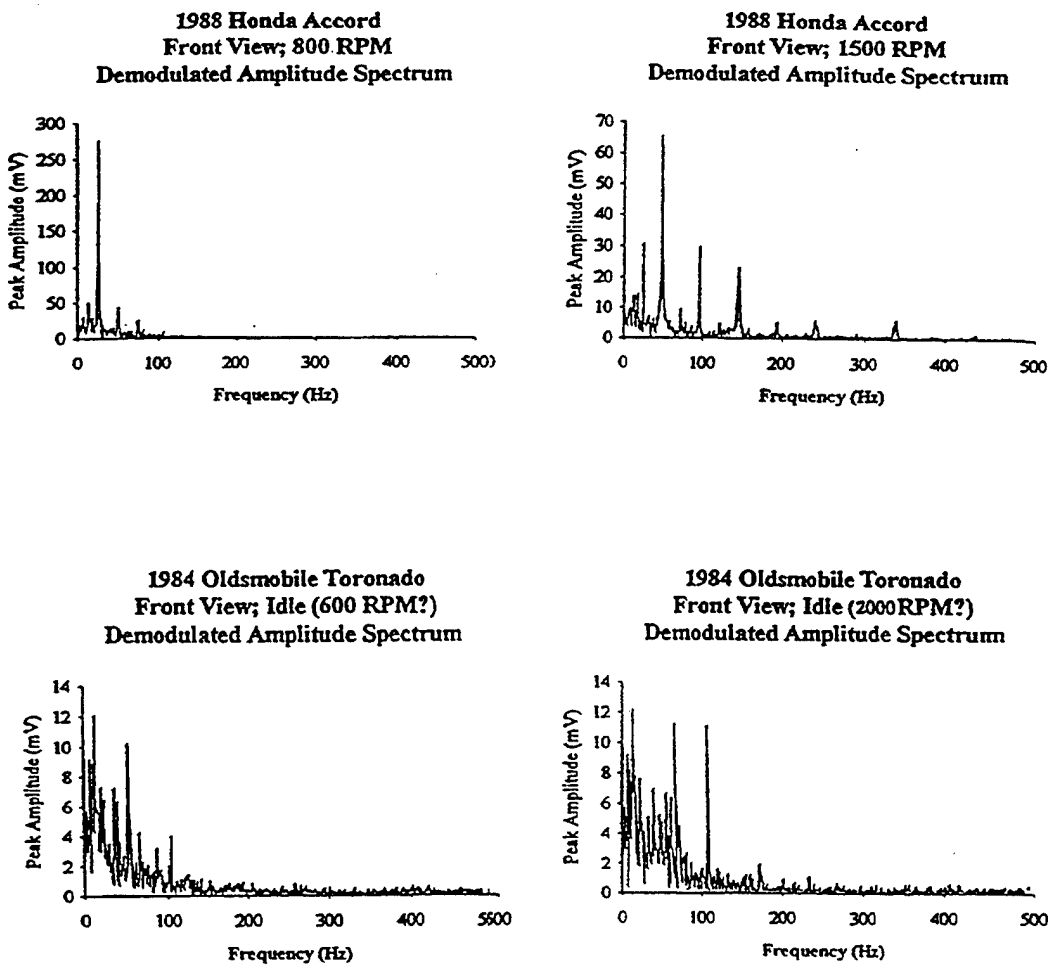


Fig. 5.5: Vibration data illustrating non-cooperative target identification.

Fig. 5.5 is an example of the data which was obtained using two different automobiles as targets. The major peaks correspond to the number of combustion detonations per second or the number of combustion cycles per crankshaft revolution. Note that for the Honda, the number of combustion cycles per revolution is two, thus at 800 RPM, the frequency should be 26.7 Hz, in very close agreement with location of the peak near 25 Hz. For the Oldsmobile, it has four combustion cycles per crankshaft revolution, thus at 600 RPM, it has a peak near 40 Hz. It was observed that while these peaks moved around with RPM, they became much larger upon passing through a structural resonance of the automobile.

In conclusion, CREOL, in conjunction with Schwartz Electro-Optics, demonstrated that this coherent, eye-safe laser radar technology can be used for non-cooperative vehicle identification in a real-world environment.

5.3 References

1. Private communication, R. Mongeon.

2. P. Gatt, T. Costello, C. Vogt, and C. M. Stickley, "Laser radar spectrum of a rotating random rough surface", Proc. SPIE 1936, 165-179, (1993).

3. M. Acharekar, P. Gatt and L. Mizerka, "Laser Vibration Sensor", Proc. SPIE 2472, 2-11 (1995).

5.4 Publications, Presentations and Theses

1. P. Gatt, T. Costello, C. Vogt, and C. M. Stickley, "Laser radar spectrum of a rotating random rough surface", Proc. SPIE 1936, 165-179, (1993).

2. M. Acharekar, P. Gatt and L. Mizerka, "Laser Vibration Sensor", Proc. SPIE 2472, 2-11 (1995).

3. P. Gatt, M. Acharekar, L. Mizurka, and D. Heimmermann, "Two-micron Laser Vibration Sensor", LASERS '94, Quebec, Quebec, Canada, December 1994.

4. C. M. Stickley, P. Gatt, T. Costello and C. L. Vogt, "Coherent Laser Radar Spectrum of a Rough Rotating Cylindrical Surface", presented at the 7th Conference on Coherent Laser Radar, Paris, France, 1993.

5.5 Appendix 5A

P. Gatt, T. Costello, C. Vogt, and C. M. Stickley, "Laser radar spectrum of a rotating random rough surface", Proc. SPIE 1936, 165-179, (1993).

Laser radar spectrum of a rotating random rough surface

Philip Gatt, Tom P. Costello, Chie L. Vogt and C.M. Stickley

Center for Research in Electro-Optics and Lasers

University of Central Florida

12424 Research Parkway, Orlando, FL 32826

407-658-6800

ABSTRACT

A theoretical model is presented for the mutual coherence function that describes the power spectrum of the intermediate frequency (IF) signal resulting from coherent laser radar detection of a fully illuminated rough rotating cylinder. Preliminary calculations for the correlation length and bandwidths of the coherently detected dynamic speckle fields show that Doppler effects on speckle dynamics are on the same order of magnitude as speckle translation effects, and that scatterer exchange effects are negligible. We show that for a fully diffuse object, the shape of the spectrum (and autocorrelation function) is Gaussian. The spectrum appears to be the sum of two Gaussian functions when the scattering object has a near-specular component, one due to the specularly reflected field and the other due to the diffusely reflected field. While the IF BW of the signal resulting from the diffusely scattered field depends upon the cylinder's radius, the IF BW of the specularly reflected field depends upon the width of a window function, which is wavelength dependent.

1. INTRODUCTION

IF power spectra from coherently detected moving objects are known to be broadened by amplitude modulation (AM) and phase modulation (PM). The AM and PM effects are produced by dynamic laser speckle phenomena. Laser speckle is the random phase and amplitude variation of the reflected optical field from a coherently illuminated rough surface.

Static speckle is the interference pattern resulting from the phasor summation of the temporally coherent (narrowband) light reflected from many random scatterers on a static object surface. The individual scattered fields have random amplitudes and phases. The resulting interference pattern constitutes speckle and is usually characterized by its spatial correlation function.

Dynamic speckle is a spatio-temporal phenomenon, resulting from relative motion between the object, the transmitter, and the observation point. Dynamic speckle may be characterized by the field mutual coherence function (MCF), which describes the temporal correlation of the fields of two spatially separated points. The temporal characteristics of the speckle field at a single point in the observation plane may be of interest, however, as in the case of a laser radar system with a single detector. In this case, the field temporal autocorrelation function (ACF) provides the necessary information. In addition, the power spectral density

0-8194-1172-8/93/$6.00

function (PSDF), which is the Fourier transform of the ACF, alternatively provides the necessary information.

Dynamic speckle manifests itself in two distinct phenomena: "speckle translation" and "speckle boiling". Speckle translation is the transverse motion of an otherwise static speckle field across the observation plane. Speckle boiling describes the evolution (emergence and disappearance) of individual speckle cells. Various object dynamics cause speckle translation and boiling.

Object dynamics that cause speckle translation include lateral motion of the object, the illuminating beam, and the detector. These dynamics and object rotation cause scatterer exchange, which is a source of speckle boiling. Doppler effects from object rotation cause speckle boiling as well. The relative strengths of the translation and boiling phenomena in a given speckle field depend heavily on the illumination function, the system geometry, and the type of object motion.

In this paper we report some of our early findings regarding the influence of a plane wave illuminated rotating cylindrical object on the shape of the PSDF and ACF of a heterodyne-detected speckle field. We compare and contrast the effects of speckle boiling and the speckle translation.

Experimental measurements were conducted at two wavelengths, one and ten microns. These diverse wavelengths provide a comparison of speckle dynamics for an object which appears fully diffuse at the shorter wavelength yet has a prominent specular component at the longer wavelength.

2. SYSTEM DESCRIPTION

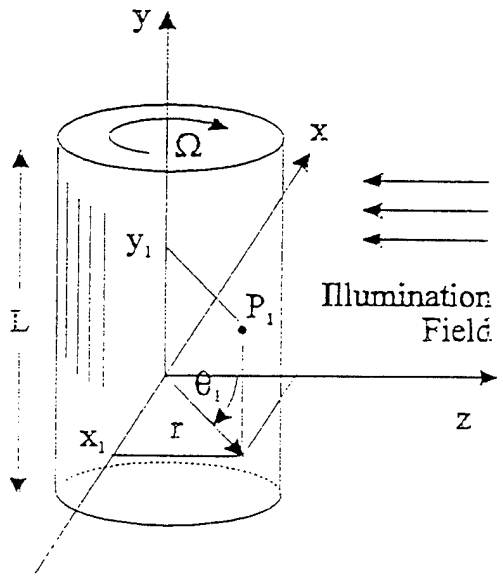


Figure 1. Rotating rough cylinder under illumination along the z axis

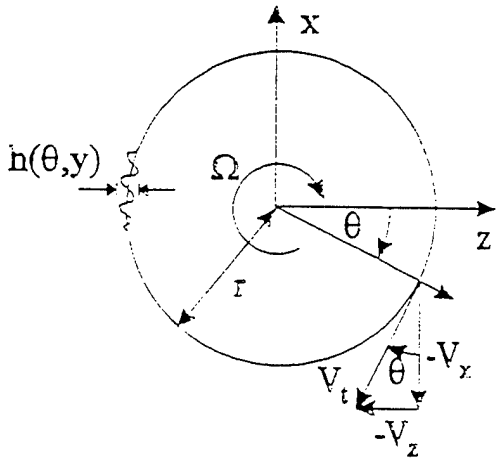


Figure 2. Cross section of the rotating rough cylinder

The geometry of the specific case under consideration is depicted by Figures 1 and 2. A cylinder of radius r and length L is rotating about its longitudinal axis (the y axis) at an angular rate Ω. Collimated plane wave coherent illumination from a monostatic ladar transceiver fully illuminates the unresolved surface of the cylinder along the line of sight (the z axis) transverse to the cylinder's axis. The surface height fluctuations are described by the zero mean random variable $h(\theta,y)$, which is superimposed on the cylinder's surface.

3. THEORETICAL PREDICTIONS

We consider two approaches to the theoretical characterization of the IF PSDF. The first is a statistical model, which when solved will result in a complete description of the PSDF. The second approach is based on a heuristic analysis, which serves to provide an approximate description of the PSDF.

3.1. Statistical Model

The statistical model is based on the Fourier transform relation between the PSDF and the ACF (a.k.a. self coherence function (SCF)), which is found by evaluating the mutual coherence function (MCF) at zero spatial separation distance. The mutual coherence function (MCF) is a common descriptor of the spatio-temporal properties of a random process. The MCF at the receiver is derived via a Fresnel-Fraunhofer propagation of the MCF of the reflected field just above the rough object surface. The reflected field is a function of the object surface characteristics as well as the illuminating beam characteristics.

In general, the MCF of an analytic signal is given by[1]

$$\Gamma_U(P_1, P_2; \tau) = \left\langle \tilde{U}(P_1; t+\tau)\, \tilde{U}^*(P_2; t) \right\rangle, \tag{1}$$

which is the expected value of the product of the complex field at two spatially separated points at two instances in time. $\tilde{U}$ represents the analytic signal of a single polarization of the electric field immediately above the object surface. This field is the product of the incident illumination with the object diffuse reflectance and an object induced phase function. The incident illumination is assumed to be a plane wave, given by

$$\tilde{U}(P;t) = A \exp[-j(2\pi f t + \phi)], \tag{2}$$

where A is the constant plane wave field amplitude, f is the optical frequency, and ϕ is the phase (assumed to be zero but shown here for completeness).

The total additional phase of the field caused by the object, just after reflection, contains three components. These components are related to the surface macro shape, the Doppler phenomena, and the random surface height.

$$\phi(P;t) = \phi_{sur}(P;t) + \phi_d(P;t) + \phi_h(P;t), \tag{3}$$

where,

$$\phi_{sur}(P;t) = [4\pi/\lambda]\, r\cos(\theta), \tag{4}$$

$$\phi_d(P;t) = -\left[8\pi^2/\lambda\right] r f_{rot} t \sin(\theta), \tag{5}$$

$$\phi_h(P;t) = \left[4\pi/\lambda\right] h(\theta, y; t)\cos(\theta), \tag{6}$$

and f_{rot} is the target rotation frequency measured in Hz. Substituting appropriate expressions for each of the components into Eq. (1) yields the following formidable expression for the MCF at the object.

$$\begin{aligned}
\Gamma_U(P_1, P_2; \tau) &= \left\langle \tilde{U}(P_1; t+\tau)\, \tilde{U}^*(P_2; t) \right\rangle \qquad (7)\\
&= \left\langle A \exp[-j2\pi f(t+\tau)] A^* \exp[j2\pi f t] \right\rangle && \text{Incident Illumination}\\
&\times \left\langle R(P_1; t+\tau) R^*(P_2; t) \right\rangle && \text{Target Reflectance}\\
&\times \left\langle \exp\left[\frac{j4\pi r}{\lambda}\right][\cos(\theta_2) - \cos(\theta_1)] \right\rangle && \text{Surface Curvature}\\
&\times \left\langle \exp\left[\frac{j8\pi^2 r f_{rot}}{\lambda}\right][(t+\tau)\sin(\theta_1) - t\sin(\theta_2)] \right\rangle && \text{Doppler Shift}\\
&\times \left\langle \exp\left[\frac{j4\pi}{\lambda}\right][\cos(\theta_2) h(\theta_2 + \Omega t, y_2) \right. && \text{Surface Height}\\
&\qquad \left. -\cos(\theta_1) h(\theta_1 + \Omega(t+\tau), y_1)] \right\rangle && \text{Deviations}
\end{aligned}$$

As previously stated, the MCF of the field in the receiver plane is found by a diffraction propagation of the MCF in the target plane. Goodman[1] shows that this is a fourth order spatial integration, which, for some special configurations, is given by

$$\begin{aligned}
\Gamma_U(Q_1, Q_2; \tau) \approx \frac{1}{(\lambda z)^2} \iint_{\Sigma_1} \iint_{\Sigma_2} & dx_{P_1} dy_{P_1} dx_{P_2} dy_{P_2}\ \Gamma(P_1, P_2; \tau)\\
& \times \exp\left(\frac{j2\pi}{\lambda z} (x_{Q_2} x_{P_2} + y_{Q_2} y_{P_2} - x_{Q_1} x_{P_1} - y_{Q_1} y_{P_1}) \right).
\end{aligned} \tag{8}$$

The ACF (or self coherence function (SCF)) is obtained by evaluating the propagated MCF at $Q_1 = Q_2$.

$$\begin{aligned}
\Gamma_U(\tau) = \Gamma_U(Q_1, Q_1; \tau) \approx \frac{1}{(\lambda z)^2} \iint_{\Sigma_1} \iint_{\Sigma_2} & dx_{P_1} dy_{P_1} dx_{P_2} dy_{P_2}\ \Gamma(P_1, P_2; \tau)\\
& \times \exp\left(\frac{j2\pi}{\lambda z} (x_{Q_1}(x_{P_2} - x_{P_1}) + y_{Q_1}(y_{P_2} - y_{P_1})) \right).
\end{aligned} \tag{9}$$

Finally, according to the Wiener-Khinchine theorem, the temporal power spectrum is obtained via a Fourier transform,

$$S_\Gamma(f) = \mathcal{F}[\Gamma_U(\tau)]. \tag{10}$$

An exact solution, however, is intractable due to the spatial non-stationarity of the phase term, specifically the $\cos(\theta_i)h(\theta_i+\Omega t_i;y_i)$ term, in the mutual coherence function (cf., Eq. (7)). Conceptually, this non-stationarity is due to the spatial dependence of the projected height function along the line of sight. For example, an on-axis point (i.e., $\theta = 0$) with a height deviation h imposes a $4\pi h/\lambda$ phase deviation on the reflected field, while that same point when rotated 90 degrees has no impact on the phase of the reflected field.

In order to obtain a solution based upon the above statistical analysis, simplifying assumptions which are not too restrictive to erode the accuracy of the result must be imposed. At a minimum these assumptions must reduce the non-stationary phase term to a stationary phase term by ignoring the $\cos(\theta)$ dependence. The impact of such a simplification, however, has not yet been fully analyzed.

3.2. Heuristic Analysis

A solution based on the statistical analysis given above, would yield the complete description of the PSDF. Even if appropriate simplifying assumptions are imposed, however, the mathematics will still be quite formidable and physical insight may be difficult to obtain from the solution. An alternative approach based upon a heuristic analysis of the problem, though not yielding the complete description of the PSDF, will provide a measure of the correlation time and the bandwidth of the spectrum.

Our heuristic analysis is based on the assumption that the translation and boiling effects have separable Gaussian PSDF's which, when convolved, produce a Gaussian composite PSDF. Wang [2] predicted time domain correlation functions for a similar though distinct problem. In his analysis the object was a stationary flat surface and the coherent transceiver was translated or scanned over the object. He showed that the resultant autocorrelation function is the product of two Gaussian functions respectively attributed to speckle translation and speckle boiling effects. In the frequency domain, then, corresponding individual Gaussian PSDF's would also be attributable to the speckle dynamics; and the composite PSDF, which is the convolution of the translation and boiling PSDF's is also Gaussian.

We will attempt to predict each component PSDF bandwidth by estimating a correlation time associated with the physical mechanisms responsible for the translation and boiling effects, then relate the correlation time to a bandwidth using the definitions described in the next section. The bandwidth of the composite PSDF, which is measurable, then should be obtainable from the component PSDF's assuming they are Gaussian.

3.2.1. Definition of Bandwidth and Correlation Time

Several definitions of bandwidth (BW) and correlation time (τ_c) are found in the literature, and the relationship between the two always depends on the shape of the PSDF or ACF. Wang [2] suggests that the PSDF and the ACF for dynamic speckle are Gaussian. A normalized (to unit total power) Gaussian PSDF may be written as

$$\hat{S}_V(f) = \frac{2\sqrt{ln2}}{\sqrt{\pi}\Delta f} exp\left[-\left[\left(2\sqrt{ln2}\right)\frac{f-\bar{f}}{\Delta f} \right]^2 \right]. \qquad (11)$$

where Δf is the full width half maximum (FWHM) BW. Using Eq. (10), the corresponding normalized ACF function is easily determined to be

$$\hat{\Gamma}_U(\tau) = \exp\left[-\left(\frac{\pi\Delta f\tau}{2\sqrt{\ln 2}}\right)^2\right]\exp(-j2\pi\bar{f}\tau). \tag{12}$$

Of the many definitions of correlation time (e.g., the 1/e point of a Gaussian ACF, the first zero in a Sinc ACF, etc.), Goodman's[3] is most common for arbitrary correlation functions and is given by

$$\tau_c \equiv \int_{-\infty}^{\infty}\left|\hat{\gamma}_U(\tau)\right|^2 d\tau. \tag{13}$$

The function $\hat{\gamma}_U(\tau)$ is the magnitude of the normalized ACF, $\hat{\Gamma}_U(\tau)$, and is given by

$$\hat{\gamma}_U(\tau) = \left|\hat{\Gamma}_U(\tau)\right| = \exp\left[-\left(\frac{\pi\Delta f\tau}{2\sqrt{\ln 2}}\right)^2\right]. \tag{14}$$

Using Goodman's definition, we arrive at the following correlation time associated with a random process which has a Gaussian PSDF, with FWHM bandwidth Δf:

$$\tau_c = \frac{1}{\Delta f}\sqrt{\frac{2\ln 2}{\pi}} = \frac{0.664}{\Delta f}. \tag{15}$$

It is interesting to note that this definition of correlation time yields $\hat{\gamma}_U(\tau_c) = \exp(-\pi/2)$ (or 0.21), as contrasted with the commonly used exp(-1) (or 0.37) point on the normalized autocorrelation function.

In the following sections, we discuss the effects of three object dynamics associated with a rotating cylinder on the BW and correlation time of the coherently detected field.

3.2.2. Effect of Object Rotation on Speckle Translation (Rotation-induced speckle translation)

As the object rotates with angular velocity Ω, the speckle field rotates at an angular velocity of 2Ω about its longitudinal object axis. These rotating speckles induce an amplitude and phase modulation on the coherently detected IF signal. These modulations result in a degradation of the temporal correlation of the IF signal which can be characterized by the correlation time as described above. At the receiver, the speckle field has a transverse velocity of

$$V_{trans} = 2\Omega z = 4\pi f_{rot}\, z, \tag{16}$$

where z is the object to transceiver distance and f_{rot} is the rotational frequency of the object. For a rectangular object cross section, the average speckle length at the receiver is given by the Van Cittert-Zernike theorem to be

$$L_{trans} = \frac{2\lambda z}{D} = \frac{\lambda z}{r}, \tag{17}$$

where r is the object radius. The ratio of the correlation length to the transverse velocity yields the following correlation time

$$\tau_{trans} = \frac{L_{trans}}{V_{trans}} = \frac{\lambda}{4\pi r f_{rot}}, \tag{18}$$

Using Eq.(15), we can relate the correlation time to the BW by

$$\Delta f_{trans} = \frac{\sqrt{2\ln 2/\pi}}{\tau_{trans}} = 4\sqrt{2\pi \ln 2}\,\frac{r f_{rot}}{\lambda} = 8.35\frac{r f_{rot}}{\lambda}. \tag{19}$$

This result indicates that the bandwidth associated with object rotation-induced speckle translation is linearly proportional to the rotation rate and the radius of the cylinder, and inversely proportional to the wavelength.

3.2.3. Effect of Scatter Exchange on Speckle Boiling

The impact of rotation-induced scatter exchange on the detected field is that of speckle boiling, which is another form of amplitude and phase modulation of the IF signal. As an object rotates, new scatterers are introduced into the illuminated region and produce a different speckle field with an entirely new complex field distribution at the receiver plane.

Precise definition of the object rotation angle necessary to decorrelate the speckle field at the receiver is difficult, but a conservative estimate would be one quarter rotation ($\pi/2$). We therefore use the following estimators

$$\tau_{exc} = \frac{1}{4 f_{rot}}, \tag{20}$$

$$\Delta f_{exc} = \frac{\sqrt{2\ln 2/\pi}}{\tau_{exc}} = 4\sqrt{2\ln 2/\pi}\, f_{rot} = 2.66 f_{rot}. \tag{21}$$

Comparing the bandwidths for the scatter-exchange process with those produced by rotation-induced translation, we infer that for the case of a fully illuminated rotating object, rotation-induced speckle translation dominates. This result would be quite different if the object was illuminated with a small focused spot [4]. In this case, scatter-exchange-induced speckle boiling can easily exceed the IF broadening due to rotation-induced speckle translation.

3.2.4. Effect of Doppler Shifts on Speckle Boiling

For a rotating object, the light scattered from any point will be Doppler shifted in an amount commensurate with that points instantaneous radial velocity (V_z depicted in Figure 2). Thus, the detected signal will be a random sum of an ensemble of scattered fields with varying frequencies, amplitudes and phases. Such a sum results in an amplitude and phase modulation of the detected IF known as Doppler broadening which, in terms of speckle dynamics, results in speckle boiling.

The frequency shift induced by the Doppler phenomena is easily predicted. It is well known that the Doppler shift of light reflected from a moving object is given by

$$f_d = \frac{2V_z}{\lambda}, \tag{22}$$

where V_z is the radial velocity of the illuminated point on the object surface. For a rotating cylindrical object this radial velocity is given by

$$V_z(\theta) = -\Omega r \sin(\theta) \tag{23}$$

The maximum velocity difference across the rotating cylinder is found by straightforward analysis to be given by

$$\Delta V_{max} = \left|2\,V_z(\theta)_{\max}\right| = 2\Omega r = 4\pi r\, f_{rot}\,. \tag{24}$$

The maximum Doppler frequency spread is found by combining the above two equations

$$\Delta f_{d_{max}} = \frac{2\Delta V_{max}}{\lambda} = \frac{8\pi r f_{rot}}{\lambda}\,. \tag{25}$$

To compute the FWHM BW Doppler broadening we must assume something about the shape of the PSDF due to Doppler alone. This would require knowledge of the Bidirectional Reflectance Distribution Function (BRDF) of the object. We have analyzed various cases for hypothetical BRDF's, all of which yield an IF FWHM BW on the order of 1/2 to 1/8 of the maximum Doppler frequency deviation. Thus, for our heuristic analysis we assume a nominal factor of 1/4.

$$\Delta f_{Doppler} = \frac{1}{4}\Delta f_{d\,max} = \frac{2\pi r f_{rot}}{\lambda} \tag{26}$$

The correlation time is then given by Eq.(15)

$$\tau_{Doppler} = \sqrt{\frac{\ln 2}{\pi^3}}\frac{\lambda}{r f_{rot}} = 0.2\frac{\lambda}{r f_{rot}}\,. \tag{27}$$

Comparing Eqs. (19) and (26), it is obvious that the two dominant object dynamics (i.e., rotation-induced speckle translation and Doppler-induced speckle boiling) produce nearly identical IF bandwidths. Although the heuristic analysis is by nature inexact, it does provide us with a relative measure of the effects of object dynamics on the speckle dynamics and the associated PSDF/ACF.

3.2.5. Combined Effects of Doppler and Object Rotation-Induced Translation

The combined effect of speckle boiling and speckle translation is a convolution of the component PSDF's or a multiplication of the ACF's (Wang [2]). Thus the combined effects of Doppler-induced boiling and rotation-induced translation is a convolution of the two individual effects, and for the case of Gaussian component PSDF's, the resultant PSDF is Gaussian with a FWHM BW equal to the root-sum-square (RSS) of the component FWHM BW's.

4. PRELIMINARY EXPERIMENTAL RESULTS

4.1. Experimental Description

Experiments were conducted to verify the laboratory scale theoretical predictions presented in the prior section. Two coherent detection ladar systems with wavelengths of 1.06 and 10.6 microns were used to illuminate a rough, rotating cylinder, while the power spectra of the raw IF signals were measured with a spectrum analyzer. The frequency broadened spectra were recorded and analyzed.

The 1.06 micron coherent ladar testbed utilizes a monostatic offset homodyne configuration. The system consisted of a 50 mW, CW Nd:YAG source, a 27 MHz acousto-optic modulator frequency shifter in the transmit path, and a 100 micron InGaAs p-i-n detector in conjunction with a 100 MHz transimpedance amplifier. The target was a 2.4 mm diameter sandblasted steel cylinder mounted from one end in a collet-type spindle and rotated at 10 Hz (600 RPM) with its rotational axis perpendicular to the ladar transceiver's line of sight. The illuminating beam spot size (1/e), and thus the effective target length, was approximately 4 mm in diameter, with a phase front radius of curvature of approximately 50 m.

The 10.6 micron testbed was also configured as a monostatic offset homodyne ladar. This testbed consisted of a 10 W, CW CO_2 source, a 27 MHz acousto-optic modulator frequency shifter in the transmit path, and a 100 micron HgCdTe detector. The same target was configured identically in both systems.

Power spectra measurements were taken with each system for two target rotational speeds (10 and 20 Hz), at a target to receiver distance of 2 meters. The results of these experiments are presented in the next section.

4.2. One Micron System Measurements

Two measured 1.06 micron power spectra, corresponding to 10 Hz and 20 Hz rotation rates, are superimposed in Figure 3. These traces consist of 100 frame video averages of the raw IF spectrum. The PSDF's are indeed Gaussian, as demonstrated by the Gaussian curve fits to the raw data. The measured FWHM of these curves are compared to those predicted by the heuristic analysis in Table 1. Note that doubling the rotational frequency from 10 to 20 Hz approximately doubles (within measurement tolerances) the FWHM bandwidth of the measured PSDF. The center frequency amplitude is reduced as well, tending to conserve total power (integrated PSDF area).

Table 1 Measured vs. theoretical FWHM bandwidths for the two target rotation rates of the 1 micron experiment.

f_{rot}	10 Hz	20 Hz
Measured	90 kHz	160 kHz
Theoretical rotation-induced translation	95 kHz	190 kHz
Theoretical Doppler-induced boiling	71 kHz	142 kHz
Theoretical RSS	118 kHz	236 kHz

The measured FWHM BW values in Table 1 are within 25% to 30% of the theoretical combination (RSS) of the translation and Doppler effects. This is a reasonable agreement given the inexact nature of the heuristic analysis. Experimental isolation of at least one of the component PSDF's will be necessary to further refine the accuracy of the theoretical estimates.

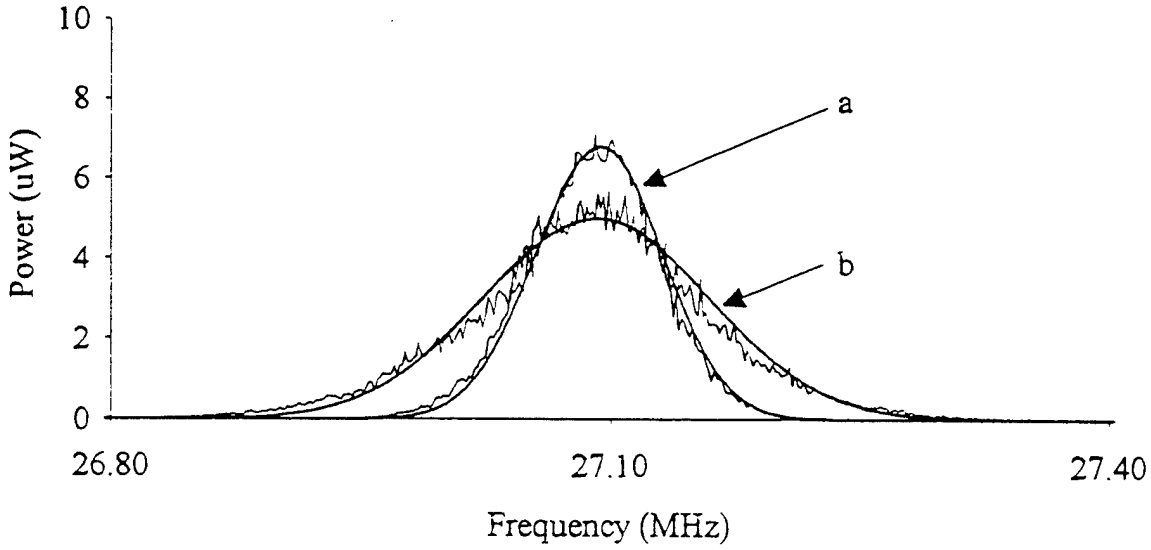


Figure 3. 1.06 micron power spectrum a) 10 Hz; b) 20 Hz, with superimposed Gaussian fits, for 2 meter range to target.

Figure 4 shows the autocorrelation functions for the two data sets illustrated in Figure 3. For the raw data, an FFT was employed to compute the ACF from the PSDF. The Gaussian traces (solid lines) shown in Figure 4 are analytic Fourier transforms of the Gaussian fits from Figure 3. The excellent agreement between the transforms of the raw data and the transforms of the analytic curve fits confirms the Gaussian nature of the PSDF.

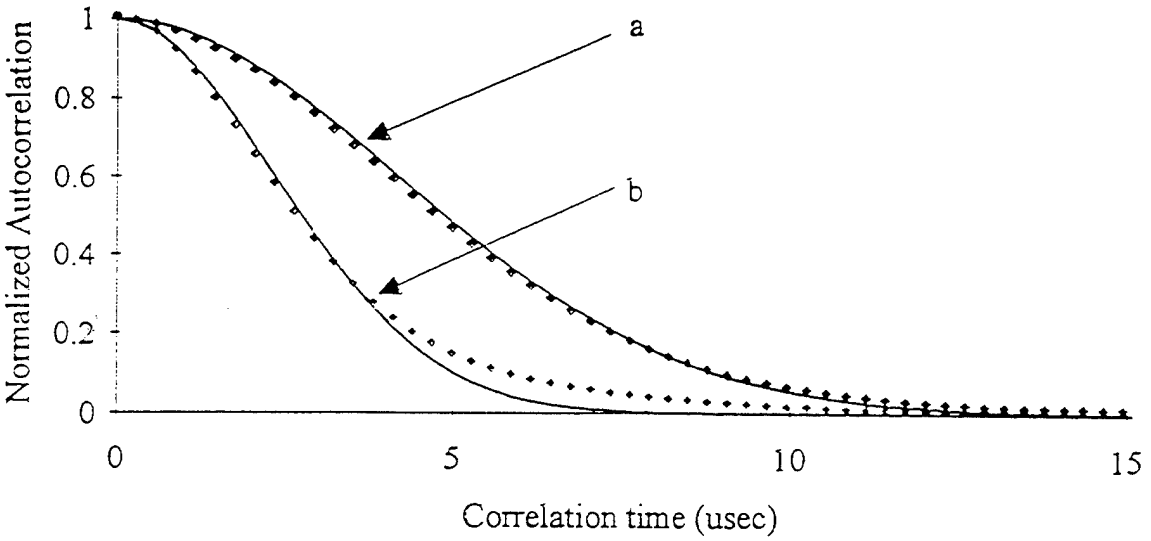


Figure 4. 1.06 micron autocorrelation function a) 10 Hz; b) 20 Hz, with analytic Fourier transforms of the Gaussian PSDF fits, for 2 meter range to target.

4.3. Ten Micron System Measurements

Identical measurements were conducted with the one and ten micron testbeds, with significantly different results. Figure 5 shows measured PSDF traces for a target rotating at 10 Hz (trace a) and 20 Hz (trace b) in the ten micron system. The smooth curves are attempts at Gaussian fits. As evident in Figure 5, the full span of neither trace could be fitted well. The Gaussian curves shown in Figure 5 only fit the central lobe, which we refer to as the narrowband component for its narrow relative bandwidth. In the outer region of the measured data (referred to as the broadband component) a much lower amplitude Gaussian fit is also apparent (cf., Figure 6).

4.3.1. Narrowband Component

The narrowband region of the measured data constitutes the bulk of the signal power. For both rotation rates, the FWHM of each trace did not scale by the inverse of the wavelength ratio as predicted by our heuristic analysis. Each FWHM bandwidth is significantly less than 1/10 of those measured at one micron. For example at 10 Hz, the measured one micron FWHM BW should scale from 90 kHz to 9 kHz. However, the narrowband component of the ten micron spectrum of Figure 5 shows little more than 2 kHz of FWHM BW.

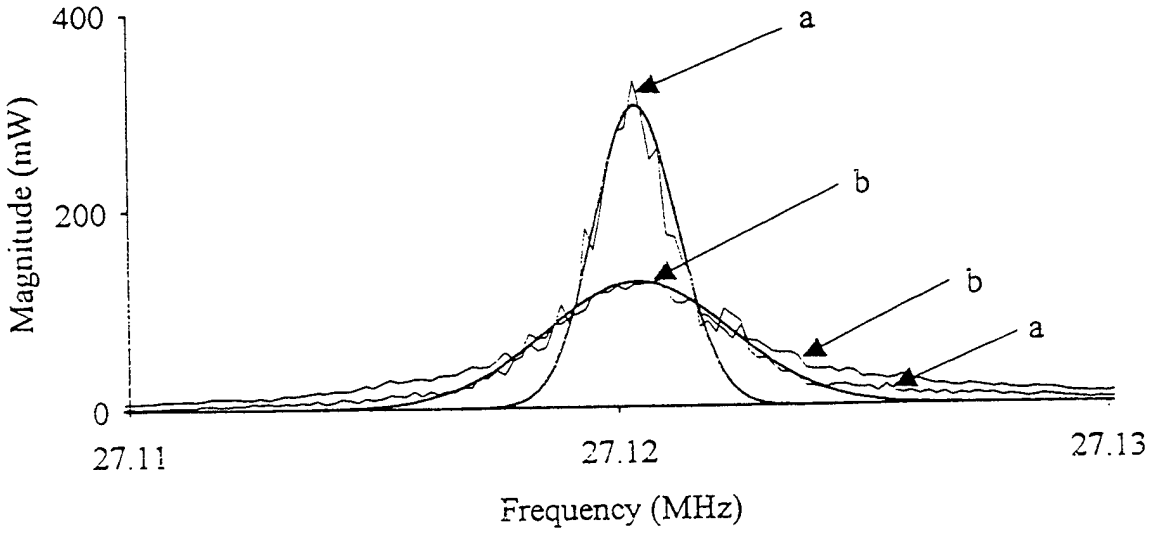


Figure 5. 10.6 micron power spectrum a) 10 Hz; b) 20 Hz, for 2 meter range to target with superimposed Gaussian fits for the narrowband components.

4.3.2. Broadband Component

The broadband region of the PSDF is magnified in Figure 6. In this region the Gaussian fits scale with the one micron data. That is, for 10 Hz. rotation, the FWHM BW scales from 90 kHz at one micron to 9 kHz at ten microns. Therefore excellent agreement is observed between the full one micron spectrum and the broadband component of the ten micron spectrum.

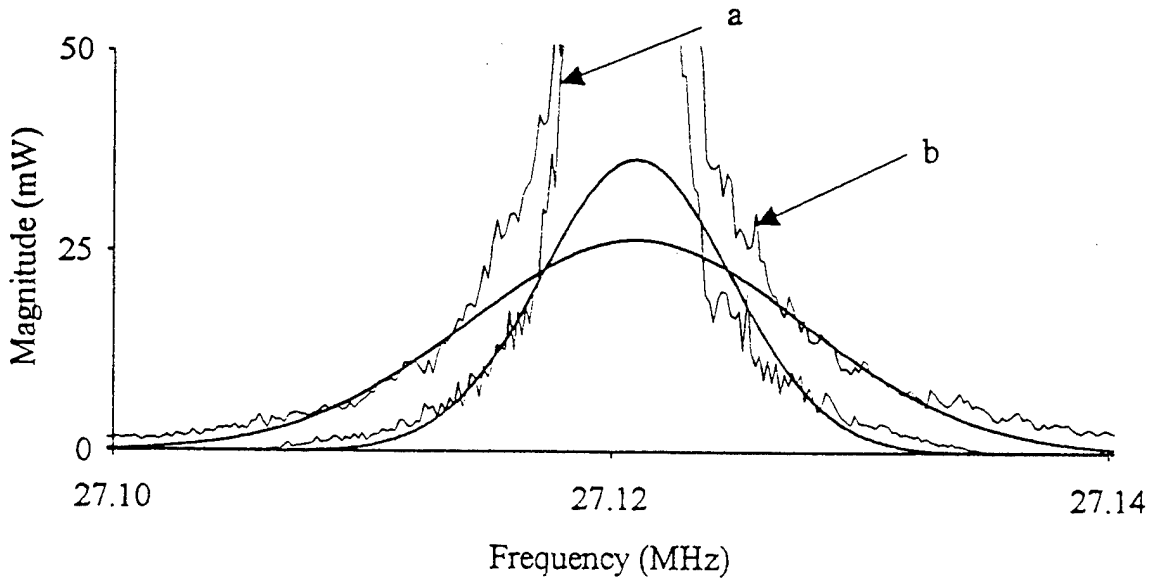


Figure 6. Enlarged 10.6 micron power spectrum a) 10 Hz; b) 20 Hz, for a 2 meter target with superimposed Gaussian fits for the broadband components.

4.3.3. Discussion of the Ten Micron Results

Table 2 summarizes our results for the ten micron experiment. Given the dual Gaussian nature of the measured PSDF's, little agreement is expected between the heuristic analysis predictions and the measured results. The FWHM BW of the braodband component was in good agreement with the hueristic analysis however the FWHM BW of the narrowband component is not.

Table 2 Measured vs. theoretical FWHM bandwidths for the10.6 micron experiment for narrowband and broadband components.

f_{rot}	10 Hz	20 Hz
Theoretical rotation-induced translation	9.5 kHz	19 kHz
Theoretical Doppler-induced boiling	7.1 kHz	14.2 kHz
Theoretical RSS (assuming diffuse reflection)	11.8 kHz	23.6 kHz
Measured Broadband Component	9 kHz	16 kHz
Measured Narrowband Component	2 kHz	4.5 kHz

We speculate that the two regions of Gaussian structure (narrowband and broadband) in the ten micron data arise from the target's greater specular reflectivity at the longer wavelength. Associated with the increased specualar reflectivity would be a peak in the angular intensity distribution of the field reflected from the cylinder. Such a peak would be narrow in angular extent and would be quickly diminished to diffuse levels with displacement from the specular

reflection angle. This near-specular reflection would give rise to a windowing effect in which a significant portion of the reflected energy is scattered from a narrow section of the cylinder surface. The narrow "window" produces a distinct speckle field with transversely longer speckles than those produced by the energy diffusely scattered from the full diameter of the target.

This dual scale speckle field is easily demonstrated. Figures 7 and 8 show two distinct portions of a far field speckle pattern produced by illuminating a slightly polished metal cylinder (not our target cylinder) with a HeNe laser (0.6328 microns). This configuration produces a wide speckle band transverse to the cylinder axis. The pattern of Figure 7 was located in the horizontal plane of the illumination (with the cylinder axis vertical) and slightly off the illumination axis. The speckle cells exhibited are relatively long with many extending over the entire transverse extent of the field of view. Above and below this in-plane field is the pattern of Figure 8, which exhibits much shorter speckle cells characteristic of diffuse scattering.

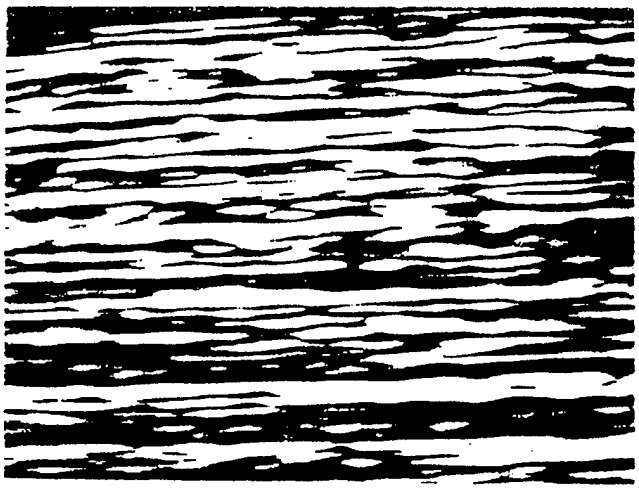


Figure 7. Near-specular portion of the speckle field resulting from HeNe illumination of a slightly polished cylinder.

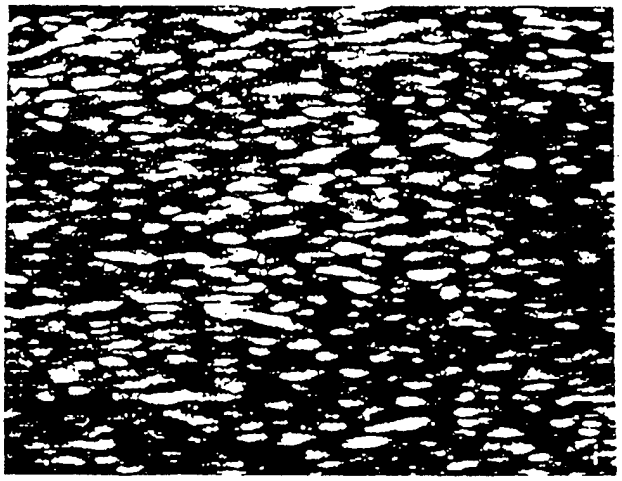


Figure 8. Diffuse portion of the same speckle field shown in Figure 7 obtained by tilting the cylinder with respect to the y axis.

The effect of the longer speckle cells in Figure 7 on the translation-associated component of the PSDF bandwidth is evident from our heuristic analysis. Equation 18 gives the linear relationship between the average speckle length and the correlation time. The longer speckles increase the correlation time and decrease the bandwidth of the translation effects. The effect of the longer speckle cells on the Doppler-induced boiling bandwidth is also evident from the analysis. Equation 25 illustrates the linear relationship between the cylinder surface's radial velocity relative to the ladar receiver and the maximum Doppler-induced boiling bandwidth. The maximum radial velocity within the narrow "window", which must be located at the cylinder's center-plane, is much smaller than that associated with the full cylinder diameter. The boiling BW is correspondingly reduced.

Since the speckle fields resulting from the near-specular and the diffuse reflected field components are spatially superimposed, the two PSDF's are additive (each the product of the convolutional relationship discussed earlier), giving rise to the two regions of Gaussian structure exhibited by the ten micron data. The relative amplitudes of these two regions would be commensurate with the relative strength of the diffuse and near-specular reflections from the object. Thus at one micron, where the reflected field is predominately diffuse only one Gaussian structure is observed while at ten microns two are present.

5. CONCLUSIONS

We have shown that a statistical model utilizing the MCF of the detected field is required to describe the shape and BW of the IF PSDF. However, for a plane wave illuminated rotating cylindrical target, no tractable solution to the MCF exists due to a spatial non-stationarity in the reflected field. On the other hand, a heuristic analysis of the problem yields significant insight into the BW of the PSDF.

Our heuristic analysis assumed the IF PSDF may, in general, be modelled as the convolution of two Gaussian component PSDF's, one arising from speckle translation effects and the second arising from speckle boiling effects. For the specific case of a plane wave illuminated rotating cylinder, our analysis estimated the BW of the translational-effect PSDF based on speckle size at the receiver and object rotational speed. We also concluded that the boiling-effect PSDF was dominated by Doppler effects, and estimated the associated BW based on the maximum surface velocity of the rotating cylinder. The estimated BW's for both component PSDF's were comparable. The BW of the composite PSDF is shown to be the RSS of the component BW's since the composite spectrum is the convolution of two Gaussian component spectra.

Measured composite PSDF's for the one micron system were Gaussian in shape, and the associated BW values were reasonably consistent with the predictions of the heuristic analysis. Measured composite PSDF's for the ten micron system appeared to be the sum of two Gaussian spectrums of distinctly different aspect ratio. We speculated that one arose from the diffuse portion of the reflected field and the other from a near-specular portion of the reflected field. We believe the BW of the diffusely reflected field is dependent upon the physical radius of the cylinder while the BW due to the specularly reflected field is dependent upon a narrower longitudinal

section of the cylinder's illuminated surface. Lastly, we observed that the relative amplitudes of the two Gaussian spectra appear to be commensurate with the relative power of the two reflected fields (near-specular vs. diffuse) at the observation point.

6. REFERENCES

[1] Goodman, J. W., Statistical Optics, John Wiley and Sons, Inc., 1985

[2] Wang, Jon Y., "Lidar signal fluctuations caused by beam translations and scan," Applied Optics, Vol. 25, No. 17, Sept. 1986

[3] Goodman, J. W., Statistical Optics, John Wiley and Sons, Inc., 1985; Eq. 5.1-28

[4] Erdman, Joachim C., "Speckle field of curved, rotating surfaces of Gaussian roughness illuminated by a laser spot," JOSA, Vol. 66, No. 11, Nov. 1976

7. ACKNOWLEDGMENTS

This research has been supported by the Strategic Defense Initiative Organization, Innovative Science and Technology Office through the Office of Naval Research under contract N00014-92-C-0123

5.6 Appendix 5B

M. Acharekar, P. Gatt and L. Mizerka, "Laser Vibration Sensor", Proc. SPIE 2472, 2-11 (1995).

Laser Vibration Sensor

By

Madhu Acharekar
Schwartz Electro-Optics,Inc.
3404 N. Orange Blossom Trail
Orlando, FL 32804
(407) 298-1802

Philip Gatt
Center for Research and Education in Optics and Lasers (CREOL)
University of Central Florida
Orlando, FL 32826
(407) 658-6800

&

Lawrence Mizerka
Night Vision & Electronic Sensors Directorate
10221 Burbeck Road, Suite 430
Fort Belvoir, VA 22060-5806
(703) 704-3666

ABSTRACT

An eye safe Ho:YAG laser operating CW at 2.09 μm was used as the transmitter for this laser radar (ladar) system. The diode pumped Ho:YAG laser provided up to 100 mW of single frequency output power. An InGaAs detector was selected as the off-set homodyne receiver for vibration signatures. This paper describes the design and characteristics of the laser vibration sensor that is based on laser Doppler radar. Initially, the sensor was calibrated using a piezoelectric loudspeaker, then vibrational signatures were obtained for automobiles having four and eight cylinder engines. This two micron laser Doppler radar permitted high-resolution vibrational signatures detection, remotely, at ranges up to 50 meters.

Keywords: Laser Doppler radar, Laser Radar. Ladar, Eye Safe Ladar, Vibration Sensor, Combat Identifrcation, Identification Friend or Foe, Ladar for IFF

1. INTRODUCTION

Recently, SEO completed a study of Non-Cooperative Combat Identification (NCCI) for the U.S. Army under contract DAAB07-93-C-U004. In that program, a diode pumped eye-safe Ho:YAG laser was used with an offset homodyne receiver to obtain high-resolution vibration signatures remotely for identification of friend or foe (IFF). This paper describes the design and characteristics of the laser vibration sensor developed under this contract.

0-8194-1825-0/95/$6.00

carrier-to-noise ratio for the laser sensor. As is common in the literature, we make a distinction between the carrier-to-noise ratio (CNR) and the signal-to-noise ratio (SNR). The CNR is defined as the ratio of the intermediate frequency (IF) carrier power to the noise power, while the SNR is the ratio of the signal power to the noise power. The SNR depends not only on the CNR, but also on the type of signal (i.e., range, velocity, vibration) and the signal processing architecture (i.e., demodulators, range gates, signal bandwidth, analog to digital converters, etc.). For reliable system operation the minimum CNR should be on the order of 10 dB. For a properly designed (i.e., shot-noise limited) heterodyne receiver the CNR is given by

$$CNR = \frac{\eta \gamma P_S}{h\nu\beta} \tag{1}$$

where η is the detector quantum efficiency, γ is the heterodyne or mixing efficiency, P_s is the received optical signal power, h is Plank's constant, ν is the optical frequency, and β is the receiver bandwidth. The detector quantum efficiency is the ratio of photoelectrons generated by the photovoltaic detector to the number of incident photons. The quantum efficiency of the room temperature InGaAs detector utilized in this program was measured to be 70%. The heterodyne efficiency is the overlap integral between the local oscillator field and the signal field.[2]

$$\gamma = \frac{\left[\iint_A U_S(x,y)\, U_{lo}^*(x,y)\, dxdy\right]^2}{\iint_A [U_S(x,y)]^2\, dxdy \iint_A [U_{lo}(x,y)]^2\, dxdy} \tag{2}$$

Here U_s and U_{lo} are the signal and local oscillator fields respectively, and the integration is performed over the detector surface area. For a uniform local oscillator field and an Airy signal field the optimum efficiency is on the order of 71.6% and occurs when the detector diameter is matched to 72% of the first Airy zero (i.e., 2.44 λf/D where f is focal length and D is diameter of the telescope)[3]. The receiver bandwidth is critical to the receiver performance. Excessive bandwidth results in excess noise and a reduced CNR, while insufficient bandwidth distorts the signal. The received laser power can be calculated directly from the ladar equation. The ladar equation is a radiometric calculation that relates the received signal power to the ladar, target, and atmospheric parameters. The ladar equation has two forms: one for resolved targets, where the beam spot size is smaller than the target, and the other for unresolved targets. For the tactical scenario of interest the targets cross section is on the order of 10 square meters and the spot size is on the order of 0.1 square meters. Therefore, the target is resolved by the beam and the corresponding ladar equation is given by[4]

$$P_S = P_x \varepsilon_x \varepsilon_r \rho T^2 \frac{A_r}{R^2 \Omega_t} \tag{3}$$

where P_x is the laser power, ε_x is the transmitter optical efficiency, ε_r is the receiver optical efficiency, ρ is the target reflectance, T is the one way transmission, A_r is the receiver aperture area, R is the range to the target, and Ω_t is the solid angle of the target backscatter. For a resolved target the received power is inversely proportional to the range squared. If the target is smaller than the beam spot size, then the unresolved form of the ladar equation applies. This form is given by

$$P_s = P_x \epsilon_x \epsilon_r \rho T^2 \frac{A_r}{R^2 \Omega_t} \frac{A_t}{R^2 \Omega_x} \quad (4)$$

where A is the target cross sectional area and Ω_t is the transmit beam solid angle. For an unresolved target the received power is proportional to the range to the negative fourth power.

The CNR was calculated to be greater than 10 dB for a 100 mW CW coherent ladar with a 5 cm transmit/receive aperture, a system bandwidth of 100 kHz, a combined optics and detection efficiency of 10%, and a 10% reflectivity target of 1 m^2 at range of 1 km. This calculation was based on a standard clear day (23 km visibility). The same calculation for a hazy day (5 km visibility) indicates that 250 mW is necessary. However, with a 10 cm aperture, assuming turbulence is not a limiting factor, less than 100 mW is required for either atmospheric condition. Therefore, a 10 cm receiver aperture and 100 mW transmitter power was selected to obtain the required range.

3. LADAR EXPERIMENTAL SET-UP

A brief description of the ladar experimental setup is given in this section.

3.1 Diode Pumped CW Holmium Laser

A diode pumped eye-safe Ho:YAG laser, providing a 100 mW output power was used as the CW transmitter for this Doppler radar application. The design of the laser includes cooling of the laser crystal from -20 to -40 °C to obtain CW lasing. The Ho:YAG crystal was cooled using a six stage thermoelectric (TE) cooler.

A SDL-2462 diode was used for pumping the Ho:YAG crystal. The SDL-2462 diode was terminated using a 200 μm core fiber with a 0.5 numerical aperture. The diode provided 786 nm output wavelength when cooled to 6 °C, and an output wavelength of 790 nm at room temperature. The SDL-2462 laser diode was provided with an internal TE cooler. The laser diode was mounted on a heat sink plate which was cooled using a circulating coolant. The diode temperature was maintained between 0 and 6 °C to obtain an output wavelength of 786 nm. The laser diode output was directly coupled to the crystal by placing the fiber very close to the crystal. A hemispherical resonator cavity consisting of two mirrors, one high reflectivity (HR) flat mirror coated directly on the Ho:YAG crystal, and a curved (10 cm radius of curvature, 98% reflectivity at 2.1 μm) output coupler, were used. The internal surface of the crystal was anti-reflection (AR) coated for 2.1 μm and the external surface was coated for HR at 2.1 μm and high transmission over the pump region of 786 nm. Tuning of the Ho:YAG gain spectrum is achieved by manual rotation of an uncoated, 1 mm thick, fused silica etalon inserted in the resonator. The spectral and spatial profiles of the transmitter were monitored with a scanning confocal interferometer (Burleigh Instruments CFT-500; resolution ~3 MHz) and a laser beam analyzer (Spiricon LBA-100A with a Cohu 4800 silicon CCD camera), respectively. The maximum power was determined by the current limit of the diode laser input power. The laser was designed, fabricated and tested at SEO.

3.2 Laser Vibration Sensor

A conventional optics transceiver was used to obtain Doppler signals from a variety of targets. Recently, there has been much interest in the use of coherent detection for long range fiber optic communication systems. This has motivated the development of various fiber components needed to implement fiber based coherent detection receivers. The application of fiber optics technology to coherent ladar systems is compelling for several reasons. Firstly, fiber optic components are much more compact and lightweight than conventional optics. This is an important consideration for systems that may eventually be deployed in the military environments. Secondly, if properly constructed, fiber optic systems are less likely to go out of alignment, especially under conditions of vibration or high acceleration. Finally, fiber optics is compatible with the use of integrated opto-electronic components such as modulators and switches, which have characteristically lower power consumption and better high frequency response than conventional discrete devices. It is interesting to note that 10 micron Doppler ladars do not have a fiber technology equivalent to that which is available for systems operating with wavelengths within the silica window.

3.2.1 Ladar Transceiver Optics

The experimental laser vibration sensor is depicted in Figure 1. An offset homodyne configuration is employed to measure the vibration signatures of stationary objects as shown here.

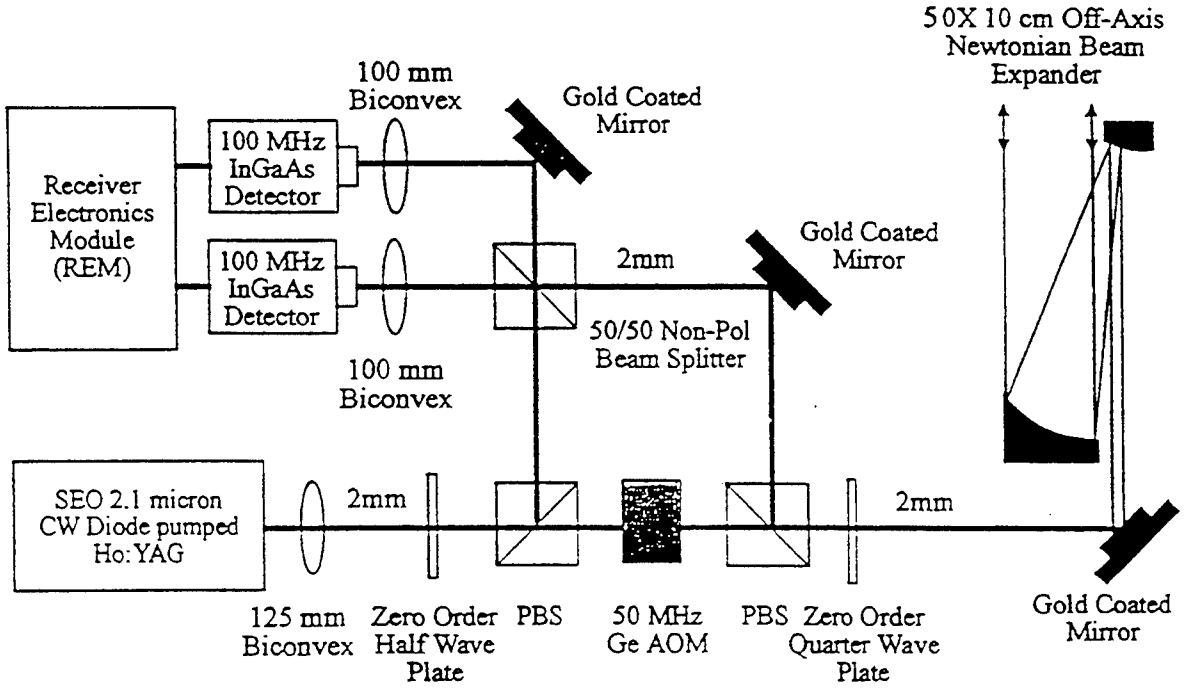


Figure 1: 2 Micron Laser Doppler Radar Experimental Set-up

To accommodate moving targets, a frequency tracking IF local oscillator would be required. In the offset homodyne configuration, a portion of the laser power is extracted for the vertically polarized local oscillator by the variable power divider. This variable power divider consists of a half wave plate and a polarizing beam splitter. The remaining horizontally polarized laser energy is frequency shifted (single side band modulation, suppressed carrier) by 27.12 MHz using the acousto-optic modulator. This energy is then transmitted through the optical duplexer and directed towards the target using the scan mirror.

The optical duplexer consists of a polarizing beam splitter and a quarter wave plate. The outgoing beam is converted to a right-hand circular polarization by the quarter-wave plate, and upon reflection from the target becomes predominantly left-hand circularly polarized. After passing through the quarter-wave plate once more, the polarization is converted to vertical. The vertically polarized signal field is then diverted to the receiver via the polarizing beam splitter. The local oscillator and the return signal fields are combined in the 50/50 non-polarizing beam splitter. Large-scale heterodyne ladars typically utilize a beam combiner with a larger mix ratio (i.e., 10/90). However, for a laboratory ladar which utilizes the same laser for the LO and the signal, a 50/50 ratio is often desirable, especially if a dual-balanced detector receiver is utilized. The high cost of graded InGaAs detectors, precluded the use of a dual-balanced detector configuration in this program. Once the two beams have been combined, they are focused onto the small (75 micron) graded dopant InGaAs detector. The focal length of the lens was selected to maximize the heterodyne efficiency.

3.2.2 Ladar Receiver Electronics

The laser vibration sensor receiver electronics are depicted in Figure 2. The photocurrent from an off-the-shelf graded dopant InGaAs detector is converted to a voltage with a low-noise transimpedance preamplifier. The gain of this DC coupled preamplifier is lower for a coherent receiver (1 kΩ in this case) than for a direct detection receiver, since higher bandwidths are required and receiver saturation by the strong (1 mA) LO photocurrent must be prevented.

The output of the preamplifier is a voltage which is then AC coupled, filtered and amplified with a wide-band RF amplifier. A 20 dB automatic gain control (AGC) circuit was designed to help stabilize the IF amplitude. The stabilized IF signal is then frequency demodulated using a commercially available wideband (10 MHz) frequency discriminator. At this point the signal, which is proportional to the vibration velocity, is integrated to generate a phase demodulated signal proportional to the target vibration displacement. The integrator provides a gain of 40 (26 dB) for DC signals and has an integration cut-on frequency of 1 Hz (i.e., an integration time of $1/[2\pi]$sec). A Hewlett Packard HP8560A 2.9 GHz RF Spectrum Analyzer recorded the raw IF vibrations spectra and a Tektronix DSA602 Signal Processing Analyzer recorded the demodulated waveforms. The DSA602 is a high speed, maximum 2 GS/s, 8 bit digitizer, with a maximum sensitivity of 10 mV per 8 bits. Frequency spectra of the demodulated signals were generated via a custom FFT algorithm. This algorithm was scaled to register the peak value of a sinusoid to facilitate the data analysis. In the automobile experiments, 1024 point data records were digitized at a rate of 1 KS/s. Therefore the frequency resolution of the FFT algorithm was 1000/1024 or just less than 1 Hz.

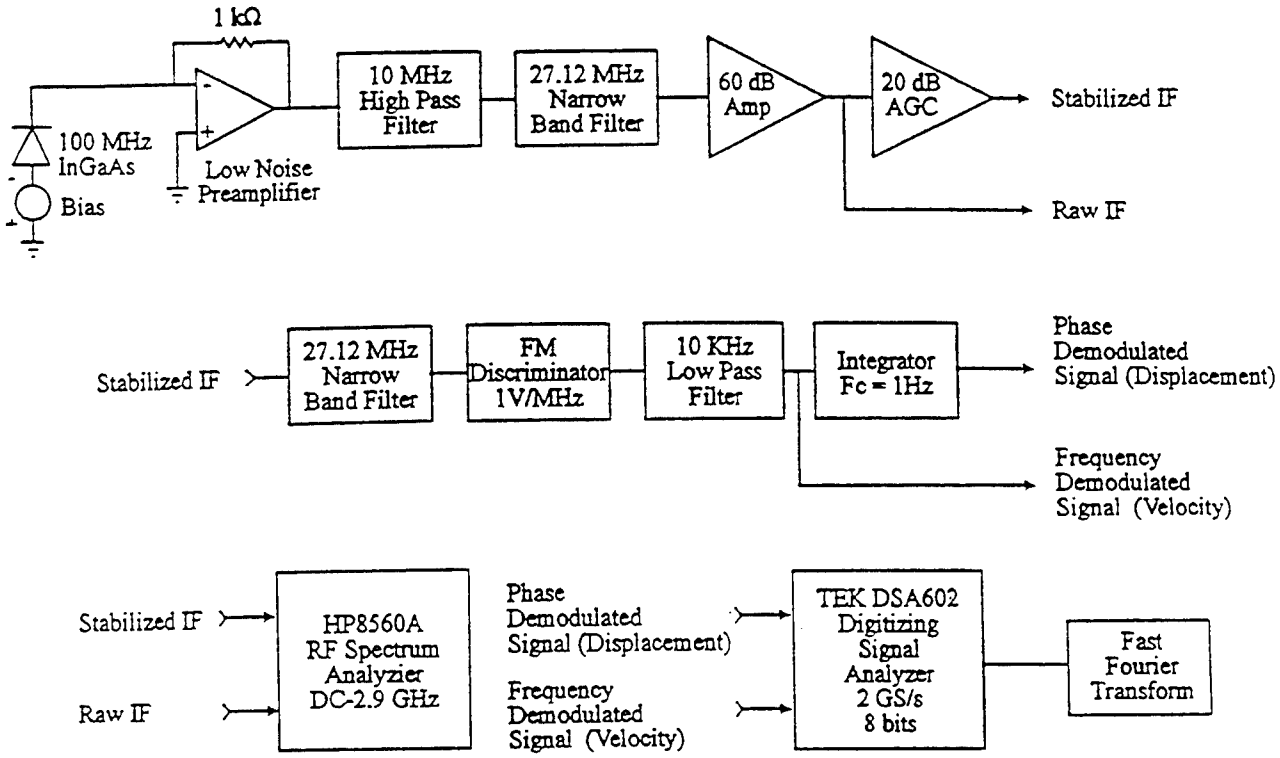


Figure 2: Vibration Sensor Electronics, Block Diagram

4. VIBRATION SENSOR PERFORMANCE

Measurements of automobile spectra were conducted with the 2.1 micron laser vibration sensor test-bed. These experiments were conducted in two phases. In the first phase, the complete system (optics and electronics) was calibrated using a piezoelectric loudspeaker as a target. Results of the calibration experiment indicated that vibration amplitude sensitivities on the order of 10 nm could be readily discerned. Spectra were collected for stationary automobiles in the second phase. In these experiments clearly unique spectral resonance frequencies were identified for two different automobiles. This indicates the usefulness of this sensor to identify different vehicles by their vibration spectra. The purpose of these experiments were to characterize the 2 μm laser vibration sensor for Non-Cooperative Combat Identification (NCCI) by using high resolution vibration signatures of targets.

4.1 Piezoelectric Loudspeaker Calibration Experiments

The calibration experiments consisted of modulating a piezoelectric loudspeaker with a monotone sinusoidal displacement function. The purpose of these calibration experiments was to determine the characteristic curve which relates the demodulated voltage amplitude to the target displacement amplitude and demonstrate the fidelity of the sensor.

Results of Piezoelectric Loudspeaker Experiments: A variety of experiments were conducted at different modulation indices and modulation frequencies. By computing the Fast Fourier Transform of the demodulated time waveform, accurate measurements of the demodulated amplitude are possible.

$$V_{P_{mV}} = 0.225\beta = 1.35D_{P_{\mu m}} \quad (5)$$

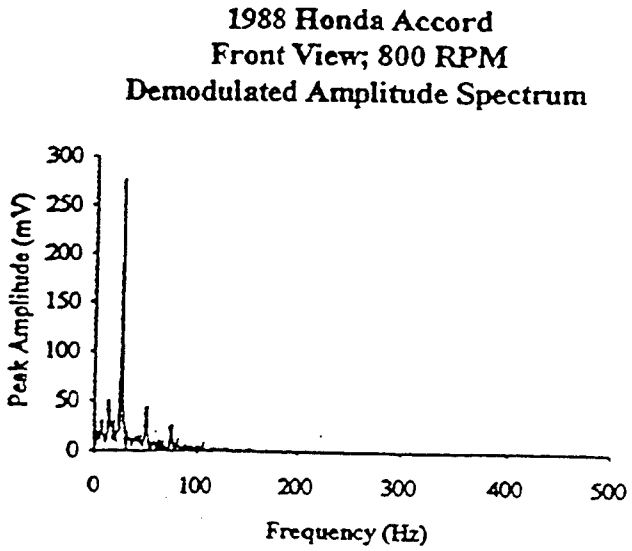
Where, V_p is the resonant frequency amplitude measured in mV and D_p is the corresponding target displacement in micrometers (μm). Furthermore, the peak voltage was found not to be a function of the modulation frequency, indicating the high-fidelity of the integrator. Sub-micrometer displacement accuracy were achieved using the ladar. It may be noted that the ladar can be easily converted from a non-cooperative to a cooperative sensor by simply providing a piezoelectric transducer (PZT) on the combat vehicle with a changeable frequency codes for the PZT.

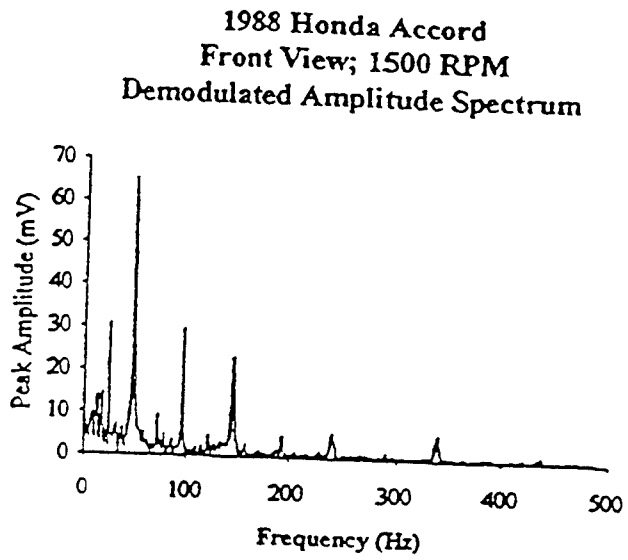
4.2 Automobile Vibration Sensing Experiments

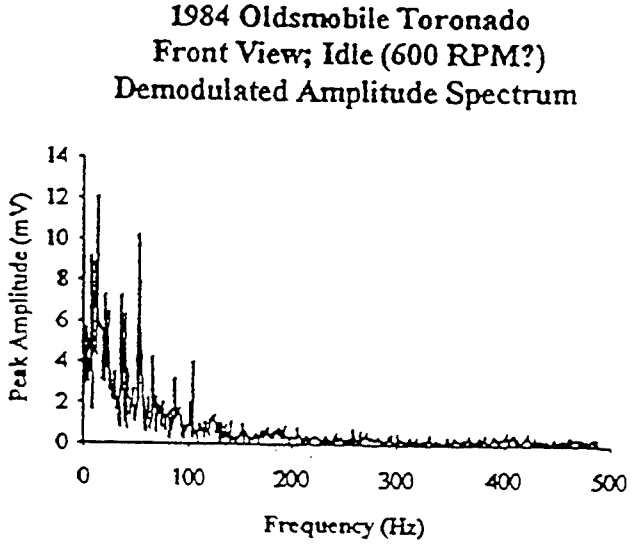
The portable 2.1 micron laser vibration sensor was deployed outside the Ladar Laboratory, on October 24, 1993 to collect data from a 1988 Honda Accord (4 Cyl.) and a 1984 Oldsmobile Toronado (8 Cyl.). A scan mirror and a boresighted HeNe Laser were used to direct the 2.1 micron beam onto a red plastic automobile retroreflector mounted on the front fender of the vehicle. The plastic reflector was needed to obtain enough CNR since a wideband RF filter (1 MHz bandwidth) was utilized. The measured RF spectra were much more narrow band than anticipated and consequently narrower filters could have been deployed, eliminating the need to use the plastic reflectors.

Results of Automobile Vibration Sensing: The data shown in Figures 3, displays the demodulated amplitude spectrum of two automobiles at two different RPMs. From the figure, one can observe that the two automobiles exhibit unique resonance frequencies which appear at definite harmonics of a base frequency. For the Honda Accord, resonance frequencies occur at multiples of 24 Hz, while the Oldsmobile Toronado exhibited resonance frequencies at multiples of 13 Hz.

The strength of the resonance frequencies depends on the engine RPM and other mechanical factors. When the frequency of combustion cycles is near a particular resonance, a strong peak is observed at that frequency. For example, with the Honda, two combustion cycles occur per crankshaft revolution; therefore, the combustion frequency is twice the RPM. Thus, at an engine speed of 800 RPM the combustion frequency is 26 Hz and the data shows the strongest resonance at 25 Hz. On the other hand at 1500 RPM the combustion frequency is 50 Hz and the strongest peak occurs at 48 Hz. Note that the resonance at 24 Hz is still present in the 1500 RPM data. For the Oldsmobile Toronodo, the combustion frequency is four times the engine RPM; this corresponds to a combustion frequency of 40 Hz at 600 RPM and 133 Hz at 2000 RPM. The data also demonstrates that as the RPM is increased energy is transferred from the lower resonance frequencies to the higher resonance frequencies. The widest RF bandwidth encountered during these experiments was less than 20 kHz.







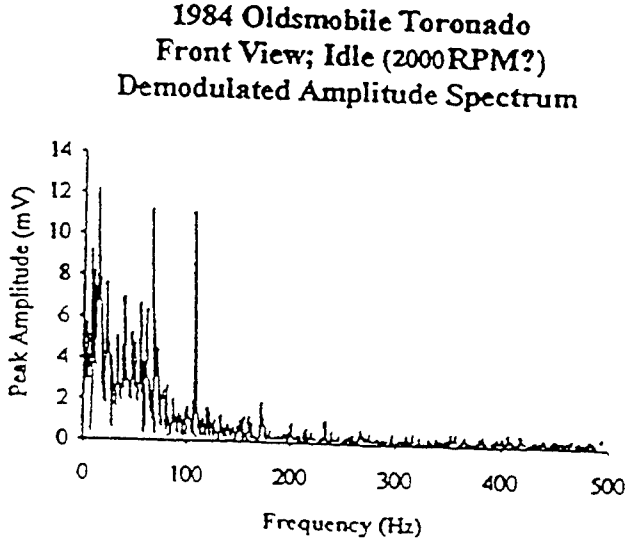


Figure 3: Vibration Data Using NCCI system

5. CONCLUSIONS

The following conclusions may be drawn based on the results obtained during the program:

1) SEO demonstrated an innovative approach for a Non-Cooperative Combat Identification (NCCI) sensor using an eye-safe laser operating at 2.1 μm. We did not, however, design the target identification algorithm which would consist of a data base of known target spectra and a correlation algorithm.
2) A single frequency, diode pumped, CW, Ho:YAG laser providing 100 mW output power was used for this laboratory NCCI system. The laser design is rugged and compact. The laser output power can be scaled to 2 W using a 10 W pump laser to make an operational NCCI system.
3) The numerical model developed during the program will allow SEO to optimize the system for the battlefield obscurant and the atmospheric turbulence.

6. ACKNOWLEDGEMENT

This work was conducted under Contract #DAAB07-93-C-U004 from Night Vision & Electronic Sensor Directorate, Fort Belvoir, VA. The laser work was conducted at SEO, Orlando, Florida while the Doppler radar work was conducted at the Center for Research in Electro-Optics and Lasers (CREOL) associated with the University of Central Florida (UCF) also located in Orlando, Florida.

7. REFERENCES

1. Desert Storm's Night Vision and Electro-Optical Equipment Suitability Survey (U), February 1992, Fort Belvoir, Virginia

2. D. Fink, "Coherent Detection Signal-to-Noise," Applied Optics, Vol. 14, No. 3, pp. 670-789, Mar. 1975

3. D. Fink, "Coherent Detection Signal-to-Noise," Applied Optics, Vol. 14, No.3, pp. 670-689, Mar. 1975

4. C.D. Bachman, "Laser Radar systems and Techniques, Massachusetts:Artech House, Inc., 1979

6.0 RECIPROCAL PATH SCATTERING THROUGH RANDOM MEDIA

James E. Harvey, Ph.D

6.1 Background

The phenomenon of retroreflection from rough surfaces or suspended particles, [1-3] the opposition effect, [4-6] enhanced backscatter (EBS), [7-12] the double passage effect [13-16] or reciprocal path imaging [17, 18] has received a great deal of attention in recent years. This naturally-occurring, partial conjugate-wave phenomenon has been shown to be caused by constructive interference produced by multiple reciprocal scattering paths through a random refractive phase screen as illustrated in Figure 6.1.

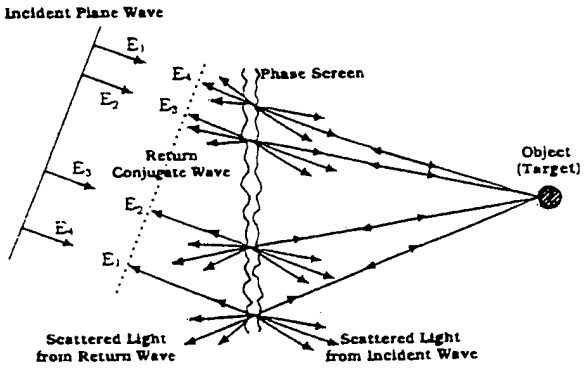


Figure 6-1: Conjugate wave formation by reciprocal path scattering through a random phase screen or turbulent medium.

If we consider two arbitrary ray paths (say E_1 and E_4) orthogonal to the incident plane wave, they will be scattered by the random phase screen as indicated schematically in Figure 6.1. One scattered component in each of the resulting angular spectrum of ray paths will be directed toward the target and reflected (scattered) back along the incoming path of a companion scattered component. They will each be scattered a second time on the return trip when they encounter the random phase screen. Again, one scattered component in each of the resulting angular spectrum of ray paths will be directed back precisely anti-parallel to the incident companion ray path. Assuming that the temporal variations of the random phase screen are slow relative to the round-trip time-of-flight, light propagating along any two reciprocal scattering paths will clearly experience identical phase delays due to their double passage through the same two points on the random phase screen. Since the above situation holds for a myriad of possible ray pairs associated with the incident wavefront, a multitude of conjugate waves (perfect plane waves) are formed (each having its own total phase delay) and propagate back toward the transmitting aperture along with all of the returned scattered radiation.

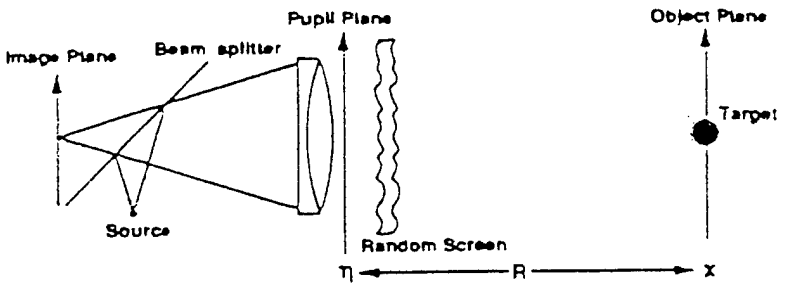


Figure 6.2: Monostatic laser imaging configuration used to demonstrate the enhanced backscatter phenomenon.

A monostatic laser imaging configuration such as that illustrated schematically in Figure 6.2 is required for the reciprocal path effect to exist and produce a conjugate return wave. The resulting irradiance pattern in the image plane is a time-varying speckle pattern caused by the random phase screen. However, by averaging over many realizations of the speckle pattern, a narrow enhancement will form on the smoothed scattered light distribution. Figure 6.3(a) illustrates a computer simulation of this phenomena representing the average of 110 realizations of the speckle pattern.[10] The target was assumed to be an optically smooth mirror in the far field of the transmitting aperture. A deliberate pointing error (incident angle) of approximately 0.02 degrees is evident from the position of the enhancement relative to the center of the scattering function.

Figure 6.3(b) shows the experimental results of a laboratory experiment in which a propane burner was used to create the time-varying random phase screen. A helium-neon laser was used as a source for illuminating the target. The target was a narrow rectangular mirror whose width was small compared to the size of the illuminating beam. After reflecting from the target in the object plane, the return radiation passed through the random phase screen a second time, was focused onto the image plane, and recorded with a CCD camera with a 1.0 sec exposure time. By translating the target slightly from the center of the illuminating beam, the recorded image was qualitatively very similar to the previous computer simulation. The greater width of the measured enhancement is due to diffraction from the finite extent of the incident beam.

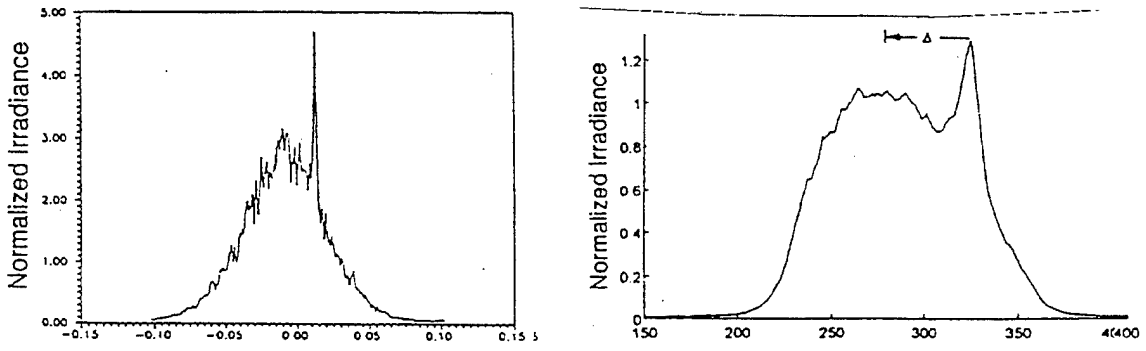


Figure 6.3: (a) Computer simulation of the EBS phenomenon (110 realizations).
(b) Measured irradiance profile through the enhancement.

One must be careful not to confuse the backscatter enhancement with an *image* of the target. The scattering function is the (turbulence-induced) blurred image of the target. If two separate targets were illuminated and imaged by the tracking telescope, there would not be two separate enhancements; instead, there would be two separate scattering functions with a single enhancement on the boresight of the telescope as illustrated schematically in Figure 6.4(a). Experimental data indicating a single enhancement due to two separate targets is illustrated in Figure 6.4(b).

A direct optical image is not formed by this EBS process; however, information concerning the spatial frequency content of the image does exist in the reciprocal path radiation, and image reconstruction algorithms can be implemented to recover an image of an extended object.[14, 15, 18] We introduce the descriptive phrase *reciprocal path scattering* (RPS) when referring to this general conjugate wave phenomenon, and reserve the phrase *reciprocal path imaging* (RPI) to describe attempts to exploit this phenomenon for obtaining diffraction-limited images of extended objects actively illuminated through random media. Furthermore, polarizing sparse array configurations have been described which yield uniform spatial frequency response in RPI applications.[17, 18]

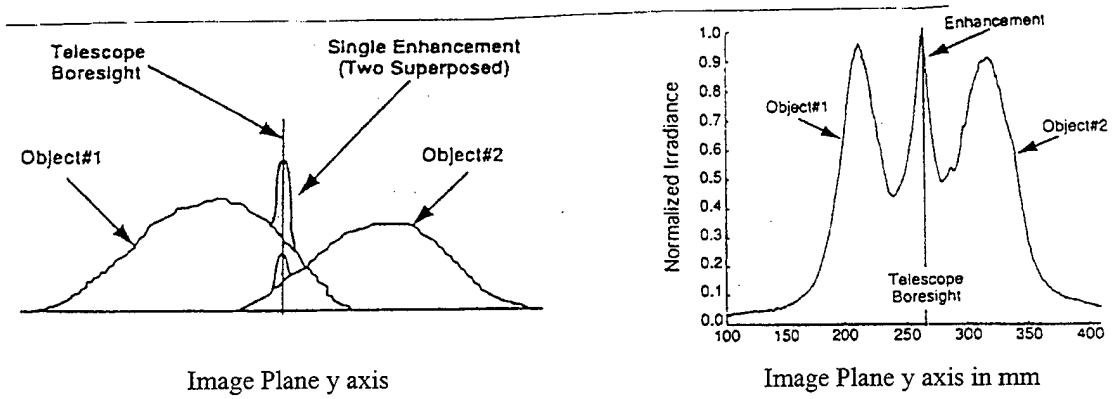


Figure 6.4: Two separate targets result in a single backscatter enhancement. (a) Schematic illustration, (b) Experimentally measured data.

Since the enhancement lies precisely on the boresight of the optical system (the return conjugate wave is a precise retroreflection) and the scattering function is centered on the geometrical image of the target, this RPS phenomena seems particularly suited to pointing and tracking applications.

6.2 Accomplishments

In this program we have extended the concept presented by Solomon and Dainty[14] to include subapertures of finite size for use in reciprocal path imaging applications where high resolution images of extended objects with fine detail are required (i.e., imaging laser radar applications). If there is no *a priori* knowledge concerning the spatial frequency content of the object, it is not sufficient merely to have a non-redundant array of subapertures. The array configuration and the size of the subapertures must be carefully chosen to optimize the spatial frequency response of the system. A method is presented for constructing both one-dimensional and two-

dimensional sparse array configurations which result in a modulation transfer function (MTF) exhibiting continuous and uniform spatial frequency response out to an arbitrarily high cut-off spatial frequency.

We have also exploited the reciprocal path scattering (RPS) phenomenon for precision pointing and tracking applications through a random media such as atmospheric turbulence. It allows interferometric sensitivity in pointing accuracy to be achieved when using a dual aperture and polarization techniques for suppressing unwanted scattered radiation. This new precision pointing and tracking technique has potential applications in ground-to-space laser communications, laser power beaming to satellites, and theater missile defense scenarios.

6.2.1 Reciprocal Path Imaging (RPI) of Extended Objects

In [15], Solomon and Dainty extend the use of polarization for noise suppression in reciprocal path imaging applications to a multiple pinhole aperture, and discuss the possibility of obtaining estimates of many spatial frequencies simultaneously. They depicted the six (6) pinhole aperture illustrated in Figure 6.5. Here the "plus" sign denotes those subapertures with quarter wave plates placed at +45° relative to the linearly polarized illuminating beam, and the "minus" sign denotes those subapertures with quarter wave plates placed at -45°. In this double-pass process, those spatial frequencies (defined by the separation between the pinholes) formed by pinholes having dissimilar quarter wave plates will be imaged by the system, and those formed by pinholes having similar quarter wave plates will be eliminated due to the resulting orthogonal polarizations.

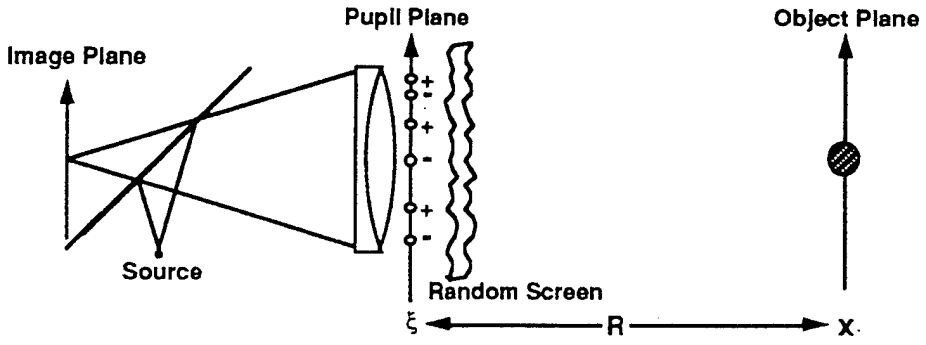


Figure 6.5: A schematic of reciprocal path imaging through a random screen.

They then stated that the number of spatial frequencies defined by a non-redundant polarizing aperture of this type is

$$\left(\frac{n}{2}\right)^2 \quad \text{for even } n \tag{6.1}$$

and

$$\frac{n+1}{2}\ \frac{n-1}{2} \quad \text{for odd } n, \tag{6.2}$$

whereas, for a conventional n pinhole non-redundant aperture, the number of spatial frequencies is given by

$$\frac{n(n-1)}{2} \quad \text{for any } n. \tag{6.3}$$

It is clear that in the limit of large n, the polarizing aperture only images one-half as many spatial frequencies as the conventional (non-polarizing) array.

Since [13] was published, it has become clear that the above technique will indeed produce a stable fringe pattern for each pair of pinholes which have a "plus" and "minus" sign associated with it. However, if more than two pinholes are used, there will also be time-varying cross-terms which corrupt the stable terms.[19] This results in an unstable fringe pattern and time-averaging is again required to obtain a simple relationship between the object spectrum and the image spectrum.

If a continuous spatial frequency response is desired for a given imaging application, one needs to determine the optimum array configuration and then increase the size of the subapertures until the various passbands overlap. For conventional (non-polarizing) sparse imaging arrays, it has been shown that a non-redundant configuration of four square subapertures (n = 4) results in a uniform spatial frequency response out to a cut-off spatial frequency of $L/\lambda f$.[18] Figure 6.6 illustrates a five-element (n = 5) polarizing array which exhibits an identical (non-normalized) MTF except at very low spatial frequencies.

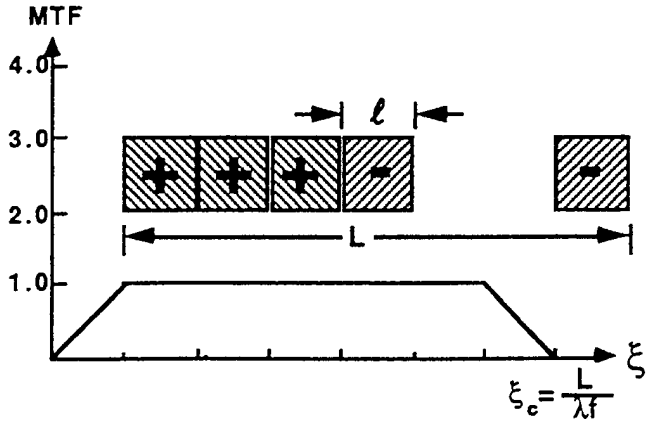


Figure 6.6: A one-dimensional non-redundant polarizing array of square subapertures and the corresponding MTF for reciprocal path imaging applications.

The MTF is given by the autocorrelation of the pupil function. The autocorrelation integral can be interpreted as the area under the product of the pupil function with a shifted version of itself.[20] However, the value of the MTF is now given not merely by the overlapped pupil area, but by the overlapped area of subapertures of opposite signs (polarizations). Hence, at the origin the MTF has a value of zero instead of the usual maximum value proportional to the total area of the pupil.

The (non-normalized) MTF dropping to zero at the origin implies, via the central ordinate theorem of Fourier transform theory, that there is no energy in the point spread function. However, this behavior is merely indicative of the fact that the polarizing aperture is effectively "opaque" to reciprocal path radiation carrying zero spatial frequency information. This absence of zero spatial frequency information can presumably be readily restored from the conventional incoherent image produced by a single subaperture.

Barakat has stated that for $n \geq 5$, it is not possible (even in principle) to obtain this desirable behavior for conventional imaging arrays.[21] However, for RPI applications utilizing polarization for noise suppression, this

behavior can be extended to an arbitrarily high cut-off spatial frequency.[17] A prescription for constructing a one-dimensional array of subapertures exhibiting a completely uniform spatial frequency response is as follows: (1) arrange a string of $n/2$ ($(n+1)/2$ if n is odd) adjacent square subapertures (of width l) with a given polarization, (2) followed by one subaperture of the opposite polarization, (3) followed by $n/2 - 1$ ($(n-1)/2$ if n is odd) blank spaces of width l, (4) followed by one subaperture of the opposite polarization. Repeat steps (3) and (4) until all subapertures have been arranged in this configuration. This will result in an MTF which is constant out to a spatial frequency given by

$$\left(\frac{n}{2}\right)^2 \frac{l}{\lambda f} \qquad \text{for even } n \tag{6.4}$$

and

$$\left(\frac{n+1}{2}\ \frac{n-1}{2}\right)\frac{l}{\lambda f} \qquad \text{for odd } n \tag{6.5}$$

This behavior is consistent with Eqs.(6-1) and (6-2). A nine element non-redundant polarizing array of square subapertures and its associated MTF is shown in Figure 6-7.

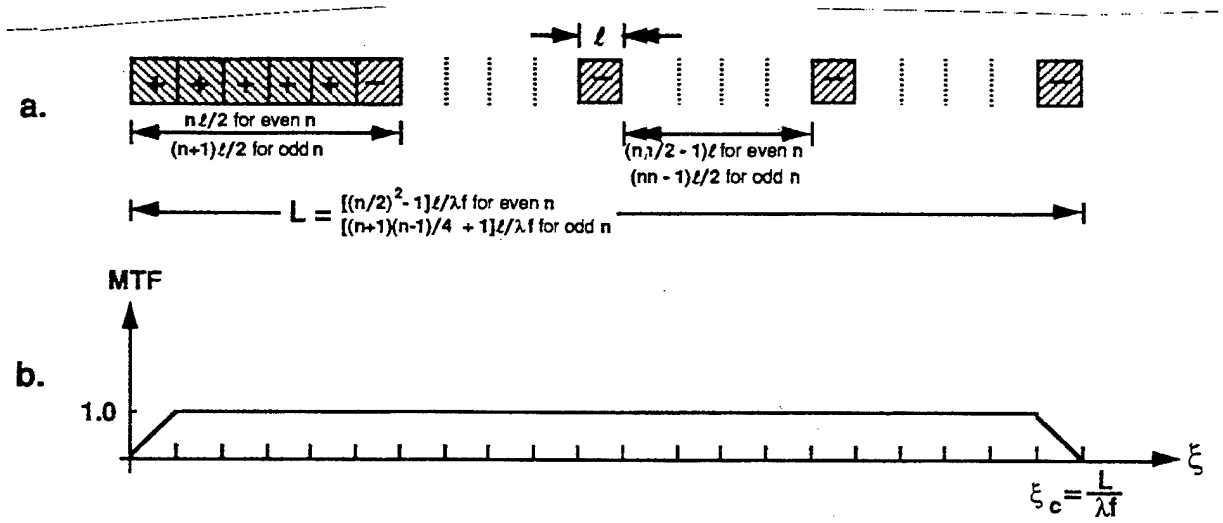


Figure 6.7: (a) A nine-element one-dimensional, non-redundant, polarizing, imaging array of square subapertures, and (b) the corresponding uniform modulation transfer function for reciprocal path imaging applications.

The above behavior for the MTF is readily verified by performing an analytical, numerical, or graphical autocorrelation of the above polarizing aperture. The above prescription for the optimum array configuration was empirically determined to yield the desirable MTF characteristics rather than analytically derived.

For RPI applications utilizing orthogonal polarizations for suppression of noise, the two-dimensional, non-redundant, polarizing array of square subapertures illustrated in Figure 6.8 will produce an MTF with uniform spatial frequency response. This configuration is also extendable to arbitrarily high cut-off spatial frequency (arbitrarily high resolution) by merely increasing the size of the array.

Again the uniform behavior for the above MTF is readily verified by performing an analytical, numerical, or graphical autocorrelation of the above polarizing aperture. The value of the MTF is given by the overlapped area of subapertures of opposite sign (polarization) as one pupil is shifted past the other.

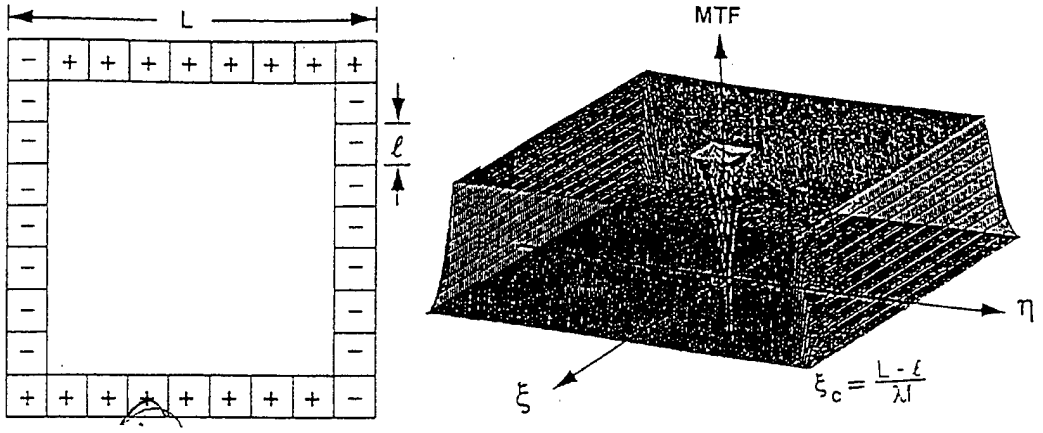


Figure 6.8: (a) A two-dimensional, polarizing array of square subapertures, (b) and its associated MTF exhibiting uniform response over an extended range in spatial frequency space.

This aperture and its MTF is reminiscent of the conventional incoherent MTF of a highly obscured annual aperture. However, it again exhibits a notch at the origin due to the fact that the polarizing aperture is effectively "opaque" to reciprocal path radiation carrying zero spatial frequency information.

The cut-off spatial frequency is given by

$$\xi_c = (L-l)/\lambda f, \tag{6.6}$$

and the number n of subapertures of size ℓ required to achieve that cut-off spatial frequency is given by

$$n = 4(L-l)/l . \tag{6.7}$$

Clearly, the fill factor of the array rapidly decreases with increasing n.

We have thus developed a unique and original prescription for constructing both one-dimensional and two-dimensional dilute array configurations that yield an MTF exhibiting continuous and uniform spatial frequency response up to an arbitrarily high cut-off spatial frequency. Clearly the size of the subapertures for a given application will be limited by the size of the individual phase cells in the random phase screen. The precise effect of specific phase screen characteristics upon the achievable image quality, and the optimum subaperture size (and therefore number) will be the subject of following studies.

6.2.2 Interferometric Sensitivity in Pointing and Tracking Accuracy

For a pointing and tracking application, the enhancement lies precisely on the boresight of the tracking telescope (the return conjugate wave is a precise retro-reflection) and the scattering function is centered on the geometrical image of the target. If the correlation width of the wavefront perturbations are small enough (relative to the wavelength) to produce a scattering cone sufficient to illuminate the target from anywhere within the

transmitting aperture, and if the return radiation from the target fills the telescope aperture, the width of the enhancement will be equal to the diffraction-limited point spread function (PSF) of the telescope.[22] In addition, if the correlation width of the phase errors is small relative to the aperture diameter (there are many effective scattering centers), the height of the enhancement will be approximately twice the height of the scattering function at the location of the enhancement.[8, 23] Figure 6.9 illustrates the experimentally measured intensity profile through the enhancement as the target is displaced relative to the telescope boresight. Note that the enhancement remains fixed on the telescope boresight while the scattering function moves with the geometrical image of the target.

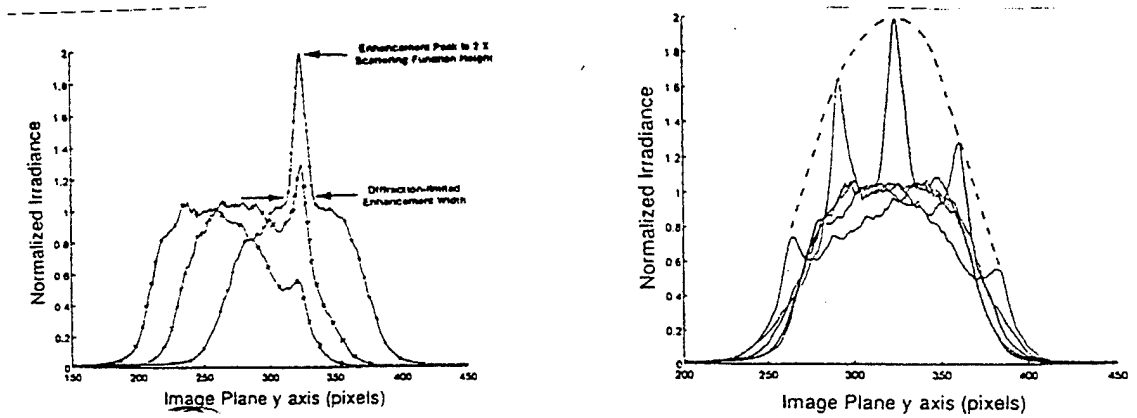


Figure 6.9: (a) Intensity profile through enhancement as target is displaced from the boresight; (b) Superposition of scattering functions indicating envelope of enhancement heights.

The pointing error is given by the displacement between the enhancement and the center of the broad scattering function. This could perhaps be measured two different ways. If we try to measure the displacement directly, we can determine it no more accurately than we can determine the center of the broad scattering function. On the other hand, if we try to maximize the height of the enhancement as we vary the boresight angle, we again have a slowly-varying function that just follows the broad scattering function as illustrated in Figure 6.9. Hence, the presence of the enhancement does not initially appear to improve our pointing capability.

However, if we use a dual aperture as illustrated in Figure 6-10, the enhancement takes the form of a Young's interference pattern.[24] The polarization beamsplitter and 1/4 wave plate allow us to separate the reciprocal path radiation which is insensitive to the random phase variations caused by atmospheric turbulence from the common path radiation which does experience significant scatter effects. The double-pass radiation consists of the following four components; AA(ll), BA($\perp$), AB($\perp$), and BB(ll): where the polarizations indicated in the parentheses are parallel or perpendicular to the y-axis. If the width of the individual subapertures is less than the average turbulence cell size (Fried's r_o)[25], then BA($\perp$) and AB($\perp$) constitute the reciprocal path radiation producing the intensity enhancement; whereas AA(ll) and BB(ll) constitute the common-path radiation which results in the broad scattering function. The polarizing beamsplitter reflects the reciprocal path radiation towards Lens#2 and the CCD camera and transmits the common-path radiation.

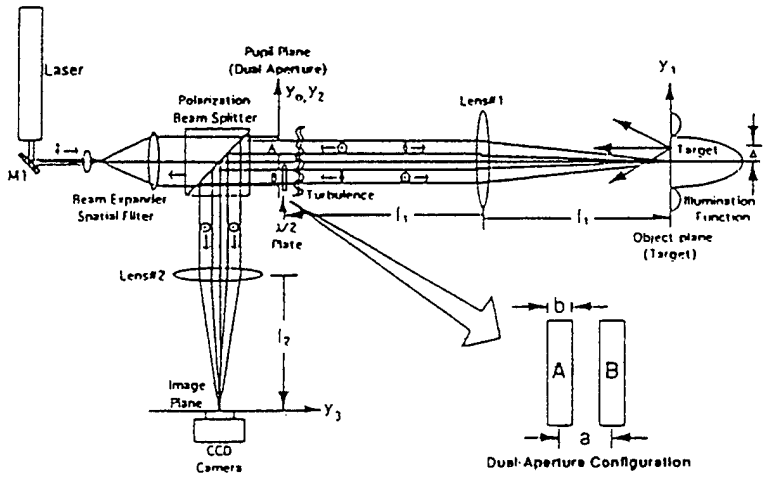


Figure 6.10: Optical layout for precision tracking, utilizing the enhanced backscatter phenomenon with a dual aperture.

Figure 6-11 illustrates a measured irradiance profile (one second exposure time) superposed upon a theoretical prediction for a situation where there is an appreciable pointing error. The agreement between theory and experiment is excellent. The asymmetry of the three-fringe pattern is a measure of the pointing error and can be quantified by defining the following asymmetry parameter

$$R = \frac{I_3 - I_1}{I_3 + I_1} \tag{6.8}$$

where I_1, I_2, and I_3 are the peak irradiance of the three fringes as illustrated in Figure 6.11. This asymmetry parameter can be accurately measured, and is extremely sensitive to the pointing error.

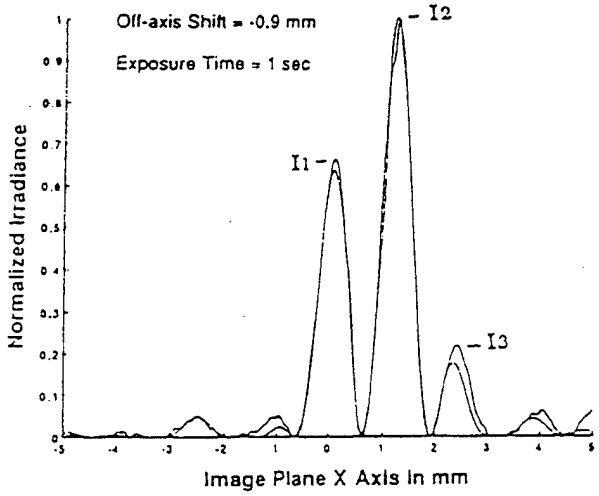


Figure 6.11: Theoretical EBS function (caused by a dual aperture) superposed upon an experimental measurement in the presence of turbulence (one second exposure time). The asymmetry in the side lobes due to pointing error is evident.

Since the Young's fringes remain stable even in the presence of temporally varying random phase errors, the above asymmetry parameter can be plotted versus the telescope pointing error to produce a calibration curve

which indicates that interferometric pointing sensitivity can be achieved with this reciprocal path scattering technique. Figure 6.12 illustrates this interferometric sensitivity, as well as the 2p ambiguity that accompanies many interferometric measurements.

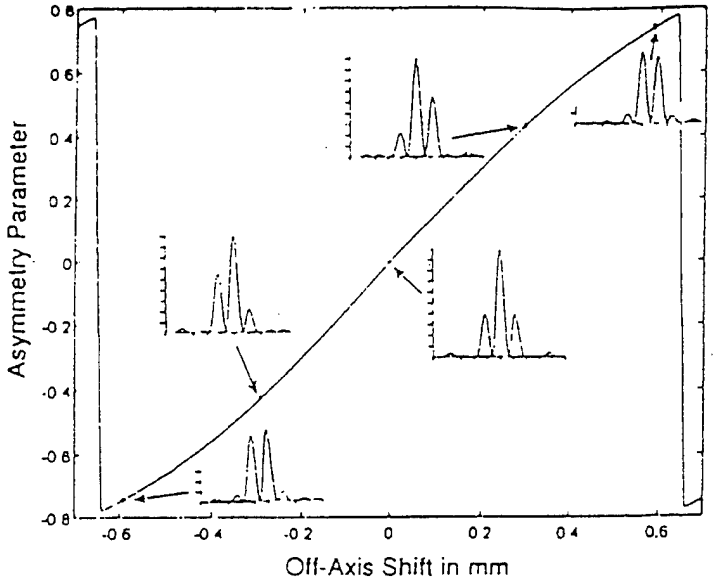


Figure. 6.12: Pointing error calibration curve indicating interferometric accuracy.

An extensive series of CCD images were recorded as the target in the object plane was sequentially displaced by a precise amount with a translation stage fitted with a micrometer adjustment. A propane burner was used to produce the random phase screen and each image was recorded with a ten (10) second exposure time. The asymmetry parameter was calculated and superposed upon the pointing error calibration curve as illustrated in Figure 6-13. Note that the displacement of the target has been divided by the focal length of Lens#1 to convert the ordinate of the calibration curve to angular units.

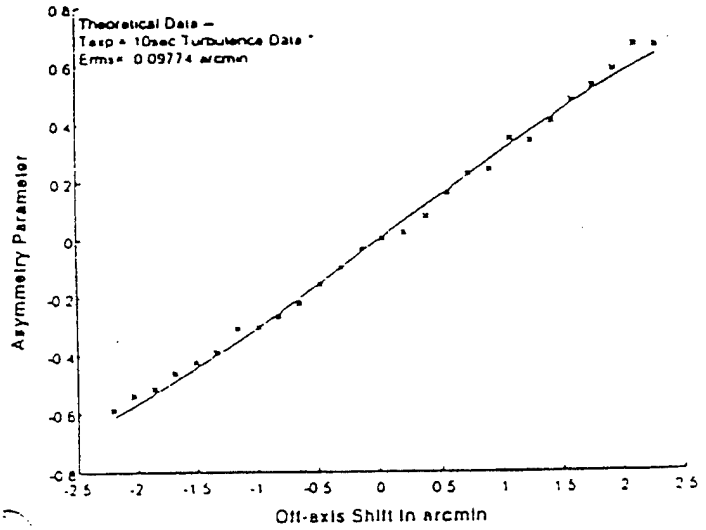


Figure 6.13: Comparison of experimental data with theoretical predictions.

A small adjustment in the slope of the experimental curve was made to correct for a systematic error (average wavefront tilt across the aperture) in the experimental data, and the root-mean-square (rms) error in the measured asymmetry parameter for this set of data was calculated

$$R_{rms} = \sqrt{\frac{(R_{ex} - R_{th})^2}{N}}. \tag{6.9}$$

Dividing this by the slope of the calibration curve yields the accuracy to which the pointing error can be determined by this method. The rms error in the measurement of the telescope pointing error for these twenty-seven (27) data points is less than a tenth of an arc min

$$rms\ Error = 0.09774\ arc\ min. \tag{6.10}$$

This is best expressed as a fraction of the angular width of the diffraction-limited point spread function (PSF) of the telescope aperture. In our laboratory experiment, $\lambda = 0.6328\mu m$ and $D = a + b = 0.65mm$; hence

$$\frac{\alpha_{rms}}{\alpha_{diff}} = \frac{0.09774\ arc\ min}{2.44\,\lambda/D} = 0.012. \tag{6.11}$$

We have thus demonstrated with this laboratory experiment that the EBS phenomenon can be exploited to measure telescope pointing errors through a random media or turbulent atmosphere to *approximately 1% of the angular width of the diffraction-limited PSF of the pointing telescope.*

6.3 Conclusions

We have reviewed the enhanced backscatter, or reciprocal path scattering, phenomenon with regard to pointing, tracking, and imaging applications through atmospheric turbulence. In particular, we developed a unique and original prescription for constructing both one-dimensional and two-dimensional dilute polarizing array configurations that yield an MTF exhibiting continuous and uniform spatial frequency response up to an arbitrarily high cut-off spatial frequency. Clearly the size of the subapertures for a given application will be limited by the size of the individual phase cells in the turbulent atmosphere. The precise effect of specific atmospheric characteristics upon the achievable image quality, and the optimum subaperture size (and therefore number) will be the subject of following studies. In addition, introducing a monostatic dual aperture with subaperture sizes of the order of Fried's r_0 causes the reciprocal path scattered light to produce an intensity enhancement in the form of a temporally stable Young's interference pattern (even in the presence of time-varying turbulence). By choosing the width to separation ratio of the dual aperture appropriately, only three interference fringes exist within the central lobe of the envelope function. The interference fringes are fixed with respect to the telescope boresight and the envelope function moves with the pointing error; hence, the asymmetry of the interference pattern is a direct measure of the telescope pointing error. Polarization techniques are exploited to suppress noise from the common path scattered radiation. Hence, the asymmetry can be measured accurately enough to achieve interferometric sensitivity in pointing error

measurements through atmospheric turbulence. A laboratory proof-of-concept experiment has demonstrated an rms measurement error of approximately 1% of the angular width of the diffraction-limited PSF of the pointing telescope.

We are currently planning to perform field tests to demonstrate precision one-dimensional tracking through long-range distributed atmospheric turbulence. These field tests will be performed at BMDO's ISTEF over horizontal ranges of 200, 500, and 1000 meters. A four (4) watt Argon ion laser and a dual aperture will be used for these experiments. Equipment is being assembled at CREOL on a 3 ft. X 5 ft. optical breadboard to be transported to ISTEF for the planned field experiments.

A logical growth path for this technology would be to perform a two-dimensional ten (10) kilometer precision pointing demonstration at ISTEF, then conduct a precision tracking experiment on a high-altitude balloon, and finally to demonstrate precision tracking of actual space satellites. Two likely applications for this technology is the precision pointing and tracking through a turbulent atmospheric for Airborne Laser (ABL) weapon systems and for high data rate ground-to-space laser communication links.

6.4 References

1. J. S. Preston, "Retro-reflexion by diffusing surfaces", Nature **213**, 1007-1008 (1967).

2. T. S. Trowbridge, "Retroreflection from rough surfaces", J. Opt. Soc. Am. **68**, 1225-1242 (1978).

3. Y. Kuga and A Ishimaru, "Retroreflectance from a dense distribution of spherical particles", J. Opt. Soc. Am. A **1**, 831-835 (1984).

4. L. W. Stockham and T. J. Love, "Investigation of the opposition effect in integrating spheres", J. Opt. Soc. Am. **60**, 251-254 (1970).

5. W. W. Montgomery and R. H. Kohl "Opposition Effect Experimentation", Opt. Lett. **5**, 546-548 (1980).

6. Z.H. Gu, R. S. Dummer, A. A. Maradudin, and A. R. McGurn, "Experimental study of the opposition effect in the scattering of light from a randomly rough metal surface", Appl. Opt. **28**, 537-543 (1989).

7. L. Tsang and A Ishimaru, "Backscatter enhancement of random discrete scatterers", J. Opt. Soc. Am. A **1**, 836-839 (1984).

8. E. Jakeman, "Enhanced backscattering through a deep random phase screen", J. Opt. Soc. Am. A **5**, 1638-1648 (1988).

9. P. R. Tapster, A. R. Weeks and E. Jakeman, "Observation of backscattering enhancement through atmospheric phase screen", J. Opt. Soc. Am. A **6**, 517-522 (1989).

10. G. Welch and R. L. Phillips, "Simulation of enhanced backscatter by a phase screen", J. Opt. Soc. Am. A **7**, 578-584 (1990).

11. M. Nieto-Vesperinas, "Enhanced backscattering", Optics and Photonics News, 50-52 (Dec. 1990).

12. D. A. de Wolf, "Backscatter enhancement: random continuum and particles", J. Opt. Soc. Am. A **8**, 465-471 (1991).

13. T. Mavroidis and J. C. Dainty, "Imaging after double passage through a random screen", Opt. Lett. **15**, 857-859 (1990).

14. C. J. Solomon, J. C. Dainty, and R. G. Lane, "Double passage imaging through a random phase screen using a non-redundant aperture", J. Mod Opt. **10**, 1993-2008 (1991).

15. C. J. Solomon and J. C. Dainty, "Use of polarisation in double passage imaging through a random screen", Optics. Comm. **87**,207-211 (1992).

16. A. N. Bogaturov, A. S. Gurvich, V. A. Myakinin, J. C. Dainty, C. J. Solomon, and N. J. Wooder, "Use of polarization in interferometry after double passage through turbulence", Opt. Lett. **17**, 757-759 (1992).

17. J. E. Harvey and A. Kotha, "Sparse Array Configurations Yielding Uniform MTF's in Reciprocal Path Imaging Configurations', Opt. Com. **106**,178-182 (March 1994).

18. J. E. Harvey, A. Kotha and R. L. Phillips, "Image Characteristics in Applications Utilizing Dilute Subaperture Arrays , Appl. Opt. **34**, 2983-2992 (1995).

19. C. J. Solomon, Blackett Laboratory, Imperial College, London SW7 2BZ, UK (personal communications, 1994).

20. J. D. Gaskill, *Linear Systems, Fourier Transforms, and Optics*, (John Wiley and Sons, New York 1978), Chap. 7, p. 208.

21. R. Barakat, "Dilute aperture diffraction imagery and object reconstruction", Opt. Eng. **29**, 131-139 (1990).

22. A. Dogariu, G. D. Boreman, and M. Dogariu, "Enhanced Backscattering from a Volume-scattering Medium behind a Phase Screen", Opt. Lett. **20**, 1665-1667 (1995).

23. R. C. Heileman and R. L. Phillips, "Experimental Measurements of statistical Properties of Scattered Light due to Double Passage through a Random Phase Screen", Proc. SPIE **1968**-71, 662-672 (1993).

24. J. E. Harvey, S. P. Reddy, and R. L. Phillips, "Precision Pointing and Tracking through Random Media by Exploiting the Enhanced Backscatter Phenomenon", Appl. Opt. **35**, 4220-4228 (1996).

25. D. L. Fried, "Optical Resolution through a Randomly Inhomogeneous Medium for Very Long and Very Short Exposures", J. Opt. Soc. Am **56**, 1372-1379 (1966).

6.5 Publications, Presentations and Theses

1. J. E. Harvey and A. Kotha, "Sparse Array Configurations Yielding Uniform MTF's in Reciprocal Path Imaging Applications", Optics Comm. 106, 178-182, (15 March 1994).

2. J. E. Harvey, A. Kotha, and R. L. Phillips, "Image Characteristics in Applications Utilizing Dilute Subaperture Arrays", Appl. Opt. 34, 2983-2992, (1995).

3. J. E. Harvey and C. Ftaclas, "Field-of-view Limitations of Phased Telescope Arrays", Appl. Opt. 34, 5787-5798, (1995).

4. J. E. Harvey, S. P. Reddy and R. L. Phillips, "Precision Pointing and Tracking through Random Media by Exploiting the Enhanced Backscatter Phenomenon", Appl. Opt. **35**, 4220-4228 (20 July 1996).

5. J. E. Harvey, A. Kotha, and R. L. Phillips, "Fundamental Limitations of Reciprocal Path Imaging through the Atmosphere with Dilute Subaperture Arrays", Proc. SPIE, 2222-62, (1994).

6. A. Kotha, J. E. Harvey, and R. L. Phillips, "Reciprocal Path Imaging: A Technique for the Mitigation of Image Degradation due to Atmospheric Turbulence", Proc SPIE, 2222-61 (1994).

7. J. E. Harvey, S. P. Reddy, and R. L. Phillips, "Precision Pointing and Tracking through Atmospheric Turbulence for Hard Target Imaging Applications", presented at the OSA sponsored Topical Meeting on Coherent Laser Radar, Keystone, Colorado, July 1995.

8. J. E. Harvey, S. P. Reddy, and R. L. Phillips, "Reciprocal Path Scattering through Random Media: Applications to Pointing, Tracking and Imaging", presented at the Annual OSA Meeting, Portland, Oregon, September 1995.

9. P. Reddy, J. E. Harvey, and R. L. Phillips, "Pointing, Tracking, and Acquisition through Atmospheric Turbulence using Reciprocal Path Techniques", presented at SPIE's International Symposium on Optical Engineering in Aerospace Sensing in Orlando, FL (1996); published in Proc. SPIE **2739-04** (1996).

10. J. E. Harvey, R. L. Phillips, C. M. Stickley, and A. Kotha, "Reciprocal Path Imaging through the Atmosphere with Sparse Array Receivers", presented at the 7th Conference on Coherent Laser Radar, Paris, France (1993).

11. S P. Reddy, "Enhanced Backscatter with a Dual Aperture: Precision Pointing and Tracking through a Random Media", M.S. Thesis, University of Central Florida, December 1995.

6.6 Appendix 6A

J. E. Harvey and A. Kotha, "Sparse Array Configurations Yielding Uniform MTF's in Reciprocal Path Imaging Configurations", Opt. Com. **106**,178-182 (March 1994).

Optics Communications 106 (1994) 178–182
North-Holland

OPTICS
COMMUNICATIONS

Sparse array configurations yielding uniform MTFs in reciprocal path imaging applications

J.E. Harvey and Anita Kotha
Center for Research in Electro-Optics and Lasers (CREOL), The University of Central Florida, 12424 Research Parkway, Orlando, FL 32826, USA

Received 14 June 1993; revised manuscript received 8 November 1993

A method is presented for constructing both one-dimensional and two-dimensional configurations of sparse subaperture arrays which result in uniform spatial frequency response with arbitrarily high spatial resolution for reciprocal path imaging applications (i.e., imaging laser radar applications).

1. Introduction

The phenomenon of retroreflection from rough surfaces or suspended particles [1–3], the opposition effect [4–6], enhanced backscatter [7–12], or the double passage effect [13–15] have received a great deal of attention in recent years. This naturally-occurring conjugate-wave phenomenon has been shown to be caused by constructive interference between reciprocal multiple scattering paths. We therefore use the descriptive phrase *reciprocal path imaging* (RPI) to describe attempts to exploit this phenomenon for obtaining diffraction-limited images of extended objects obscured by a random phase screen such as a turbulent atmosphere.

Solomon and Dainty have discussed the use of polarization for noise suppression in reciprocal path imaging through a random screen [15]. Bogaturov et al. have experimentally demonstrated the effectiveness of using orthogonal polarizations to separate the common path radiation from the reciprocal path radiation in a simple Young's interference experiment [16]. This experiment yielded stable, high-contrast interference fringes even in the presence of time-varying turbulence.

In this paper we extend the concept presented by Solomon and Dainty [15] to include subapertures of finite size for use in reciprocal path imaging applications where high resolution images of extended objects with fine detail are required (i.e., imaging laser radar applications). If there is no a priori knowledge concerning the spatial frequency content of the object, it is not sufficient merely to have a non-redundant array of subapertures. The array configuration and the size of the subapertures must be carefully chosen to optimize the spatial frequency response of the system. A method is presented for constructing both one-dimensional and two-dimensional sparse array configurations which result in a modulation transfer function (MTF) exhibiting continuous and uniform spatial frequency response out to an arbitrarily high cut-off spatial frequency.

2. The use of polarization for noise suppression in RPI applications

In ref. [15] Solomon and Dainty extend the use of polarization for noise suppression in reciprocal path imaging applications to a multiple pinhole aperture, and discuss the possibility of obtaining estimates of many spatial frequencies simultaneously. They depicted the six-pinhole aperture illustrated in fig. 1, where the "plus" sign denotes those subapertures with quarter wave plates placed at $+45°$ and the "minus" sign denotes those subapertures with quarter wave plates placed at $-45°$. It is clear from their analysis that those spatial frequencies (defined

Elsevier Science B.V.
SSDI 0030-4018(93)E0582-Z

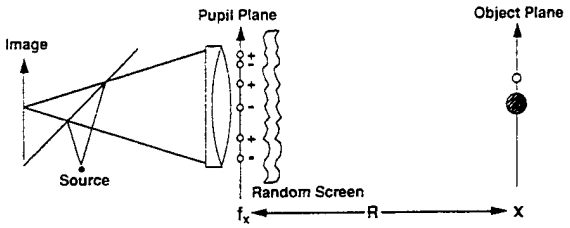


Fig. 1. A schematic of reciprocal path imaging through a random screen.

by the separation between the pinholes) having different signs at the two defining pinholes will be imaged by the system and those having the same sign will be eliminated.

They then stated that the number of spatial frequencies defined by a non-redundant polarizing aperture of this type is

$$\left(\frac{n}{2}\right)^2 \quad \text{for even } n \tag{1}$$

and

$$\frac{n+1}{2}\frac{n-1}{2} \quad \text{for odd } n\,, \tag{2}$$

whereas, for a conventional n-pinhole non-redundant aperture, the number of spatial frequencies is given by

$$\frac{n(n-1)}{2} \quad \text{for any } n\,. \tag{3}$$

It is clear that in the limit of large n, the polarizing aperture only images one-half as many spatial frequencies as the conventional (non-polarizing) array.

Since ref. [15] was published, it has become clear that the above technique will indeed produce a stable fringe pattern for each pair of pinholes which have a "plus" and "minus" sign associated with it. However, if more than two pinholes are used, there will also be time-varying cross-terms which corrupt the stable terms [17]. This results in an unstable fringe pattern and time-averaging is again required to obtain a simple relationship between the object spectrum and the image spectrum. This is not a serious restriction for many practical (laser radar) applications where, in addition to atmospheric turbulence, we have a rapidly-moving target-induced speckle pattern sweeping past our receiver.

3. Spatial frequency response of sparse arrays

For time-averaged images of coherently illuminated rough and moving objects, the MTF is given by the autocorrelation of the complex pupil function. Figure 2a shows this autocorrelation function for the pinhole array illustrated in fig. 1. Note that this MTF consists of nine discrete spatial frequencies as predicted by eq. (1). Finite size subapertures will result in a narrow passband of spatial frequencies at the location of each of the discrete spatial frequencies as shown in fig. 2b.

If a continuous spatial frequency response is desired for a given imaging application, one needs to determine the optimum array configuration and then increase the size of the subapertures until these passbands overlap. For conventional (non-polarizing) sparse imaging arrays, Harvey has shown that the non-redundant configuration of four square subapertures ($n=4$) illustrated in fig. 3a results in a uniform spatial frequency response out to a cut-off spatial frequency of $7L/\lambda f$ [18]. Figure 3b illustrates a five-element ($n=5$) polarizing array which exhibits an identical (non-normalized) MTF except at very low spatial frequencies.

The (non-normalized) MTF dropping to zero at the origin implies, via the central ordinate theorem of Fourier transform theory, that there is no energy in the point spread function. However, this behavior is merely indicative of the fact that the polarizing aperture is effectively "opaque" to reciprocal path ra-

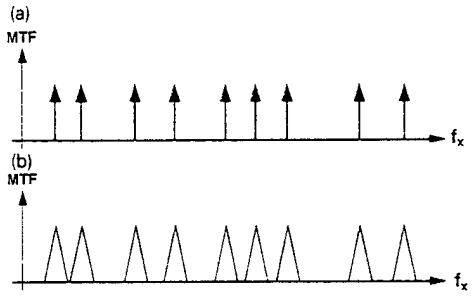


Fig. 2. (a) MTF of the pinhole array illustrated in fig. 1. (b) MTF for finite subapertures for the same array.

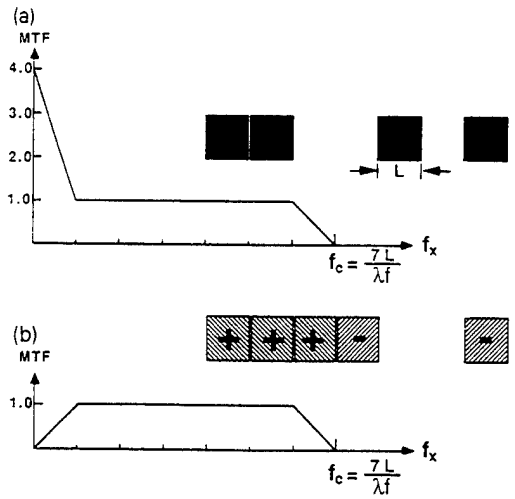


Fig. 3. A one-dimensional non-redundant imaging array of square subapertures and the corresponding (non-normalized) modulation transfer function for (a) conventional imaging, and (b) reciprocal path imaging.

diation carrying zero spatial frequency information.

4. One-dimensional array configurations for RPI applications

Barakat has stated that for $n \geqslant 5$, it is not possible (even in principle) to obtain this desirable behavior for conventional imaging arrays [19]. However, for RPI applications utilizing polarization for noise suppression, this behavior can be extended to an arbitrarily high cut-off spatial frequency. A method for constructing a one-dimensional array of subapertures exhibiting a completely uniform spatial frequency response is as follows: (1) arrange a string of $n/2$ ($(n+1)/2$ if n is odd) adjacent square subapertures (of width L) with a given polarization, (2) followed by one subaperture of the opposite polarization, (3) followed by $n/2-1$ $(n-1)/2$ if (n is odd) blank spaces of width L, (4) followed by one subaperture of the opposite polarization. Repeat steps (3) and (4) until all subapertures have been arranged in this configuration. This will result in an MTF which is constant out to a spatial frequency given by

$$\left(\frac{n}{2}\right)^2 \frac{L}{\lambda f} \quad \text{for even } n \tag{4}$$

and

$$\left(\frac{n+1}{2}\,\frac{n-1}{2}\right)\frac{L}{\lambda f} \quad \text{for odd } n\,. \tag{5}$$

This behavior is consistent with eqs. (1) and (2). A nine-element non-redundant polarizing array of square subapertures and its associated MTF is shown in fig. 4.

Again we note that, whereas the MTF of a conventional imaging array exhibits a peak at the origin with a plateau consisting of various sidebands extending to the cut-off spatial frequency, these polarizing arrays inherently exhibit a notch (zero modulation) at the origin of the MTF. This absence of zero spatial frequency information can presumably be readily restored from the conventional incoherent image produced by a single subaperture.

5. Two-dimensional array configurations for RPI applications

Golay developed a family of two-dimensional, non-redundant arrays with uniform coverage in the spatial frequency plane for conventional imaging applications [20]. For RPI applications utilizing orthogonal polarizations for suppression of noise, the two-dimensional, polarizing array of square subapertures illustrated in fig. 5 will produce an MTF with uniform spatial frequency response. This behavior is reminescent of the conventional incoherent MTF of a highly obscured annular aperture. However, it again exhibits a notch at the origin due to the fact that the polarizing aperture is effectively "opaque" to recip-

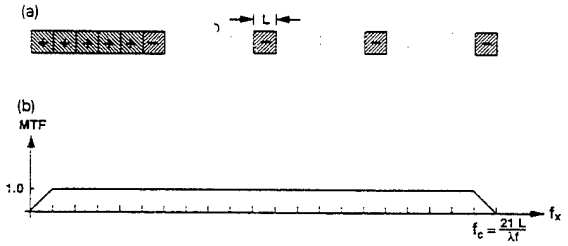


Fig. 4. (a) A nine-element one-dimensional, non-redundant, polarizing, imaging array of square subapertures, and (b) the corresponding uniform modulation transfer function for reciprocal path imaging applications.

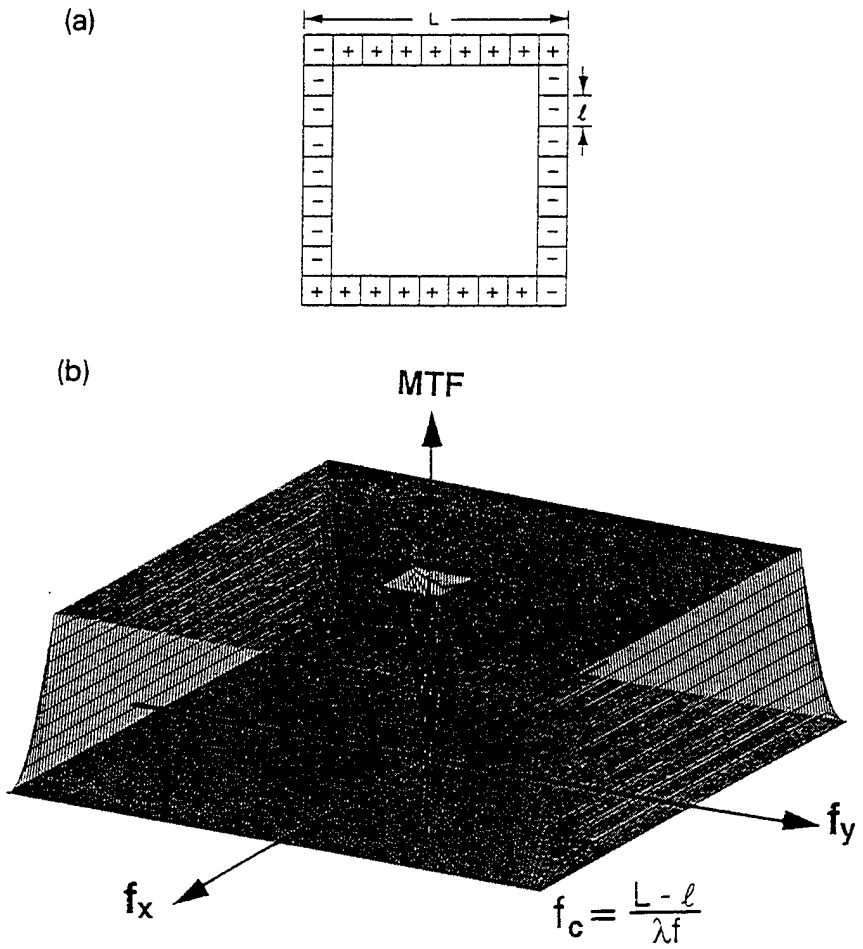


Fig. 5. (a) A two-dimensional, polarizing array of square subapertures, and (b) its associated MTF exhibiting uniform response over an extended range in spatial frequency space.

rocal path radiation carrying zero spatial frequency information. Like the one-dimensional array above, this configuration is also extendible to arbitrarily high cut-off spatial frequency (arbitrarily high resolution) by merely increasing the size of the array.

The cut-off spatial frequency of the MTF characterizing the image quality achievable with this array is given by

$$f_c = \frac{L-l}{\lambda f}, \tag{6}$$

and the number n of subapertures of size l required to achieve that cut-off spatial frequency is given by

$$n = 4(L/l - 1). \tag{7}$$

6. Conclusion

We have described a procedure for constructing both one-dimensional and two-dimensional sparse array configurations which result in a modulation transfer function (MTF) exhibiting continuous and uniform spatial frequency response out to an arbitrarily high cut-off spatial frequency when utilizing orthogonal polarizations for noise suppression in RPI applications. RPI is an approach for obtaining dif-

fraction-limited images of extended objects obscured by a random phase screen such as turbulent atmosphere by exploiting the enhanced backscattering phenomenon (the opposition effect) due to multiple scattering from rough surfaces or suspended particles and double passage through random refractive errors.

Acknowledgements

This research has been supported by the Strategic Defense Initiative Organization, Innovative Science and Technology Office through the Office of Naval Research under contract N00014-92-C0123.

References

[1] J.S. Preston, Nature 213 (1967) 1007.
[2] T.S. Trowbridge, J. Opt. Soc. Am. 68 (1978) 1225.
[3] Y. Kuga and A. Ishimaru, J. Opt. Soc. Am. A 1 (1984) 831.
[4] L.W. Stockham and T.J. Love, J. Opt. Soc. Am. 60 (1970) 251.
[5] W.W. Montgomery and R.H. Kohl, Optics Lett. 5 (1980) 546.
[6] Z.H. Gu, R.S. Dummer, A.A. Maradudin and A.R. McGurn, Appl. Optics 28 (1989) 537.
[7] L. Tsang and A. Ishimaru, J. Opt. Soc. Am. A 1 (1984) 836.
[8] E. Jakeman, J. Opt. Soc. Am. A 5 (1988) 1638.
[9] P.R. Tapster, A.R. Weeks and E. Jakeman, J. Opt. Soc. Am. A 6 (1989) 517.
[10] G. Welch and R.L. Phillips, J. Opt. Soc. Am. A 7 (1990) 578.
[11] M. Nieto-Vesperinas, Optics Photonics News (Dec. 1990) 50.
[12] D.A. de Wolf, J. Opt. Soc. Am. A 8 (1991) 465.
[13] T. Mavroidis and J.C. Dainty, Optics Lett. 15 (1990) 857.
[14] C.J. Solomon, J.C. Dainty and R.G. Lane, J. Mod. Optics 10 (1991) 1993.
[15] C.J. Solomon and J.C. Dainty, Optics Comm. 87 (1992) 207.
[16] A.N. Bogaturov, A.S. Gurvich, V.A. Myakinin, J.C. Dainty, C.J. Solomon and N.J. Wooder, Opt. Lett. 17 (1992) 757.
[17] C.J. Solomon, private communications.
[18] J.E. Harvey and R.A. Rockwell, Opt. Eng. 27 (1988) 762.
[19] R. Barakat, Opt. Eng. 29 (1990) 131.
[20] M.J.E. Golay, J. Opt. Soc. Am. 61 (1971) 272.

6.7 Appendix 6B

J. E. Harvey, A. Kotha and R. L. Phillips, "Image Characteristics in Applications Utilizing Dilute Subaperture Arrays , Appl. Opt. **34**, 2983-2992 (1995).

Image characteristics in applications utilizing dilute subaperture arrays

James E. Harvey, Anita Kotha, and Ronald L. Phillips

When a large aperture is synthesized with an array of smaller subapertures for high-resolution imaging applications, it is important not only to arrange the subapertures to achieve minimal spatial frequency redundancy but also to choose the size of the subapertures (i.e., the dilution ratio) necessary to achieve the best possible image quality. Spurious or ghost images often occur even for nonredundant dilute subaperture arrays. We show that array configurations producing a uniform modulation transfer function will not exhibit these undesirable ghost images. A prescription that is unique and original (to the best of our knowledge) is then presented for constructing both one-dimensional and two-dimensional configurations of dilute subaperture arrays that results in a uniform spatial frequency response with an arbitrarily high spatial resolution for reciprocal path-imaging applications.

Key words: Imaging with sparse arrays, dilute subaperture arrays, nonredundant arrays, modulation transfer functions of dilute arrays, synthetic aperture optics.

1. Introduction

Optical imaging systems utilizing an array of small subapertures to synthesize a conventional large-aperture imaging system have been studied extensively.[1–13] These efforts have been motivated by the phenomenal success of radio astronomers in producing high-resolution images of cosmic radio sources by utilizing dilute antenna arrays.[14] There are a variety of engineering and scientific reasons for considering such systems. Different applications might call for close-packed arrays[15,16] or spatially dilute (sparse, thinned) arrays.[17,18] For dilute arrays, the array configuration (or distribution of subapertures) is of particular importance. It is important not only to arrange the subapertures to achieve minimal spatial frequency redundancy,[19–22] but also to choose the size (i.e., dilution ratio) of the subapertures necessary to achieve the best possible image quality.[23] Spurious or ghost images often occur even for nonredundant dilute subaperture arrays.[24]

Barakat recently stated that, in his opinion, too much emphasis has been placed on the subaperture positioning problem and not enough has been placed on the resultant diffraction imagery.[24] He went on to calculate the diffraction images of a two-bar target for a variety of nonredundant array configurations, and he showed the resulting miserable images, complete with accompanying multiple ghost images.

Here we first review the linear systems theory of image formation and establish the fact that the modulation transfer function (MTF) is the image-quality criterion of choice for a wide variety of applications. The nature of the MTF for redundant and nonredundant subaperture arrays is briefly discussed. We then extend the research of Barakat by calculating and illustrating the point spread function (PSF), MTF, and diffraction images of two-bar targets formed by subaperture arrays with varying dilution ratios. These results demonstrate that the bothersome multiple ghost images reported by Barakat vanish if we optimize the array configuration to produce a MTF that exhibits uniform spatial frequency response.

Finally, we discuss the specific application of reciprocal path imaging (RPI) and recent attempts to exploit this phenomenon to achieve diffraction-limited images of extended objects obscured by a random phase screen, such as a turbulent atmosphere. In particular, we present a prescription for constructing both one-dimensional and two-dimensional polarizing dilute array configurations that result in a MTF that exhibits continuous and uniform spatial frequency response up to an arbitrarily high cutoff spatial frequency when utilized for these RPI applications.

The authors are with the Center for Research in Electro-Optics and Lasers, University of Central Florida, 12424 Research Parkway, Orlando, Florida 32826.

Received 15 November 1993; revised manuscript received 28 October 1994.

0003-6935/95/162983-10$06.00/0.

© 1995 Optical Society of America.

2. Diffraction Imagery with Dilute Subaperture Configurations

Linear systems theory and the associated system transfer function have been widely applied to the characterization of imaging systems.[25–30] If field-dependent aberrations are ignored, the resulting isoplanatic imaging system can be treated as a linear, shift-invariant system. The output of a linear, shift-invariant system is given by the convolution of the input with the system impulse response, or on the application of the convolution theorem of Fourier transform theory, the output spectrum of the system is given by the product of the input spectrum with the system transfer function.[30]

A. Transfer Function Characterization of Imaging Systems

For incoherent imaging systems, the optical transfer function (OTF) is given by the normalized autocorrelation of the complex pupil function[29]:

$$\mathrm{OTF}(\xi, \eta) = \frac{\int_{-\infty}\int^{\infty} P(x', y')P(x' - x, y' - y)\mathrm{d}x'\mathrm{d}y'}{\int_{-\infty}\int^{\infty} |P(x', y')|^2\mathrm{d}x'\mathrm{d}y'},$$

$$x = \lambda f\xi, y = \lambda f\eta, \tag{1}$$

where λ is the wavelength and f is the effective focal length of the system. The modulus of the OTF is called the MTF. For diffraction-limited (aberration-free) imaging systems, this has the geometrical interpretation of being equal to the area of overlap between two displaced pupil functions normalized by the total area of the pupil. The MTF of a conventional imaging system can be thought of as a plot of modulation (or contrast) in the image as a function of spatial frequency. It has a value of unity at the origin (zero spatial frequency) and decreases to zero at the cutoff spatial frequency. This cutoff spatial frequency is given by $\xi_c = L/\lambda f$ for an aperture of width L. The limiting resolution is defined as the reciprocal of this cutoff spatial frequency.

The PSF of an imaging system is given by the squared modulus of the Fourier transform of the complex pupil function,

$$\mathrm{PSF}(x, y) = |A(\xi, \eta)|^2_{\xi=x/\lambda f,\eta=y/\lambda f},$$

$$\text{where } A(\xi, \eta) = \mathcal{F}\{P(x, y)\}. \tag{2}$$

From the autocorrelation theorem of Fourier transform theory, the OTF and the PSF are Fourier transforms of each other.[30] These relationships are illustrated schematically in Fig. 1, along with a variety of commonly used image-quality criteria.

The MTF has gained almost universal acceptance as the image-quality criterion of choice for a wide range of imaging applications involving diffraction-limited optics, or for those suffering from image degradation caused by aberrations. Further image degradation caused by image motion has also been modeled by an appropriate transfer function for image jitter and drift.[31,32]

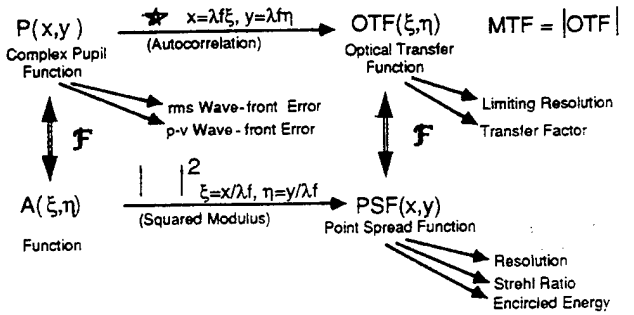


Fig. 1. Relationship between the complex pupil function, OTF, and the PSF.

B. Redundant and Nonredundant Arrays

A potential benefit of dilute subaperture arrays lies in the fact that the resolution of an imaging system is inversely proportional to the diameter of its aperture. In addition, in the basic synthetic aperture concept, when the area of a single aperture is redistributed by the segmentation of the aperture and then the appropriate separation of the segments, the theoretical resolution of that aperture is improved.

Dilute aperture imaging systems will, in general, produce MTF's that exhibit a central peak and a number of discrete spatial frequency passbands surrounded by substantial regions within which the MTF has a value of zero. If one has *a priori* knowledge of the spatial frequency content of an object, then the subaperture array can be tailored to produce a passband in the region of interest. However, for general purpose observations the irretrievable loss of midspatial frequency information is unacceptable, and the practical resolution limit is given not by the reciprocal of the cutoff spatial frequency but rather by the reciprocal of the maximum spatial frequency (ξ_R), below which no zeros occur in the MTF.[23]

The term redundancy refers to the fact that the spatial frequency information may be received redundantly through multiple bandpass filters. Each bandpass filter results from the autocorrelation of one or more equally spaced pairs of subapertures. By merely rearranging the subapertures within a given maximum separation, we can eliminate redundancies and obtain a more uniform coverage below the cutoff spatial frequency. This minimizes the region of zero modulation and improves the possibility of image enhancement by transfer function restoration techniques.

Reference 23 provides a tutorial discussion of redundant and nonredundant arrays, in which several two-dimensional dilute subaperture arrays are illustrated along with their corresponding MTF's. It is emphasized that an even number of subapertures arranged in a circle is a redundant array because there are multiple pairs of subapertures with the same separation and orientation. The MTF consists of a central peak surrounded by a relatively few

prominent passbands, whereas, an odd number of subapertures arranged in a circle produces a MTF with no redundant passbands (although equally spaced, no two pairs of subapertures have the same orientation). Examples from the family of Golay two-dimensional, nonredundant subaperture arrays are also illustrated. They have MTF passbands that are equally spaced, thus providing the most uniform spatial frequency coverage below the cutoff spatial frequency for a fixed aperture area.[21]

Barakat showed that even nonredundant arrays of dilute subapertures can produce miserable images, replete with accompanying multiple ghost images.[24] As we stated in Section 1, in Barakat's opinion, too much emphasis has been placed on the subaperture positioning problem and not enough has been placed on the resultant diffraction imagery. We therefore proceed to show that it is important not only to arrange the subapertures to achieve minimal spatial frequency redundancy, but also to choose the size of the subapertures necessary to achieve the best possible image quality.

We make parametric calculations of the PSF, the MTF, and the diffraction images of a two-bar target as produced by different dilute subaperture arrays. We define a parameter called the dilution ratio as the width, ℓ, of a subaperture divided by the maximum possible width, ℓ_{max}, such that no two subapertures in the array will overlap. This dilution ratio can thus vary from zero to unity for a given subaperture configuration:

$$\text{dilution ratio} \equiv \beta = \ell/\ell_{max}. \tag{3}$$

Figure 2 illustrates several different one-dimensional, four-element, nonredundant arrays spanning a total width L. With each dilution ratio there is an associated fill factor, defined as the ratio of the area of the subapertures in an array to the total area of the synthesized aperture spanned by the array. This

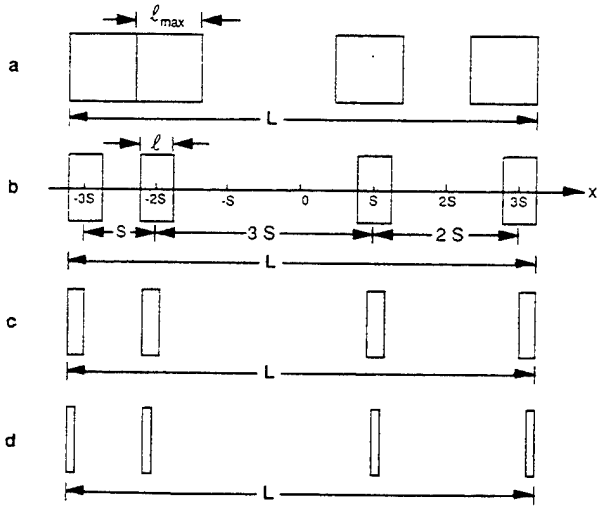


Fig. 2. One-dimensional, four-element, nonredundant array of subapertures spanning a total distance of L, with dilution ratios of a, 1.00; b, 0.50; c, 0.25; d, 0.125.

will be used as a measure of diluteness:

fill factor

$$= \frac{\text{area of subapertures}}{\text{area of full aperture spanned by subapertures}}. \tag{4}$$

For an arbitrary dilution ratio, β, the subaperture width in the four-element nonredundant arrays illustrated in Fig. 2 is given by

$$\ell = \beta \ell_{max} = \beta L/7, \tag{5}$$

and the minimum center-to-center separation for subapertures (of width ℓ) in this configuration spanning a distance L is given by

$$S = (L - \ell)/6. \tag{6}$$

The array in Fig. 2a thus has a dilution ratio of unity (note that the widths of the two subapertures on the left side of the array are equal to their separation and therefore cannot be increased without the subapertures physically overlapping). The fill factor is equal to $4/7$ or 0.571. The array in Fig. 2b exhibits a dilution ratio of 0.50 and a fill factor of 0.286. Figure 2c exhibits a dilution ratio of 0.25 and a fill factor of 0.143, and the final array has a dilution ratio of 0.125 and a fill factor of 0.071. The goal in many cases is to find an array configuration and dilution ratio that will yield a satisfactory image quality with the smallest possible fill factor.

We now proceed to calculate and display the diffraction-limited PSF, the MTF, and the diffraction images of a two-bar target for the set of four-element nonredundant subaperture arrays illustrated in Fig. 2.

C. Point Spread Functions Resulting from Dilute Subaperture Configurations

Because the convolution of a delta function with any function merely replicates that function at the location of the delta function, the aberration-free pupil function of this four-element array can be expressed as

$$P(x, y) = \text{rect}\left(\frac{x}{\ell}\right) * [\delta(x + 3S) + \delta(x + 2S) + \delta(x - S) + \delta(x - 3S)], \tag{7}$$

or we can write this as the convolution of a rectangle function of width ℓ with an even impulse pair (whose Fourier transform we know to be a cosine function) plus a shifted rectangle function convolved with another even impulse pair:

$$P(x, y) = \text{rect}\left(\frac{x}{\ell}\right) * [\delta(x + 3S) + \delta(x - 3S)] + \text{rect}\left(\frac{x + S/2}{\ell}\right) * [\delta(x + 3S/2) + \delta(x - 3S/2)]. \tag{8}$$

Note that we have adopted the symbolic representa-

tion of mathematical functions and operations popularized by Bracewell,[28] Goodman,[29] and Gaskill.[30] This notation has become widely used by electrical and optical engineers because of the ease and convenience with which certain mathematical manipulations can be performed when one is dealing with Fourier analysis and linear systems theory. A brief review defining those functions, operations, and theorems used here is provided in Appendix A.

Aside from some multiplicative phase factors that disappear when we take the squared modulus, we can now write the one-dimensional profile of the amplitude spread function as[29]

$$A(\xi) = \left(\frac{1}{i\lambda f}\right)\{\ell\ \mathrm{sinc}(\ell\xi)2\cos(2\pi 3S\xi) + \ell\ \mathrm{sinc}(\ell\xi)\exp[i2\pi(S/2)\xi]2\cos[2\pi(3/2)S\xi]\}, \tag{9}$$

$$A(\xi) = \left(\frac{2\ell}{i\lambda f}\right)\mathrm{sinc}(\ell\xi)\{\cos(2\pi 3S\xi) + \exp[i2\pi(S/2)\xi]\cos[2\pi(3/2)S\xi]\}. \tag{10}$$

We now obtain the x profile of the resulting PSF by taking the squared modulus of $A(\xi)$ and evaluating at $\xi = x/\lambda f$:

$$\mathrm{PSF}(x) = \left[\frac{4\ell^2}{\lambda^2 f^2}\,\mathrm{sinc}^2(\ell x/\lambda f)\right]\{\cos^2(2\pi 3Sx/\lambda f) + \cos^2[2\pi(3/2)Sx/\lambda f] + 2\cos(2\pi 3Sx/\lambda f)\cos[2\pi(3/2)Sx/\lambda f] \times \cos[2\pi(S/2)x/\lambda f]\}. \tag{11}$$

This function consists of two parts: a sinc^2 envelope function determined solely by the width of a single subaperture, and an interference term that modulates the intensity distribution under that envelope function. Figure 3 schematically illustrates both the pupil function and the corresponding PSF for a dilution ratio of $\beta = 1.0$, 0.5, 0.25, and 0.125. Note that the more dilute the array, the broader the envelope function, and the more interference peaks in the resulting PSF. The PSF curves have all been normalized to the same peak intensity. The behavior illustrated for the highly dilute arrays in Fig. 3 will obviously result in spurious multiple ghost images of fine detail in the image of an extended object.

D. Modulation Transfer Functions Resulting from Dilute Subaperture Configurations

The normalized autocorrelation of the pupil function described by Eq. (7) is given by

$$\mathrm{MTF}(\xi) = \mathrm{tri}\left(\frac{\xi}{\ell/\lambda f}\right) * \left\{\delta(\xi) + \frac{1}{4}\sum_{n=1}^{6}\left[\delta\left(\xi - \frac{nS}{\lambda f}\right) + \delta\left(\xi + \frac{nS}{\lambda f}\right)\right]\right\}. \tag{12}$$

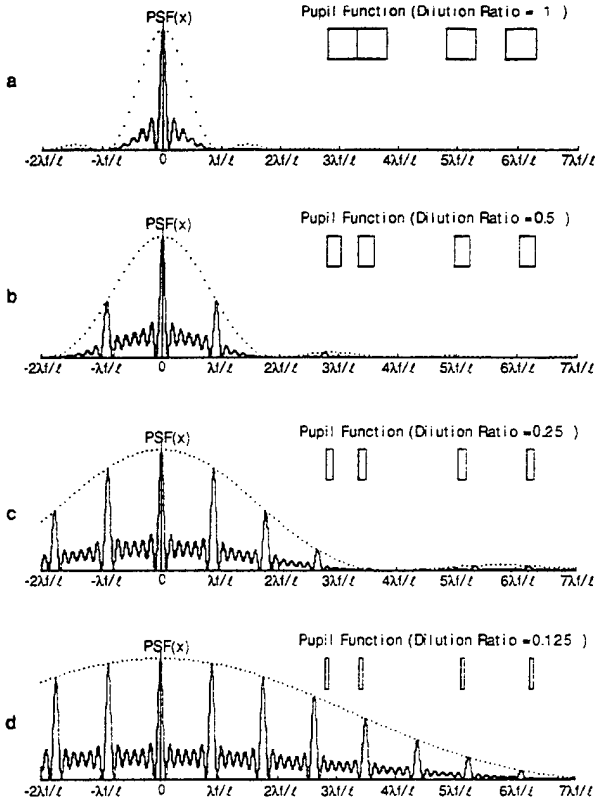


Fig. 3. Graphic illustration of the pupil function and the corresponding diffraction-limited PSF for dilution ratios of 1.00, 0.50, 0.25, and 0.125.

Figure 4 illustrates the pupil function and the corresponding MTF for dilution ratios of 1.0, 0.5, 0.25, and 0.125. Note that the cutoff spatial frequency is constant for all cases because we held the distance spanned by the array fixed. The very small dilution ratios result in a MTF consisting of widely separated discrete passbands. There are gaps (values of zero) between the passbands for dilution ratios less than 0.5. For $\beta > 0.5$ the passbands overlap, eliminating all gaps in the MTF below the cutoff spatial frequency. For $\beta = 1.0$ the spatial frequency response is perfectly uniform between $L/7\lambda f \le \xi \le 6L/7\lambda f$.

E. Images Resulting from Dilute Subaperture Configurations

Before we investigate the diffraction images of a two-bar target as produced by dilute subaperture configurations, let us first look at the images resulting from conventional full apertures of different sizes. The diffraction-limited image of an extended object is obtained when the geometrical image is convolved with the diffraction-limited PSF. Figure 5 illustrates the two-bar target along with the PSF and the diffraction image resulting from one-dimensional apertures of width $\ell_{\max}$, $2\ell_{\max}$, $4\ell_{\max}$, and $7\ell_{\max} = L$. Normalizing the x axis by $\lambda f/4\ell_{\max} = 7\lambda f/4L$, we choose an object defined by

$$\mathrm{Object}(x) = \mathrm{rect}\left(\frac{x - 0.5}{0.5}\right) + \mathrm{rect}\left(\frac{x + 0.5}{0.5}\right). \tag{13}$$

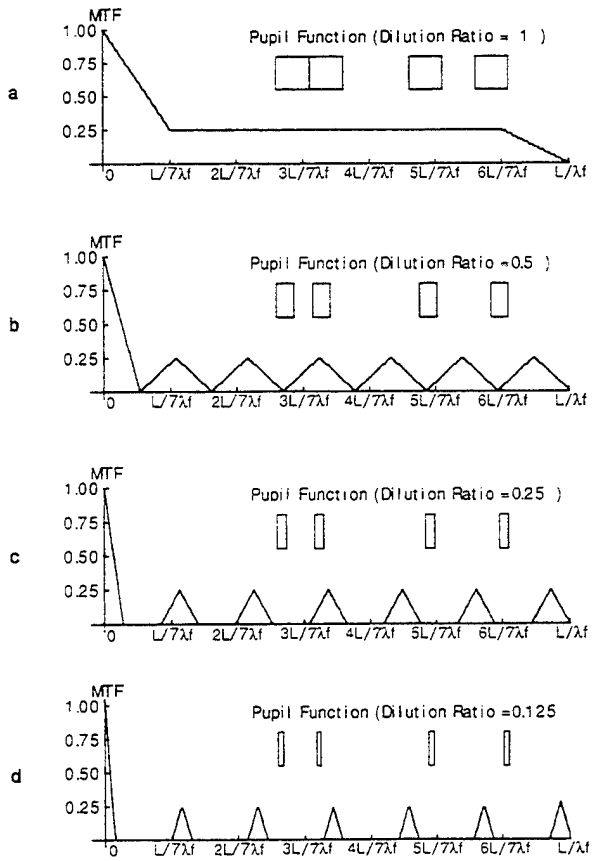


Fig. 4. Graphic illustration of the pupil function and the corresponding diffraction-limited MTF for dilution ratios of 1.00, 0.50, 0.25, and 0.125.

From Fig. 5a it is evident that an aperture of width ℓ_{max} does not even come close to resolving the two bars making up the object. Figure 5c indicates that it takes an aperture at least $4\ell_{max}$ wide to just barely resolve the target, and Figure 5d illustrates the well-resolved image obtained with a full aperture of width L.

Figure 6 now illustrates the diffraction-limited images produced by dilute subaperture arrays spanning the total width L. These images were calculated when a numerical convolution of the PSF was performed with the two-bar target (object) illustrated in Fig. 5. The dilute arrays producing these images are identical to those of Figs. 3 and 4.

F. Discussion of Imaging Behavior

The above results illustrate that the severity of the multiple ghost images reported by Barakat increases with decreasing dilution ratio. These dilute arrays all have the same optimum nonredundant configuration (uniform distribution of passbands), and they all span the same total width. They differ only in the degree of diluteness (dilution ratio or fill factor). A careful comparison of these diffraction images with the MTF's of the corresponding arrays indicates that the ghost images vanish if we optimize the array to produce an MTF exhibiting uniform spatial frequency response. One can do this for a given nonredundant array configuration merely by increasing the

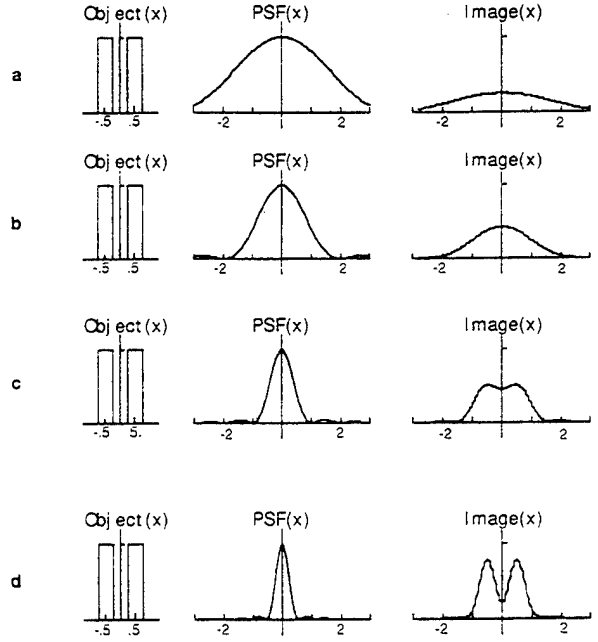


Fig. 5. Illustration of the object, the PSF, and the image produced by apertures of width a, ℓ_{max}; b, $2\ell_{max}$; c, $4\ell_{max}$; d, $7\ell_{max} = L$.

subaperture size. The results of the previous analysis indicate that highly dilute arrays should be avoided if high-resolution images are to be obtained of extended objects containing fine detail with high contrast. In this situation, multiple ghost images produced by a given feature could easily be mistaken for actual features in the object.

The phenomenal success of radio astronomers in achieving extremely high-fidelity and high-dynamic-range images by using very dilute telescope arrays is

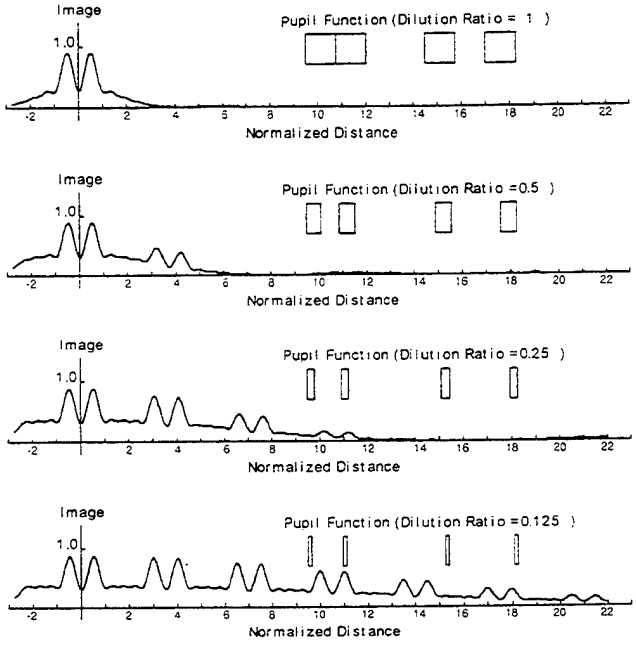


Fig. 6. Graphic illustration of the pupil function and the corresponding diffraction images of a two-bar target for dilution ratios of 1.00, 0.50, 0.25, and 0.125.

due to the fact that their application provides them with (a) high signal-to-noise observations (there is no photon-noise penalty for subdividing the aperture or for low-visibility objects), (b) astronomical observations that allow for long exposure times with aperture synthesis by rotation of the Earth, reconfiguration of the array, or both, and (c) the luxury of time to perform postdetection processing with many powerful image deconvolution algorithms. Many tactical and strategic defense applications of interest do not have any of these advantages.

We now apply the results of the previous analysis to the special case of RPI applications with dilute array receivers. In this particular application, polarizing subaperture arrays can be utilized to obtain diffraction-limited information concerning objects obscured by a random phase screen. A prescription is presented for constructing both one-dimensional and two-dimensional configurations of dilute polarizing subaperture arrays that result in a uniform spatial frequency response with arbitrarily high spatial resolution.

3. Optimum Subaperture Array Configurations for Reciprocal Path Imaging Applications

The phenomenon of retroreflection from rough surfaces or suspended particles,[33–35] the opposition effect,[36–38] enhanced backscatter,[39–44] or the double-passage effect[45–47] has received a great deal of attention in recent years. This naturally occurring, partial conjugate-wave phenomenon has been shown to be caused by constructive interference between reciprocal multiple-scattering paths as illustrated in Fig. 7.

We introduce the descriptive phrase reciprocal path scattering when referring to this general phenomenon, and we reserve the phrase reciprocal path imaging to describe attempts to exploit this phenomenon for obtaining diffraction-limited images of extended objects obscured by a random phase screen, such as a turbulent atmosphere.

Solomon and Dainty discussed the use of polarization for noise suppression in RPI through a random phase screen.[47] A direct optical image is not obtained with this process; however, diffraction-limited information concerning the object is obtained in the form of the modulus of the Fourier spectrum of the object. Hence the actual image must be reconstructed from that information. Bogaturov *et al.* experimentally demonstrated the effectiveness of using orthogonal polarizations to separate the common path radiation from the reciprocal path radiation in a simple Young's interference experiment.[48] This experiment yielded stable, high-contrast interference fringes even in the presence of time-varying turbulence.

From the previous analysis, it is clear that an MTF which is uniform up to the required cut-off spatial frequency is highly desirable. We therefore apply the previous analysis and extend the concept of polarizing subapertures presented in Ref. 47 to include subapertures of finite size for use in RPI applications where high resolution images of extended objects are required.

A. Use of Polarization for Noise Suppression in Reciprocal Path Imaging Applications

Solomon and Dainty[47] extended the use of polarization for noise suppression in RPI applications to a multiple-pinhole aperture, and they discussed the possibility of obtaining estimates of many spatial frequencies simultaneously. They depicted the six-pinhole aperture illustrated in Fig. 8. Here the plus sign denotes those subapertures with quarter-wave plates placed at $+45°$ relative to the linearly polarized illuminating beam, and the minus sign denotes those subapertures with quarter-wave plates placed at $-45°$. In this double-pass process, those spatial frequencies (defined by the separation between the pinholes) formed by pinholes having dissimilar quarter-wave plates will be imaged by the system, and those formed by pinholes having similar quarter-wave plates will be eliminated because of the resulting orthogonal polarizations.

Solomon and Dainty then stated that the number of spatial frequencies defined by a nonredundant polarizing aperture of this type is

$$\left(\frac{n}{2}\right)^2 \quad \text{for even } n, \tag{14}$$

$$\frac{n+1}{2}\frac{n-1}{2} \quad \text{for odd } n, \tag{15}$$

whereas for a conventional n-pinhole nonredundant

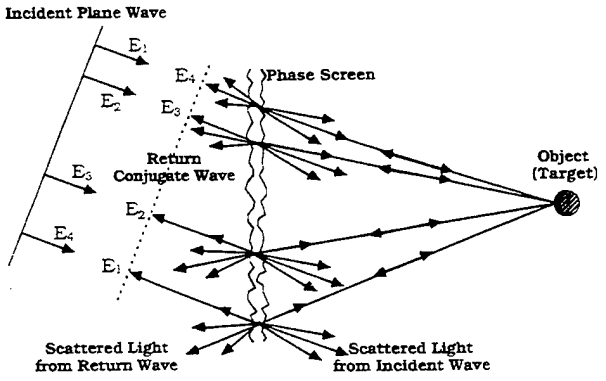


Fig. 7. Conjugate-wave formation by reciprocal path scattering through a random phase screen or turbulent medium.

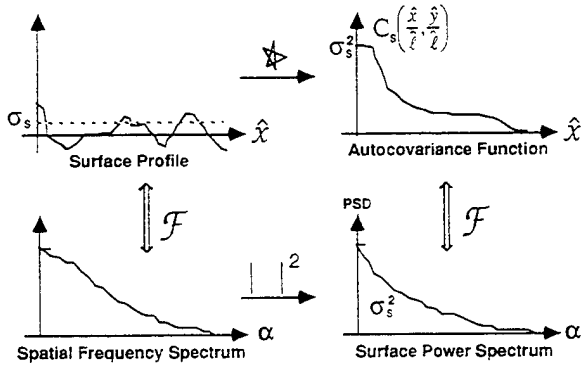


Fig. 8. Schematic of RPI through a random screen.

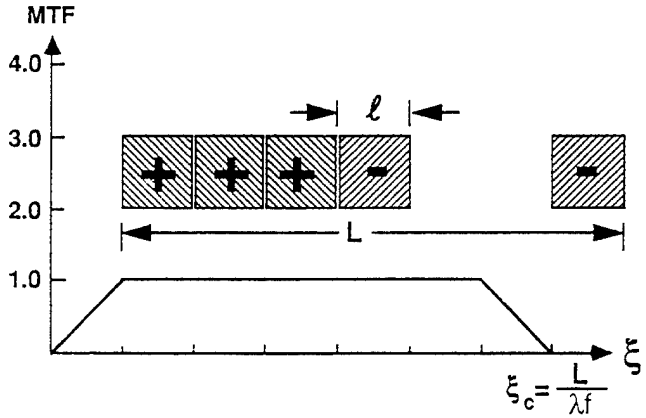


Fig. 9. One-dimensional nonredundant polarizing array of square subapertures and the corresponding MTF for RPI applications.

aperture, the number of spatial frequencies is given by

$$\frac{n(n-1)}{2} \quad \text{for any } n. \tag{16}$$

It is clear that in the limit of large n, the polarizing aperture only images one half as many spatial frequencies as the conventional (nonpolarizing) array.

Since Ref. 45 was published, it has become clear that the above technique will indeed produce a stable fringe pattern for each pair of pinholes that has a plus and minus sign associated with it. However, if more than two pinholes are used, there will also be time-varying cross terms that corrupt the stable terms.[49] This results in an unstable fringe pattern, and time averaging is again required to obtain a simple relationship between the object spectrum and the image spectrum.

B. One-dimensional Array Configurations for Reciprocal Path Imaging Applications

If a continuous spatial frequency response is desired for a given imaging application, one needs to determine the optimum array configuration and then increase the size of the subapertures until the various passbands overlap. For conventional (nonpolarizing) sparse imaging arrays, we have shown that the nonredundant configuration of four square subapertures ($n = 4$) illustrated in Fig. 4a results in a uniform spatial frequency response out to a cutoff spatial frequency of $L/\lambda f$. Figure 9 illustrates a five-element ($n = 5$) polarizing array that exhibits an identical (nonnormalized) MTF except at very low spatial frequencies.

The MTF is given by the autocorrelation of the pupil function [recall Eq (1)]. The autocorrelation integral can be interpreted as the area under the product of the pupil function with a shifted version of itself.[30] However, the value of the MTF is now given not merely by the overlapped pupil area, but instead by the overlapped area of subapertures of opposite signs (polarizations). Hence, at the origin the MTF has a value of zero instead of the usual maximum value proportional to the total area of the pupil.

The (nonnormalized) MTF dropping to zero at the origin implies, by means of the central ordinate theorem of Fourier transform theory, that there is no energy in the PSF. However, this behavior is merely indicative of the fact that the polarizing aperture is effectively opaque to reciprocal path radiation carrying zero spatial frequency information. This absence of zero spatial frequency information can presumably be readily restored from the conventional incoherent image produced by a single subaperture.

Barakat[24] stated that for $n \geq 5$, it is not possible (even in principle) to obtain this desirable behavior for conventional imaging arrays. However, for RPI applications utilizing polarization for noise suppression, this behavior can be extended to an arbitrarily high cutoff spatial frequency.[50] A prescription for constructing a one-dimensional array of subapertures exhibiting a completely uniform spatial frequency response is as follows: (a) arrange a string of $n/2$

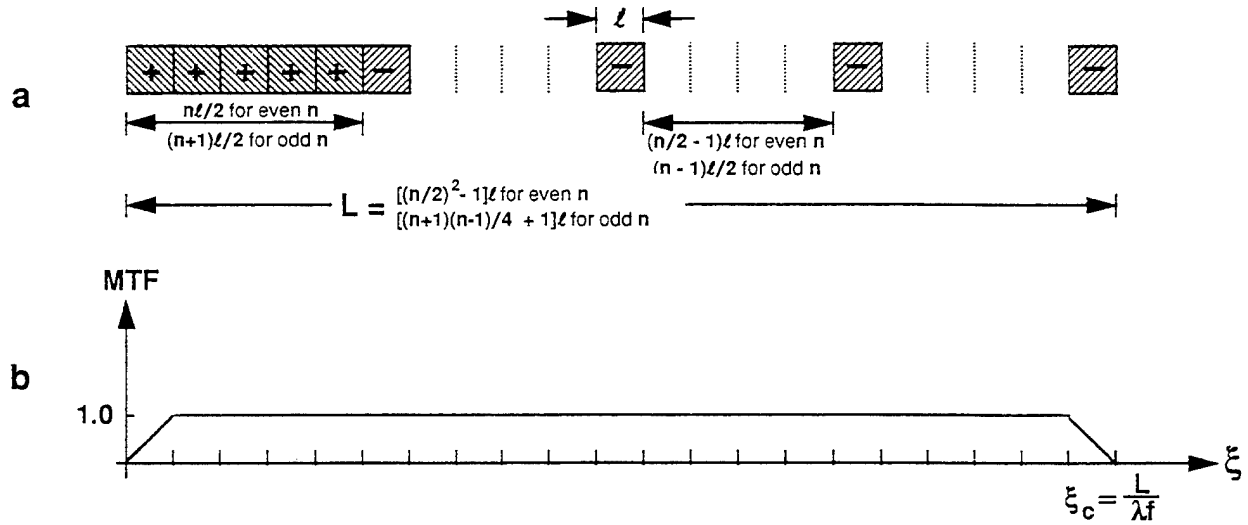


Fig. 10. a, Nine-element one-dimensional, nonredundant, polarizing imaging array of square subapertures; b, the corresponding uniform MTF for RPI applications.

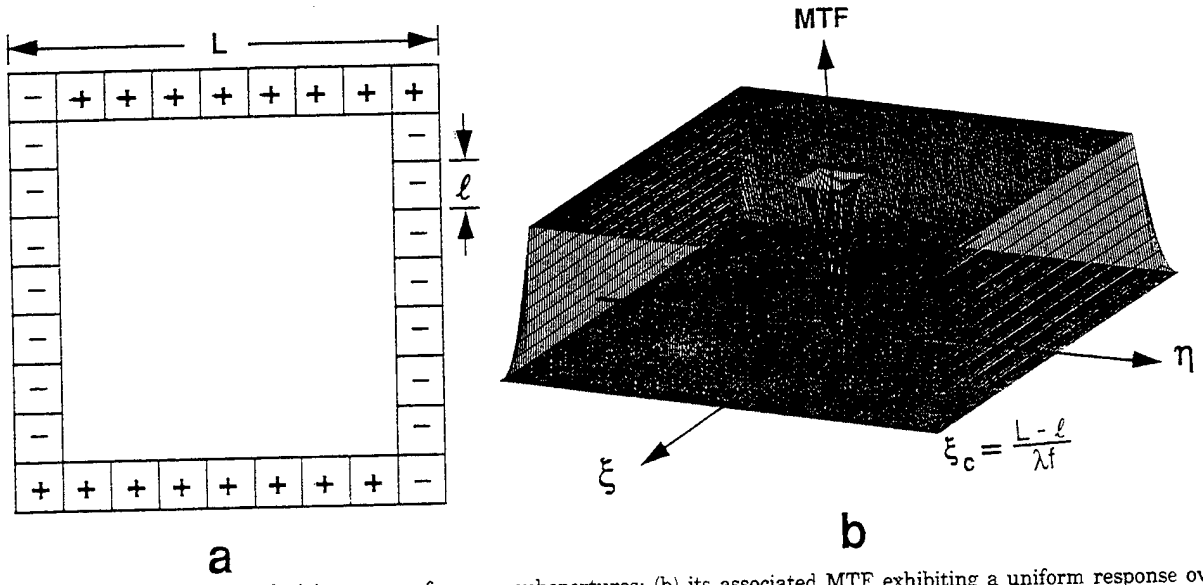


Fig. 11. a, Two-dimensional polarizing array of square subapertures; (b) its associated MTF exhibiting a uniform response over an extended range in spatial frequency space.

[$(n + 1)/2$ if n is odd] adjacent square subapertures (of width ℓ) with a given polarization, (b) follow this with one subaperture of the opposite polarization, (c) follow this with $n/2 - 1$ [$(n - 1)/2$ if n is odd] blank spaces of width ℓ, (d) follow this with one subaperture of the opposite polarization. Repeat steps (c) and (d) until all subapertures have been arranged in this configuration. This will result in an MTF that is constant out to a spatial frequency given by

$$\left(\frac{n}{2}\right)^2 \frac{\ell}{\lambda f} \qquad \text{for even } n, \tag{17}$$

$$\left(\frac{n+1}{2}\frac{n-1}{2}\right)\frac{\ell}{\lambda f} \qquad \text{for odd } n. \tag{18}$$

This behavior is consistent with conditions (14) and (15). A nine-element nonredundant polarizing array of square subapertures and its associated MTF is shown in Fig. 10.

The above behaivor for the MTF is readily verified when an analytical, numerical, or graphical autocorrelation of the above polarizing aperture is performed. The above prescription for the optimum array configuration was empirically determined to yield the desirable MTF characteristics rather than analytically derived.

C. Two-dimensional Array Configurations for Reciprocal Path Imaging Applications

For RPI applications utilizing orthogonal polarizations for suppression of noise, the two-dimensional, nonredundant, polarizing array of square subapertures illustrated in Fig. 11 will produce a MTF with a uniform spatial frequency response. One can also extend this configuration to an arbitrarily high cutoff spatial frequency (arbitrarily high resolution) merely by increasing the size of the array. Again, the uniform behavior for the above MTF is readily verified when an analytical, numerical, or graphical autocorrelation of the above polarizing aperture is performed. The value of the MTF is given by the overlapped area of subapertures of opposite sign (polarization) as one pupil is shifted past the other.

This aperture and its MTF is reminiscent of the conventional incoherent MTF of a highly obscured annular aperture. However, it again exhibits a notch at the origin because the polarizing aperture is effectively opaque to reciprocal path radiation carrying zero spatial frequency information. The cutoff spatial frequency is given by

$$\xi_c = (L - \ell)/\lambda f, \tag{19}$$

and the number n of subapertures of size ℓ required to achieve that cutoff spatial frequency is given by

$$n = 4(L - \ell)/\ell. \tag{20}$$

Clearly, the fill factor of the array rapidly decreases with increasing n.

4. Summary

After a brief discussion of redundant and nonredundant arrays, a linear systems approach was utilized to calculate the image characteristics of dilute aperture configurations. Detailed parametric performance predictions of the PSF, the MTF, and the diffraction-limited image of a two-bar target were calculated for a four-element, nonredundant subaperture configuration as a function of dilution ratio. Multiple ghost images were seen to occur for the highly dilute arrays, and it was observed that a uniform spatial frequency response is required to eliminate these undesirable image characteristics.

Hence, when utilizing dilute aperture configurations to produce direct optical images of extended objects containing fine detail, we find it is not sufficient merely to use an optimum nonredundant array configuration. The size of the subapertures, or the dilution ratio, is also a crucial parameter affecting the resulting image quality. Furthermore, the uniformity of the MTF representing a subaperture array is an excellent indicator of the achievable image quality.

Finally, we applied the above analysis to RPI applications and developed a prescription for constructing both one-dimensional and two-dimensional dilute array configurations that yield an MTF exhibiting a continuous and uniform spatial frequency response up to an arbitrarily high cutoff spatial frequency. Clearly, the size of the subapertures for a given application will be limited by the size of the individual phase cells in the random phase screen. The precise effect of specific phase screen characteristics on the achievable image quality, and the optimum subaperture size (and therefore number) will be the subject of following studies.

Appendix A. Symbolic Representation of Mathematical Functions

The symbolic representation of mathematical functions is as follows:

$$\text{rectangle function} \equiv \text{rect}\left(\frac{x}{\ell}\right) = \begin{cases} 0 & \text{for } |x| > \frac{\ell}{2}, \\ \frac{1}{2} & \text{for } |x| = \frac{\ell}{2}, \\ 1 & \text{for } |x| < \frac{\ell}{2}; \end{cases}$$

$$\text{triangle function} \equiv \text{tri}\left(\frac{x}{\ell}\right) = \begin{cases} 1 - |x| & \text{for } |x| < \ell, \\ 0 & \text{for } |x| \geq \ell; \end{cases}$$

$$\text{sinc function} \equiv \text{sinc}\left(\frac{x}{\ell}\right) = \frac{\sin(\pi x/\ell)}{\pi x/\ell};$$

$$\text{Dirac delta function} \equiv \delta(x - a).$$

The convolution operation is

$$f(x) * g(x) = \int_{-\infty}^{\infty} f(\alpha)g(x - \alpha)\mathrm{d}\alpha,$$

and the cross-correlation operation is

$$f(x) \star g(x) = \int_{-\infty}^{\infty} f(\alpha)g(\alpha - x)\mathrm{d}\alpha.$$

The Fourier transform pairs are defined as

$$F(\xi, \eta) = \int_{-\infty}^{\infty}\int f(x, y)\exp[-i2\pi(x\xi + y\eta)]\mathrm{d}x\mathrm{d}y,$$

$$f(x, y) = \int_{-\infty}^{\infty}\int F(\xi, \eta)\exp[i2\pi(x\xi + y\eta)]\mathrm{d}\xi\mathrm{d}\eta.$$

Examples of these are as follows:

$$f(x, y) \Leftrightarrow F(\xi, \eta), \quad \text{or} \quad F(\xi, \eta) = \mathcal{F}\{f(x, y)\},$$
$$\text{rect}(x, y) \Leftrightarrow \text{sinc}(\xi, \eta),$$
$$\text{tri}(x, y) \Leftrightarrow \text{sinc}^2(\xi, \eta),$$
$$2\cos(2\pi ax) \Leftrightarrow \delta(\xi - a) + \delta(\xi + a),$$
$$\delta(x) \Leftrightarrow 1.$$

The similarity theorem is

$$\text{given} \quad f(x) \Leftrightarrow \mathrm{F}(\xi),$$
$$\text{then} \quad f(x/a) \Leftrightarrow |a|F(a\xi);$$

the shift theorem is

$$\text{given} \quad f(x) \Leftrightarrow F(\xi),$$
$$\text{then} \quad f(x - x_0) \Leftrightarrow \exp(-i2\pi x_0\xi)F(\xi);$$

and the convolution theorem is

$$\text{given} \quad f(x) \Leftrightarrow F(\xi), \quad g(x) \Leftrightarrow G(\xi),$$
$$\text{then} \quad f(x) * g(x) = F(\xi)G(\xi).$$

This research was supported by the Ballistic Missile Defense Organization, Innovative Science and Technology Directorate, through the U.S. Office of Naval Research under contract N00014-92-C-0123.

References

1. *Synthetic Aperture Optics* (National Academy of Sciences–National Research Council, Washington, D.C., 1967).
2. A. B. Meinel, "Aperture synthesis using independent telescopes," Appl. Opt. **9,** 2501–2504 (1970).
3. M. W. Stockton, ed., "Symposium on synthetic aperture optics," Tech. Rep. 58 (Optical Sciences Center, University of Arizona, Tucson, Ariz., 1970).
4. J. W. Goodman, "Synthetic aperture optics," in *Progress in Optics*, E. Wolf, ed. (North-Holland, Amsterdam, 1970), Vol. 8, pp. 3–48.
5. B. M. Boyce, "The role of imaging processing in synthetic aperture systems and aberrated axially symmetric systems," in *A Symposium on Sampled Images* (Perkin-Elmer, Norwalk, Conn., 1971), pp. 313–327.
6. A. B. Meinel, R. R. Shannon, F. L. Whipple, and F. J. Low, "A large multiple-mirror telescope project," in *Instrumentation in Astronomy,* L. Larmore and R. W. Poindexter, eds., Proc. Soc. Photo-Opt. Instrum. Eng. **28,** 155–160 (1972).
7. G. M. Sanger, T. E. Hoffman, and M. A. Reed, "Some design aspects of a multiple mirror telescope," in *Instrumentation in*

Astronomy, L. Larmore and R. W. Poindexter, eds., Proc. Soc. Photo-Opt. Instrum. Eng. **28,** 161–171 (1972).
8. H. P. Gush, "Optical imaging using aperture synthesis," J. Opt. Soc. Am. **69,** 187–191 (1979).
9. J. S. Fender, ed., *Synthetic Aperture Systems*, Proc. Soc. Photo-Opt. Instrum. Eng. **440,** 1–172 (1984).
10. R. B. Johnson, W. L. Wolfe, and J. S. Fender, eds., *Infrared, Adaptive, and Synthetic Aperture Optical Systems*, Proc. Soc. Photo-Opt. Instrum. Eng. **643,** 121–243 (1986).
11. W. A. Traub, "Combining beams from separated telescopes," Appl. Opt. **25,** 528–532 (1986).
12. J. M. Beckers, "Field of view considerations for telescope arrays," in *Advanced Technology Optical Telescopes III*, L. D. Barr, ed., Proc. Soc. Photo-Opt. Instrum. Eng. **628,** 255 (1986).
13. J. S. Fender, ed., *Multiple-Aperture Optical Systems*, special issue of Opt. Eng. **27,** 705–800 (1988).
14. G. W. Swenson, Jr., "Radio-astronomy precedent for optical interferometer imaging," J. Opt. Soc. Am. **3,** 1311–1319 (1986).
15. J. E. Harvey, M. J. MacFarlane, and J. L. Forgham, "Design and performance of ranging telescopes: monolithic versus synthetic aperture," Opt. Eng. **24,** 183–188 (1985).
16. M. H. Krim, "Applications of replicated glass mirrors to large segmented optical systems," in *Large Optics Technology*, G. M. Sanger, ed., Proc. Soc. Photo-Opt. Instrum. Eng. **571,** 60–75 (1986).
17. W. A. Traub and W. F. Davis, "Coherent optical system of modular imaging collectors (COSMIC) telescope array: astronomical goals and preliminary image reconstruction results," in *Advanced Technology Optical Telescopes*, L. D. Barr and G. Burbidge, eds., Proc. Soc. Photo-Opt. Instrum. Eng. **332** 164–175 (1982).
18. J. E. Harvey, A. B. Wissinger, and A. N. Bunner, "A parametric study of various synthetic aperture telescope configurations for coherent imaging applications," in *Infrared, Adaptive, and Synthetic Aperture Optical Systems*, J. S. Fender, R. B. Johnson, and W. L. Wolfe, eds., Proc. Soc. Photo-Opt. Instrum. Eng. **643,** 194–207 (1985).
19. F. Gori and G. Guattari, "Imaging systems using linear arrays of nonequally spaced elements," Phys. Lett. A **32,** 38–39 (1970).
20. F. Gori and G. Guattari, "Optical analog of a nonredundant array," Phys. Lett. A **32,** 446–447 (1970).
21. M. J. E. Golay, "Point arrays having compact nonredundant autocorrelations," J. Opt. Soc. Am. **61,** 272–273 (1971).
22. F. D. Russell and J. W. Goodman, "Nonredundant arrays and postdetection processing for aberration compensation in incoherent imaging," J. Opt. Soc. Am. **61,** 182–191 (1971).
23. J. E. Harvey and R. A. Rockwell, "Performance characteristics of phased array and thinned aperture optical telescopes," Opt. Eng. **27,** 762–768 (1988).
24. R. Barakat, "Dilute aperture diffraction imagery and object reconstruction," Opt. Eng. **29,** 131–139 (1990).
25. R. V. Shack, "Outline of practical characteristics of an image forming system," J. Opt. Soc. Am. **46,** 755–757 (1956).
26. E. Ingelstam, "Nomenclature for Fourier transforms of spread functions," J. Opt. Soc. Am. **51,** 1441 (1961).
27. F. D. Smith, "Optical image evaluation and the transfer function," Appl. Opt. **2,** 335–350 (1963).
28. R. N. Bracewell, *The Fourier Transform and its Applications* (McGraw-Hill, New York, 1965), Chap. 4, p. 51.
29. J. W. Goodman, *Introduction to Fourier Optics* (McGraw-Hill, New York, 1968), Chap. 2, pp. 13–21.
30. J. D. Gaskill, *Linear Systems, Fourier Transforms, and Optics* (Wiley, New York, 1978), Chap. 7, p. 208.
31. S. C. Som, "Analysis of the effect of linear smear on photographic images," J. Opt. Soc. Am. **61,** 859–864 (1971).
32. W. Swindell, "A noncoherent optical analog image processor," Appl. Opt. **9,** 2459–2469 (1970).
33. J. S. Preston, Retro-reflexion by diffusing surfaces," Nature **213,** 1007–1008 (1967).
34. T. S. Trowbridge, "Retroreflection from rough surfaces," J. Opt. Soc. Am. **68,** 1225–1242 (1978).
35. Y. Kuga and A. Ishimaru, "Retroreflectance from a dense distribution of spherical particles," J. Opt. Soc. Am. A **1,** 831–835 (1984).
36. L. W. Stockham and T. J. Love, "Investigation of the opposition effect in integrating spheres," J. Opt. Soc. Am. **60,** 251–254 (1970).
37. W. W. Montgomery and R. H. Kohl, "Opposition effect experimentation," Opt. Lett. **5,** 546–548 (1980).
38. Z. H. Gu, R. S. Dummer, A. A. Maradudin, and A. R. McGurn, "Experimental study of the opposition effect in the scattering of light from a randomly rough metal surface," Appl. Opt. **28,** 537–543 (1989).
39. L. Tsang and A. Ishimaru, "Backscatter enhancement of random discrete scatterers," J. Opt. Soc. Am. A **1,** 836–839 (1984).
40. E. Jakeman, "Enhanced backscattering through a deep random phase screen," J. Opt. Soc. Am. A **5,** 1638–1648 (1988).
41. P. R. Tapster, A. R. Weeks, and E. Jakeman, "Observation of backscattering enhancement through atmospheric phase screen," J. Opt. Soc. Am. A **6,** 517–522 (1989).
42. G. Welch and R. L. Phillips, "Simulation of enhanced backscatter by a phase screen," J. Opt. Soc. Am. A **7,** 578–584 (1990).
43. M. Nieto-Vesperinas, "Enhanced backscattering," Opt. Photon. News **1,** 50–52 (1990).
44. D. A. de Wolf, "Backscatter enhancement: random continuum and particles," J. Opt. Soc. Am. A **8,** 465–471 (1991).
45. T. Mavroidis and J. C. Dainty, "Imaging after double passage through a random screen," Opt. Lett. **15,** 857–859 (1990).
46. C. J. Solomon, J. C. Dainty, and R. G. Lane, "Double passage imaging through a random phase screen using a nonredundant aperture," J. Mod. Opt. **10,** 1993–2008 (1991).
47. C. J. Solomon and J. C. Dainty, "Use of polarisation in double passage imaging through a random screen," Optics. Commun. **87,** 207–211 (1992).
48. A. N. Bogaturov, A. S. Gurvich, V. A. Myakinin, J. C. Dainty, C. J. Solomon, and N. J. Wooder, "Use of polarization in interferometry after double passage through turbulence," Opt. Lett. **17,** 757–759 (1992).
49. C. J. Solomon, Blackett Laboratory, Imperial College, London SW7 2BZ, UK (personal communication), 1994.
50. J. E. Harvey and A. Kotha, "Sparse array configurations yielding uniform MTF's in reciprocal path imaging configurations," Opt. Commun. **106,** 178–182 (1994).

6.8 Appendix 6C

J. E. Harvey, S. P. Reddy, and R. L. Phillips, "Precision Pointing and Tracking through Random Media by Exploiting the Enhanced Backscatter Phenomenon", Appl. Opt. **35**, 4220-4228 (1996).

Precision pointing and tracking through random media by exploitation of the enhanced backscatter phenomenon

James E. Harvey, Stephen P. Reddy, and Ronald L. Phillips

The active illumination of a target through a turbulent medium with a monostatic transmitter–receiver results in a naturally occurring conjugate wave caused by reciprocal scattering paths that experience identical phase variations. This reciprocal path-scattering phenomenon produces an enhanced backscatter in the retroverse direction (precisely along the boresight of the pointing telescope). A dual aperture causes this intensity enhancement to take the form of Young's interference fringes. Interference fringes produced by the reciprocal path-scattering phenomenon are temporally stable even in the presence of time-varying turbulence. Choosing the width-to-separation ratio of the dual apertures appropriately and utilizing orthogonal polarizations to suppress the time-varying common-path scattered radiation allow one to achieve interferometric sensitivity in pointing accuracy through a random medium or turbulent atmosphere. Computer simulations are compared with laboratory experimental data. This new precision pointing and tracking technique has potential applications in ground-to-space laser communications, laser power beaming to satellites, and theater missile defense scenarios.

Key words: Precision pointing and tracking, reciprocal path scattering, enhanced backscatter.
© 1996 Optical Society of America

1. Introduction

The ultimate success or failure of many military, space, and scientific programs relies on the accuracy of optical pointing and tracking subsystems. Optical tracking technology advanced rapidly during World War II, when gun directors, optical sighting systems, and infrared devices were developed. Since that time, the missile and space age has become the prime motivation for the continued development of optical tracking systems. Although this research field has historically been driven by defense-related applications, it is now poised for many dual-use and commercial applications, such as laser communication links, laser radar applications, and laser power beaming to satellites. There seems to be a continual demand for higher pointing accuracy in the presence of various environmental error sources, such as atmospheric turbulence and vibration-induced jitter. A brief review of the historical development of optical tracking was provided in a special issue of *Applied Optics* in 1966.[1] A 1993 issue of *Optical Engineering* brings us up to date on more recent technology developments, with a guest editorial and 20 separate papers on acquisition, pointing, and tracking.[2]

This paper deals specifically with a new technique that exploits the enhanced backscatter (EBS) phenomenon that results from reciprocal scattering paths when the active illumination of a target through a random refractive medium is used. The use of a dual aperture permits interferometric sensitivity in pointing accuracy to be achieved when polarization techniques are employed to suppress unwanted scattered radiation. First the EBS phenomenon is reviewed; then the use of polarization for noise suppression during the use of a dual aperture is discussed; finally, the pointing sensitivity is modeled and compared with the results of laboratory experiments.

2. Enhanced Backscatter Phenomenon

The phenomenon of retroreflection from rough surfaces or suspended particles,[3–5] the opposition effect,[6–8] enhanced backscatter,[9–14] the double-passage

The authors are with the Center for Research and Education in Optics and Lasers, University of Central Florida, P.O. Box 162700, Orlando, Florida 32816.

Received 30 August 1995; revised manuscript received 7 February 1996.

0003-6935/96/214220-09$10.00/0
© 1996 Optical Society of America

effect,[15–18] or reciprocal path imaging[19,20] has received a great deal of attention in recent years. This naturally occurring, partial conjugate-wave phenomenon has been shown to be caused by constructive interference produced by multiple reciprocal scattering paths through a random refractive phase screen, as illustrated in Fig. 1.

If we consider two arbitrary ray paths (say E_1 and E_4) orthogonal to the incident plane wave, they will be scattered by the random phase screen as indicated schematically in Fig. 1. One scattered component in each of the resulting angular spectrum of ray paths will be directed toward the target and reflected (scattered) back along the incoming path of a companion scattered component. They will each be scattered a second time on the return trip when they encounter the random phase screen. Again, one scattered component in each of the resulting angular spectrum of ray paths will be directed back precisely antiparallel to the incident companion ray path. Assuming that the temporal variations of the random phase screen are slow relative to the round trip time of flight, we see that light propagating along any two reciprocal scattering paths will clearly experience identical phase delays because of their double passage through the same two points on the random phase screen. Because this situation holds for a myriad of possible ray pairs associated with the incident wave front, a multitude of conjugate waves (perfect plane waves) are formed (each having its own total phase delay) and propagate back toward the transmitting aperture, along with all of the returned scattered radiation.

A monostatic laser imaging configuration such as that illustrated schematically in Fig. 2 is required for the reciprocal path effect to exist and produce a conjugate return wave. The resulting irradiance pattern in the image plane is a time-varying speckle pattern caused by the random phase screen. However, when many realizations of the speckle pattern are averaged, a narrow enhancement will form on the smoothed scattered light distribution. Figure 3 illustrates a computer simulation of this phenomena

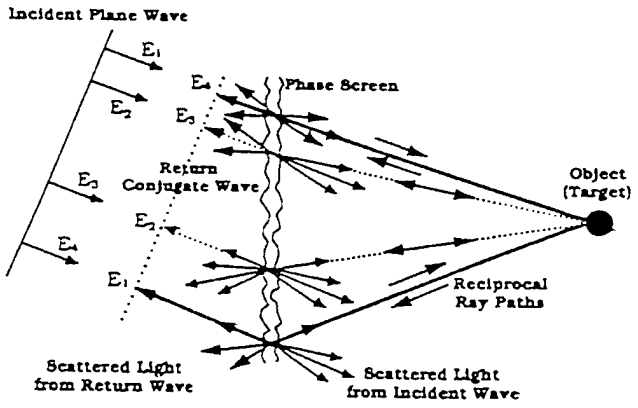


Fig. 1. Conjugate-wave formation by RPS through a random phase screen or turbulent medium.

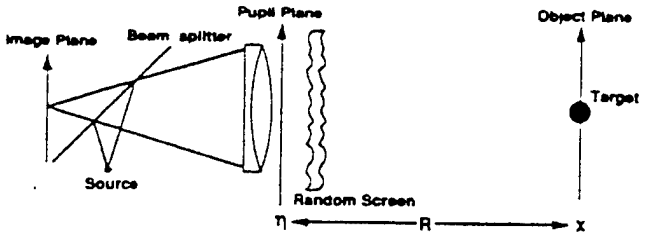


Fig. 2. Monostatic laser imaging configuration used to demonstrate the EBS phenomenon.

that represents the average of 110 realizations of the speckle pattern.[12] The target was assumed to be an optically smooth mirror in the far field of the transmitting aperture. A deliberate pointing error (incident angle) of approximately 0.02° is evident from the position of the enhancement relative to the center of the scattering function.

Figure 4 shows the experimental results of a laboratory experiment in which a propane burner was used to create the time-varying random phase screen. A helium–neon laser was used as a source for illuminating the target. The target was a narrow rectangular mirror whose width was small compared with the size of the illuminating beam. After reflecting from the target in the object plane, the return radiation passed through the random phase screen a second time, was focused onto the image plane, and was recorded with a CCD camera with a 1.0-s exposure time. By translating the target slightly from the center of the illuminating beam, the recorded image was qualitatively very similar to the previous computer simulation. The greater width of the measured enhancement is due to diffraction from the finite extent of the incident beam.

One must be careful not to confuse the backscatter enhancement with an *image* of the target. The scattering function is the (turbulence-induced) blurred image of the target. If two separate targets were illuminated and imaged by the tracking telescope,

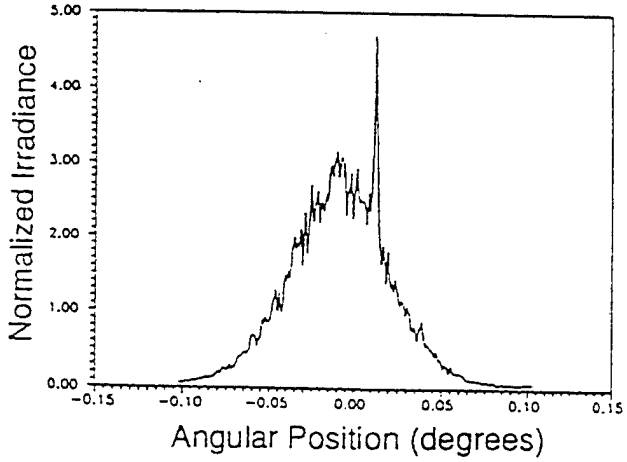


Fig. 3. Computer simulation of the EBS phenomenon (110 realizations).

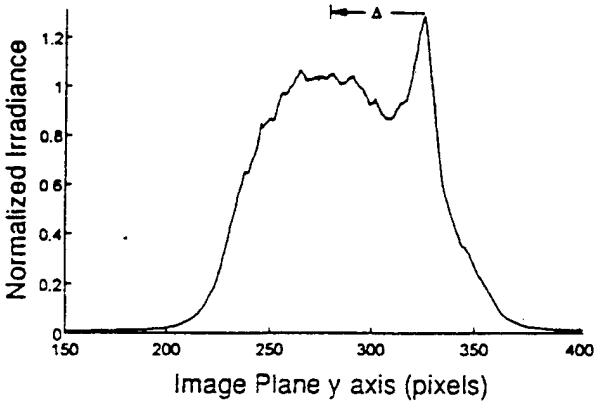


Fig. 4. Measured irradiance profile through the enhancement (1.0-s exposure time).

there would not be two separate enhancements; instead, there would be two separate scattering functions with a single enhancement on the boresight of the telescope, as illustrated schematically in Fig. 5. Experimental data indicating a single enhancement caused by two separate targets are illustrated in Fig. 6.

A direct optical image is thus not formed by this EBS process; however, information concerning the spatial frequency content of the image does exist in the reciprocal path radiation, and image reconstruction algorithms can be implemented to recover an image of an extended object.[16,18–19] We introduce the descriptive phrase reciprocal path scattering (RPS) when referring to this general conjugate-wave phenomenon, and we reserve the phrase reciprocal path imaging to describe attempts to exploit this phenomenon for obtaining diffraction-limited images of extended objects actively illuminated through a random medium. Furthermore, polarizing sparse array configurations have been described that yield a uniform spatial frequency response in reciprocal path imaging applications.[19,20]

For a pointing and tracking application, the enhancement lies precisely on the boresight of the

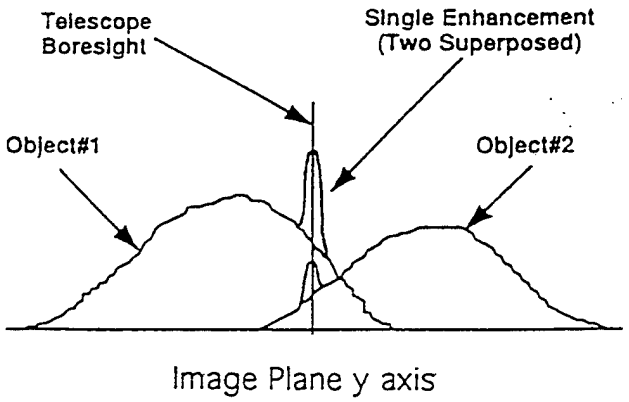


Fig. 5. Schematic illustration, showing that two separate targets result in two broad scattering functions and a single backscatter enhancement.

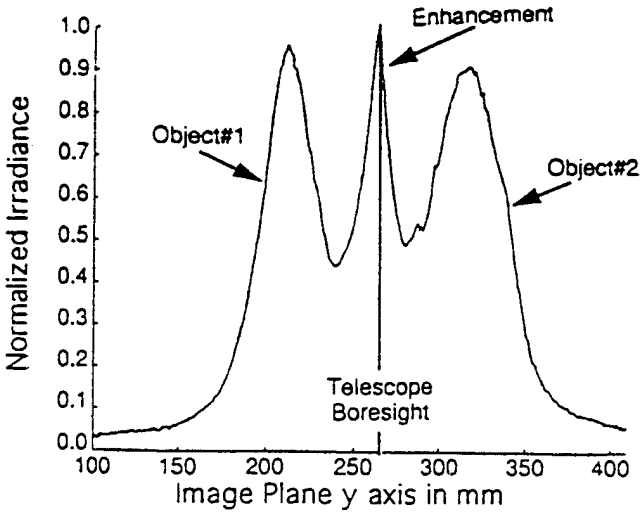


Fig. 6. Experimentally measured profile, indicating two separate scattering functions and a single backscatter enhancement.

tracking telescope (the return conjugate wave is a precise retroreflection), and the scattering function is centered on the geometrical image of the target. If the correlation width of the wave-front perturbations is small enough (relative to the wavelength) to produce a scattering cone sufficient to illuminate the target from anywhere within the transmitting aperture, and if the return radiation from the target fills the telescope aperture, then the width of the enhancement will be equal to the diffraction-limited point spread function (PSF) of the telescope.[21] In addition, if the correlation width of the phase errors is small relative to the aperture diameter (there are many effective scattering centers), the height of the enhancement will be approximately twice the height of the scattering function at the location of the enhancement.[10,22] Figure 7 illustrates the experimentally measured intensity profile through the

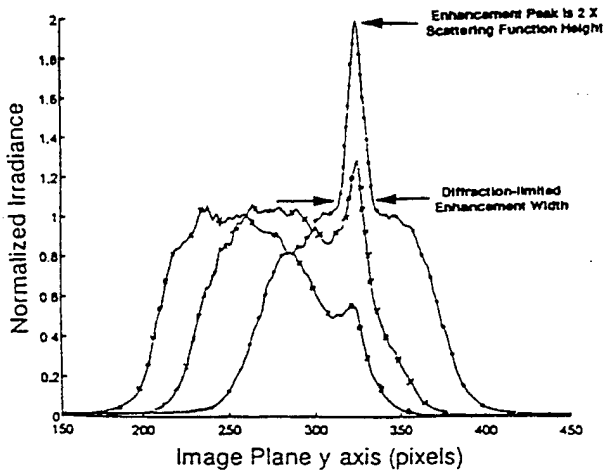


Fig. 7. Experimentally measured intensity profile through the enhancement as the target is displaced from the boresight. Note that the enhancement remains fixed on the telescope boresight and the scattering function shifts with the geometrical image of the target.

enhancement as the target is displaced relative to the telescope boresight. Note that the enhancement remains fixed on the telescope boresight while the scattering function moves with the geometrical image of the target.

The pointing error is given by the displacement between the enhancement and the center of the broad scattering function. This could perhaps be measured two different ways. On one hand, if we try to measure the displacement directly, we can determine it no more accurately than we can determine the center of the broad scattering function. On the other hand, if we try to maximize the height of the enhancement as we vary the boresight angle, we again have a slowly varying function that just follows the broad scattering function as illustrated in Fig. 8. Hence, the presence of the enhancement does not initially appear to improve our pointing capability. However, we see in Section 3 that this enhancement can indeed be manipulated to yield a very precise measure of the pointing error.

3. Interferometric Pointing Sensitivity through a Random Media

We now proceed to describe a technique that exploits the EBS phenomenon discussed above to achieve precision pointing accuracy when a dual aperture is used for active illumination of the target and polarization techniques are used for suppressing the unwanted (common-path) scattered radiation.

A. Polarization Suppression of Noise in Young's Interference Experiment

Solomon and Dainty have discussed the use of polarization for noise suppression in double-passage imaging through a random phase screen.[18] A direct optical image is not obtained with this process; however, diffraction-limited information concerning the object is obtained in the form of the squared modulus of the Fourier spectrum of the object.

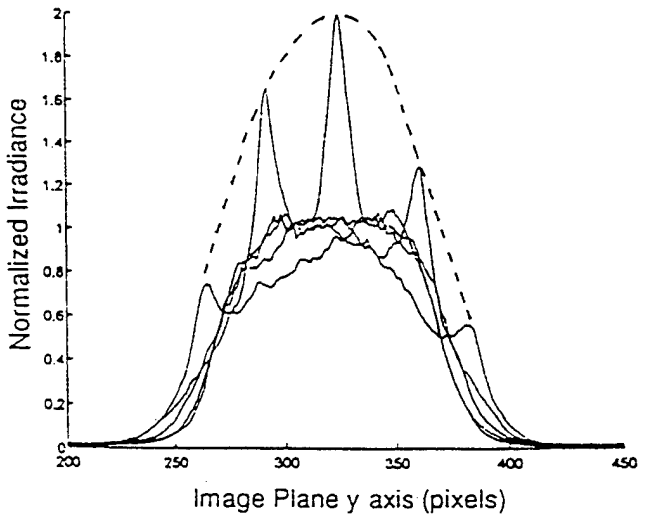


Fig. 8. Superposition of experimentally measured scattering functions, indicating the envelope of enhancement heights.

Hence, the actual image must be reconstructed from that information. Bogaturov *et al.*[23] have experimentally demonstrated the effectiveness of using orthogonal polarizations to separate the common-path radiation from the reciprocal path radiation in a simple Young interference experiment. This experiment yielded stable, high-contrast interference fringes even in the presence of time-varying turbulence.

B. EBS in the Form of Young's Interference Fringes

Consider the use of a dual aperture made up of two rectangular subapertures of width b, separated by a distance a in the pupil plane of the previous optical setup, with a half-wave plate following one subaperture as illustrated in Fig. 9. A horizontally polarized (parallel to the y axis) incident laser beam is transmitted through the polarizing beam-splitter cube and the dual aperture before passing through the random phase screen. If subaperture B is followed by a $\lambda/2$ plate whose axis is oriented at 45°, the polarization of that beam is rotated by 90° to the vertical orientation (perpendicular to the y axis). The target is thus illuminated by two separate beams of orthogonal polarizations.

On the return trip, each subaperture receives scattered light that originated from both subapertures. The $\lambda/2$ plate again rotates the polarization of the return radiation passing through subaperture B by 90°. The double-pass radiation thus consists of the following four components: $AA(\parallel)$, $BA(\perp)$, $AB(\perp)$, and $BB(\parallel)$; where the polarizations indicated in the parentheses are parallel or perpendicular to the y axis. If the width of the individual subapertures is less than the average turbulence cell size (Fried's r_o),[24] then $BA(\perp)$ and $AB(\perp)$ constitute the reciprocal path radiation producing the enhanced backscatter, whereas $AA(\parallel)$ and $BB(\parallel)$ constitute the common-path radiation that results in the broad scattering function. The polarizing beam splitter reflects the reciprocal path radiation toward lens 2 and the CCD camera and transmits the common-path radiation.

Assuming the spatial phase variations are negligible over the width of the subaperture ($b < r_o$), we can derive the expected irradiance distribution in the image plane in a straightforward manner by using conventional Fourier optical techniques to propagate from the transmitting pupil plane (x_0, y_0), to the object plane (x_1, y_1), to the receiving pupil plane (x_2, y_2), and finally to the image plane (x_3, y_3). As we stated above, the target is illuminated by two separate beams of orthogonal polarizations. We thus calculate the complex amplitude distribution emerging from the receiving pupil plane for the two components of the reciprocal path radiation separately, and then we add and propagate to the image plane.

Starting with a unit amplitude plane wave normally incident upon subaperture A (which lies in the

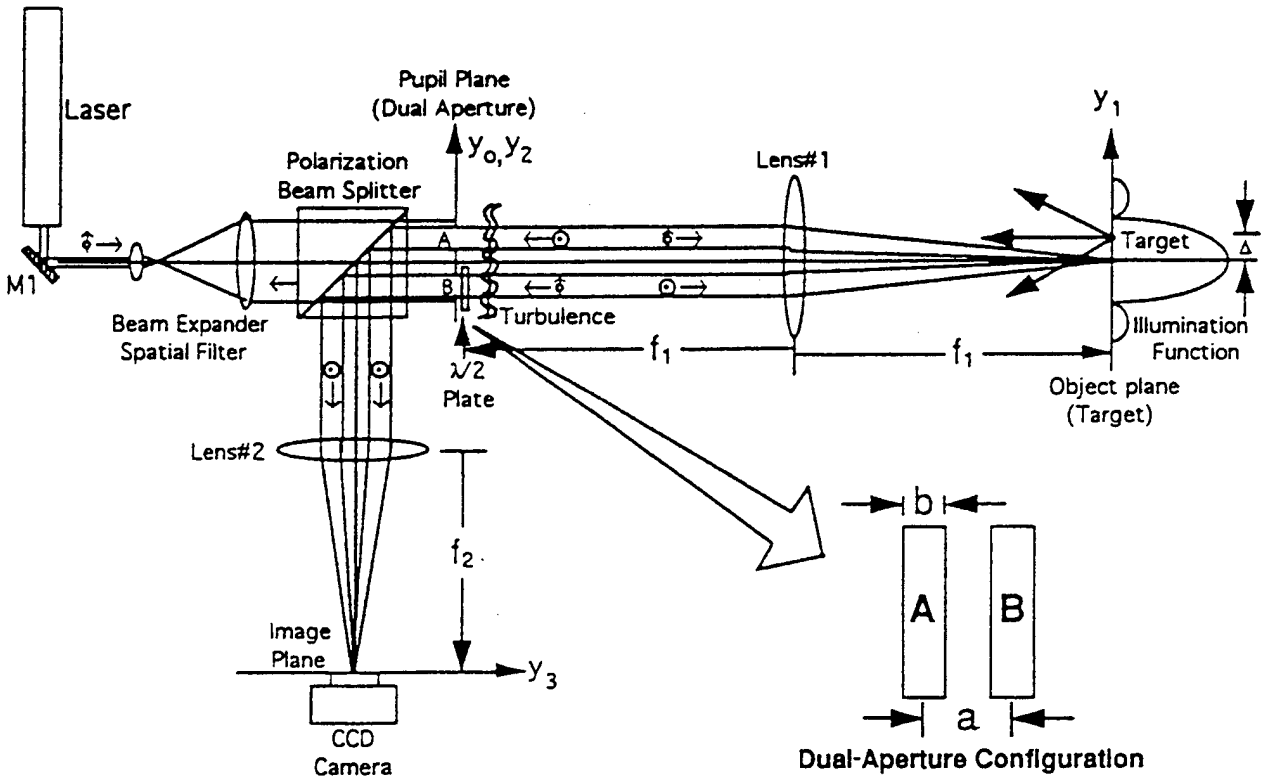


Fig. 9. Optical layout for precision tracking, utilizing EBS with a dual aperture.

front focal plane of lens 1) and using the symbolic notation of mathematical functions and operations popularized by Bracewell,[25] Goodman,[26] and Gaskill,[27] we can express the complex amplitude illuminating the target as[27]

$$U^-_{AB:\perp}(y_1) = \frac{\exp(i2kf_1)}{i\lambda f_1}\,\mathcal{F}\left\{\mathrm{rect}\left(\frac{y_0 - a/2}{b}\right)\right\}_{\eta = y_1/\lambda f_1}, \quad (1)$$

where $\mathcal{F}[\ ]$ is the Fourier transform operator and η is the reciprocal variable in Fourier transform space. Applying the shift theorem and the similarity theorem of Fourier transform theory,[25,26] we can write

$$U^-_{AB\perp}(y_1) = \frac{\exp(i2kf_1)}{i\lambda f_1}\, b\,\mathrm{sinc}\left(\frac{y_1}{\lambda f_1/b}\right)\exp\left(-i2\pi\,\frac{a}{2}\,\frac{y_1}{\lambda f_1}\right). \quad (2)$$

If an unresolved target is located a distance Δ from the optical axis, it can be represented by a shifted delta function, and the optical disturbance reflected from the target can be written as

$$\begin{aligned} U^-_{AB\perp}(y_1) &= U^-_{AB:\perp}(y_1)\delta(y_1 - \Delta) \\ &= \frac{\exp(i2kf_1)}{i\lambda f_1}\, b\,\mathrm{sinc}\left(\frac{y_1}{\lambda f_1/b}\right) \\ &\quad \times \exp\left(-i2\pi\,\frac{a}{2}\,\frac{y_1}{\lambda f_1}\right)\delta(y_1 - \Delta). \end{aligned} \quad (3)$$

Note that we use U^- to represent the complex amplitude distribution incident upon a plane (mask or target) and U^+ to represent the complex amplitude distribution emerging from that plane. The return radiation incident upon the receiving pupil plane mask is now given by the same Fourier transform relationship expressed earlier in Eq. (1):

$$U^-_{AB:\perp}(y_2) = \frac{\exp(i2kf_1)}{i\lambda f_1}\,\mathcal{F}\{U^+_{AB:\perp}(y_1)\}_{\eta = y_2/\lambda f_1}. \quad (4)$$

Substituting Eq. (3) into Eq. (4) and applying the convolution theorem and the scaling property of the convolution operation,[26] we obtain

$$\begin{aligned} U^-_{AB\perp}(y_2) &= -\exp(i4kf_1)\mathrm{rect}\left(\frac{y_2 + a/2}{b}\right) \\ &\quad * \exp\left(\frac{-i2\pi\Delta y_2}{\lambda f_1}\right), \end{aligned} \quad (5)$$

where the asterisk denotes the convolution operation.

The reciprocal path radiation (complex amplitude distribution) emerging from subaperture B is now given by

$$U^+_{AB:\perp}(y_2) = U^-_{AB\perp}(y_2)\mathrm{rect}\left(\frac{y_2 + a/2}{b}\right) \quad (6)$$

or

$$U_{AB,2}(y_2) = -\exp(i4kf_1) \times \left[\operatorname{rect}\left(\frac{y_2 + a/2}{b}\right) * \exp\left(\frac{-i2\pi\Delta y_2}{\lambda f_1}\right)\right] \times \operatorname{rect}\left(\frac{y_2 + a/2}{b}\right), \quad (7)$$

which can be written as

$$U_{AB,2}(y_2) = -\exp(i4kf_1) \times \left\{\left[\operatorname{rect}\left(\frac{y_2}{b}\right) * \exp\left(\frac{-i2\pi\Delta y_2}{\lambda f_1}\right)\right]\operatorname{rect}\left(\frac{y_2}{b}\right)\right\} * \delta\left(y_2 + \frac{a}{2}\right). \quad (8)$$

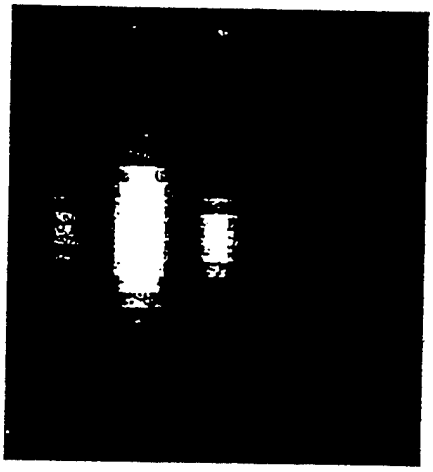
Similarly, if we start with subaperture B and perform the same operations, the reciprocal path radiation emerging from subaperture A is given by

$$U_{BA,2}(y_2) = -\exp(i4kf_1) \times \left\{\left[\operatorname{rect}\left(\frac{y_2}{b}\right) * \exp\left(\frac{-i2\pi\Delta y_2}{\lambda f_1}\right)\right]\operatorname{rect}\left(\frac{y_2}{b}\right)\right\} * \delta\left(y_2 - \frac{a}{2}\right). \quad (9)$$

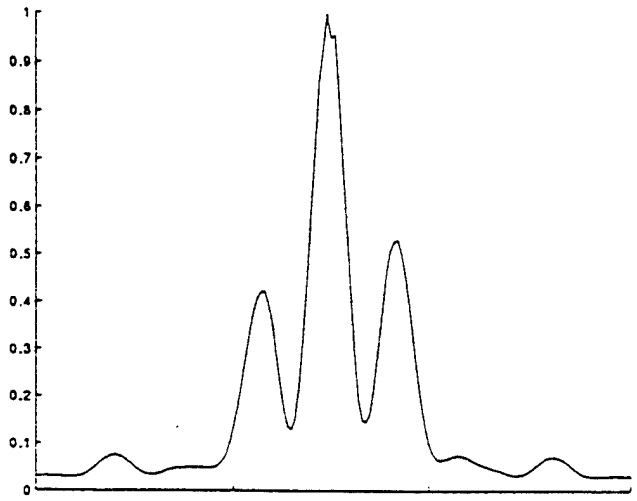
Because they are now of the same polarization we can add the complex amplitudes:

$$U^+(y_2) = U_{AB,2}(y_2) + U_{BA,2}(y_2). \quad (10)$$

Because of the distributive property of the convolu-

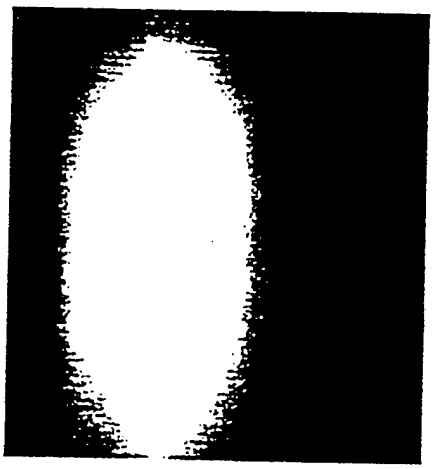


CCD Image

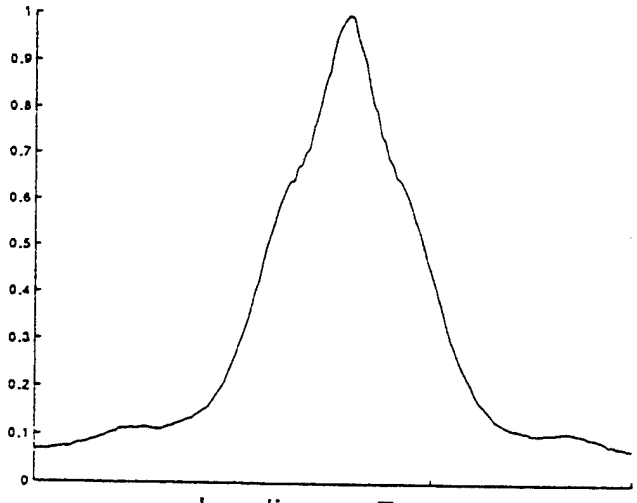


Irradiance Profile

(a)



CCD Image



Irradiance Profile

(b)

Fig. 10. CCD image and irradiance profile of (a) monostatic measurements, demonstrating static fringes over a 10-s exposure time; (b) bistatic measurements, demonstrating time-varying (blurred) fringes.

tion operation,[27]

$$U^{-}(y_2) = -\exp(i4kf_1) \times \left\{\left[\operatorname{rect}\left(\frac{y_2}{b}\right) * \exp\left(\frac{-i2\pi\Delta y_2}{\lambda f_1}\right)\right]\operatorname{rect}\left(\frac{y_2}{b}\right)\right\} * \left[\delta\left(y_2 - \frac{a}{2}\right) + \delta\left(y_2 + \frac{a}{2}\right)\right]. \quad (11)$$

Because the dual aperture also lies in the front focal plane of lens 2, the complex amplitude distribution incident upon the final image plane is now given by

$$U^{-}(y_3) = \frac{\exp(i2kf_2)}{i\lambda f_2}\,\mathcal{F}\{U^{-}(y_2)\}_{\eta = y_3/\lambda f_2}. \quad (12)$$

Substituting Eq. (11) into Eq. (12), we get

$$U^{-}(y_3) = \frac{-K}{i}\left[b\,\operatorname{sinc}\left(\frac{y_3}{\lambda f_2/b}\right)\delta\left(y_3 + \frac{\Delta f_2}{f_1}\right) * b\,\operatorname{sinc}\left(\frac{y_3}{\lambda f_2/b}\right)\right]2\cos\left(\pi\frac{y_3}{\lambda f_2/a}\right), \quad (13)$$

where $K = \exp[ik(2f_2 + 4f_1)]$. Applying the shift theorem and the sifting property of the delta function[25] to the convolution in the square brackets, we obtain

$$U^{-}(y_3) = \frac{-K2b^2}{i}\operatorname{sinc}\left(\frac{\Delta}{\lambda f_1/b}\right)\operatorname{sinc}\left(\frac{y_3 + \Delta f_2/f_1}{\lambda f_2/b}\right) \times \cos\left(\pi\frac{y_3}{\lambda f_2/a}\right), \quad (14)$$

and the irradiance distribution in the image plane is given by

$$I_3(y_3) = |U^{-}(y_3)|^2 = 4b^4\operatorname{sinc}^2\left(\frac{\Delta}{\lambda f_1/b}\right) \times \left[\operatorname{sinc}^2\left(\frac{y_3 + \Delta f_2/f_1}{\lambda f_2/b}\right)\cos^2\left(\pi\frac{y_3}{\lambda f_2/a}\right)\right]. \quad (15)$$

Hence, not only is the unwanted scattered light completely suppressed by the polarizing beam splitter, but the backscattered enhancement now takes the form of a target-position-dependent constant multiplied by Young's interference pattern (term in square brackets) consisting of $\cos^2$ interference fringes with a sinc^2 envelope function. Furthermore, it is apparent that the interference fringes are independent of the target position (the zero-order fringe remains fixed on the telescope boresight regardless of pointing error), and the envelope function is always centered on the image of the target, i.e., the envelope function shifts relative to the fringes by an amount proportional to the pointing error. The peak irradiance clearly decreases with increasing pointing error, and it goes to zero when the illuminating sinc^2 function misses the target.

If we choose the separation of the two rectangular subapertures to be twice their width ($a = 2b$), then there will be only three cosine fringes under the central lobe of the sinc^2 envelope function. This fringe pattern remains quite stable even in the presence of severe turbulence produced by the propane burner. Actually, only the nulls (dark fringes) are completely stable. The peaks fluctuate slightly in amplitude (and perhaps slightly in position); however, these effects are quite small and the relative stability is demonstrated by the CCD image and irradiance profile illustrated in Fig. 10(a). A bistatic measurement (separate transmitting and receiving apertures) in which the EBS phenomenon does not occur and in which the time-varying fringes are completely washed out is shown in Fig. 10(b) for comparison. Note that a 10-s exposure time was used to record these images.

Figure 11 illustrates a measured irradiance profile (1-s exposure time) superposed on a theoretical prediction calculated from Eq. (15) for a situation in which there is an appreciably pointing error. The agreement between theory and experiment is excellent. The asymmetry of the three-fringe pattern is a measure of the pointing error and can be quantified by defining the following asymmetry parameter:

$$R = \frac{I_3 - I_1}{I_3 + I_1}, \quad (16)$$

where I_1, I_2, and I_3 are the peak irradiances of the three fringes as illustrated in Fig. 11. This asymmetry parameter can be accurately measured, and it is quite sensitive to the pointing error.

C. Interferometric Pointing Sensitivity

Because Young's fringes remain stable even in the presence of temporally varying random phase errors, the above asymmetry parameter can be plotted versus the telescope pointing error to produce a

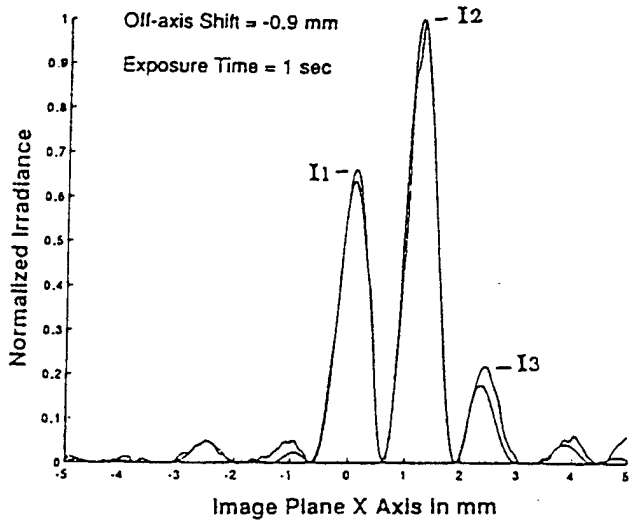


Fig. 11. Theoretical EBS function (caused by a dual aperture) superposed on an experimental measurement in the presence of turbulence (1-s exposure time). The asymmetry in the sidelobes caused by the pointing error is evident.

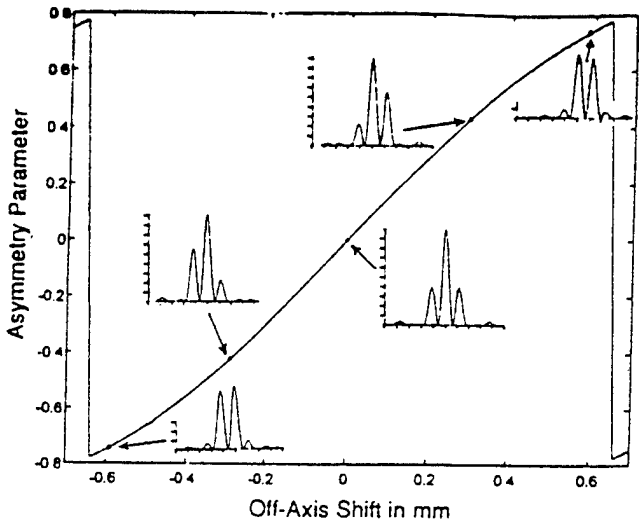


Fig. 12. Pointing error calibration curve, indicating interferometric accuracy.

calibration curve, which indicates that interferometric pointing sensitivity can be achieved with this RPS technique. Figure 12 illustrates this interferometric sensitivity, as well as the 2π ambiguity that accompanies many interferometric measurements.

4. Comparison of Experimental Data with Computer Simulations

An extensive series of CCD images were recorded as the target in the object plane was sequentially displaced by a precise amount with a translation stage fitted with a micrometer adjustment. A propane burner was used to produce the random phase screen, and each image was recorded with a 10-s exposure time. The asymmetry parameter was calculated and superposed on the pointing error calibration curve, as illustrated in Fig. 13. Note that the displacement of the target has been divided by the focal length of lens 1 to convert the ordinate of the calibration curve to angular units.

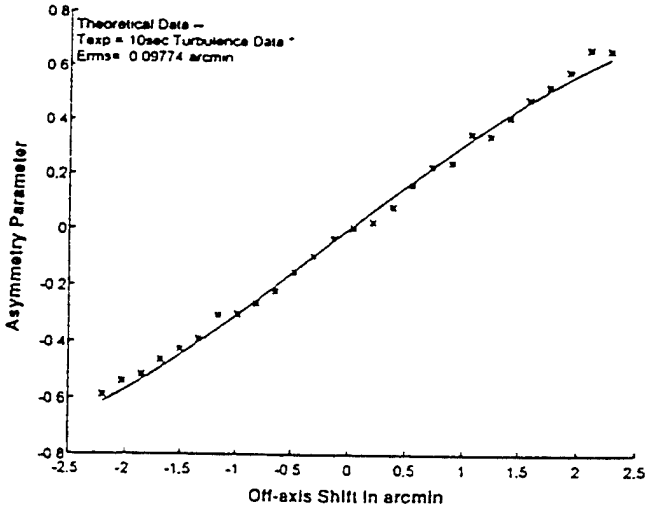


Fig. 13. Comparison of experimental data with theoretical predictions.

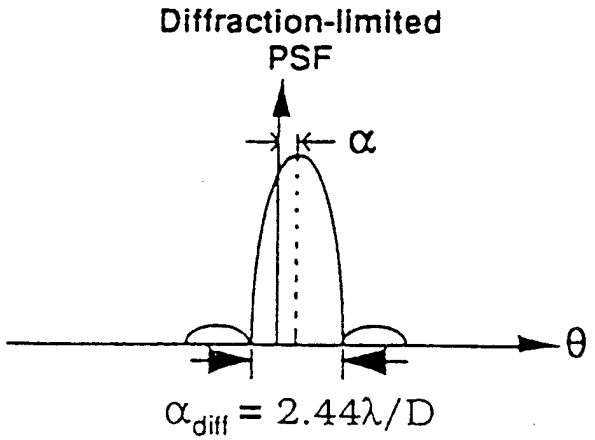


Fig. 14. Diffraction-limited angular PSF of the pointing telescope.

A small adjustment in the slope of the experimental curve was made to correct for a systematic error (average wave-front tilt across the aperture) in the experimental data, and the root-mean-square (rms) error in the measured asymmetry parameter for this set of data was calculated:

$$R_{rms} = \left[\frac{(R_{ex} - R_{th})^2}{N}\right]^{1.2}, \tag{17}$$

Dividing this by the slope of the calibration curve yields the accuracy to which the pointing error can be determined by this method. The rms error in the measurement of the telescope pointing error for these 27 data points is less than a tenth of an arcmin:

$$\text{rms error} = 0.09774 \text{ arcmin}. \tag{18}$$

This is best expressed as a fraction of the angular width of the diffraction-limited PSF of the telescope aperture, as illustrated in Fig. 14. In our laboratory experiment, $\lambda = 0.6328$ μm and $D = a + b = 0.65$ mm; hence

$$\frac{\alpha_{rms}}{\sigma_{diff}} = \frac{0.09774 \text{ arcmin}}{2.44\,\lambda/D} = 0.012. \tag{19}$$

We have thus demonstrated with this laboratory experiment that the EBS phenomenon can be exploited to measure telescope pointing errors through a random media or turbulent atmosphere to approximately 1% of the angular width of the diffraction-limited PSF of the pointing telescope.

5. Summary and Conclusions

After reviewing the enhanced backscatter phenomenon, we noted that the mere presence of a diffraction-limited intensity enhancement is not sufficient to allow the accurate measurement of telescope pointing errors through random media. However, introducing a monostatic dual aperture with subaperture sizes of the order of Fried's r_o causes the RPS light to produce an intensity enhancement in the form of a temporally stable Young interference pattern (even

in the presence of time-varying turbulence). By choosing the width-to-separation ratio of the dual aperture appropriately, we see that only three interference fringes exist within the central lobe of the envelope function. The interference fringes are fixed with respect to the telescope boresight and the envelope function moves with the pointing error; hence, the asymmetry of the interference pattern is a direct measure of the telescope pointing error. In addition, by using polarization techniques to suppress noise from the common-path scattered light, we can measure this asymmetry accurately enough to achieve interferometric sensitivity in pointing error measurements through random refractive media. A laboratory proof-of-concept experiment has demonstrated an rms measurement error of approximately 1% of the angular width of the diffraction-limited PSF of the pointing telescope.

This enhanced backscatter phenomenon should exist for any target that reflects (or scatters) light back through a monostatic aperture. However, many real situations of interest might involve uncooperative targets (highly absorbing or diffusely reflecting) from which the enhanced backscattered signal is greatly diminished, thus resulting in severe signal-to-noise problems. Field experiments with various targets over a 1-km range are currently being planned to determine the conditions under which this is a workable pointing and tracking technique.

This research was supported by the Ballistic Missile Defense Organization, Innovative Science and Technology Directorate through the U.S. Office of Naval Research under contract N00014-92-C-0123.

References

1. Six papers in a special section on optical tracking, Appl. Opt. 5, 481–532 (1966).
2. Twenty papers in a special section on acquisition, pointing, and tracking," Opt. Eng. 32, 2647–2811 (1993).
3. J. S. Preston, "Retro-reflexion by diffusing surfaces," Nature 213, 1007–1008 (1967).
4. T. S. Trowbridge, "Retroreflection from rough surfaces," J. Opt. Soc. Am. 68, 1225–1242 (1978).
5. Y. Kuga and A. Ishimaru, "Retroreflectance from a dense distribution of spherical particles," J. Opt. Soc. Am. A 1, 831–835 (1984).
6. L. W. Stockham and T. J. Love, "Investigation of the opposition effect in integrating spheres," J. Opt. Soc. Am. 60, 251–254 (1970).
7. W. W. Montgomery and R. H. Kohl, "Opposition effect experimentation," Opt. Lett. 5, 546–548 (1980).
8. Z. H. Gu, R. S. Dummer, A. A. Maradudin, and A. R. McGurn, "Experimental study of the opposition effect in the scattering of light from a randomly rough metal surface," Appl. Opt. 28, 537–543 (1989).
9. L. Tsang and A. Ishimaru, "Backscatter enhancement of random discrete scatterers," J. Opt. Soc. Am. A 1, 836–839 (1984).
10. E. Jakeman, "Enhanced backscattering through a deep random phase screen," J. Opt. Soc. Am. A 5, 1638–1648 (1988).
11. P. R. Tapster, A. R. Weeks, and E. Jakeman, "Observation of backscattering enhancement through atmospheric phase screen," J. Opt. Soc. Am. A 6, 517–522 (1989).
12. G. Welch and R. L. Phillips, "Simulation of enhanced backscatter by a phase screen," J. Opt. Soc. Am. A 7, 578–584 (1990).
13. M. Nieto-Vesperinas, "Enhanced backscattering," Opt. Photon. News 1(12), 50–52 (1990).
14. D. A. de Wolf, "Backscatter enhancement: random continuum and particles," J. Opt. Soc. Am. A 8, 465–471 (1991).
15. T. Mavroidis and J. C. Dainty, "Imaging after double passage through a random screen," Opt. Lett. 15, 857–859 (1990).
16. C. J. Solomon, J. C. Dainty, and R. G. Lane, "Double passage imaging through a random phase screen using a non-redundant aperture," J. Modern Opt. 10, 1993–2008 (1991).
17. W. T. Rhodes and G. Welch, "Determination of a coherent wave field after double passage through a diffuser," J. Opt. Soc. Am. A 9, 341–343 (1991).
18. C. J. Solomon and J. C. Dainty, "Use of polarisation in double passage imaging through a random screen," Opt. Commun. 87, 207–211 (1992).
19. J. E. Harvey and A. Kotha, "Sparse array configurations yielding uniform MTF's in reciprocal path imaging applications," Opt. Commun. 106, 178–182 (1994).
20. J. E. Harvey, A. Kotha, and R. L. Phillips, "Image characteristics in applications utilizing dilute subaperture arrays," Appl. Opt. 34, 2983–2991 (1995).
21. A. Dogariu, G. D. Boreman, and M. Dogariu, "Enhanced backscattering from a volume-scattering medium behind a phase screen," Appl. Opt. (to be published).
22. R. C. Heileman and R. L. Phillips, "Experimental measurements of statistical properties of scattered light due to double passage through a random phase screen," in *Atmospheric Propagation and Remote Sensing II*, A. Kohnle and W. B. Miller, eds., Proc. SPIE 1968, 662–672 (1993).
23. A. N. Bogaturov, A. S. Gurvich, V. A. Myakinin, J. C. Dainty, C. J. Solomon, and N. J. Wooder, "Use of polarization in interferometry after double passage through turbulence," Opt. Lett. 17, 757–759 (1992).
24. D. L. Fried, "Optical resolution through a randomly inhomogeneous medium for very long and very short exposures," J. Opt. Soc. Am. 56, 1372–1379 (1966).
25. R. N. Bracewell, *The Fourier Transform and its Applications* (McGraw-Hill, New York, 1965), Chap. 6, pp. 74, 101, 104.
26. J. W. Goodman, *Introduction to Fourier Optics* (McGraw-Hill, New York, 1968), Chap. 2, p. 9.
27. J. D. Gaskill, *Linear Systems, Fourier Transforms, and Optics* (Wiley, New York, 1978), Chap. 10, pp. 159, 166, 194, 196, 417.

7.0 IMAGE RECONSTRUCTION

Harley R. Myler, Ph.D. and Arthur R. Weeks, Ph.D.

7.1 Background

This chapter of the report discusses the development of a suite of reconstruction algorithms for the processing of laser speckle images. Laser speckle images provide a measurement of the Fourier transform magnitude of an illuminated object while the phase information associated with the magnitude is lost. In order to be able to reconstruct the original image of the object, the phase information must be reconstructed.

The reconstruction of a signal from the knowledge of its Fourier transform magnitude remains a very difficult problem and many algorithms have been developed in the past two decades to solve this problem. Currently, the most successful and practical approach in solving the phase retrieval problem are the iterative algorithms that seek numerical solutions that minimize the distance between the measurement and its estimate rather than finding an exact solution. These algorithms suffer from several drawbacks that set a limit to the size and complexity of the images that can be reconstructed. A major disadvantage of the iterative algorithms is the stagnation problem where the algorithm becomes trapped in a local minimum. The second problem is the slow convergence of the algorithm and the third is its computational cost.

In this program, two multiresolution adaptations of a basic iterative image reconstruction algorithm were developed that enable the algorithm to avoid stagnation, improve global convergence and dramatically reduce computational complexity [see Appendices 7A and 7B and reference Rabadi, 1995].

The first of these approaches is based on Burt's pyramid and the second is based on wavelet decomposition. These two methods have the following advantages: 1) they provide a rough and quick estimate of the solution at a low resolution that may later be refined as the algorithm progresses. This coarse-to-fine strategy enables the algorithm to avoid stagnation by providing a better initial guess and giving the algorithm a higher likelihood of arriving at a global minimum, 2) they can improve the convergence rate by decomposing the search space into orthogonal subspaces that can reduce the low frequency component of the error responsible for the slow convergence of the algorithm and, 3) since the number of independent variables to be processed at each coarser level are less than that at the full resolution grid, the reduction dramatically reduces the computational cost of the algorithm and ensures a faster convergence rate.

The computer simulations that we have processed indicate that the multiresolution approaches developed here are not only capable of avoiding stagnation and reducing the computational complexity of the iterative algorithm, but also produce reconstructions that are superior in quality to the single-grid approaches and are more robust in the presence of noise. The multiresolution methods show a great potential for future research to develop new algorithms in order to solve different image reconstruction and signal recovery problems. Some possible approaches to achieve better performance of these algorithms are discussed in the conclusion to this section.

7.2 Accomplishments

7.2.1 Multiresolution Error Reduction (MRER) Approach to Image Reconstruction

Iterative reconstruction techniques are currently the most effective approaches in solving a number of difficult signal reconstruction and recovery problems, and all of these algorithms suffer from stagnation, slow convergence rate and computational complexity that limits their use in future practical applications.

We developed a new multiresolution adaptation of the Error-Reduction algorithm that attempts to solve the above problems (i.e., stagnation, slow convergence and computational complexity) by employing the concept of pyramids [Burt, 1983, Rosenfeld, 1984]. This method is based on decomposing the problem of image reconstruction onto different resolutions. By using the solution obtained from the iterative algorithm at a lower resolution as an initial guess for the next finer level, then, following a coarse-to-fine strategy, this approach will enable the iterative algorithm to escape local minima by providing a better initial phase estimate while dramatically reducing the computational cost. The concept of pyramid decomposition will be introduced, followed by a detailed description of the algorithm. Finally, we will compare the performance of the developed algorithm with well-known iterative techniques and demonstrate the effectiveness of the algorithm in terms of convergence, robustness and computational efficiency as applied to both synthetic and real-world images.

Although the Error-Reduction algorithm [Gerchberg, 1972] and its modification by Fienup [Fienup, 1982] are currently the most effective approaches in solving the phase retrieval problem, we have shown that they still suffer from stagnation. This problem is due to the large number of attraction basins that cause the estimate to be trapped into a local minimum when the initial guess is not close enough to the global minimum. It has been shown [Nieto-Vesperinas, 1986] that the number of local minima in the phase retrieval problem increases dramatically with increasing the number of pixels in the image. Therefore, as the size of the image increases, it becomes difficult to find a good initial point that is close to the global minimum. An additional limitation of the iterative algorithm is the intensive computation due to the pair of 2D DFT's required for each cycle of the algorithm. This computational complexity is a function of the number of iterations required, as well as the number of pixels in the image to be reconstructed.

If several iterations of the Error-Reduction algorithm are performed on reduced sizes (lower resolutions) of the image, then the number of local minima, as well as the computations required, will be reduced. This indicates that it is more likely to find a good initial point that is close to the global minimum by coarser grid iteration rather than a finer grid where a smaller number of independent variables is processed. Then, by locally interpolating the result to a finer grid and using this result as the new initial estimate, rapid convergence can be achieved. Therefore, the multiresolution scheme can be expected to speed up the convergence and reduce the total computational cost of the algorithm as we will demonstrate.

Pyramids, in general, are data structures that provide successively condensed representation of the input information [Rosenfeld, 1984]. What is condensed may be simply image intensity, so that the successive levels of the pyramid are reduced-resolution versions of the input image. However, this condensed representation also represents a smoothed or subsampled version of any information structure presented at the input, so that each level of the pyramid represents a reduced *entropy* of the input. Pyramids support fast coarse-fine search strategies that are robust, compact and computationally efficient. This structure has been used to speed up different types of image and

low-vision operations, such as image segmentation, feature extraction and motion analysis [Rosenfeld, 1984]. Here, we will examine pyramids in the context of image reconstruction from three different perspectives, speed of convergence, robustness and computational efficiency.

The simplest type of image pyramid is constructed by repeatedly averaging the image intensities in nonoverlapping blocks of pixels. Given an input image $\boldsymbol{I_o}$ of size $N \times N$, where $N=2^p$, applying the averaging process yields a reduced image size $\boldsymbol{I_1}$ of size $2^{p-1} \times 2^{p-1}$. Applying the process again to the reduced image yields a still smaller image $\boldsymbol{I_2}$ of size $2^{p-2} \times 2^{p-2}$ and so on till we reach $\boldsymbol{I_{M-1}}$ (called the *apex* or top of the pyramid) of size 1×1 which represents the average value of the input image. If we imagine these images stacked on top of one another, they constitute an exponentially tapering $\boldsymbol{M}$-level pyramid of images. In order to develop a multiresolution algorithm, several components must be specified [Rosenfeld, 1984]:

1. The number of levels and the size of the grid at each level.
2. A *restriction* operation that maps from a fine level to a coarser grid.
3. A *prolongation* operation that maps a solution at a coarse level to a finer grid.
4. A *coordination scheme* that specifies the number of iterations at each level and the sequence of prolongations and restrictions.

In this work, we will use a simple version of the restriction and prolongation operations, where weighted averaging is used for restriction and linear interpolation is used for prolongation. The coordination scheme is a simple coarse-to-fine algorithm, where the prolongated coarse solution is used as a starting point for the next finer level.

Let I_o be the original image at the bottom of an M-level pyramid P_M ,and let I_{M-1} be the top level of P_M, then for $0 < l < M$, we can define the restriction operation R on P_M such that,

$$I_l = R\ I_{l-1} \tag{7.1}$$

where I_l is a reduced-resolution version of I_{l-1} . In our case, R can be defined as,

$$I_l(i,j) = \sum_{m=1}^{4}\sum_{n=1}^{4} w(m,n) I_{l-1}(2i+m-2, 2j+n-2) \qquad \begin{matrix} 1 \le l < M \\ 0 \le i,j < N/2^l \end{matrix} \tag{7.2}$$

where the weighting function $w(m,n)$ is the generating kernel (or the convolution mask). Since the number of sampled pixels has been reduced by half in each dimension, this kernel plays the role of a smoothing process (a low-pass filter that removes from the input spectrum all the components with frequency larger than half the sampling frequency) to ensure meeting the Nyquist rate condition and avoid non-appropriate subsampling. This weighting function $w(m,n)$ is chosen subject to four constraints [Rosenfeld, 1984]:

1. Separability: $w(m,n) = w_1(m) \cdot w_2(n)$.
2. Normalization: $\sum\sum w(m,n) = 1$.
3. Symmetry: $w(5-m,n) = w(m,5-n)$.
4. Equal contribution to the next level: $\sum_{i=0}^{1}\sum_{j=0}^{1} w(2i+m, 2j+n) = \frac{1}{4}$, $(m,n)=1,2$.

Among different masks, the following 4×4 kernel is widely used:

$$\begin{bmatrix} B^2 & AB & AB & B^2 \\ AB & A^2 & A^2 & AB \\ AB & A^2 & A^2 & AB \\ B^2 & AB & AB & B^2 \end{bmatrix} \tag{7.3}$$

The above four constraints are satisfied if:

1. $4\times(A+B)^2 = 1 \Rightarrow A+B = \pm ½$,
2. Symmetry is automatically satisfied,
3. $|A| \geq |B|$,
4. $(A+B)^2 = ¼ \Rightarrow A+B = \pm ½$,
5. $w = (B\ A\ A\ B)$.

Several pairs (A,B) satisfy this set of conditions, among them is the equal contribution without overlapping kernel for which $(A,B) = (1/2,0)$,

$$W = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0.25 & 0.25 & 0 \\ 0 & 0.25 & 0.25 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \tag{7.4}$$

In a similar manner, we define a prolongation operation P on P_M such that,

$$I_l = P\ I_{l+1} \tag{7.5}$$

which, in our case, can be defined as,

$$I_{l-1}(i,j) = \sum_{m=1}^{4}\sum_{n=1}^{4} w(m,n)\, I_l\left(\frac{i+m-2}{2}, \frac{j+n-2}{2}\right) \qquad \begin{matrix} 1 \leq l < M \\ 0 \leq i,j < N/2^l \end{matrix} \tag{7.6}$$

where only terms for which $(i+m-2)/2$ and $(j+n-2)/2$ are integers contribute to the sum, and $w(m,n)$ is the generating kernel given in Eq.(7.4) above. In this operation, the generating kernel will filter out the duplicated part of the spectrum that resulted from down sampling and project the rest back to a finer grid.

The first step in the image reconstruction procedure outlined here is to construct a bottom-up pyramid, where each level represents an approximation of the measured Fourier transform magnitude at that level. But since a precise measurement for the Fourier transform modulus at all the pyramid levels (i.e. $|F_l(u,v)|$, $0<l<M$) is not available, an estimate of these measurements is sought. An approximation of the Fourier transform magnitude at coarse grids can be obtained by first down sampling the measured autocorrelation function followed by Fourier transforming the subsampled autocorrelation and taking the square-root of its real part. Using the restriction operation **R** defined by Eq. (2), the above procedure can be expressed as,

$$\left|\tilde{F}_l(u,v)\right| = \sqrt{Re\left[\Im\left\{\mathcal{R}\ \ \Gamma_{l-1}(x,y)\right\}\right]} \qquad 1 \leq l < M \tag{7.7}$$

where $G_o(x,y) = G(x,y)$.

It is worthwhile to mention at this point that the approximation of the Fourier transform magnitude at different resolutions results in the construction of a *Fourier pyramid* where each level represents an approximation of the Fourier transform magnitude of a low-pass version of the image to be reconstructed. In order to show that this is indeed the case, the autocorrelation function of the ring shape, $\Gamma_o(x,y)$, was subsampled and compared with the autocorrelation of the subsampled (low-pass) version of the ring-shape image at different levels of the pyramid. Fig. 7.3 illustrates this comparison and it is clear that this makes a good approximation of the autocorrelation function of a low-pass version of the image as measured by the normalized root mean-square (NRMS) error defined by

$$E_{\Gamma} = \frac{\sum_m \sum_n \left| \Gamma_{\tilde{f}}(m,n) - \tilde{\Gamma}_f(m,n) \right|^2}{\sum_m \sum_n \left| \Gamma_{\tilde{f}}(m,n) \right|^2} \tag{7.8}$$

Here, $\Gamma_{\tilde{f}}(m,n)$ is the autocorrelation function of a subsampled version of the original image and $\tilde{\Gamma}_f(m,n)$ is the subsampled autocorrelation of the original image. The values of the NRMS error due to this approximation is shown in Table 7.1.

TABLE 7.1 NRMS Errors on Reconstruction

	NRMS Error	
	Direct Approach	Pyramid Approach
Reconstruction using the actual autocorrelation	0.40559	0.19052
Reconstruction using the autocorrelation estimated from averaging 1000 speckle realizations	0.45025	0.32803
Reconstruction using the autocorrelation estimated from averaging 100 speckle realizations	0.48256	0.40611

After forming the Fourier pyramid, the reconstruction procedure proceeds from the pyramid top to the bottom by performing several cycles of the basic Error-Reduction algorithm on the lowest-resolution level ($l=M-1$) until the number of iterations reaches an arbitrary limit that depends on the complexity of the image. The reconstructed image $g_{M-1}(i,j)$ is then transferred to the next level of the pyramid ($l=M-2$) with double the resolution by applying the interpolation process **P** defined by Eq.(7.6), that is,

$$g_{k-1} = \mathcal{P}\ g_k \qquad 1 \le l < M \tag{7.9}$$

The same basic iterative procedure is performed on the new pyramid level using the last image estimate computed from the previous level as the new initial guess for the algorithm. This process is repeated following a coarse-to-fine strategy, until level zero of the pyramid is reached, where the image is now reconstructed at the original resolution. The above procedure is described by the following steps:

Step 1:

Construct the Fourier pyramid as given by Eq.(7.7).

Step 2:

Start at the highest level of the pyramid desired ($l=M-1$), using an initial random guess of size $2^{-l}N \times 2^{-l}N$ as a starting point.

Step 3:

Perform k number of iterations of the error reduction algorithm, where K is the total number of iterations to be performed, using the approximated Fourier transform magnitude at this level(i.e. $|F_l(u,v)|$) as the Fourier domain constraint, while reducing the object's support to constraint to $2^{-l}D_f$

Step 4:

Interpolate the reconstructed image $g_l(x,y)$ to obtain $g_{l-1}(x,y)$ using Eq.(7.6), and use it as the new initial guess.

Step 5:

Go to the next finer level : $l = l-1$.

Step 6:

If $l < 0$ stop, otherwise go to **Step 3**.

The number of levels of the pyramid M for an image of size $N \times N$ is typically chosen to be equal to $\log_2 N$. In addition, the coordination scheme is chosen such that an equal number of iterations, $k=K/M$, is performed at each level of the pyramid, where K is the total number of iterations performed throughout the reconstruction procedure.

The most obvious advantage of the pyramid representation is that it provides a means of reducing the computational cost of the iterative algorithm by using a divide-and-conquer principle. To investigate the reduction in overall computational cost that results from using the multiresolution approach, we calculate the number of real additions and multiplications (due only to the pair of 2D DFT's since the imposition of constraints involves far less computations) required to perform K iterations of the error reduction algorithm on an $N \times N$ image. If the 1D FFT is used to compute the 2D DFT in a row-column decomposition, then the evaluation of each 2D DFT requires $8N^2 \log_2 N$ real multiplication's and $12N^2 \log_2 N$ real additions [Dudgeon, 1984], a total of:

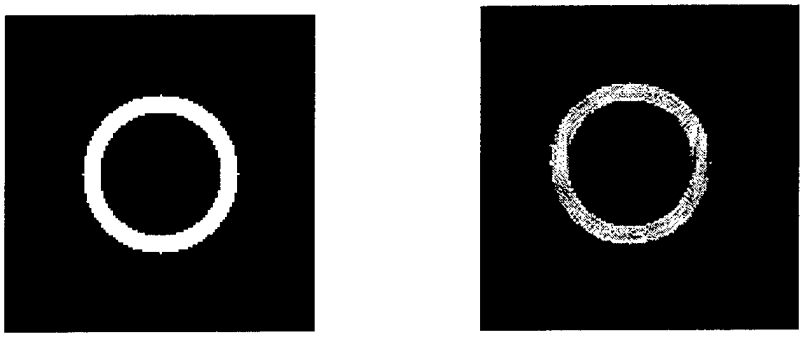
$$20KN^2 \log_2 N \tag{7.10}$$

algebraic operations for K iterations of the algorithm. When performing the same number of iterations on an M-level pyramid (where $k=K/M$ number of iterations is performed at each level), the total number of algebraic operations reduces to :

$$\frac{20 \cdot K}{M} \sum_{l=0}^{M-1} \left(\frac{N}{2^l} \right)^2 \log_2 \left(\frac{N}{2^l} \right) \tag{7.11}$$

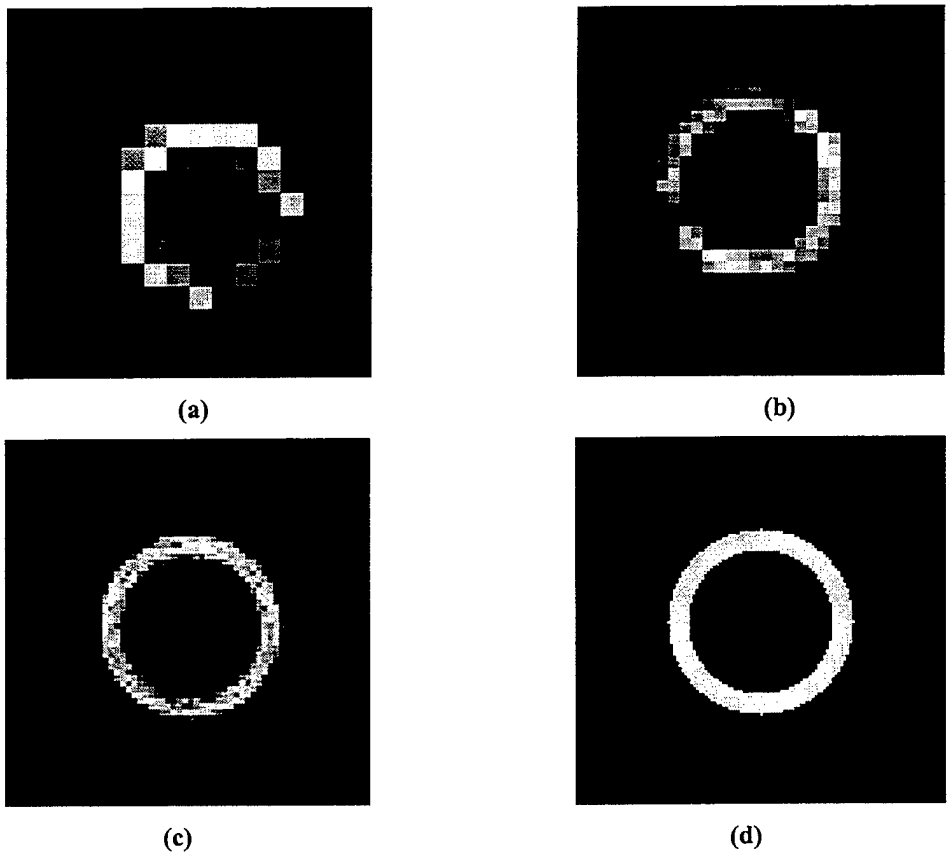
This indicates that a considerable reduction on the computation is achieved by this decomposition of the problem.

In the experiments that we performed, the pyramid approach produced reconstructions that are superior in quality (as measured by the residual error) to the direct iterative method. To illustrate the procedure of the multiresolution pyramid approach, a synthetic image of a ring-shape shown in Figure 7.1(a) was used. Figure 7.1(b) shows the reconstructed image after 100 iterations of the error-reduction algorithm on a single grid (i.e. direct approach), where the actual (noiseless) autocorrelation function was used to obtain the Fourier transform magnitude of the ring-shape image.



(a) (b)

Figure 7.1: Example of image reconstruction from the actual autocorrelation function (the direct approach): (a) Original image (128×128) of the ring-shape object, and (b) Image reconstructed after 100 iterations of the Error-Reduction algorithm on a single grid.



(a) (b)

(c) (d)

Figure 7.2: Example of image reconstruction from the actual autocorrelation function (the multiresolution approach): (a) Image reconstructed after 25 iterations on level-3 (16×16), (b) Image reconstructed after 25 iterations on level-2 (32×32, (c) Image reconstructed after 25 iterations on level-1 (64×64), (d) Image reconstructed after 25 iterations on level-0 (128×128).

Figure 7.2 shows a 4-level Fourier pyramid constructed as described by Eq.(7.7), where each level of the pyramid represents an approximation of the Fourier transform magnitude at different resolution (128×128, 64 ×64, 32×32, ... etc.). Again, the actual (noiseless) autocorrelation function of the ring-shape image was used to obtain the Fourier transform magnitude. each level of the 4-level pyramid by performing 25 iterations of the error reduction algorithm at each level (a total of 100 iterations), where Figure 7.3(a) shows the 16×16 pixel images reconstructed

at level 3 of the pyramid after 25 iterations of the error reduction algorithm. By interpolating this image as given by Eq.(7.6), a new image with 32×32 pixel resolution can be obtained. This new image can now be used as the initial guess for the error reduction algorithm. Figure 7.3(b) shows the reconstructed image at level 2 of the pyramid after 25 iterations of the error reduction algorithm. The same procedure is repeated again to obtain the reconstructed image shown in Figure 7.3(c). Figure 7.3(d) shows the image reconstructed at the original resolution, i.e., level zero of the pyramid.

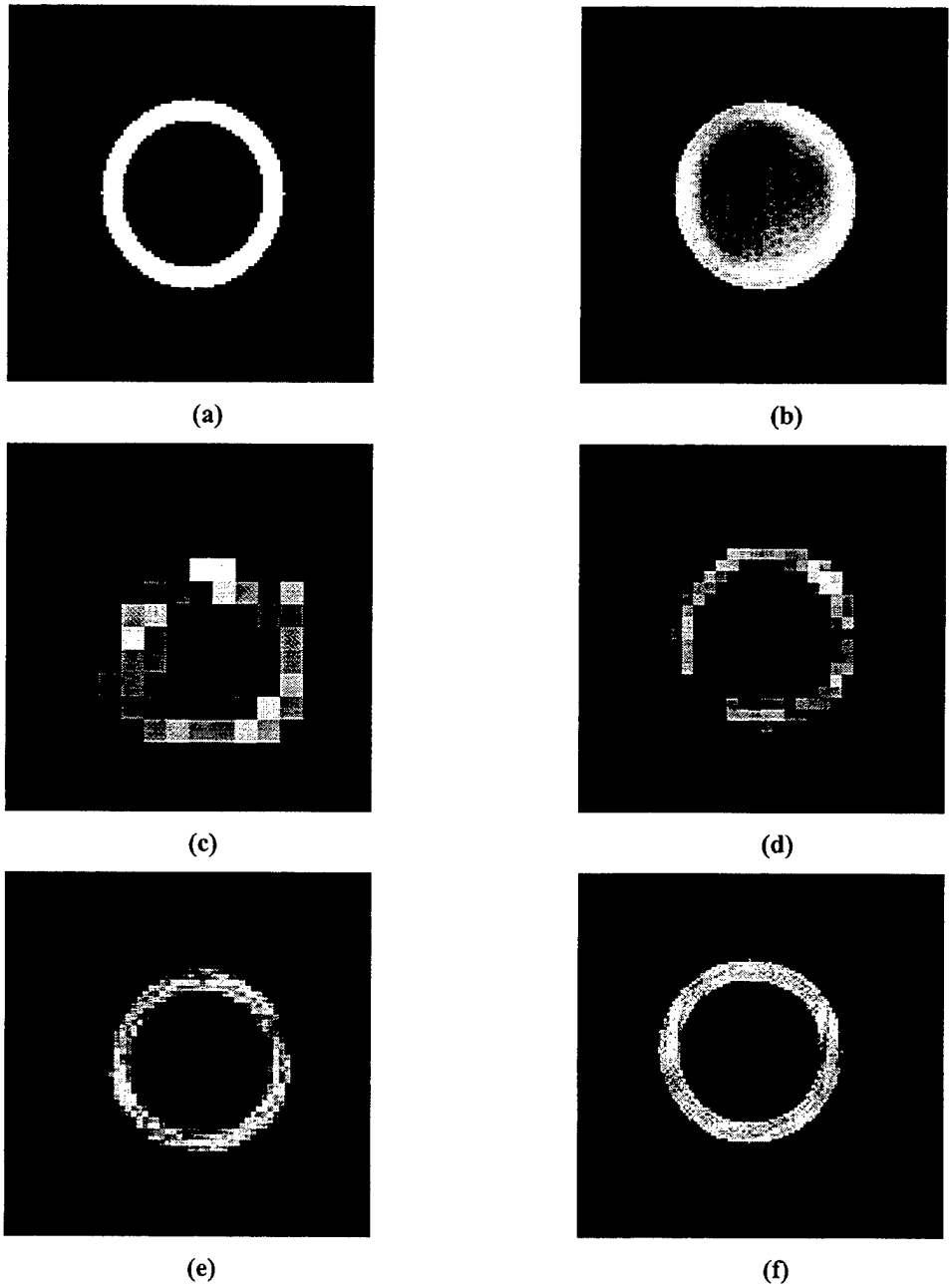


Figure 7.3: Example of image reconstruction from the noisy estimate of the auto-correlation function obtained from 1000 speckle pattens (the direct approach vs. the multiresolution approach): (a) Original image of the ring-shape object. (b) Image reconstructed after 100 iterations of the direct Error-Reduction, (c) Image reconstructed after 25 iterations on level-3 (16×16), (d) Image reconstructed after 25 iterations on level-2 (32×32), (e) Image reconstructed after 25 iterations on level-1 (64×64), (f) Image reconstructed after 25 iterations on level-0 (128×128).

To investigate the performance of the proposed approach in the presence of noise, we repeated the above procedure using a noisy estimate of the autocorrelation. Figure 7.4(a) shows the ring-shape test image. Figure 7.4(b) shows the reconstructed image after 100 iterations of the error reduction algorithm on a single grid, where an estimate of the autocorrelation function obtained from averaging 1000 speckle realizations was used to obtain the magnitude of the Fourier transform of the image. Figure 7.4(c), 7.4(d) 7.4(e) and 7.4(f) show the reconstruction at each level of the 4-level pyramid by performing 25 iterations of the error reduction algorithm at each level (a total of 100 iterations).

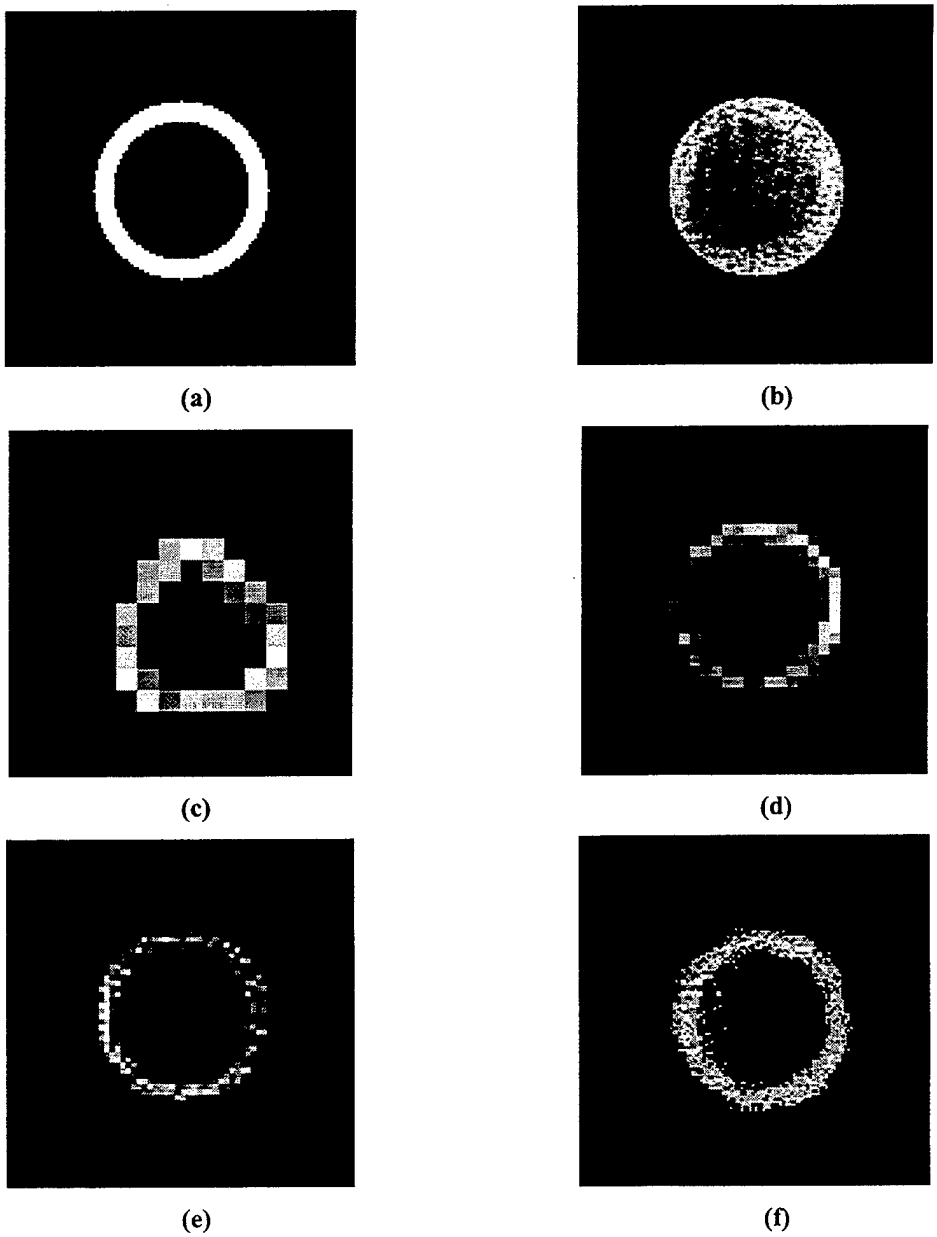


Figure 7.4: Example of image reconstruction from the noisy estimate of the autocorrelation function obtained from 100 speckle patterns (the direct approach vs. the multiresolution approach): (a) Original image of the ring-shape object., (b) Image reconstructed after 100 iterations of the Error-Reduction algorithm on a single grid (the direct approach), (c) Image reconstructed after 25 iterations on level-3 (16×16),(d) Image reconstructed after 25 iterations on level-2 (32×32), (e) Image reconstructed after 25 iterations on level-1 (64×64), (f) Image reconstructed after 25 iterations on level-zero (128×128).

We repeated the same procedure using the same ring-shape image, where a noisier estimate of the autocorrelation function was obtained from averaging 100 speckle realizations, was used to obtain the Fourier transform modulus. Figure 7.5(a) shows the original image, and Figure 7.5(b) shows the reconstructed image after 100 iterations of the error reduction algorithm performed at a single grid. The results of the multiresolution approach at different levels of the pyramid using the same autocorrelation function estimate (from 100 speckle realizations) are shown in Figs. 7.5(c), (d), (e) and (f).

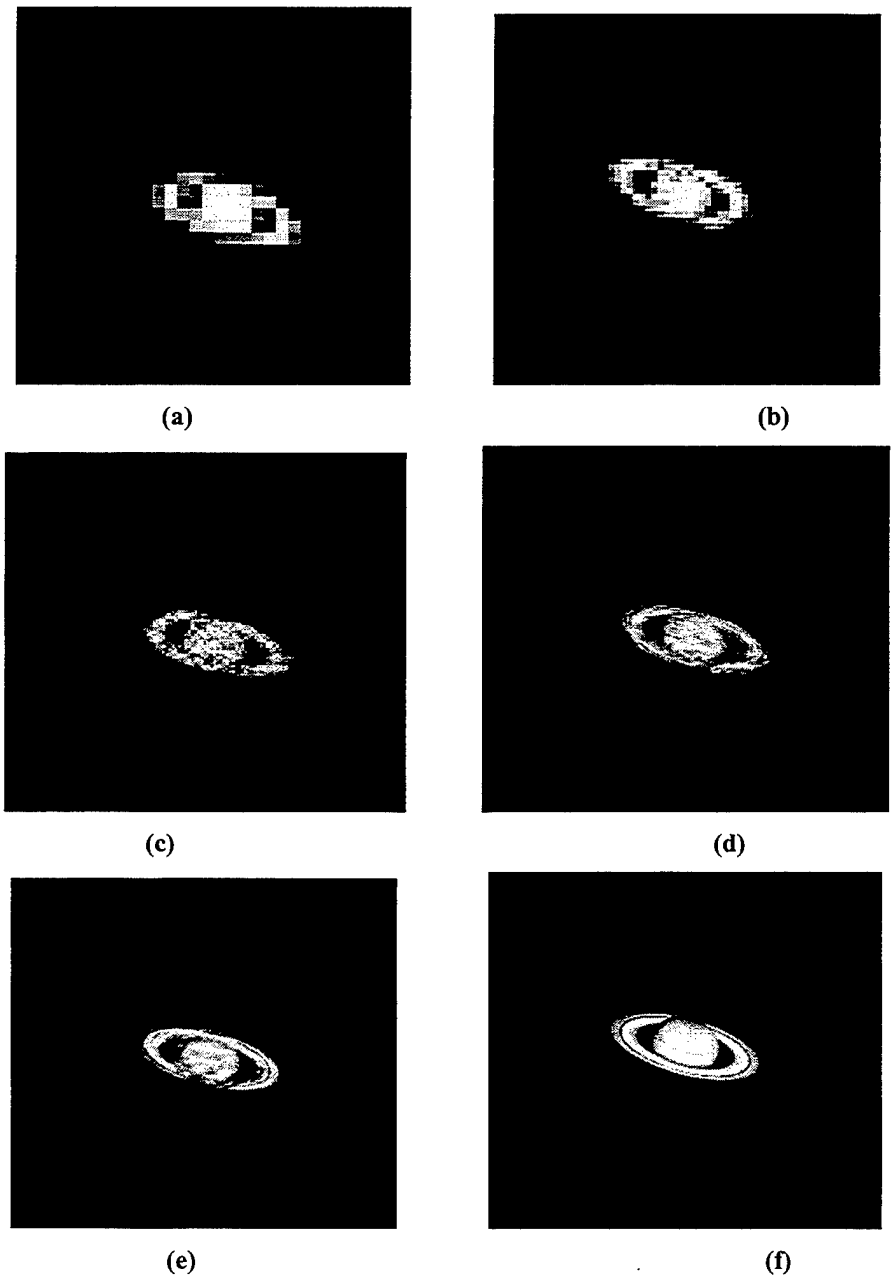


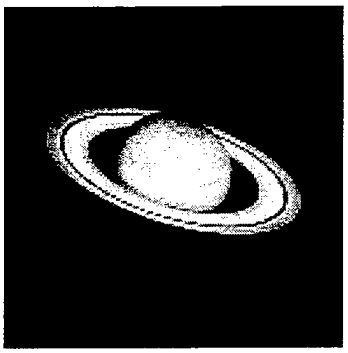
Figure 7.5: Example of image reconstruction on 5-level pyramid (the multiresolution approach): (a) Image reconstructed after 200 iterations on level-4 (16×16); (b) Image reconstructed after 200 iterations on level-3(32×32); (c) Image reconstructed after 200 iterations on level-2 (64×64); (d) Image reconstructed after 200 iterations on level-1 (128×128); (e) Image reconstructed after 200 iterations on level-0 (256×256); and (f) Original Image.

We compare the quality of the reconstructions obtained by using the multiresolution approach with that using the direct approach from both the noiseless and noisy autocorrelation function. Table 7.1 shows the numerical values of the NRMS error of these reconstructions, which in this case defined as

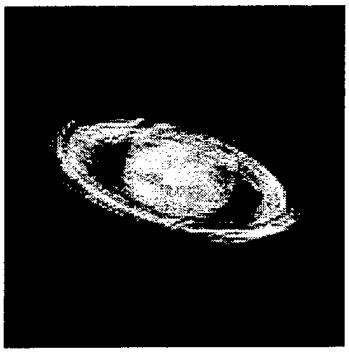
$$E_f = \frac{\sum_u \sum_v |F(u,v) - G(u,v)|^2}{\sum_u \sum_v |F(u,v)|^2} \tag{7.12}$$

It is clear from these results that the proposed approach produced reconstructions that are superior to the direct error-reduction approach from both noiseless and noisy measurements of the autocorrelation function. This indicates that the multiresolution approach is more robust in the presence of noise.

A more complex image (Saturn) of size 256×256 shown in Figure 7.6(a) is used to evaluate the performance of the multiresolution approach. The image reconstructed after 800 iterations of the Error-Reduction algorithm on a single grid is shown in Figure 7.6(b).



(a) *Original image of Saturn.*



(b) *Image reconstructed after 800 iterations of the original ER algorithm*

Figure 7.6: Original image Vs. reconstructed image using the Error-Reduction algorithm.

7.2.2 Wavelet Decomposition Approach to Image Reconstruction

Wavelet decomposition is an alternative approach to multiresolution pyramids that is able to decompose the $L^2(\mathbf{R}^2)$ space into orthogonal subspaces and provide a clear model of signal decomposition is the wavelet representation. Wavelets and the multiresolution decomposition of Mallat [Daubechies, 1988, Mallat, 1989] can provide a very efficient framework in solving many problems such as image compression [Mallat, 1989], image restoration [Donoho, 1992], partial differential equations [Amaratunga, 1993] and fast numerical techniques [Beylkin, 1991].

As previously discussed, the existing iterative algorithms are currently the most effective approach in solving the image reconstruction problem. However, for reasonably large images, solving this problem at the original resolution is a very difficult task due to stagnation and computational complexity. As in many engineering problems, we often require a quick rough estimate of the solution at preliminary stage, which may later be refined as the design or investigation progresses. Wavelets have the capability of providing a multilevel description of the solution. The multiresolution property of wavelets, along with their orthogonality and localization properties, means

that we may obtain an initial coarse description of the solution with little computational effort and then successively refine the solution with a minimum of extra effort. This philosophy can be used in our case to decompose the image reconstruction problem into different resolutions, where solving the problem at a coarse resolution will provide us with a rough approximation of solution that can be used as an initial guess for the next higher resolution.

We have adapted a wavelet decomposition algorithm that is more efficient, has a faster convergence rate, is more robust in the presence of noise and requires far less computations than the existing algorithms currently used to solve the phase retrieval and image reconstruction problem.

In the phase retrieval problem we need to reconstruct an image (spatial domain information) given a measurement of the Fourier transform magnitude of that image (partial spectral information). Hence, we need to obtain a clear representation of the wavelet decomposition in the frequency domain which, in turn, will give us a vital interpretation of how the frequency spectrum of a signal is divided between different subspaces. Thus, it is important to study the wavelet decomposition in the frequency domain.

Let $\mathbf{A}_m$ be the operator that approximates a signal $f(x)$ at resolution m. This operator is:

1. A linear operator, defined as

$$A_m f(x) = 2^{-m/2} \sum_{n=-\infty}^{\infty} \left\langle f(u), \varphi(2^{-m}u - n) \right\rangle \varphi(2^{-m}x - n) \tag{7.13}$$

The approximation $\mathbf{A}_m f(x)$ of the signal $f(x)$ at the resolution m, is thus characterized by the set of inner products which is denoted by

$$A_m^d f = \left(\left\langle f(u), \varphi_{m,n}(u) \right\rangle \right)_{n \in Z} \tag{7.14}$$

$\mathbf{A}_m^d f$ is called a *discrete approximation* of $f(x)$ at the resolution m.

2. A projection operator on a particular vector space V_m. That is:

$$f(x) \approx f^m(x) = A_m f(x) \tag{7.15}$$

where $f^m(x) \in V_m$.

3. Among all the approximated functions of $f(x)$, $f^m(x) = \mathbf{A}_m f(x)$ is the function which is the best similar to $f(x)$, i.e.

$$\forall g(x) \in V_m, \qquad \left\| g(x) - f(x) \right\| \geq \left\| A_m f(x) - f(x) \right\| \tag{7.16}$$

Hence, the operator $\mathbf{A}_m$ is an orthogonal projection on the vector space V_m.

4. The approximation $\mathbf{A}_m f(x)$ is not modified if we approximate it a gain at the resolution m. That is,

$$A_m f^m(x) = f^m(x) \tag{7.17}$$

We will show now that this operator can be interpreted as a convolution with a low-pass filter followed by decimation.

Let $f(x)$ be a signal with finite support and finite resolution (most signals satisfy these conditions), and let $f(x) \in V_m$, for some integer m. Without loss of generality, we can set $m=0$ and $f(x) \in V_o$. Hence

$$f(x) = f^0(x) = A_o f(x) \tag{7.18}$$

We can compute the discrete approximation of *f(x)* at any resolution *m* by

$$f^m = A_m^d f^0 = \left\langle f(u), \varphi_{m,n}(u) \right\rangle \tag{7.19}$$

From the dilation equation, $\varphi(x) = \sqrt{2} \sum_{n=-\infty}^{\infty} h_n \varphi(2x-2)$ we know that

$$\varphi_{m+1,n}(x) = \sum_{k=-\infty}^{\infty} \left\langle \varphi_{m+1,k}(u), \varphi_{m,k}(u) \right\rangle \varphi_{m,k}(u) \tag{7.20}$$

By taking the inner product of f and $j_{m+1,n}(x)$, we obtain

$$\left\langle f, \varphi_{m+1,n} \right\rangle = \left\langle f, \sum_{k=-\infty}^{\infty} \left\langle \varphi_{m+1,k}, \varphi_{m,k} \right\rangle \varphi_{m,k}(u) \right\rangle \tag{7.21}$$

rearranging the terms

$$\left\langle f, \varphi_{m+1,n} \right\rangle = \sum_{k=-\infty}^{\infty} \left\langle \varphi_{m+1,k}, \varphi_{m,k} \right\rangle \left\langle f, \varphi_{m,k}(u) \right\rangle \tag{7.22}$$

We can compute the next step of the approximation using

$$f^{m+1}(n) = \left\langle f, \varphi_{m+1,n} \right\rangle = \sum_{k=-\infty}^{\infty} \overline{h}(2n-k) f^m(n) \tag{7.23}$$

where $\overline{h}$ is a mirror filter (i.e. $\overline{h}(n) = h(-n)$) of an FIR filter *h(n*. (7.23) above shows that the approximation of a signal at resolution *m*+1 can be obtained by convolving the approximation f^m at resolution *m* with a low-pass FIR filter and keeping every other sample of the output. All the discrete approximations of f^m for *m* > 0 of *f(x)* can thus be computed from f^0 by repeating this process of convolution followed by decimation. In the same manner, one can show that the detailed signal d^m at each resolution *m* can be captured by applying the operator $\mathbf{D}_m$ to the signal

$$\delta_m(n) = D_m^d f(x) = \sum_k \overline{g}(2n-k) \delta_m(n) \tag{7.24}$$

where $\overline{g}(n)$ is the mirror filter (i.e. $\overline{g}(n) = g(-n)$) of *g(n)* defined the Daubechies dilation equation (Daubechies, 1992).

To illustrate this procedure, Figure 7.7 shows discrete approximations of the ring shaped object at four resolutions. These results have been computed with the algorithm previously described. Notice that the number of points at each approximation *m* is equal to $2^{-m}N$ points.

To illustrate what happens in the frequency domain when we decompose the signal into different resolutions, lets take the Fourier transform of the approximated signal f^m at a certain resolution *m* given by Eq.(7.23)

$$\sum_n f^{m+1}(n) e^{-2\pi i n \xi} = \sum_n \sum_k h(2n-k) f^m(n) e^{-2\pi i n \xi} \tag{7.25}$$

$$F^{m+1}(\xi) = \sum_n f^m(n) \sum_k h(2n-k) e^{-2\pi i n \xi} \tag{7.26}$$

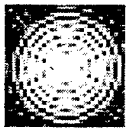
$$F^{m+1}(\xi) = \sum_n f^m(n) e^{-2\pi i n \frac{\xi}{2}} \sum_k h(k) e^{-2\pi i k \frac{\xi}{2}} \tag{7.27}$$

$$F^{m+1}(\xi) = F^m(\tfrac{\xi}{2}) H(\tfrac{\xi}{2}) \tag{7.28}$$

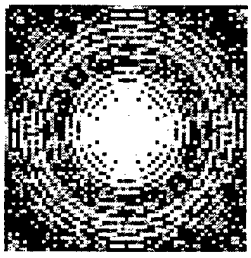
The final result given by Eq.(7.28) indicates that the approximation at resolution m+1 is a low-pass version of the approximation at resolution m.



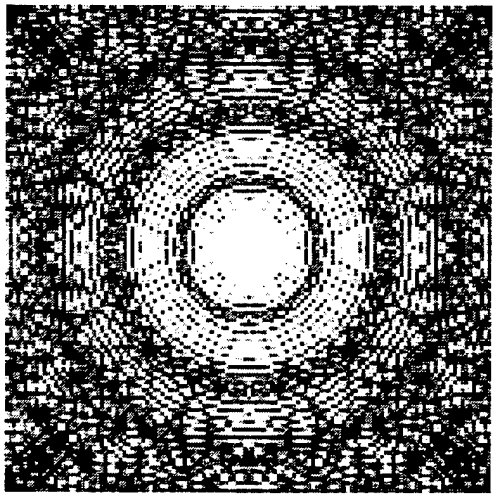
(a)



(b)



(c)



(d)

Figure 7.7: The Fourier Pyramid of the ring-shape object: (a) Level-3 of 16×16 pixels resolution; (b) Level-2 of 32×32 pixels resolution; (c) Level-1 of 64×64 pixels resolution; and (d) Level-0 of 128×128 pixels resolution.

Wavelets and the multiresolution decomposition can provide a very efficient framework in solving the image reconstruction problem stated before. We have seen previously that the iterative algorithm is currently the most effective approach in solving the image reconstruction problem. However, for reasonably large images, solving this problem at the original resolution is a very difficult task due to stagnation and computational complexity. As in many engineering problems, we often require a quick rough estimate of the solution at preliminary stage, which may later be refined as the design or investigation progresses. As shown, wavelets have the capability of providing a multilevel description of the solution. The multiresolution property of wavelets, along with their orthogonality and localization properties, means that we may obtain an initial coarse description of the solution with little computational effort and then successively refine the solution with a minimum of extra effort. This philosophy can be used in our case to decompose the image reconstruction problem into different resolutions, where solving the problem at a coarse resolution will provide us with a rough approximation of solution that can be used as an initial guess for the next higher resolution. The Error-Reduction algorithm described before will be used to furnish a solution at each resolution.

The original Error-Reduction (ER) algorithm introduced by Gerchberg and Saxton [Gerchberg, 1972, Saxton, 1978] is described in the literature. Here we will develop a multiresolution adaptation of this algorithm and show that this algorithm provides reconstructions that are superior in quality to the original algorithm and more efficient in terms of the number of computations required.

Let $|F(u,v)|$ be the magnitude of the Fourier transform of the unknown image $f(x,y) \in L^2(\mathbf{R}^2)$. Furthermore, let $f(x,y) \in V_m$ for some $m \in Z$. Without loss of generality we can assume that $f(x,y) \in V_o$. The goal is to obtain a coarse approximation $f^m \in V_m$ $(m > 1)$ of $f(x,y)$. Once the approximation f^m is obtained, it can be used as an initial estimate for the Error-Reduction algorithm to obtain a new finer approximation $f^{m-1} \in V_{m-1}$ of $f(x,y)$. This procedure is repeated until m=0, where the image is reconstructed at the original resolution.

In order to reconstruct f^m we need to use the Fourier transform magnitude at the m^{th} resolution. However, since a precise measurement for the Fourier transform modulus at all resolutions, i.e., $|F^m(u,v)|$, $0<m<M$) is not available, an estimate of these measurements is sought. An estimate can be obtained by first down sampling the measured autocorrelation function of the image followed by Fourier transforming the subsampled autocorrelation and taking the square-root of its real part. This can provide a good approximation of the Fourier modulus at that resolution; however, the coarser the approximation is (i.e. the larger m is), the less accurate the estimate $|F^m(u,v)|$ will be. This problem can be easily solved using wavelets. Since as we have shown that any coarse approximation f^m of $f(x,y)$ can be obtained by applying the discrete approximation operator to $f(x,y)$ which is equivalent to a convolution with an FIR filter followed by decimation (see Eq.(5.33)). In the frequency domain, we have shown that this operation results in decomposing the frequency domain into different subbands, that is

$$F^{m+1}(\xi,\nu) = H(\tfrac{\xi}{2},\tfrac{\nu}{2})F^m(\tfrac{\xi}{2},\tfrac{\nu}{2}) \tag{7.29}$$

where $H(z,n)$ is an FIR filter. Therefore, given $|F^0(u,v)|$ one can easily obtain $|F^1(u,v)|$ by

$$|F^1(\xi,\nu)| = |H(\tfrac{\xi}{2},\tfrac{\nu}{2})| \;\; |F^0(\tfrac{\xi}{2},\tfrac{\nu}{2})| \tag{7.30}$$

and hence, we can obtain any $|F^m(u,v)|$ for $m > 0$, and no approximation is required.

The approximate $f^m \in V_m$ can be obtained by performing several cycles of the basic Error-Reduction algorithm on this resolution until the number of iterations reaches an arbitrary limit that depends on the complexity of the image. The reconstructed image f^m is then expanded by a factor of two to produce f_o^m. The expanding operation (denoted as $\mathbf{A}_m^e$) is performed with a classical interpolation procedure. We put a zero between each sample of f^m and filter the resulting image with a low-pass filter. In our case, the low-pass filter used is the FIR low-pass filter H defined previously. By applying the expanding operator to the reconstructed image, we obtain

$$f_e^m = A_m^e f^m \tag{7.31}$$

where f_e^m becomes the initial guess used to reconstruct the image at a higher resolution f^{m-1}. Starting with this new initial guess and the Fourier modulus at this resolution $|F^{m-1}(u,v)|$, the same basic iterative procedure is performed on the new resolution level m-1.

This process is repeated following a coarse-to-fine strategy, until resolution level m=0 is reached, where the image is now reconstructed at the original resolution. The above procedure is described by the following steps:

Step 1:

Starting with an FIR filter coefficients h, derive the scaling function j numerically.

Step 2:

Starting with the $|F|=|F|$, compute the Fourier modula at all required resolutions : $|F^m|$ $0 < m \leq M$ using Eq.(7.30)

Step 3:

Set $m=M$, where M is maximum number of resolution levels desired.

Step 4:

Construct a random image f_o^m of size $2^{-m}N \times 2^{-m}N$.

Step 5:

Perform k iterations of the iterative algorithm starting with f_o^m as an initial guess for the algorithm.

Step 6:

Expand the resulting image from step 5 above using Eq.(7.31) to obtain f_e^m.

Step 7:

Initialize : $f_o^{m-1} = f_e^m$.

Step 8:

Set $m = m$-1.

Step 9:

If $m \geq 0$ Goto step 5, otherwise stop.

where the FIR filter coefficients h used here is Daub4.

The iterative algorithm used in step 5 can be either of the well-known Error-Reduction or input-output iterative algorithms. The object-domain constraints are typically the non-negativity of the object and its spatial

support. The spatial support of the object, $S_f(x,y)$, is typically obtained from the support of the object's autocorrelation function. Since the support of a 2D object cannot, in general, be determined uniquely from the support of its autocorrelation, we estimated the support of the object to be the smallest circle $C_f(x,y)$ with a diameter D_f such that $S_f(x,y) \subseteq C_f(x,y)$, where D_f is half the diameter of the object's autocorrelation function D_G. At each resolution level, we estimated the spatial support by a circle with diameter equal to $2^{-m}D_f$, where m is the resolution level.

In this case however, we are able to obtain a the spatial support of the object from the reconstructed image at the lowest resolution f_e^{M-1} by a simple thresholding procedure, which provides a tighter spatial support that improves the performance of the algorithm. This will be shown via an example in the following section.

The computational requirement for this algorithm is comparable with that of the pyramid approach. Since the main contributor for the computational cost of the existing iterative algorithms is the pair of 2D DFTs require to alternatively transfer the data from one domain to the other, we calculate the number of real additions and multiplications required by K number of cycles of the Error-Reduction algorithm. If the 2D DFT was computed using the 1D FFT in a row-column decomposition, then the number of real additions and multiplications required for a single cycle of the Error-Reduction algorithm operating on an $N \times N$ image is equal to $2 \times N^2 \times \log_2 N^2$, and hence for K number of iterations the total number of algebraic operations required will be equal to:

$20 \times K \times N^2 \log_2 N$

If the number of pixels in the original image f is equal to $N \times N$, then the approximation of the image at each resolution level $\mathbf{A}^d_m f$ will have $2^{-m}N \times 2^{-m}N$ (for $m>0$). If we assume that we started the algorithm at the lowest resolution $m=M$, then the number of algebraic operations required for the same number of the Error-Reduction algorithm (due only to the pair of 2D DFT since the imposition of constraints requires far less computation) will be equal to:

$$\frac{10 \cdot K \cdot N^2}{M} \sum_{m=0}^{M-1} 2^{-m} \log_2 (2^{-m} N)$$

assuming that the same number of iterations was perform at each resolution level. This shows that the multiresolution approach can dramatically reduce the computational cost of the existing techniques.

To illustrate the performance of the wavelet approach described, we used an image (Planet Saturn) of size 128×128 shown in Figure 7.8(a). The image reconstructed after 800 iterations of the original Error-Reduction algorithm is shown in Figure 7.8(b).

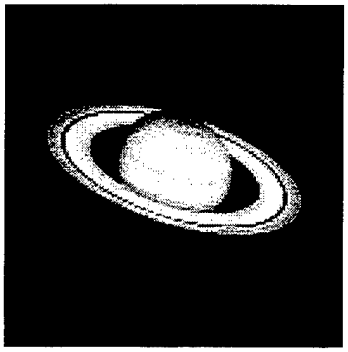
Figure 7.9 shows the spectrum decomposition of the Fourier magnitude obtained by Eq.(7.30), where the number of decomposition stages M was set equal to 3. Notice that at each resolution this decomposition divides the magnitude spectrum of the image into different frequency subbands. At the lowest resolution, only the low-frequency components are maintained while higher frequencies are attenuated. At each higher resolution, more higher frequency components are added, where at the original resolution the magnitude spectrum contains all the frequency components of the image.

Starting at the lowest resolution level (M=3), Figure 7.10(a) shows the reconstructed image f^3 obtained after performing 200 iterations of the Error-Reduction algorithm. At this stage, a circle with radius equal to $2^{-3}D_f$

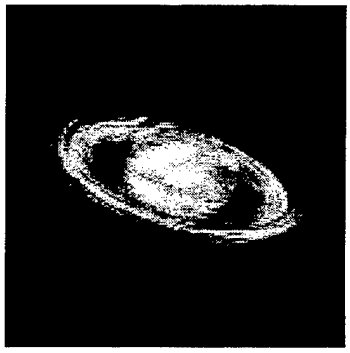
shown in Figure 7.10(b) was used as an estimate of the spatial support of the image, where D_f is half the diameter of the object's autocorrelation function D_G . Using this image and a simple thresholding procedure we obtained a tighter spatial support shown in Figure 7.10(c).

Expanding the reconstructed image and using it as the new initial guess for the iterative algorithm along with the new spatial support, we can obtain a new image at a higher resolution. This procedure is repeated for all resolution level till we reach the resolution m=0, where the image is reconstructed at the original resolution. Figure 7.9 shows the reconstructed images at four consecutive resolutions. In order to visually illustrate the superior quality of the image reconstructed using the multiresolution approach, Figure 7.11 shows the original image, the image reconstructed after 800 iterations of the original ER algorithm, and the image reconstructed using the multiresolution approach for which 200 iterations were performed at each of the four resolutions.

By examining the reconstruction procedure illustrated by Figure 7.11, one can see that at the lowest resolution, only the image features associated with low frequency components are reconstructed, and therefore the image appears to be very blurred. As the reconstruction procedure progresses, more high frequency components are reconstructed and more detailed features of the image start to appear.



(a) Original image of Planet Saturn.



(b) Image reconstructed after 800 iterations of the original ER algorithm

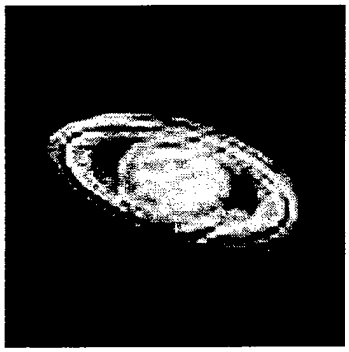
Figure 7.8: Original image vs. reconstructed image using the Error-Reduction algorithm.



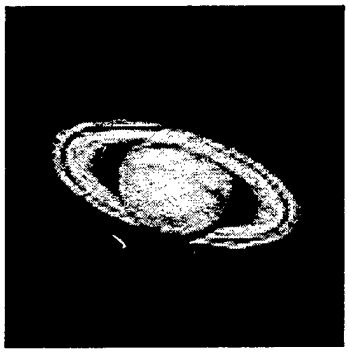
(a) Image reconstructed at the lowest resolution (m=3) after 200 iterations of the Error-Reduction algorithm.



(b) Image reconstructed at the 2nd resolution level (m=2) after 200 iterations of the Error-Reduction algorithm.

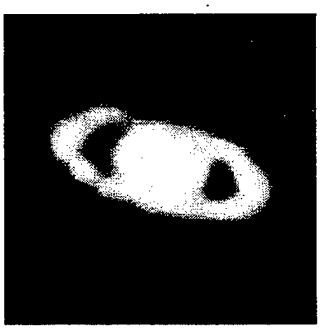


(c) Image reconstructed at the 1st resolution level (m=1) after 200 iterations of the Error-Reduction algorithm.

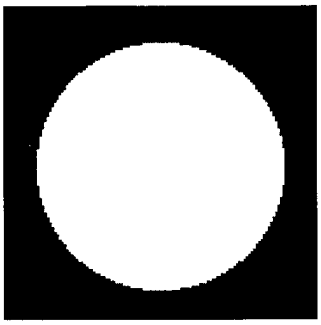


(d) Image reconstructed at the original resolution (m=0) after 200 iterations of the Error-Reduction algorithm.

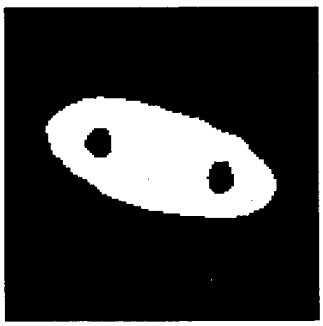
Figure 7.9: Multiresolution Reconstruction Procedure.



(a) Image reconstructed at the lowest resolution f^m



(b) Initial spatial support used.

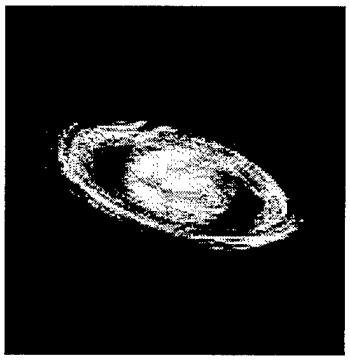


(c) Tighter support obtained from the image reconstructed at the lowest resolution shown in (a).

Figure 7.10: Reconstruction of Saturn image (a) using circle (b) for support. Tighter support (c) developed from initial reconstruction (a) and circle (b).



(a) Original image



(b) Image reconstructed using the original ER algorithm



(c) Image reconstructed using the Wavelet approach

Figure 7.11: Reconstruction of Saturn image (a) with direct ER approach (b) vs. the wavelet approach (c).

7.3 Conclusions

We have presented two new approaches to solve the problem of image reconstruction from the magnitude of its Fourier transform, or what is called, the phase retrieval problem, namely: the pyramid decomposition approach. and the Multiresolution Error-Reduction (MRER) approach. We have shown that this problem has become a subject of extensive interest from both fundamental and practical points of view. This problem emerges in many different disciplines, one of which is a new and growing field known as Unconventional Imaging. This technique can obtain high resolution images of objects at very long ranges without the use of large diameter optical elements by using coherent illumination.

Among all the phase retrieval algorithms currently used, the iterative algorithms remain the most efficient and practical approach in solving the problem. However, all these algorithms suffer from fundamental drawbacks that limits their use in future practical applications. In addition, we have shown that the Multiresolution phase retrieval algorithms developed from this research have shown significant advantages over their single-grid counterparts. Our analysis and computer simulation have revealed that the pyramid algorithm and the wavelet decomposition algorithm proposed here not only can overcome the stagnation problem of the existing iterative algorithms, but also provide a faster convergence rate while dramatically reducing the computational cost of the current algorithms.

In the following, we will summarize the major disadvantages of the current iterative techniques and the reasons why the pyramid decomposition approach and the Multiresolution Error-Reduction (MRER) approaches developed here demonstrated a superior performance, faster convergence and far less computations :

1. The stagnation problem, where the algorithm becomes trapped in a local minimum and failed to reach the global minimum, is the fundamental drawback of the current iterative techniques. This is a common problem in most nonlinear optimization techniques since the global convergence theory of these optimization procedures is based on the assumption that the initial starting point is close to the global minimum. Currently, all the iterative algorithms employed to solve the image reconstruction problem lack a good criterion for determining where to start the iterative procedure, and all these algorithms suggested the use of random numbers as an initial guess to give the algorithm an unbiased start. Clearly, in the absence of a good criterion to obtain a good starting point for the iterative procedure, the stagnation problem will remain a fundamental limitation of the current algorithms.

   When the problem is decomposed into different resolution levels, however, one can start the iterative procedure by performing a certain number of iterations at the lowest resolution level desired. This will furnish a quick rough estimate of the solution that can be used later as starting point at higher resolution. Since the number of variables at a coarse level are much less than that at the original resolution level, then the number of local minimum will be less that will give the algorithm a higher likelihood of arriving at a global minimum and escape local minima and hence avoid stagnation.
2. Iterative algorithms, in general, have a very slow convergence and the scheme appears to stall. There are two main reasons for this phenomenon [Briggs, 1987], first, all conventional iterative procedures has rapid decrease in the error during the early iterations due to the efficiency of these procedures to eliminate the oscillatory modes (high frequency components) of the error. Once the oscillatory modes have been removed, the iteration is much less effective in reducing the remaining smooth modes (low frequency components) of the error. Second, since most iterative algorithms operate on the *fine-grid*, that is the original lattice on which the image is defined, and therefore, the variables to be estimated, in this case the pixels of the image, are not sufficiently decoupled.

   The fact that the smooth modes of the error on a fine grid looks less smooth on a coarse grid [Briggs, 1987], suggests that the coarse grid relaxation will remove the low frequency components of the error responsible for the slow convergence of the iterative procedure. In addition, the decomposition of the original space into orthogonal subspaces, such as in the wavelet case, ensures that the variables at each resolution level are sufficiently uncorrelated, a fact that has been shown to provide a faster convergence of the iterative algorithms [Donoho, 1992].
3. The other drawback of the current algorithms is their high computational cost due to the pair of 2D DFTs performed in each cycle of the iteration. This computational burden is a function of the number of pixels in the image and the number of iterations required for the algorithm to converege to a solution. However, since, in a coarse grid, there are few unknown variables to be updated, the multiresolution decomposition makes the iterative procedure computationally less expefnsive.

4. Finally, is issue of robustness of the algorithm in the presence of noise. We have shown that the multiresolution approaches are less sensitive to noise than the current algorithms. The reason for that is the fact that the energy in most real-world images concentrates in relatively low-frequencies/resolutions and, hence, the SNR tends to be lower in high resolutions. When the noise power is large, the SNR in high resolutions would be so low in the image that the reconstruction process at the original resolution level would be rendered useless.

The computer simulation results presented in this here showed that indeed the multiresolution methods provide reconstructions that are superior in quality to their single-grid counterparts. For instance, the pyramid approach to used to reconstruct the image of planet Saturn yielded a reconstruction with 2.6 times less residual (NRMS) error that the direct approach. The wavelet decomposition approach produced an image with three times less residual error. On the other hand, both the pyramid approach and the multiresolution approach were able to achieve 68.1% reduction in computations over the existing algorithms. Both approaches performed better than the direct approaches in the presence of noise.

Clearly, the approaches presented here offer great advantages over the current algorithms employed in solving the problem of image reconstruction from the measurement of the magnitude of its Fourier transform. The multiresolution approaches, and particularly the wavelet representation, show considerable potential for future research in the phase retrieval problem and open the door for the development of more efficient algorithms to solve different signal recovery problems. Discussion of greater detail of the MRER methods may be found in [Rabadi, 1995].

7.4 References

1. Amaratunga and J. R. Williams, "Wavelet based Green's function approach to 2D PDEs," *Engineering Computations*," Vol.**10**, 349-367, 1993.
2. Beylkin, R. R. Coifman and V. Rokhlin, "Fast wavelet transforms and numerical algorithms I," *Commun. Pure and Appl. Math.*, vol.**44**, 141-183, 1991.
3. Briggs, W. L. A Multigrid Tutorial, SIAM, Philadelphia 1987.
4. Burt P. J. and E. H. Adelson, "The Laplacian pyramid as a compact image code," *IEEE Trans Comm.*, Vol.**31**, 532-540, 1983.
5. Daubechies, I. "Orthonormal bases of compactly supported wavelets," *Commun. Pure and Appl. Math.*, vol. **41**, 909- 996, Nov. 1988.
6. Daubechies, I., "Ten lectures on wavelets", CBMS-NSF Reg. Conf. Ser. Appl. Math. **61**, SIAM, Philadelphia, 1992.
7. Donoho, D. L. "Nonlinear solution of linear inverse problems by wavelet-vaguelette decomposition," *manuscript*, June 1992.
8. Dudgeon D. E. and R. M. Mersereau, Multidimensional Digital Signal Processing, Prentice-Hall, Englewood Cliffs, New Jersey, 1984.
9. Fienup, J.R., "Phase Retrieval Algorithms: a Comparison," *Applied Optics,* Vol. 21, No. 15, pp 2758-2769, 1982.
10. Gerchberg R. W. and W. O. Saxton, "A practical algorithm for the determination of phase from image and diffraction plane pictures," *Optik* **35**, 237-246, 1972.
11. Nieto-Vesperinas, M. "A Study of The Performance of Nonlinear Least-Square Optimization Methods in The Problem of Phase Retrieval," *Optica ACTA* **33**, 713-722, 1986.

12. Rabadi, W. A., Multiresolution Image Processing, PhD Dissertation, University of Central Florida, December, 1995.

13. Rosenfeld, A. Multiresolution Image Processing and Analysis, Berlin: Springer-Verlag, 1984.

14. Saxton, W.O. Computer Techniques for Image Processing in Electron Microscopy, Academic Press, New York, 1978.

7.5 Publications, Presentations and Theses

1. W. A. Rabadi, H. R. Myler, A. R. Weeks, and K. J. Gamble, "Pyramid Framework for Image Reconstruction from Nonimaged Laser Speckle", Proc. SPIE, **2484**, 309-320, July 1995.

2. W. A. Rabadi, "Multiresolution Image Reconstruction", Ph.D. Thesis, University of Central Florida, September 1995.

3. W. A. Rabadi, H. R. Myler and A. R. Weeks, "A Multiresolution Algorithm for Image Reconstruction from the Magnitude of its Fourier Transform", Opt. Eng. **35**, 1015-1024, (1996).

4. G. E. Hague, A. R. Weeks and H. R. Myler, "Histogram Equalization of 24-bit Color Images in the Color Difference (C-Y) Color Space", J. Elec. Imag. **4**, 15-22, (1995).

5. W. A. Rabadi, H. R. Myler and A. R. Weeks, "Large Scale Image Transforms on a Hyercube Supercomputer", Proc. of the 2nd Congress of the Network of Arab Scientists and Technologies Abroad (ASTA), Amman, Jordan, August 1994.

7.6 Appendix 7A

Rabadi W.A., H. R. Myler and A. R. Weeks, *"An iterative multiresolution algorithm for image reconstruction from the magnitude of its Fourier transform," Optical Engineering*, 35:4, 1996.

Iterative multiresolution algorithm for image reconstruction from the magnitude of its Fourier transform

Wissam A. Rabadi
Texas Instruments
Digital Signal Processing Group
12203 Southwest Freeway
Stafford, Texas 77251

Harley R. Myler, MEMBER SPIE
Arthur R. Weeks, MEMBER SPIE
University of Central Florida
Department of Electrical and Computer Engineering
Orlando, Florida 32816-2450

Abstract. Iterative algorithms are currently the most effective approaches to solving a number of difficult signal reconstruction and recovery problems, and all of these algorithms suffer from stagnation and computational complexity. We propose a new multiresolution iterative approach that employs the concept of a multiresolution pyramid. This method attempts to solve the problem of image reconstruction from the measurement of the image's Fourier modulus by decomposing the problem onto different resolution grids, which enables the iterative algorithm to avoid stagnation by providing a better initial guess and enabling a higher likelihood of arriving at a global minimum while dramatically reducing the computational cost. Results on both synthetic and real-world images are shown; a performance comparison with the direct iterative algorithm demonstrates the effectiveness of our approach in terms of convergence, robustness and computational efficiency. © *1996 Society of Photo-Optical Instrumentation Engineers.*

Subject terms: image reconstruction; phase retrieval; multiresolution pyramid; iterative algorithms; laser speckle.

Paper 30055 received May 25, 1995; revised manuscript received Oct. 6, 1995; accepted for publication Oct. 18, 1995.

1 Introduction

Image reconstruction from the measurement of the magnitude of the image's Fourier transform (or equivalently, from its autocorrelation function) has become the subject of intense interest from both fundamental and practical points of view.[1–8] This problem emerges in many different disciplines, including astronomy, x-ray crystallography, electron microscopy and optical design.[3] In applications, several versions of the image reconstruction problem appear, depending on how much is known and what is to be determined. The version of an image reconstruction problem we study here comes from an application in speckle imaging known as imaging correlography.[9,10] Imaging correlography is based on the fact that, by illuminating an object with a laser that has adequate spatial and temporal coherence, an estimate of the autocorrelation (correlogram) of the object's brightness distribution can be obtained from the measurements of the (nonimaged) backscattered laser speckle intensity. Since the Fourier transform of the autocorrelation of the object brightness function is equivalent to the square modulus of the Fourier transform of the brightness function,[11] an image of the object can be reconstructed if the phase associated with the measured Fourier transform magnitude can be determined, i.e., phase retrieval.

In this context, the image reconstruction problem can be stated as follows: Given the magnitude of the Fourier transform, $|F(u,v)|$, of an image $f(x,y)$, and some a priori information on $f(x,y)$, reconstruct the image $f(x,y)$, or equivalently, reconstruct the phase $\Phi(u,v)$ of $F(u,v)$. Several approaches have been proposed to solve this problem. These methods can be divided into two major categories. The first is an analytical approach that is based on the theory of the uniqueness of multidimensional spectral factorization.[2,3] The major advantage of this method is the theoretical elegance and completeness of the algorithm. Unfortunately, spectral factorization is highly sensitive to noise and is computationally intensive, factors that make it impractical for real-time applications. The second method of solving the phase retrieval problem is known as the iterative approach,[4–7] which seeks numerical solutions that minimize the distance between the measurement and its estimate rather than an attempt to find an exact solution. The first practical iterative method to be reported was the error-reduction algorithm of Gerchberg and Saxton,[5] later extended by Fienup.[6] The main advantage of this method is the ease of implementation on a digital computer and the fast convergence at the beginning of the iterations. However, in many cases these algorithms stagnate before reaching a solution and become trapped in a local minimum.[8] An additional limitation of iterative-type algorithms is the intensive computation required, which increases exponentially with increasing size of the image to be reconstructed.[9–11]

We present a multiresolution adaptation of the error-reduction algorithm that attempts to solve the stagnation and computational complexity problems by employing the concept of pyramids.[7,12] This approach is based on decomposing the problem of image reconstruction onto different resolutions. By using the solution obtained from the iterative algorithm at a lower resolution as an initial guess for the next finer level, then, following a coarse-to-fine strategy, the approach enables the iterative algorithm to escape

local minima by providing a better initial phase estimate while dramatically reducing the computational cost.

This paper is organized as follows: In Sec. 2 we briefly discuss the image correlography technique as described in Ref. 9 and show the computer simulation results of estimating the autocorrelation function of an object from the associated laser speckle patterns. We also provide a brief review of similar techniques using multiresolution to solve the reconstruction problem. In Sec. 3 we revisit the original iterative algorithm first proposed by Gerchberg and Saxton, then in Sec. 4 we introduce the concept of the multiresolution pyramid. The adaptation of the pyramid decomposition to the iterative algorithms as applied to the image reconstruction problem is presented in Sec. 5. Section 6 discusses the computational efficiency of the algorithm, while the performance of the algorithm compared with the existing iterative approaches is demonstrated in Sec. 7; finally, in Sec. 8 conclusions are drawn.

2 Image Correlography

Assume that an object with an optically rough surface is illuminated with a laser beam whose coherence length is at least twice as long as the object is deep. The backscattered intensity in the far field will produce a so-called fully developed speckle pattern.[13] Let the field reflected at the object's surface be given by[14]

$$f_r(x,y)=|f_0(x,y)|\exp[j\Phi(x,y)], \tag{1}$$

where $|f_0(x,y)|$ is the object's field amplitude reflectivity, and $\Phi(x,y)$ is the phase of the reflected field. The observed speckle pattern $I(u,v)$ may be expressed as the square modulus of the Fourier transform of $f(x,y)$, that is,[14]

$$I(u,v)=|\mathfrak{F}\{f_r(x,y)\}|^2. \tag{2}$$

The autocorrelation function of the illuminated object's brightness distribution function $\Gamma(x,y)$ can then be obtained from the average spectrum of the laser speckle intensity patterns, which can be implemented as[10]

$$\Gamma(x,y)=\frac{1}{N}\sum_{n=1}^{N}|\mathfrak{F}^{-1}\{I_n(u,v)-\bar{I}_n\}|^2, \tag{3}$$

where $I_n(u,v)$ is the n'th realization of the observed speckle intensity pattern (where each speckle realization can be obtained by slightly rotating the object), $\bar{I}_n$ is the mean intensity of that realization and N is the total number of speckle realizations used. Therefore, by using Eq. (3), one can estimate the autocorrelation function of the object.

A measure of the object's Fourier transform modulus can be directly obtained by Fourier transformation of the autocorrelation function followed by taking the square root. One can then reconstruct an incoherent image of the object from the estimated Fourier modulus by using a phase retrieval technique.

An example of this technique is shown in Figure 1, where a synthetic image (a ring shape) shown in Figure 1(a) was used to generate a sequence of speckle patterns. A typical speckle pattern associated with this object is shown in Figure 1(b). [This is a computer-generated speckle pattern produced by the UCFLadar© Simulator developed at the University of Central Florida, Center of Research and Education in Optics and Laser (CREOL), Orlando, Florida.] Figure 2(a) shows the actual autocorrelation function of the ring-shaped image. Figures 2(b), 2(c) and 2(d) illustrate the estimated autocorrelation function from 10, 100 and 1000 speckle frames, respectively. Reference 10 presents an excellent treatment of image correlography from both the theoretical and the experimental points of view.



Fig. 1 (a) Original image (128×128) of a ring-shaped object and (b) speckle pattern associated with the ring-shaped object.

3 Iterative Transform Algorithms

This family of algorithms follows the philosophy of the Gerchberg–Saxton algorithm known as the error-reduction algorithm.[5] All of these algorithms incorporate a similar idea—iterate between the spatial and frequency domains while successively satisfying a set of constraints in both.

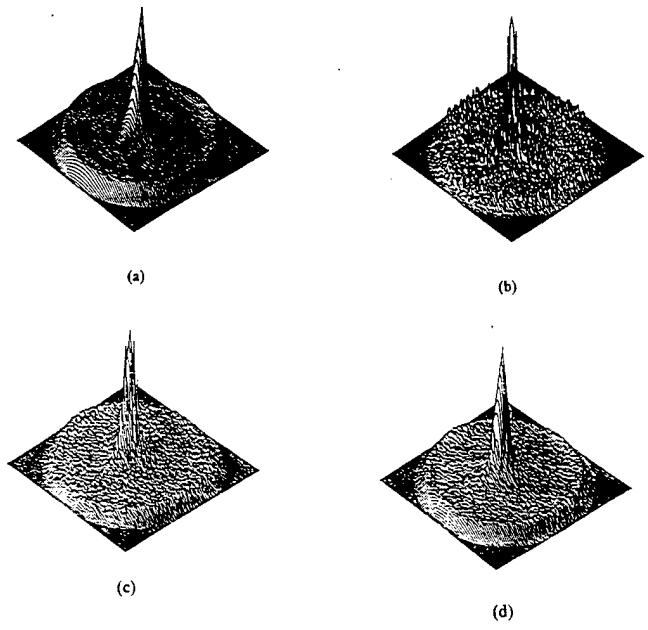


Fig. 2 Example of estimating the autocorrelation function from a sequence of speckle patterns: (a) actual autocorrelation function of the ring-shaped object, (b) autocorrelation function of the ring-shaped object estimated from averaging 10 speckle patterns, (c) autocorrelation function of the ring-shaped object estimated from averaging 100 speckle patterns, and (d) autocorrelation function of the ring-shaped object estimated from averaging 1000 speckle patterns.

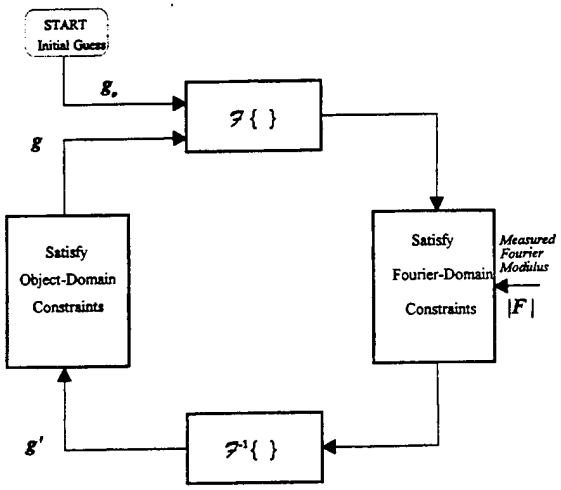


Fig. 3 A block diagram of the error-reduction algorithm.

Figure 3 shows a flow chart representing the error-reduction algorithm process, which can be described in the follow four steps[4]:

1. Fourier transform $g_k(x,y)$, to yield $G_k(u,v)$, where k is the iteration number and $g_0(x,y)$ is a random guess
2. make the minimum changes in $G_k(u,v)$ that allow it to satisfy the Fourier domain constraints to form $G'_k(u,v)$ an estimate of $F(u,v)$
3. inverse Fourier transform $G'_k(u,v)$ yielding $g'_k(x,y)$
4. make the minimum changes in $g'_k(x,y)$ that allow it to satisfy the object-domain constraints to form $g_{k+1}(x,y)$.

The object-domain constraints are typically the non-negativity of the object, and its spatial support object is a mask outside of which the image is constrained to be equal to zero. The spatial support for the reconstruction can be obtained from the support of the object's autocorrelation function since the diameter of the object is just half of the diameter of the autocorrelation.[6]

Although this algorithm and its modification by Fienup[6] are currently the most effective approaches to solving the phase retrieval problem, they still suffer from stagnation. This problem is due to the large number of attraction basins that cause the estimate to be trapped in a local minimum when the initial guess is not close enough to the global minimum. It has been shown[8] that the number of local minima in the phase retrieval problem increases with increasing number of pixels in the image; therefore, as the size of the image increases, it becomes difficult to find a good initial point that is close to the global minimum. An additional limitation of the iterative algorithm is the intensive computation due to the pair of 2D discrete Fourier transforms (DFTs) required for each cycle of the algorithm. This computational complexity is a function of the number of iterations required as well as the number of pixels in the image to be reconstructed.

Several attempts have been proposed to overcome the drawbacks of the error-reduction algorithm. One modification of the error-reduction algorithm is the hybrid input–output (HIO) algorithm[6] that, in practice, exhibits better convergence properties and appears to be able to escape local minima in the function being minimized. However, the HIO algorithm lacks a rigorous analysis of its convergence and can diverge when close to the true image. Other approaches[15] have used a small aperture telescope in conjunction with intensity measurements over a large aperture to obtain a low-resolution image that can be used to help in reconstructing a fine-resolution image by providing an estimate of the Fourier phase. Clearly, this approach can improve the performance of the iterative algorithm by providing a better start; however, it requires additional optical elements in order to obtain the initial low-resolution image.

We present a new multiresolution adaptation of the error-reduction algorithm described previously by decomposing the problem into smaller arrays. If several iterations of the error-reduction algorithm are performed on reduced sizes (lower resolutions) of the image, then the number of local minima, as well as the computations required, will be reduced. This indicates that it is more likely to find a good initial point that is close to the global minimum by coarser grid iterations rather than with a finer grid where a smaller number of independent variables is processed. Then, by locally interpolating the result to a finer grid and using this result as the new initial estimate, rapid convergence can be achieved. Therefore, the multiresolution scheme can be expected to speed up the convergence and reduce the total computational cost of the algorithm, as we will demonstrate.

4 Multiresolution Pyramid as a Computational Tool

Pyramids, in general, are data structures that provide a successively condensed representation of the input information.[12] What is condensed may be image intensity, so that the successive levels of the pyramid are reduced-resolution versions of an input image. However, this condensed representation also represents a smoothed or subsampled version of any information structure presented at the input, so that each level of the pyramid represents a reduced entropy of the input. Pyramids support fast coarse-fine search strategies that are robust, compact and computationally efficient. This structure has been used to speed up different types of image and low-vision operations, such as image segmentation, feature extraction and motion analysis.[12] Here, we will examine pyramids in the context of image reconstruction from three different perspectives—speed of convergence, robustness and computational efficiency.

The simplest type of image pyramid is constructed by repeatedly averaging the image intensities in nonoverlapping blocks of pixels. Given an input image $\mathbf{I_0}$ of size $N \times N$, where $N = 2^p$, applying the averaging process yields a reduced image size $\mathbf{I_1}$ of size $2^{p-1} \times 2^{p-1}$. Applying the process again to the reduced image yields a still smaller image $\mathbf{I_2}$ of size $2^{p-2} \times 2^{p-2}$ and so on till we reach $\mathbf{I}_{M-1}$ (called the *apex* or top of the pyramid) of size 1×1, which represents the average value of the input image. If we imagine these images stacked on top of one another, they constitute an exponentially tapering M-level pyramid

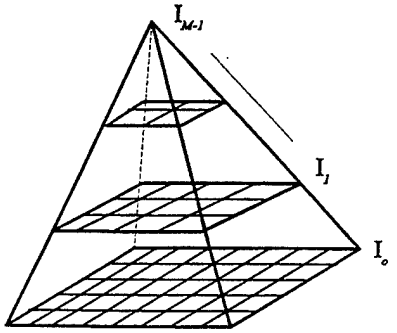


Fig. 4 Example of an M-level pyramid: P_M.

of images, as depicted graphically in Figure 4. In order to develop a multiresolution algorithm, several components must be specified:[16]

1. the number of levels and the size of the grid at each level
2. a restriction operation that maps from a fine level to a coarser grid
3. a prolongation operation that maps a solution at a coarse level to a finer grid
4. a coordination scheme that specifies the number of iterations at each level and the sequence of prolongations and restrictions.

In this work, we used a simple version of the restriction and prolongation operations, where weighted averaging is used for restriction and linear interpolation is used for prolongation. The coordination scheme is a simple coarse-to-fine algorithm, where the prolongated coarse solution is used as a starting point for the next finer level.

Let I_0 be the original image at the bottom of an M-level pyramid P_M, and let I_{M-1} be the top level of P_M; then for $0 < l < M$, we can define the restriction operation $\mathscr{R}$ on P_M such that,

$$I_l = \mathscr{R} I_{l-1}, \tag{4}$$

where I_l is a reduced-resolution version of I_{l-1}. In our case, $\mathscr{R}$ can be defined as

$$I_l(i,j) = \sum_{m=1}^{4}\sum_{n=1}^{4} w(m,n) I_{l-1}(2i+m-2, 2j+n-2)$$
$$1 \leq l < M \quad 0 \leq i,j < N/2^l, \tag{5}$$

where the weighting function $w(m,n)$ is the generating kernel (or the convolution mask) and $I(i,j)$ is causal, that is $I(i,j) = 0$ for $i < 0$ or $j < 0$. Since the number of sampled pixels has been reduced by half in each dimension, this kernel plays the role of a smoothing process (a low-pass filter that removes from the input spectrum all the components with a frequency larger than half the sampling frequency) to ensure meeting the Nyquist rate condition and avoid nonappropriate subsampling. This weighting function $w(m,n)$ is chosen subject to four constraints:[12]

1. separability: $w(m,n) = w_1(m) \cdot w_2(n)$
2. normalization: $\Sigma\Sigma w(m,n) = 1$
3. symmetry: $w(5-m,n) = w(m,5-n)$
4. equal contribution to the next level: $\Sigma_{i=0}^{1}\Sigma_{j=0}^{1} w(2i+m,2j+n) = 1/4, \ (m,n) = 1,2.$

The fourth constraint stipulates that each pixel at level l contribute the same total weight to the pixels at level $l+1$ in order to avoid distortion of the image. Although there exist different masks that satisfy this set of constraints, we found the following 4×4 kernel to be satisfactory:

$$W = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0.25 & 0.25 & 0 \\ 0 & 0.25 & 0.25 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}. \tag{6}$$

In a similar manner, we define a prolongation operation $\mathscr{P}$ on P_M such that,

$$I_l = \mathscr{P} I_{l+1}, \tag{7}$$

which, in our case, can be defined as

$$I_{l-1}(i,j) = \sum_{m=1}^{4}\sum_{n=1}^{4} w(m,n) I_l\left(\frac{i+m-2}{2}, \frac{j+n-2}{2}\right)$$
$$1 \leq l < M \quad 0 \leqslant (i,j) < N/2^l, \tag{8}$$

where only terms for which $(i+m-2)/2$ and $(j+n-2)/2$ are integers contribute to the sum, and $w(m,n)$ is the generating kernel given in Eq. (6). In this operation, the generating kernel will filter out the duplicated part of the spectrum that resulted from downsampling and will project the rest back to a finer grid.

5 Description of the Procedure

The first step in the image reconstruction procedure outlined here is to construct a bottom-up pyramid in which each level represents an approximation of the measured Fourier transform magnitude at that level. Since a precise measurement for the Fourier transform modulus at all the pyramid levels (i.e. $|F_l(u,v)|$, $0 < l < M$) is not available, an estimate of these measurements is sought. An approximation of the Fourier transform magnitude at coarse grids can be obtained by first downsampling the measured autocorrelation function, followed by Fourier transforming the subsampled autocorrelation and taking the square root of its real part. Using the restriction operation $\mathscr{R}$ defined by Eq. (5), the above procedure can be expressed as

$$|\tilde{F}_l(u,v)| = \sqrt{Re(\mathfrak{F}\{\mathscr{R}\Gamma_{l-1}(x,y)\})}^{1/2} \qquad 1 \leqslant l < M, \tag{9}$$

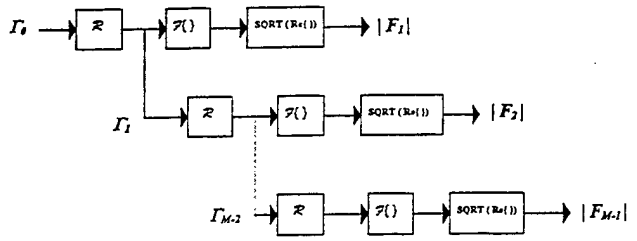


Fig. 5 Block diagram representing the process of constructing an M-level Fourier pyramid.

where $\Gamma_0(x,y) = \Gamma(x,y)$. The block diagram shown in Figure 5 illustrates this procedure. It is worthwhile to mention at this point that the approximation of the Fourier transform magnitude at different resolutions results in the construction of a Fourier pyramid in which each level represents an approximation of the Fourier transform magnitude of a low-pass version of the image to be reconstructed. In order to show that this is indeed the case, the autocorrelation function of the ring shape, $\Gamma_0(x,y)$, was subsampled and compared with the autocorrelation of the subsampled (low-pass) version of the ring-shaped image at different levels of the pyramid. Figure 6 illustrates this comparison and it should be clear that this is a good approximation of the autocorrelation function of a low-pass version of the image as measured by the normalized root-mean-square (NRMS) error defined by

$$E_\Gamma = \frac{\Sigma_m \Sigma_n |\Gamma_f(m,n) - \tilde{\Gamma}_f(m,n)|^2}{\Sigma_m \Sigma_n |\Gamma_f(m,n)|^2}. \tag{10}$$

Here, $\Gamma_f(m,n)$ is the autocorrelation function of a subsampled version of the original image and $\tilde{\Gamma}_f(m,n)$ is the subsampled autocorrelation of the original image.

After forming the Fourier pyramid, the reconstruction procedure proceeds from the pyramid top to the bottom by performing several cycles of the basic error-reduction algorithm on the lowest resolution level ($l = M-1$) until the number of iterations reaches an arbitrary limit that depends on the complexity of the image. The reconstructed image $g_{M-1}(i,j)$ is then transferred to the next level of the pyramid ($l = M-2$) with double the resolution by applying the interpolation process $\mathcal{P}$ defined by Eq. (8); that is,

$$g_{k-1} = \mathcal{P} g_k \qquad 1 \leq l < M. \tag{11}$$

The same basic iterative procedure is performed on the new pyramid level using the last image estimate computed from the previous level as the new initial guess for the algorithm. This process is repeated following a coarse-to-fine strategy, until level zero of the pyramid is reached, where the image is now reconstructed at the original resolution. This procedure is described by the following steps:

Step 1: Construct the Fourier pyramid as given by Eq. (9).

Step 2: Start at the highest level of the pyramid desired ($l = M-1$), using an initial random guess of size $2^{-l}N \times 2^{-l}N$ as a starting point.

Step 3: Perform k number of iterations of the error-reduction algorithm, where K is the total number of iterations to be performed, using the approximated Fourier transform magnitude at this level (i.e. $|F_l(u,v)|$) as the Fourier domain constraint.

Step 4: Interpolate the reconstructed image $g_l(x,y)$ to obtain $g_{l-1}(x,y)$ using Eq. (8), and use it as the new initial guess.

Step 5: Go to the next finer level: $l = l-1$.

Step 6: If $l < 0$ stop, otherwise go to Step 3.

The number of levels of the pyramid M for an image of size $N \times N$ is typically chosen to be equal to $\log_2 N$. In addition, the coordination scheme is chosen so that an equal number of iterations, $k = K/M$, is performed at each level of the pyramid, where K is the total number of iterations performed throughout the reconstruction procedure. This is a simple and straightforward coordination scheme and one may choose more optimal schemes so that the number of iterations at lower resolutions is less than that at the next resolution level since at lower resolutions the number of unknowns to be updated is lower. In addition, the object-domain support used in step 3 was obtained from the subsampled version of the autocorrelation function at each level of the pyramid. In order to avoid any problems that occur with the diameter constraint, we followed the strategy suggested by Fienup[4] for choosing the support mask in which in the first few iterations a smaller support mask is used and then this mask is enlarged to the correct support.

6 Computational Efficiency of the Algorithm

The most obvious advantage of the pyramid representation is that it provides a means of reducing the computational cost of the iterative algorithm by using a divide-and-conquer principle. To investigate the reduction in overall computational cost that results from using the multiresolution approach, we calculate the number of real additions and multiplications (due only to the pair of 2D DFTs since the imposition of constraints involves far fewer computations) required to perform K iterations of the error-reduction algorithm on an $N \times N$ image. If the 1D fast Fourier transform (FFT) is used to compute the 2D DFT in a row-column decomposition, then the evaluation of each 2D DFT requires $4N^2\log_2 N$ real multiplications and $6N^2\log_2 N$ real additions,[17] a total of:

$$20KN^2\log_2 N \tag{12}$$

algebraic operations for K iterations of the algorithm. When performing the same number of iterations on an M-level pyramid (where $k = K/M$ is the number of iterations performed at each level), the total number of algebraic operations reduces to:

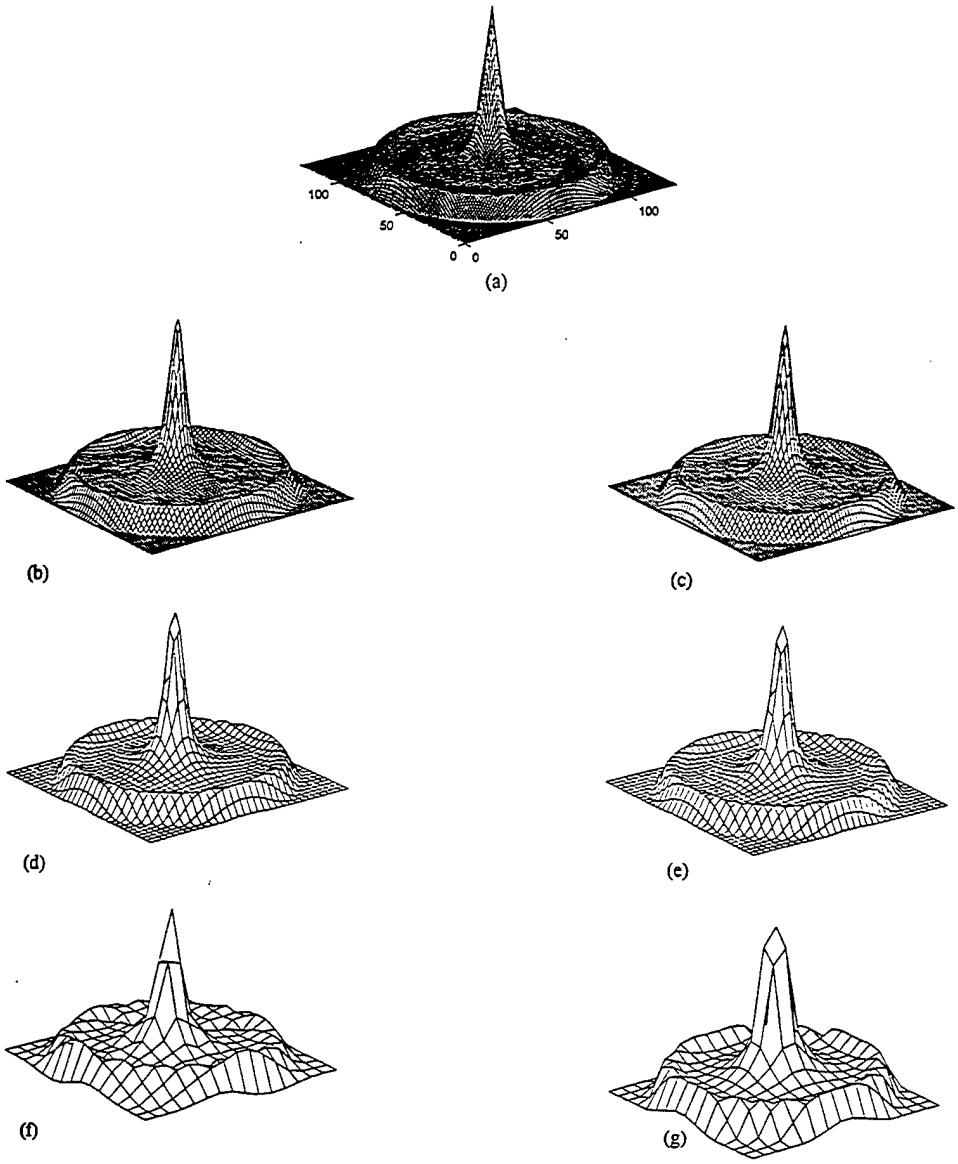


Fig. 6 Approximation of the autocorrelation function of the subsampled image of a ring-shaped object by a subsampled version of the autocorrelation function of the object: (a) actual autocorrelation function of the ring-shaped object (128×128); (b) autocorrelation function of the subsampled version of the ring-shaped image (64×64); (c) subsampled version of the autocorrelation function of the ring-shaped object (64×64). The NRMS error between (b) and (c) is equal to 0.078; (d) autocorrelation function of the subsampled version of the ring-shaped object image (32×32); (e) subsampled version of the autocorrelation function of the ring-shaped object (32×32). The NRMS error between (d) and (e) is equal to 0.157; (f) autocorrelation function of the subsampled version of the ring-shaped object image (16×16); and (g) subsampled version of the autocorrelation function of the ring-shaped object (16 ×16). The NRMS error between (f) and (g) is equal to 0.306.

$$\frac{20 \cdot K}{M} \sum_{l=0}^{M-1} \left(\frac{N}{2^l}\right)^2 \log_2\left(\frac{N}{2^l}\right). \tag{13}$$

This indicates that a considerable reduction in the computation is achieved by this decomposition of the problem.

7 Simulation Results

In the experiments that we performed, the pyramid approach produced reconstructions that are superior in quality (as measured by the residual error) to the direct iterative method. To illustrate the procedure of the multiresolution pyramid approach, a synthetic image of the ring shape



Fig. 7 Example of image reconstruction from the actual autocorrelation function (the direct approach): (a) original image (128×128) of the ring-shaped object and (b) image reconstructed after 100 iterations of the error-reduction algorithm on a single grid.

shown in Figure 7(a) was used. Figure 7(b) shows the reconstructed image after 100 iterations of the error-reduction algorithm on a single grid (i.e. direct approach) where the actual (noiseless) autocorrelation function was used to obtain the Fourier transform magnitude of the ring-shaped image. Figure 8 shows a 4-level Fourier pyramid constructed as described by Eq. (9) in Sec. 5, where each level of the pyramid represents an approximation of the Fourier transform magnitude at a different resolution (128×128, 64×64, 32×32, ... etc.). Again, the actual (noiseless) autocorrelation function of the ring-shaped image was used to obtain the Fourier transform magnitudes at each level of the 4-level pyramid by performing 25 iterations of the error-reduction algorithm at each level (a total of 100 iterations), where Figure 9(a) shows the 16×16 pixel images reconstructed at level 3 of the pyramid after 25 iterations of the error-reduction algorithm. By interpolating this image as given by Eq. (8), a new image with 32×32 pixel resolution can be obtained. This new image can now be used as the initial guess for the error-reduction algorithm. Figure 9(b) shows the reconstructed image at level 2 of the pyramid after 25 iterations of the error-reduction algorithm. The same procedure is repeated again to obtain the reconstructed image shown in Figure 9(c). Figure 9(d) shows the image reconstructed at the original resolution (i.e. level 0 of the pyramid).

To investigate the performance of the proposed approach in the presence of noise, we repeated the above procedure using a noisy estimate of the autocorrelation. Figure 10(a) shows the ring-shaped test image. Figure 10(b) shows the reconstructed image after 100 iterations of the error-reduction algorithm on a single grid, where an estimate of the autocorrelation function obtained from averaging 1000

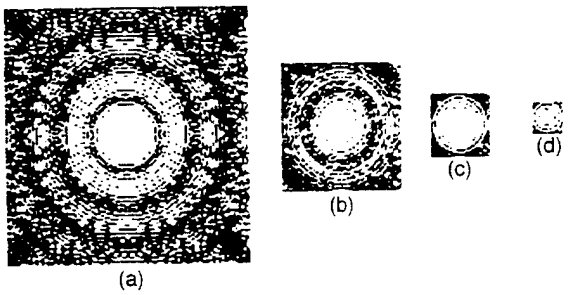


Fig. 8 The Fourier pyramid of the ring-shaped object: (a) level 3 of 16×16 pixel resolution, (b) level 2 of 32×32 pixel resolution, (c) level 1 of 64×64 pixel resolution, and (d) level 0 of 128×128 pixel resolution.

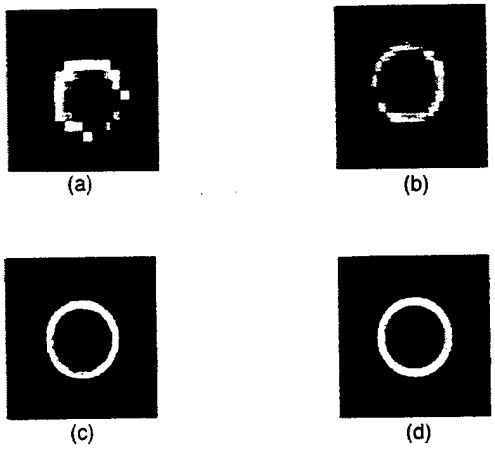


Fig. 9 Example of image reconstruction from the actual autocorrelation function (the multiresolution approach): (a) image reconstructed after 25 iterations on level 3 (16×16), (b) image reconstructed after 25 iterations on level 2 (32×32), (c) image reconstructed after 25 iterations on level 1 (64×64), and (d) image reconstructed after 25 iterations on level 0 (128×128).

speckle realizations (shown in Figure 2(d)) was used to obtain the magnitude of the Fourier transform of the image. Figures 10(c), 10(d), 10(e) and 10(f) show the reconstruction at each level of the 4-level pyramid by performing 25 iterations of the error-reduction algorithm at each level (a total of 100 iterations).

We repeated the same procedure using the same ring-shaped image, where a noisier estimate of the autocorrelation function obtained from averaging 100 speckle realizations, shown in Figure 2(c), was used to obtain the Fourier transform modulus. Figure 11(a) shows the original image, and Figure 11(b) shows the reconstructed image after 100 iterations of the error-reduction algorithm performed at a single grid. The results of the multiresolution approach at different levels of the pyramid using the same autocorrela-

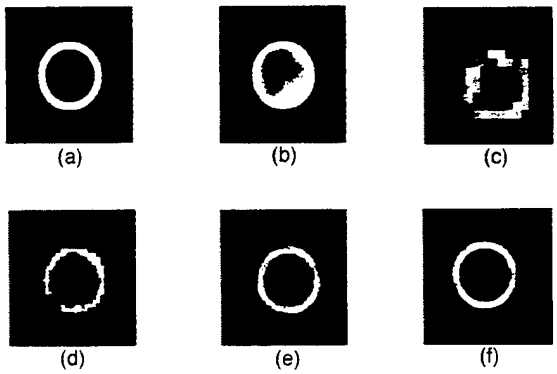


Fig. 10 Example of image reconstruction from the noisy estimate of the autocorrelation function obtained from 1000 speckle patterns (the direct approach versus the multiresolution approach): (a) original image of the ring-shaped object, (b) image reconstructed after 100 iterations of the error-reduction algorithm on a single grid (the direct approach), (c) image reconstructed after 25 iterations on level 3 (16×16), (d) image reconstructed after 25 iterations on level 2 (32×32), (e) image reconstructed after 25 iterations on level 1 (64×64), and (f) image reconstructed after 25 iterations on level 0 (128×128).

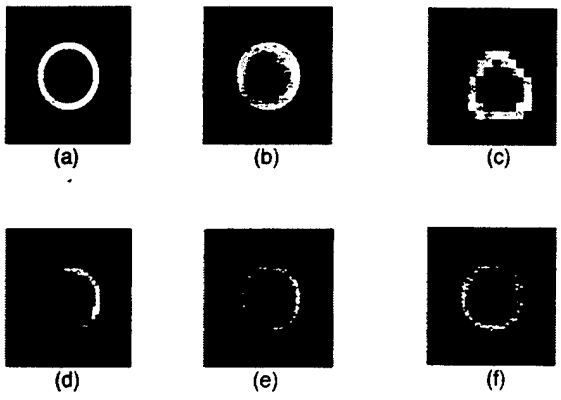


Fig. 11 Example of image reconstruction from the noisy estimate of the autocorrelation function obtained from 100 speckle patterns (the direct approach versus the multiresolution approach): (a) original image of the ring-shaped object, (b) image reconstructed after 100 iterations of the error-reduction algorithm on a single grid (the direct approach), (c) image reconstructed after 25 iterations on level 3 (16×16), (d) image reconstructed after 25 iterations on level 2 (32 ×32), (e) image reconstructed after 25 iterations on level 1 (64×64), and (f) image reconstructed after 25 iterations on level 0 (128×128).

tion function estimate (i.e., from 100 speckle realizations) are shown in Figures 11(c), 11(d), 11(e) and 11(f). We compared the quality of the reconstructions obtained by using the multiresolution approach with those using the direct approach from both the noiseless and noisy autocorrelation function. Table 1 shows the numerical values of the NRMS error of these reconstructions, which in this case is defined as[4]

$$E_f=\frac{\Sigma_u\Sigma_v|F(u,v)-G(u,v)|^2}{\Sigma_u\Sigma_v|F(u,v)|^2}. \tag{14}$$

These results indicate that the multiresolution approach produced reconstructions that are superior to the direct error-reduction method from both noiseless and noisy measurements of the autocorrelation function. This indicates that the multiresolution approach is more robust in the presence of noise.

A more complex image (the planet Saturn) of size 256 ×256 shown in Figure 12(a) was used to evaluate the performance of the multiresolution approach. The image reconstructed after 1000 iterations of the error-reduction algorithm on a single grid is shown in Figure 12(b). Figure 13 shows the reconstructed image obtained by using the multiresolution approach implemented on a 5-level pyramid, with 200 error-reduction iterations performed at each level (a total of 1000 error-reduction iterations), where Fig-

Table 1 NRMS error for reconstructions obtained using the multiresolution approach versus that obtained using the direct approach.

	NRMS Error	
	Direct Approach	Pyramid Approach
Reconstruction using the actual autocorrelation	0.40559	0.19052
Reconstruction using the autocorrelation estimated from averaging 1000 speckle realizations	0.45025	0.32803
Reconstruction using the autocorrelation estimated from averaging 100 speckle realizations	0.48256	0.40611

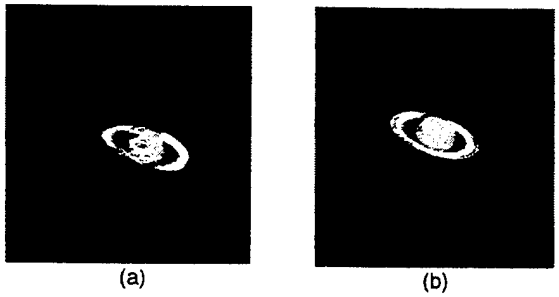


Fig. 12 Example of image reconstruction on a single grid: (a) original image of planet Saturn (256×256) and (b) image reconstructed after 1000 iterations of the error-reduction algorithm.

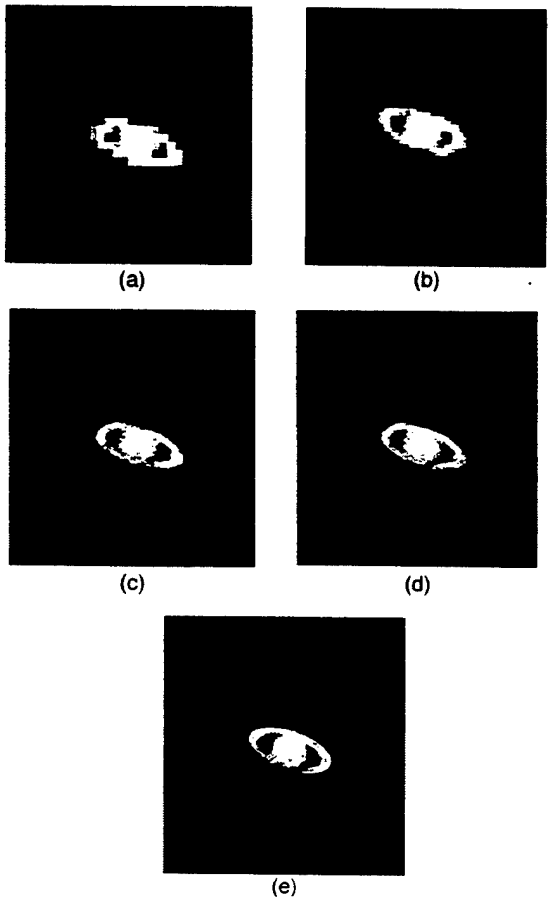


Fig. 13 Example of image reconstruction on a 5-level pyramid (the multiresolution approach): (a) image reconstructed after 200 iterations on level 4 (16×16), (b) image reconstructed after 200 iterations on level 3 (32×32), (c) image reconstructed after 200 iterations on level 2 (64×64), (d) image reconstructed after 200 iterations on level 1 (128×128), and (e) image reconstructed after 200 iterations on level 0 (256×256).

ures 13(a) through 13(e) show the images reconstructed on levels 5 through 0 of the pyramid, respectively. The final result, shown in Figure 13(e), is the image reconstructed at the original resolution. Simple calculations using Eqs. (12) and (13) show that a 68.1% reduction in computation is achieved, while the NRMS error is reduced by a factor of 2.6.

8 Conclusions

A multiresolution approach to the phase retrieval problem has been described that can significantly improve the performance of existing iterative methods. This approach is based on decomposing the problem of image reconstruction into different resolutions. By using the solution obtained from the iterative algorithm at a lower resolution as an initial guess for the next finer level and then following a coarse-to-fine strategy, one can produce reconstructions that are superior in quality to the direct error-reduction algorithm in terms of the residual error. From a computational efficiency perspective, we have shown that our approach dramatically reduces the computational cost.

Results of the multiresolution approach presented here suggest that the method may have potential in solving many of the phase retrieval problems experienced with current iterative approaches such as stagnation and computational complexity. However, the performance of this algorithm in overcoming other practical difficulties, such as twin-image formation and striping effects, will require further research. In addition, we are working toward better strategies of choosing the image-domain support constraints.

One alternative to the pyramid approach uses wavelet theory in the multiresolution decomposition.[18,19] It has been shown[14] that the phase retrieval problem can be solved by finding the least-squares solution of a set of nonlinear equations that can be implemented using a Gauss–Newton method that has been extensively analyzed. In addition, convergence of the algorithm has been shown. This method involves finding the pseudo-inverse of a Hessian matrix of size $N^2 \times N^2$, which would be required for reconstructing an image of size $N \times N$. This approach becomes computationally expensive and sometimes numerically unstable due to round-off errors for large N, which makes it impractical as the size of the image becomes larger. In the wavelet basis, however, large and dense matrices become sparse[20] and various numerical schemes can be used to find the matrix inverse efficiently. Our current research is directed toward investigating this wavelet approach to solving the image reconstruction problem.

Acknowledgment

This research was performed at the University of Central Florida (UCF) Center for Research and Education in Optics and Lasers (CREOL) under Office of Naval Research Contract N00014-92-C-0123.

References

1. J. R. Fienup, "Reconstruction and synthesis applications of an iterative algorithm," in *Transformation in Optical Signal Processing, Proc. SPIE* **373**, 147–160 (1981).
2. M. H. Hayes, "The reconstruction of a multidimensional sequence from the phase or magnitude of its Fourier transform," *IEEE Trans. Acoust., Speech, Sig. Proc.*, **ASSP-30**, 140–154 (1982).
3. R. G. Lane, W. R. Fright, and R. H. T. Bates, "Direct phase retrieval," *IEEE Trans. Acoust., Speech, Sig. Proc.*, **ASSP-35**, 520–525 (1987).
4. J. C. Dainty and J. R. Fienup, "Phase retrieval and image reconstruction for astronomy," in *Image Recovery: Theory and Applications*, H. Stark, Ed., Chapter 7, pp. 231–275, Academic Press, Orlando, FL (1987).
5. R. W. Gerchberg and W. O. Saxton, "A practical algorithm for the determination of phase from image and diffraction plane pictures," *Optik* **35**, 237–246 (1972).
6. J. R. Fienup, "Phase retrieval algorithms: A comparison," *Appl. Opt.* **21**, 2758–2769 (1982).
7. W. A. Rabadi, H. R. Myler, A. R. Weeks, and K. J. Gamble, "Pyramid framework for image reconstruction from nonimaged laser speckle," *Proc. SPIE* **2484**, 309–320 (1995).
8. M. Nieto-Vesperinas, "A study of the performance of nonlinear least-squares optimization methods in the problem of phase retrieval," *Optica Acta* **33**, 713–722 (1986).
9. P. S. Idell, J. R. Fienup, and R. S. Goodman, "Image synthesis from nonimaged laser speckle patterns," *Opt. Lett.* **12**, 858–860 (1987).
10. D. G. Voelz, J. D. Gonglewski, and P. S. Idell, "Image synthesis from nonimaged laser-speckle patterns: Comparison of theory, computer simulation and laboratory results," *Appl. Opt.* **30**, 3333–3344 (1991).
11. R. N. Bracewell, *The Fourier Transform and Its Applications*, 2nd ed., McGraw-Hill, New York (1978).
12. A. Rosenfeld, *Multiresolution Image Processing and Analysis*, Springer-Verlag, Berlin (1984).
13. J. C. Dainty, *Laser Speckle and Related Phenomena*, Springer-Verlag, Heidelberg (1984).
14. J. D. Gaskill, *Linear Systems, Fourier Transform and Optics*, Wiley, New York (1978).
15. J. R. Fienup and A. M. Kowalczyk, "Phase retrieval for a complex-valued object by using a low-resolution image," *J. Opt. Soc. Am.* **7**(3), 450–458 (1990).
16. R. Szeliski, *Bayesian Modeling of Uncertainty in Low-level Vision*, Kluwer Academic Publishers, Boston (1989).
17. D. E. Dudgeon and R. M. Mersereau, *Multidimensional Digital Signal Processing*, Prentice-Hall, Englewood Cliffs, NJ (1984).
18. S. G. Mallat, "A theory of multiresolution signal decomposition: The wavelet representation," *IEEE Trans. Pattern Anal. Machine Intell.* **11**, 674–693 (1989).
19. W. A. Rabadi, *Multiresolution Image Reconstruction*, PhD Diss., University of Central Florida, Orlando, FL (1995).
20. G. Beylkin, R. R. Coifman, and V. Rokhlin, "Fast wavelet transforms and numerical algorithms I," *Comm. Pure Appl. Math.* **XLIV**, 141–183 (1991).



Wissam A. Rabadi received his BSc in electrical engineering from the University of Jordan in Amman, Jordan in 1990; his MS in electrical engineering from the University of Detroit in Detroit, Michigan in 1991; and his PhD in electrical engineering from the University of Central Florida in Orlando, Florida in 1995. He is at present a member of the Digital Signal Processing Group at Texas Instruments in Houston, Texas. His interests include signal processing, image recovery, wavelets and multiresolution techniques, parallel processing and visual communication. Dr. Rabadi is a member of IEEE and SPIE.



Harley R. Myler received a BSEE from Virginia Military Institute in 1975 and MSEE and PhD degrees from New Mexico State University in 1981 and 1985. He is currently an associate professor in the Department of Electrical and Computer Engineering at the University of Central Florida in Orlando. He has published two books and more than 30 scientific papers in the areas of imaging, machine intelligence and computer simulation studies. Dr. Myler is a member of HKN, TBP, SPIE, and IEEE.



Arthur R. Weeks received his PhD in electrical engineering from the University of Central Florida in Orlando in 1987. He is an assistant professor of electrical and computer engineering at the University of Central Florida. His interests include the reduction of noise within images using adaptive nonlinear filters, color image processing, and the use of artificial neural networks in pattern recognition. Dr. Weeks is an associate editor of the *Journal of Electronic Imaging* and is a member of the IEEE, SPIE, and Tau Beta Pi.

7.7 Appendix 7B

Rabadi W. A., H. R. Myler, A. R. Weeks, K. J. Gamble, "Pyramid framework for image reconstruction from nonimaged laser speckle" *Proc. SPIE* **2484**, 1995

Pyramid framework for image reconstruction form nonimaged laser speckle

W. A. Rabadi, H. R. Myler, A. R. Weeks, K. J. Gamble

Electrical and Computer Engineering Department
University of Central Florida, Orlando, Florida 32816-0450
email: war@engr.ucf.edu

ABSTRACT

A multiresolution approach for image reconstruction from the magnitude of its Fourier transform has been developed and implemented by employing the concept of pyramid sampling. In this approach several iterations of the error reduction algorithm are performed at each level of the pyramid using a coarse-to-fine strategy, resulting in improved convergence and reduced computational cost.

Keywords: image reconstruction, phase retrieval, multiresolution, pyramids, laser-speckle

1. INTRODUCTION

The problem of phase retrieval from the intensity measurements has become a subject of extensive interest from both the fundamental and practical points of view. This problem emerges in many different disciplines such as X-ray crystallography, electron microscopy [1] and image reconstruction of coherently illuminated objects from measurements of backscattered laser-speckle intensity.[2,3] Given an image $f(x,y) \in L^2[I^2]$, then this image and its spectrum $F(u,v)$ constitute a Fourier transform pair:

$$F(u,v) = \Im\{f(x,y)\} = |F(u,v)|e^{j\Phi(u,v)} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)e^{j2\pi(ux+vy)}dxdy \quad (1)$$

where $|F(u,v)|$ is the modulus and $\Phi(u,v)$ is the phase of $F(u,v)$. The phase retrieval problem can be stated as, given $|F(u,v)|$ and knowledge of the support and/or nonnegativity of $f(x,y)$, reconstruct the image $f(x,y)$ or equivalently reconstruct the phase $\Phi(u,v)$ of $F(u,v)$. Several approaches have been proposed to solve this problem. These approaches can be divided into two major categories, the first, is the analytical approach that is based on the theory of uniqueness of multidimensional spectral factorization.[4] The major advantage of this approach is the theoretical perfectness and elegance of the algorithm. Unfortunately, this approach is highly sensitive to noise and computationally intensive which make it impractical for real time

0-8194-1837-4/95/$6.00

applications. The second approach used in solving the phase retrieval problem is the iterative approach[5], which seeks numerical solutions that minimize the distance between the measurement and its estimate. The first practical iterative algorithm is the Error-Reduction algorithm, first proposed by Gerchberg and Saxton,[6] and later extended by Fienup.[1] The main advantage of this method is the ease of implementation on a digital computer, and the fast convergence at the beginning of the iterations. However, in many cases these algorithms stagnate before reaching a solution, where the iterative algorithm becomes trapped in a local minimum. An additional limitation of the iterative algorithms is the intensive computation required, which increases exponentially as with increasing the size of the image to be reconstructed, as well as the number of iterations required. Several algorithms have been proposed to solve these shortcomings of the iterative algorithm mentioned above.[7,8]

We present a multiresolution adaptation of the Error-Reduction algorithm by employing the concept of pyramids[9,10] that will be shown to reduce the computational complexity, and provide a better initial phase estimate that enables the iterative algorithm escape local minima.

2. THE ERROR-REDUCTION ALGORITHM

Iterative algorithms follow a similar idea, iterate between the spatial and frequency domains while successively satisfying the constraints in both. Figure[1] shows a flow chart representing the steps of the ERA, and can be described in the follow four steps: [5]

1. Fourier transform $g_k(x,y)$ an estimate of $f(x,y)$ yielding $G_k(u,v)$.
2. Make the minimum changes in $G_k(u,v)$ which allow it to satisfy the Fourier domain constraints to form $G'_k(u,v)$ an estimate of $F(u,v)$.
3. Inverse Fourier transform $G'_k(u,v)$ yielding $g'_k(x,y)$.
4. Make the minimum changes in $g'_k(x,y)$ that allow it to satisfy the object-domain constraints (support and/or nonnegativity) to form $g'_{k+1}(x,y)$.

As we mentioned before, the main disadvantage of this algorithm and all iterative approaches is that in many cases it stagnate before reaching a solution. This stagnation problem is due to the large number of attraction basins which cause the estimate to be trapped into a local minimum if the initial guess is not close to the global minimum. It has been shown[11] that the number of local minima in the phase retrieval problem increases dramatically with the number of pixels in the image, therefore, as the size of the image increases, it becomes difficult to find a good initial point which is close to the global minimum, and hence the stagnation problem in the iterative phase retrieval algorithm arises. An additional limitation of the iterative algorithm is the intensive computation due to the pair of 2D DFT's required for each cycle of the algorithm. This computational complexity is a function of the number of iterations required, as well as the number of pixels in the image to be reconstructed. If several iterations of the Error-Reduction algorithm are performed on reduced sizes of the image, then the number of local

minima, as well as the computations required will be reduced. This indicates that it is more likely to find a good initial point that is close to the global minimum by coarser grid iteration rather than a finer grid where a smaller number of independent variables is processed. Therefore, the multiresolution scheme is able to speed up the convergence and to reduce the total computational cost of the algorithm.

3. MULTIRESOLUTION PYRAMID AS A COMPUTATIONAL TOOL

Pyramids, in general, are data structures that provide successively condensed representation of the input information.[10] What is condensed may be simply image intensity, so that the successive levels of the pyramid are reduced-resolution versions of the input image. However, this condensed representation may also represents a smoothed or subsampled version of any information structure presented at the input, so that each level of the pyramid represents a reduced *entropy* of the input. Pyramids support fast coarse-fine search strategies which are robust, compact and computationally efficient. This structure has been used to speed up different types of image and low-vision operations, such as image segmentation, feature extraction and motion analysis.[10] We will examine pyramids from the last of these perspectives, i.e. computational efficiency.

4. PYRAMID ARCHITECTURE

The simplest type of image pyramid is constructed by repeatedly averaging the image intensities in nonoverlapping blocks of pixels.[10] Given an input image I_o of size $2^n \times 2^n$ (called the *base* or the bottom of the pyramid), applying the averaging process yields a reduced image size I_1 of size $2^{n-1} \times 2^{n-1}$. Applying the process again to the reduced image yields a still smaller image I_2 of size $2^{n-2} \times 2^{n-2}$ and so on till we reach I_{M-1} (called the *apex* or top of the pyramid) of size 1×1 which represents the average value of the input image. If we imagine these images stacked on top of one another, they constitute an exponentially tapering M-level pyramid of images, as depicted in Figure[2]. In order to develop a multiresolution algorithm, several components must be specified:

- The number of levels and the size of the grid at each level.
- A *restriction* operation that maps from a fine level to a coarser grid.
- A *prolongation* operation that maps a solution at a coarse level to a finer grid.
- A *coordination scheme* that specifies the number of iterations at each level and the sequence of prolongations and restrictions.

In this work, we will use a simple version of the restriction and prolongation operations, where averaging is used for restriction and linear interpolation is used for prolongation. The coordination scheme is a simple coarse-to-fine algorithm, where the prolongated coarse solution is used as a starting point for the next finer level.

Let I_o be the original image at the bottom of an M-level pyramid P_M ,and let I_{M-1} be the top level of P_M, then for $0 < k < M$, we can define the restriction operation $\mathcal{R}$ on P_M as,

$$I_k = \mathcal{R}\, I_{k-1} \tag{2}$$

where I_k is a reduced-resolution version of I_{k-1} . In our case, $\mathcal{R}$ can be defined as,

$$I_k(i,j) = \frac{1}{4}\sum_{m=2i}^{2i+1}\sum_{n=2j}^{2j+1} I_{k-1}(m,n) \qquad \begin{matrix} 1 \le k < M \\ 0 \le i,j < N/2^k \end{matrix} \tag{3}$$

Similarly, we define a prolongation operation $\mathcal{P}$ on P_M as,

$$I_k = \mathcal{P}\, I_{k+1} \tag{4}$$

which, in the case of linear interpolation, is given by,

$$I_{k-1}(i,j) = 4\sum_{n=-2}^{2}\sum_{m=-2}^{2} I_k\left(\frac{i+n}{2}, \frac{j+m}{2}\right) \qquad \begin{matrix} 1 \le k < M \\ 0 \le i,j < N/2^k \end{matrix} \tag{5}$$

where only terms for which $(i+n)/2$ and $(j+m)/2$ are integers contribute to the sum.

5. DESCRIPTION OF THE PROCEDURE

We start by constructing a bottom-up pyramid, where each level represents an approximation of the given Fourier transform magnitude. The reconstruction procedure proceeds from the pyramid top to the bottom, using a coarse-to-fine strategy, and performing an arbitrary number of iterations at each level. This can be implemented as follows: given an $N \times N$ array, where $N=2^p$, an M-level pyramid is constructed, where the bottom, or zero, level ($k=0$) is the given Fourier transform magnitude, and each higher level of the pyramid represents an approximation of $|F(u,v)|$ obtained by repeatedly averaging the Fourier transform magnitude values within a 2×2 non-overlapping window using the restriction operation $\mathcal{R}$ defined by Eq.(3), that is,

$$\left|\tilde{F}_k\right| = \mathcal{R}\left|\tilde{F}_{k-1}\right| \qquad 1 \le k < M \tag{6}$$

where $|\tilde{F}_0(i,j)| = |F(i,j)|$. As a result, each array in this sequence is a smoothed subsampled version of its predecessor. Several cycles of the basic error-reduction algorithm are then performed on the lowest-resolution level ($k=M-1$) until the number of iterations reaches an arbitrary limit that depends on the complexity of the image. The reconstructed image $g_{M-1}(i,j)$ is

then transferred to the next level of the pyramid ($k=M-2$) with double the resolution by applying the interpolation process $\mathcal{P}$ defined by Eq.(5), that is,

$$g_{k-1} = \mathcal{P}\, g_k \qquad 1 \le k < M \tag{7}$$

The same basic iterative procedure is performed on the new pyramid level using the last image estimate computed from the previous level as an initial guess for the algorithm. This process is repeated following a coarse-to-fine strategy, until reaching level zero of the pyramid, where the image is reconstructed at the original resolution.

6. RESULTS AND DISCUSSION

The most obvious advantage of the pyramid representation is that it provides a means of reducing the computational cost of the iterative algorithm by using a divide-and-conquer principle. To investigate the reduction in the computational cost that results from adopting the multiresolution approach, we calculate the number of real additions and multiplications (due only to the pair of 2D DFT's since the imposition of constraints involves far less computations) required to perform K iterations of the error reduction algorithm on an $N \times N$ image. If the 1D FFT is used to compute the 2D DFT in a row-column decomposition, then the evaluation of each 2D DFT requires $8N^2 \log_2 N$ real multiplication's and $12N^2 \log_2 N$ real additions,[12] a total of $20KN^2 \log_2 N$ algebraic operations for K iterations of the algorithm. When performing the same number of iterations on an M-level pyramid, using the multiresolution approach, the total number of algebraic operations reduces to :

$$\frac{20 \cdot K}{M} \sum_{l=0}^{M-1} \left(\frac{N}{2^l}\right)^2 \log_2\left(\frac{N}{2^l}\right). \tag{8}$$

In the experiments that we performed, our approach produced reconstructions that are superior in quality (as measured by the residual norm) to the direct iterative method. To illustrate the procedure of the multiresolution pyramid approach, the Fourier transform magnitude of a 128×128 image of a " T " shape shown in Figure[3] is used. Figure[4] illustrates the magnitude-only synthesis using random phase. The reconstructed image after 100 iterations of the error reduction algorithm on a unigrid (direct approach) is shown in Figure [5]. Figure[6] shows a 4-levels pyramid constructed by restricting (subsampling) the given Fourier transform magnitude $|F(i,j)|$ at each level of the pyramid. Figure[7] shows the images reconstructed at different levels of a 4-levels pyramid with 25 Error-Reduction iterations performed at each level of the pyramid (a total of 100 iterations of the Error-Reduction algorithm).

A more complex image (image of Saturn) of size 256×256 (shown in Figure[8]) is used to evaluate the performance of the multiresolution approach. Figure[9] shows the magnitude-only synthesis using random phase. The image reconstructed after 800 iterations of the error

reduction algorithm on a single grid is shown in Figure[10]. Figure[11] shows the reconstructed image obtained by using the multiresolution approach implemented on a 4-level pyramid, with 200 error reduction iterations performed at each level. Simple calculations using Eq.(8) shows that 68.1% reduction in computation is achieved, while the residual error is reduced by a factor of 2.6.

An alternative to the pyramid approach is the wavelet theory proposed by Mallat.[13] Since the error reduction algorithm is actually the least square solution of a set of nonlinear equations, it can be implemented by a Gauss-Newton method. This involves finding the pseudo-inverse of a Hessian matrix of size $N^2 \times N^2$, which would be required for reconstructing an image of size $N \times N$. [11] This approach becomes computationally expensive and sometimes numerically unstable due to round-off errors for large N, which makes it impractical as the size of the image becomes larger. In the wavelet basis, however, large and dense matrices become sparse and various numerical schemes can be used to find the matrix inverse. Currently, we are investigating this wavelet approach to solving the image reconstruction problem.

7. CONCLUSIONS

A multiresolution approach to the phase retrieval problem has been described that can significantly improve the performance of existing iterative methods. From a computational efficiency perspective, we have shown that our approach drastically reduces the computational cost, and produces reconstructions that are superior in quality to the direct iterative algorithms in terms of the residual error.

8. ACKNOWLEDGMENT

This research was performed at the University of Central Florida (UCF) Center for Research and Education in Optics and Lasers (CREOL) under ONR contract N00014-92-C-0123.

9. REFERENCES

1. J. R. Fienup, "Reconstruction and synthesis applications of an iterative algorithm," In *Transformation in optical signal processing, Proc. SPIE* **373**, 1981, pp. 147-160.
2. H. R. Myler, A. R. Weeks, W. A. Rabadi, K. J. Gamble, " Visualization of laser-speckle behavior in time," Submitted to the *IEEE Trans. Visual. and Comp. Graph.*, Dec. 1994.
3. D. G. Voelz, J. D. Gonglewski, and P.S. Idell, " Image synthesis from nonimaged laser-speckle patterns : comparison of theory, computer simulation and laboratory results, " *Appl. Opt.*, **30**, pp. 3333-3344, Aug. 1991.
4. M. H. Hayes, " The reconstruction of a multidimensional sequence from the phase or magnitude of its Fourier transform," *IEEE Trans. Acoust., Speech, Sig. Proc.*. **ASSP-30**, pp. 140-154, 1982.

5. H. Stark, Image Recovery: Theory and Applications, Ch. 7, Academic Press, Orlando, 1987.
6. R.W. Gerchberg and W.O. Saxton, "A practical algorithm for the determination of phase from image and diffraction plane pictures," *Optik*, vol. **35**, pp. 237-246, 1972.
7. Z. Wu and A.H. Tewfik, " Multigrid algorithm for image reconstruction from its Fourier Modulus," in *Proc. ICASSP 91 Conf.* (Toronto, Canada), May 1991, pp. 2985-2988.
8. N.R. Guivens, Jr. and P.D. Henshaw, "Reduced computational algorithm for phase retrieval," in *Digital Image Synthesis and Inverse Optics, Proc. SPIE.*, vol. **1351**, pp. 661-672, 1990.
9. W. A. Rabadi, H. R. Myler, A. R. Weeks, " A multiresolution algorithm for image reconstruction from the magnitude of its Fourier transform, " Submitted to the *IEEE Signal Processing Letters*, Jan. 1995.
10. A. Rosenfeld, Multiresolution Image Processing and Analysis, Ch. 2, Berlin:Springer-Verlag,1984.
11. M. Nieto-Vesperinas, " A study of the performance of nonlinear least-square optimization methods in the problem of phase retrieval," *Optica Acta*, vol. **33**, pp. 713-722, 1986.
12. D.E. Dudgeon and R.M. Mersereau, Multidimensional Digital Signal Processing, Ch. 2, Englewood Cliffs, NJ: Printice-Hall, 1984.
13. S. A. Mallat, "A theory of multiresolution signal decomposition: the wavelet representation," *IEEE Trans. Pattern Anal. Machine Intell.*, **11**, 1989, pp. 674-693.

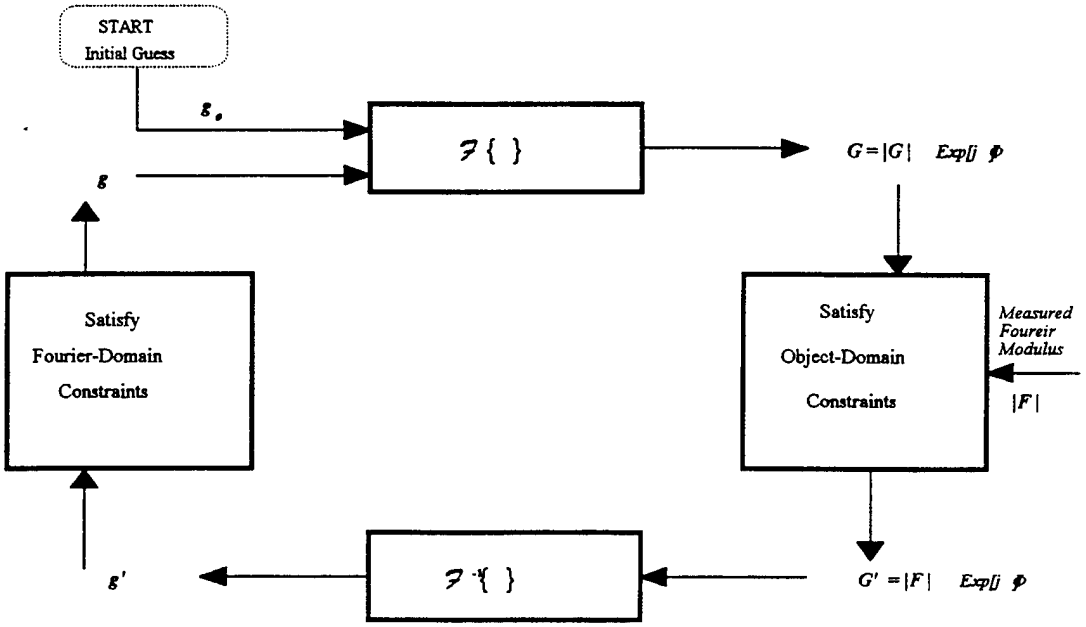


Figure 1 : *Block Diagram of the Error-Reduction Algorithm*

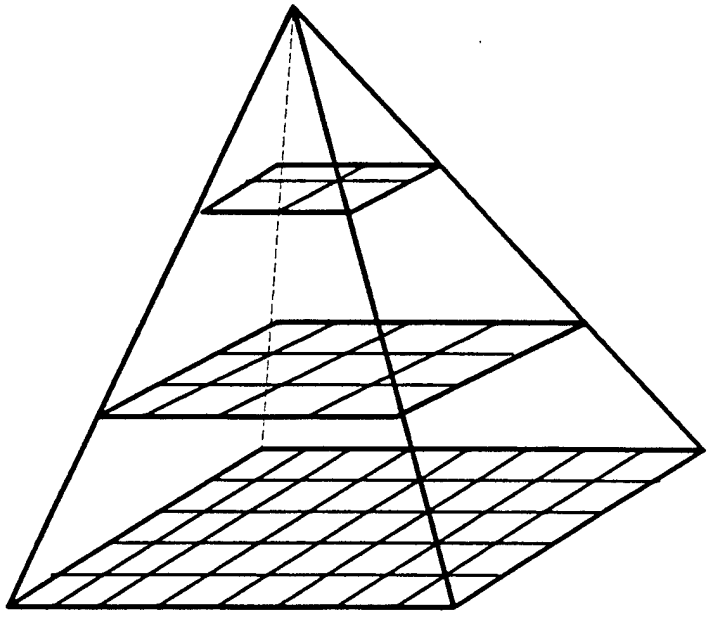


Figure 2 : *Multiresolution Pyramid*

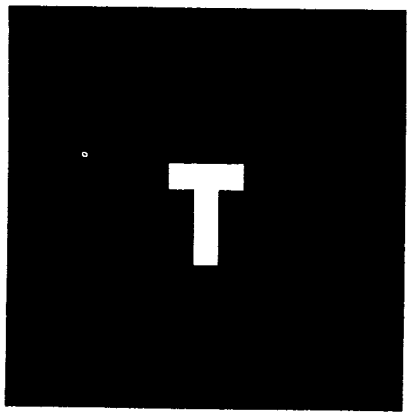


Figure 3:. *Original Image: T-Shape (128x128)*

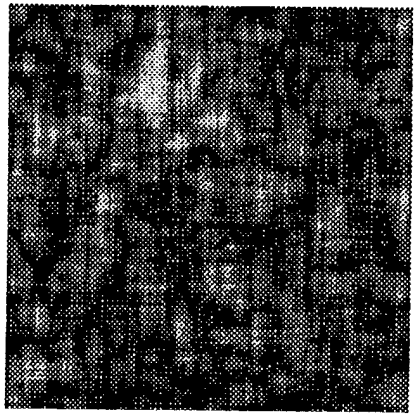


Figure 4: *Magnitude-only synthesis using random phase*

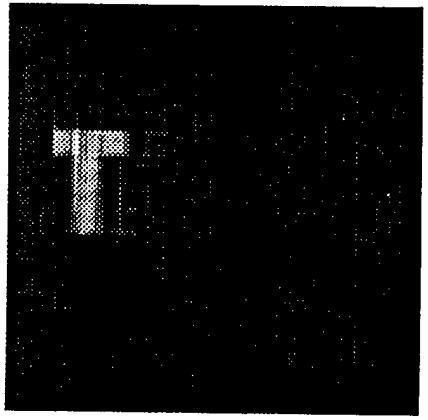


Figure 5: *Reconstruction on a single grid (128x128) after 100 iterations of the ERA*

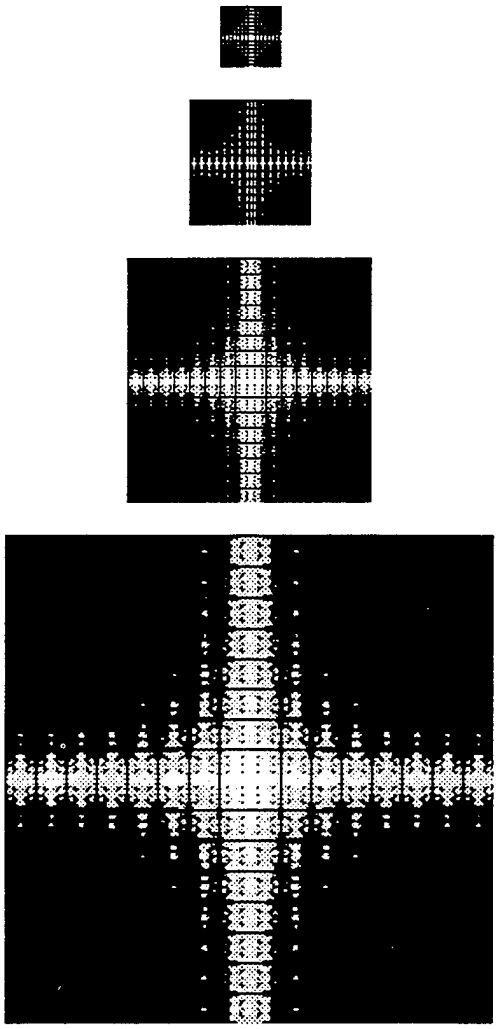
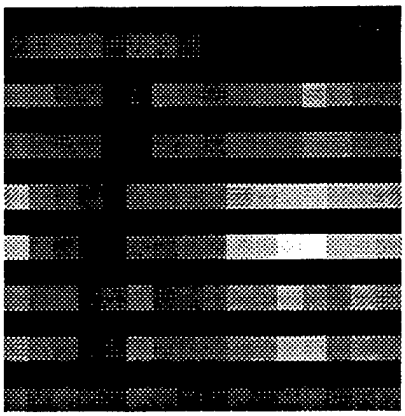


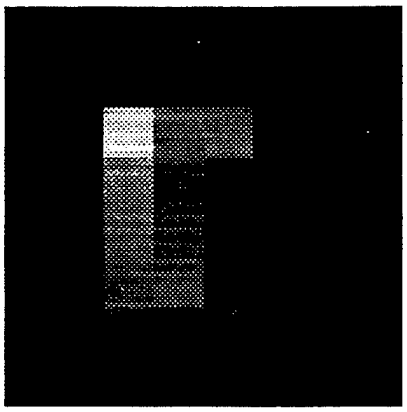
Figure 6 : *The Construction of a 4-Level Pyramid of the Foureir Transform Magnitude of the T-Shape Image.*



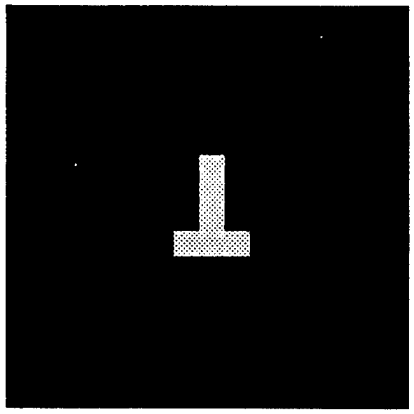
a- *Reconstruction on level 3 (16x16), 25 ER iterations.*



b- *Reconstruction on level 2 (32x32), 25 ER iterations.*



c- *Reconstruction on level 1 (64x64), 25 ER iterations.*



d- *Reconstruction on level 0 (128x128), 25 ER iterations.*

Figure 7 : *Reconstruction at Different Resolutions*

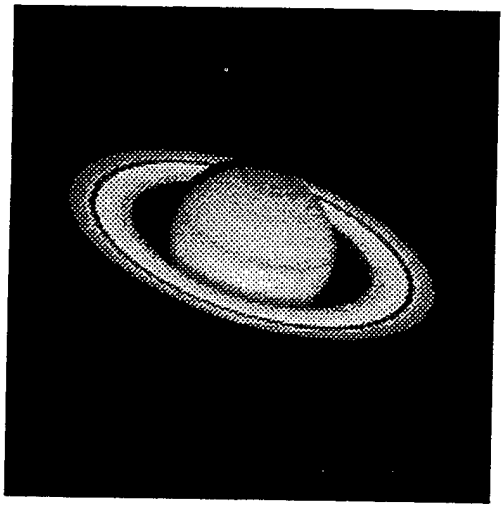


Figure 8 : *Original image of Saturn*

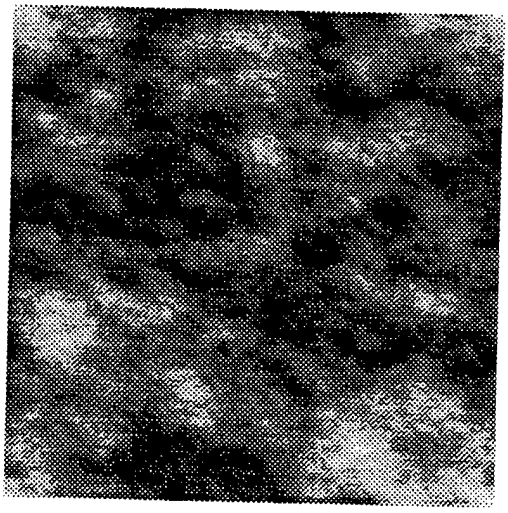


Figure 9 : *Magnitude-only synthesis using random phase.*

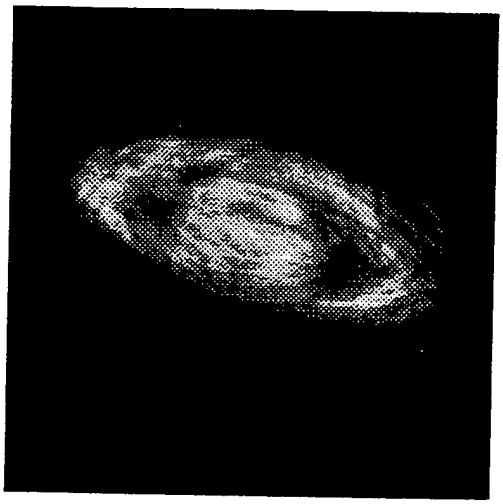


Figure 10 :*Reconstructed image after 800 iterations of the error reduction algorithm on single grid (direct approach).*

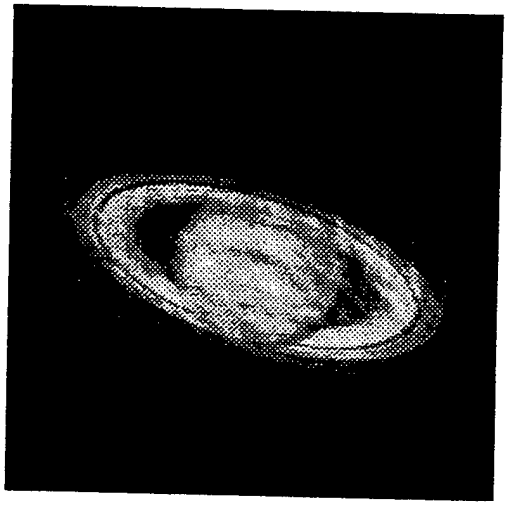


Figure 11 : *Reconstructed image obtained using the multiresolution approach on a 4-level pyramid, with 200 iterations of the error reduction algorithm performed at each level.*

8. RECIPROCAL PATH SCATTERING AND LASER RADAR

Robert A. Murphy, Ph.D. and R. L. Phillips, Ph.D.

8.1 Background - what is RPS?

One of the technical areas investigated by the CREOL Laser Radar Laboratory under the BMDO contract is reciprocal path scattering (RPS). RPS is a natural consequence for the monostatic receiver geometry when a random medium (or media) scatters the radiation during the round-trip propagation from the transmitter to the receiver. RPS has been observed due to scattering from rough surfaces, round-trip scattering through atmospheric turbulence, as well as scattering from suspended particulates. For BMDO applications, both scattering from rough surface targets and round-trip scattering through atmospheric turbulence produce RPS for monostatic laser radar transceivers.

When a receiver and transmitter share a common optic axis, field correlation between the different transmit and return paths imposed by scattering from random media produces a backscatter enhancement (BE) effect. This effect produces a pronounced peak in the angular distribution of scattered light in the anti-incident direction. BE is also referred to as backscatter amplification and enhanced backscattering in the scientific literature. Two distinct scattering phenomena, unique to scattering by random media, are capable of producing backscatter enhancement: spatially coherent reciprocal path scattering and spatially incoherent random focusing. The theoretical investigations performed by the CREOL Laser Radar Laboratory are concerned with backscatter enhancement that is produced entirely by RPS.

In general, scattering by random media produces an incoherent return, i.e. there is no fixed relative phase between transversely translated versions of a field in the transceiver plane; and the spatial distribution of the resulting intensity component fluctuates randomly about a mean value. Reciprocal path scattering, however, produces a narrow, pronounced coherent peak in the backscatter direction that is superimposed upon the incoherent component of the angular distribution of the average backscattered intensity. This peak is produced by the constructive interference of wave pairs that traverse reciprocal, time-reversed, scattering paths from the transmitter to the monostatic receiver that are imposed by randomly scattering media.

RPS provides the opportunity to partially mitigate the adverse effects of incoherent scattering by random media. Incoherent scattering by turbulence and rough surface targets produces amplitude and phase fluctuations in the backscattered field that degrade the performance of imaging systems by randomizing the spatial frequency information for target position and target identification, i.e. discrimination between the target and decoys. The spatial coherence of RPS may be exploited to recapture near diffraction limited target spatial frequency information for coherent reciprocal path imaging (RPI) applications. Since turbulence scattering permits a greater range of backscattered spatial frequencies to reach the transceiver, coherent RPI techniques actually capture more spatial frequency content in the presence of atmospheric turbulence than would be received in the absence of turbulence. In other words, refractive scattering by the atmospheric turbulence layer may someday actually be used as an extended collection aperture for coherent RPI systems.

8.2 Accomplishments

An analytically tractable phenomenological model was derived by Murphy [1] for scattering by the combination of atmospheric turbulence and a rough surface target. Phenomenological models were heuristically developed for the atmospheric turbulence reciprocal path scattering (AT-RPS) and the rough surface RPS (RS-RPS) processes individually and then for the combined process RPS (CP-RPS). These models are based upon the scattering geometry, and physical surface and turbulence parameters.

The model for rough surface scattering assumes the gamma probability distribution for the random surface heights and is based upon the rms surface slope and the geometry of the scattering processes that occur at the rough surface. The average backscattered intensity is assumed to consist of two components: an incoherent component produced by the self-interference of single scattered (SS) and double scattered (DS) field components, and a coherent component produced by RS-RPS. The average total backscattered intensity from the random rough surface is modeled as that produced by the superposition of three random arrays, two representing incoherent SS and DS events, and the other representing coherent RS-RPS.

The model for round-trip propagation through atmospheric turbulence is based upon the turbulence field transverse spatial coherence lengths over the up-link and down-link propagation paths. A reference plane is first determined at which the turbulence field transverse spatial coherence lengths for the up-link and down-link propagation paths are equal; this field transverse spatial coherence length defines the dimensions of the turbulence coherent propagation cells in the reference plane. The strength of the coherent AT-RPS peak is dependent upon the number of these cells within the transverse area of the up-link beam at the reference plane. The transverse length of the coherent AT-RPS peak in the transceiver plane is modeled as the main Airy diffraction lobe produced by an effective coherent AT-RPS aperture at the reference plane radiating in the direction of the transceiver.

The model for the CP-RPS produced by the combination of saturated atmospheric turbulence and a strong scattering rough surface provides physical insight into the resulting scattering based upon physical surface and turbulence parameters. The coherent CP-RPS enhancement factor is derived and expressed in terms of the coherent intensity enhancements produced by AT-RPS and RS-RPS individually. This relationship permits the prediction of real world coherent CP-RPS backscattered intensity enhancement based upon laboratory measurements of RS-RPS and calibrated field measurements of AT-RPS. Under the assumptions of this analysis, the coherent CP-RPS intensity peak is almost entirely produced by the average coherent intensity contribution due to field components that traverse reciprocal scattering paths both through the atmospheric turbulence and at the rough surface.

The modification to the traditional laser radar equation (LRE) [2] first incorporates the detrimental effect of beam spreading by atmospheric turbulence; then the effect of the backscatter enhancement produced by turbulence is added. The effect of the backscatter enhancement produced by scattering by rough targets is next modeled in the absence of turbulence; and finally, the analysis models the combination of turbulence beam spreading, turbulence backscatter enhancement, and rough surface backscatter enhancement.

The results of this analysis indicate that the traditional LRE is too oversimplified for practical power budget calculations when a monostatic transceiver geometry is employed. As expected, beam spreading by atmospheric turbulence significantly reduces the return signal photons incident upon the transceiver; this effect is

only partially mitigated when turbulence backscatter enhancement is taken into account. When both turbulence and rough target backscatter enhancement occur, the number of return signal photons incident upon the transceiver exceeds that predicted by the traditional LRE, even in the presence of turbulence beam spreading. This analysis illustrates that the traditional LRE can be off by more than an order of magnitude when the first order effects of turbulence and rough target scattering are considered.

8.3 Conclusions

A rigorous model for the scattering produced by the combination of atmospheric turbulence and rough surfaces is ultimately desired for RPI applications [3-10] that includes weak turbulence scattering and oblique incidence for the rough target scattering. The dissertation "Scattering from Rough Surfaces and Atmospheric Turbulence in Monostatic Laser Radar Systems" provides the foundation for the development of a rigorous CP-RPS model. As demonstrated by the dissertation, the mathematics necessary to accomplish this goal are very complicated even for the most simple atmospheric turbulence and rough surface scattering models. A logical course of action to obtain the desired goal is to progressively build rigor into the foundation presented in the dissertation based upon experimental data. This may be accomplished by initially adopting a second-order Kirchoff Approximation method [11- 13] for RS-RPS as well as a second-order Rytov solution method for AT-RPS. Finally, the extinction theorem method could then be applied to model the scattering from both random media. With active research support, this research may then someday provide the capability to achieve exceptional imaging performance through the Earth's atmosphere.

Several new issues have come to light since the publication of the modified LRE that accounts for turbulence beam spreading as well as turbulence and rough target backscatter enhancement. The general trend of the results are probably still valid, i.e. the LRE is too oversimplified for even first order estimates when the monostatic transceiver geometry is applied. With a minimum level of effort (approximately 12 man months) these new issues could be incorporated that would correctly formulate the modified LRE for monostatic transceiver applications. This would be a great benefit to system designers in general as well as benefit RPI applications.

8.4 References

1. Murphy, R. A., Scattering from Rough Surfaces and Atmospheric Turbulence in Monostatic Laser Radar Systems", 1995. University of Central Florida Ph.D. Dissertation.

2. Murphy, R. A., P. Gatt, and R. L. Phillips, "Reciprocal Path Scattering from Surfaces and Turbulence in Monostatic Laser Radar - Theory", SPIE 1993 OE/Aerospace Sensing Symposium: Program on Propagation and Imaging - Atmospheric Propagation and Remote Sensing II, No. 1968-70, (1993).

3. Harvey, J. E. and A. Kotha, "Sparse Array Configurations Yielding Uniform MTF's in Reciprocal Path Imaging Applications", Optical Communications 106:178-182, (1994).

4. Solomon, C.J. and J. C. Dainty, "Imaging a Coherently Illuminated Object Through a Random Screen By Using A Dilute Aperture", Journal of the Optical Society of America A 9:1385-1390, (1992).

5. Solomon, C. J., R. G. Lane, T. Mavroidis, and J. C. Dainty, "Double Passage Imaging Through a Random Screen using a Non-redundant Aperture", Journal of Modern Optics 38:1993-2008, (1992).

6. Dainty, J. C., T. Mavroidis, and C. J. Solomon, "Double Passage Imaging Through Turbulence", SPIE Propagation Engineering: Fourth in a Series vol. 1487:2-9, (1991).

7. Jakeman, E., "Active Imaging Through a Random Phase Screen", Journal of Physics D 24:227-232, (1991).

8. Mavroidis, T., C. J. Solomon, and J. C. Dainty, "Imaging a Coherently Illuminated Object After Double Passage Through a Random Screen", Journal of the Optical Society of America A 8:1003-1013, (1991).

9. Mavroidis, T. and J. C. Dainty, "Imaging After Double Passage Through a Random Screen", Optics Letters 15:857-859, (1990).

10. Mavroidis, T., J. C. Dainty, and M. J. Northcott, "Imaging of Coherently Illuminated Objects Through Turbulence: Plane-wave Illumination", Journal of the Optical Society of America A 7:348-355, (1990).

11. Ishimaru, A. and J. S. Chen, "Scattering from Very Rough Surfaces Based on the Modified Second-order Kirchoff Approximation with Angular and Propagation Shadowing", Journal of the Acoustic Society of America 88:1877-1883, (1990).

12. Chen, J. S. and A. Ishimaru, "Numerical Simulation of the Second-order Kirchoff Approximation from Very Rough Surfaces and a Study of Backscattering Enhancement", Journal of the Acoustic Society of America 88:1846-1850, (1990).

13. Bruce, N. C., and J. C. Dainty, "Multiple Scattering from Random Rough Surfaces Using the Kirchhoff Approximation", Journal of the Optical Society of America 38:579-590, (1991).

8.5 Publications, Presentations, and Theses

1. R. A. Murphy and R. L. Phillips, "Atmospheric Turbulence Reciprocal Path Scattering Effects for a Ground-Based Monostatic Laser Radar Tracking a Space Target", Proc. SPIE 2828, 52, (1996).

2. R. A. Murphy and R. L. Phillips, "Reciprocal Path Scattering Effects for a Ground-Based Monostatic Laser Radar Tracking a Space Target Through Turbulence", Submitted to App. Opt.

3. R. A. Murphy and R. L. Phillips, "Reciprocal Path Scattering due to the Combination of Atmospheric Turbulence and Rough Surfaces", Proc. SPIE 2471, 56 (1995).

4. R. A. Murphy, P. Gatt, and R. L. Phillips, "Reciprocal Path Scattering from Surfaces and Turbulence in Monostatic Ladar - Theory", Proc. SPIE 1968, 650-661, 1993.

5. R. A. Murphy, "Scattering by Rough Surfaces and Atmospheric Turbulence in Monostatic Laser Radar Systems", Ph.D Thesis, University of Central Florida, May 1995.

8.6 Appendix 8A

R. A. Murphy, P. Gatt, and R. L. Phillips, "Reciprocal Path Scattering from Surfaces and Turbulence in Monostatic Ladar - Theory", Proc. SPIE 1968, 650-661 (1993).

Reciprocal Path Scattering from Surfaces and Turbulence in Monostatic Ladar - Theory

Robert A. Murphy, Philip Gatt, and Ronald L. Phillips
Laser Radar and Wave Propagation Laboratories
Center for Research in Electro-Optics and Lasers (CREOL)
University of Central Florida
12424 Research Parkway
Orlando, Fl. 32826

1. Abstract

A model is presented which quantifies the combined effect of rough surface and turbulence double passage reciprocal path scattering (RPS) on the power received by a ground-based monostatic laser radar. This is not a rigorous analysis; instead a phenomenological modification to the traditional laser radar equation (LRE) for power linkage is presented, which includes the first order effects of atmospheric turbulence as well as rough surface and double passage RPS.

2. Introduction

A monostatic ladar system is one in which the transmitter and receiver optical paths utilize the same aperture for beam transmission and reception, respectively. The transmitted beam propagates through the turbulent atmosphere to the target, is back-scattered by the target surface, and then propagates downward through the atmosphere in the direction of the receiving aperture. The two distinct scattering processes encountered by the propagating beam are due to rough surface scattering by the target, and scattering caused by atmospheric refractive index variations over the path due to atmospheric turbulence. Reciprocal time-reversed multiple scattering paths exist for each scattering process due to EM reciprocity; the combination of these processes sequentially produces coherent (in-phase) wave segment pairs at the receiver. The coherent component of the received power contains diffraction-limited information which is exploited by active coherent imaging applications[1].

At laser wavelengths, the surface irregularities of physical targets are comparable to the wavelength of the illuminating beam. This surface roughness contributes significantly to the scattered cross section of the target. Multiple scattering processes are experienced by wave segments incident on the surface. Reciprocal path coherent wave segment pairs exist for multiply scattered wave segments within the valleys of the rough target surface[2], as well as for surface wave resonant polaritons[3]. Coherent backscattered intensity has been observed as large as 80 times the diffuse backscattered incoherent intensity[4].

With respect to beam propagation, the Earth's lower atmosphere is a random inhomogeneous medium due to local temperature gradients caused by solar heating. The refractive index of the atmosphere is a function of the local temperature. Temperature differences between the ground and the air produce turbulence cells (or eddies), each with slightly different temperatures. These eddies are effectively individual scatterers producing both refractive and diffractive scattering events. Coherent reciprocal path wave segment pairs exist for double passage (transmitter to target - target to receiver) propagation through atmospheric turbulence[5]. When the target is in the far-field of the turbulence layer, turbulence double passage produces a backscattered signal which is at maximum 3 dB above the diffuse backscattered incoherent intensity, as confirmed by numerical experiments[6].

0-8194-1204-X/93/$6.00

Rough surface and turbulence double passage RPS processes each produce coherent backscatter individually. We hypothesize that in combination, the interrelation of the two processes is multiplicative, with turbulence double passage being the governing scattering process for the relative magnitudes of the coherent and incoherent intensities sensed by the receiver. This is because the scattering processes occur sequentially; and the return passage through turbulence is the last scattering process that occurs during the round trip propagation from the transmitter to the receiver.

Glossary

λ = Laser wavelength = 1.064 μm
P_x = Laser output power
ε = Transceiver optical efficiency
D_r = Monostatic transceiver aperture diameter = 10 cm
A_r = Aperture area
θ_t = Transmitted Gaussian beam divergence half angle
T = One-way average atmospheric transmissivity
P_r = Power received by detector
FOV = Receiver field of view is assumed to be defined by θ_t
R = Space target altitude = 510 km
A_t = Target area
ρ = Target reflectivity
Ω_b = Target backscatter projected solid angle

Neglecting turbulence:

A_x = Transmit beam area at target
A_b = Backscatter area at receiver

Including turbulence:

R_1 = Phase screen altitude = 10 km
r_o = Atmospheric coherence diameter = 10 cm
θ_{ps} = Turbulence forward scatter angle
A_{ps} = Transmit beam area at phase screen
A_{xt} = Transmit beam area at target
A_{bps} = Backscatter area at phase screen
A_{br} = Effective aperture accepted backscatter area at receiver

Including rough target RPS:

θ_c = Coherent backscatter full angle beamwidth
θ_d = Diffuse backscatter full angle beamwidth
A_{bc} = Coherent backscatter area at receiver
A_{bd} = Diffuse backscatter area at receiver
A_{bcps} = Coherent backscatter area at phase screen
A_{bdps} = Diffuse backscatter area at phase screen
F_{sc} = Rough surface coherent backscattered power factor
F_{sd} = Rough surface diffuse backscattered power factor

Including turbulence DP-RPS

A_{bdp} = Backscattered area over which reciprocal paths exist = A_r
F_{tc} = Double passage coherent backscattered power factor
F_{td} = Double passage diffuse backscattered power factor

3. Traditional Laser Radar Equation

The traditional formulation of the ladar range equation (LRE) for the power received P_r from an unresolved target[7] is

$$P_r = P_x \frac{A_t}{R^2 \Omega_x} \varepsilon\rho \frac{A_r}{R^2 \Omega_b} T^2 \tag{1}$$

where Ω_x is the transmit beam projected solid angle and the other relevant quantities are defined in the Glossary. The laser radar range equation assumes that the transmitted and backscattered fields propagate into solid angles with constant transverse cross-sectional power density. The range equation assumes that the target is in the farfield of the transmitted beam and the receiver aperture is smaller than the area of the backscatter at the receiver.

Meaningful insight is obtained by rewriting the ladar range equation in terms of solid angle ratios which can then be converted into area ratios

$$P_r = P_x \varepsilon\rho T^2 \frac{\Omega_t}{\Omega_x} \frac{\Omega_r}{\Omega_b} = P_x \varepsilon\rho T^2 \frac{A_t}{A_x} \frac{A_r}{A_b} \tag{2}$$

where A_x is the cross sectional area of the transmitted beam at the space target and A_b is the projected area of the target backscatter at the receiver. Thus for an unresolved target, the received power depends upon the ratio of the illuminated target area to the projected transmitted beam area at the target, as well as the ratio of the receiver aperture area to the projected target backscatter area. Neglecting turbulence, the transmitted Gaussian beam divergence half angle is aperture limited by the relationship $\theta_t = 2\lambda/(\pi D_r)$. Figure 1 illustrates the geometry for the LRE.

4. Theoretical Development

In this section the modifications to the LRE will be presented. First the LRE will be modified to include atmospheric turbulence. Expressions will then be developed to include the effects of turbulence double passage RPS (DP-RPS), rough target, valley wave RPS (VW-RPS), and rough target, surface wave RPS (SW-RPS). Finally, the reciprocal path modified LRE (RPMLRE) expression is presented which is defined as including the combined effects of rough target RPS and turbulence DP-RPS. Since reciprocal scattering paths exist only within the transmit beam, the assumption will be made that the detector field of view (FOV) is matched to the transmitted beam for the theoretical development to follow.

4.1. Atmospheric Turbulence Model

Atmospheric turbulence is worst near the ground but occurs for altitudes ranging to 100 km. For ground-based operation against space targets, turbulence is a fraction of the total one-way propagation distance from the transmitter to the target. The remainder of the propagation path is considered to be free space.

It is well known that atmospheric refractive index fluctuations due to an extended continuous inhomogeneous medium are sufficiently modelled by a thin, transparent, random phase screen[8], which for this scenario is located beyond the Rayleigh range of the transmitted beam yet close to the transceiver with respect to the target. We assume the transceiver aperture is atmosphere limited, with a diameter defined by

the atmospheric coherence diameter[9] r_0. Taking $r_0 = 10$ cm as a reasonable estimate for zenith viewing at night, and a transmit aperture three times the beam waist ($\omega_0 = 3.33$ mm), the Rayleigh range of the transmitted beam is $z_R = \pi\omega_0^2/\lambda = 3.2$ km. A phase screen altitude of 10 km is thus appropriate under these conditions. For targets located within the lower atmosphere, a phase screen model may not be sufficient, and a turbulence profile may be required.

In terms of the LRE, the phase screen forward scatter reduces the received power by expanding the transmitted and backscattered solid angles. The transmitted beam is incident upon the phase screen and forward scattered within a cone of propagation directions defined by the full angle θ_{ps}. With the receiver FOV matched to the transmit beam, the forward illuminated area of the phase screen A_{ps} can be modelled as an effective projected receiving aperture, as illustrated by Figure 2. The target backscatter incident upon the phase screen is again forward scattered by θ_{ps} towards the receiver aperture. Only the portion of the target backscatter incident upon A_{ps} can enter the FOV of the receiver. The target backscatter incident upon A_{ps} is forward scattered by the phase screen into an area A_{br} at the receiver, as shown. Under these assumptions, the LRE including turbulence, yet neglecting RPS effects is

$$P_r = P_x \varepsilon \rho T^2 \frac{A_t}{A_{xt}} \frac{A_{ps}}{A_{bpsd}} \frac{A_r}{A_{br}} \tag{4}$$

where A_{bpsd} is the diffuse backscatter projected area at the phase screen.

4.2. Reciprocal Path Scattering Model

To model the effects of coherent RPS, the LRE assumption that constant power density is diffusely backscattered into a solid angle is expanded to include both diffuse and coherent scattered power components, each with constant power density across its respective solid angle. Conservation of energy is preserved by the inclusion of diffuse and coherent scattered power ratio factors, which sum to unity. Double passage RPS will first be modelled, followed by rough surface RPS. The RPMLRE model for the combination of rough surface RPS with DP-RPS will then be presented.

At this point a discussion pertaining to the enhancement factor, as referred to in the literature, is necessary for clarity. The enhancement factor f is defined to be the ratio of the total backscattered intensity to the diffuse backscattered intensity

$$f = \frac{I_{coh} + I_{diff}}{I_{diff}} \tag{5}$$

For example, an enhancement factor of five implies a coherent backscattered intensity four times the diffuse backscattered intensity.

4.2.1. Turbulence Double Passage RPS

The reciprocal path requirement of DP-RPS implies that backscattered wave segments exist with cyclic partners only within the detector FOV defined by the transmit beam. The receiver aperture thus defines the maximum area over which DP-RPS is possible. With the receiver FOV matched to the transmit beam, the projection of the receiver aperture at the phase screen can be modelled as the effective reciprocal path reception aperture, as shown in Figure 3. DP-RPS reciprocal paths require the backscattered area over which cyclic paths exist A_{bdp} to be less than or equal to the receiver aperture area.

The LRE including turbulence with DP-RPS is given by

$$P_r = P_x \varepsilon \rho T^2 \frac{A_t}{A_{xt}} \frac{A_{ps}}{A_{bpsd}} \left(F_{td} \frac{A_r}{A_{br}} + F_{tc} \frac{A_{bdp}}{A_r} \right) \tag{6}$$

where F_{td} and F_{tc} are the diffuse (non-cyclic) and coherent (cyclic) turbulence double passage scattered power factors, respectively, which satisfy the restriction $F_{td} + F_{tc} = 1$

$$F_{td} = \frac{A_{br}}{A_{br} + A_r (f_t - 1)} \quad \text{and} \quad F_{tc} = \frac{A_r (f_t - 1)}{A_{br} + A_r (f_t - 1)} \tag{7}$$

where f_t is the turbulence double passage enhancement factor. For example, when the far-field turbulence enhancement factor is maximum $f_t = 2$, the diffuse and coherent forward scatter intensity strengths are equal.

4.2.2. Rough Surface RPS

Coherent RPS from rough surfaces occurs due to resonant, time-reversed multiple scattering events. Several distinct phenomena produce coherent backscatter from rough surfaces[10]; and the simplest of these will briefly be discussed. For moderate and very rough surfaces cyclic multiple scattering within the valleys of the surface irregularities is the predominant RPS mechanism. This RPS mechanism will be referred to as "valley wave" VW-RPS. For slightly rough surfaces, however, the slopes are too small for significant multiple scattering within the valleys. In this case, reciprocal surface wave multiple scattering events (resonant polariton localization) produce coherent RPS from conductive targets. This RPS mechanism will similarily be referred to as "surface wave" SW-RPS. Although not addressed in this analysis, multiply scattered bulk volume waves are similarly responsible for RPS from slightly rough dielectric surfaces.

In contrast to the diffuse backscatter, the coherent backscatter propagates within a narrow cone of propagation directions that is centered about the retro direction. The strength and angular width of the coherent backscatter is a function of both the surface roughness and complex permittivity. For VW-RPS, the angular width of the coherent backscatter is approximated by the diffraction-limited divergence of an effective aperture whose diameter is proportional to the mean free path between the multiple scattering events. In this case, the mean free path is nominally a wavelength or less; and the full angular width of the coherent VW-RPS peak is typically several degrees. The angular width of the SW-RPS peak depends upon the polariton localization length, which can be several times the mean free path[11]. The full angular extent of the SW-RPS peak is typically a degree, or less.

For the LRE, in addition to diffuse backscatter with area A_{bd}, the RPS from a rough target surface generates coherent backscatter at the receiver with projected area A_{bc}, as illustrated by Figure 4. Incorporating the coherent rough surface RPS, the traditional LRE (Eq. 2) becomes

$$P_r = P_x \varepsilon \rho T^2 \frac{A_t}{A_x} \left(F_{sd} \frac{A_r}{A_{bd}} + F_{sc} \frac{A_r}{A_{bc}} \right) \tag{8}$$

where the rough surface diffuse and coherent power backscatter factors F_{sd} and F_{sc} sum to unity, and are

given by

$$F_{sd} = \frac{A_{bd}}{A_{bd} + A_{bc}(f_s - 1)} \quad \text{and} \quad F_{sc} = \frac{A_{bc}(f_s - 1)}{A_{bd} + A_{bc}(f_s - 1)} \tag{9}$$

where f_s is the rough surface enhancement factor.

The rough surface enhancement factor varies substantially. Measured values typically range from slightly greater than unity to two, although factors of ten are not uncommon. A combination of several coherent backscatter mechanisms is believed to be responsible for the enhancement factor of 80 observed from a Teflon-coated polished silver grating[4].

4.3. Combined Turbulence Double Passage RPS and Rough Surface RPS

The effect of combining the two turbulence and rough surface RPS phenomena is multiplicative by consideration of the combined reciprocal scattering possibilities. Combining Eqs. 6 and 8, the Reciprocal Path Modified Ladar Range Equation (RPMLRE) is determined to be

$$P_r = P_x \varepsilon \rho T^2 \frac{A_t}{A_{xt}} \left(F_{td} \frac{A_r}{A_{br}} + F_{tc} \frac{A_{bdp}}{A_r} \right) \left(F_{sd} \frac{A_{ps}}{A_{bpsd}} + F_{sc} \frac{A_{ps}}{A_{bpsc}} \right) \tag{10}$$

Maximum reciprocal path possibilities occur when $A_{ps} = A_{bdp}$ as shown by Figure 5.

Four terms result from the multiplication; and they determine the relative strengths of the diffuse and coherent backscattered power received. The term containing the product $F_{td}F_{sd}$ corresponds to diffuse target backscatter which is diffusely forward scattered by the turbulence as illustrated by Figure 6. Figure 7 shows the $F_{tc}F_{sd}$ term representing the diffusely backscattered wave segments from the target which exist with cyclic partners due to turbulence double passage, and thus contribute to the total coherent power. The coherent target backscatter which is incoherently forward scattered by turbulence is described by the $F_{td}F_{sc}$ term as demonstrated in Figure 8. Finally, $F_{tc}F_{sc}$ represents the coherent target backscattered coherently forward scattered through the turbulence, which is illustrated by Figure 9.

The diffuse power received is the sum of the two terms which are multiplied by F_{td}; and the coherent received power is the sum of the two terms which are multiplied by F_{tc}. The turbulence double passage phenomena is thus the governing RPS process. Double passage through turbulence either preserves or incoherently disturbs the individual cyclic wave segment pairs which are coherently backscattered from the target. Similarly, double passage either preserves the incoherence of diffuse target backscatter, or combines individual diffuse target backscattered wave segments into coherent cyclic pairs.

5. Summary

Within the constraints of this analysis, rough surface RPS in combination with turbulence double passage RPS increases the predicted received power of a monostatic system over that predicted by the traditional LRE. The theoretical formulation is constructed such that the relative magnitudes of the coherent and diffuse received power components are easily separated for RPS exploitation purposes.

6. Future Work

A follow-on paper is in process, in which the received power from a rough target and a mirror (glint) target will be compared using the various LRE formulations presented in this paper. A simple method will be presented to approximate the turbulence and rough surface RPS power scatter factors from measured data.

7. Acknowledgements

The authors would like to thank Dr. C. M. Stickley (CREOL) and Drs. J. C. Jafolla and Z. Gu (Surface Optics Corporation) for valuable technical discussions. This research has been supported by the Strategic Defense Initiative Organization, Innovative Science and Technology Office through the Office of Naval Research under contract N00014-92-C-0123, and by the Advanced Research Project Agency, through the Office of Naval Research under contract N00014-92-J-1906.

8. References

1. T. Mavroidis, C. J. Solomon, and J. C. Dainty, "Imaging a coherently illuminated object after double passage through a random screen," *J. Opt. Soc. Am. A*, Vol. 8, No. 7, pp. 1003-1013, July 1991.
2. E. R. Mendez and K. A. O'Donnell, "Observation of depolarization and backscattering enhancement in light scattering from Gaussian random surfaces," *Opt. Comm.*, Vol. 61, No. 2, pp. 91-95, Jan. 1987.
3. V. Celli, A. A. Maradudin, A. M. Marvin, A. R. McGurn, "Some aspects of light scattering from a randomly rough metal surface," *J. Opt. Soc. Am. A*, Vol. 2, No. 12, pp. 2225-2239, Dec. 1985.
4. Technical conversations with Drs. J. C. Jafolla and Z. Gu, Surface Optics Corporation.
5. V. A. Banakh and V. L. Mironov, Lidar in a Turbulent Atmosphere, pp. 25-56, V. E. Zuev, ed., Artech House, Boston, 1987.
6. G. Welsh and R. P. Phillips, "Simulation of enhanced backscatter by a phase screen," *J. Opt. Soc. Am. A*, Vol. 7, No. 4, pp. 578-584, April 1990.
7. C. G. Bachman, Laser Radar Systems and Techniques, pp. 9-11, Artech House, Dedham, 1979.
8. H. G. Booker, J. A. Fergusen, and H. O. Vats, "Comparison between the extended medium and the phase-screen scintillation theories," *J. At. and Terr. Phys.*, Vol. 47, No. 4, pp. 381-399, 1985.
9. D. L. Fried, "Optical heterodyne detection of an atmospherically distorted signal wave front," *IEEE Proc.*, Vol. 55, No. 1, pp. 57-67, Jan. 1967.
10. J. C. Jafolla, "The phenomenology and measurements of retroreflectance," *Surface Optics Corporation SOC-R677-001-0392,* prepared for Department of the Air Force, March 1992.
11. A. R. McGurn and A. A. Maradudin, "Localization effects in the elastic scattering of light from a randomly roughl surface," *J. Opt. Soc. Am. B*, Vol. 4, No. 6, pp. 910-926, June 1987.

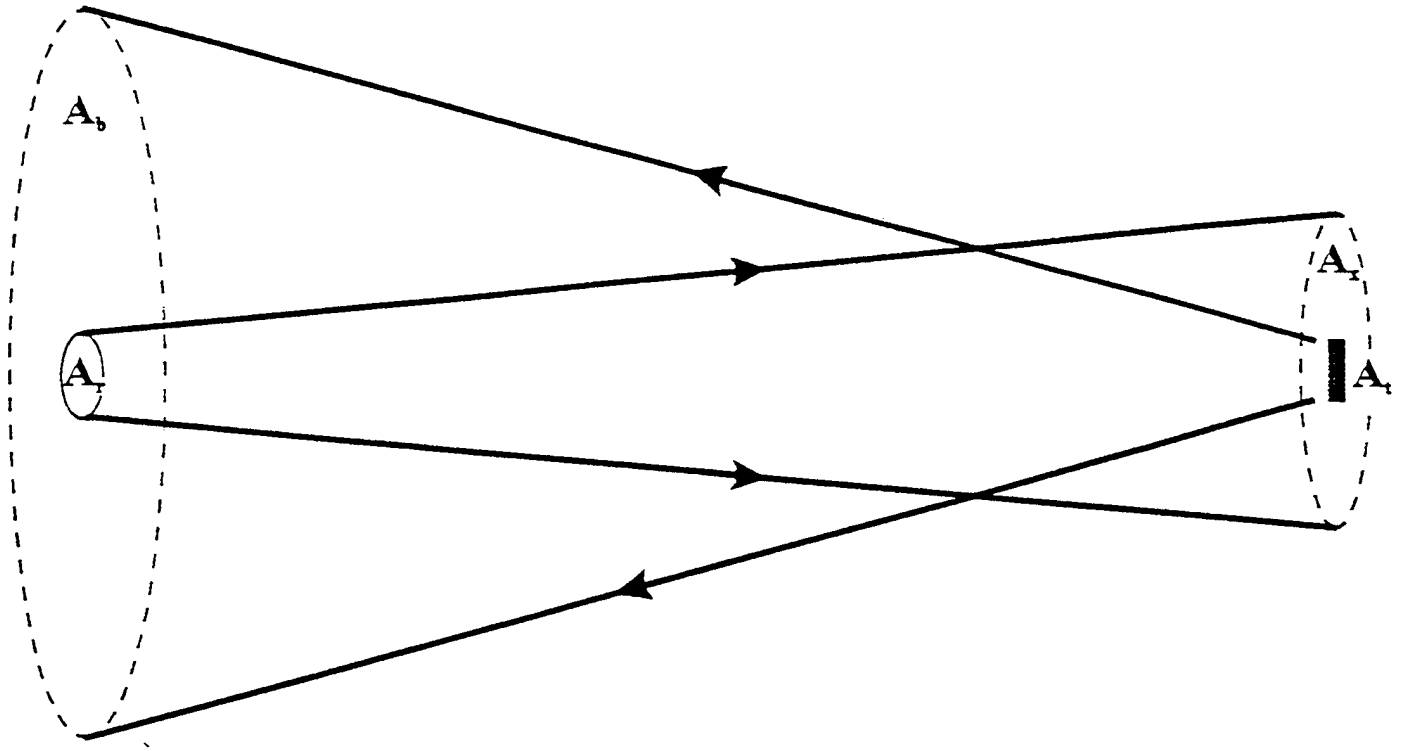


Figure 1. Traditional ladar range equation (LRE) for an unresolved target.

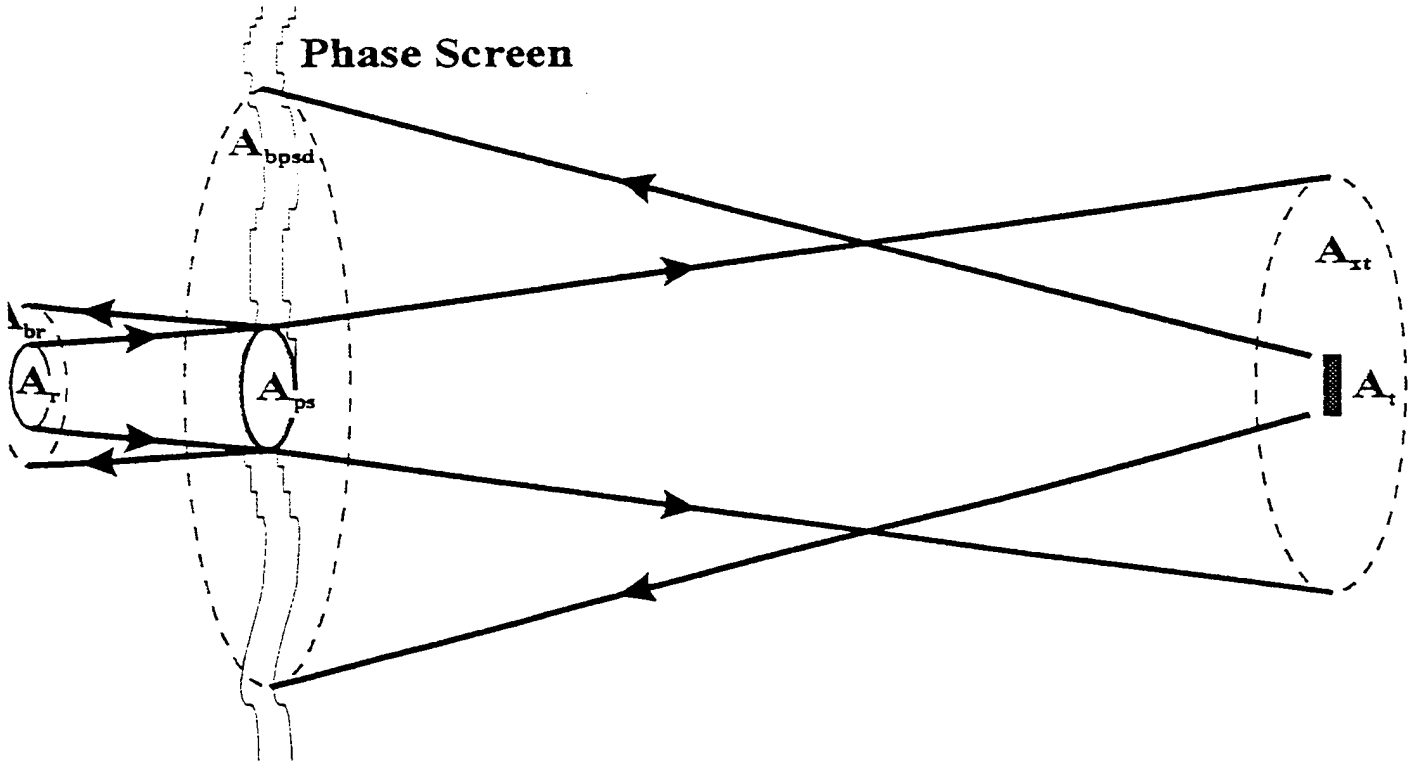


Figure 2. LRE including thin random phase screen model for atmospheric turbulence, assuming receiver FOV is defined by transmit beam.

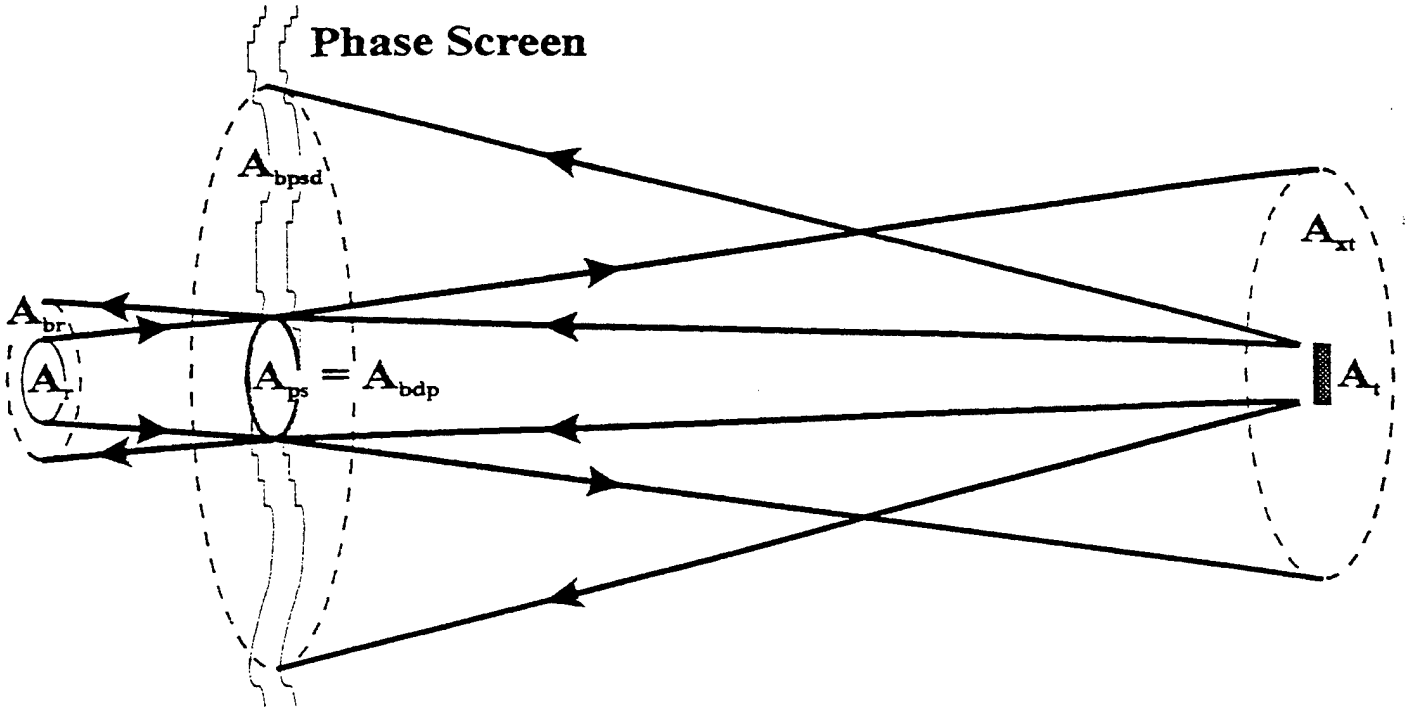


Figure 3. LRE including turbulence double passage reciprocal path scattering DP-RPS, where $F_{td} + F_{tc} = 1$. We assume $A_{bdp} = A_{ps}$.

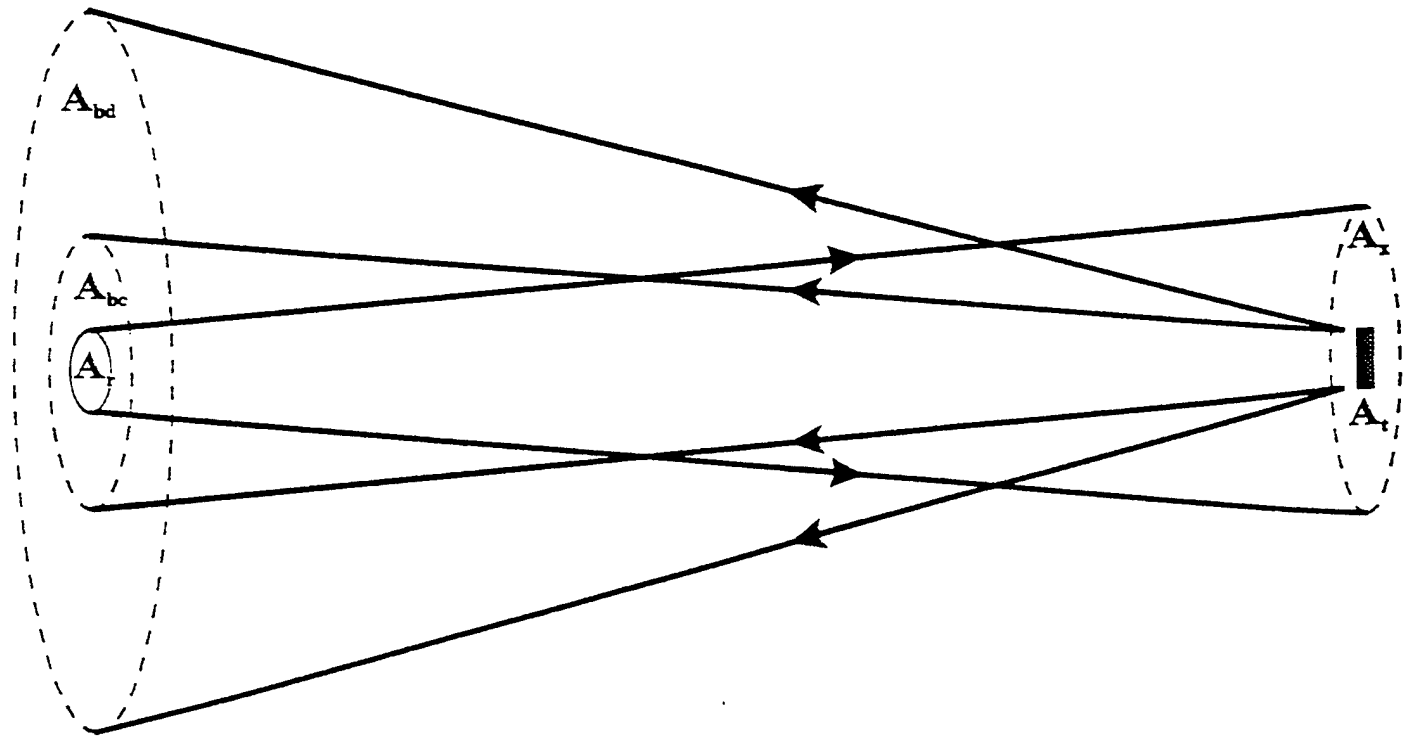


Figure 4. LRE including target rough surface RPS, where $F_{sd} + F_{sc} = 1$.

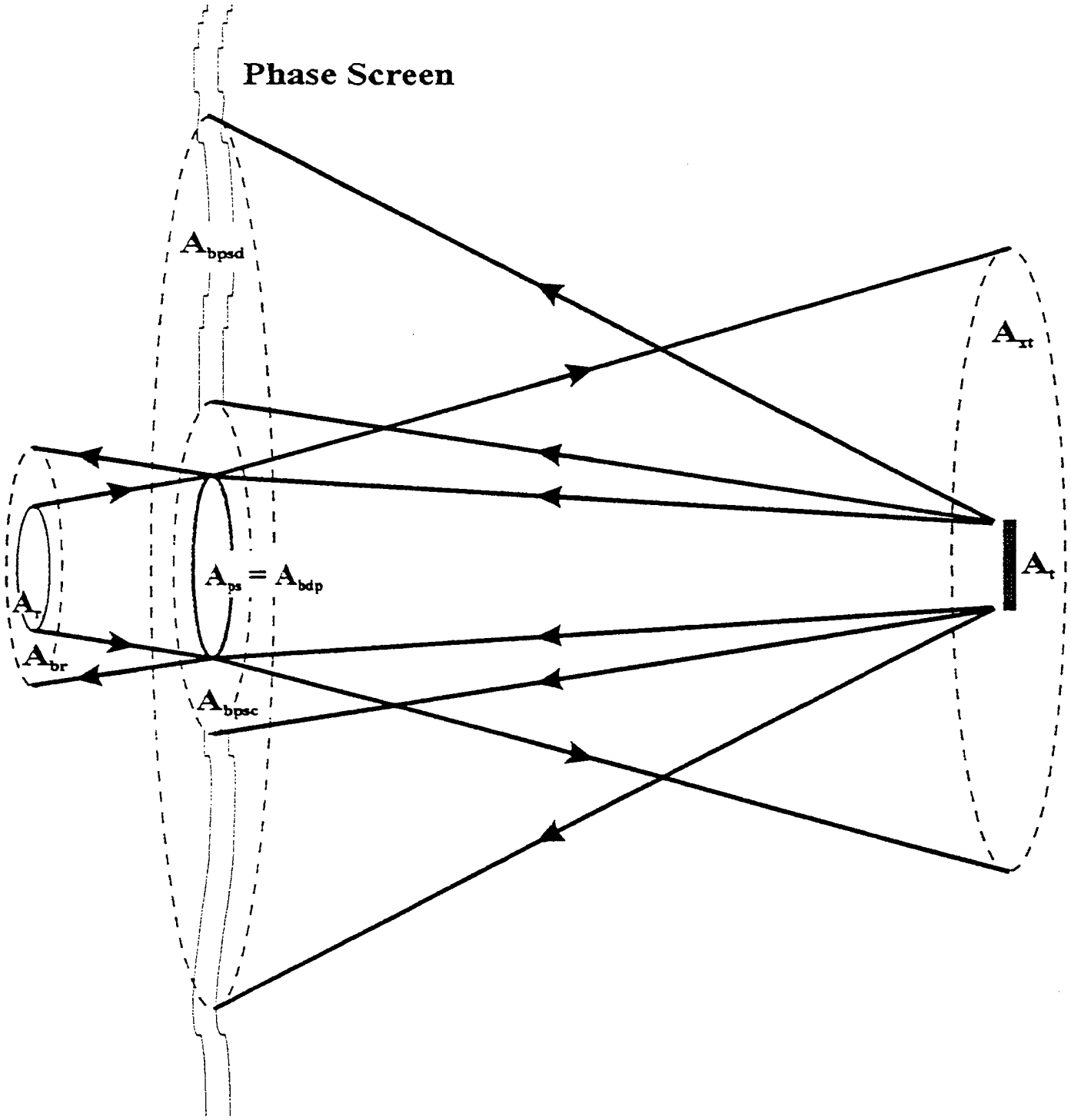


Figure 5. Reciprocal Path Modified Ladar Range Equation (RPMLRE) including combined rough surface and turbulence double passage RPS.

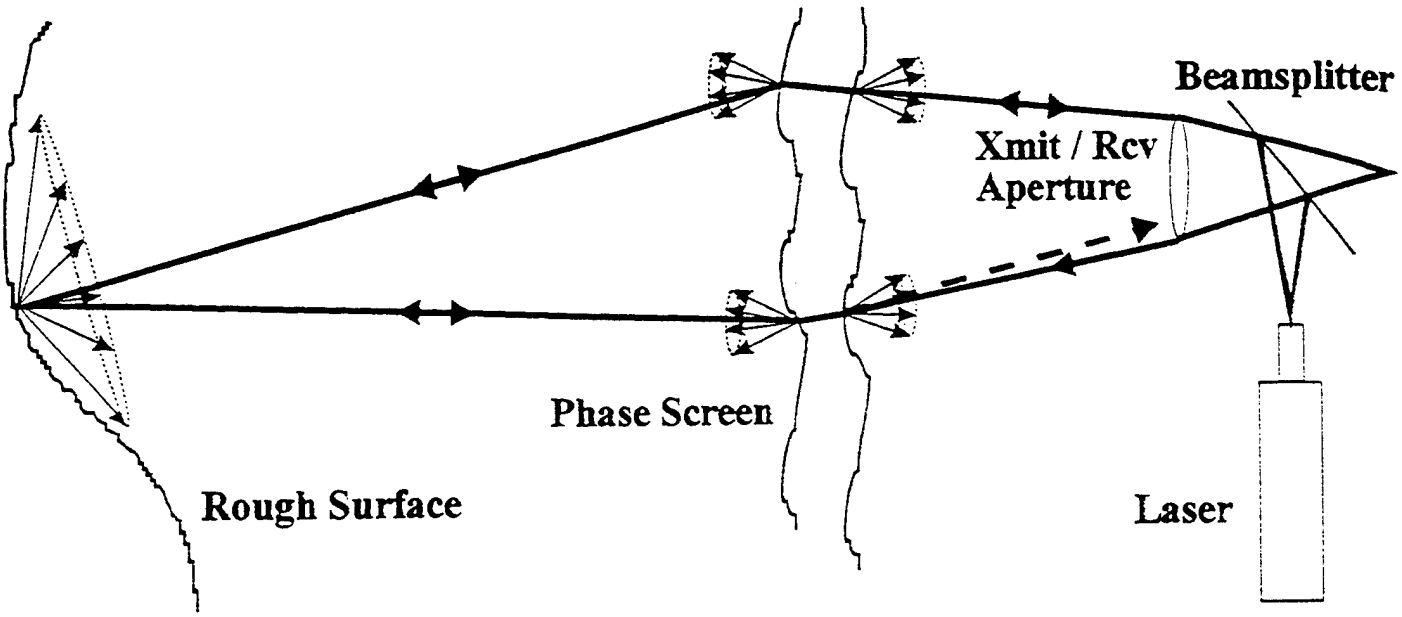


Figure 6. $F_{td}F_{sd}$ term represents diffuse target backscatter which is incoherently forward scattered through the turbulence. Diffuse power received.

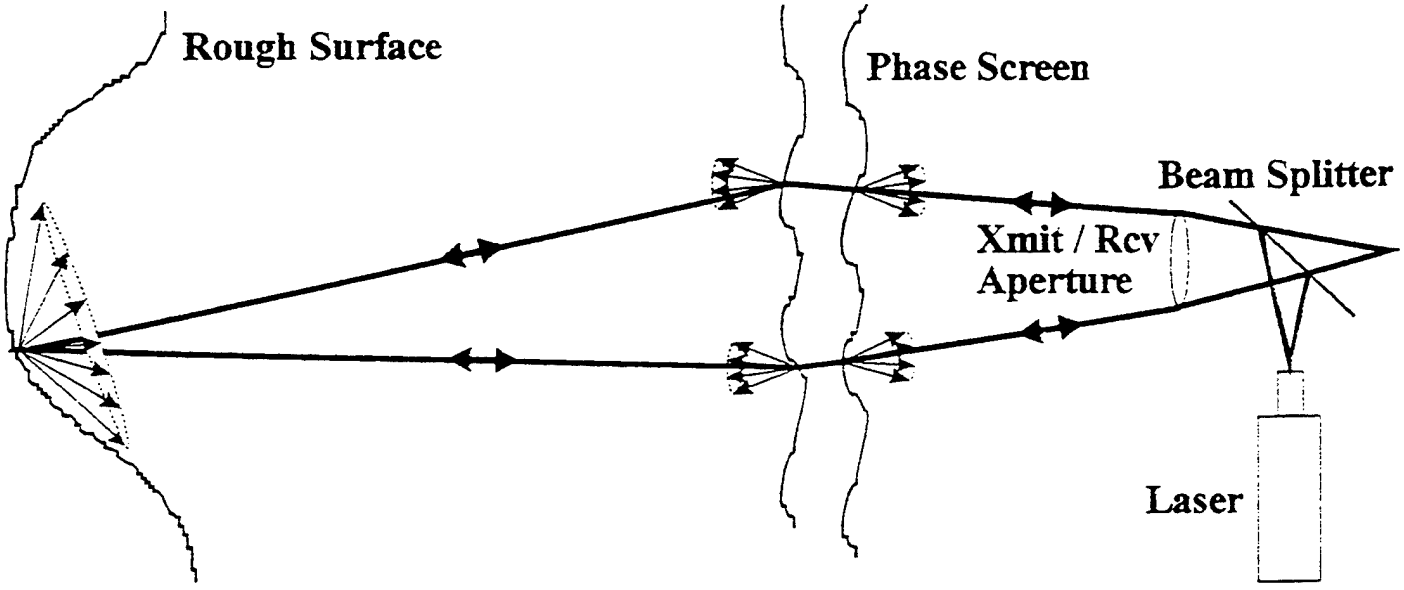


Figure 7. $F_{tc}F_{sd}$ term represents diffuse target backscatter which exist with reciprocal path partners due to turbulence double passage. Coherent power received.

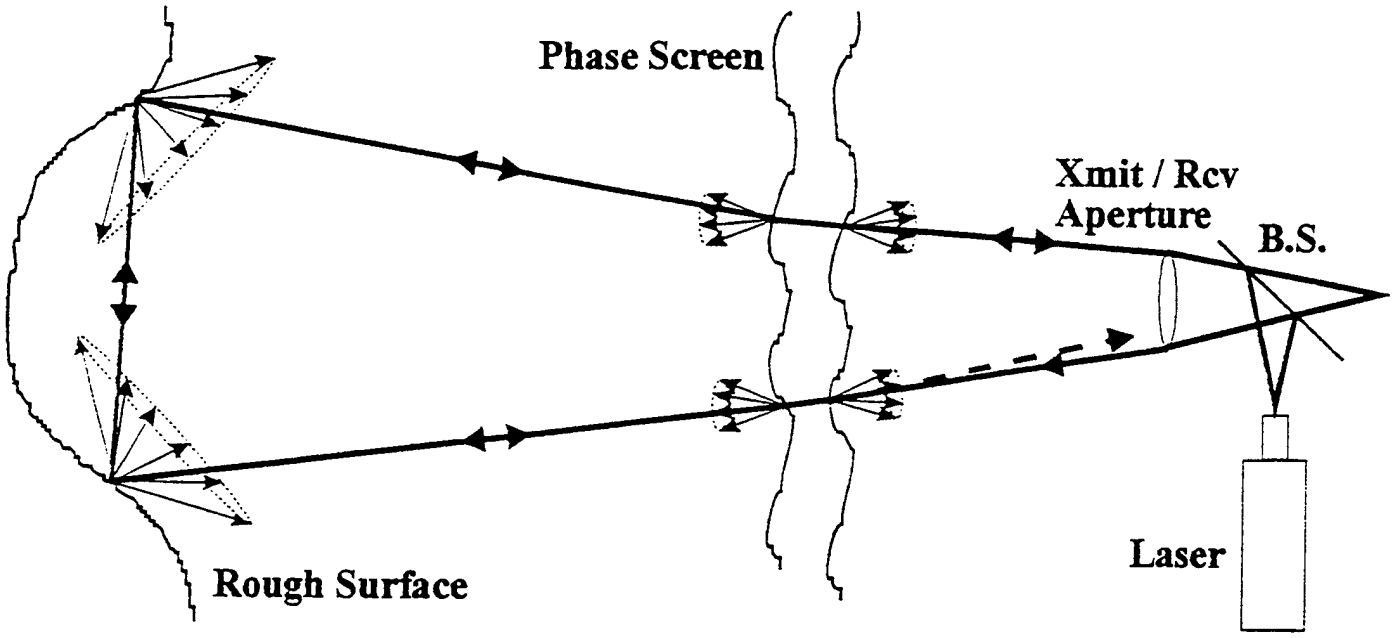


Figure 8. $F_{td}F_{sc}$ term represents coherent target backscattered wave segments whose coherence with cyclic partners is lost due to temporal atmospheric evolution.
Diffuse power received.

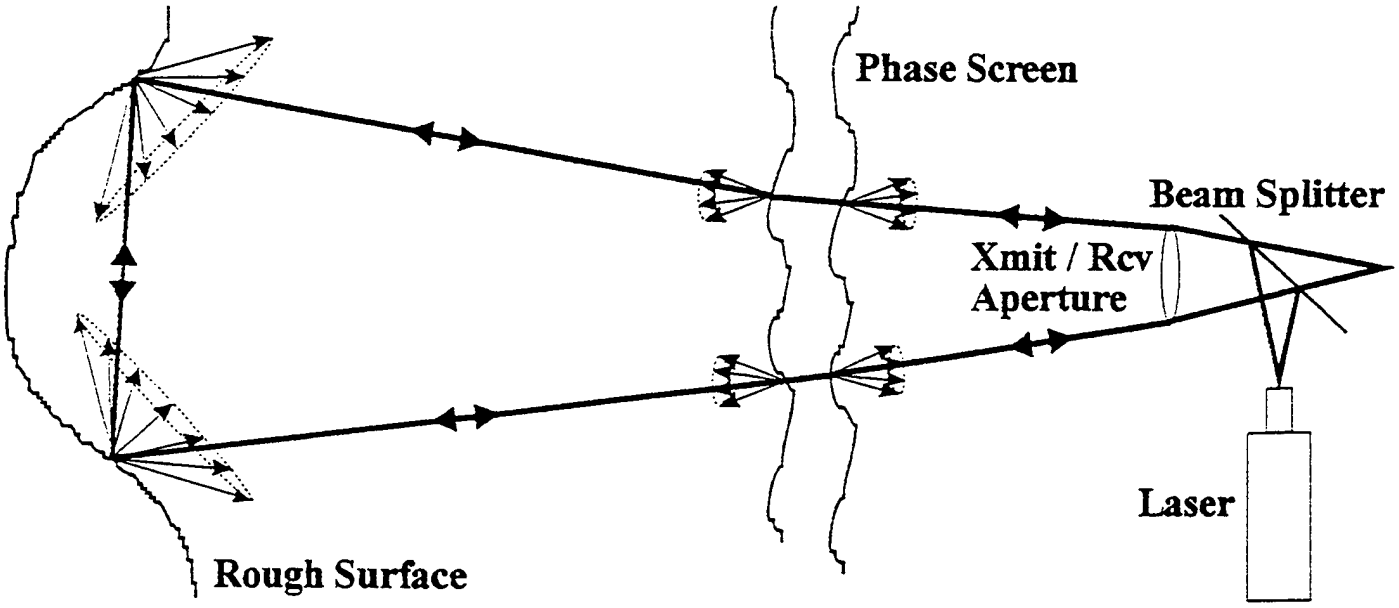


Figure 9. $F_{tc}F_{sc}$ term represents coherent target backscatter which is coherently forward scattered through the turbulence.
Coherent power received.

8.7 Appendix 8B

R. A. Murphy and R. L. Phillips, "Atmospheric Turbulence Reciprocal Path Scattering Effects for a Ground-Based Monostatic Laser Radar Tracking a Space Target", Proc. SPIE 2828, 52, (1996).

Atmospheric Turbulence Reciprocal Path Scattering Effects for a Ground-based Monostatic Laser Radar Tracking a Space Target

R. A. Murphy
Southwest Research Institute
613 Russell Parkway, Warner Robins, GA 31088

and

R. L. Phillips
Center for Research and Education in Optics and Lasers
University of Central Florida, Orlando, FL 32816

ABSTRACT

A phenomenological model is developed for the strength and spatial width of the backscattered coherent intensity peak produced by reciprocal path scattering through atmospheric turbulence. The model is applied to a ground-based monostatic laser radar tracking a space target under optically saturated atmospheric turbulence conditions. The models for the amplitude and width of the RPS peak are based upon the spatial coherence widths of the propagating fields over the up-link and down-link paths within the atmospheric turbulence as well as the cross-sectional area of the up-link beam.

Keywords: atmospheric turbulence, reciprocal path scattering, enhanced backscattering, and backscatter amplification.

1. INTRODUCTION

A ground-based monostatic laser radar transmits a Gaussian beam through atmospheric turbulence for the initial portion of the up-link propagation path to a space-borne target. The scattered fields from the target then propagate back through the same turbulence for the last segment of the down-link path to the transceiver. Propagation through turbulence degrades the spatial coherence of the initially coherent beam and causes it to spread more than that due to diffraction[1-3]. Turbulence scattering also causes the intensity to fluctuate erratically[4-10] at the transceiver and induces fluctuations in the received angle of arrival information for target position; this distorts an object's image in an optical system. Atmospheric turbulence reciprocal path scattering (RPS), however, produces a spatially coherent average intensity component at the transceiver that may be exploited by reciprocal path imaging (RPI) applications[11-13] to partially mitigate the adverse effects of scattering by turbulence.

The atmospheric turbulence scattering theory necessary for the development of a round-trip propagation model is first briefly reviewed. The phenomenological model for round-trip propagation through saturated turbulence including RPS is then derived and its performance evaluated under a variety of propagation environments. An example of the calculations required to evaluate the performance of this model is also included.

2. ATMOSPHERIC TURBULENCE SCATTERING THEORY

The scintillation index σ_I^2 (the normalized variance of the intensity) provides a measure of the turbulence strength[1-10], the effects of which are commonly classified as $\sigma_I^2 < 1$, $\sigma_I^2 \sim 1$, and $\sigma_I^2 > 1$, representing weak, strong, and saturated turbulence conditions, respectively. A theoretical measure of the strength of turbulence scattering is the Rytov variance[1,5,14], σ_R^2. For a plane wave with wavenumber k along a horizontal propagation path (uniform structure constant C_n^2) with length L, σ_R^2 is

$$\sigma_R^2 = 1.23\, C_n^2\, k^{7/6}\, L^{11/6} \tag{1}$$

As the scattering increases, fluctuations in the scattered intensity increase; the conditions of saturated turbulence ($\sigma_I^2 > 1$) are the subject of this analysis.

0-8194-2216-9/96/$6.00

In addition to intensity scintillation, turbulence inhomogeneities produce other undesirable effects upon a propagating laser beam[1-8]. Forward scattering of a beam propagating through turbulence inhomogeneities causes it to spread beyond the free space diffraction-limited divergence[1-3]. This spreading is important to characterize for power budget calculations since up-link beam spreading reduces the power density of the beam incident upon the target. Beam spreading may be severe for the up-link propagation path through saturated turbulence; however, turbulence induced spreading is usually negligible relative to the target scattering for the down-link echo path. Propagation through turbulence also produces a reduction in spatial and temporal coherence, beam wander, and beam break-up into speckle. Beam wander is caused by propagation through large turbulence cells, which causes the spatial location of the beam to vary at the target. Beam break-up into speckle is the result of propagating through refractive inhomogeneities that are smaller than the beam.

The formulation of the expressions that characterize the beam propagation through turbulence is facilitated by first considering the free space propagation parameters[1,8,9,15] for a diffraction-limited up-link beam. The diffraction-limited Gaussian beam transmitter parameters are given by

$$\Omega_o = 1 - \frac{L}{R_o} \qquad \text{and} \qquad \Omega = \frac{2L}{k\,\omega_o^2} \tag{2}$$

where ω_o and R_o are the transmitted beam radius and radius of curvature, respectively. For a collimated Gaussian beam with $R_o = \infty$, we have $\Omega_o = 1$. Here L is the vertical or slant path length from the transceiver at an altitude h_o to the target at an altitude H at the zenith angle ξ_{sp}

$$L = \frac{H - h_o}{\cos(\xi_{sp})} \tag{3}$$

The diffraction-limited Gaussian beam characteristics after up-link propagation through free space are then conveniently expressed in terms of the target plane beam parameters

$$W = \omega_o \sqrt{\Omega_o^2 + \Omega^2} \quad , \quad \Lambda = \frac{\Omega}{\Omega_o^2 + \Omega^2} = \frac{2L}{k W^2} \quad , \text{and} \quad \Theta = \frac{\Omega_o}{\Omega_o^2 + \Omega^2} \tag{4}$$

where W is the diffraction-limited beam radius at the target plane.

For Gaussian beam propagation through atmospheric turbulence, an effective propagating beam may be defined that accounts for the beam spreading and beam wander produced by turbulence[1-3]. The long term average effective beam spot size (beam radius), for a beam with a spot size W in free space, including beam spread and beam wander for the up-link propagation path through turbulence, W_l, is derived from the average on-axis intensity and is given by[1,9,16]

$$W_l = W \sqrt{1 + 6.72\, k^{7/6}\, \Lambda\, m_0^{6/5} \left[(H - h_o) \sec(\xi_{sp}) \right]^{11/5}} \qquad \text{where} \qquad m_0 = \int_{h_o}^{H} C_n^2(h)\, dh \tag{5}$$

Here we integrate the structure constant C_n^2 as a function of the altitude, h. For modeling the propagation of Gaussian beams upward and downward through the atmosphere, we use the Hufnagel-Valley day model for the dependence of C_n^2 upon altitude[17]

$$C_n^2(h) = 0.00594 \left(\frac{ws}{27} \right)^2 (10^{-5} h)^{10} \exp\left(\frac{-h}{1000} \right) + 2.7 \times 10^{-16} \exp\left(\frac{-h}{1500} \right) + C_{n,o}^2 \exp\left(\frac{-h}{100} \right) \tag{6}$$

where h is in meters (m), ws is a high altitude wind speed (m/s), and $C_{n,o}^2$ is the ground level structure constant in $m^{-2/3}$. Also of interest is the short term average beam spot size W_s that includes beam spread but not beam wander, i.e. the physical radius of the effective propagating beam

$$W_s = \sqrt{W_l - 1.96 W^2 \left(\frac{1}{2\omega_o} \right)^{1/3} k\, L\, \Lambda\, m_0} \tag{7}$$

The short-term average up-link Gaussian beam spot size defines the transverse area through which strictly reciprocal round-trip propagation paths may occur.

When an initially coherent laser beam propagates through atmospheric turbulence it suffers a reduction in spatial coherence. The transverse field spatial coherence width, denoted l_{turb}, describes the transverse distance about the center of a beam (symmetric about the optic axis) over which the wavefront remains spatially coherent[1,5,18] during a single passage

through turbulence inhomogeneities. The turbulence field spatial coherence width is defined as the transverse separation distance between two points on a wavefront at which the field mutual coherence function[4] decreases to $1/e$ of its maximum value, and may be determined from the wave structure function. For any fixed path length l_{turb} decreases as the turbulence scattering strength increases.

For a vertical or slant propagation path through the atmosphere, l_{turb} is different over the up-link and down-link propagation paths. This is due to C_n^2 being strongest near the ground and decreasing with altitude up to about 20 km[7,17]. The beam propagating over the up-link path experiences the strongest refractive inhomogeneities in the initial portions of the path to the target, whereas the strongest turbulence refractive inhomogeneities are encountered over the down-link path just prior to arrival at the transceiver. The up-link field transverse spatial coherence width is denoted l_{up} and its value at the target plane denoted l_{tgt}. Similarily the down-link field transverse spatial coherence width is l_{down} and at the transceiver plane l_{tcvr}.

The up-link field spatial coherence width for a path through saturated turbulence with zenith angle ξ_{sp} is given by[1,16]

$$l_{up} = \left[1.46\,k^2\left(m_2 + 0.622\,\Lambda_t^{5/6}\,m_0\right)\right]^{-3/5}\cos^{3/5}\left(\xi_{sp}\right) \qquad \text{where} \qquad m_2 = \left(\frac{1}{H-h_o}\right)^{5/3}\int_{h_o}^{H}\left(h-h_o\right)^{5/3}C_n^2\left(h\right)dh \tag{8}$$

and where m_0 is given by Eq. 6. The turbulence effective beam parameter at the target plane in Eq. 8

$$\Lambda_t = \frac{\Lambda}{1 + 1624\left(\sigma_R^2\right)^{6/5}\Lambda} \tag{9}$$

is formulated in terms of the plane wave Rytov variance[1,14,16] for a slant propagation path through the atmosphere

$$\sigma_R^2 = 2.25\,k^{7/6}\sec^{11/6}\left(\xi_{sp}\right)H^{5/6}\left(\frac{1}{H-h_o}\right)^{5/3}\int_{h_o}^{H}\left(1-\frac{h-h_o}{H-h_o}\right)^{5/6}C_n^2\left(h\right)dh \tag{10}$$

For a target that is far beyond the turbulence layer, the down-link field transverse spatial coherence length of a spherical wave is accurately approximated by the plane wave expression[1,16]

$$l_{down} = \left(1.46\,k^2\,m_0\right)^{-3/5}\cos^{3/5}\left(\xi_{sp}\right) \tag{11}$$

The spatial coherence width models for the up-link and down-link waves as a function of altitude are illustrated in Figure 1.

3. ATMOSPHERIC TURBULENCE RECIPROCAL PATH SCATTERING

When a receiver and transmitter share a common optic axis, field correlation between up-link and down-link propagation paths through different turbulence inhomogeneities produces a pronounced peak in the angular distribution of scattered light in the anti-incident direction that is known as the backscatter enhancement (BE) effect[5,6]. BE is also referred to as backscatter amplification and enhanced backscattering in the scientific literature. Enhanced backscatter produced by atmospheric turbulence was first predicted by Belen'kii[19,20] and experimentally verified by Gurvich and Kashkarov[20,21]. Since then, BE has been the subject of many experiments[5,6,21-24] and publications[25-29].

The strength of the BE produced by round-trip propagation through atmospheric turbulence depends upon the turbulence scattering strength as well as the size and nature of the target[5,6,30]. In general, the enhancement factor, K (the factor above that received from the target for a bistatic receiver) depends upon the path integrated turbulence strength[5,6,31] as conceptually illustrated by Figure 2. In the region of weak scattering, K increases from unity (no enhancement) as the scattering strength increases due to coherent reciprocal path scattering (RPS) events. For strong scattering conditions K reaches a maximum[5,6,25,26] due to incoherent random focusing (IRF), with values of K typically three and up to five and six. For very strong scattering, $\sigma_R^2 >> 1$, the intensity scintillations become saturated and the random focusing becomes negligible such that K asymptotically decreases to two. Although the enhancement factor does have a dependence upon the nature of the target (mirror, rough surface, or corner reflector), here we take the target to be a 1 m^2 Lambertian surface.

Reciprocal path scattering produces a narrow, pronounced coherent peak in the backscatter direction that is superimposed upon the incoherent component of the angular distribution of the average backscattered intensity. For this analysis, we alternatively define the coherent RPS enhancement factor $f_{c,AT}$ to theoretically quantify the coherent enhancement due entirely to RPS (neglecting IRF) to be given by

$$f_{c.AT} = \frac{\langle I_c \rangle + \langle I_{ic} \rangle}{\langle I_{ic} \rangle} \tag{12}$$

The coherent RPS enhancement factor is defined to be the sum of the coherent and incoherent average intensity contributions (neglecting IRF) normalized by the average incoherent intensity contribution. As illustrated by Figure 2, $f_{c.AT}$ increases from unity for weak turbulence scattering to a maximum value of two for saturated turbulence scattering.

The RPS peak is produced by the constructive interference of wave pairs that traverse reciprocal, time-reversed, scattering paths through different turbulence inhomogeneities from the transmitter to the monostatic receiver[5,6], as illustrated by Figure 3. The transverse width of the coherent RPS intensity peak in the transceiver plane is designated $l_{c.AT}$, which is not the same parameter as l_{tcvr} given in Equation 11. The RPS coherence width $l_{c.AT}$ defines the transverse distance in the transceiver pupil plane over which positive intensity correlation exists between the up-link and down-link propagation paths[5,6,18]. For weak turbulence scattering[5], $l_{c.AT.weak}$ is approximately the radius of the first Fresnel zone of a spherical wave propagating from the target to the transceiver $l_{c.AT.weak} \approx \sqrt{\lambda L}$. For saturated turbulence scattering[5,6,18], the coherent RPS width $l_{c.AT}$ is approximately the field transverse spatial coherence width at the transceiver l_{tcvr}, i.e. $l_{c.AT} \approx l_{tcvr}$. In this case, the coherent RPS peak that is produced by the cross-interference of reciprocal path pairs is spatially coherent only over the distance that the individual fields themselves are spatially coherent. The width of the RPS peak is thus narrower for saturated turbulence scattering than for weak scattering conditions.

4. PHENOMENOLOGICAL MODEL FOR RPS THROUGH THE ATMOSPHERE

Based upon the Gaussian beam model and turbulence model in Section 2, we now develop a model for the RPS that is produced by the up-link and down-link propagation through the atmosphere. In our analysis we assume the transmitted beam is collimated. The waist, ω_o, of the collimated transmit beam is chosen to be $1/\pi$ times the transceiver aperture diameter D_x to yield near total (99%) power transmission through the circular aperture[32], i.e. $\omega_o = D_x / \pi$.

In this analysis we assume the target surface to have a Lambertian spatial distribution such that the backscattered field from is entirely due to the superposition of the single scattering events on the surface of the target (no multiple scattering)[33-36]. In this case the average intensity backscattered target propagates into a hemisphere with a cosine angular dependence as measured from the surface normal. We assume the target is in space well beyond the atmosphere hence the scattered waves propagate a large distance through free space and then downward through the atmospheric turbulence. We develop a simple model for general analytical predictions based upon the turbulence propagation phenomenology to predict the amplitude and width of the peak of the angular scattering distribution due to RPS. The amplitude of the RPS peak is formulated in terms of the up-link and down-link transverse spatial coherence widths of the field within the atmosphere, which depend upon the integrated effect of the turbulence on the beam. The spatial width of the RPS intensity peak in the transceiver plane is modeled as the Airy diffraction pattern produced by the down-link wave propagation from an effective circular aperture located at a reference plane (whose location is discussed below) within the atmospheric turbulence.

Since the RPS peak is produced by the cross-interference of components of the field that traverse different up-link and down-link propagation paths through the extended atmospheric turbulence in reciprocal directions[5,6], we initially examine the up-link and down-link transverse spatial coherence widths l_{up} and l_{down}, given by Eqs. 8 and 11to develop a phenomenological model for RPS. After an initial decrease, l_{up} increases as a beam propagates upward through the atmosphere due to diffraction and a decrease in turbulence with altitude. In contrast, the down-link field spatial coherence width, l_{down}, decreases monotonically as the field propagates through the atmosphere. As shown in Figure 1, at some plane normal to the propagation path within the atmosphere, $l_{up}(h) = l_{down}(h) = l_e$. The altitude at which these spatial coherence widths are equal, H_e, is referred to as the location of the equal coherence plane, which is obtained by iterative numerical integration of Eqs. 8 and 11 over the propagation paths. The short term average up-link beam spot size, W_e, at the altitude H_e, which is determined by the evaluation of Eq. 7, defines the physical radius of the up-link beam in the equal coherence plane and the transverse area A_e through which strictly reciprocal propagation paths occur, as illustrated by Figure 4.

We now imagine that the backscattered intensity components arriving at the transceiver plane have propagated through distinct turbulence coherence cells in the equal coherence plane for the up-link and down-link paths. The area A_{cell}, of a coherence cell in this plane is defined by the transverse spatial coherence width, l_e, at the altitude H_e. These coherence cells

are taken to be distributed within the physical cross-sectional area, A_∞, of the up-link beam at the altitude H_∞, as illustrated in Figure 5. Figure 6 illustrates the possible up-link and down-link paths through several coherence cells with area, A_{cell}. As discussed previously, field components that make the round-trip to and from the target through the same coherence cell contribute to the average incoherent scattered intensity. The RPS peak is produced by the cross interference of field components that make the round-trip through different cells in the equal coherence plane but in reciprocal directions.

From Eq. 12, the enhancement factor for this model is given by[5]

$$f_{c,AT} = 2 - \frac{1}{n_{cells}} \tag{13}$$

where n_{cells} is the number of discrete coherence cells within the cross-sectional area of the up-link beam at the equal coherence plane. A simple approximation for n_{cells} is given by $n_{cells} \approx A_\infty / A_{cell}$; however, this neglects the fact that the cells are non-overlapping, as illustrated previously in Figure 5, and thus overestimates n_{cells}. Since the spatial coherence width is maximum when measured symmetrically about the beam optic axis and decreases at the beam edges, the spatial coherence area A_{cell} should decrease near the beam edge[5,16]. The assumption of identical cell areas within the entire beam cross-section in the equal coherence plane thus partially compensates for the non-overlapping cell areas in our approximation.

The diameter of the effective coherent RPS aperture $D_{AT,eff}$, is simply the diameter of the transverse up-link beam area at the equal coherence plane, through which reciprocal paths exist, i.e. $D_{AT,eff} = 2W_\infty$. With this model the predicted transverse width, l_{RPS}, of the coherent RPS peak in the transceiver aperture plane is

$$l_{RPS} = 2.44 \frac{\lambda L_\infty}{D_{AT,eff}} = 1.22 \frac{\lambda L_\infty}{W_\infty} \tag{14}$$

where L_∞ is the propagation path length through the atmosphere to the equal coherence plane, i.e. $L_\infty = H_\infty / \cos(\xi_{sp})$, and ξ_{sp} is the zenith angle.

5. NUMERICAL EXAMPLE OF RPS IN TRACKING A SPACE TARGET THROUGH TURBULENCE

An example for the phenomenological model of RPS for strong scattering conditions is provided in this section. We first present the system and propagation parameters, then the average incoherent intensity backscattered to the transceiver by a Lambertian scattering target is determined. The necessary calculations are then performed to predict the strength and spatial width of the average coherent RPS intensity component in the transceiver plane.

Suppose we have a ground-based monostatic laser radar operating at 1.064 μm and a 1 m^2 space target that is just above the horizon. A collimated beam exits the transceiver aperture at an altitude of h_o = 2 m above the ground and propagates toward a target at 300 km (H = 300 km) at a zenith angle of ξ_{sp} = 86°. The structure constant of the turbulence at ground-level is $C_{n,o}^2 = 1.9 \times 10^{-14}$ $m^{-2/3}$ and the high altitude wind speed at 10 km is ws = 21 m/s. The scintillation index of the intensity fluctuations on the axis of the beam is first verified to be saturated in the target plane. Taking the transceiver aperture to be D_x = 1 cm, the up-link field transverse spatial coherence width in the plane of the target is determined from Eq. 8 to be $l_{trgt} = 1476\,m$ and the down-link field transverse spatial coherence width in the transceiver plane is calculated from Eq. 11 to be $l_{tcvr} = 1.13\,cm$. The physical beam radius (i.e. the short term average spot size which neglects beam wander effects) of the up-link beam in the target plane is from Eq. 7, given by $W_s = 493.672\,m$.

The spatially incoherent component of the average backscattered intensity $\langle I_{ic}(\theta_s) \rangle$ produced by the self-interference of the fields incident at the transceiver plane is assumed to have the cosine angular distribution produced by a Lambertian scattering target[33-36] with area A_{trgt} in a homogeneous medium. The fields backscattered from the Lambertian target that propagate through the same up-link and down-link cell in the equal coherence plane produce an incoherent average intensity contribution in the transceiver pupil plane that is expressed as

$$\left\langle I_{ic,SC}(\theta_s) \right\rangle = \frac{n_{cells} A_{trgt}}{L^2} I_t \cos(\theta_s) \tag{15}$$

where the intensity of the beam incident upon the target I_t is a function of W_s. Similarily, the average incoherent intensity contribution produced by the self-interference of the scattered fields that propagate through different up-link and down-link coherence cells in the equal coherence plane is expressed as

$$\left\langle I_{ic.DC}\left(\theta_s\right)\right\rangle = \frac{n_{cells}\left(n_{cells}-1\right)A_{trgt}}{L^2} I_t \cos\left(\theta_s\right) \tag{16}$$

The average incoherent intensity component in the aperture plane of the transceiver is the sum of these contributions given by Eqs. 15 and 16

$$\left\langle I_{ic}\left(\theta_s\right)\right\rangle = \left\langle I_{ic.SC}\left(\theta_s\right)\right\rangle + \left\langle I_{ic.DC}\left(\theta_s\right)\right\rangle = \frac{n_{cells}^2 A_{trgt}}{L^2} I_t \cos\left(\theta_s\right) \tag{17}$$

To derive an expression for the average coherent RPS intensity component, a series of calculations must be performed to predict the strength and transverse width of the coherent RPS peak in the transceiver's aperture plane. We first calculate the altitude of the equal coherence plane. By successive numerical iteration of Eqs. 8 and 11, this altitude is determined to be $H_o = 10.49\ km$ and the field spatial coherence width in the equal coherence plane is $l_{up_o} = l_{down_o} = l_o = 9.57\ cm$. Since the strength of the coherent RPS peak relative to the incoherent background depends upon the number of cells in the equal coherence plane and the spatial coherence width is $l_o = 9.57$ cm, the area of the coherence cells is $A_{cell} = \pi l_o^2 = 287\ \text{cm}^2$. The short term average up-link beam spot size at the equal coherence plane, which defines the area over which reciprocal propagation paths may occur, is calculated from Eq. 7 to be $W_o = 17.21$ m and the corresponding area is $A_o = \pi W_o^2 = 930.64\ \text{m}^2$. The number of equal coherence plane cells through which reciprocal paths may occur is then $n_{cells} \approx A_o / A_{cell} \approx 32378$ and the coherent RPS enhancement factor is from Eq. 13 given by

$$f_{c.AT} = \frac{\left\langle I_{tot}\left(\theta_s = 0\right)\right\rangle}{\left\langle I_{ic}\left(\theta_s = 0\right)\right\rangle} = \frac{\left\langle I_{c.AT}\left(\theta_s = 0\right)\right\rangle + \left\langle I_{ic}\left(\theta_s = 0\right)\right\rangle}{\left\langle I_{ic}\left(\theta_s = 0\right)\right\rangle} = 2 - \frac{1}{n_{cells}} = 1.99997 \tag{18}$$

The transverse width of the coherent RPS peak in the transceiver plane $l_{c.AT}$, which is defined by the turbulence field transverse spatial coherence length at the transceiver l_{tcvr} for saturated turbulence scattering[5,6,18], is modeled as the zeroth order lobe of the Airy diffraction pattern produced by an effective coherent RPS aperture located in the equal coherence plane. The diameter of the effective coherent RPS aperture is taken to be twice the spot size of the up-link beam in the equal coherence plane. The predicted width of the coherent RPS peak in the transceiver aperture plane is given by Eq. 14 to be $l_{RPS} = 1134\ cm$. This simple estimate for the width of the coherent RPS peak provides excellent agreement with the value above $l_{tcvr} = 1.133$ cm obtained numerically from Eq. 11.

Using this approach, the average intensity component in the transceiver aperture plane[37] that is produced by the effective coherent RPS aperture is given by

$$I_{eff}\left(\rho_o\right) = \left(\frac{k W_o^2}{2 L_o}\right)^2 \left[2\frac{J_1\left(\frac{k W_o}{L_o}\rho_o\right)}{\frac{k W_o}{L_o}\rho_o}\right]^2 \tag{19}$$

where L_o is the propagation path length to the equal coherence plane. The effective coherent RPS aperture is utilized in this model; however, only to predict the transverse width of the coherent RPS peak in the transceiver aperture plane, thus the amplitude factor of this effective aperture is normalized to unity. Since we desire the transverse width of the coherent RPS peak to ultimately be expressed in terms of the scattering angle from the target, Eq. 19 is rewritten as

$$I_{eff}\left(\theta_s\right) \propto \left\{2\frac{J_1\left[\frac{k W_o}{L_o} L\tan\left(\theta_s\right)\right]}{\frac{k W_o}{L_o} L\tan\left(\theta_s\right)}\right\}^2 \tag{20}$$

The average coherent RPS intensity component produced by the cross-interference of scattered fields from the Lambertian surface that propagate in reciprocal directions through different up-link and down-link coherence cells in the equal coherence plane is therefore given by

$$\left\langle I_{c.AT}\left(\theta_s\right)\right\rangle = \frac{n_{cells}\left(n_{cells}-1\right)A_{trgt}}{L^2} I_t \cos\left(\theta_s\right)\left\{2\frac{J_1\left[\frac{kW_o}{L}L\tan\left(\theta_s\right)\right]}{\frac{kW_o}{L}L\tan\left(\theta_s\right)}\right\}^2 \tag{21}$$

and is shown in normalized form by Figure 7 versus radial position in the transceiver aperture plane.

The average total observed intensity for this example is now formulated from the results of the preceding discussions. The average total intensity is from Eqs. 17 and 21

$$\left\langle I_{tot}\left(\theta_s\right)\right\rangle = \left\langle I_{ic}\left(\theta_s\right)\right\rangle + \left\langle I_{c.AT}\left(\theta_s\right)\right\rangle = \frac{n_{cells}^2 A_{trgt}}{L^2} I_t \cos\left(\theta_s\right) + \frac{n_{cells}\left(n_{cells}-1\right)A_{trgt}}{L^2} I_t \cos\left(\theta_s\right)\left\{2\frac{J_1\left[\frac{kW_o}{L}L\tan\left(\theta_s\right)\right]}{\frac{kW_o}{L}L\tan\left(\theta_s\right)}\right\}^2$$

$$= \frac{n_{cells}^2 A_{trgt}}{L^2} I_t \cos\left(\theta_s\right)\left\{1+\frac{\left(n_{cells}-1\right)}{n_{cells}}\left\{2\frac{J_1\left[\frac{kW_o}{L}L\tan\left(\theta_s\right)\right]}{\frac{kW_o}{L}L\tan\left(\theta_s\right)}\right\}^2\right\} \tag{22}$$

It is interesting to note that this expression agrees correctly with the previous expression[5] Eq. 13 for the coherent RPS enhancement factor

$$f_{c.AT} = \frac{\left\langle I_{tot}\left(\theta_s=0\right)\right\rangle}{\left\langle I_{ic}\left(\theta_s=0\right)\right\rangle} = 1+\frac{n_{cells}-1}{n_{cells}} = 2-\frac{1}{n_{cells}} \tag{23}$$

This model is seen to provide an excellent approximation to the effect of RPS on the average total observed intensity for the propagation environment of this example, as illustrated by Figure 8.

The performance of this model is now investigated for other propagation environments by varying the zenith angle, the ground-level structure constant of the refractive index fluctuations, and the high altitude wind speed at 10 km. The various cases that are examined and presented in Table 1 are organized as follows. The ground-level structure constant $C_{n,o}^2$ assumes the values 1.0×10^{-14} $m^{-2/3}$, 5.0×10^{-14} $m^{-2/3}$, and 1.0×10^{-13} $m^{-2/3}$ and the high altitude wind speed *ws* assumes the values 21 m/s and 30 m/s, respectively. For each of the possible combinations of ground level structure constant $C_{n,o}^2$ and high altitude wind speed *ws* above, the slant path zenith angle ξ_{sp} is incremented in degrees when it is large enough to produce saturated turbulence effects at the target and until the up-link field transverse spatial coherence width at the target plane decreases to one meter.

6. CONCLUSION

The results of this analysis indicate that the phenomenological model for RPS presented in this analysis provides surprisingly accurate predictions for a variety of saturated turbulence propagation environments. One may wonder what benefit is provided by the phenomenological model presented, since the strength of the RPS peak approaches two under saturated turbulence conditions, and the width of the RPS peak is predicted by the down-link spatial coherence width. The answer is that the phenomenological turbulence RPS model herein may be combined with another phenomenological model for the RPS produced by backscattering from a non-Lambertian rough target. Since Lambertian surfaces are only a theoretical modeling tool and do not occur naturally, a laser radar tracking a space target receives RPS contribution from both atmospheric turbulence scattering as well as rough target scattering[38]. The phenomenological model for rough target RPS, as well as the combination of atmospheric turbulence RPS and rough target RPS is the subject of future publications.

Table 1. Atmospheric turbulence RPS model performance for various propagation environments.

Ground-level structure constant $C_{n,o}^2$ ($m^{-2/3}$)	High altitude wind speed ws (m/s)	Slant path zenith angle ξ_{sp} (deg)	Transceiver diameter D_x (mm)	Equal coherence altitude H- (km)	Numerical RPS peak width l_{tcvr} (mm)	Phenomenological RPS peak width l_{RPS} (mm)
1.0×10^{-14}	21	86	15	10.5	14.92	16.7
		87	12	10.6	12.56	13.46
		88	10	10.5	9.85	11.1
	30	85	16	11	16.2	17.92
		86	14	11.1	14.17	15.67
5.0×10^{-14}	21	85	8	10.4	7.92	8.87
		86	7	10.4	6.93	7.76
		87	6	10.4	5.83	6.63
		88	4	10.4	4.58	4.53
	30	84	8	11	8.71	9.02
		85	7	11.1	7.81	7.92
		86	6	11	6.83	6.8
1.0×10^{-13}	21	85	5	10.2	5.39	5.62
		86	4	10.2	4.71	5.49
		87	4	10.3	3.97	4.43
		88	3	10.3	3.11	3.35
	30	84	6	10.9	5.96	6.66
		85	5	10.9	5.35	5.61
		86	4	11	4.68	4.55

ACKNOWLEDGEMENTS

The authors wish to thank Dr. L. C. Andrews (University of Central Florida) and Dr. V. A. Banakh (Russian Academy of Sciences) for valuable technical discussions. This research, performed at the Center for Research and Education in Optics and Lasers, University of Central Florida, has been supported by the Ballistic Missile Defense Organization, Innovative Science and Technology Division through the Office of Naval Research under contract N00014-92-C-0123, monitored by Dr. William Stachnik, and by the Advanced Research Projects Agency through the Office of Naval Research under contract N00014-92-J-1906, monitored by Dr. Vernon Smiley.

REFERENCES

1. L. C. Andrews, W. B. Miller, and J. C. Ricklin, "Spatial Coherence of a Gaussian-beam wave in weak and strong optical turbulence," J. Opt. Soc. Am. A **11**, 1653-1660 (1994).

2. M. S. Belen'kii and V. L. Mironov, "Phase fluctuations when focusing light in a turbulent atmosphere," Radiophys. Quantum Electron. **12**, 1096-1101 (1983).

3. M. S. Belen'kii and V. L. Mironov, "Mean diffracted rays of an optical beam in a turbulent medium," J. Opt. Soc. Am. **70**, 159-163 (1980).

4. J. W. Goodman, *Statistical Optics* (Wiley - Interscience, New York, 1985) pp.128 and 361-464.

5. V. A. Banakh and V. L. Mironov, *Lidar in a Turbulent Atmosphere*, V. E. Zuev, ed. (Artech, Boston, MA 1987) pp. 25-96.

6. Y. N. Barabanenkov, Y. A. Kravtsov, V. D. Ozrin, and A. I. Saichev, "Enhanced backscattering in optics," *Progress in Optics XXIX*, E. Wolf, ed. (North Holland, Amsterdam, Netherlands 1991) pp. 65-197.

7. L. C. Andrews and R. L. Phillips, "Optical scintillations and fade statistics for a satellite communication system," Accepted for publication in *Appl. Opt.*.

8. W. B. Miller, J. C. Ricklin, and L. C. Andrews, "Effects of refractive index spectral model on the irradiance variance of a Gaussian beam," J. Opt. Soc. Am. A **11**, 2719-2726 (1994).

9. W. B. Miller, J. C. Ricklin, and L. C. Andrews, "Log-amplitude variance and wave structure function: a new perspective for Gaussian beams." J. Opt. Soc. Am. A **10**, 661-672 (1993).
10. M. S. Belen'kii, "Effect of atmospheric turbulence on heterodyne lidar performance," Appl. Opt. **32**, 5368-5372 (1993).
11. J. E. Harvey and A. Kotha, "Sparse array configurations yielding uniform MTF's in reciprocal path imaging applications," Opt. Comm. **106**, 178-182 (1994).
12. J. C. Dainty, T. Mavroidis, and C. J. Solomon, "Double passage imaging through turbulence." SPIE Propagation Engineering: Fourth in a Series, Vol. 1487, 2-9 (1991).
13. T. Mavroidis, J. C. Dainty, and M. J. Northcott, "Imaging of coherently illuminated objects through turbulence: plane-wave illumination." J. Opt. Soc. Am. A 7, 348-55 (1990).
14. V. I. Tatarskii, *Wave Propagation in a Turbulent Medium* (McGraw-Hill, New York 1961).
15. L. C. Andrews, W. B. Miller, and J. C. Ricklin, "Geometrical representation of Gaussian beams propagating through complex paraxial optical systems," Appl. Opt. **32**, 5918-5929 (1993).
16. L. C. Andrews, (technical conversation, December 23, 1994, saturated turbulence beam spread and coherence length).
17. R. E. Hufnagel, "Atmospheric turbulence" *The Infrared Handbook*, W. L. Wolfe and G. J. Zissis, ed., (The Infrared Information and Analysis Center, Environmental Research Institute of Michigan, Ann Arbor, MI 1989) pp. 6.1-6.20.
18. V. A. Banakh, (e-mail, December, 9, 1994, and technical letter, May, 17, 1994, dependence of backscatter amplification coherence width upon turbulence strength).
19. M. S. Belen'kii, "Diffraction of optical radiation by a reflecting disk in a turbulent atmosphere." Kvantovaya Electron. **N5**, 38-45 (1972).
20. V. A. Banakh, (e-mail, November, 15, 1994, and technical discussion, April 7, 1994, history of atmospheric turbulence RPS).
21. A. S. Gurvich and S. S. Kashkarov, "Amplification of scattering in a turbulent medium," Izv. Vyssh. Uchebn. Zaved. Radiofiz. **20**, 794-796 (1977).
22. V. I. Tatarski, "Some new aspects in the problem of waves and turbulence." Radio Science **22**, 859-865 (1987).
23. Y. A. Kravtsov and A. I. Saichev, "Effects of double passage of waves in randomly inhomogeneous media." Sov. Phys. Usp. **25**, 494-508 (1982).
24. J. H. Churnside and J. J. Wilson, "Enhanced backscatter of a reflected beam in atmospheric turbulence. Appl. Opt. **32**, 265 -2655 (1993).
25. E. Jakeman, J. P. Frank, and G. J. Balmer, "The effect of enhanced backscattering on target detection," Proceedings of the Agard Meeting on Atmospheric Propagation Effects Through Natural and Man-made Obscurants for Visible to Millimeter-Wave Propagation (Mallovea, Italy1993).
26. P. R. Tapster, A. R. Weeks, and E. Jakeman, "Observation of backscattering enhancement through an atmospheric phase screen," J. Opt. Soc. Am. A **6**, 517-522 (1989).
27. B. S. Agrovskii, A. N. Bogaturov, A. S. Gurvich, S. V. Kireev, and V. A. Myakinin, "Enhanced backscattering from a plane mirror viewed through a turbulent phase screen." J. Opt. Soc. Am. A **8**, 1142-1147 (1991).
28. Y. A. Kravtsov and A. I. Saichev, "Properties of coherent waves reflected in a turbulent medium." J. Opt. Soc. Am. A **2**, 2100-2105 (1985).
29. J. F. Holmes, "Enhancement of backscattered intensity for a bistatic lidar operating in atmospheric turbulence." Appl. Opt. **30**, 2643-2646 (1991).
30. V. A. Banakh, (e-mail, December, 7, 1994, dependence of backscatter amplification upon target dimensions).
31. The curve for K including IRF in Figure 2 is reprinted with permission from Figure 2.2 of reference 6.
32. A.E. Siegman, *Lasers* (University Science Books, Mill Valley, CA 1986) pp. 80-115 and 665-667.
33. A. V. Jelalian, *Laser Radar Systems* (Artech, Boston, MA 1992) pp. 3-10.
34. L. G. Shirley, E. D. Ariel, G. R. Hallerman, H. C. Payson, and J. R. Vivilecchia, "Advanced techniques for target discrimination using laser speckle." Lincoln Laboratory J. **5**, 380-391 (1992).
35. J. C. Stover, *Optical Scattering: Measurement and Analysis*, (McGraw-Hill, New York 1990) pp. 23-44 and 157-159.
36. C. G. Bachman, *Laser Radar Systems and Techniques* (Artech, Dedham, MA 1979) pp. 9-40.
37. J. W. Goodman, *Introduction to Fourier Optics* F. Terman, H. Heffner, and A. E. Siegman, ed., (McGraw-Hill, San Francisco, CA 1968) pp. 30-65.
38. R. A. Murphy, *Scattering from Rough Surfaces and Atmospheric Turbulence in Monostatic Laser Radar Systems* (Ph.D. dissertation, University of Central Florida, 1995).

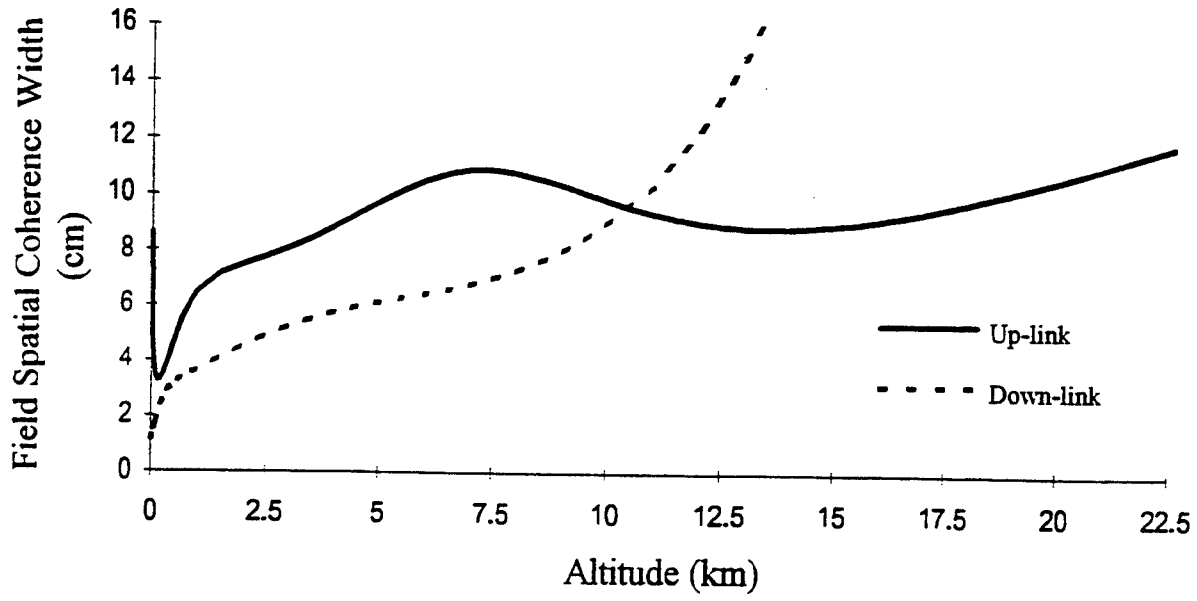


Figure 1. Comparison of the up-link and down-link field spatial coherence widths versus altitude for a 1.064 μm laser radar with a transceiver aperture diameter D_x = 1 cm at an altitude h_o = 2m for a path zenith angle ξ_{zp} = 86°, a ground level structure constant $C_{n,o}^2 = 1.9 \times 10^{-14}$ $m^{-2/3}$, a high altitude wind speed ws = 21 m/s, and a target altitude H = 300 km.

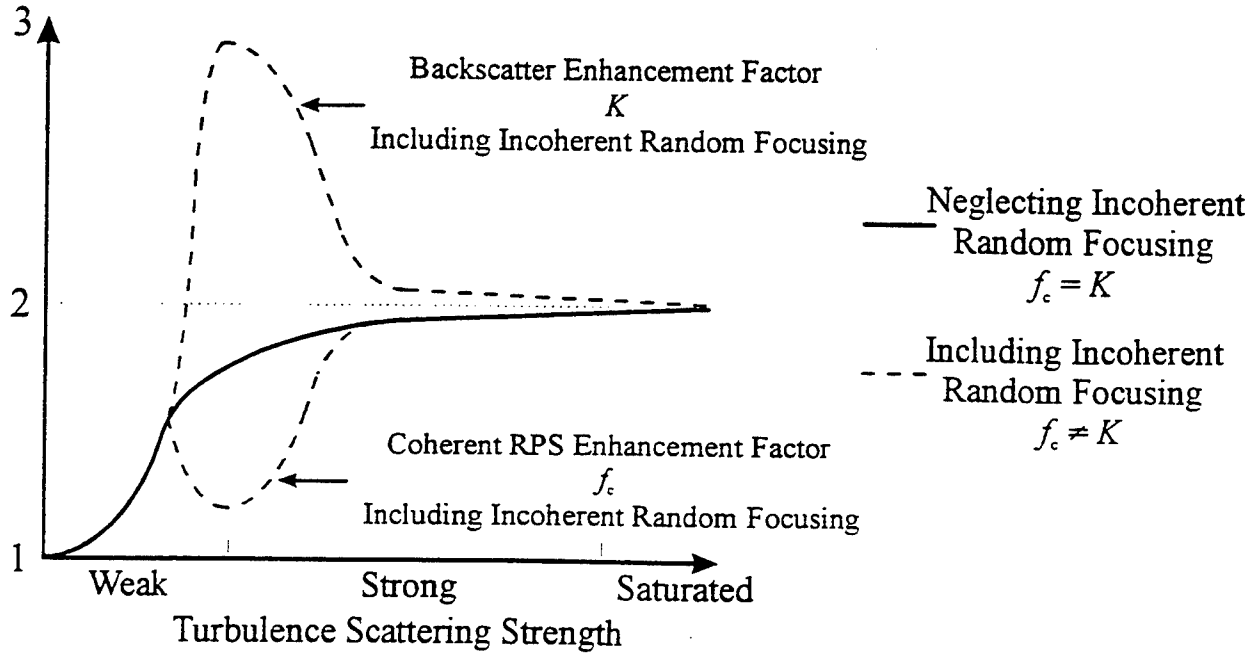


Figure 2. Comparison of the traditional backscatter amplification factor K with the coherent RPS enhancement factor $f_{c,AT}$ when incoherent random focusing is neglected (solid) and included (dashed)[6,31]. K and $f_{c,AT}$ are equivalent when incoherent random focusing is neglected.

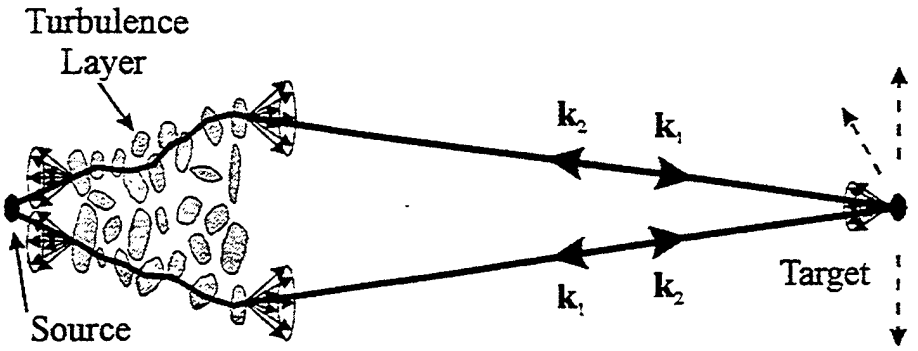


Figure 3. Coherent RPS is produced by the cross-interference of fields that propagate in reciprocal directions through different turbulence inhomogeneities over the up-link and down-link propagation paths to and from the target, respectively.

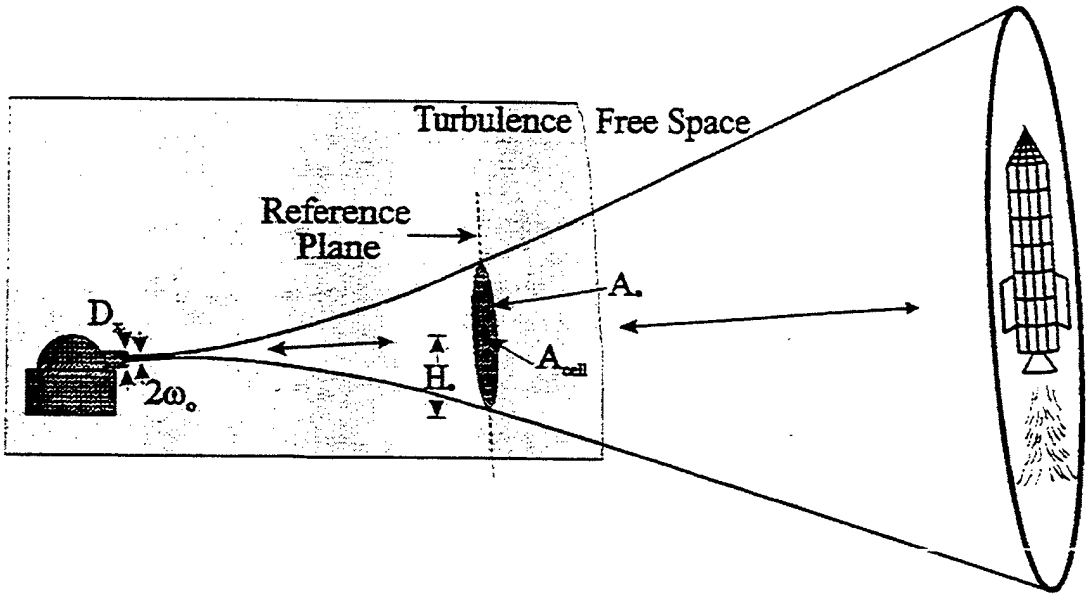


Figure 4. The area of the up-link beam $A_\bullet$ defines the cross-sectional area at the equal coherence plane through which strictly reciprocal propagation paths may occur. The location of the equal coherence plane within the atmosphere is obtained by determining the altitude $H_\bullet$ for the round-trip propagation path at which the up-link and down-link field spatial coherence widths are equal.

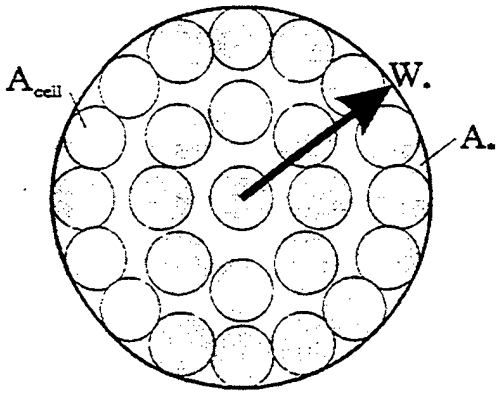


Figure 5. The coherence cells are identical and randomly distributed within the cross-sectional area of the up-link beam $A_\bullet$ at the equal coherence plane. The area of a coherence cells A_{cell} is defined by the up-link and down-link field spatial coherence length $l_\bullet$ at the equal coherence plane.

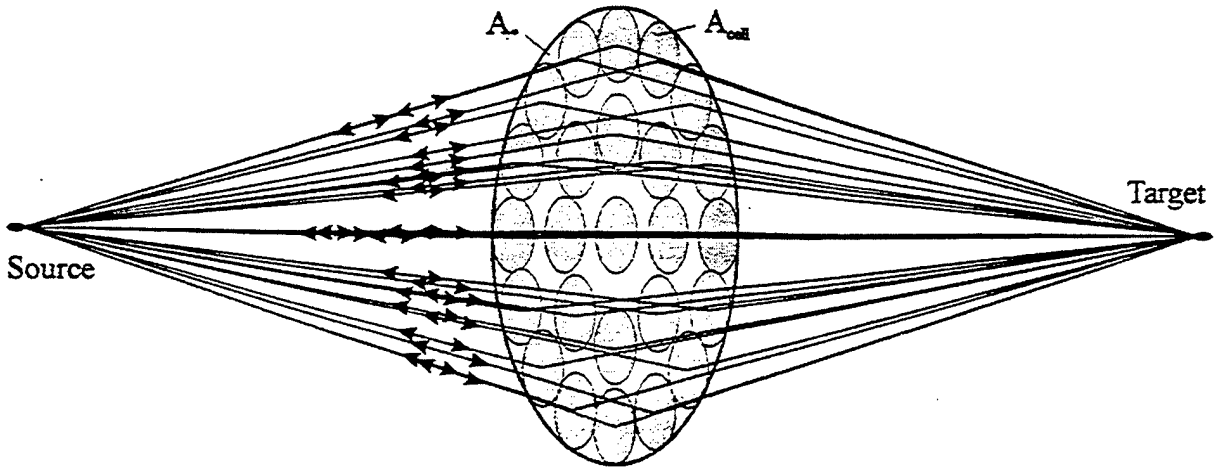


Figure 6. The up-link and down-link round-trip propagation paths to and from the target through the turbulence occur either through the same coherence cell, or through different coherence cells in the equal coherence plane.

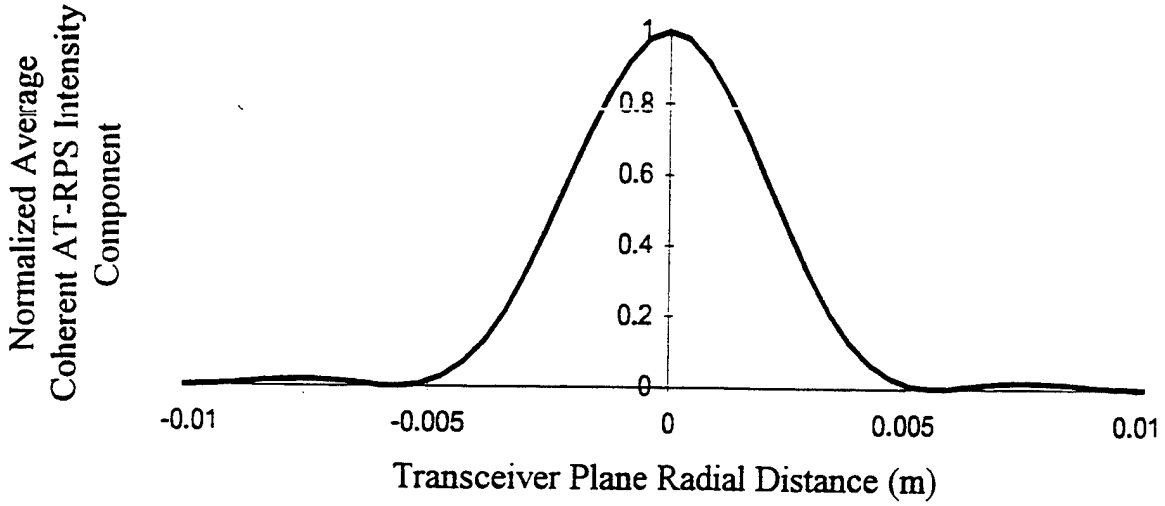


Figure 7. Normalized average coherent RPS intensity component versus radial position in the transceiver aperture plane.

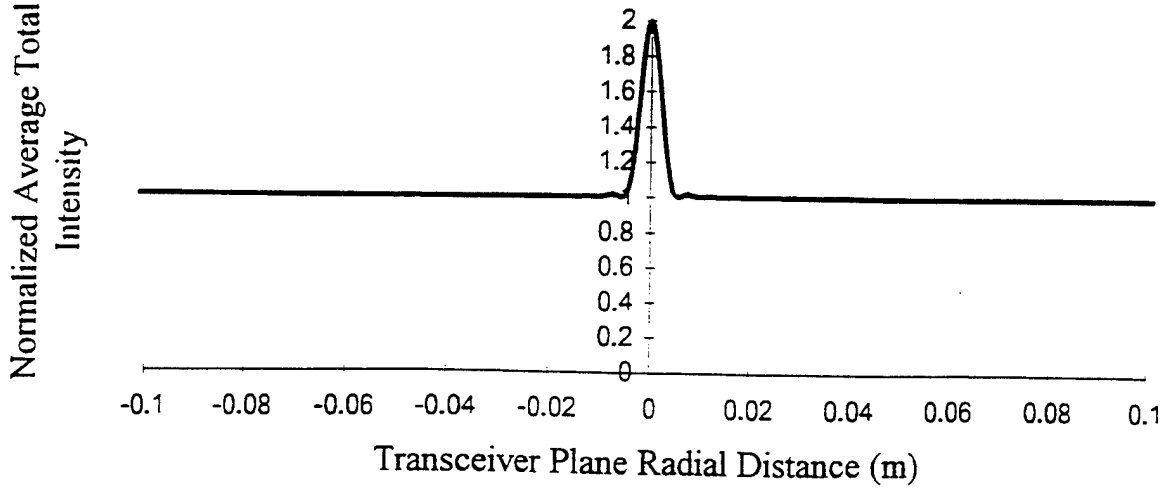


Figure 8. Normalized average total intensity versus radial position in the transceiver aperture plane.