Phase Diversity for Segmented and Multi-Aperture Systems

by

Matthew R. Bolcar

Submitted in Partial Fulfillment

of the

Requirements for the Degree

Doctor of Philosophy

Supervised by

Professor James R. Fienup

The Institute of Optics
The College
School of Engineering and Applied Sciences

University of Rochester
Rochester, New York

2008

Curriculum Vitae

The author was born in Phillipsburg, NJ on October 19, 1979. He attended Cornell University from 1998 to 2002, and graduated with a Bachelor of Science degree in Applied and Engineering Physics, with a concentration in Optics. He came to the University of Rochester in the Fall of 2002 and began graduate studies in Optics under the direction of Professor James R. Fienup. He received a Graduate Student Research Program Fellowship from NASA Goddard Space Flight Center in 2005, 2006, and 2007. His technical advisor at NASA Goddard Space Flight Center was Dr. Bruce H. Dean.

Acknowledgements

My success today is only possible through the support and efforts of many people.

First and foremost, I cannot give enough thanks to my advisor, Professor Fienup, for always having his door open to discuss questions or ideas regarding my research, thoughts on my future career, or just the news of the day. I started this journey with very little background in wavefront sensing, but he was quick to guide me toward a topic that I found challenging and fascinating. He was always open-minded and careful to give suggestions and not orders, allowing me the freedom to explore (and occasionally get lost) on my own. I could not ask for a better mentor. And while I saw more of his red pen then I'd sometimes like to admit, his tireless efforts in revising my papers have not only improved the quality of my work, but taught me to be a better writer.

There are several people responsible for the financial support of my graduate research. I am grateful to Rick Kendrick at Lockheed Martin Corporation for supporting my early research on phase diversity for multi-aperture systems.

I am also extremely thankful to NASA Goddard Space Flight Center, and specifically Bruce Dean, for support during my GSRP Fellowship. Bruce was always ready to offer a challenging problem for me to tackle or an enthusiastic discussion to work through a question.

During the past few years, I've had the great honor of meeting many people who are pioneers of their fields: Bob Gonsalves, Rick Paxman, Rick Lyon, Chuck Matson, Scott Acton and many others. I am thankful to them for many great

discussions and their encouragement and guidance has always inspired me and kept me motivated.

I am grateful to my committee, Prof. Duncan Moore, Prof. Miguel Alonso, and Prof. Gaurav Sharma, for their advice and guidance in directing my research.

The administrative staff of The Institute of Optics has made life easier during the last few years. Joan Christian, Noelene Votens, Lissa Cotter, Lori Russel, Gina Kern, Gayle Thompson, Marie Banach, Betsy Benedict, and Maria Schnitzler have made sure I've gotten paid, reimbursed, kept me stocked with supplies, helped me order equipment, reserved rooms, helped plan events or were just willing to chat during a moment of procrastination. I am indebted to them for all of their hard work. Special thanks go to Per Adamson, for lending me his HeNe laser for so many outreach demonstrations.

The past few years would have been quite literally impossible without the support of many friends.

To my colleagues in Prof. Fienup's group: I owe more to each of them than I can express. Thanks to Abbie and Manuel, for entertaining me when I would wander into their office with a question or to just waste time. Thanks to Tom and Greg for making sure my computer always worked and my data was always saved, but especially for their collaboration on so many projects and adventures. And thanks to Sapna, for her thoughtfulness and cheerfulness that made even the most ordinary day seem special. Thanks are also due to Sam Thurman, for always being willing to help

me with my early work in multi-aperture systems and my later work in image reconstruction. His expertise has been a valuable asset to my graduate career.

Special thanks are owed to my housemates, Nathan Clark, Graham Marsh, Sean Anderson, Brad Deutsch, Emily Clark and Kathleen Adelsberger, for injecting a healthy mixture of excitement and paranoia into my life. The adventures we have had and the pranks they have staged will *never* be forgotten.

Thanks to Jessica Morgan and Jessica DeGroote for being my study group during the first two years. Sometimes you are like sisters to me and, much to my chagrin, sometimes like mothers, but you will always be my friends.

Thanks to my many friends outside of Rochester: Asheen Phansey, Chris Konek, Frances Spalding, Joel Chestnutt, Herb Wilmoth, Robert Kovacs and Sandy Veresink, for being a home away from home and reminding me that there is, in fact, life outside of graduate school.

I owe a great deal of thanks to my martial arts instructors, Master Ray Mondschein and Shihan Cyndy Jones, for their instruction, friendship, and for providing a place for me to train and channel the frustrations of a bad day.

An unending amount of thanks go to Sara Johnston for always being there with a ready ear or a hearty laugh. I am lucky to have met you.

The greatest thanks are saved for my family.

To my Aunts Madeline and Maggie and Uncle Joe, thanks for always supporting me, for always being interested in my pursuits, and for teaching me the importance of family, and that life should be fun.

To my sister Jen, thanks for sharing my childhood and being my first and best friend. Thank you for providing music to keep me inspired and encouragement to keep me motivated.

And most importantly, thank you to my parents, Judy and Paul. You have always supported me, letting me choose the path I wanted to take and never questioning my decisions. It is only through your love, work, and sacrifice that I have had the opportunity to do as I dreamed. I owe my success to you, and dedicate this thesis to you. Thank you.

Abstract

As telescopes become larger, segmented and multi-aperture designs are being implemented to meet cost, size and weight constraints. These systems require alignment of the segments or sub-apertures to within fractions of a wavelength. We investigate the performance of phase diversity, a technique of image-based wavefront sensing, for characterizing and aligning segmented and multi-aperture systems. Supporting work developing the core phase-diversity algorithm is also presented.

The modification of phase diversity to incorporate a broadband object model is discussed. Through digital simulation, we show a benefit to using a broadband, gray-world algorithm, as opposed to the conventional monochromatic algorithm, when bandwidths greater than 20% are used for imaging. We further demonstrate that knowledge of the gray-world object spectrum is not required to achieve improved performance.

Using digital simulation, three regularization techniques for phase-diversity metrics are compared to the conventional phase-diversity algorithm. For low signal-to-noise ratio (SNR), we demonstrate an improvement in both the phase estimation accuracy and convergence properties of the algorithm when a regularization based on the object and noise power spectra is used.

We present a novel implementation of phase diversity unique to segmented and multi-aperture systems that utilizes individual sub-aperture piston phases in the pupil. Through digital simulation, we compare sub-aperture piston phase diversity to conventional focus diversity and show that, for high SNR and optimal values of

diversity, sub-aperture piston phase diversity yields higher phase estimation accuracy and comparable image reconstructions.

Using large amounts of sub-aperture piston phase diversity we derive a modified phase-diversity metric and gradients that enable estimation of the average object spectral coefficients, in addition to the unknown phase and object. We successfully demonstrate the algorithm through digital simulation and explore limitations on spectral resolution and sampling.

An experimental setup is described and used to demonstrate monochromatic and broadband versions of focus diversity and sub-aperture piston phase diversity. A hexagonally segmented MEMs deformable mirror is used as the system under test and a modified digital projector is used to provide an extended object source. Phase-shifting interferometry is used to corroborate phase estimates from the phase-diversity algorithm. Agreement between the estimated phase and measured phase to 0.06 λ RMS is achieved. Images are also reconstructed using the estimated phases.

A new method of performing multi-field wavefront sensing that directly estimates a field-dependent model of the wavefront using phase-diverse phase retrieval is presented. Using digital simulations and information-theoretic Cramer-Rao bounds, we show an advantage to using the new method over a conventional technique in which the field-dependent wavefront is extrapolated or interpolated from a series of wavefronts across samples of the field of view.

This work demonstrates that phase diversity is a useful method of aligning segmented and multi-aperture systems. Furthermore, the architecture of such systems

offers unique opportunities for the implementation of phase diversity and for recovering the system phase, an image of the object and the average spectrum of the object.

Table of Contents

List of Tables

List of Figures

List of Acronyms

AO	Adaptive Optics
AU	Arbitrary Units
AVIRIS	Airborne Visual Infrared Imaging Spectrometer
CCD	Charge-Coupled Device
CMOS	Complementary Metal-Oxide-Semiconductor
CRB	Cramer-Rao Bound
DFT	Discrete Fourier Transform
DIRS	Digital Imaging and Remote Sensing
DL	Diffraction Limited
DLP™	Digital Light Processing™
DM	Deformable Mirror
DMD	Digital Mirror Device
EELT	European Extremely Large Telescope
FFT	Fast Fourier Transform
FIM	Fisher Information Matrix
FOV	Field of View
FTIS	Fourier Transform Imaging Spectroscopy
GMT	Giant Magellan Telescope
GS	Gerchberg-Saxton
HeNe	Helium-Neon
HET	Hobby-Eberly Telescope
HST	Hubble Space Telescope
IFFT	Inverse Fast Fourier Transform
ITA	Iterative Transform Algorithm
JMAP	Joint Maximum *A Posteriori*
JWST	James Webb Space Telescope
mAP	Marginal *A Posteriori*
MEMS	Micro-Electro-Mechanical Systems
MGS	Misell-Gerchberg-Saxton
MIDAS	Multi-Instrument Distributed Aperture Sensor
MMT	Multiple Mirror Telescope
MTF	Modulation Transfer Function
NRMSE	Normalized Root-Mean-Squared Error
NSR	Noise-to-Signal Ratio
OPD	Optical Path Difference
OPL	Optical Path Length

OTF	Optical Transfer Function
PD	Phase Diversity
PDPR	Phase-Diverse Phase Retrieval
PR	Phase Retrieval
PSF	Point Spread Function
PV	Peak-to-Valley
PZT	Piezo-electric Transducer
RG	Reduced Gaussian
RMS	Root-Mean-Square
RMSE	Root-Mean-Squared Error
ROI	Region-of-Interest
RRG	Regularized Reduced Gaussian
SAFIR	Single Aperture Far-Infrared
SALT	South African Large Telescope
SAPPD	Sub-Aperture Piston Phase Diversity
SNR	Signal-to-Noise Ratio
TMT	Thirty Meter Telescope
TPF-I	Terrestrial Plane Finder - Interferometer
WFS	Wavefront Sensing

1 Introduction

1.1 Segmented and Multi-Aperture Systems

The angular resolution of a diffraction-limited imaging system is given by

$$\theta = 1.22\frac{\lambda}{D} \tag{1.1}$$

where λ is the wavelength and D is the diameter of the entrance pupil of the imaging system [1]. Equation (1.1) implies that the resolution of the system can be improved either by decreasing the wavelength of observation or increasing the diameter of the aperture of the system. The wavelength is typically fixed by the illumination of the object or by the phenomenology being studied. This leaves the diameter of the aperture as the only independent variable.

As a result, telescopes are getting bigger. The drive to increase both the light collecting efficiency and imaging resolution of ground-based and space-based observatories has pushed the limits of size and weight for monolithic primary mirrors. Not only are large primary mirrors increasingly difficult to fabricate and mount, but for space-based systems they may be impossible to launch. For example, the Hubble Space Telescope (HST) represents the largest monolithic-primary payload capable of being stowed in current launch vehicles.

In response, technology is heading in the direction of segmented and multi-aperture systems as the next generation of telescopes. Segmented systems, such as the Keck Observatory in Hawaii [2], the Thirty Meter Telescope (TMT) [3] and the James Webb Space Telescope (JWST) [4] to be launched after 2013, use an array of

actuated hexagonal segments to create a primary mirror that is easier to fabricate than a monolithic mirror of equivalent size, and, in the case of the JWST, capable of being folded and stowed in a launch vehicle and deployed on orbit. Proposed multi-aperture systems, such as the Terrestrial Planet Finder Interferometer (TPF-I) [5], Keck Interferometer [6], and MIDAS [7] use an array of afocal telescopes that are combined to achieve a resolution comparable to a primary mirror equivalent to the entire array size. An example of a multi-aperture system is shown in Figure 1-1.

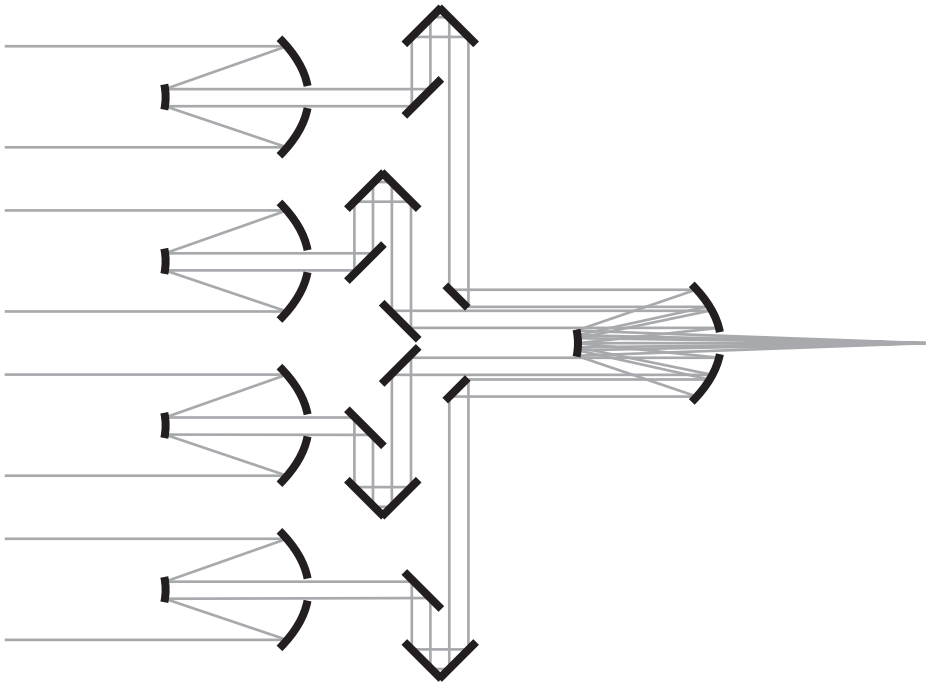


Figure 1-1: Schematic of a multiple telescope array where the light from each sub-aperture is brought together at a common focus. A series of optical delay lines are used to maintain equal optical path length between each sub-aperture. If the path lengths are kept equivalent, the system has a resolution determined by the diameter of the entire array, rather than the diameter of an individual sub-aperture. This drawing is provided by Samuel T. Thurman and is used with permission.

1.1.1 History: Multi-aperture Systems

One of the earliest examples of a multi-aperture system is Michelson's experiment at Mt. Wilson in 1921, where he and Pease measured the angular diameter of α-Orionis [8-10]. The primary mirror of the Mt. Wilson 100-inch reflector was used to combine two beams from separate apertures at the end of the Cassegrain cage. The visibility of the fringes created by the interfering beams was monitored as the distance, d, between the two apertures was varied. Michelson showed [9] that the smallest distance, d_o, for which the fringe visibility is zero is given by

$$d_o = \frac{0.61\lambda_o}{\theta} \tag{1.2}$$

where λ_o is the center wavelength and θ is the angular diameter of the source. The accuracy of Michelson's method was largely dependent on very precise control of the optical path lengths (OPLs) between the two sub-apertures. Unequal path lengths reduced the fringe visibility and therefore introduced error in the measurement of d_o.

Later examples of multi-aperture interferometers include the stellar intensity interferometer of Hanbury Brown and Twiss [9, 10]. This system made use of the Hanbury Brown-Twiss effect [9, 11] to determine a star's angular diameter from the correlation coefficient of photoelectric signals as a function of the distance, d, between the sub-apertures [12]. In 1975, Labeyrie expanded on Michelson's stellar interferometer with the first true separate-aperture optical interferometer (as opposed to Michelson's adaptation of the Mt. Wilson 100-inch reflector) [10, 13]. In

Labeyrie's instrument, the light from two separate telescopes was brought together on an optical table and interfered.

These early multi-aperture systems required that the object being observed was spatially incoherent and unresolved by a single sub-aperture. According to the van Cittert-Zernike theorem [9], light from such a source will have some partial coherence in the far field. The purpose of these early systems was to measure the partial coherence function of light from a distant star. The amplitude of the partial coherence function is related to the fringe visibility and the phase is related to the position of the intensity maxima in the fringe pattern. The intensity distribution of the star can, in theory, be obtained from the Fourier transform of the complex partial coherence function.

A different class of multi-aperture system is those that observe objects that are extended and resolved by each sub-aperture. One of the earliest examples of such a system was the Multiple Mirror Telescope (MMT) in 1985 which consisted of six sub-apertures, each with a diameter of 1.83 meters [14, 15]. The light from all the sub-apertures is brought to a common focus to form an image of the object being observed. If the OPL for each sub-aperture is equal, then the resulting image has a resolution equivalent to that of a monolithic mirror of the same size as the sub-aperture array, or 6.9 meters in the case of the MMT.

The first experiment to demonstrate rapid active control of the OPL for each sub-aperture was performed using Phasar by the Air Force Weapons Laboratory in 1988 [16, 17]. The experiment sensed and corrected for piston, tip and tilt

misalignments between the sub-apertures. On a much larger scale, the Keck Interferometer combines the light of the two 10-meter Keck Telescopes, each of which is itself a segmented system, to perform both multi-aperture interferometry and multi-aperture imaging [6].

1.1.2 History: Segmented Systems

Segmented systems share the same purpose as multi-aperture systems – to reduce weight, cost and fabrication limits – but have a different heritage. The first segmented systems were the twin Keck 10-meter telescopes in Hawaii, which saw first light in 1993 and 1996 [18]. Thirty-six 1.8-meter hexagonal segments comprise each of the primary mirrors. Each segment is lighter, cheaper and easier to fabricate than a 10-meter monolithic mirror. Furthermore, the smaller segments experience reduced gravity sag compared to a larger primary mirror.

Other current segmented systems include the Hobby-Eberly Telescope (HET) in Texas [19] and the South African Large Telescope (SALT) located in Karoo, South Africa [20]. Each of these systems has 91 hexagonal segments making up a single 10-meter primary mirror.

1.1.3 Future Segmented and Multi-aperture Systems

With the success of these segmented and multi-aperture systems, numerous others have been proposed and have even begun construction. For ground-based systems, the Giant Magellan Telescope (GMT) will combine 6 circular sub-apertures using a

common secondary mirror [21]. The Thirty Meter Telescope (TMT) will use 492 segments to create a 30-meter primary mirror. The European Extremely Large Telescope (EELT) will use more than 900 segments to create a 42-meter primary mirror [22].

Since segmented or multi-aperture systems can be folded into a compact package or constructed modularly, they are ideal for space-based observatories. Perhaps the best example is the James Webb Space Telescope, to be launched after 2013. This observatory will have a 6.6-meter primary mirror comprised of 18 hexagonal segments. The observatory will be folded for launch and deployed on orbit. Other proposed space-based systems include the Single Aperture Far-Infrared (SAFIR) observatory [23], the Terrestrial Planet Finder – Interferometer (TPF-I) [5], the stellar imager [24] and MIDAS [7]. While the JWST is well into construction and testing, the other missions are still in the initial design phases.

1.2 Wavefront Sensing for Segmented and Multi-aperture Systems

Regardless of whether the system is a multi-aperture interferometer, multi-aperture imager or segmented system, a crucial part of the design is control of the OPL between the segments or sub-apertures [16, 17]. For multi-aperture interferometers, if the OPLs are not equivalent, fringe visibility rapidly decreases. This ultimately limited the experiments of Michelson and Pease in 1921. For multi-aperture imaging systems and segmented systems, if the OPLs are not equivalent, the light from each

sub-aperture does not coherently combine at the detector plane and the resolution will be determined by the sub-aperture size and not that of the full sub-aperture array.

Modern segmented and multi-apertures systems are designed to allow for precise control of the sub-aperture OPLs. Multi-aperture systems utilize optical delay lines to equalize the path lengths between each sub-aperture. Segmented systems mount each segment on a hexapod to allow for control of six degrees of freedom: piston, tip, tilt, x-translation, y-translation and clocking. In the case of the JWST a separate actuator is used to control intra-segment radius of curvature. It is necessary to know the position of each segment or sub-aperture to within a small fraction of a wavelength in order to apply the appropriate corrections to the actuators. There are several classes of wavefront sensor that are capable of providing such knowledge.

1.2.1 Interferometry

Laser interferometry may be an option for initial alignment and commissioning of segmented and multi-aperture systems, however given the size and complexity of modern systems it is increasingly difficult and costly to fabricate appropriate null surfaces. Interferometers are sensitive to vibration, turbulence and temperature instabilities, making it difficult to house the end-to-end test in an interferometry-friendly environment. Furthermore, frequent recalibration of the segmented or multi-aperture system is likely necessary due to drift in the mechanical mounting and actuating systems, requiring that sensing and alignment be performed repeatedly *in situ*.

1.2.2 Shack-Hartmann, Curvature and Pyramid Sensors

Shack-Hartmann [25-29], curvature [27, 30] and pyramid [31-33] wavefront sensors have proven useful for large ground-based systems where rapid sensing of atmospheric turbulence is required. Generally these wavefront sensors only work for point source objects and do not perform well for wavefronts having high complexity. Laser guide stars can provide an isolated point source for ground-based systems, but do not work for space-based systems since they use the properties of the sodium layer of the upper atmosphere. Shack-Hartmann wavefront sensors have been shown to work with extended scenes when used in a correlation mode [34], but require bright, high contrast objects that only fill very a limited field of view.

While Shack-Hartmann, curvature and pyramid sensors may be well suited for ground-based systems sensing atmospheric turbulence, they are less attractive for space-based systems. All three of these sensors require additional hardware and optics to obtain an image of the pupil of the system where wavefront or wavefront-slope measurements are made – a detriment to space-based systems where size and weight are critical issues. These sensors also do not work well for highly complex wavefronts, and in the case of the Shack-Hartmann wavefront sensor, cannot sense pure piston errors between segments and sub-apertures. As many future space-based missions will be based on segmented or multi-aperture architectures, these methods of wavefront sensing will not be sufficient.

1.2.3 Image-based Wavefront Sensing

Image-based wavefront sensing is a method that uses images captured by an optical system to estimate the wavefront in the pupil of the system. This differs from interferometric methods that require a reference wavefront, matched to the nominal optical prescription of the system, to form an interference pattern with the wavefront produced by the system under test. Image-based wavefront sensing further differs from Shack-Hartmann, pyramid and curvature wavefront sensors that require obtaining an image of the pupil of an optical system to measure wavefronts.

While image-based wavefront sensing requires significantly less hardware than other methods, it does require more computation to obtain wavefront estimates and is therefore not well suited to ground-based atmospheric sensing applications which require kHz speeds. However, image-based wavefront sensing has become an enabling technology for many segmented and multi-aperture systems where wavefront sensing and control will occur on longer time scales. Image-based wavefront sensing is not generally limited in the complexity of wavefronts being estimated, and is capable of sensing pure piston errors between segments and sub-apertures, making it an ideal option for future space-based missions. For example, not only will image-based wavefront sensing be an integral part of the end-to-end ground testing of the JWST, but it will also be the primary method of commissioning and maintaining fine alignment of the observatory on orbit [35-37].

In the literature, one comes across several names for the various methods of image-based wavefront sensing. Often, these names are used interchangeably and

confusion ensues. Therefore, we begin by making clear the definitions and terminology used in this thesis.

Phase retrieval (PR) refers to a basic form of image-based wavefront sensing in which a *single* image of a *point source* is used to estimate the unknown phase parameters of an optical system.

Phase-diverse phase retrieval (PDPR) is similar in concept to phase retrieval except *multiple* images of a *point source*, each differing by a known phase term, are used to estimate the unknown phase parameters of an optical system.

Phase diversity (PD) refers to a method of image-based wavefront sensing where *multiple* images of an *unknown extended object or scene* are used to estimate *both* the unknown phase parameters and the unknown object. The multiple images differ by some known phase term.

Henceforth, when we use the term "phase retrieval", it will include "phase-diverse phase retrieval", which is more commonly used since the additional measurements provide needed robustness. Both methods are essentially the same algorithmically (with some extra bookkeeping needed for multiple images), and both methods estimate the same quantities. In the next two sections, phase retrieval and phase diversity will be discussed in detail.

1.3 Phase Retrieval

1.3.1 History

The concept of phase retrieval has been around for many decades, particularly in the field of X-ray crystallography. The first practical algorithm for non-crystallographic data was in the early 1970's when Gerchberg and Saxton first proposed it for use in electron microscopy [38, 39]. The Gerchberg-Saxton (GS) algorithm was iterative in nature, involving the propagation of a complex field between a diffraction plane and an image plane while applying measured amplitude constraints to the field in each plane. Misell later adapted the GS algorithm to iterate between two out-of-focus image planes, satisfying the constraint of the two measured image intensities [40, 41]. Later, Southwell [42], Gonsalves *et al.* [43, 44] and Fienup [45] would expand the techniques to include parametric search algorithms in which a nonlinear optimization routine attempts to minimize an objective function. Gonsalves and Chidlaw also suggested the use of multiple images, introducing the concept of phase-diverse phase retrieval to the field [44].

Phase retrieval gained popularity in 1991 when it successfully characterized the primary mirror errors of the in-orbit Hubble Space Telescope [46, 47]. Around this same time, phase retrieval was proposed to align NASA's Next Generation Space Telescope, now called the James Webb Space Telescope (JWST). Today, in addition to being an enabling technology for the JWST, phase retrieval is being applied to many fields, including imaging through atmospheric turbulence, X-ray diffraction crystallography, and optical metrology [48-50]. Many algorithm advancements have

been made including a new form of iterative transform algorithm that is capable of dealing with wrapped phases [37], and advancements to nonlinear optimization routines that extend the algorithm capture range and can account for line-of-sight and figure jitter [51, 52].

1.3.2 Algorithm Description

Most phase retrieval algorithms can be classified in one of two categories: the iterative transform methods of Gerchberg, Saxton and Misell, or the parametric search methods of Southwell, Gonsalves and Fienup. An example of a basic optical system on which phase retrieval would be performed is show in Figure 1-2.

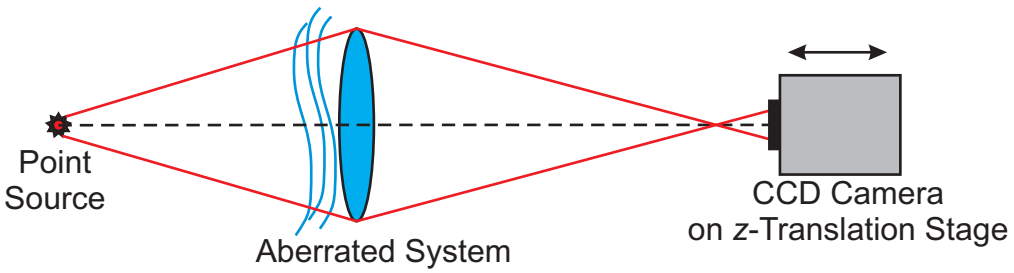


Figure 1-2: Example phase retrieval system. A point source is imaged by an aberrated system onto a CCD camera. The CCD camera is mounted on a *z*-translation stage to allow for multiple out-of-focus images to be collected.

Sometimes called Iterative Transform Algorithms (ITA), Gerchberg-Saxton (GS), or Misell-Gerchberg-Saxton (MGS), iterative methods are perhaps the simplest implementation of phase retrieval. The algorithm works as a projection-onto-sets

algorithm in which the complex field is iteratively propagated between two planes while constraints are applied to the field in each plane.

An example flow diagram of the ITA is shown in Figure 1-3. In this example, the complex pupil field $P(x,y)$ is propagated to the image plane field $h(u,v)$ by the operator $\mathbf{P}\{...\}$ which may consist of a series of Fourier or Fresnel integrals. In each plane, a constraint is placed on the field amplitude. For example, in the pupil plane, one might assume that the field amplitude is uniform and limited by the aperture function of the system, $|P(x,y)|$. In the image plane, one detects an intensity point spread function (PSF), $I(u,v)$, which is related to the field amplitude by a square root operation and is used as the amplitude constraint in the image plane. Eventually, the algorithm finds a set of phase functions $\phi(x,y)$ and $\theta(u,v)$ such that the two complex fields, $|P(x,y)|\exp[i\phi(x,y)]$ and $[I(u,v)]^{(1/2)}\exp[i\theta(u,v)]$, simultaneously satisfy the constraints in both planes.

By comparison, parametric search algorithms are significantly more complicated in implementation yet offer greater flexibility of the system modeling. A parametric search algorithm seeks to minimize an objective function with respect to a set of unknown phase parameters. These unknown phase parameters may be individual pixel values, or coefficients of a polynomial set, such as the Zernike polynomials. An example of an objective function is a basic squared sum of differences given by

$$E(\boldsymbol{\alpha}) = \sum_{k=1}^{K} \sum_{u,v} \left[\left| h_k(u,v;\boldsymbol{\alpha}) \right| - \sqrt{I_k(u,v)} \right]^2 \tag{1.3}$$

where $I_k(u,v)$ is the detected PSF intensity, K is the number of detected PSF intensities, and $h_k(u,v;\boldsymbol{\alpha})$ is an estimate to the coherent impulse response that depends on the unknown phase parameters $\boldsymbol{\alpha}$.

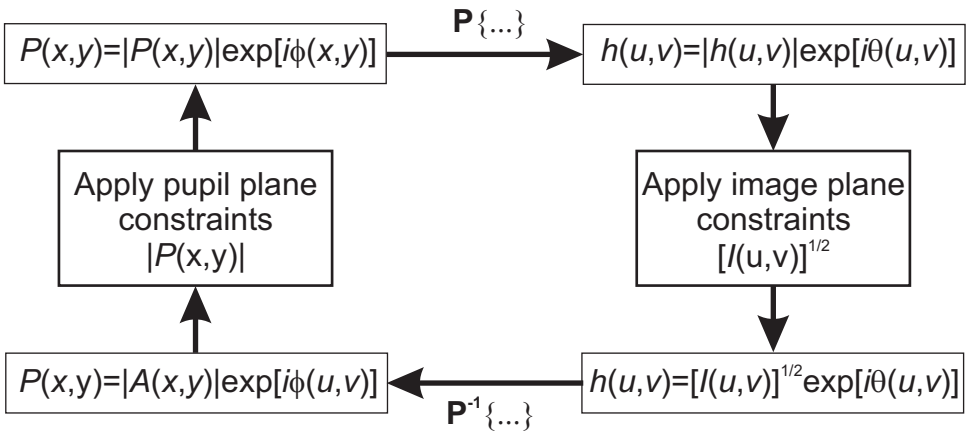


Figure 1-3: Flow diagram for the ITA algorithm. Starting in the upper left, a guess of the phase, $\phi(x, y)$, is made and used with the known pupil amplitude, $|P(x, y)|$, to construct the complex pupil. The field is propagated to the image plane where the computed phase, $\theta(u, v)$, is kept and the amplitude is replaced by $[I(u, v)]^{1/2}$, the square root of the measured intensity. The resulting field is then back-propagated to the pupil plane where the pupil plane amplitude constraint is once again applied. The process continues until a phase pair $\phi(x, y)$ and $\theta(u, v)$ are found such that the complex fields satisfy the constraints in each domain.

The error metric in Eq. (1.3) is minimized with respect to the parameters $\boldsymbol{\alpha}$ by a nonlinear optimization routine such as steepest descent, conjugate gradient, simplex, Newton or quasi-Newton, Levenberg-Marquardt, simulated annealing, or even genetic algorithm. Typically, the gradient search algorithms are the most efficient for this type of problem. Analytic gradients of Eq. (1.3) can be readily calculated, thus avoiding the costly computation and errors associated with finite

difference approximations. Additional terms can be included in the error metric to account for experimental unknowns such as detector bias, gain, bad pixels, jitter, broadband light, etc. Furthermore, the wavefronts can be parameterized in terms of polynomials to naturally enforce a smoothness constraint while avoiding the need to unwrap phases.

Extensive work on types of metrics and implementation of the nonlinear optimization has been done [45, 53, 54] and will not be discussed here. This section has given only a very brief description of the phase retrieval procedure to serve as a background to the subject of the rest of this thesis, phase diversity.

1.4 Phase Diversity

1.4.1 History

Phase diversity was introduced in 1979 by Gonsalves and Chidlaw when they suggested applying the techniques of phase-diverse phase retrieval to a system where the source was not a point, but rather an extended object or scene [44]. Paxman and Fienup applied phase diversity to the problem of phasing a multi-aperture system in 1988 [55]. Later work by Paxman, Schulz and Fienup casts phase diversity as a statistical problem based on maximizing a likelihood function related to the dominant noise mechanism of the system [56]. Some work has been done to implement phase diversity as an iterative problem not unlike the ITA, though it is generally better suited to nonlinear optimization routines [57].

In 1994, Kendrick *et al* demonstrated in experiment the phasing of a segmented aperture using phase diversity [58]. Not only were they able to sense the errors in the segmented system, but the estimated phase was then used to actuate the mirror and correct the wavefront. In 2000, Seldin *et al* presented both simulation and experimental results of using a broadband version of phase-diverse phase retrieval to sense the misalignments in a multi-aperture system [59]. They also extended the broadband derivation to phase diversity. Major advancements have been made in increasing the speed of phase diversity algorithms using dedicated hardware [60, 61], yet they have not reached the speeds necessary to sense atmospheric variations for correcting ground-based observatories with adaptive optics.

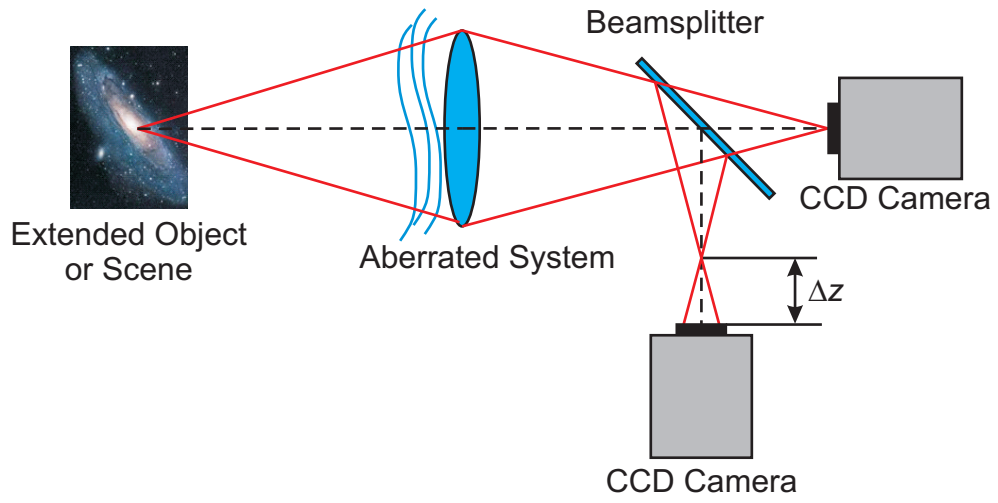


Figure 1-4: Example phase diversity system. In this implementation two detector planes, one of them translated out of focus by a distance Δz, are used to simultaneously capture diversity images. The object being imaged is no longer a point source but rather an extended object.

1.4.2 Concept

Figure 1-4 shows a typical phase diversity system in which an extended object is imaged by an aberrated system onto two different detectors, one in the conventional focal plane and one defocused by some known amount Δz. For this system the unknown parameters consist of two quantities: the object parameters, $\boldsymbol{f}$, which are typically the pixel values $f(u, v)$, and a set of phase parameters, $\boldsymbol{\alpha}$. The phase parameters are generally coefficients of a basis set, a linear sum of which comprises the phase of the system,

$$\phi(x, y; \boldsymbol{\alpha}) = \frac{2\pi}{\lambda} \sum_{j=1}^{J} \alpha_j Z_j(x, y) \tag{1.4}$$

where λ is the wavelength, J is the number of terms in the parameterization and $Z_j(x,y)$ is the j^{th} member of the basis set. The basis set could be a series of delta functions to represent individual pixel values, in which case J could number in the tens or hundreds of thousands, or the basis set could be the Zernike polynomials which may adequately represent the phase of the system with fewer than one hundred terms. The complex pupil of the system is then given by

$$P(x, y; \boldsymbol{\alpha}) = \left|P(x, y)\right| \exp\left[i\phi(x, y; \boldsymbol{\alpha})\right]. \tag{1.5}$$

The complex pupil function is related to the coherent impulse response function by a Fresnel-like transform, given by

$$h_k(u,v;\boldsymbol{\alpha}) = \exp\left[i\frac{\pi D_k}{\lambda B_k}\left(u^2+v^2\right)\right]\iint P(x,y;\boldsymbol{\alpha})\exp\left[i\frac{\pi A_k}{\lambda B_k}\left(x^2+y^2\right)\right] \times\exp\left[-i\frac{2\pi}{\lambda B_k}(xu+yv)\right]dxdy \tag{1.6}$$

where (u, v) are the image plane pixel indices and A_k, B_k, and D_k are the elements of the ABCD ray-transfer matrix that relates the pupil plane to the k^{th} image plane [62]. Defocus is introduced by changing the propagation distance and therefore the elements of the ABCD matrix, principally the A_k/B_k term. The incoherent point spread function (PSF) is then given by

$$s_k(u,v;\boldsymbol{\alpha}) = \left|h_k(u,v;\boldsymbol{\alpha})\right|^2. \tag{1.7}$$

Using the concepts of Fourier optics [63] the unknown parameters $\boldsymbol{f}$ and $\boldsymbol{\alpha}$ can be related to the intensities collected in each image plane by the basic imaging equation,

$$d_k(u,v) = f(u,v) * s_k(u,v;\boldsymbol{\alpha}) + n_k(u,v) \tag{1.8}$$

where $d_k(u, v)$ is the k^{th} detected noisy intensity, $n_k(u, v)$ is an additive noise term, and the asterisk, $*$, denotes a convolution operation. Note the PSF depends on the unknown phase parameters because it is the squared magnitude of the propagated complex pupil. Throughout this thesis, we will use (u, v) to denote coordinates in the image plane, (x, y) to denote coordinates in the pupil plane, (f_u, f_v) to denote spatial frequencies in the Fourier domain counterpart to the image plane and (f_x, f_y) to denote spatial frequencies in the Fourier domain counterpart to the pupil plane.

As shown by Paxman *et al*. [56], the problem of estimating $\boldsymbol{f}$ and $\boldsymbol{\alpha}$ can be cast as a statistical problem that depends on the dominant noise mechanism. The two limiting noise regimes of Poisson photon noise and Gaussian detector noise will be discussed separately.

1.4.3 Poisson Noise Regime

When noise is present, there is a probability associated with detecting a specific value at any given pixel in an image. If the system is photon-limited, then the probability follows a Poisson distribution, given by

$$\mathbf{P}\left[d_k\left(u_o,v_o\right);\boldsymbol{f},\boldsymbol{\alpha}\right]=\frac{g_k\left(u_o,v_o\right)^{d_k(u_o,v_o)}\exp\left[-g_k\left(u_o,v_o\right)\right]}{d_k\left(u_o,v_o\right)!} \tag{1.9}$$

where (u_o,v_o) is a specific pixel in the image array and $g_k(u,v)$ is the "infinite SNR" or noise-free image given by

$$g_k\left(u,v\right)=f\left(u,v\right)*s_k\left(u,v;\boldsymbol{\alpha}\right) \tag{1.10}$$

where $g_k(u, v)$ is in units of detected photo-electrons. The dependence of $g_k(u,v)$ on $\boldsymbol{f}$ and $\boldsymbol{\alpha}$ is dropped for compactness of notation.

Since the noise at individual pixels and of individual images is statistically independent, the total probability associated with detecting a set of K diversity images is given by

$$\mathbf{P}\left[\left\{d_k\left(u,v\right)\right\};\boldsymbol{f},\boldsymbol{\alpha}\right]=\prod_{k=1}^{K}\prod_{u,v}\frac{g_k\left(u,v\right)^{d_k(u,v)}\exp\left[-g_k\left(u,v\right)\right]}{d_k\left(u,v\right)!}. \tag{1.11}$$

Equation (1.11) is a likelihood function that represents the probability of detecting a set of images given $\boldsymbol{f}$ and $\boldsymbol{\alpha}$. In practice, the phase diversity algorithm maximizes the log-likelihood function which is obtained by taking the natural logarithm of Eq. (1.11),

$$\mathbf{L}\left[\left\{d_k(u,v)\right\};\boldsymbol{f},\boldsymbol{\alpha}\right]=\sum_{k=1}^{K}\sum_{u,v}\left\{d_k(u,v)\ln\left[g_k(u,v)\right]-g_k(u,v)-\ln\left[d_k(u,v)!\right]\right\}. \tag{1.12}$$

Similar to the objective functions used in phase retrieval, the log-likelihood function of Eq. (1.12) can be maximized with respect to $\boldsymbol{f}$ and $\boldsymbol{\alpha}$ by a nonlinear optimization routine. Gradient-based searches such as conjugate gradient, Newton, quasi-Newton and Levenberg-Marquardt tend to be the most efficient. To avoid the costly and potentially error-prone computation of finite differences, the analytic gradients of Eq. (1.12) with respect to the unknown parameter sets $\boldsymbol{f}$ and $\boldsymbol{\alpha}$ can be readily calculated.

1.4.4 Gaussian Noise Regime

When the dominant noise mechanism is detector read noise, then the detected images are given by Eq. (1.8) where $n_k(u,v)$ is an additive, zero-mean Gaussian random variable. As with the Poisson case, a probability is associated with detection at a single pixel (u_o,v_o) in the k^{th} image,

$$\mathbf{P}\left[d_k(u_o,v_o);\boldsymbol{f},\boldsymbol{\alpha}\right]=\frac{1}{\left(2\pi\sigma_n^2\right)^{1/2}}\exp\left\{-\frac{\left[d_k(u_o,v_o)-g(u_o,v_o)\right]^2}{2\sigma_n^2}\right\} \tag{1.13}$$

where σ_n^2 is the variance associated with the detector read noise and $g(u_o,v_o)$ is given by Eq. (1.10). Assuming statistical independence between detection of individual images and at individual pixels, the probability associated with detecting an entire set of images is given by

$$\mathbf{P}\left[\left\{d_k(u,v)\right\};\boldsymbol{f},\boldsymbol{\alpha}\right]=\prod_{k=1}^{K}\prod_{u,v}\frac{1}{\left(2\pi\sigma_n^2\right)^{1/2}}\exp\left\{-\frac{\left[d_k(u,v)-g(u,v)\right]^2}{2\sigma_n^2}\right\}. \quad (1.14)$$

Again, the natural logarithm is taken to obtain the log-likelihood function,

$$\begin{aligned}\mathbf{L}\left[\left\{d_k(u,v)\right\};\boldsymbol{f},\boldsymbol{\alpha}\right]=-\sum_{k=1}^{K}\sum_{u,v}\Bigg\{&\frac{1}{2}\ln\left(2\pi\sigma_n^2\right)\\&+\frac{1}{2\sigma_n^2}\left[d_k(u,v)-g(u,v)\right]^2\Bigg\}.\end{aligned} \quad (1.15)$$

Notice that aside from an additive constant and scale factor, Eq. (1.15) resembles a least squares type of metric, similar to those used in phase-diverse phase retrieval. As with the Poisson noise scenario, a nonlinear optimization routine is used to maximize this likelihood function with respect to the unknown parameters.

1.4.5 The Reduced Gaussian Metric

Consider the search space involved in maximizing Eqs. (1.12) and (1.15). If the phase is parameterized as polynomial coefficients and is limited to J terms, and an $M\times N$ array of pixels of the object is estimated, the dimensionality of the search space of the optimization is given by $M\times N+J$. For example, a moderately sized problem including 45 polynomial coefficients and a 256×256 pixel region of the

object requires searching a 65,581-dimensional space, 99.9% of which is due to the estimation of the object.

Gonsalves first proposed a method of reducing the search space for the least squares type of optimization of the Gaussian log-likelihood function with two diversity images [44]. Later, Paxman, *et al.* generalized the method for an arbitrary number of diversity images and showed that the same simplification was not possible for the log-likelihood function of the Poisson noise case [56].

The reduced Gaussian (RG) metric is more easily obtained by re-writing Eq. (1.15) in the Fourier domain by using Parseval's theorem [63],

$$\mathbf{L}(\boldsymbol{f},\boldsymbol{\alpha}) = -\sum_{k=1}^{K}\sum_{f_u,f_v}\left|D_k(f_u,f_v) - F(f_u,f_v)S_k(f_u,f_v;\boldsymbol{\alpha})\right|^2 \qquad (1.16)$$

where an arbitrary constant and scale factor have been dropped, (f_u, f_v) are Fourier domain spatial frequency indices, and D, F and S are the Fourier transforms of d, f and s, respectively. Notice that S is the optical transfer function (OTF) of the system [63] which depends explicitly on $\boldsymbol{\alpha}$, the unknown phase parameters. There exists an object Fourier transform, $F_{\max}(f_u, f_v)$, that maximizes Eq. (1.16) for a given set of phase parameters, $\boldsymbol{\alpha}$. $F_{\max}(f_u,f_v)$ are found analytically by taking the derivative of Eq. (1.16) with respect to the object pixel values and setting the derivative equal to zero. After some algebra, $F_{\max}(f_u, f_v)$ is found to be

$$F_{\max}\left(f_u, f_v\right) = \begin{cases} \dfrac{\sum_{k=1}^{K} D_k\left(f_u, f_v\right) S_k^*\left(f_u, f_v; \boldsymbol{\alpha}\right)}{\sum_{\ell=1}^{K} \left|S_\ell\left(f_u, f_v; \boldsymbol{\alpha}\right)\right|^2}, & \sum_{\ell=1}^{K} \left|S_\ell\left(f_u, f_v; \boldsymbol{\alpha}\right)\right|^2 \neq 0 \\ 0, & \sum_{\ell=1}^{K} \left|S_\ell\left(f_u, f_v; \boldsymbol{\alpha}\right)\right|^2 = 0 \end{cases} \tag{1.17}$$

where the second line accounts for the pixel indices where the solution might diverge by using any value that preserves the Hermitian property of $F_{\max}(f_u, f_v)$. This is similar to a multi-frame inverse filter. Substituting Eq. (1.17) into Eq. (1.16) and simplifying gives the RG metric

$$\mathbf{L}_{RG}\left(\boldsymbol{\alpha}\right) = \sum_{f_u, f_v \in \chi} \frac{\left|\sum_{j=1}^{K} D_j\left(f_u, f_v\right) S_j^*\left(f_u, f_v; \boldsymbol{\alpha}\right)\right|^2}{\sum_{\ell=1}^{K} \left|S_\ell\left(f_u, f_v; \boldsymbol{\alpha}\right)\right|^2} - \sum_{f_u, f_v} \sum_{k=1}^{K} \left|D_k\left(f_u, f_v\right)\right|^2 \tag{1.18}$$

where

$$\chi = \left\{ \left(f_u, f_v\right) \in \sum_{k=1}^{K} \left|S_k\left(f_u, f_v; \boldsymbol{\alpha}\right)\right|^2 \neq 0 \right\}. \tag{1.19}$$

Equation (1.18) represents a simplified likelihood function that depends only on the unknown phase parameters. After determining the optimal phase parameters by nonlinear optimization, Eq. (1.17) or various other image reconstruction methods such as a Wiener filter may be used to compute the unknown object parameters (i.e. an image of the object).

As stated earlier, the RG metric cannot be found for the case of Poisson noise statistics. For that condition, it is not possible to find a closed form solution of

$F_{\max}(f_u, f_v)$, and the object parameters cannot be removed from the optimization. However, for reasonably sized images, the optimization problem is still tractable on a modern desktop computer.

1.4.6 Reduced Gaussian Metric Gradients

The gradients of Eq. (1.18) with respect to the unknown phase coefficients can be readily computed. For a given phase coefficient, α_o, the gradient is given by

$$\frac{\partial \mathbf{L}_{RG}(\boldsymbol{\alpha})}{\partial \alpha_o} = \frac{\partial}{\partial \alpha_o} \sum_{f_u, f_v \in \chi} \frac{\left| \sum_{j=1}^{K} D_j(f_u, f_v) S_j^*(f_u, f_v; \boldsymbol{\alpha}) \right|^2}{\sum_{\ell=1}^{K} \left| S_\ell(f_u, f_v; \boldsymbol{\alpha}) \right|^2} \tag{1.20}$$

where the derivative of the second term in Eq. (1.18) evaluates to zero. Distributing the derivative through and simplifying gives

$$\frac{\partial \mathbf{L}_{RG}(\boldsymbol{\alpha})}{\partial \alpha_o} = \sum_{f_u, f_v \in \chi} \sum_{k=1}^{K} Y_k(f_u, f_v) \frac{\partial S_k(f_u, f_v; \boldsymbol{\alpha})}{\partial \alpha_o} + c.c. \tag{1.21}$$

where *c.c.* denotes the complex conjugate of the preceding terms and $Y_k(f_u, f_v)$ is given by

$$Y_k(f_u, f_v) = \begin{cases} \dfrac{\left(\sum_{j=1}^{K} D_j S_j^* \right) \left(\sum_{\ell=1}^{K} |S_\ell|^2 \right) D_k^* - \left| \sum_{j=1}^{K} D_j S_j^* \right|^2 S_k^*}{\left(\sum_{\ell=1}^{K} |S_\ell|^2 \right)^2}, & f_u, f_v \in \chi \\ 0, & f_u, f_v \notin \chi \end{cases} \tag{1.22}$$

where the dependence on (f_u, f_v) has been dropped for brevity.

In order to evaluate the derivative of the OTF in Eq. (1.21), we must first recall its dependency on the aberration parameters, $\boldsymbol{\alpha}$. The OTF is the Fourier transform of the incoherent point spread function, but can also be written as the autocorrelation of the generalized pupil function [63]

$$S_k\left(f_u, f_v\right) = \sum_{f_u', f_v'} H_k\left(f_u', f_v'\right) H_k^*\left(f_u' - f_u, f_v' - f_v\right). \tag{1.23}$$

The generalized pupil function, $H_k(f_u, f_v)$, is given by

$$H_k\left(f_u, f_v\right) = P_k\left(f_u, f_v\right) \exp\left[i \frac{\pi A_k}{\lambda B_k} \left(f_u^2 + f_v^2 \right) \right] \tag{1.24}$$

where $P_k(f_u, f_v)$ is the complex pupil function given in Eq. (1.5) and A_k and B_k are the elements of the ABCD ray-transfer matrix that relates the pupil plane to the k^{th} image plane [62]. Using Eqs. (1.4) and (1.5), the explicit dependence of the OTF on the unknown phase parameters, $\boldsymbol{\alpha}$, becomes clear.

As detailed in [56], the derivative in Eq. (1.21) can now be evaluated to yield

$$\begin{aligned} \frac{\partial S_k\left(f_u, f_v\right)}{\partial \alpha_o} = \frac{i2\pi}{\lambda} \sum_{f_u', f_v'} Z_o\left(f_u', f_v'\right) \Big[& H_k\left(f_u', f_v'\right) H_k^*\left(f_u' - f_u, f_v' - f_v\right) \\ & - H_k^*\left(f_u', f_v'\right) H_k\left(f_u' + f_u, f_v' + f_v\right) \Big] \end{aligned} \tag{1.25}$$

where Z_o is the basis function corresponding to the phase coefficient α_o. Substituting Eq. (1.25) into Eq. (1.21) and performing a fair bit of algebra yields

$$\begin{aligned} \frac{\partial \mathbf{L}_{RG}}{\partial \alpha_o} = \frac{8\pi}{\lambda} \operatorname{Im} \Bigg\{ & \sum_{k=1}^{K} \sum_{f_u', f_v'} Z_o\left(f_u', f_v'\right) H_k\left(f_u', f_v'\right) \\ & \times \sum_{f_u, f_v \in \chi} Y_k^*\left(f_u, f_v\right) H_k^*\left(f_u' - f_u, f_v' - f_v\right) \Bigg\}. \end{aligned} \tag{1.26}$$

Throughout this thesis, various metrics will be introduced that will require the derivatives to be calculated in a similar way as was done here and in [56]. For each metric, the procedure is the same and the form of the gradient will not change much from Eq. (1.26). Variations to the phase diversity metric appear in the gradient as a variation in the $Y_k(f_u, f_v)$ term. This makes it simple to interchange metrics that are similar in form to Eq. (1.20) in a phase diversity algorithm.

1.5 Outline

This thesis will detail research that has advanced the state-of-the-art in phase diversity algorithms as they pertain to segmented and multi-aperture systems.

In Chapter 2 we will present improvements made to the conventional phase diversity algorithm. A method for extending phase diversity to account for broadband objects will be discussed and compared to a monochromatic algorithm. Various regularization techniques will also be discussed and compared.

Chapter 3 will introduce a new method of implementing phase diversity that is specific to segmented and multi-aperture systems. Making use of the segment or sub-aperture actuation, piston phases can be added to individual sub-apertures to create a known diversity phase. This method is compared to the conventional method of focus diversity.

Chapter 4 will build on the use of sub-aperture piston phase diversity and show how it can be used to estimate object spectral content. The broadband algorithm in Chapter 2 is used extensively.

Chapter 5 will detail laboratory work performed with a segmented MEMs deformable mirror. The conventional monochromatic focus diversity algorithm is tested as a baseline. Results using the broadband focus diversity algorithm and the broadband sub-aperture piston phase diversity algorithm are also presented.

Chapter 6 will present additional work related to phase retrieval for telescopes. A phase-diverse phase retrieval algorithm that accounts for field-dependent aberrations is presented and Cramer-Rao bounds are used to evaluate its performance.

In Chapter 7 we will summarize the thesis and conclude. A plan for future research on these topics will also be presented.

A list of acronyms can be found on page xxix.

The novel elements of this thesis are: (1) experimental demonstration of a broadband phase diversity algorithm, (2) a derivation of and simulation and experimental demonstration of sub-aperture piston phase diversity for segmented and multi-aperture systems, (3) a new method of using phase diversity to estimate object spectral content, (4) an investigation of regularization techniques for phase diversity metrics, and (5) a theoretical and simulation-based comparison of multi-field wavefront sensing using phase-diverse phase retrieval.

1.6 References

1. M. Born, and E. Wolf, *Principles of Optics, 7th Ed.* (Cambridge University Press, Cambridge, UK, 1999).

2. J. N. Aubrun, K. Lorell, T. Mast, and J. Nelson, "Dynamic analysis of the actively controlled segmented mirror of the W. M. Keck ten-meter telescope," Control Systems Magazine, IEEE **7**, 3-10 (1987).

3. L. Stepp, and S. Strom, "The Thirty-Meter Telescope project design and development phase," in *Second Workshop on Extremely Large Telescopes, Proc. SPIE*, A. L. Ardeberg, and T. Andersen, eds., **5382** (2004), p. 67.

4. J. Contreras, and P. Lightsey, "Optical design and analysis of the James Webb Space Telescope: optical telescope element," in *Novel Optical Systems Design and Optimization VII, Proc. SPIE*, J. M. Sasian, R. J. Koshel, P. K. Manhart, and R. C. Juergens, eds., **5524** (2004), p. 30.

5. C. Beichman, D. Coulter, C. Lindersmith, and P. Lawson, "Selected mission architectures for the terrestrial planet finder (TPF): large, medium and small," (2002).

6. M. Colavita, and P. Wizinowich, "Keck Interferometer: progress report," in *Interferometry in Optical Astronomy, Proc. SPIE*, P. J. Lena, and A. Quirrenbach, eds., **4006** (2000), pp. 310-320.

7. D. Stubbs, A. Duncan, J. Pitman, R. Sigler, R. Kendrick, J. Chilese, and E. Smith, "Multiple instrument distributed aperture sensor (MIDAS) science payload concept," in *Optical, Infrared, and Millimeter Space Telescopes, Proc. SPIE*, J. C. Mather, ed. **5487** (2004), p. 1444.

8. A. Michelson, and F. Pease, "Measurement of the Diameter of Alpha-Orionis by the Interferometer," Proceedings of the National Academy of Sciences **7**, 143-146 (1921).

9. L. Mandel, and E. Wolf, *Optical Coherence and Quantum Optics* (Cambridge University Press, Cambridge, UK, 1995).

10. T. Pauls, "Origins of sparse aperture imaging," in *Aerospace Conference, Proc. IEEE*, **3** (2001), pp. 1421-1427.

11. R. Brown, and R. Twiss, "Interferometry of the Intensity Fluctuations in Light. I. Basic Theory: The Correlation between Photons in Coherent Beams of Radiation," Proceedings of the Royal Society of London. Series A, Mathematical and Physical Sciences (1934-1990) **242**, 300-324 (1957).

12. R. Brown, and R. Twiss, "A Test of a New Type of Stellar Interferometer on Sirius," Nature **178**, 1046-1048 (1956).

13. A. Labeyrie, "Interference fringes obtained on Vega with two optical telescopes," SPIE Milestone Series MS **139**, 179-183 (1997).

14. E. Hege, J. Beckers, P. Strittmatter, and D. McCarthy, "Multiple mirror telescope as a phased array telescope," Appl. Opt. **24**, 2565-2576 (1985).

15. M. Lloyd-Hart, R. Dekany, D. Sandler, D. Wittman, R. Angel, and D. McCarthy, "Progress in diffraction-limited imaging at the Multiple Mirror Telescope with adaptive optics," J. Opt. Soc. Am. A **11**, 846-846 (1994).

16. L. Weaver, J. Fender, and C. De Hainaut, "Design considerations for multiple telescope imaging arrays," Optical Engineering **27**, 730-735 (1988).

17. J. Fender, and R. Carreras, "Demonstration of an optically phased telescope array," Optical Engineering **27**, 706-711 (1988).

18. W. M. K. Observatory, http://www.keckobservatory.org/about.php.

19. L. Ramsey, M. Adams, T. Barnes III, J. Booth, M. Cornell, J. Fowler, N. Gaffney, J. Glaspey, J. Good, and G. Hill, "Early performance and present status of the Hobby-Eberly Telescope," in *Advanced Technology Optical/IR Telescopes VI, Proc. SPIE*, L. M. Stepp, ed. **3352** (1998), p. 34.

20. R. Stobie, J. Meiring, and D. Buckley, "Design of the Southern African Large Telescope (SALT)," in *Optical Design, Materials, Fabrication, and Maintenance, Proc. SPIE*, P. Dierickx, ed. **4003** (2000), p. 355.

21. M. Johns, J. Angel, S. Shectman, R. Bernstein, D. Fabricant, P. McCarthy, and M. Phillips, "Status of the Giant Magellan Telescope(GMT) project," in *Ground-based Telescopes, Proc. SPIE*, J. M. Oschmann, ed. **5489** (2004), pp. 441-453.

22. R. Gilmozzi, and J. Spyromilio, "The European Extremely Large Telescope (E-ELT)," The Messenger **127**, 11-19 (2007).

23. P. Harvey, G. Rieke, D. Lester, and D. Benford, "Single aperture far-infrared observatory (SAFIR)," in *IR Space Telescopes and Instruments, Proc. SPIE*, J. C. Mather, ed. **4850** (2003), p. 1097.

24. K. Carpenter, C. Schrijver, R. Lyon, L. Mundy, R. Allen, J. Armstrong, W. Danchi, M. Karovska, J. Marzouk, and L. Mazzuca, "Stellar Imager(SI) mission concept," in *Future EUV/UV and Visible Space Astrophysics Missions and Instrumentation, Proc. SPIE*, J. C. Blades, and O. H. W. Siegmund, eds., **4854** (2003), pp. 293-302.

25. D. Dayton, B. Pierson, B. Spielbusch, and J. Gonglewski, "Atmospheric structure function measurements with a Shack–Hartmann wave-front sensor," Opt. Lett **17**, 1737-1739 (1992).

26. R. Lane, and M. Tallon, "Wave-front reconstruction using a Shack-Hartmann sensor," Appl. Opt. **31**, 6902-6908 (1992).

27. F. Roddier, "Error propagation in a closed-loop adaptive optics system: a comparison between Shack-Hartmann and curvature wave-front sensors," Optics communications **113**, 357-359 (1995).

28. B. Platt, "History and Principles of Shack-Hartmann Wavefront Sensing," J. of Refractive Surgery **17**, 573-577 (2001).

29. D. R. Neal, J. Copland, and D. A. Neal, "Shack-Hartmann wavefront sensor precision and accuracy," in *Advanced Characterization Techniques for Optical, Semiconductor, and Data Storage Components, Proc. SPIE*, A. Duparr, and B. Singh, eds., **4779** (2002), pp. 148-160.

30. J. Fienup, B. Thelen, R. Paxman, and D. Carrara, "Comparison of Phase Diversity and Curvature Wavefront Sensing," in *Adaptive Optical System Technologies, Proc. SPIE*, D. B. R. K. Tyson, ed. **3353** (1998), pp. 930-940.

31. S. Esposito, A. Riccardi, and O. Feeney, "Closed-loop performance of pyramid wavefront sensor," in *Laser Weapons Technology, Proc. SPIE*, T. D. Steiner, and P. H. Merritt, eds., **4034** (2000), pp. 184-189.

32. S. Esposito, and A. Riccardi, "Pyramid Wavefront Sensor behavior in partial correction Adaptive Optic systems," Astronomy & Astrophysics **369**, 9-12 (2001).

33. J. A. Johnson, R. Kupke, D. Gavel, and B. Bauman, "Pyramid wavefront sensing: theory and component technology development at LAO," in *Advances in Adaptive Optics II, Proc. SPIE*, **6272** (2006), pp. 4-11.

34. L. Poyneer, "Scene-based Shack-Hartmann wave-front sensing: analysis and simulation," Appl. Opt. **42**, 5807-5815 (2003).

35. D. Acton, T. Towell, J. Schwenker, J. Swensen, D. Shields, E. Sabatke, L. Klingemann, A. Contos, B. Bauer, and K. Hansen, "Demonstration of the James Webb Space Telescope commissioning on the JWST testbed telescope," in *Space Telescopes and Instrumentation I: Optical, Infrared, and Millimeter, Proc. SPIE*, J. C. Mather, H. A. MacEwen, and M. W. M. de Graauw, eds., **6265** (2006), p. 62650R.

36. A. Contos, D. Acton, P. Atcheson, A. Barto, P. Lightsey, and D. Shields, "Aligning and maintaining the optics for the James Webb Space Telescope (JWST) on-orbit: the wavefront sensing and control concept of operations," in *Space Telescopes and Instrumentation I: Optical, Infrared, and Millimeter, Proc. SPIE*, J. C. Mather, H. A. MacEwen, and M. W. M. de Graauw, eds., **6265** (2006), p. 62650X.

37. B. Dean, D. Aronstein, J. Smith, R. Shiri, and D. Acton, "Phase retrieval algorithm for JWST Flight and Testbed Telescope," in *Space Telescopes and Instrumentation I: Optical, Infrared, and Millimeter, Proc. SPIE*, J. C. Mather, H. A. MacEwen, and M. W. M. d. Graauw, eds., **6265** (2006), p. 626511.

38. R. Gerchberg, and W. Saxton, "Phase determination for image and diffraction plane pictures in the electron microscope," Optik **34**, 275-284 (1971).

39. R. Gerchberg, and W. Saxton, "A practical algorithm for the determination of phase from image and diffraction plane pictures," Optik **35**, 237-246 (1972).

40. D. Misell, "A method for the solution of the phase problem in electron microscopy," J. Phys. D **6**, L6-L9 (1973).

41. D. Misell, "An examination of an iterative method for the solution of the phase problem in optics and electron optics: I. Test calculations," J. Phys. D **6**, 2200-2216 (1973).

42. W. Southwell, "Wave-front analyzer using a maximum likelihood algorithm," J. Opt. Soc. Am **67** (1977).

43. R. Gonsalves, "Phase retrieval from modulus data," J. Opt. Soc. Am **66**, 961–964 (1976).

44. R. Gonsalves, and R. Chidlaw, "Wavefront sensing by phase retrieval," in *Applications of Digital Image Processing III, Proc. SPIE*, **207** (1979).

45. J. Fienup, "Phase-retrieval algorithms for a complicated optical system," Appl. Opt. **32**, 1737-1746 (1993).

46. J. Fienup, "Phase retrieval for the Hubble Space Telescope using iterative propagation algorithms," in *Applications of Digital Image Processing XIV, Proc. SPIE*, A. G. Tescher, ed. **1567** (1991), pp. 327-332.

47. J. Fienup, J. Marron, T. Schulz, and J. Seldin, "Hubble Space Telescope characterized by using phase-retrieval algorithms," Appl. Opt. **32**, 1747-1767 (1993).

48. R. Irwan, and R. Lane, "Phase retrieval with prior information," J. Opt. Soc. Am. A **15**, 2302-2311 (1998).

49. R. Millane, and W. Stroud, "Reconstructing symmetric images from their undersampled Fourier intensities," J. Opt. Soc. Am. A **14**, 568-579 (1997).

50. G. Brady, and J. Fienup, "Phase retrieval as an optical metrology tool," Optifab: Technical digest, SPIE Technical Digest TD03, 139-141 (2005).

51. T. Zielinski, and J. R. Fienup, "Extending Wavefront Sensing Capture Range for Segmented Systems through Tip and Tilt Estimation," in *Frontiers in Optics*, (Optical Society of America, 2006), p. FMF4.

52. T. Zielinski, and J. R. Fienup, "Compensation for Dynamic Optomechanical Vibrations in a Phase-Diverse Phase Retrieval Algorithm," in *Frontiers in Optics*, (Optical Society of America, 2008), p. FMF5.

53. J. Fienup, "Phase retrieval for undersampled broadband images," J. Opt. Soc. Am. A **16**, 1831-1837 (1999).

54. J. Fienup, "Invariant error metrics for image reconstruction," Appl. Opt. **36** (1997).

55. R. G. Paxman, and J. R. Fienup, "Optical misalignment sensing and image reconstruction using phase diversity," J. Opt. Soc. Am. A **5**, 914 (1988).

56. R. G. Paxman, T. J. Schulz, and J. R. Fienup, "Joint estimation of object and aberrations by using phase diversity," J. Opt. Soc. Am. A **9**, 1072 (1992).

57. N. Baba, H. Tomita, and N. Miura, "Iterative reconstruction method in phase-diversity imaging," Appl. Opt. **33** (1994).

58. R. L. Kendrick, D. S. Acton, and A. L. Duncan, "Phase-diversity wave-front sensor for imaging systems," Appl. Opt. **33**, 6533-6546 (1994).

59. J. H. Seldin, R. G. Paxman, V. G. Zarifis, L. Benson, and R. E. Stone, "Closed-loop wavefront sensing for a sparse-aperture, phased-array telescope using broadband phase diversity," in *Imaging Technology and Telescopes, Proc. SPIE*,

J. W. Bilbro, J. B. Breckinridge, R. A. Carreras, S. R. Czyzak, M. J. Eckart, R. D. Fiete, and P. S. Idell, eds., **4091** (2000), pp. 48–63.

60. J. A. Georges III, P. Dorrance, K. Gleichman, J. Jonik, D. Liskow, H. Lapprich, V. Naik, S. Parker, R. Paxman, M. Warmuth, A. Wilson, and T. Zaugg, "High-speed closed-loop dual deformable-mirror phase-diversity testbed," in *Advanced Wavefront Control: Methods, Devices, and Applications V, Proc. SPIE*, **6711** (2007), pp. 05-11.

61. J. S. Smith, B. H. Dean, and S. Haghani, "Distributed computing architecture for image-based wavefront sensing and 2D FFTs," in *Advanced Software and Control for Astronomy, Proc. SPIE*, **6274** (2006), p. 627421.

62. A. E. Siegman, *Lasers* (University Science Books, Sausalito, CA, 1986).

63. J. Goodman, *Introduction to Fourier Optics, 3rd Ed.* (Roberts & Co, 2005).

2 Improvements to the Phase Diversity Algorithm

In this Chapter we will present modifications made to the conventional phase-diversity algorithm. These modifications are not specific to segmented or multi-aperture systems but rather generally advance the state-of-the-art in phase diversity.

In Section 2.1, we will discuss a broadband phase-diversity algorithm first developed by Seldin *et al* [1]. We will present our implementation of the broadband phase-diversity algorithm and compare it the conventional monochromatic equivalent.

In Section 2.2, we will discuss regularization factors in the phase-diversity error metric. While many regularization techniques have been proposed in the literature, we will discuss in detail three such techniques and compare each of them with the conventional phase diversity algorithm.

2.1 Broadband Phase Diversity

2.1.1 Introduction

In reality it is unlikely that an object of interest will be truly monochromatic. Conventional phase diversity methods either assume a monochromatic object or simply ignore the effects due to broadband light. Narrowband spectral filters must be used to enforce the monochromatic constraint for conventional phase diversity, reducing the amount of light available for imaging and therefore reducing the signal-to-noise ratio (SNR). Furthermore, broadband imaging may be part of the science requirement, prohibiting the use of such filters. For example, in Fourier transform

imaging spectroscopy (FTIS), it is the spectral content of the image that is directly of interest [2, 3].

In this section we discuss a broadband phase-diversity algorithm developed by Seldin *et al* [1]. We present our implementation of the algorithm as well as results comparing the broadband algorithm to the conventional monochromatic algorithm.

2.1.2 The Gray-World Approximation

Typical phase-diversity algorithms use nonlinear optimization techniques to match intensities, computed from a model estimate of the system, to the collected data. This is usually accomplished by maximizing a likelihood function depending on the dominant noise mechanism of a system. For example, in the Gaussian noise regime, the log-likelihood function is given in the Fourier domain by

$$\mathbf{L}(\boldsymbol{f},\boldsymbol{\alpha}) = -\sum_{k=1}^{K}\sum_{f_u,f_v}\left|D_k(f_u,f_v) - F(f_u,f_v)\,S_k(f_u,f_v;\boldsymbol{\alpha})\right|^2 \qquad (2.1)$$

where $\boldsymbol{f}$ are the unknown object parameters (typically $N \times N$ object pixel values), $\boldsymbol{\alpha}$ are the unknown phase parameters (typically J polynomial coefficients), K is the number of diversity images, (f_u, f_v) are the Fourier-domain spatial-frequency coordinates, D is the Fourier transform of the detected images, F is the Fourier transform of the current estimate of the object, and S_k is the current estimate of the system optical transfer function (OTF) for the k^{th} diversity image. Equation (2.1) is to be maximized with respect to the unknown quantities $\boldsymbol{f}$ and $\boldsymbol{\alpha}$.

As shown in Section 1.4.5, the search space of the nonlinear optimization procedure can be reduced by removing the object pixel values explicitly from the metric. The reduced Gaussian metric is then given as

$$\mathbf{L}_{RG}(\boldsymbol{\alpha}) = \sum_{f_u, f_v \in \chi} \frac{\left| \sum_{j=1}^{K} D_j(f_u, f_v) S_j^*(f_u, f_v; \boldsymbol{\alpha}) \right|^2}{\sum_{\ell=1}^{K} \left| S_\ell(f_u, f_v; \boldsymbol{\alpha}) \right|^2} - \sum_{u,v} \sum_{k=1}^{K} \left| D_k(f_u, f_v) \right|^2 . \quad (2.2)$$

In the presence of broadband light, Eq. (2.1) becomes

$$\mathbf{L}(\boldsymbol{f}, \boldsymbol{\alpha}) = -\sum_{k=1}^{K} \sum_{f_u, f_v} \left| D_k(f_u, f_v) - \sum_{\lambda} F_\lambda(f_u, f_v) S_{k,\lambda}(f_u, f_v; \boldsymbol{\alpha}) \right|^2 \quad (2.3)$$

where λ indexes the spectral bands of the object. Notice that the unknown object parameters and the unknown phase parameters are coupled by the summation over wavelength. This coupling prevents us from obtaining a closed-form solution of the reduced Gaussian metric. In this case, a simultaneous optimization over the unknown object parameters and unknown phase parameters must be performed.

An alternative is to consider the approximation

$$f_\lambda(u,v) \approx \Phi_\lambda f(u,v), \quad (2.4)$$

known as the gray-world approximation, which assumes that every pixel in the object has the same spectrum, defined by the spectral coefficient Φ_λ. At first the gray-world approximation may seem naive; however, for regions of images of natural scenery, the gray-world approximation proves to be quite robust over moderately large spectral bandwidths. For example, Figure 2-1 shows a portion of an AVIRIS data set

[4]. This data set is truly multi-spectral, with each pixel having a unique spectrum ranging from 495 nm to 1,506 nm. Figure 2-2 shows the average pixel spectrum with the error bars representing a single standard deviation. For moderate bandwidths in this spectrum, the gray-world approximation is fairly robust. This approximation may also be true for some astronomical images.

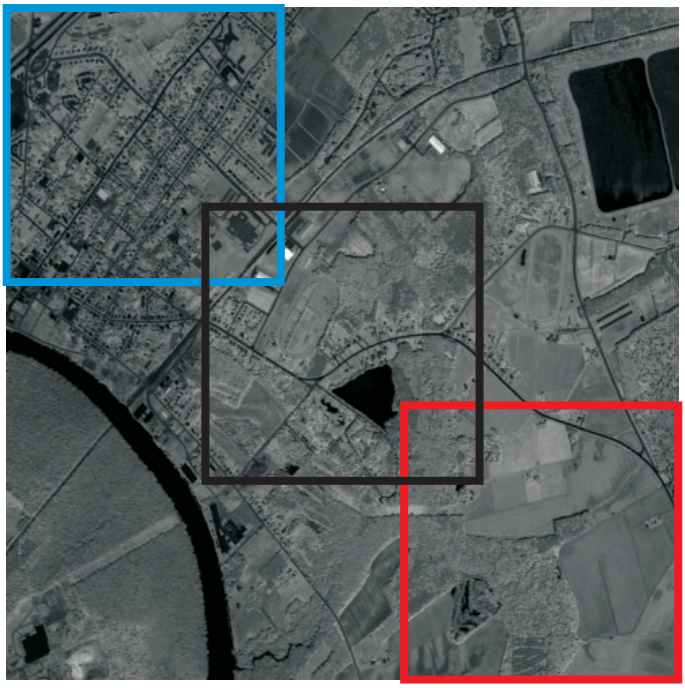


Figure 2-1: A 500 × 500 pixel portion of an AVIRIS data set. The spectral band corresponding to a wavelength of 1 μm is shown. Highlighted regions correspond to the average pixel spectra in Figure 2-2.

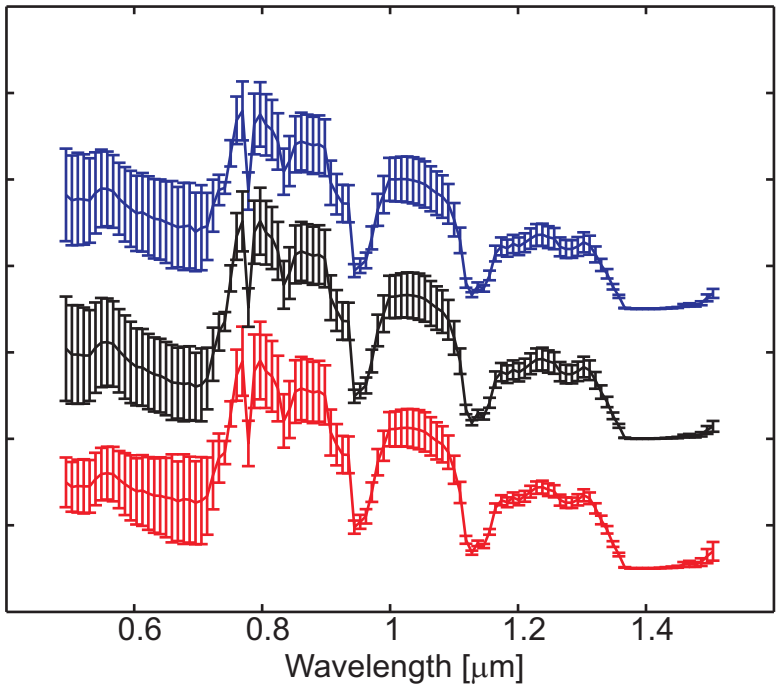


Figure 2-2: The average spectrum over a 1.011 μm bandwidth centered about 1 μm of every pixel in the regions shown in Figure 2-1. Error bars represent a single standard deviation. The average spectra are displaced vertically for clarity.

Under the gray-world approximation, Eq. (2.3) becomes

$$\mathbf{L}(\boldsymbol{f},\boldsymbol{\alpha}) = -\sum_{k=1}^{K}\sum_{f_u,f_v}\left|D_k(f_u,f_v) - F(f_u,f_v)\sum_{\lambda}\Phi_\lambda S_{k,\lambda}(f_u,f_v;\boldsymbol{\alpha})\right|^2, \qquad (2.5)$$

and the unknown object parameters and unknown phase parameters are decoupled, allowing a reduced Gaussian metric to be computed

$$\mathbf{L}_{RG}(\boldsymbol{\alpha}) = \sum_{f_u,f_v\in\chi} \frac{\left|\sum_{j=1}^{K} D_j(f_u,f_v)\sum_{\lambda}\Phi_\lambda S^*_{j,\lambda}(f_u,f_v;\boldsymbol{\alpha})\right|^2}{\sum_{\ell=1}^{K}\left|\sum_{\lambda}\Phi_\lambda S_{\ell,\lambda}(f_u,f_v;\boldsymbol{\alpha})\right|^2 - \sum_{u,v}\sum_{k=1}^{K}\left|D_k(f_u,f_v)\right|^2} \qquad (2.6)$$

where χ restricts the summation over spatial frequencies were the denominator does not equal zero.

The process of calculating analytic gradients of Eq. (2.6) is very similar to that presented in Section 1.4.6 except now the OTF terms, $S_k(f_u, f_v)$, become the gray-world spectral OTFs given by

$$S_k\left(f_u, f_v; \boldsymbol{\alpha}\right) = \sum_{\lambda} \Phi_\lambda S_{k,\lambda}\left(f_u, f_v; \boldsymbol{\alpha}\right), \tag{2.7}$$

and Eq. (1.25) becomes

$$\begin{aligned} \frac{\partial S_k\left(f_u, f_v; \boldsymbol{\alpha}\right)}{\partial \alpha_o} &= \sum_{\lambda} \Phi_\lambda \frac{\partial S_{k,\lambda}\left(f_u, f_v; \boldsymbol{\alpha}\right)}{\partial \alpha_o} \\ &= \sum_{\lambda} \Phi_\lambda \frac{i2\pi}{\lambda} \sum_{f_u', f_v'} Z_o\left(f_u', f_v'\right) \\ &\qquad \times \Big[H_{k,\lambda}\left(f_u', f_v'\right) H_{k,\lambda}^*\left(f_u' - f_u, f_v' - f_v\right) \\ &\qquad\quad - H_{k,\lambda}^*\left(f_u', f_v'\right) H_{k,\lambda}\left(f_u' + f_u, f_v' + f_v\right) \Big] \end{aligned} \tag{2.8}$$

where an additional summation over wavelength is required. Similarly, Eq. (1.26) becomes

$$\begin{aligned} \frac{\partial \mathbf{L}_{RG}}{\partial \alpha_o} = 8\pi \operatorname{Im} \Bigg\{ & \sum_{k=1}^{K} \sum_{\lambda} \frac{\Phi_\lambda}{\lambda} \sum_{f_u', f_v'} Z_o\left(f_u', f_v'\right) H_{k,\lambda}\left(f_u', f_v'\right) \\ & \times \sum_{f_u, f_v \in \chi} Y_k^*\left(f_u, f_v\right) H_{k,\lambda}^*\left(f_u' - f_u, f_v' - f_v\right) \Bigg\} \end{aligned} \tag{2.9}$$

where $Y_k(f_u, f_v)$ is still given by Eq. (1.22) with the exception of the gray-world spectral OTF substitution given by Eq. (2.7). Equation (2.9) is similar to Eq. (25) in [1].

2.1.3 Comparison of Broadband and Monochromatic Phase Diversity

A series of digital simulations was performed to evaluate the performance of conventional phase diversity when a broadband object is present but is assumed to be monochromatic, i.e. when Eq. (2.2) is used as the objective function. The performance results were then compared to the case when the gray-world approximation is made and Eq. (2.6) is used for two scenarios: when the spectral coefficients, Φ_λ, are known, and when the spectral coefficients are assumed to be uniform. The number of spectral coefficients used to sample the bandwidth is also varied.

The object used for this study was taken from a set of AVIRIS data which consists of spectroscopic images with a spectral resolution of 10 nm. Images were simulated for a triarm-9 aperture having 0.22 λ_o RMS of wavefront error. The wavefront error consisted of Zernike terms up to 6^{th} order on the global aperture and up to 4^{th} order on each sub-aperture. Figure 2-3(a) shows the triarm-9 aperture and Figure 2-3(b) shows the wavefront error. Five diversity images were simulated, using [–2, –1, 0, 1, 2] λ_o peak-to-valley (PV) of quadratic focus diversity. Each set of images had a center wavelength of 1 μm and bandwidths between 0 and 1.011 μm.

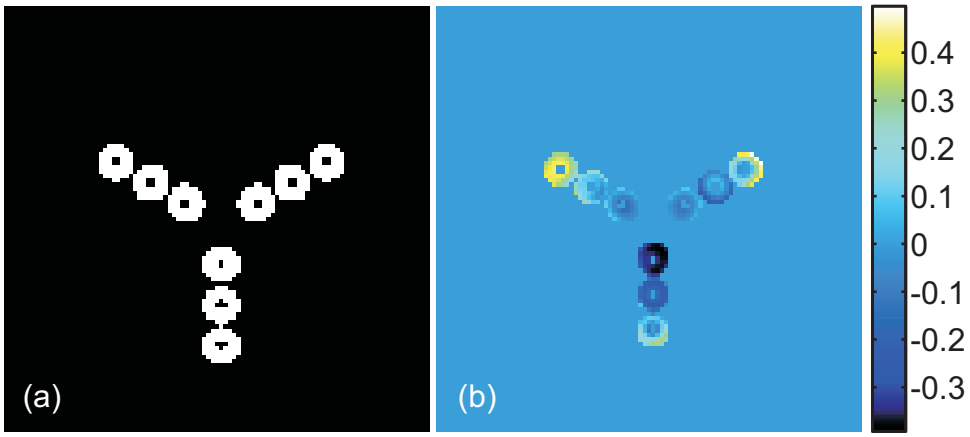


Figure 2-3: (a) Triarm-9 aperture. (b) Wavefront error realization used for image simulation, 0.22 λ_o RMS at a wavelength of 1 μm, consisting of Zernike polynomials up to 6th order over the global aperture and up to 4th order on each sub-aperture. Scale has units of μm.

Realistic noise effects were included in the simulation. The peak pixel in each image was set to 80% of a 50,000 e^- well-depth. Intensity-dependent Poisson noise was included along with zero-mean, additive Gaussian noise with a standard deviation of 15 e^-. The images were then quantized to 12 bits to simulate analog-to-digital conversion.

Strehl ratio was used as a metric to compare the performance of the two methods since it is insensitive to global piston, tip, and tilt phase errors, which the phase diversity algorithm cannot estimate and which do not affect image quality. Given a true wavefront optical path difference (OPD), W_{true}, and the wavefront OPD estimated by the phase-diversity algorithm, W_{est}, we calculate the residual OPD error as $W_{res} = W_{true} - W_{est}$. A pupil function is constructed from the residual OPD error as

$$P_{res}(x,y) = \left|P(x,y)\right| \exp\left[i\frac{2\pi}{\lambda}W_{res}(x,y)\right]. \quad (2.10)$$

If W_{res} is equal to zero, then the pupil function represents a diffraction-limited system. A series of monochromatic impulse response functions is computed by propagating $P_{res}(x, y)$ to the image plane for each wavelength in the bandwidth. The squared-magnitude of the impulse response functions are then summed together to compute $s_{res}(u, v)$, a broadband incoherent point spread function. The Strehl ratio quantifying the residual error is then defined as

$$S.R. = \frac{\max_{u,v}\left[s_{res}(u,v) \right]}{s_{DL}(0,0)} \tag{2.11}$$

where $s_{DL}(u, v)$ is the broadband incoherent PSF of a diffraction-limited system. If W_{res} is equal to zero, then

$$s_{res}(u,v) = s_{DL}(u,v) \tag{2.12}$$

and the residual Strehl ratio is equal to unity. As the error in estimating the wavefront increases, W_{res} becomes non-zero and energy is spread away from the core of the PSF, reducing the maximum value of $s_{res}(u, v)$. Thus as the error in wavefront estimation increases, the residual Strehl ratio decreases. Using Marechal's approximation [5],

$$S.R. \approx \exp\left(-\sigma^2\right) \tag{2.13}$$

where σ is the RMS wavefront error in radians, a Strehl ratio of 0.80 corresponds to an RMS wavefront error of 0.075 λ_o. To accurately evaluate the maximum in Eq. (2.11), the PSFs were upsampled by a factor of 100 using a fast, efficient upsampling algorithm by matrix-multiply DFT [6, 7].

Figure 2-4 shows the results from the first set of simulations in which it was assumed the gray-world spectral coefficients were known. Depending on the bandwidth, between 21 and 111 wavelengths were used to simulate the imagery. However, only three, seven or eleven wavelengths were used to sample the bandwidth in the phase diversity algorithm. These results are compared to using just the center wavelength of 1 μm and assuming the object is monochromatic. The monochromatic algorithm achieved Strehl ratios greater than 0.8 for bandwidths up to 60%, and greater than 0.95 for bandwidths up to 25%. However, using only three wavelengths to sample the bandwidth allowed the phase-diversity algorithm to estimate the phase to a Strehl better than 0.8 for bandwidths of 100%, and a Strehl better than 0.95 for bandwidths up to 50%. For these bandwidths, using more than seven wavelengths to sample the bandwidth did not result in a significant increase in phase estimation accuracy. This behavior would likely depend on the object spectral content, as more complicated spectra would require more coefficients to be adequately represented.

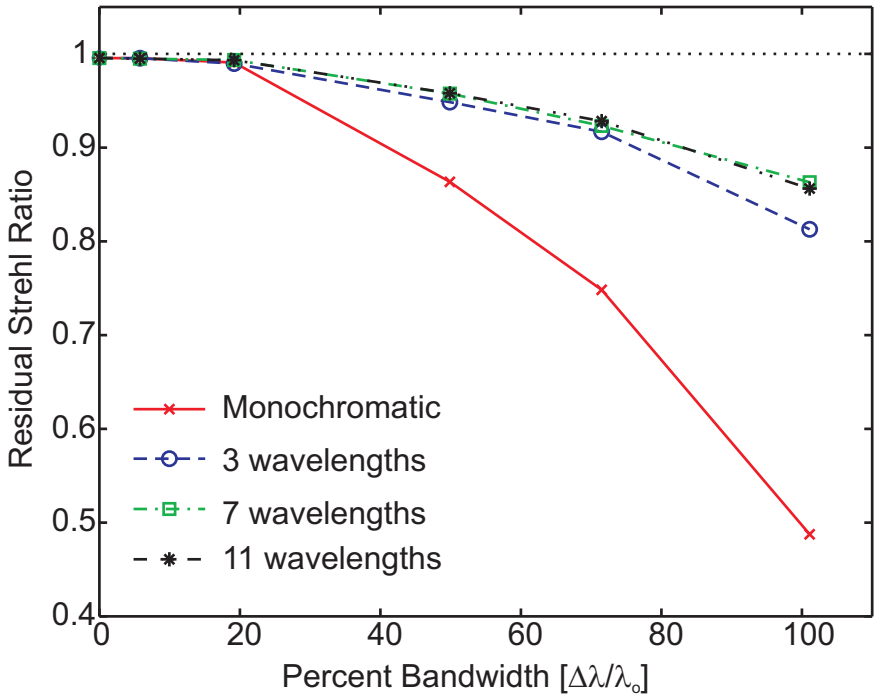


Figure 2-4: Phase estimation error vs. percent bandwidth. For bandwidths exceeding 60%, the monochromatic algorithm performs poorly, while the broadband algorithm can achieve Strehl ratios greater than 0.80 for bandwidths ($\Delta\lambda/\lambda_o$) of 100%. The broadband algorithm used here assumed the gray-world spectral coefficients were known.

The data from Figure 2-4 corresponding to the monochromatic algorithm and the algorithm using seven known coefficients is repeated in Figure 2-5 which also includes the case when the seven coefficients are not known and are assumed to be uniform. As expected, when the gray-world spectral coefficients are known, the best phase estimation results. However, even with no knowledge of the spectral coefficients, Strehl ratios greater than 0.8 can be achieved for bandwidths up to 90%, and Strehl ratios greater than 0.95 can be achieved for bandwidths up to 40%.

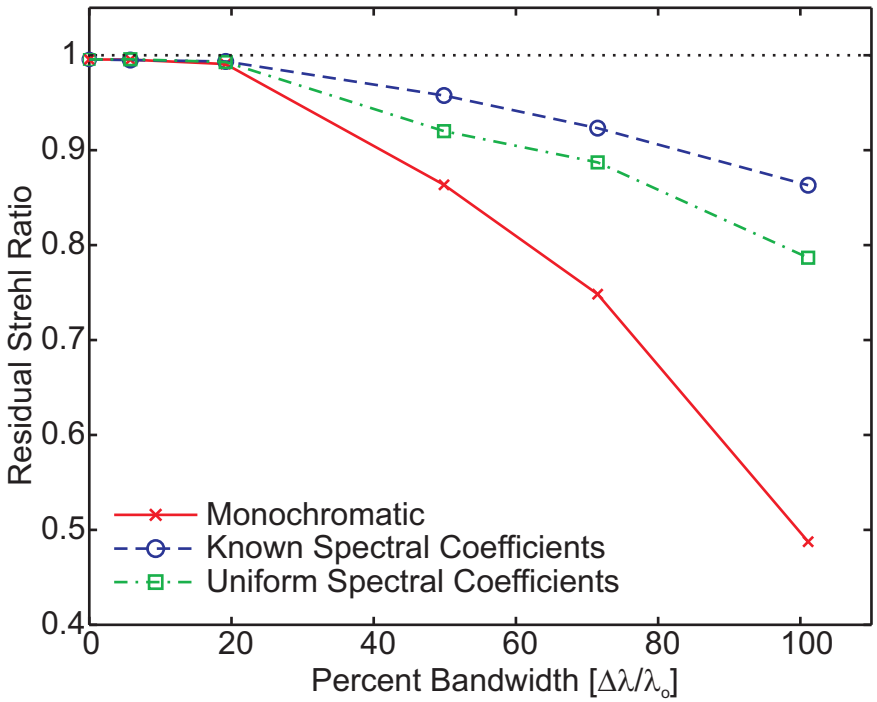


Figure 2-5: Seven spectral coefficients were used to sample the bandwidth. The blue dashed line assumed the spectral coefficients were known, while the green dash-dot line assumed the spectrum was uniform. The monochromatic algorithm is also plotted in red for comparison.

2.1.4 Summary

We have tested a broadband phase-diversity algorithm based on a gray-world assumption on the object being observed. The broadband algorithm was compared to a conventional monochromatic phase-diversity algorithm. The monochromatic algorithm performs moderately well, resulting in an estimation error smaller than 0.075 λ_o RMS for bandwidths up to 60%. However, using only three known spectral coefficients to sample the bandwidth, the broadband algorithm was able to achieve greater accuracy for bandwidths of 100%.

Assuming no knowledge of the spectral coefficients and assuming a uniform spectrum sampled by seven coefficients resulted in a phase estimation error smaller than 0.075 λ_o RMS for bandwidths up to 90%.

2.2 Comparison of Regularization Techniques for Phase Diversity

2.2.1 Introduction

Consider the reduced Gaussian (RG) metric of Eq. (2.2), reprinted here:

$$\mathbf{L}_{RG}(\boldsymbol{\alpha}) = \sum_{f_u, f_v \in \chi} \frac{\left| \sum_{j=1}^{K} D_j(f_u, f_v) S_j^*(f_u, f_v; \boldsymbol{\alpha}) \right|^2}{\sum_{\ell=1}^{K} \left| S_\ell(f_u, f_v; \boldsymbol{\alpha}) \right|^2} - \sum_{u,v} \sum_{k=1}^{K} \left| D_k(f_u, f_v) \right|^2 \quad (2.2)$$

where $\boldsymbol{\alpha}$ is a vector of the unknown phase parameters, (f_u, f_v) are spatial frequency coordinates, K is the number of diversity images, D_k is the Fourier transform of the k^{th} detected image, S_k is the k^{th} optical transfer function (OTF), and χ is the set of spatial frequencies at which the denominator does not equal zero. Limiting the summation over χ acts as a regularization term since it masks regions of the frequency domain where we do not expect there to be any meaningful signal. Any non-zero values outside of the passband of the OTF must be noise, and therefore will not contribute to and likely hinder the estimation of the system phase or object parameters. Furthermore, summing over χ prevents the value of the error metric from diverging for regions within the passband of the OTF where aberrations or aperture sparsity may cause the OTF to be zero-valued.

A limitation of the reduced Gaussian metric is its behavior in situations of low signal-to-noise ratio (SNR). For example, at high spatial frequencies, $D_k(f_u, f_v)$, the Fourier transform of the detected image, will likely have low SNR since both the object Fourier transform and the OTF will have small values there. Since the signal is dominated by noise for these spatial frequencies, it might be desirable to limit, but perhaps not exclude, their contribution to the estimation process with an intelligently chosen spatial-frequency-domain regularization.

Many regularization techniques have been developed in the literature, including methods that define simple weighting functions that act on the frequency data to more sophisticated methods that utilize SNR information via the object and noise power spectra [8-13]. In this section we will present three such SNR-based regularization techniques and compare their performance to the reduced Gaussian metric of Eq. (2.2). The first regularized metric is based on a Wiener-filter model of the object and was suggested to us by Rick Paxman of General Dynamics. The second and third regularized metrics are based on a Bayesian maximum *a posteriori* derivation and were first published by Blanc *et al*. in 2003 [8].

2.2.2 The Multi-frame Wiener Filter Regularization

In Section 1.4.5 it was shown that in order to obtain the RG metric of Eq. (2.2), an inverse-filtered version of the object was substituted into the Gaussian log-likelihood metric. If instead we substitute a Wiener-filtered version of the object, then a regularized reduced Gaussian (RRG) metric can be derived.

Yaroslavsky and Caulfield [14] defined the k^{th} frame of a multi-frame Wiener filter as

$$W_k(f_u, f_v) = \frac{S_k^*(f_u, f_v)\left[\Psi_O(f_u, f_v) \Big/ \Psi_{N,k}(f_u, f_v)\right]}{1 + \sum_{m=1}^{K} \left|S_m(f_u, f_v)\right|^2 \left[\Psi_O(f_u, f_v) \Big/ \Psi_{N,m}(f_u, f_v)\right]} \tag{2.14}$$

where $S_k(f_u, f_v)$ is the k^{th} OTF, K is the number of frames, $\Psi_O(f_u, f_v)$ is the object power spectrum, and $\Psi_{N,k}(f_u, f_v)$ is the noise power spectrum of the k^{th} frame. Using this filter, the reconstructed Fourier data can be computed using the K detected frames as

$$\hat{F}(f_u, f_v) = \sum_{k=1}^{K} D_k(f_u, f_v) W_k(f_u, f_v) \tag{2.15}$$

where $D_k(f_u, f_v)$ is the Fourier transform of the k^{th} detected image.

Assumptions can be made that will simplify Eq. (2.14). First, we assume the noise power spectrum, $\Psi_{N,k}(f_u, f_v)$, does not vary from frame to frame. This assumption is valid if the images are collected by the same detector within a moderately short period of time during which the noise statistics of the system do not vary. Secondly, we assume that the noise power spectrum is uniform with respect to spatial frequency. This assumption is valid because the power spectrum of an uncorrelated random process is uniform. Under these two assumptions, $\Psi_{N,k}(f_u, f_v)$ becomes Ψ_N and Eq. (2.14) can be simplified to

$$W_k(f_u, f_v) = \frac{S_k^*(f_u, f_v)}{c\dfrac{\Psi_N}{\Psi_O(f_u, f_v)} + \sum_{m=1}^{K} \left|S_m(f_u, f_v)\right|^2}. \tag{2.16}$$

where we have also included a constant, c, that can be used to "tune" the filter to emphasize either image sharpening or noise suppression.

Using Eqs. (2.15) and (2.16), $\hat{F}(f_u, f_v)$, an estimate of the object, is formed and substituted into the log-likelihood function of Eq. (2.1). After simplifying, the RRG metric is given by

$$\mathbf{L}_{RRG}(\boldsymbol{\alpha}) = \sum_{f_u, f_v \in \chi} \left\{ \left| \sum_{j=1}^{K} D_j S_j^* \right|^2 \frac{\sum_{m=1}^{K} |S_m|^2 + 2c \frac{\Psi_N}{\Psi_O}}{\left[\sum_{\ell=1}^{K} |S_\ell|^2 + c \frac{\Psi_N}{\Psi_O} \right]^2} \right\} \tag{2.17}$$

where the dependence on (f_u, f_v) has been dropped for brevity. The summation limit over χ is retained because it limits the number of pixels over which the calculations must be performed, although mathematically it has no effect. Inside of the region defined by χ, the noise-to-signal ratio (NSR) defined by (Ψ_N/Ψ_O) appropriately weights the spatial frequencies.

2.2.3 The Joint Maximum *A Posteriori* Metric

Blanc *et al*. derived a metric very similar to the RRG metric using a Bayesian maximum *a posteriori* interpretation of the phase diversity problem [8]. Bayes' theorem defines the probability of an object $\mathbf{f}$ and a set of phase parameters $\boldsymbol{\alpha}$ resulting in a set of detected images, $\{d_k(u, v)\}$, as

$$P\left[\mathbf{f}, \boldsymbol{\alpha} \mid \{d_k(u,v)\}\right] = \frac{P\left[\{d_k(u,v)\} \mid \mathbf{f}, \boldsymbol{\alpha}\right] P[\mathbf{f}] P[\boldsymbol{\alpha}]}{P\left[\{d_k(u,v)\}\right]}. \tag{2.18}$$

We assume $P[\{d_k(u,v)\}]$ is equal to unity. The object and phase can be estimated by maximizing Eq. (2.18) with respect to $\boldsymbol{f}$ and $\boldsymbol{\alpha}$.

Blanc *et al*. make several assumptions on the noise, object and phase statistics and form a log-likelihood function similar to that derived by Paxman *et al*. [15]. An object estimate is calculated by differentiating the log-likelihood function with respect to the object pixel values and setting it equal to zero. The object estimate is then substituted back into the log-likelihood function to find an object-independent metric.

Blanc *et al*. derive their metric for the specific case of $K = 2$ diversity images. We generalize the joint maximum *a posteriori* (JMAP) metric for arbitrary K and drop additive constant and multiplicative scale factors to obtain the JMAP metric

$$\mathbf{L}_{JMAP}(\boldsymbol{\alpha}) = \sum_{f_u, f_v \in \chi} \frac{\left| \sum_{j=1}^{K} D_j S_j^* \right|^2}{\Psi_N \left[\sum_{m=1}^{K} |S_m|^2 + \frac{\Psi_N}{\Psi_O} \right]}. \tag{2.19}$$

2.2.4 The Marginal *A Posteriori* Metric

The second metric developed by Blanc *et al.* is the marginal *a posteriori* metric (mAP) [8]. The mAP metric begins with the same Bayesian *a posteriori* log-likelihood function as the JMAP metric. However, instead of determining the object estimate from the derivative, the object is marginalized, or integrated out of the probability. To marginalize a quantity is to compute a marginal probability law by

summing over all possible values of the quantity. Under these circumstances, Eq. (2.18) becomes

$$P\left[\boldsymbol{\alpha}\,|\,\{d_k(u,v)\}\right] = \int \frac{P\left[\{d_k(u,v)\}\,|\,f,\boldsymbol{\alpha}\right]P[f]P[\boldsymbol{\alpha}]}{P\left[\{d_k(u,v)\}\right]}df\,. \tag{2.20}$$

Using this technique, the object is integrated out of the *a posteriori* log-likelihood and the mAP metric is obtained. Again, we extend the original derivation to account for an arbitrary number of diversity images and drop additive constants:

$$\mathbf{L}_{mAP}(\boldsymbol{\alpha}) = \sum_{f_u,f_v\in\chi} \frac{\left|\sum_{j=1}^{K} D_j S_j^*\right|^2}{\Psi_N\left[\sum_{m=1}^{K}|S_m|^2 + \frac{\Psi_N}{\Psi_O}\right]} + \sum_{f_u,f_v\in\chi} \ln\left\{\sum_{\ell=1}^{K}|S_\ell|^2 + \frac{\Psi_N}{\Psi_O}\right\}. \tag{2.21}$$

Notice Eq. (2.21) is equivalent to the JMAP metric of Eq. (2.19) except for the addition of the natural logarithm term.

2.2.5 Analytic Gradients of New Metrics

As was discussed in Section 1.4.6, the gradients of the RRG, JMAP and mAP metrics are calculated exactly the same way as for the RG metric. A detailed derivation for the case of the RRG metric is given in Appendix A. For each, the final gradient expression is given by

$$\frac{\partial \mathbf{L}_{met}}{\partial \alpha_o} = \frac{8\pi}{\lambda}\,\mathrm{Im}\left\{\sum_{k=1}^{K}\sum_{f_u',f_v'} Z_o(f_u',f_v')H_k(f_u',f_v') \times \sum_{f_u,f_v\in\chi} Y_{met,k}^*(f_u,f_v)H_k^*(f_u'-f_u,f_v'-f_v)\right\} \tag{2.22}$$

where only the $Y_{met,k}(f_u, f_v)$ term changes for each metric. For the case of the RRG metric, the Y-term is given by

$$Y_{RRG,k} = \frac{\left(\sum_{j=1}^{K} D_j S_j^*\right)\left(\sum_{m=1}^{K}|S_m|^2 + 2c\frac{\Psi_N}{\Psi_O}\right)D_k^*}{\left[\sum_{\ell=1}^{K}|S_\ell|^2 + c\frac{\Psi_N}{\Psi_O}\right]^2} + \frac{\left|\sum_{j=1}^{K} D_j S_j^*\right|^2 S_k^*}{\left[\sum_{\ell=1}^{K}|S_\ell|^2 + c\frac{\Psi_N}{\Psi_O}\right]^2} - 2\frac{\left|\sum_{j=1}^{K} D_j S_j^*\right|^2 \left(\sum_{m=1}^{K}|S_m|^2 + 2c\frac{\Psi_N}{\Psi_O}\right)}{\left[\sum_{\ell=1}^{K}|S_\ell|^2 + c\frac{\Psi_N}{\Psi_O}\right]^3}. \tag{2.23}$$

For the JMAP metric, the Y-term is given by

$$Y_{JMAP,k} = \frac{\left(\sum_{j=1}^{K} D_j S_j^*\right)D_k^*}{\left[\sum_{\ell=1}^{K}|S_\ell|^2 + \frac{\Psi_N}{\Psi_O}\right]} - \frac{\left|\sum_{j=1}^{K} D_j S_j^*\right|^2 S_k^*}{\Psi_N\left[\sum_{\ell=1}^{K}|S_\ell|^2 + \frac{\Psi_N}{\Psi_O}\right]^2}. \tag{2.24}$$

Finally, for the mAP metric, the Y-term is given by

$$Y_{mAP,k} = \frac{\left(\sum_{j=1}^{K} D_j S_j^*\right)D_k^*}{\left[\sum_{\ell=1}^{K}|S_\ell|^2 + \frac{\Psi_N}{\Psi_O}\right]} - \frac{\left|\sum_{j=1}^{K} D_j S_j^*\right|^2 S_k^*}{\Psi_N\left[\sum_{\ell=1}^{K}|S_\ell|^2 + \frac{\Psi_N}{\Psi_O}\right]^2} - \frac{\Psi_N S_k^*}{\left[\sum_{\ell=1}^{K}|S_\ell|^2 + \frac{\Psi_N}{\Psi_O}\right]}. \tag{2.25}$$

2.2.6 Estimating the Object and Noise Power Spectra

All three of the regularized metrics require *a priori* knowledge of the object and noise power spectra. However, one of the fundamental assumptions of phase diversity is that the object is not known. Instead the object and noise power spectra are estimated from the detected images, assuming no prior on the object or noise.

We use the same implementation developed by Thurman and Fienup in [16]. The object power spectrum is assumed to be radially symmetric and of the form

$$\Psi_O(f_u, f_v) = \begin{cases} \sqrt{2}A^2 & \rho(f_u, f_v) = 0 \\ A^2 \rho^{-2\gamma} & \rho(f_u, f_v) \neq 0 \end{cases} \tag{2.26}$$

where A and γ are constants and $\rho(f_u, f_v)$ is given by

$$\rho(f_u, f_v) = \left[f_u^2 + f_v^2 \right]^{1/2}. \tag{2.27}$$

The object power spectrum parameters A and γ and the noise power spectrum Ψ_N are then estimated by minimizing the metric

$$E = \sum_{\substack{f_u, f_v \\ \rho(f_u, f_v) \neq 0}} \frac{1}{\rho(f_u, f_v)} \left\{ \frac{\left| D(f_u, f_v) \right|^2}{\left| S(f_u, f_v) \right|^2 \Psi_O(f_u, f_v) + \Psi_N} + \ln\left[\left| S(f_u, f_v) \right|^2 \Psi_O(f_u, f_v) + \Psi_N \right] \right\}. \tag{2.28}$$

For this estimation, only the Fourier transform of the zero-diversity image and the zero-diversity OTF are used.

Since the OTF is required in estimating the object and noise power spectrum, a round-robin method must be used with a phase diversity algorithm that utilizes one of the regularized metrics. Initially, a guess of small random values is made to the

unknown phase parameters and an OTF is computed. The OTF is then used to estimate starting guesses to the object and noise power spectra. These values are then used in one of the regularized metrics in a phase-diversity algorithm to refine the estimate of the phase. After some moderate number of iterations, the refined phase estimate is used to compute a new OTF and another optimization is performed to update the estimation of the object and noise power spectra. The process repeats until the phase estimation converges. One final estimate of the object and noise power spectra is then performed for use in a multi-frame Wiener filter, as in Eq. (2.16), to reconstruct an image of the object.

2.2.7 Comparison of Regularization Techniques

A series of digital simulations was run to compare the three regularized metrics with the conventional RG metric for various levels of SNR. The object used for this study was again taken from a set of AVIRIS data which consists of spectroscopic images with a spectral resolution of 10 nm. Images were simulated for a triarm-9 aperture having 0.19 λ_o RMS of wavefront error. The wavefront error consisted of Zernike terms up to 6^{th} order on the global aperture and up to 2^{nd} order on each sub-aperture. Ten different wavefront realizations were randomly generated. Figure 2-6(a) shows the triarm-9 aperture and Figure 2-6(b) shows a representative wavefront error. Three diversity images were simulated, using [–2, 0, 2] λ_o peak-to-valley (PV) of quadratic focus diversity. Each set of images had a center wavelength of 1 μm and a bandwidth of 19 nm. The broadband phase diversity metric of Section 2.1 was used assuming

three spectral coefficients equal to unity. Figure 2-7 shows a panchromatic version of the object.

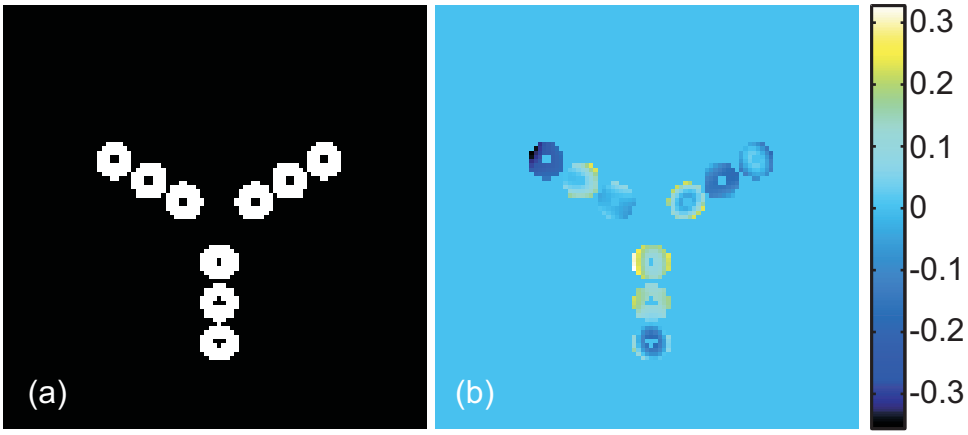


Figure 2-6: (a) A binary Triarm-9 aperture mask, (b) a representative wavefront. Colorbar has units of waves

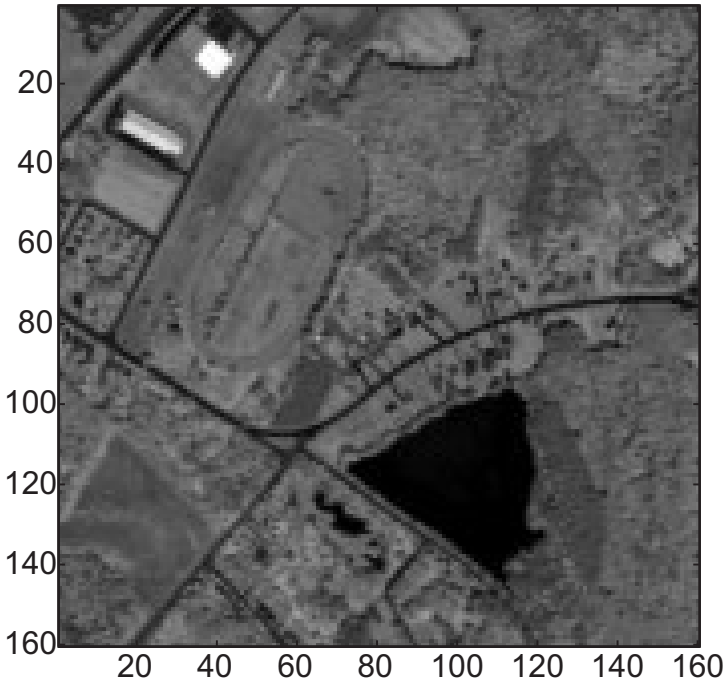


Figure 2-7: Example panchromatic object.

The average pixel SNR was varied by setting the image peak pixel to values between 100 e^- and 50,000 e^-. Intensity-dependent Poisson noise was then added along with zero-mean, additive Gaussian noise with a standard deviation of 15 e^-. A dark current of 0.1 e^-/s was assumed over a 1 second integration time. Table 2-1 shows the peak pixel photon count and corresponding SNRs. For each of the 10 wavefront realizations, 10 independent noise realizations were simulated.

Table 2-1: Number of photons in the peak pixel of each image and the corresponding average pixel SNR.

Peak Pixel Photons (e^-)	100	500	1000	2500	5000	7500	10,000	50,000
Average Pixel SNR	5	17	26	44	63	78	90	204

The average pixel SNR is defined as

$$SNR = \frac{1}{MN} \sum_{u,v} \frac{g(u,v)}{\sqrt{g(u,v) + \sigma_{dark}^2 + \sigma_{read}^2}} \tag{2.29}$$

where M and N are the number of rows and columns in the array, $g(u,v)$ is the noise-free image given by Eq. (1.10), σ_{dark}^2 is the variance associated with the dark current and σ_{read}^2 is the variance associated with the read noise.

We use three criteria to compare the metrics. The first is the Strehl ratio, as defined in Section 2.1.3, Eq. (2.11), which measures the accuracy of the phase estimation. We also consider the convergence properties of each metric. Finally, the number of iterations required before each metric reaches its stopping criteria is also compared. Since the time required to evaluate each metric and its gradients is nearly

identical, the number of iterations is a good measurement of the amount of computation required by each metric.

Figure 2-8 shows the results with respect to phase estimation, plotting the best Strehl ratio of the 100 trials for each metric at each noise level. For low SNR (few photons in the peak pixel), it is clear that the conventional RG metric did poorly. All three of the regularized metrics yielded more accurate phase estimation, with the RRG metric as the most accurate and the JMAP metric nearly equal. Both metrics resulted in a residual Strehl ratio greater than 0.80 which corresponds to an RMS error of 0.075 λ_o.

Figure 2-9 shows a detail of the curve at high SNR. All four metrics performed comparably well, achieving Strehl ratios greater than 0.99 and differing only in the third decimal place. The RG and mAP metric barely outperformed the JMAP and RRG metric.

Figure 2-10 shows a scatter plot of all 100 trials for each metric at each noise level. For low SNR it is clear that convergence of the algorithm is highly variable, with each metric spanning a large range of Strehl ratios. However, both the RRG and JMAP metric occasionally achieved residual Strehl ratios of greater than 0.8 which, using Marechal's approximation [5], corresponds to an RMS wavefront error of 0.075 λ_o. This is commonly thought to be an acceptable level of phase error, although a system with an RMS error of 0.075 λ_o should be measured by something with an estimation error several times better than 0.075 λ_o RMS.

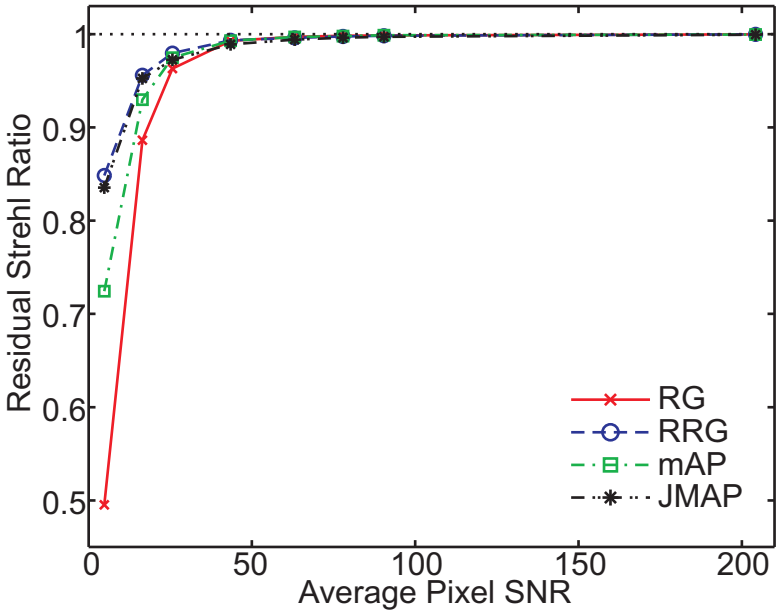


Figure 2-8: Phase estimation accuracy for the four metrics. The best Strehl ratio from 100 trials was plotted for each metric at each noise level. At low SNR (fewer photons in the peak pixel), the conventional RG metric performs poorly. At high SNRs, all four metrics perform comparably well.

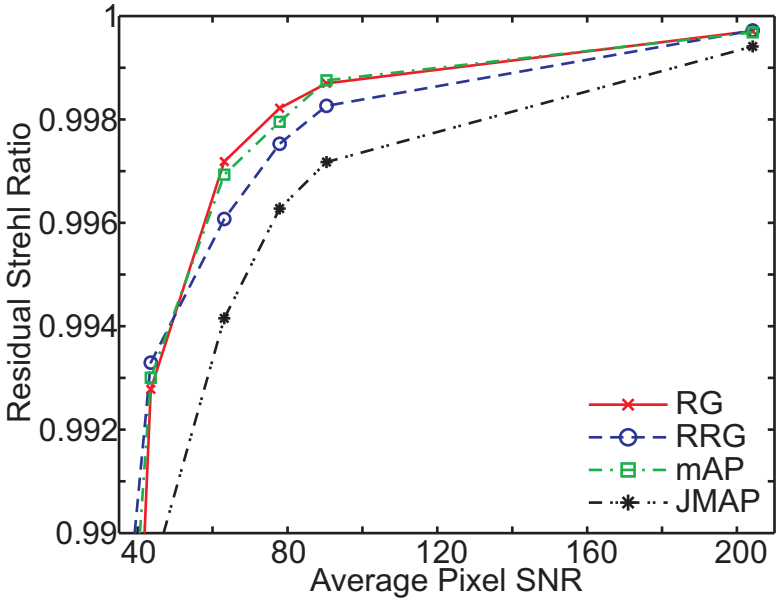


Figure 2-9: Detail of high SNR region of Figure 2-8. While all four metrics achieve Strehl ratios of greater than 0.99, the conventional RG metric and mAP metric achieve the most accurate phase estimations.

For high SNR, all four metrics showed improved convergence with the exception of a few trials of the three regularized metrics. The outlier points correspond to situations in which the estimation of the object and noise power spectra stagnated and became stuck in a local minimum of the search space. When this occurs, unrealistic values are estimated for the object and noise power spectra and consequently the phase-diversity algorithm fails to accurately estimate the phase parameters.

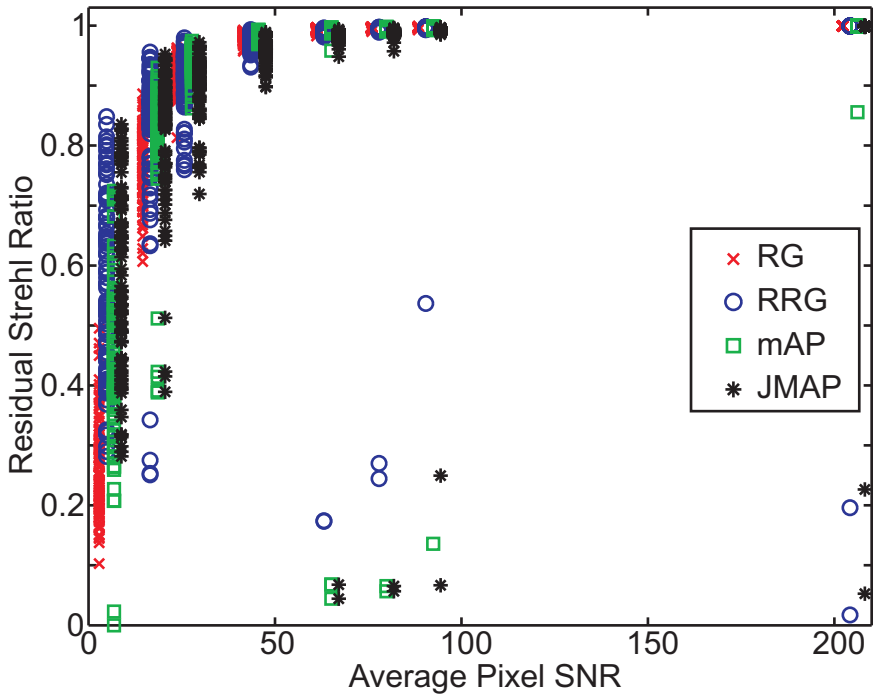


Figure 2-10: Scatter plot of all 100 trials for each metric at each noise level. For low SNR all four metrics converge to a very broad range of Strehl ratios. For high SNR, the range is much more compact, with the exception of instances of the regularized metrics when the object and noise power spectrum estimation failed.

In real world scenarios, one would not have the true wavefront error to compute the residual Strehl ratio. Instead, knowledge of convergence must be obtained from the objective function value computed in the phase-diversity algorithm. Figures 2-11 through 2-14 show scatter plots of the residual Strehl ratio vs. objective function value for 10 different starting guesses for each error metric. In general, lower objective function values correspond to higher residual Strehl ratios, indicating good phase estimation. For the RG, RRG and mAP metrics, the trials corresponding to an SNR of 17 are the only outliers in which the highest residual Strehl ratio does not correspond to the lowest objective function value. We do not understand why these particular trials deviated from the trend.

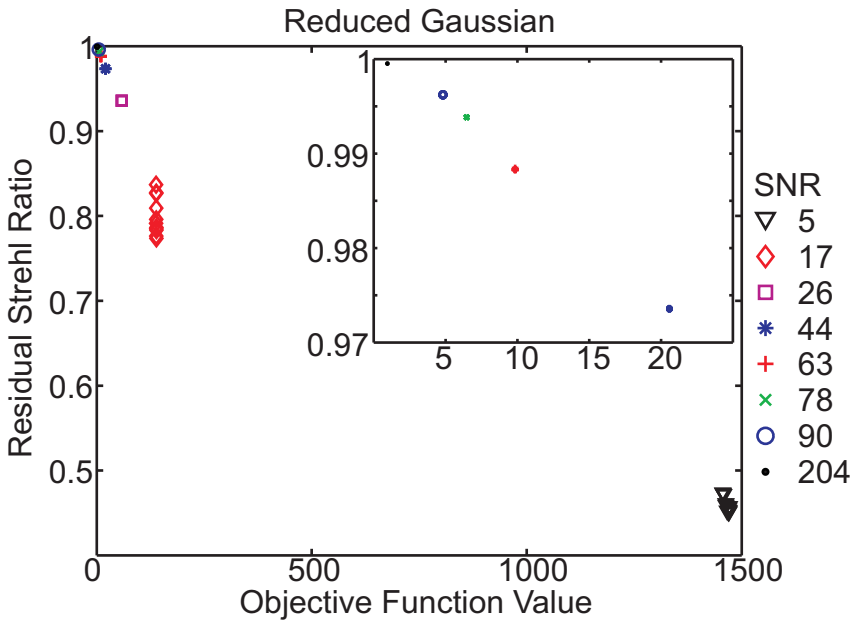


Figure 2-11: Residual Strehl ratio vs. objective function value for 10 random starting guesses of the reduced Gaussian metric. The inset shows a detail of the high residual Strehl ratio regime. Except for the SNR level of 17, a lower objective function value corresponds to a higher residual Strehl ratio.

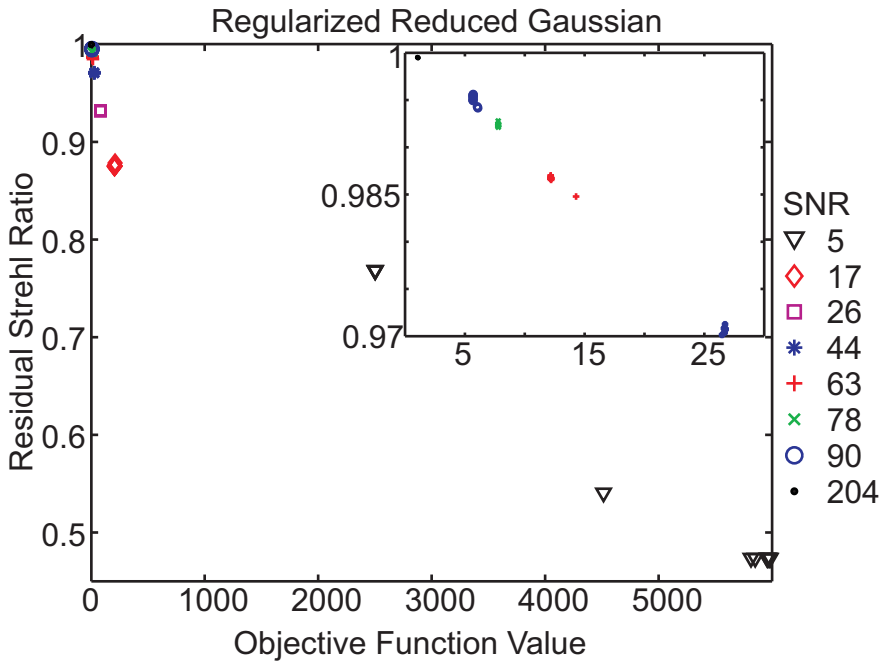


Figure 2-12: Residual Strehl ratio vs. objective function value for 10 random starting guesses of the regularized reduced Gaussian metric. The inset shows a detail of the high residual Strehl ratio regime. Except for the SNR level of 17, a lower objective function value corresponds to a higher residual Strehl ratio.

For both the mAP and JMAP metrics the highest SNR trials obtain high residual Strehl ratio values for a broad range of objective function values. Therefore, even if we do not select the lowest objective function value, we are still assured good phase estimation.

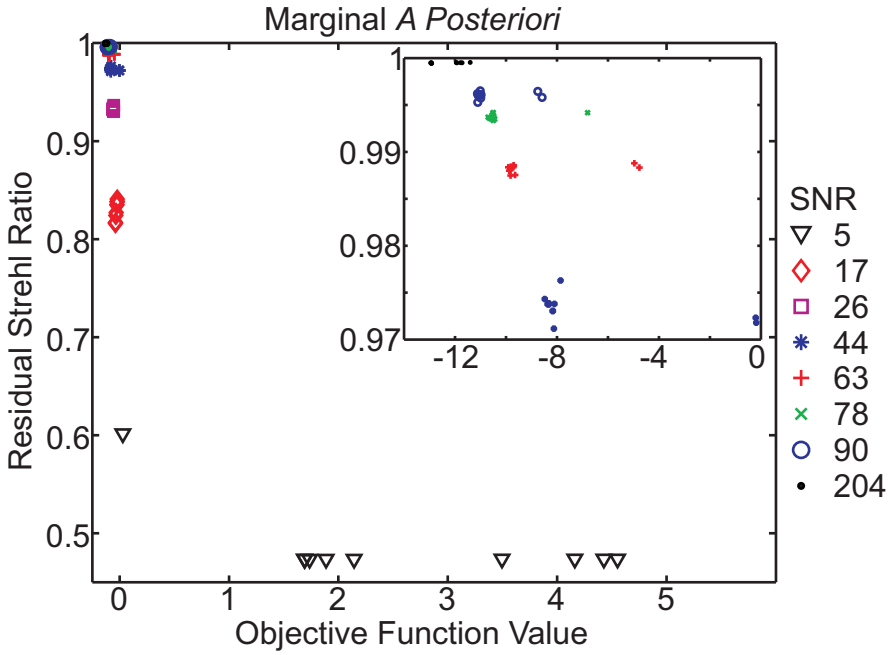


Figure 2-13: Residual Strehl ratio vs. objective function value for 10 random starting guesses of the marginal *a posteriori* metric. The inset shows a detail of the high residual Strehl ratio regime. The lowest objective function value does not necessarily correspond to the highest residual Strehl ratio; however, for high SNR there is little change in the residual Strehl ratio for a range of objective function values.

Figure 2-15 shows the average number of iterations required for each metric at each noise level. This does not include the additional optimization required to estimate the object and noise power spectra. Error bars correspond to a single standard deviation. From this plot it is clear that the number of iterations required before the algorithm reaches its exit criteria is highly erratic. However, the conventional RG metric uniformly required fewer iterations to converge.

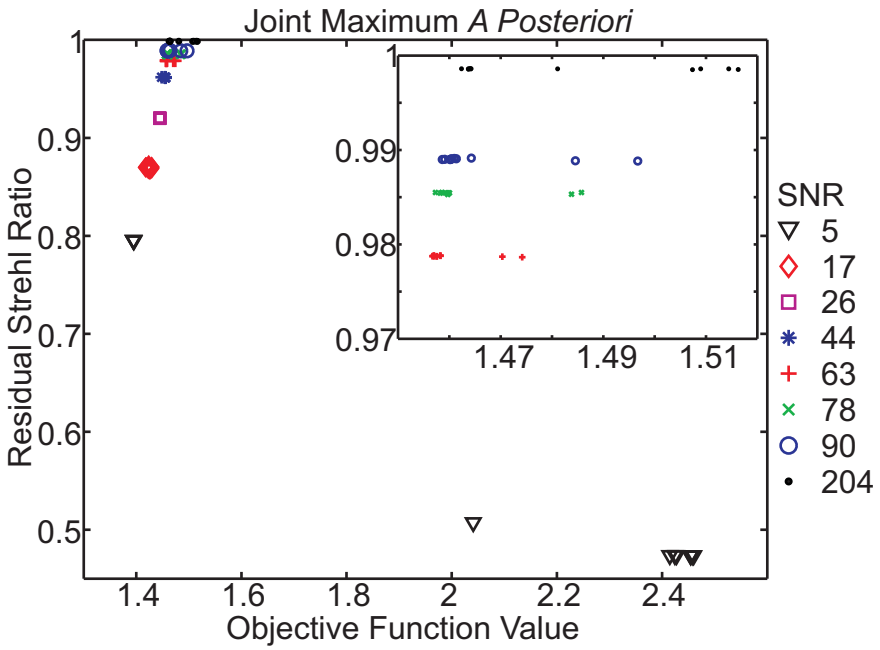


Figure 2-14: Residual Strehl ratio vs. objective function value for 10 random starting guesses of the joint maximum *a posteriori* metric. The inset shows a detail of the high residual Strehl ratio regime. The lowest objective function value does not necessarily correspond to the highest residual Strehl ratio; however, for high SNR there is little change in the residual Strehl ratio for a range of objective function values.

2.2.8 Summary

Four phase-diversity metrics were compared. The conventional RG metric and three regularized metrics were derived and analytic gradients of each were computed. The RRG metric was based on a similar derivation as the RG metric, except a Wiener-filtered version of the object was used to regularize and reduce the metric. The JMAP and mAP metric were both derived from a Bayesian maximum *a posteriori* metric by Blanc *et al* for the case of $K = 2$ diversity images and extended here to an arbitrary number of diversity images. For all three of the new metrics the object and noise power spectra were estimated from the detected images.

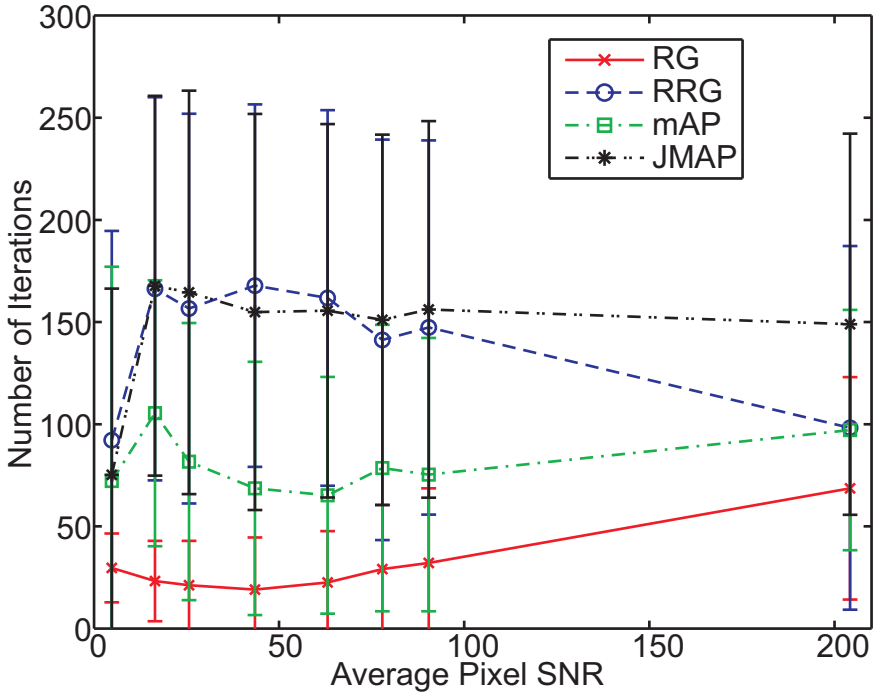


Figure 2-15: Number of iterations required before the metric reached exit criteria. Uniformly, the RG metric required fewer iterations than the three regularized metrics. These numbers do not reflect the additional optimization required to estimate the object and noise power spectra. Error bars represent a single standard deviation.

Digital simulations showed that while the RG metric uniformly required fewer iterations to converge, the RRG metric achieved the most accurate phase estimation at low SNR. For high SNR, all four metrics achieved comparable accuracy in phase estimation. We conclude that for high SNR data, the RG metric should be used as it requires less computation. However, for low SNR data, the RRG metric will more likely yield a better phase estimate.

2.3 Appendix A: Simulating Broadband Images

Some care must be taken in digitally modeling a broadband image due to the dependence of the size of pixels in the pupil on wavelength. Equations (2.30) and (2.31) show the discrete Fourier transform (DFT) as defined by MATLAB® and a discrete optical transform, respectively:

$$F(p,q) = \sum_{m,n=0}^{M-1,N-1} f(m,n) \exp\left[-i2\pi\left(\frac{mp}{M} + \frac{nq}{N}\right)\right], \tag{2.30}$$

$$F(p,q) = \sum_{m,n=0}^{M-1,N-1} f(m,n) \exp\left[-i\frac{2\pi}{\lambda z}\left(m\Delta_x p\Delta_u + n\Delta_y q\Delta_v\right)\right], \tag{2.31}$$

where (p, q) and (m, n) are pixel indices, M and N are the number of rows and columns in the array, Δ_u and Δ_v are the pixel pitches in the (u, v) plane (image), Δ_x and Δ_y are the pixel pitches in the (x, y) plane (pupil), λ is the wavelength and z is the propagation distance. The pixel pitch is defined to be the distance between pixel centers. Comparing the Fourier kernels, it is clear that

$$\begin{aligned} M &= \frac{\lambda z}{\Delta_x \Delta_u} \\ N &= \frac{\lambda z}{\Delta_y \Delta_v} \end{aligned}. \tag{2.32}$$

Of these quantities, the sample spacing in the image plane is fixed by the physical detector pixel size, λ is fixed by the spectral bandpass of the system and z is fixed by the optical system. Therefore, either the array sizes, M and N, or the pupil pixel dimensions, Δ_x and Δ_y, are free to be chosen. Re-arranging Eqs. (2.32) yields

$$\Delta_x = \frac{\lambda z}{M\Delta_u}$$
$$\Delta_y = \frac{\lambda z}{N\Delta_v} \tag{2.33}$$

which show that the digital sampling of the pupil function changes with wavelength. Longer wavelengths "see" larger pixels in the pupil plane. Figure 2-16 shows an example of this effect. In Figure 2-16(a), a cut through a quadratic phase function is shown plotted vs. pixel index for three wavelengths. Because the red pixels are larger than the blue pixels, it takes fewer of them to represent the pupil and a coarser sampling is the result. In Figure 2-16(b), the same quadratic function is plotted vs. pupil coordinate in physical units of meters. As expected, the physical size the pupil does not change with wavelength. Also apparent in Figure 2-16 is the effect that the values of the phase scale with wavelength as well. This effect is clear from the form of the complex pupil function in Eqs. (1.4) and (1.5) in which the phase in radians is given by $(2\pi/\lambda)$OPL, where OPL is the optical path length.

There are two ways to model the effect of pupil sample size scaling with wavelength [17]. The first is to use Eqs. (2.33) to calculate pixel pitch in the pupil for each wavelength. Then defining the aperture size in real units (e.g. radius = 1 m, etc.) on each coordinate set will determine the number of pixels across the aperture at each wavelength. The optical transform can then be performed by using the MATLAB® DFT in Eq. (2.30) and the resulting image plane PSF will have the appropriate sampling for each wavelength. While this method is very easily implemented, a drawback is that a copy of the pupil must be held in memory for every wavelength

being modeled, including all pupil quantities such as apertures and sub-apertures, basis functions, apodizations, etc. For large array sizes and a large number of wavelengths, computer memory can quickly become a limitation.

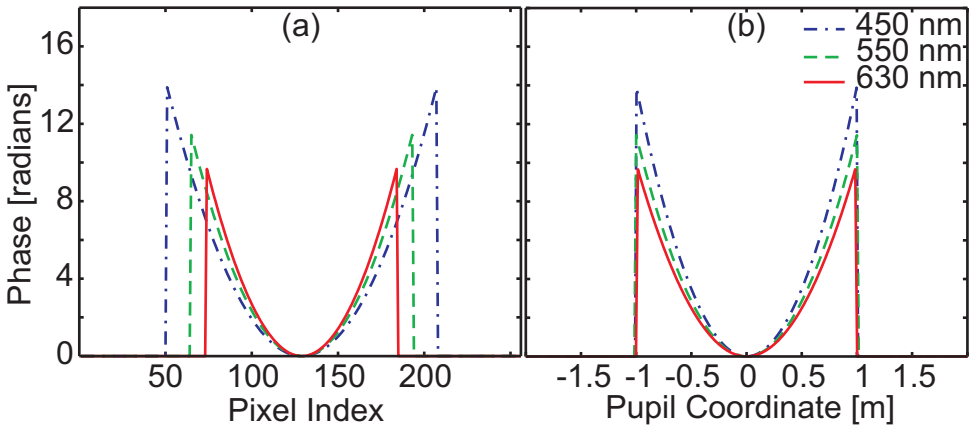


Figure 2-16: (a) A cut through a 1 μm PV quadratic phase vs. pixel index for each of three wavelengths. Longer wavelengths see larger pixels and therefore require fewer pixels to sample the pupil. The amplitude of the phase error also scales with wavelength. (b) When plotted vs. pupil coordinate it is clear that the physical aperture size does not change with wavelength.

The second method to account for the wavelength scaling is to use Eqs. (2.32) and scale the array size. In this method, the DFT of Eq. (2.30) is only performed over the number of pixels, M and N, as determined by Eqs. (2.32). The advantage to this method is that only a single copy of the pupil quantities needs to be held in memory. The arrays are zero-padded or cropped appropriately before performing the DFT to propagate the pupil plane to the image plane. The array is then cropped or zero-padded back to the desired array size. If M_o and N_o are the array dimensions for

reference wavelength λ_o, then the pad-size or crop-size for wavelength λ_ℓ can be derived from Eqs. (2.32):

$$\begin{aligned} \Delta_x \Delta_u &= \frac{\lambda_o z}{M_o} = \frac{\lambda_\ell z}{M_\ell} \\ \therefore M_\ell &= M_o \frac{\lambda_\ell}{\lambda_o}. \end{aligned} \tag{2.34}$$

We note that Eq. (2.34) is different from Eq. (4) in [17] which is in error. If calculating M_ℓ results in a non-integer array size, it is rounded to the nearest integer and λ_ℓ is recalculated. For typical array sizes, this process results in a very small change in λ_ℓ.

A disadvantage of this method is that the calculated array sizes may be prime or non-highly-composite numbers which may hinder the efficiency of calculating the DFT using a fast Fourier transform (FFT). However, on modern desktop computers using recent versions of MATLAB®, this effect is relatively small. Furthermore, if uniform spectral sampling is not required, ideal FFT lengths can be chosen and the modeled wavelengths can be calculated accordingly.

Using the second method of scaling the array size, we take the following steps to simulate a broadband image:

1) The complex field in the pupil is defined according to Eqs. (1.4) and (1.5).
2) The pupil array is zero-padded or cropped appropriately for the wavelength λ_ℓ according to Eq. (2.34).

3) The complex field in the pupil is propagated to the image plane using an FFT.
4) The resulting coherent impulse response is cropped or zero-padded back to the original array size.
5) The incoherent PSF is calculated by squaring the magnitude of the coherent impulse response.
6) The optical transfer function (OTF) is calculated by performing an FFT on the incoherent PSF.
7) The Fourier transform of the object at wavelength λ_ℓ is multiplied by the OTF at the same wavelength.
8) The product of step 7 is transformed using an inverse-FFT (IFFT). This is the noise-free monochromatic image at wavelength λ_ℓ.
9) The images for each wavelength are summed.
10) Noise and detector effects are added.

The noise and detector effects added in step 10 may include intensity dependent Poisson noise, additive Gaussian read noise, dark current, detector bias and non-uniformity, quantization due to analog-to-digital conversion, charge diffusion, integration due to finite pixel size, line-of-sight jitter and others.

2.4 Appendix B: Analytic Gradients of the RRG Metric

In this Appendix, we will provide a derivation of the analytic gradients for the RRG metric with respect to the phase parameters. The derivative with respect to phase parameter α_o is given by

$$\frac{\partial \mathbf{L}_{RRG}(\boldsymbol{\alpha})}{\partial \alpha_o} = \frac{\partial}{\partial \alpha_o} \sum_{f_u, f_v \in \chi} \left[\left| \sum_{j=1}^{K} D_j S_j^* \right|^2 \frac{\sum_{m=1}^{K} |S_m|^2 + 2c \frac{\Psi_N}{\Psi_O}}{\left(\sum_{\ell=1}^{K} |S_\ell|^2 + c \frac{\Psi_N}{\Psi_O} \right)^2} \right]. \tag{2.35}$$

Expanding the squared magnitude gives

$$\frac{\partial \mathbf{L}_{RRG}(\boldsymbol{\alpha})}{\partial \alpha_o} = \frac{\partial}{\partial \alpha_o} \sum_{f_u, f_v \in \chi} \left[\frac{\left(\sum_{k=1}^{K} D_k S_k^* \right) \left(\sum_{j=1}^{K} D_j^* S_j \right) \left(\sum_{m=1}^{K} |S_m|^2 + 2c \frac{\Psi_N}{\Psi_O} \right)}{\left(\sum_{\ell=1}^{K} |S_\ell|^2 + c \frac{\Psi_N}{\Psi_O} \right)^2} \right]. \tag{2.36}$$

Using the quotient rule gives

$$\frac{\partial \mathbf{L}_{RRG}(\boldsymbol{\alpha})}{\partial \alpha_o} = \sum_{f_u, f_v \in \chi} \left(\sum_{\ell=1}^{K} |S_\ell|^2 + c\frac{\Psi_N}{\Psi_O} \right)^{-4}$$
$$\times \left\{ \left[\left(\sum_{k=1}^{K} D_k \frac{\partial S_k^*}{\partial \alpha_o} \right) \left(\sum_{j=1}^{K} D_j^* S_j \right) \left(\sum_{m=1}^{K} |S_m|^2 + 2c\frac{\Psi_N}{\Psi_O} \right) \right.\right.$$
$$+ \left(\sum_{k=1}^{K} D_k S_k^* \right) \left(\sum_{j=1}^{K} D_j^* \frac{\partial S_j}{\partial \alpha_o} \right) \left(\sum_{m=1}^{K} |S_m|^2 + 2c\frac{\Psi_N}{\Psi_O} \right)$$
$$\left. + \left| \sum_{k=1}^{K} D_k S_k^* \right|^2 \left(\sum_{m=1}^{K} \frac{\partial |S_m|^2}{\partial \alpha_o} \right) \right] \left(\sum_{m=1}^{K} |S_m|^2 + c\frac{\Psi_N}{\Psi_O} \right)^2$$
$$-2 \left| \sum_{k=1}^{K} D_k S_k^* \right|^2 \left(\sum_{m=1}^{K} |S_m|^2 + 2c\frac{\Psi_N}{\Psi_O} \right)$$
$$\left. \times \left(\sum_{m=1}^{K} |S_m|^2 + c\frac{\Psi_N}{\Psi_O} \right) \sum_{m=1}^{K} \frac{\partial |S_m|^2}{\partial \alpha_o} \right\} \tag{2.37}$$

where

$$\frac{\partial |S_m|^2}{\partial \alpha_o} = S_m^* \frac{\partial S_m}{\partial \alpha_o} + S_m \frac{\partial S_m^*}{\partial \alpha_o}. \tag{2.38}$$

Substituting Eq. (2.38) into Eq. (2.37) and collecting terms gives

$$\frac{\partial \mathbf{L}_{RRG}(\boldsymbol{\alpha})}{\partial \alpha_o} = \sum_{f_u, f_v \in \chi} \left(\sum_{\ell=1}^{K} |S_\ell|^2 + c\frac{\Psi_N}{\Psi_O} \right)^{-4} \times \left\{ \left[\left(\sum_{j=1}^{K} D_j S_j^* \right) \left(\sum_{m=1}^{K} |S_m|^2 + 2c\frac{\Psi_N}{\Psi_O} \right) \sum_{k=1}^{K} D_k^* \frac{\partial S_k}{\partial \alpha_o} + \left| \sum_{j=1}^{K} D_j S_j^* \right|^2 \sum_{k=1}^{K} S_k^* \frac{\partial S_k}{\partial \alpha_o} \right] \left(\sum_{m=1}^{K} |S_m|^2 + c\frac{\Psi_N}{\Psi_O} \right)^2 - 2 \left| \sum_{j=1}^{K} D_j S_j^* \right|^2 \left(\sum_{m=1}^{K} |S_m|^2 + 2c\frac{\Psi_N}{\Psi_O} \right) \times \left(\sum_{m=1}^{K} |S_m|^2 + c\frac{\Psi_N}{\Psi_O} \right) \sum_{k=1}^{K} S_k^* \frac{\partial S_k}{\partial \alpha_o} \right\} + c.c. \tag{2.39}$$

where *c.c.* denotes the complex conjugate of the preceding term. Let

$$Y_{RRG,k} = \frac{\left(\sum_{j=1}^{K} D_j S_j^* \right) \left(\sum_{m=1}^{K} |S_m|^2 + 2c\frac{\Psi_N}{\Psi_O} \right) D_k^*}{\left[\sum_{\ell=1}^{K} |S_\ell|^2 + c\frac{\Psi_N}{\Psi_O} \right]^2} + \frac{\left| \sum_{j=1}^{K} D_j S_j^* \right|^2 S_k^*}{\left[\sum_{\ell=1}^{K} |S_\ell|^2 + c\frac{\Psi_N}{\Psi_O} \right]^2} - 2\frac{\left| \sum_{j=1}^{K} D_j S_j^* \right|^2 \left(\sum_{m=1}^{K} |S_m|^2 + 2c\frac{\Psi_N}{\Psi_O} \right)}{\left[\sum_{\ell=1}^{K} |S_\ell|^2 + c\frac{\Psi_N}{\Psi_O} \right]^3} \tag{2.40}$$

and Eq. (2.39) can be written as

$$\frac{\partial \mathbf{L}_{RRG}(\boldsymbol{\alpha})}{\partial \alpha_o} = \sum_{f_u, f_v \in \chi} \sum_{k=1}^{K} Y_{RRG,k}(f_u, f_v) \frac{\partial S_k(f_u, f_v)}{\partial \alpha_o} + c.c. \tag{2.41}$$

Using Eq. (2.41), the final expression for the derivative can be computed as described in Section 1.4.6. Furthermore, the analytic gradients for the JMAP and mAP metrics can be calculated using the same process here.

2.5 References

1. J. H. Seldin, R. G. Paxman, V. G. Zarifis, L. Benson, and R. E. Stone, "Closed-loop wavefront sensing for a sparse-aperture, phased-array telescope using broadband phase diversity," in *Imaging Technology and Telescopes, Proc. SPIE*, J. W. Bilbro, J. B. Breckinridge, R. A. Carreras, S. R. Czyzak, M. J. Eckart, R. D. Fiete, and P. S. Idell, eds., **4091** (2000), pp. 48–63.

2. S. Thurman, and J. Fienup, "Multi-aperture Fourier transform imaging spectroscopy: theory and imaging properties," Opt. Express **13**, 2160-2175 (2005).

3. D. A. Naylor, B. G. Gom, M. K. Tahic, and G. R. Davis, "Astronomical spectroscopy using an aliased step-and-integrate Fourier transform spectrometer," in *Millimeter and Submillimeter Detectors for Astronomy II, Proc. SPIE*, J. Zmuidzinas, W. S. Holland, and S. Withington, eds., **5498** (2004), pp. 685-694.

4. Provided through the courtesy of Jet Propulsion Laboratory, California Institute of Technology, Pasadena, California (http://aviris.jpl.nasa.gov/).

5. G. Chanan, and M. Troy, "Strehl ratio and modulation transfer function for segmented mirror telescopes as functions of segment phase error," Appl. Opt. **38** (1999).

6. M. Guizar-Sicairos, S. T. Thurman, and J. R. Fienup, "Efficient subpixel image registration algorithms," Opt. Lett. **33**, 156-158 (2008).

7. R. Soummer, L. Pueyo, A. Sivaramakrishnan, and R. J. Vanderbei, "Fast computation of Lyot-style coronagraph propagation," Opt. Express **15**, 15935-15951 (2007).

8. A. Blanc, L. M. Mugnier, and J. Idier, "Marginal estimation of aberrations and image restoration by use of phase diversity," J. Opt. Soc. Am. A **20**, 1035-1045 (2003).

9. O. M. Bucci, A. Capozzoli, and G. D'Elia, "Regularizing strategy for image restoration and wave-front sensing by phase diversity," J. Opt. Soc. Am. A **16**, 1759-1768 (1999).

10. R. L. Kendrick, D. S. Acton, and A. L. Duncan, "Phase-diversity wave-front sensor for imaging systems," Appl. Opt. **33**, 6533-6546 (1994).

11. D. J. Lee, M. C. Roggemann, B. M. Welsh, and E. R. Crosby, "Evaluation of least-squares phase-diversity technique for space telescope wave-front sensing," Appl. Opt. **36**, 9186-9197 (1997).

12. M. Lofdahl, and G. Scharmer, "Phase diverse speckle inversion applied to data from the Swedish 1-meter solar telescope," in *Innovative Telescopes and Instrumentation for Solar Astrophysics, Proc. SPIE*, S. L. Keil, and S. V. Avakyan, eds., **4853** (2003), pp. 567-575.

13. R. Paxman, J. Seldin, M. Loefdahl, G. Scharmer, and C. Keller, "Evaluation of Phase-Diversity Techniques for Solar-Image Restoration," The Astrophysical Journal **466**, 1087 (1996).

14. L. P. Yaroslavsky, and H. J. Caulfield, "Deconvolution of multiple images of the same object," Appl. Opt. **33**, 2157-2162 (1994).

15. R. G. Paxman, T. J. Schulz, and J. R. Fienup, "Joint estimation of object and aberrations by using phase diversity," J. Opt. Soc. Am. A **9**, 1072 (1992).

16. S. T. Thurman, and J. R. Fienup, "Noise histogram regularization for iterative image reconstruction algorithms," J. Opt. Soc. Am. A **24**, 608-617 (2007).

17. J. Fienup, "Phase retrieval for undersampled broadband images," J. Opt. Soc. Am. A **16**, 1831-1837 (1999).

3 Sub-aperture Piston Phase Diversity for Segmented and Multi-aperture Systems

In this Chapter we continue the development of phase diversity by utilizing the unique architecture of segmented and multi-aperture systems to introduce sub-aperture piston phase as the diversity function, as an alternative to focus diversity. We include here the effects of a broadband object and algorithm, as well as compare the performance of sub-aperture piston phase diversity (SAPPD) and focus diversity with respect to object reconstruction error.

SAPPD can be useful for Fourier transform imaging spectroscopy (FTIS) with a multi-aperture system where the sub-apertures are purposely pistoned to collect spectral data [1]. Also, SAPPD can be useful as a risk reduction method on systems where the focus diversity mechanism fails or is disabled.

In Section 3.1 we present the imaging model in the context of a segmented and multi-aperture system, incorporating the gray-world model of Section 2.1. The multi-aperture metrics and gradients are derived in Section 3.2.

In Section 3.3 we describe digital simulations that compare the performance of sub-aperture piston phase diversity with conventional focus diversity. We describe three criteria for comparison: phase estimation, computational burden and object reconstruction.

In Section 3.4 we present results of the digital simulations and conclude in Section 3.5.

3.1 Imaging Model: Segmented and Multi-aperture Systems

The detected images are modeled as

$$d_k(u,v) = \sum_{\lambda} f_\lambda(u,v) * s_{k,\lambda}(u,v) + n_k(u,v), \tag{3.1}$$

where $d_k(u, v)$ is the k^{th} detected image, (u,v) are the image plane coordinates, $f_\lambda(u, v)$ are the object pixel values at wavelength λ, $s_{k,\lambda}(u, v)$ is the k^{th} intensity point spread function (PSF) at wavelength λ, $n_k(u, v)$ is the noise in the k^{th} image and $*$ denotes a convolution. The intensity PSF is the magnitude squared of the coherent impulse response,

$$s_{k,\lambda}(u,v) = \left| h_{k,\lambda}(u,v) \right|^2, \tag{3.2}$$

which in turn is a Fresnel-like transform of the pupil

$$\begin{aligned} h_{k,\lambda}(u,v) = \exp\left[i\frac{\pi D_k}{\lambda B_k}\left(u^2+v^2\right)\right] \iint P_{k,\lambda}(x,y) \exp\left[i\frac{\pi A_k}{\lambda B_k}\left(x^2+y^2\right)\right] \\ \times \exp\left[-i\frac{2\pi}{\lambda B_k}(xu+yv)\right] dxdy \end{aligned} \tag{3.3}$$

where (x, y) are the pupil plane coordinates and A_k, B_k and D_k are the elements of the ABCD ray-transfer matrix that relates the pupil plane to the image plane for the k^{th} diversity image [2]. The generalized pupil, $P_{k,\lambda}(x, y)$, is given by a sum over the sub-aperture functions

$$\begin{aligned} P_{k,\lambda}(x,y) = \sum_{q=1}^{Q} P_{q,k,\lambda}(x,y) = \sum_{q=1}^{Q} \left| P_q(x,y) \right| \\ \times \exp\left\{ i\frac{2\pi}{\lambda}\left[W_q(x,y) + W_{q,k}^{div}(x,y)\right]\right\} \end{aligned} \tag{3.4}$$

where Q is the number of sub-apertures or segments, W_q is the unknown contribution of the phase on sub-aperture q in terms of optical path delay (OPD), and $W_{q,k}^{div}$ is the known diversity contribution to the phase on sub-aperture q in terms of OPD.

The unknown OPD, $W_q(x,y)$, can be parameterized in a number of ways. The most straightforward method is as a pixel-by-pixel phase map,

$$W_q(x,y) = \sum_{m,n} \alpha_{q,m,n} \delta(x - m\Delta x, y - n\Delta y) \tag{3.5}$$

where $\alpha_{q,m,n}$ is the magnitude of the OPD at pixel index (m, n) in the q^{th} sub-aperture, (Δx, Δy) are the pixel spacings and δ is a delta function. Another popular method is to estimate the phase by an expansion over basis functions

$$W_q(x,y) = \sum_{j=1}^{J} \alpha_{q,j} Z_{q,j}(x,y) \tag{3.6}$$

where J is the number of terms in the expansion and $Z_{q,j}(x, y)$ is the j^{th} basis function defined over the q^{th} sub-aperture. Zernike-like polynomials are commonly chosen for the expansion since each polynomial represents a balanced optical aberration. In either representation, the phase is parameterized in terms of the vector of coefficients, $\boldsymbol{\alpha}$.

The goal of phase diversity is to estimate the phase parameters, $\boldsymbol{\alpha}$, and the object pixels, $f_\lambda(u,v)$, at each wavelength λ, from the set of detected images $\{d_k(u, v)\}$. This goal may be unattainable, however, as the problem is likely under-determined. If K diversity images are measured and we assume the information in every pixel in each $M \times N$ image is non-redundant (which is almost certainly not the case), then

there are $M \times N \times K$ knowns. However, there are $M \times N \times L + J$ unknowns, where L is the number of wavelengths and J is the number of phase parameters. Therefore, taking into account that there is some redundant information in the diversity images, K may need to be several times greater than L to make the problem well-posed. The next section will address a method to mitigate this.

3.2 Nonlinear Optimization

In [3], the phase diversity problem is formulated as a nonlinear optimization in which the appropriate likelihood function is maximized for given noise statistics. We follow the same treatment here and maximize the Gaussian log-likelihood function in the Fourier domain, given by

$$\mathbf{L}\left[\left\{d_k(u,v)\right\};\boldsymbol{f},\boldsymbol{\alpha}\right] = -\sum_{k=1}^{K}\sum_{f_u,f_v}\left|D_k(f_u,f_v) - \sum_{\lambda} F_\lambda(f_u,f_v)\,S_{k,\lambda}(f_u,f_v)\right|^2 \quad (3.7)$$

where now $\boldsymbol{f} = \{f_\lambda(u,v)\}$ is an estimate of the object, K is the number of diversity images, (f_u, f_v) are spatial frequency coordinates, D_k is the Fourier transform of the k^{th} detected image, F_λ is the Fourier transform of the object estimate at wavelength λ, and $S_{k,\lambda}$ is the k^{th} optical transfer function (OTF) estimate at wavelength λ.

Maximizing Eq. (3.7), or equivalently minimizing the negative of Eq. (3.7), with respect to $\boldsymbol{f}$ and $\boldsymbol{\alpha}$ provides an estimate of the object pixel values and phase parameters. If the phase is parameterized as polynomial coefficients and is limited to J terms, and an $M \times N$ array of pixels at L spectral bands of the object is estimated, the dimensionality of the search space of the optimization is given by $M \times N \times L + J$.

For example, a moderately sized problem including 45 polynomial coefficients, 5 spectral bands and a 256×256 pixel region of the object requires searching a 327,000-dimensional space, the bulk of which is due to the estimation of the object.

Gonsalves derived a method [4] which involves substituting an estimate of the object into Eq. (3.7) that reduces the metric to a function of only the phase parameters. The most common execution of Gonsalves' method uses an inverse-filtered version of the object as an object estimate for a given phase estimate. To make this work for broadband objects, a gray-world assumption must first be made, where it is assumed that each pixel in the object has the same spectrum,

$$f_\lambda(u,v) = \Phi_\lambda f(u,v) \tag{3.8}$$

where Φ_λ is a spectral coefficient at wavelength λ. Under this assumption, the reduced Gaussian metric becomes

$$\mathbf{L}_{RG}(\boldsymbol{\alpha}) = \sum_{f_u,f_v\in\chi}\left\{\sum_{k=1}^{K}|D_k|^2 - \frac{\left|\sum_{j=1}^{K} D_j \sum_{\lambda}\Phi_\lambda S^*_{j,\lambda}(\boldsymbol{\alpha})\right|^2}{\sum_{m=1}^{K}\left|\sum_{\lambda}\Phi_\lambda S_{m,\lambda}(\boldsymbol{\alpha})\right|^2}\right\} \tag{3.9}$$

where χ is the set of pixels where the denominator does not equal zero and the dependence of D_j and $S_{j,\lambda}$ on (f_u, f_v) has been left out for brevity. Notice the object does not appear explicitly in Eq. (3.9). Instead, the object is implicitly estimated jointly with the phase parameters when the metric is minimized. After the phase is estimated, the object can be reconstructed with a Wiener-Helstrom filter [5, 6].

Numerous regularization schemes have been proposed for improving the robustness of the reduced Gaussian metric with respect to noise [7-10], three of which were reported on in Section 2.2. For this work we simulate images with moderate to high SNRs and therefore use only the regularization imposed by restricting the summation over χ. Furthermore, we assume here that the spectral coefficients Φ_λ are known *a priori*.

Typically, Eq. (3.9) is minimized with a gradient search algorithm. To assist the computation and avoid costly finite difference calculations, analytic gradients of the metric with respect to the unknown phase parameters can be computed and are given by

$$\begin{aligned}\frac{\partial \mathbf{L}_{RG}}{\partial \alpha_{\xi,j}} = 8\pi \,\mathrm{Im}\Bigg\{ & \sum_{k=1}^{K} \sum_{\lambda} \frac{\Phi_\lambda}{\lambda} \sum_{f_u', f_v'} Z_{\xi,j}\left(f_u', f_v'\right) P_{\xi,k,\lambda}\left(f_u', f_v'\right) \\ & \times \exp\left[i \frac{\pi A_k}{\lambda B_k} \left(f_u'^2 + f_v'^2 \right) \right] \\ & \times \sum_{f_u, f_v \in \chi} Y_k^*\left(f_u, f_v\right) H_{k,\lambda}^*\left(f_u' - f_u, f_v' - f_v\right) \Bigg\}\end{aligned} \tag{3.10}$$

where $H_{k,\lambda}$ are the generalized pupil functions given by

$$H_{k,\lambda}\left(f_u, f_v\right) = P_{k,\lambda}\left(f_u, f_v\right) \exp\left[i \frac{\pi A_k}{\lambda B_k} \left(f_u^2 + f_v^2 \right) \right] \tag{3.11}$$

and

$$Y_k\left(f_u, f_v\right) = \frac{\sum_{m=1}^{K}\left|\sum_{\lambda}\Phi_\lambda S_{m,\lambda}\right|^2\left(\sum_{\ell=1}^{K} D_\ell \sum_{\lambda}\Phi_\lambda S^*_{\ell,\lambda}\right)D^*_k}{\left(\sum_{m=1}^{K}\left|\sum_{\lambda}\Phi_\lambda S_{m,\lambda}\right|^2\right)^2} - \frac{\left|\sum_{\ell=1}^{K} D_\ell \sum_{\lambda}\Phi_\lambda S^*_{\ell,\lambda}\right|^2 \sum_{\lambda}\Phi_\lambda S^*_{k,\lambda}}{\left(\sum_{m=1}^{K}\left|\sum_{\lambda}\Phi_\lambda S_{m,\lambda}\right|^2\right)^2} . \quad (3.12)$$

Equations (3.10) and (3.12) are similar in form to Eqs. (25) and (26) in [11].

3.3 Digital Simulations

A triarm-9 aperture was modeled in digital simulation to compare the two types of phase diversity, quadratic and sub-aperture piston, through a series of Monte Carlo simulations. The amount of each type of diversity was varied, and for each amount of diversity three images were simulated, one image with no diversity phase and two images with equal but opposite amounts of diversity phase. Each type and magnitude of diversity was then tested at varying levels of signal-to-noise ratio (SNR).

Figure 3-1 shows the sub-aperture piston phase and focus diversity implementations used. We note that our selection of sub-apertures to be pistoned is arbitrary. Some work has been done to optimize the selection of sub-aperture groups for pistoning for FTIS [1], the imaging properties of which are similar to the application here. However, a full study of pupil geometries and sub-aperture group selection is beyond the scope of this work.

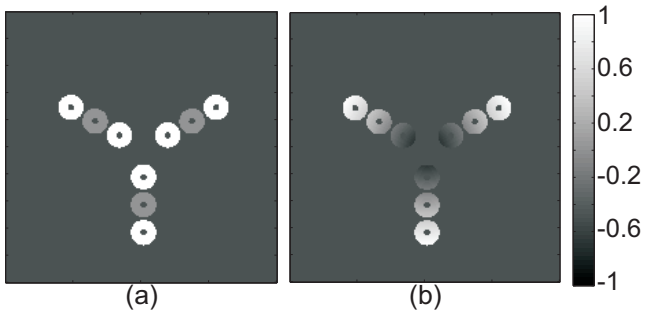


Figure 3-1: Example of phase diversity implementation: (a) sub-aperture piston diversity, (b) focus diversity. Scale has units of waves.

For these simulations, a hyperspectral AVIRIS [12] data cube was used as the object. We extracted a spectral bandwidth of 96 nm centered about 1 μm with 10 nm separation between adjacent wavelengths (L = 11 bands). Since the object was not truly gray-world, the spectrum averaged over all of the pixels was used for the spectral coefficients, Φ_λ. Figure 3-2 shows a panchromatic representation of the 250 × 250 pixel object. Note that in the simulations the images were computed without the gray-world assumption, as in Eq. (3.1), but in our reconstructions we used the gray-world assumption, a realistic model mismatch.

Five independent phase realizations were tested with an average RMS wavefront error of 0.18 λ_o which consisted of global Zernike terms up to 6th order and sub-aperture Zernike terms up to 2nd order. A representative phase realization is shown in Figure 3-3. For each phase realization, five independent noise realizations were simulated. The results were then averaged over the 25 trials.

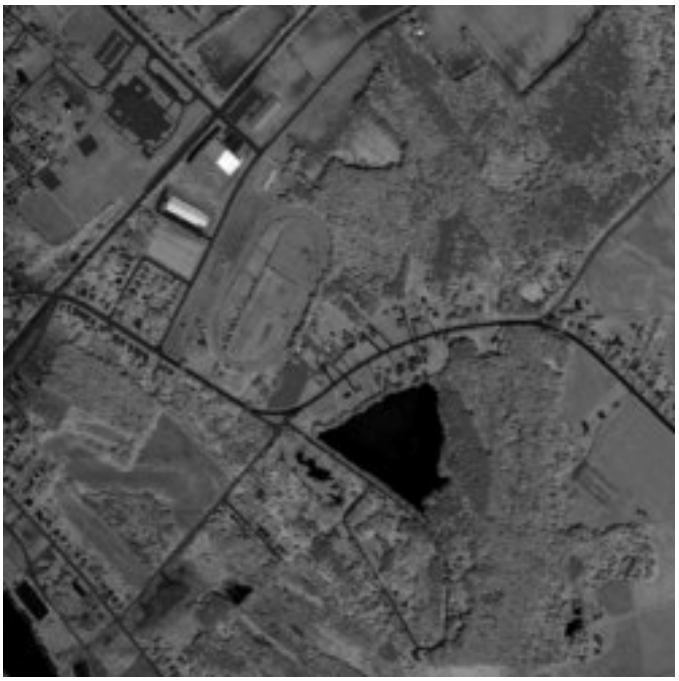


Figure 3-2: Panchromatic representation of the multi-spectral object. The object consists of 11 spectral bands, centered about 1 μm and spanning 96 nm.

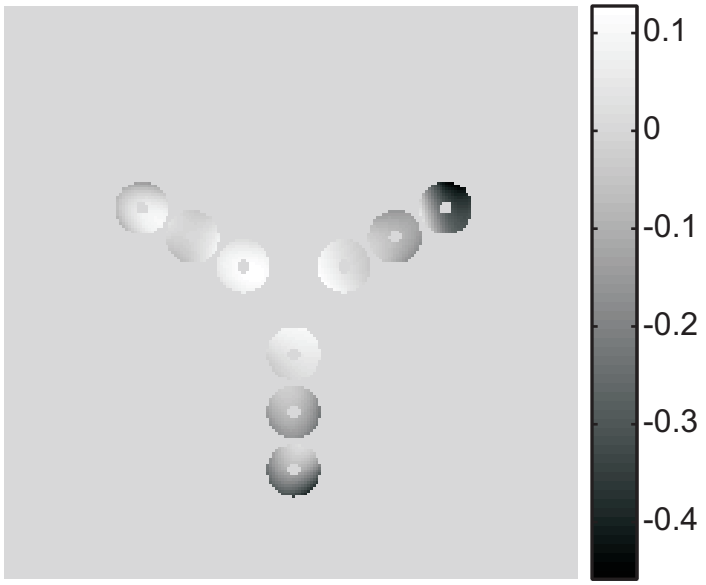


Figure 3-3: Example phase realization composed of up to 6th order Zernike terms on the global aperture and up to 2nd order Zernike terms on each sub-aperture. Scale has units of waves.

Three criteria were used here to compare the two types of diversity. The first criterion is the error in the phase estimation, which must be insensitive to global piston, tip, and tilt phase errors, which the phase diversity algorithm cannot estimate and do not affect image quality. We choose to quantify the phase error in terms of the Strehl ratio, which was presented in Section 2.1.3.

The second criterion of comparison is the amount of computation required by each type of diversity, quantified by the number of iterations required before the algorithm met the exit criteria. The amount of time involved in evaluating the error metric and its gradients was nearly identical for each type of diversity. Therefore, comparing the number of iterations provides a comparison of the computational burden of each type of diversity.

The final criterion is the error in the reconstructed object. A multi-frame Wiener-Helstrom filter is computed using the estimated phase [6, 13] and then used to reconstruct a gray-world image of the object. For a fair comparison, a series of diffraction-limited, gray-world images were simulated and noise was added,

$$d_{DL}(u,v) = \left[\sum_{\lambda} f_{\lambda}(u,v)\right] * \left[\sum_{\lambda} s_{DL,\lambda}(u,v)\right] + n(u,v) \tag{3.13}$$

with similar noise statistics as the aberrated images. The multi-frame Wiener-Helstrom filter provides an SNR benefit according to the number of frames being used. Since the diffraction-limited images have no phase diversity, three frames were simulated, each with a different noise realization, for each SNR level. This allows for

a comparable SNR benefit from the multi-frame Wiener-Helstrom filter as the aberrated diverse images would achieve.

The reconstructed, aberrated, gray-world images were then compared to the reconstructed, diffraction-limited, gray-world images by means of a normalized root-mean-squared error (NRMSE) metric that is invariant with image translation [14].

3.4 Results

Figure 3-4 shows the results for the Strehl ratio for three SNRs. Focus diversity exhibits expected behavior giving the best performance for one to two waves of focus diversity. Piston diversity shows a different behavior. As the amount of diversity increases, the phase estimation goes through a cycle. At integer values of the center wavelength of the band, the piston diversity performs very poorly, because at integer wavelength values there is little diversity, unless the spectral bandwidth is very large. In these simulations, the bandwidth was a modest 10%. However, piston diversity performs well for non-integer numbers of wavelengths. At half-integer values of the center wavelength a slight decrease in performance is seen due to the fact that for small bandwidths there are effectively only two diversity images. For example, the trials using 1.5 λ_o PV of sub-aperture piston diversity used three images with diversity values of [–1.5, 0, 1.5] λ_o of diversity. The first and last images in this set differ by exactly 3 λ_o of piston diversity and have little diversity between them. For these trials, at the optimum diversity value for each type of diversity, focus diversity

performed better than SAPPD for low SNR, but SAPPD performed better at high SNRs.

The observant reader may notice that it appears that the phase estimation performance decreases for focus diversity as the SNR increases from 74 to 120. In fact, the average residual Strehl ratio does indeed decrease between to the two highest SNR levels, as shown in Figure 3-5. However, the significant overlap between the error bars for these two SNR levels implies statistical equivalence.

Figure 3-6 shows the number of iterations required before the algorithm reached its exit criteria. The algorithm typically stops iterating when there is no significant change in either the error metric value or parameters being estimated. This corresponds to finding a minimum value of the objective function, which may or may not be the global minimum. For both types of diversity, the algorithm took fewer iterations when phase estimation was poor. This may indicate that the algorithm was getting trapped in a local minimum and that restarting the algorithm with a different initial guess might be beneficial. Overall, sub-aperture piston phase diversity took fewer iterations than conventional focus diversity.

Figure 3-7 shows the normalized root-mean-squared error in the reconstructed objects. For phase diversity at an integer number of waves of the center wavelength, where the phase estimate is poor, SAPPD gives a poor object reconstruction. However, SAPPD gives a good object reconstruction nearer to half-integer waves of diversity, and the error does not appear to increase with the amount of diversity since the suppression of the OTF does not significantly change with the magnitude of the

piston diversity. In comparison, the OTF becomes increasingly suppressed as quadratic phase diversity is increased. Overall, when both types of diversity yield a good reconstruction, the NRMSE is comparable.

Figure 3-8 shows the normalized root-mean-squared error between an object reconstructed using the estimated phase and an object reconstructed using the actual phase. This comparison shows how much of the error in the object reconstruction is due to an error in the phase estimation. The behavior is very similar to that of the error phase estimation. For low SNR, focus diversity results in the best reconstructed objects, but for higher SNRs, SAPPD results in better reconstructions. Generally, when phase estimation was poor for either type of diversity, so was object reconstruction.

Figure 3-9 shows an example reconstruction for an SNR of 74 using SAPPD. Figure 3-9(a) shows the gray-world image reconstructed using the phase estimated from the phase diversity algorithm using sub-aperture piston diversity values of [–1.5, 0, 1.5] λ_o, in good agreement with the diffraction-limited, gray-world, noisy reconstructed image show in Figure 3-9(b). For comparison, Figure 3-9(c) shows the zero-diversity aberrated imagery that was used as input to the phase diversity algorithm; the reconstructed images are greatly improved over the best measured image.

3.5 Summary

We introduced a new method of implementing phase diversity on segmented and multi-aperture systems that utilizes control of the individual segments or sub-apertures. Instead of adding defocus phase to each diversity image, a subset of the segments or sub-apertures was pistoned with respect to the others to introduce the diversity. A series of Monte Carlo simulations was run to compare sub-aperture piston phase diversity to conventional focus phase diversity. In each case, performance was compared for varying amounts of diversity and varying levels of SNR.

It was shown that focus diversity generally estimates the phase better than SAPPD for low SNR, with the reverse being true for high SNR. Piston phase diversity performed well when non-integer values of the center wavelength are chosen for the amount of diversity. In practice one could select this favorable amount of diversity. SAPPD uniformly required substantially less computation before reaching a minimum value of the error metric than focus diversity. Finally, both focus diversity and SAPPD provide comparably good object reconstructions when the phase estimation is also good.

SAPPD proves to be a useful type of phase diversity for segmented and multi-aperture systems where pistoning of sub-apertures is allowable. SAPPD may be especially useful as a risk reduction technique for systems if the primary focus diversity mechanism fails.

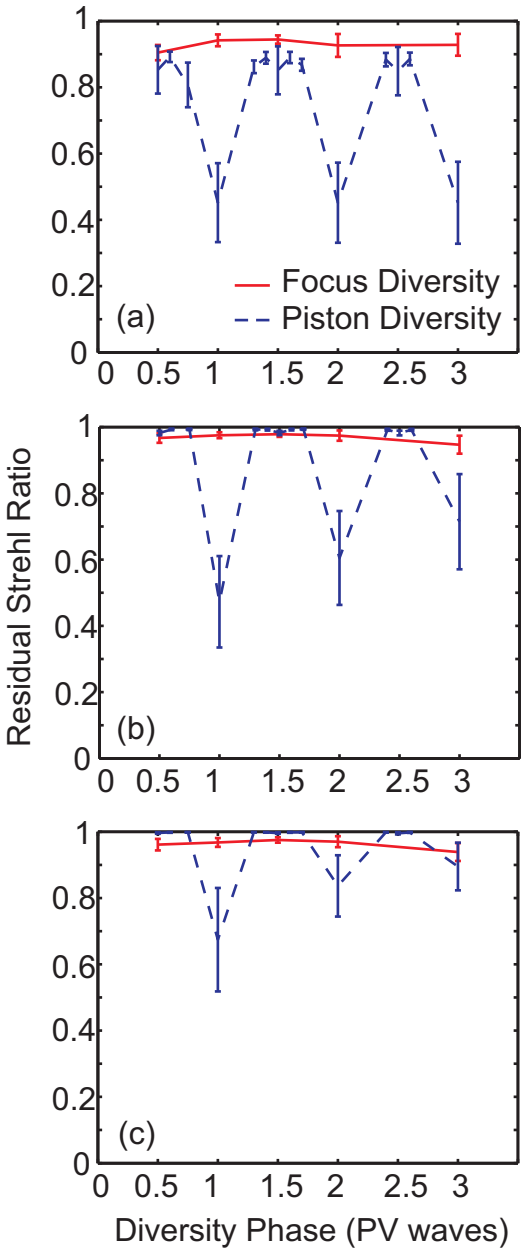


Figure 3-4: Phase estimation results in terms of Strehl ratio: (a) average pixel SNR of 20, (b) SNR 74, (c) SNR 170. Vertical axis shows Strehl ratio, horizontal axis is the peak-to-valley amount of diversity in waves. Each data point is an average of 25 trials (5 phase realizations x 5 noise realizations); error bars show a single standard deviation.

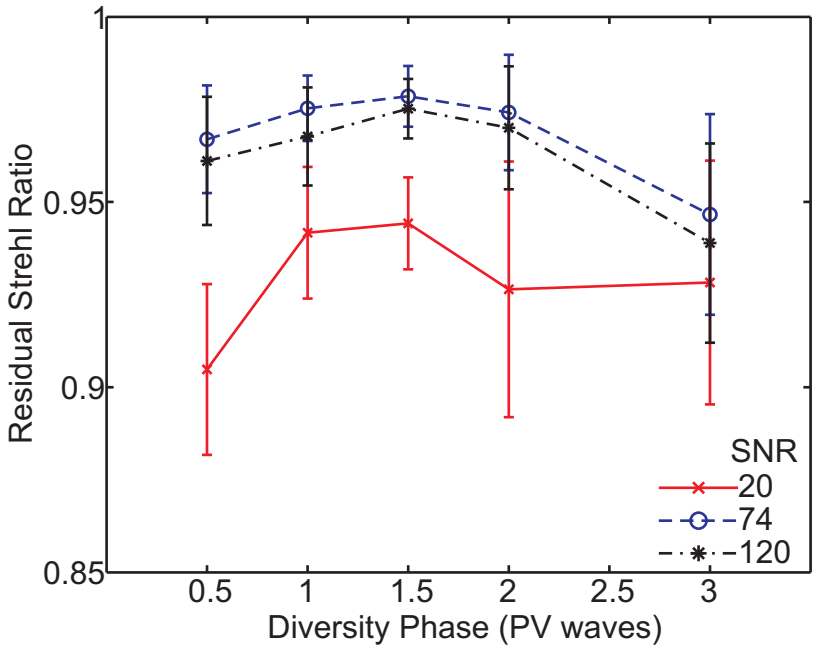


Figure 3-5: Residual Strehl ratio results for focus diversity at each SNR. While the average value for an SNR of 120 is consistently lower than the average value for an SNR of 74, the overlapping error bars indicate statistical equivalence.

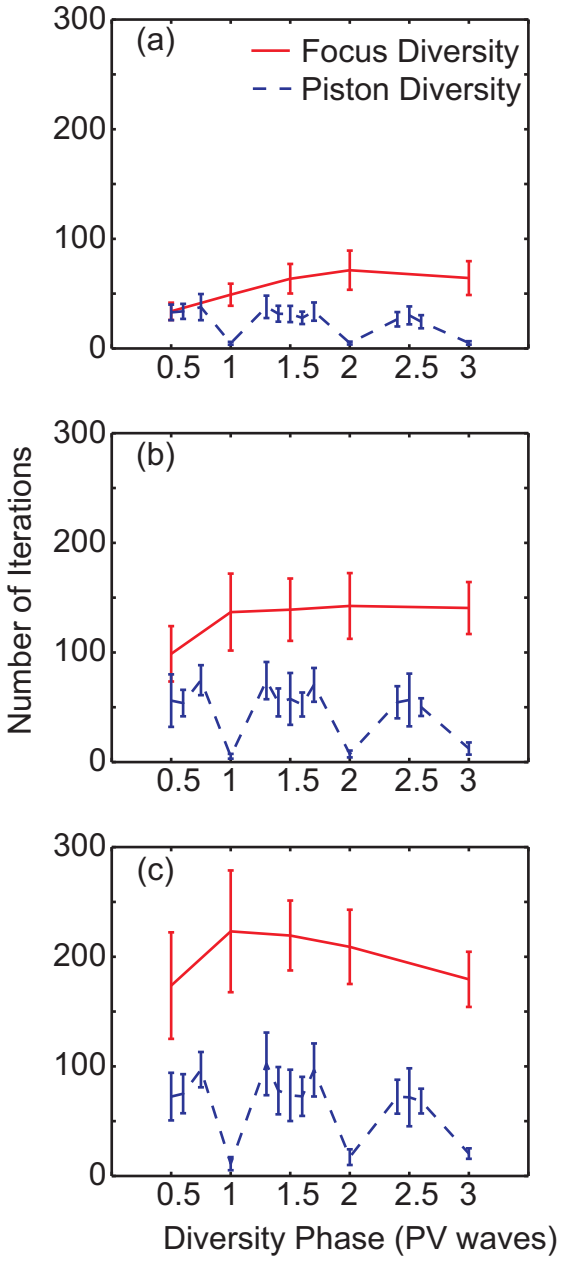


Figure 3-6: Number of iterations before the algorithm reaches the exit criteria: (a) average pixel SNR of 20, (b) SNR 74, (c) SNR 170. Vertical axis shows the number of iterations, horizontal axis is the peak-to-valley amount of diversity in waves. Each data point is an average of 25 trials; error bars show a single standard deviation.

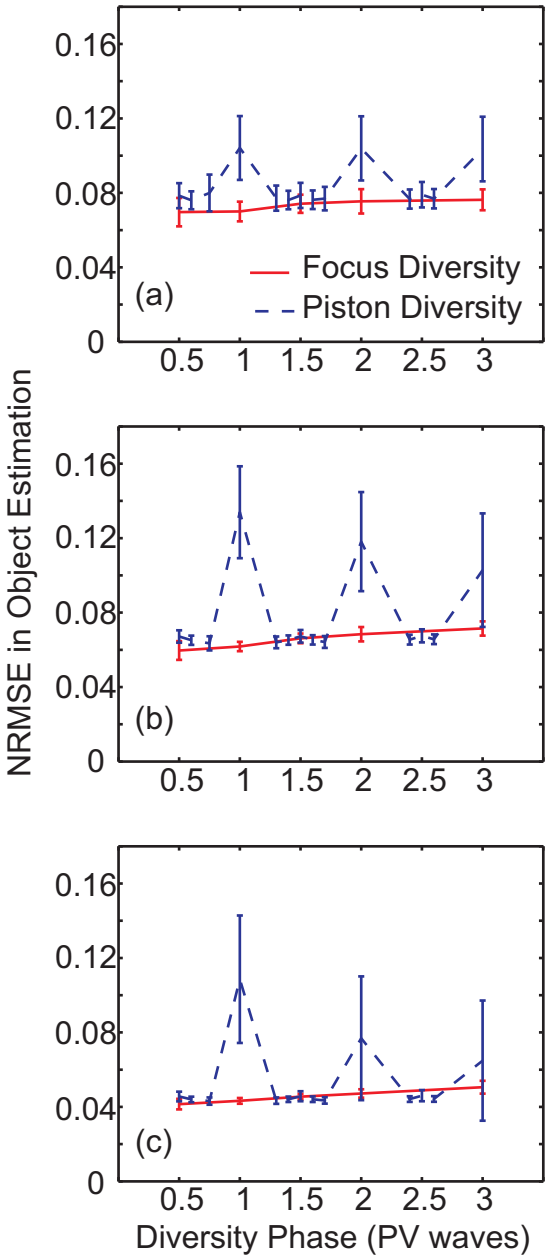


Figure 3-7: Normalized RMS error between reconstructed aberrated object and reconstructed diffraction-limited object: (a) average pixel SNR of 20, (b) SNR 74, (c) SNR 170. Vertical axis shows NRMSE, horizontal axis is the peak-to-valley amount of diversity in waves. Each data point is an average of 25 trials; error bars show a single standard deviation.

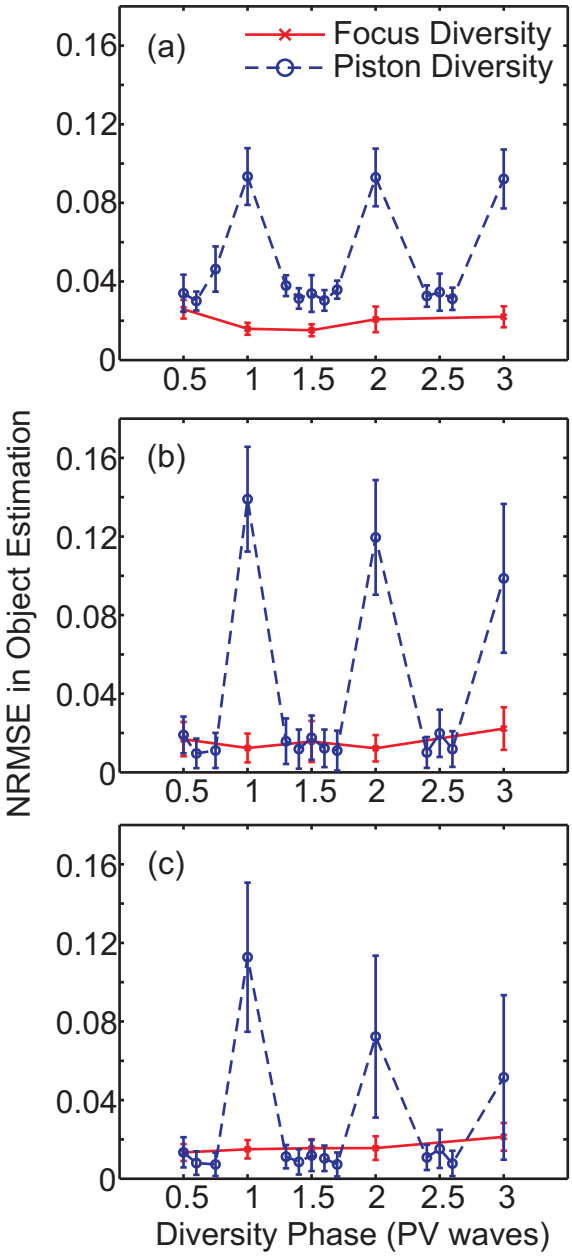
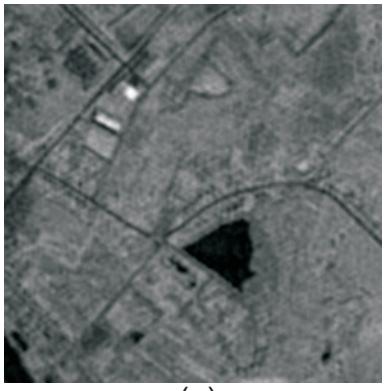
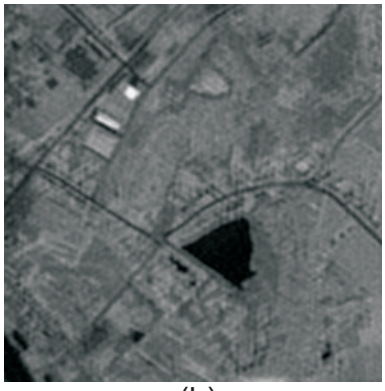


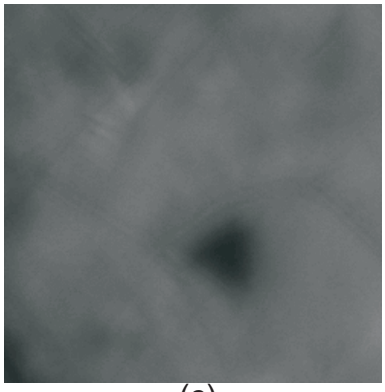
Figure 3-8: Normalized RMS error between object reconstructed with estimated phase and object reconstructed with true phase: (a) average pixel SNR of 20, (b) SNR 74, (c) SNR 170. Vertical axis shows NRMSE, horizontal axis is the peak-to-valley amount of diversity in waves. Each data point is an average of 25 trials; error bars show a single standard deviation.



(a)



(b)



(c)

Figure 3-9: Example reconstructed images: (a) reconstructed gray-world image using phase estimate from phase diversity algorithm, (b) reconstructed gray-world diffraction-limited image, (c) original aberrated, zero-diversity image used as input to phase diversity algorithm. All images are shown on the same color scale, SNR 74.

The work in this Chapter was published in Applied Optics:

M. R. Bolcar, and J. R. Fienup, "Sub-aperture piston phase diversity for segmented and multi-aperture systems," Appl. Opt. **48**, A5-A12 (2009).

3.6 References

1. S. Thurman, and J. Fienup, "Multi-aperture Fourier transform imaging spectroscopy: theory and imaging properties," Opt. Express **13**, 2160-2175 (2005).

2. A. E. Siegman, *Lasers* (University Science Books, Sausalito, CA, 1986).

3. R. G. Paxman, T. J. Schulz, and J. R. Fienup, "Joint estimation of object and aberrations by using phase diversity," J. Opt. Soc. Am. A **9**, 1072 (1992).

4. R. Gonsalves, and R. Chidlaw, "Wavefront sensing by phase retrieval," in *Applications of Digital Image Processing III, Proc. SPIE*, **207** (1979).

5. J. Goodman, *Introduction to Fourier Optics, 3rd Ed.* (Roberts & Co, 2005).

6. L. P. Yaroslavsky, and H. J. Caulfield, "Deconvolution of multiple images of the same object," Appl. Opt. **33**, 2157-2162 (1994).

7. D. J. Lee, M. C. Roggemann, B. M. Welsh, and E. R. Crosby, "Evaluation of least-squares phase-diversity technique for space telescope wave-front sensing," Appl. Opt. **36**, 9186-9197 (1997).

8. O. M. Bucci, A. Capozzoli, and G. D'Elia, "Regularizing strategy for image restoration and wave-front sensing by phase diversity," J. Opt. Soc. Am. A **16**, 1759-1768 (1999).

9. M. Lofdahl, and G. Scharmer, "Phase diverse speckle inversion applied to data from the Swedish 1-meter solar telescope," in *Innovative Telescopes and*

Instrumentation for Solar Astrophysics, Proc. SPIE, S. L. Keil, and S. V. Avakyan, eds., **4853** (2003), pp. 567-575.

10. A. Blanc, L. M. Mugnier, and J. Idier, "Marginal estimation of aberrations and image restoration by use of phase diversity," J. Opt. Soc. Am. A **20**, 1035-1045 (2003).

11. J. H. Seldin, R. G. Paxman, V. G. Zarifis, L. Benson, and R. E. Stone, "Closed-loop wavefront sensing for a sparse-aperture, phased-array telescope using broadband phase diversity," in *Imaging Technology and Telescopes, Proc. SPIE*, J. W. Bilbro, J. B. Breckinridge, R. A. Carreras, S. R. Czyzak, M. J. Eckart, R. D. Fiete, and P. S. Idell, eds., **4091** (2000), pp. 48–63.

12. Provided through the courtesy of Jet Propulsion Laboratory, California Institute of Technology, Pasadena, California (http://aviris.jpl.nasa.gov/).

13. S. T. Thurman, and J. R. Fienup, "Noise histogram regularization for iterative image reconstruction algorithms," J. Opt. Soc. Am. A **24**, 608-617 (2007).

14. M. Guizar-Sicairos, S. T. Thurman, and J. R. Fienup, "Efficient subpixel image registration algorithms," Opt. Lett. **33**, 156-158 (2008).

4 Estimation of Object Spectral Content

In Section 2.1, we presented a broadband phase diversity algorithm that made use of a gray-world assumption on the object to reduce the metric to a function of only the unknown phase parameters. It was shown that if seven gray-world spectral coefficients are used but assumed to be uniform, it is possible to accurately estimate the phase to within 0.075 λ_o for bandwidths up to 90% and to within 0.05 λ_o for bandwidths up to 70%. However, if the spectral coefficients are known accurately estimating the phase to within 0.06 λ_o RMS is possible for bandwidths exceeding 100% and to within 0.04 λ_o RMS for bandwidths of 60%.

In this Chapter, we present a method with which it is possible to estimate the gray-world spectral coefficients, in addition to the unknown phase and unknown object being observed. The sub-aperture piston phase diversity technique of Chapter 3 is shown to be critical to the success of this method. Requirements on the number of diversity images and amount of diversity are investigated.

In Section 4.1, we will present modifications to the conventional reduced Gaussian metric that aid in the spectral estimation process.

Limitations on the spectral resolution are presented in Section 4.2, while requirements on diversity sampling are discussed in Section 4.3.

Digital simulations are described in Section 4.4 and results are presented in Section 4.5. The Chapter is summarized in Section 4.6.

4.1 Spectral Estimation Metric

Recall the reduced Gaussian broadband likelihood function of Eq. (2.6), reprinted here,

$$\mathbf{L}_{RG}(\boldsymbol{\alpha}) = \sum_{f_u, f_v \in \chi} \frac{\left| \sum_{j=1}^{K} D_j(f_u, f_v) \sum_{\lambda} \Phi_\lambda S^*_{j,\lambda}(f_u, f_v; \boldsymbol{\alpha}) \right|^2}{\sum_{\ell=1}^{K} \left| \sum_{\lambda} \Phi_\lambda S_{\ell,\lambda}(f_u, f_v; \boldsymbol{\alpha}) \right|^2 - \sum_{u,v} \sum_{k=1}^{K} \left| D_k(f_u, f_v) \right|^2}, \tag{2.6}$$

and that by maximizing Eq. (2.6) with respect to the unknown phase parameters, $\boldsymbol{\alpha}$, an estimate of the phase can be obtained. In practice, the negative of Eq. (2.6) is minimized with a nonlinear optimization routine. We re-write Eq. (2.6) in error metric form as

$$E_{RG}(\boldsymbol{\alpha}) = - \sum_{f_u, f_v \in \chi} \frac{\left| \sum_{j=1}^{K} D_j(f_u, f_v) \sum_{\lambda} \Phi_\lambda S^*_{j,\lambda}(f_u, f_v; \boldsymbol{\alpha}) \right|^2}{\sum_{\ell=1}^{K} \left| \sum_{\lambda} \Phi_\lambda S_{\ell,\lambda}(f_u, f_v; \boldsymbol{\alpha}) \right|^2} \tag{4.1}$$

where we also drop the summation over the squared magnitude of the Fourier data since it is independent of $\boldsymbol{\alpha}$. With a few modifications to Eq. (4.1), it is possible to estimate the gray-world spectral coefficients, Φ_λ, along with the phase parameters, $\boldsymbol{\alpha}$. First, positivity of the spectral coefficients is enforced using the relationship

$$\Phi_\lambda = \varphi_\lambda^2 \tag{4.2}$$

where φ_λ are the parameters that are actually estimated by the algorithm. The second constraint on the spectral coefficients is derived from the normalization of the OTFs. The panchromatic, gray-world OTF is given by

$$S(f_u, f_v) = \sum_\lambda \Phi_\lambda S_\lambda (f_u, f_v). \tag{4.3}$$

By definition the panchromatic, gray-world OTF should be unity at DC [1]. If each spectral OTF, $S_\lambda(f_u, f_v)$, is also normalized to unity at DC, then it follows that

$$\begin{aligned} S(0,0) &= \sum_\lambda \Phi_\lambda S_\lambda (0,0) = 1 \\ \therefore \sum_\lambda \Phi_\lambda &= 1. \end{aligned} \tag{4.4}$$

Using Eqs. (4.2) and (4.4) we define a spectral error metric as

$$E_{spec}(\varphi_\lambda) = \left[1 - \sum_\lambda \varphi_\lambda^2 \right]^2. \tag{4.5}$$

A total error metric is formed by adding Eqs. (4.1) and (4.5):

$$E(\varphi_\lambda, \boldsymbol{\alpha}) = - \sum_{f_u, f_v \in \chi} \frac{\left| \sum_{j=1}^{K} D_j(f_u, f_v) \sum_\lambda \varphi_\lambda^2 S_{j,\lambda}^*(f_u, f_v; \boldsymbol{\alpha}) \right|^2}{\sum_{\ell=1}^{K} \left| \sum_\lambda \varphi_\lambda^2 S_{\ell,\lambda}(f_u, f_v; \boldsymbol{\alpha}) \right|^2} + \kappa \left[1 - \sum_\lambda \varphi_\lambda^2 \right]^2. \tag{4.6}$$

where a constant, κ, is included to weight the spectral error relative to the rest of the metric. The gradients of the spectral estimation error metric are derived in Appendix A. Minimizing Eq. (4.6) with respect to φ_λ and $\boldsymbol{\alpha}$ explicitly estimates the gray-world spectrum of the object and the unknown phase parameters while implicitly estimating

the object itself. Once the spectral coefficients and phase parameters are known, a gray-world, multi-frame Wiener filter [2] can be constructed as

$$W_k\left(f_u, f_v\right) = \frac{\sum_{\lambda} \varphi_\lambda^2 S_\lambda^*\left(f_u, f_v\right)}{\sum_{m=1}^{K}\left|\sum_{\lambda} \varphi_\lambda^2 S_\lambda\left(f_u, f_v\right)\right|^2 + c\dfrac{\Psi_N}{\Psi_O\left(f_u, f_v\right)}} \tag{4.7}$$

where Ψ_N and $\Psi_O(f_u, f_v)$ are the noise and object power spectra, respectively, which are estimated using the procedure discussed in Section 2.2.6. The Fourier transform of the panchromatic gray-world object can then be reconstructed as

$$F\left(f_u, f_v\right) = \sum_{k=1}^{K} W_k\left(f_u, f_v\right) D_k\left(f_u, f_v\right). \tag{4.8}$$

4.2 Spectral Resolution

For many spectrometers, a rule-of-thumb for the spectral resolution of the system is given by

$$\delta\lambda \approx \frac{\lambda_o^2}{OPD_{\max}} \tag{4.9}$$

where λ_o is the center wavelength and $OPD_{\max}$ is the maximum path difference causing the spectral interference or separation [3]. For example, in the case of Fourier Transform Imaging Spectroscopy (FTIS) with multi-aperture systems [4], groups of sub-apertures are stepped through a series of delays with respect to other sub-apertures. At each delay, a frame is gathered, building up a data cube of spatial image information in the *x* and *y* dimension and temporal information in the *t* (or path

length) dimension. Fourier transforming the data cube with respect to the temporal dimension yields spectra of the object at each pixel in the image. The spectral resolution of the system is determined by the maximum difference in the delays as the $OPD_{\max}$ term in Eq. (4.9).

An extension can be made to using phase diversity to estimate object spectral content. Here, the $OPD_{\max}$ is given by the maximum difference between diversity values for the collected image set. Two images differing by +2 λ_o and –2 λ_o of focus diversity would have an $OPD_{\max}$ of 4 λ_o, peak-to-valley (PV). If the bandwidth is centered at 1 μm, then the spectral resolution, according to Eq. (4.9) would be 0.25 μm. The diversity would need to be increased to achieve a usefully fine spectral resolution. However, defocusing an image by large amounts has a very negative effect on the imaging capabilities of the system, as well as the performance of phase diversity in general. Specifically, the OTF of the system becomes suppressed, with many nulls appearing in the mid and high spatial frequencies. Figure 4-1 shows MTFs (magnitude of the OTFs) for a diffraction-limited triarm-9 system for a 100 nm bandwidth centered on 1 μm for 0 λ_o, 2 λ_o and 20 λ_o of focus diversity.

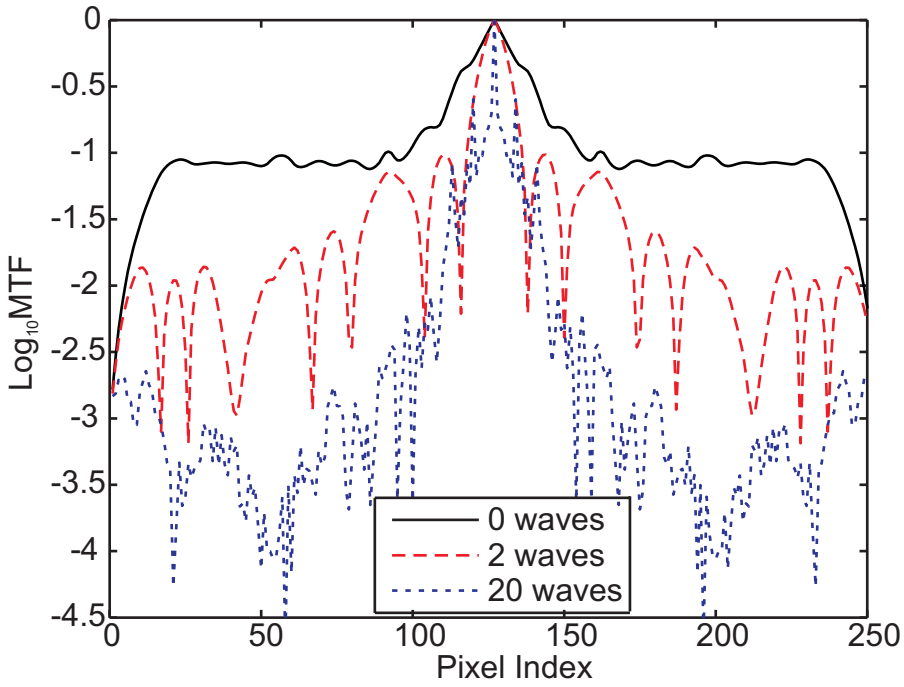


Figure 4-1: Horizontal cut through a triarm-9 MTF for 0 λ_o, 2 λ_o, and 20 λ_o of defocus. For large defocus, the mid to high spatial frequencies of the MTF are significantly depressed.

Sub-aperture piston phase diversity (SAPPD) is much better suited to the spectral estimation problem because large amounts of piston can be added to individual sub-apertures with much less suppression of the OTF than for focus diversity. This allows for a much larger $OPD_{\max}$ and therefore a much finer spectral resolution can be achieved. Figure 4-2 shows OTFs for 0 λ_o, 2 λ_o and 20 λ_o of sub-aperture piston phase diversity for a diffraction-limited triarm-9 system with a 100 nm bandwidth centered on 1 μm. Notice that on average the OTF maintains a higher value through the mid-spatial frequency range with 20 λ_o of piston diversity than a comparable amount of focus diversity. Using SAPPD to obtain diversity images with +20 λ_o and –20 λ_o of piston phase, a spectral resolution of ~0.025 μm is achievable.

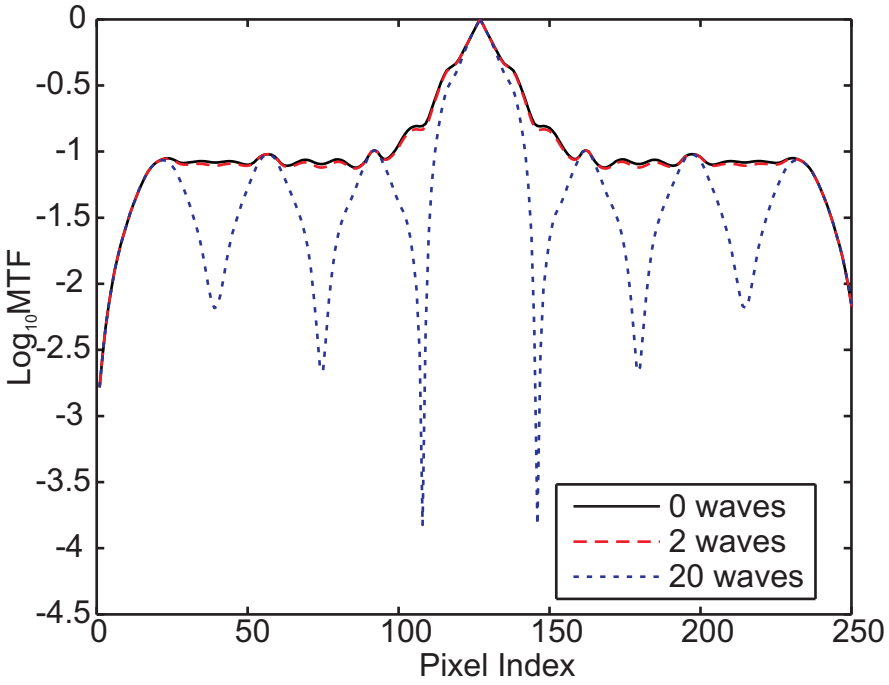


Figure 4-2: Horizontal cuts through a triarm-9 MTF for 0 λ_o, 2 λ_o and 20 λ_o of sub-aperture piston phase. Even with 20 λ_o of defocus, the MTF maintains a higher average value than with a similar amount of defocus.

4.3 OPD Sampling

For FTIS with multi-aperture systems, a sampling constraint must be enforced to ensure that the spectrum is not aliased. According to the Nyquist sampling theorem [5], the OPD sampling interval required to avoid aliasing is given by

$$\Delta OPD = \frac{c}{2\nu_{\max}} = \frac{\lambda_{\min}}{2} \tag{4.10}$$

where c is the speed of light, $\nu_{\max}$ is the maximum optical frequency of the spectrum and, conversely, $\lambda_{\min}$ is the minimum wavelength of the spectrum. It has been shown that the Nyquist constraint of Eq. (4.10) can be relaxed for objects that are band-

limited, or imaged through a spectral bandpass filter [6]. For such an object, the constraint on the OPD sampling interval is given by

$$\Delta OPD \le \frac{c}{2(\nu_{\max} - \nu_{\min})} = \frac{\lambda_{\max}\lambda_{\min}}{2(\lambda_{\max} - \lambda_{\min})} \tag{4.11}$$

where $\nu_{\min}$ and $\nu_{\max}$ are now the optical frequency limits of the bandwidth and $\lambda_{\min}$ and $\lambda_{\max}$ are the wavelength limits of the bandwidth.

4.4 Digital Simulation

Digital simulations were run to investigate the performance of the spectral estimation using SAPPD. The object used for this study was again taken from a set of AVIRIS data [7] which consists of spectroscopic images with a spectral resolution of 10 nm. A 192 nm bandwidth centered about 1 μm was extracted (L = 21 bands). With this bandwidth, Eq. (4.11) can be evaluated to determine what the maximum OPD step-size should be:

$$\Delta OPD \le \frac{(1.096\mu m)(0.904\mu m)}{2(0.192\mu m)} = 2.58\mu m\,. \tag{4.12}$$

SAPPD was implemented as shown in Figure 4-3. For this study, we chose a desired spectral resolution of 25 nm. Solving Eq. (4.9) for the maximum OPD yields

$$\begin{aligned} OPD_{\max} &\approx \frac{\lambda_o^2}{\delta\lambda} \\ &\approx \frac{(1\mu m)^2}{(0.025\mu m)} = 40\mu m. \end{aligned} \tag{4.13}$$

Therefore the piston diversity must span [–20, 20] μm. In practice we chose the interval [–20.6, 20.6] μm since the results of Section 3.4 show that fractional values of the center wavelength work best for SAPPD. We therefore expect a spectral resolution of

$$\delta\lambda = \frac{(1\ \mu m)^2}{41.2\ \mu m} = 24.3\ \text{nm}. \tag{4.14}$$

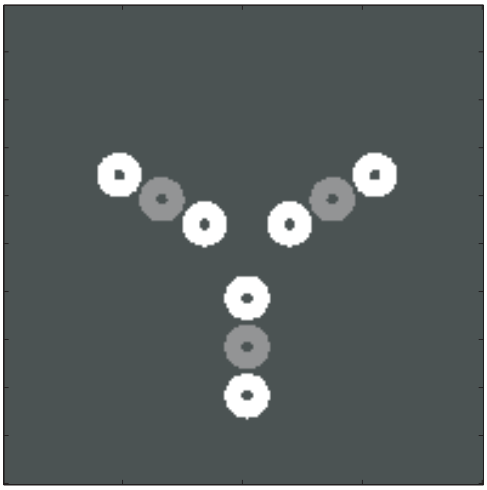


Figure 4-3: Example of SAPPD implementation used for this study.

Combining the results of Eqs. (4.12) and (4.13) and choosing 2.58 μm as the ΔOPD determines the number of OPD steps, or diversity images, required to achieve adequate spectral sampling to be

$$K = \frac{OPD_{\max}}{\Delta OPD} = 15.5. \tag{4.15}$$

We therefore require 16 diversity images in order to avoid aliasing of the spectrum.

Images were simulated to test the performance of the phase diversity algorithm under two scenarios: (1) when the OPD sampling requirement is not met (i.e. $K < 16$), and (2) when finer spectral resolution is attempted (i.e. $\delta\lambda < 24.3$ nm). Scenario (1) is tested by using fewer diversity images in the estimation process, while maintaining the maximum OPD values of [–20.6, 20.6] μm. Scenario (2) is tested by increasing the number of spectral coefficients across the 192 nm bandwidth. Table 4-1 outlines the number of diversity images used and the corresponding ΔOPD. Table 4-2 outlines the number of spectral coefficients estimated and the corresponding spectral resolution, $\delta\lambda$.

Table 4-1: Number of diversity images simulated and corresponding ΔOPD in μm. The minimum and maximum values of piston diversity were always –20.6 μm and 20.6μm, respectively. A zero-diversity image was always included in the set.

K:	19	11	7	5
ΔOPD: (μm)	2.29	4.12	6.87	10.30

Table 4-2: Number of wavelengths sampling the 192 nm bandwidth and corresponding spectral resolution in nm. The center wavelength of 1 μm was always included as one of the samples.

N_λ:	5	7	9	11
$\delta\lambda$ (nm):	38.4	27.4	21.3	17.5

Images were simulated for a triarm-9 aperture. For each set of diversity images, 5 random wavefronts were generated with an average RMS wavefront error of 0.21 λ_o. The wavefront error consisted of Zernike polynomials up to 4th order on the global aperture and up to 2nd order on each sub-aperture. A representative wavefront is shown in Figure 4-4.

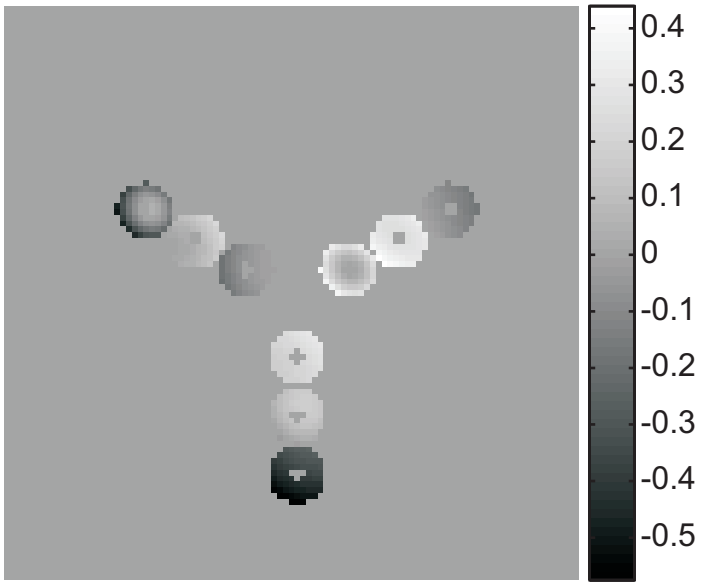


Figure 4-4: Example wavefront used in this study. Wavefront error is 0.21 λ RMS, consisting of Zernike polynomials up to 4th order on the global aperture and up to 2nd order on each sub-aperture. The scale has units of waves.

Realistic noise effects were also included. Images were normalized such that the peak pixel in each image had a value of 80% of a 50,000 e^- well-depth. Intensity-dependent Poisson noise was added, along with zero-mean, additive Gaussian noise with a standard deviation of 15 e^-. Images were then quantized to 12 bits to simulate analog-to-digital conversion. The average pixel SNR, given by Eq. (2.29), was 182. For each of the 5 independent wavefronts that were generated, 5 different noise realizations were used, for a total 25 trials for each set of diversity images.

Phase estimation accuracy was evaluated using residual Strehl ratio, as discussed in Section 2.1.3. Spectral estimation accuracy was evaluated by computing the normalized root-mean-squared error (NRMSE) between the estimated spectral

coefficients and the average spectrum of all the pixels in the object. The NRMSE is computed as

$$NRMSE = \frac{\sqrt{\sum_{\lambda}\left(\Phi_{\lambda}^{avg} - \Phi_{\lambda}^{est}\right)^2}}{\sqrt{\sum_{\lambda}\left(\Phi_{\lambda}^{avg}\right)^2}} \tag{4.16}$$

and

$$\Phi_{\lambda}^{avg} = \frac{1}{MN}\sum_{u,v} f_{\lambda}(u,v) \tag{4.17}$$

where M and N are the number of rows and columns in the array.

4.5 Results

Figure 4-5 shows the results of the simulation with respect to phase estimation. Only trials that resulted in a residual Strehl ratio greater than 0.8 (0.075 λ_o RMS) were considered successfully converged and averaged together. When more spectral coefficients are used to sample the bandwidth, which corresponds to finer spectral resolution, the phase estimation appears to do better, in agreement with the results of Section 2.1.3.

It is unclear why the Strehl ratio has minimum values for trials using 7 and 11 diversity images. Generally, it is expected that increasing the number of diversity images improves the performance of the phase-diversity algorithm. As seen in Figure 4-6, the algorithm does generally converge more often as the number of diversity images is increased.

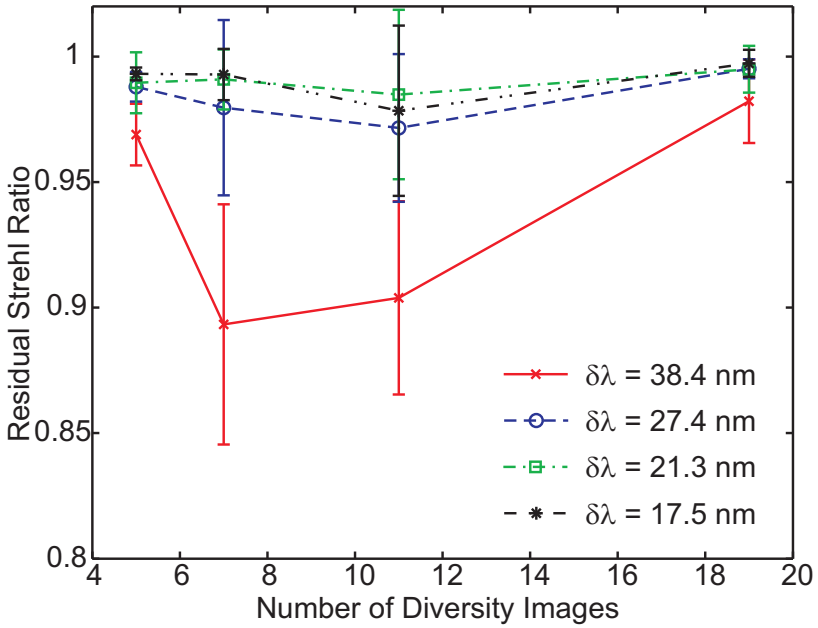


Figure 4-5: Phase estimation accuracy vs. number of diversity images for each spectral resolution. As expected, the phase estimation improves with finer spectral resolution.

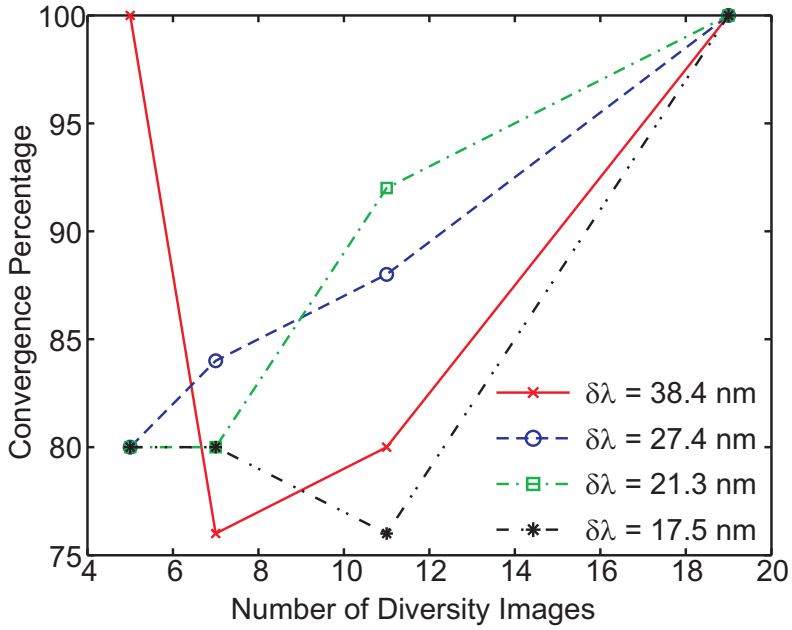


Figure 4-6: Convergence percentage vs. number of diversity images. Generally, the algorithm converges more often as the number of diversity images increases.

Figure 4-7 through Figure 4-10 show the estimated gray-world spectra for each spectral resolution and each OPD sampling. Included for comparison is the object average pixel spectrum. In Figure 4-7, when $\delta\lambda$ = 38.4 nm, the best agreement is found when the OPD sampling requirement is met, as in the lower right graph. However, even when ΔOPD is four times the required step size, as in the upper left graph, the estimated spectrum is still in good qualitative agreement with the true average pixel spectrum. Similar behavior is seen in Figure 4-8, except the best agreement appears to occur when ΔOPD is nearly twice the required step size. In Figure 4-9 and Figure 4-10, where the spectral resolution exceeds the expected limit of 24.3 nm, the agreement between the estimated spectrum and average pixel spectrum begins to deteriorate.

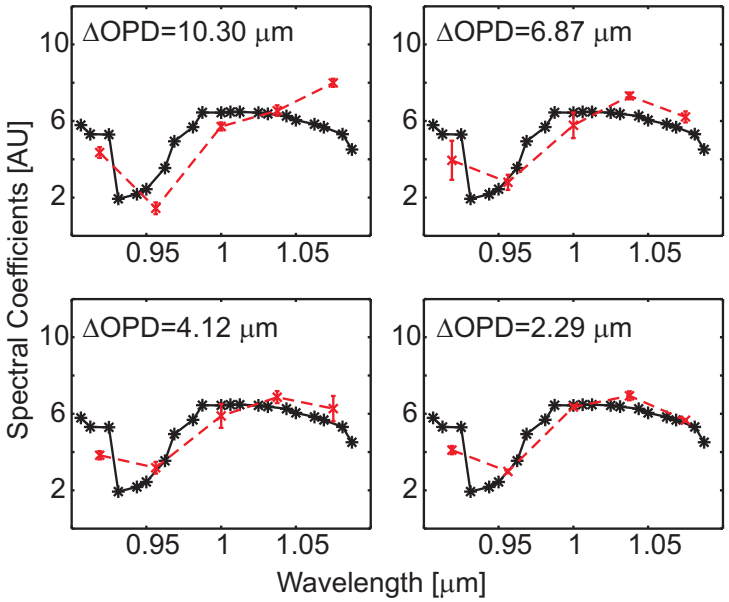


Figure 4-7: Estimated spectra for $\delta\lambda$ = 38.4 nm. The average pixel spectrum of the true object is the solid black line. The estimated spectral coefficients are the dashed red line. Error bars represent a single standard deviation.

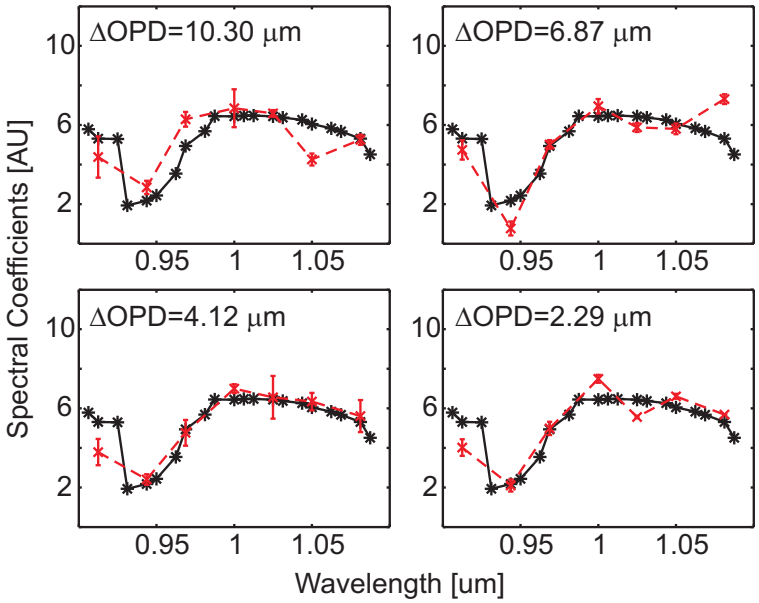


Figure 4-8: Estimated spectra for $\delta\lambda$ = 27.4 nm. The average pixel spectrum of the true object is the solid black line. The estimated spectral coefficients are the dashed red line. Error bars represent a single standard deviation.

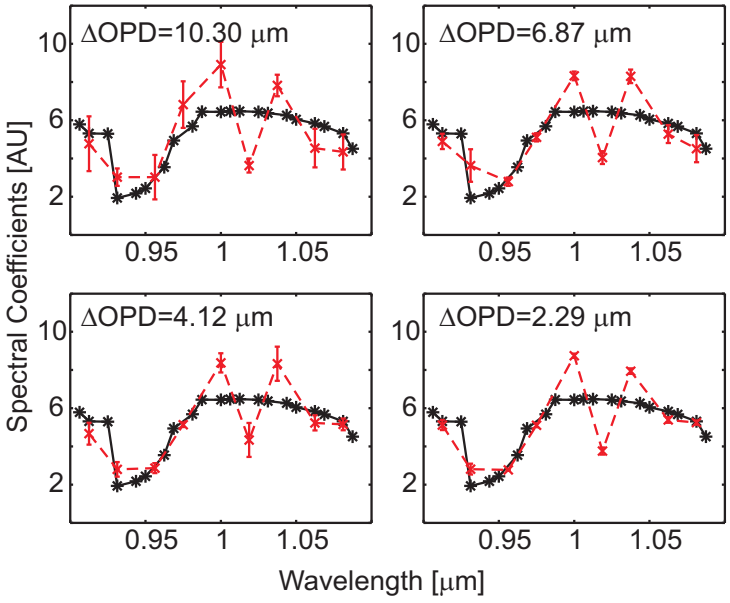


Figure 4-9: Estimated spectra for $\delta\lambda$ = 21.3 nm. The average pixel spectrum of the true object is the solid black line. The estimated spectral coefficients are the dashed red line. Error bars represent a single standard deviation.

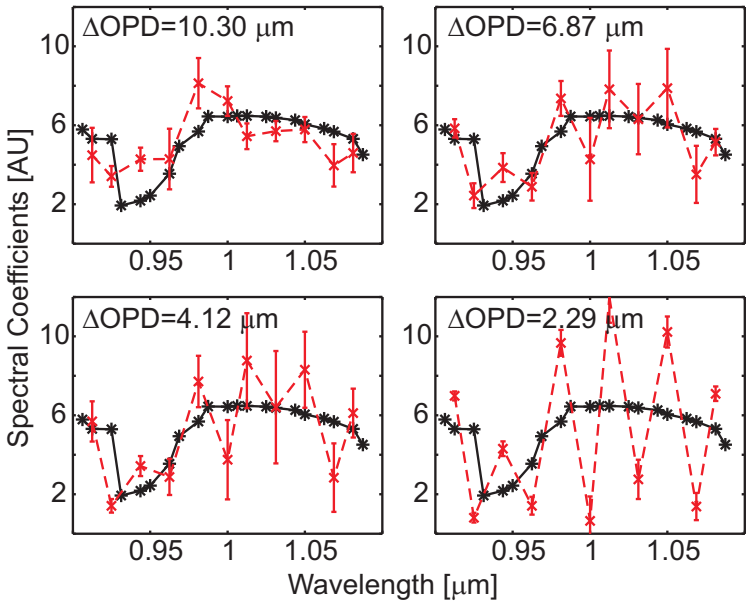


Figure 4-10: Estimated spectra for $\delta\lambda$ = 17.5 nm. The average pixel spectrum of the true object is the solid black line. The estimated spectral coefficients are the dashed red line. Error bars represent a single standard deviation.

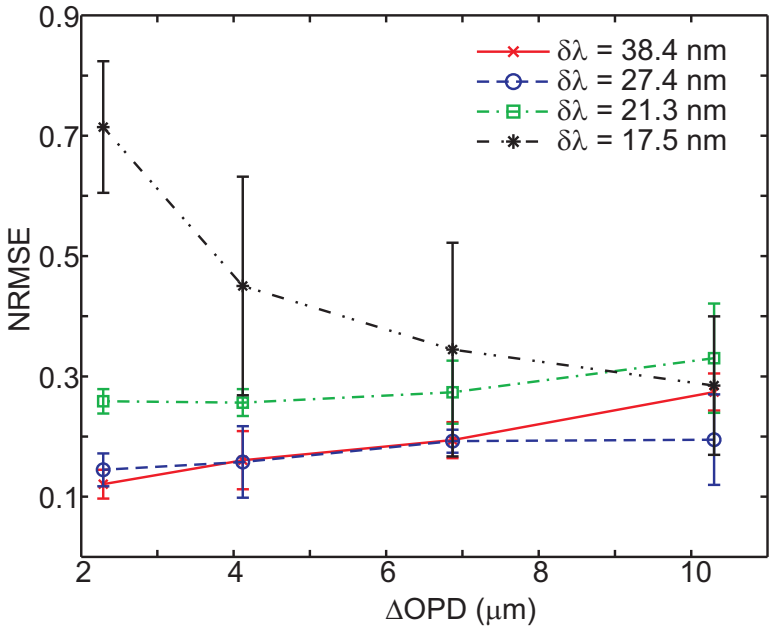


Figure 4-11: Normalized root-mean-squared error in the spectral estimation vs. OPD step size.

Figure 4-11 shows the average normalized root-mean-squared error of the estimated spectrum for each spectral resolution. The lowest error occurred when both the spectral resolution and the OPD sampling are in agreement with Eqs. (4.9) and (4.11), respectively. The error generally increases as these limits are exceeded, with the exception of the $\delta\lambda$ = 17.5 nm resolution, where the error decreases as the OPD sampling increases.

An image was reconstructed for one of the $\delta\lambda$ = 27.4 nm and ΔOPD = 4.12 μm (K = 11) trials. Using the phase and spectral coefficients estimated by the algorithm, a multi-frame Wiener filter using Eq. (4.7) was constructed. For comparison, a diffraction-limited gray-world image was also simulated according to Eq. (3.13). This image was also restored using a diffraction-limited, multi-frame Wiener filter. Figure 4-12(a) shows the restored diffraction-limited gray-world image, in good agreement with Figure 4-12(b), which shows the restored aberrated image using the estimated phase and spectrum. Figure 4-12(c) shows the unprocessed zero-diversity image input to the phase-diversity algorithm.

4.6 Summary

A method was presented with which the gray-world spectral coefficients of a broadband object can be estimated using sub-aperture piston phase diversity. Large amounts of diversity are necessary for fine spectral resolution, such that conventional focus diversity is not ideal. Pistoning sub-apertures through 40 μm of OPD allows a spectral resolution of 25 nm to be achieved. To avoid aliasing of the spectrum,

sampling requirements are placed on the diversity OPD. However, it has been shown that even when the piston OPD step size is nearly four times larger than required, good qualitative agreement between the estimated spectrum and true average pixel spectrum can be achieved.

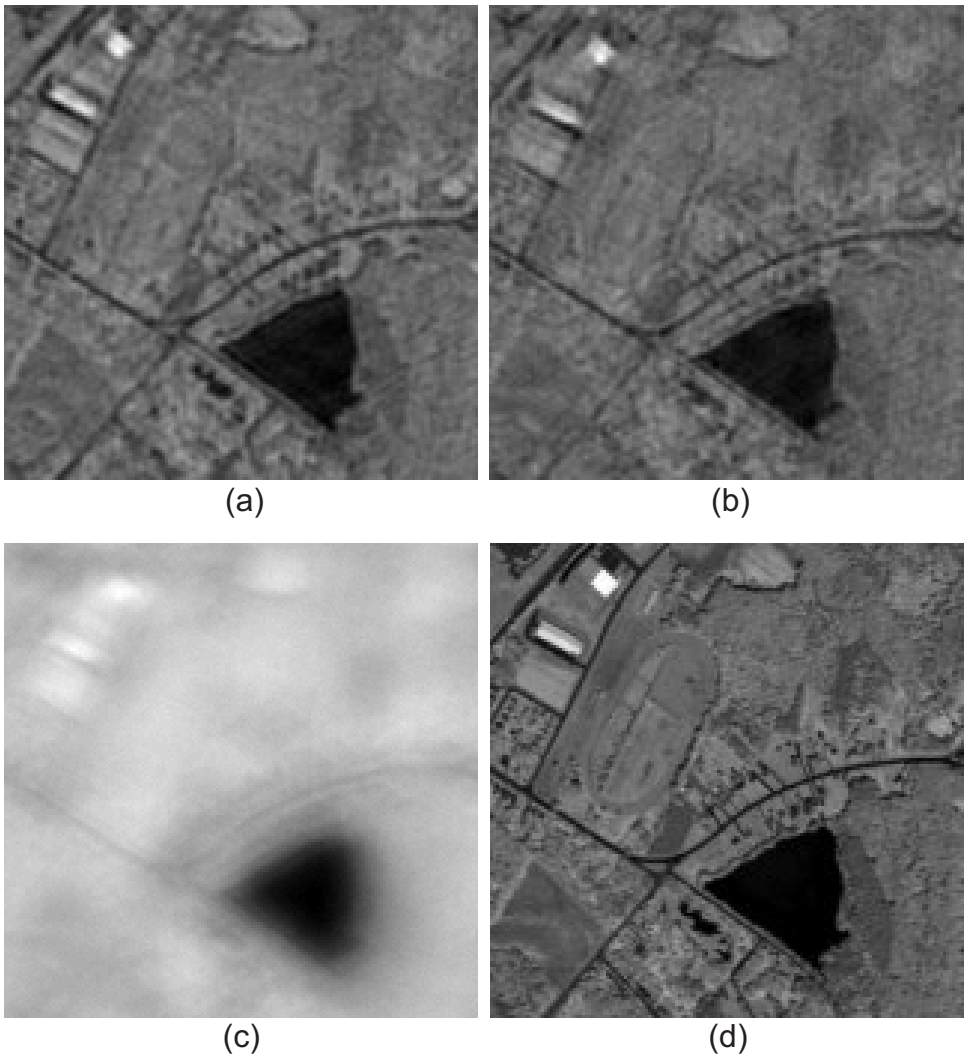


Figure 4-12: (a) Restored diffraction-limited image. (b) Restored aberrated image. (c) Zero-diversity image as input to phase-diversity algorithm. (d) Panchromatic representation of the true object.

4.7 Appendix A: Spectral Estimation Metric Gradients

To aid in the minimization process, analytic gradients of Eq. (4.6) with respect to the phase parameters and gray-world spectral coefficients can be computed. The analytic gradients with respect to the phase parameters are identical to those of the reduced Gaussian likelihood function with the exception that Φ_λ is replaced by φ_λ^2 and that the gradients are negated due to the fact that we are now minimizing an error metric rather than maximizing a likelihood. We write the derivative of Eq. (4.6) with respect to the j^{th} phase parameter in the ξ^{th} sub-aperture as

$$\frac{\partial E(\varphi_\lambda, \boldsymbol{\alpha})}{\partial \alpha_{\xi,j}} = -8\pi \operatorname{Im}\left\{ \sum_{k=1}^{K} \sum_{\lambda} \frac{\varphi_\lambda^2}{\lambda} \sum_{f_u', f_v'} Z_{\xi,j}(f_u', f_v') P_{\xi,k,\lambda}(f_u', f_v') \right. \\ \times \exp\left[i \frac{\pi A_k}{\lambda B_k} \left(f_u'^2 + f_v'^2 \right) \right] \\ \left. \times \sum_{f_u, f_v \in \chi} Y_k^*(f_u, f_v) H_{k,\lambda}^*(f_u' - f_u, f_v' - f_v) \right\} \tag{4.18}$$

where $Z_{\xi j}$ is the j^{th} member of the basis set for the ξ^{th} sub-aperture, $P_{\xi k,\lambda}$ is the complex pupil function for the ξ^{th} sub-aperture and k^{th} diversity image at wavelength λ, A_k and B_k are the elements of the ABCD ray-transfer matrix that relates the pupil plane to the image plane [8], and Y_k and $H_{k,\lambda}$ are given by Eqs. (3.12) and (3.11), respectively.

The analytic gradient of the total error metric with respect to the gray-world spectral coefficient at wavelength λ_o is given by

$$\frac{\partial E(\varphi_\lambda,\boldsymbol{\alpha})}{\partial \varphi_{\lambda_o}} = -\sum_{f_u,f_v\in\chi}\frac{\partial}{\partial\varphi_{\lambda_o}}\frac{\left|\sum_{j=1}^{K} D_j(f_u,f_v)\sum_{\lambda}\varphi_\lambda^2 S^*_{j,\lambda}(f_u,f_v;\boldsymbol{\alpha})\right|^2}{\sum_{\ell=1}^{K}\left|\sum_{\lambda}\varphi_\lambda^2 S_{\ell,\lambda}(f_u,f_v;\boldsymbol{\alpha})\right|^2} + \frac{\partial}{\partial\varphi_{\lambda_o}}\kappa\left[1-\sum_{\lambda}\varphi_\lambda^2\right]^2. \tag{4.19}$$

The first term can be calculated in the same manner as the phase parameters, as described in Section 2.4. Distributing the derivative through the first term of Eq. (4.19) and collecting terms yields

$$\text{First Term} = -\sum_{f_u,f_v\in\chi}\sum_{k=1}^{K} Y_k(f_u,f_v)\frac{\partial}{\partial\varphi_{\lambda_o}}\sum_{\lambda}\varphi_\lambda^2 S_{k,\lambda}(f_u,f_v) + c.c. \tag{4.20}$$

where $Y_k(f_u,f_v)$ is given by Eq.(3.12). Evaluating the derivative gives

$$\begin{aligned}\text{First Term} &= -\sum_{f_u,f_v\in\chi}\sum_{k=1}^{K} Y_k(f_u,f_v)\sum_{\lambda} 2\varphi_\lambda\frac{\partial\varphi_\lambda}{\partial\varphi_{\lambda_o}} S_{k,\lambda}(f_u,f_v) + c.c.\\ &= -\sum_{f_u,f_v\in\chi}\sum_{k=1}^{K} Y_k(f_u,f_v)\,2\varphi_{\lambda_o} S_{k,\lambda_o}(f_u,f_v) + c.c. \qquad (4.21)\\ &= -2\,\text{Re}\left[\sum_{f_u,f_v\in\chi}\sum_{k=1}^{K} Y_k(f_u,f_v)\,2\varphi_{\lambda_o} S_{k,\lambda_o}(f_u,f_v)\right]\end{aligned}$$

where Re[...] denotes the real part of the argument and we've used the equality

$$\frac{\partial\varphi_\lambda}{\partial\varphi_{\lambda_o}} = \delta(\lambda,\lambda_o) = \begin{cases}1, & \lambda=\lambda_o\\ 0, & \lambda\neq\lambda_o\end{cases}. \tag{4.22}$$

Evaluating the second term of Eq. (4.19) yields

$$\text{Second Term} = 4\kappa\left[1-\sum_{\lambda}\varphi_\lambda^2\right]\varphi_{\lambda_o}. \tag{4.23}$$

Combining Eqs. (4.21) and (4.23) gives

$$\frac{\partial E(\varphi_\lambda, \boldsymbol{\alpha})}{\partial \varphi_{\lambda_o}} = -2\,\mathrm{Re}\left[\sum_{f_u, f_v \in \chi} \sum_{k=1}^{K} Y_k(f_u, f_v) 2\varphi_{\lambda_o} S_{k,\lambda_o}(f_u, f_v) \right] + 4\kappa \left[1 - \sum_{\lambda} \varphi_\lambda^2 \right] \varphi_{\lambda_o}. \tag{4.24}$$

4.8 References

1. J. Goodman, *Introduction to Fourier Optics, 3rd Ed.* (Roberts & Co, 2005).

2. L. P. Yaroslavsky, and H. J. Caulfield, "Deconvolution of multiple images of the same object," Appl. Opt. **33**, 2157-2162 (1994).

3. M. Born, and E. Wolf, *Principles of Optics, 7th Ed.* (Cambridge University Press, Cambridge, UK, 1999).

4. S. Thurman, and J. Fienup, "Multi-aperture Fourier transform imaging spectroscopy: theory and imaging properties," Opt. Express **13**, 2160-2175 (2005).

5. H. Nyquist, A. Telephone, T. Co, and N. York, "Certain topics in telegraph transmission theory," Proceedings of the IEEE **90**, 280-305 (2002).

6. D. A. Naylor, B. G. Gom, M. K. Tahic, and G. R. Davis, "Astronomical spectroscopy using an aliased step-and-integrate Fourier transform spectrometer," in *Millimeter and Submillimeter Detectors for Astronomy II, Proc. SPIE*, J. Zmuidzinas, W. S. Holland, and S. Withington, eds., **5498** (2004), pp. 685-694.

7. Provided through the courtesy of Jet Propulsion Laboratory, California Institute of Technology, Pasadena, California (http://aviris.jpl.nasa.gov/).

8. A. E. Siegman, *Lasers* (University Science Books, Sausalito, CA, 1986).

5 Experimental Demonstration of Phase Diversity on a Segmented System

In previous chapters we've presented both a broadband version of the phase diversity algorithm as well as an implementation of sub-aperture piston phase diversity. Both of these algorithms have been successfully demonstrated in simulation. In this Chapter we will present experimental results that verify these simulations.

In Section 5.1, we described the experimental setup, including modes of operation, scene projector details and descriptions of the MEMs deformable mirrors used in the experiments. In Section 5.2, we discuss the experimental procedure, including processing of pupil images, collection of phase-shifting interferometry frames and collection of phase diversity images. In Section 5.3, we present results of the experiments, comparing the phases estimated by the phase-diversity algorithm with phases measured with interferometry. Images are also reconstructed using a multi-frame Wiener filter. Section 5.4 summarizes the Chapter.

5.1 Experimental Setup

An experimental test bed based on a segmented MEMs deformable mirror (DM) has been constructed and is capable of several modes of operation. The MEMs DM can be actuated to introduce a phase error into the system which can then be estimated using phase retrieval or phase diversity. Phase-shifting interferometry can be performed as an independent measurement of the phase error to corroborate the

estimates. Also, pupil images can be captured to be used in the phase retrieval and phase-diversity algorithms.

Figure 5-1 shows a schematic of the test bed. In the following section each mode of the test bed will be discussed. Details of the scene projector will be discussed, as well as capabilities of the MEMs DM. Table 5-1 lists the relevant parameters of the optics used in the setup and Table 5-2 lists the relevant distances between elements of the setup.

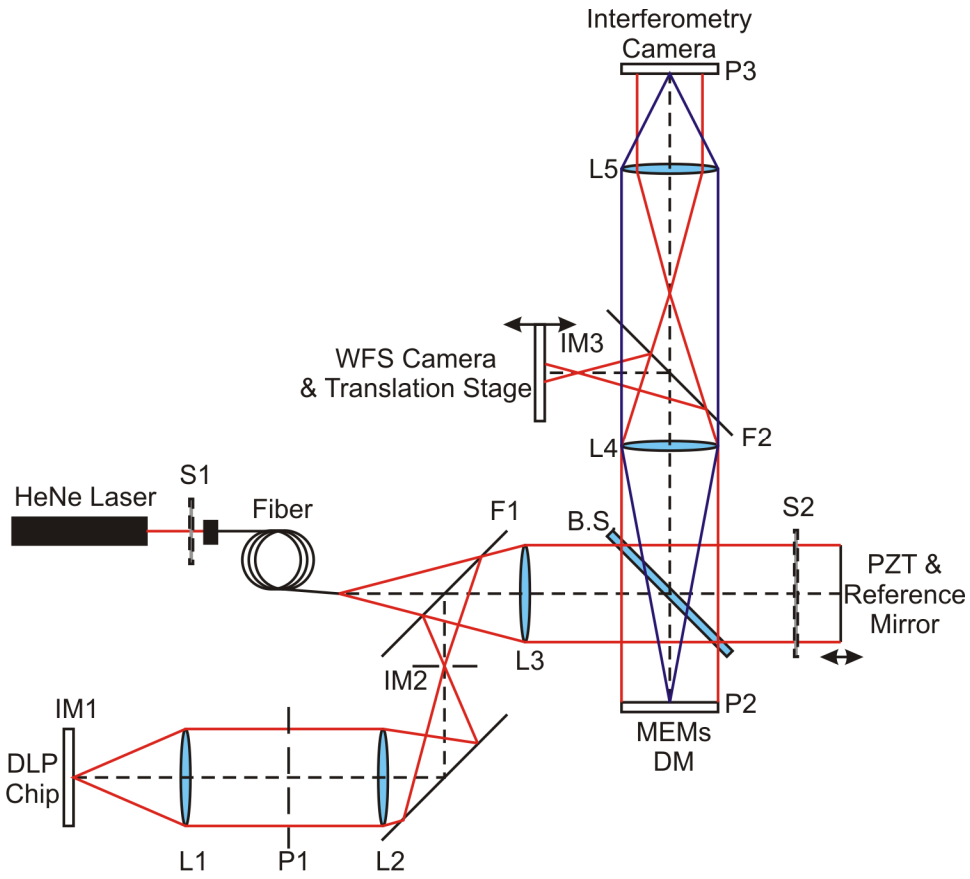


Figure 5-1: Schematic of experimental setup. Lenses are denoted by L prefix, shutters by S, flip-mirrors by F, conjugate pupil planes by P, and conjugate image planes by IM. DLP™: Digital Light Processing, PZT: piezo-electric transducer, WFS: wavefront sensing, MEMs DM: micro-electro-mechanical deformable mirror, B.S.: beam splitter.

Table 5-1: List of optics and associated parameters used in the experimental setup depicted in Figure 5-1.

Label	Description	Focal Length	Diameter	Comment
L1, L2, L4	Achromatic Doublet	150 mm	50.8 mm	Newport Part #: PAC086AR.14
L3	Achromatic Doublet	250 mm	25.4 mm	Newport Part #: PAC067AR.14
L5	Singlet	130 mm	38.1 mm	
F1, F2	Flat Metallic Mirror	N/A	25.4 mm	Newport Part #: 10Z40ER.1
P1	Iris	N/A	25.4 mm	Nominal Clear Aperture: 2.5 mm
BS	Plate Beamsplitter	N/A	50 × 70 mm	Linos Part #: G344171000

Table 5-2: Nominal distances for the experimental setup depicted in Figure 5-1.

IM1 – L1:	150 mm
L1 – P1:	150 mm
P1 – L2:	150 mm
L2 – IM2:	150 mm
IM2 – L3:	250 mm
Fiber exit – L3:	250 mm
L3 – P2:	250 mm
L3 – PZT:	250 mm
P2 – F2:	150 mm
F2 – IM3:	150 mm
F2 – L5:	280 mm
L5 – P3:	130 mm

5.1.1 Phase Retrieval

Figure 5-2 shows a schematic of the test bed when configured for collecting phase retrieval frames. A HeNe laser-fed fiber point source is collimated by lens L3 and incident on the MEMs DM. The reflected light from the DM is then focused onto the wavefront sensing (WFS) camera by lens L4 and flip-mirror F2. The WFS camera is

a Q-Imaging Retiga 2000R CCD. Table 5-3 shows the relevant specifications of this camera. The WFS camera can be scanned through focus via a computer controlled translation stage.

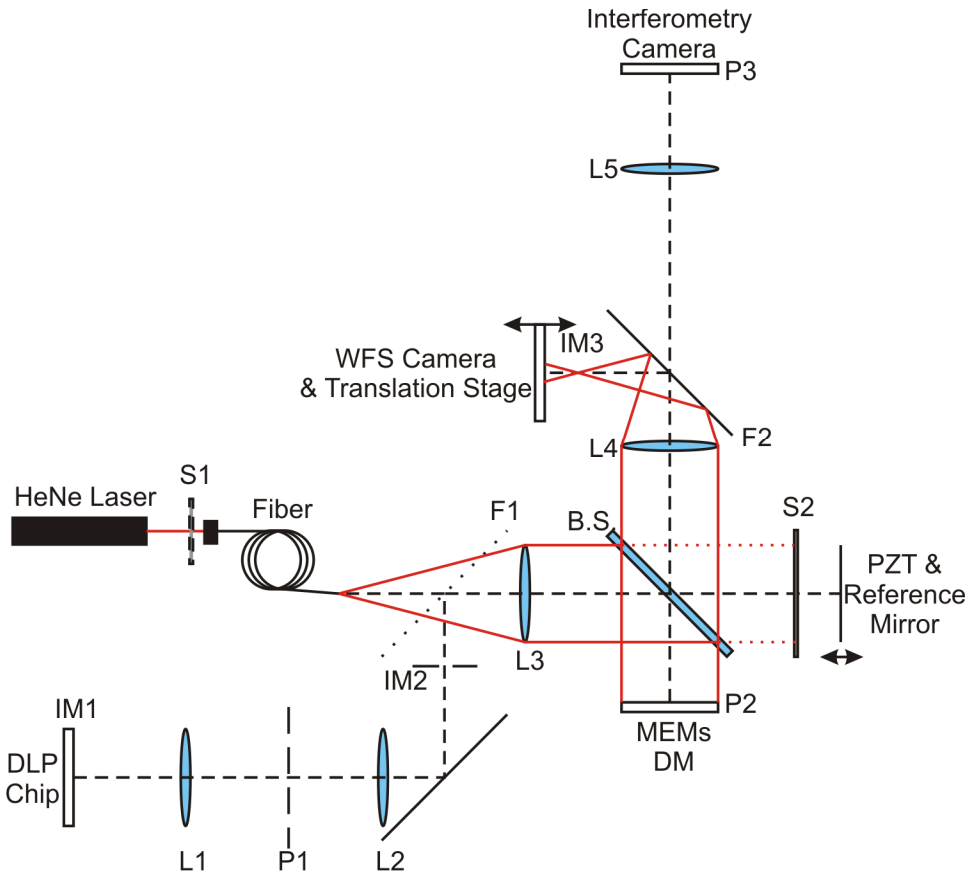


Figure 5-2: Schematic of test bed in phase retrieval mode. Shutter S2 is closed to block the reference arm of the interferometer and flip-mirror F1 is lowered to allow point source illumination.

Table 5-3: Specifications for the Q-Imaging Retiga 2000R CCD wavefront sensing camera.

Chip Size (pixels):	1600 x 1200
Chip Size (mm):	11.84 x 8.88
Pixel Size (μm):	7.4
Well Depth (e^-):	40,000
Bits:	12
Read Noise (e^-):	16
Dark Current (e^-/s/pix):	0.5

5.1.2 Phase-shifting Interferometry

Figure 5-3 shows a schematic of the test bed when configured for phase-shifting interferometry. From the configuration shown in Figure 5-2, flip-mirror F2 is lowered and shutter S2 is opened. Lowering flip-mirror F2 allows the pupil to be re-imaged by lens L3 onto the interferometry camera. The interferometry camera is a Lumenera Lu120M CMOS sensor, the relevant specifications of which are shown in Table 5-4. The reference mirror is mounted on a piezo-electric transducer (PZT) that shifts the mirror by fractions of a wavelength to collect phase-shifting frames. Thirteen frames are collected and processed using a Zygo® 13-step algorithm [1] to determine the phase error introduced by the DM.

5.1.3 Pupil Imaging

Figure 5-4 shows a schematic of the test bed when configured for pupil imaging for point-source illumination. From the configuration shown in Figure 5-3, shutter S2 is closed. The interferometry camera is conjugate to the DM pupil plane. A pupil image can be gathered when the scene projector is the illumination source simply by raising flip-mirror F1 and closing shutter S1 to block the laser illumination.

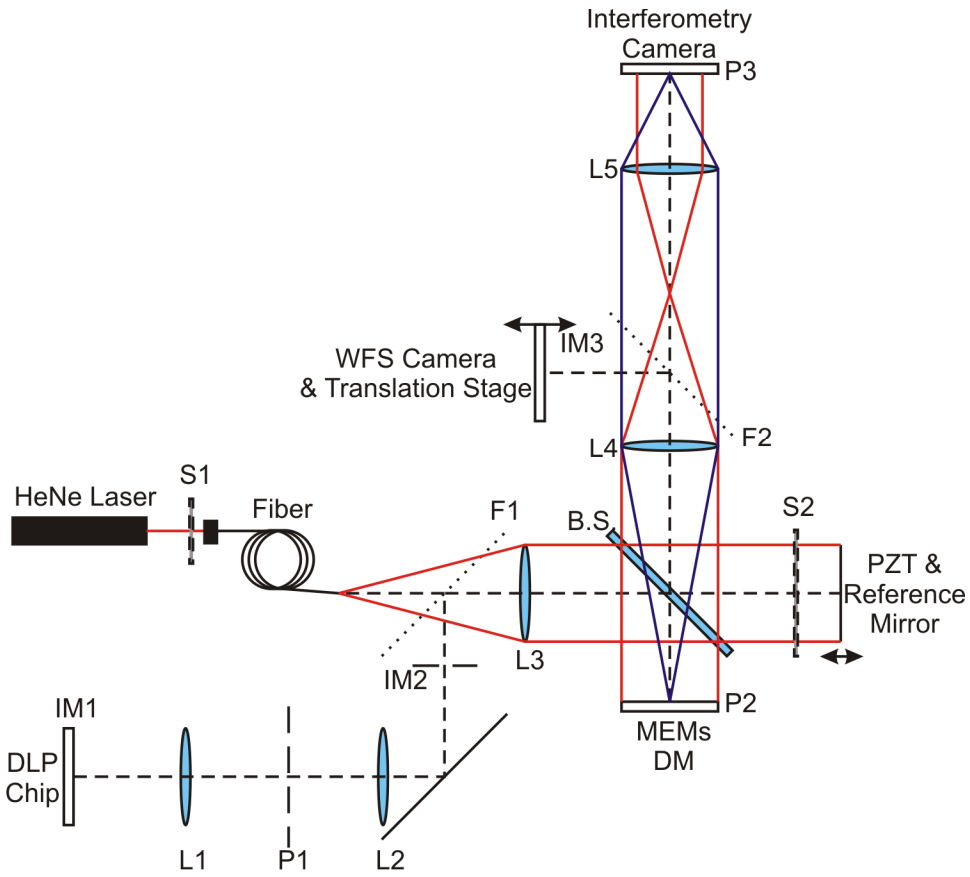


Figure 5-3: Schematic of test bed in interferometry mode. Shutter S2 is opened to access the interferometer reference arm and flip-mirror F2 is lowered to access the interferometry camera.

Table 5-4: Specifications for the Lumenera Lu120m CMOS interferometry camera.

Chip Size (pixels):	1280 x 1024
Chip Size (mm):	8.6 x 6.9
Pixel Size (μm):	6.7
Bits:	8

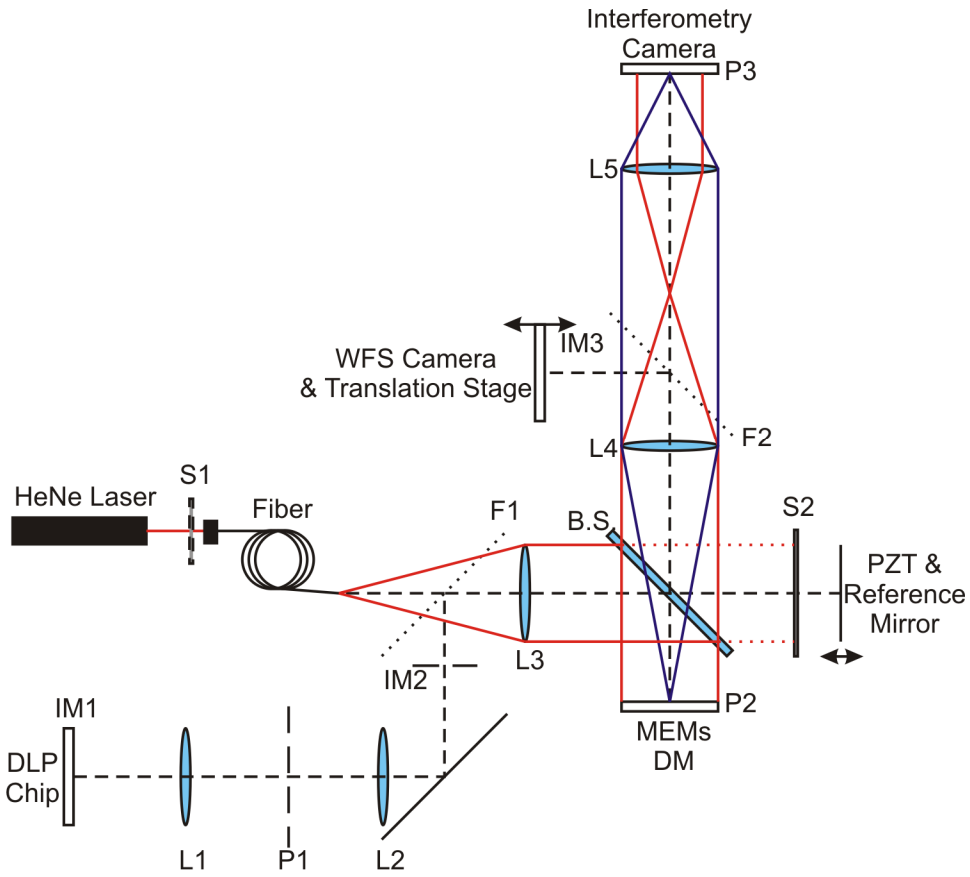


Figure 5-4: Schematic of test bed in pupil imaging mode when point-source illumination is used. Shutter S2 is closed to block the interferometer reference arm.

5.1.4 Phase Diversity

Figure 5-5 shows the test bed when configured for collecting phase diversity frames. From the configuration in Figure 5-4, flip-mirrors F1 and F2 are both raised and shutter S1 is closed. The scene projector is a modified BenQ MP622 Digital Light Processing (DLP™) projector (see Section 5.1.5). The relevant specifications of the projector are given in Table 5-5. The DLP™ chip is conjugate to the WFS camera as well as image plane IM2. A filter wheel near IM2 holds various spectral filters that

can used to create several bandwidths. Table 5-6 lists the spectral filters and the corresponding bandwidths.

Plane P1 is conjugate to the DM pupil plane. An iris in plane P1 is used as the limiting aperture of the system, restricting the pupil to the inner 19 segments of the MEMS DM.

Table 5-5: Specifications for the BenQ MP622 DLP™ scene projector.

Image Size (pixels):	1024 x 768
DLP Chip Size (mm):	11.19 x 8.40
DLP Pixel Size (μm):	10.93
Color Wheel Segments:	5 (R, G, B, Y, Clear)

Table 5-6: Color filters and associated bandwidths. The shortpass filters are combined with the 500 nm longpass filter to create the bandwidths shown.

Description	Center/Cut-off Wavelength (nm)	$\Delta\lambda$ (nm)	$\Delta\lambda/\lambda$
Laser-line Filter	632.8	1	0.2%
Bandpass Filter	550	10	1.8%
Bandpass Filter	550	40	7.3%
Bandpass Filter	550	80	14.6%
Longpass Filter	500	N/A	N/A
Shortpass Filter	600	100	18.2%
Shortpass Filter	650	150	26.1%
Shortpass Filter	700	200	33.3%

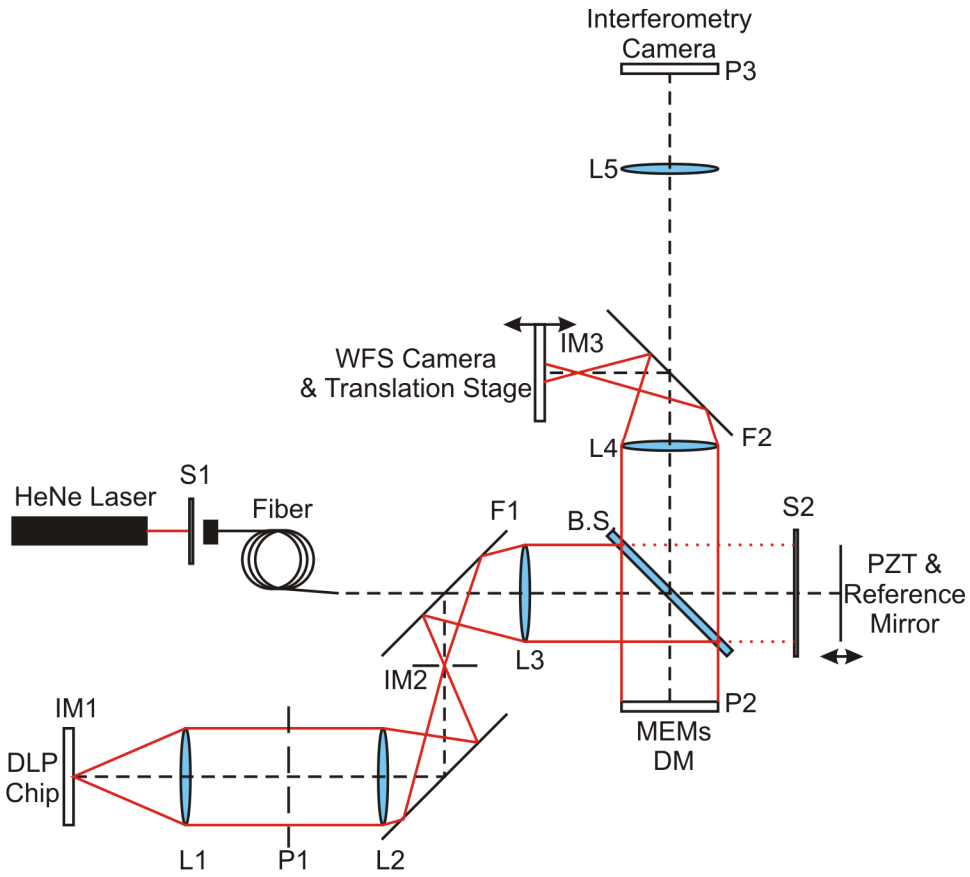


Figure 5-5: Schematic of test bed in phase diversity mode. Flip-mirror F1 is raised to allow extended scene illumination. Shutters S1 and S2 are both closed to block laser illumination and the interferometer reference arm, respectively. Flip-mirror F2 is raised to allow imaging by the WFS camera. An iris in conjugate pupil plane P1 restricts the pupil to the inner 19 segments of the DM. Spectral filters are included near plane IM2 to control the spectral bandwidth of the object being observed.

5.1.5 Scene Projector

A BenQ MP622 DLP™ projector was purchased to be used as a scene projector for the phase diversity arm of the test bed. A DLP™ projector uses a digital mirror device (DMD) to turn individual pixels on or off by tilting micro-mirrors to either

reflect light into the projection optics (“on”) or into a baffle (“off”). The mirror motion is synchronized with a spinning color wheel to produce color images.

A few modifications had to be made before the projector could be used in the test bed. Phase-diversity algorithms assume incoherent illumination of the object and therefore require uniform illumination of the pupil. Figure 5-6 shows an image of the pupil of the projector where it is clearly seen that the illumination is not uniform.

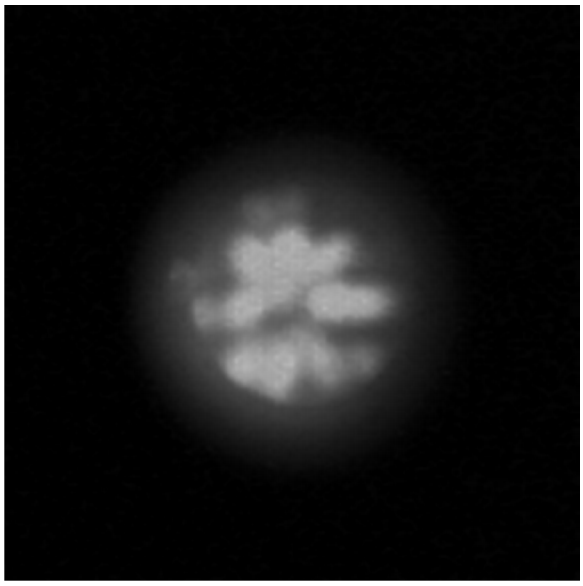


Figure 5-6: Image of projector pupil showing non-uniform illumination.

Since the manufacturer would not provide specific designs of the illumination and projection optics, the projector was dismantled to better understand its operation. Figure 5-7 shows a schematic of the projector optical path. An integrating rod is used to homogenize the source. It is assumed Köhler illumination is used and that an image of the exit port of the integrating rod is formed in the pupil of the projection

lens. A piece of ground glass was fixed at the exit port of the integrating rod to further diffuse the source illumination. Figure 5-8 shows an image of the projector pupil after the ground glass diffuser was installed, showing a marked improvement in the uniformity of the pupil illumination.

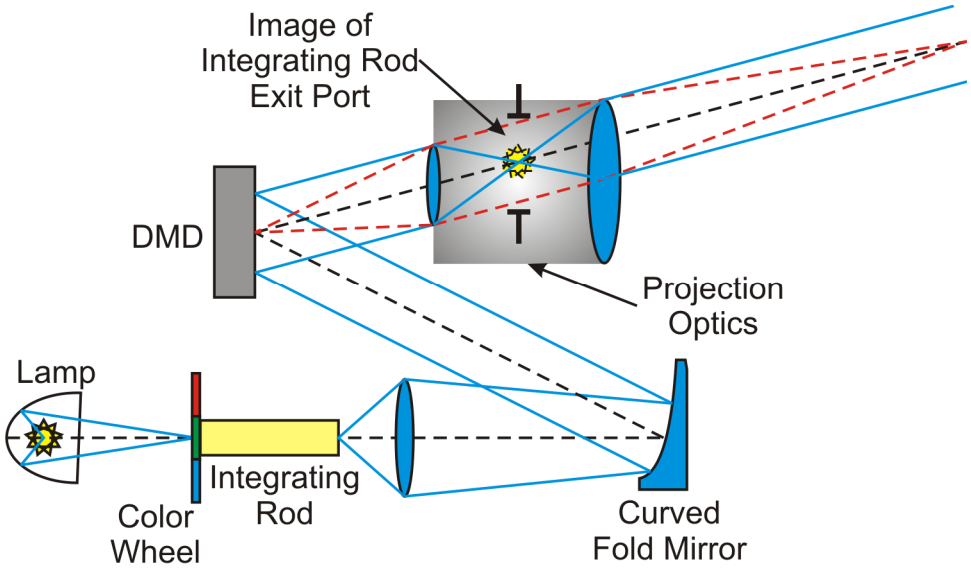


Figure 5-7: Schematic of projector optics. Köhler illumination is used such that an image of the source (exit port of integrating rod) is formed in the pupil of the projection optics. The projection optics are designed to project the image above the DMD.

The second modification to the projector was the removal of the projection optics. The projector was optimized for producing a magnified version of the DMD a large distance in front of and slightly above the projector, which does not match our application. Instead, by removing the projection optics, direct access to the DMD chip is available and an optical relay can be designed with the magnification desired for the test bed.

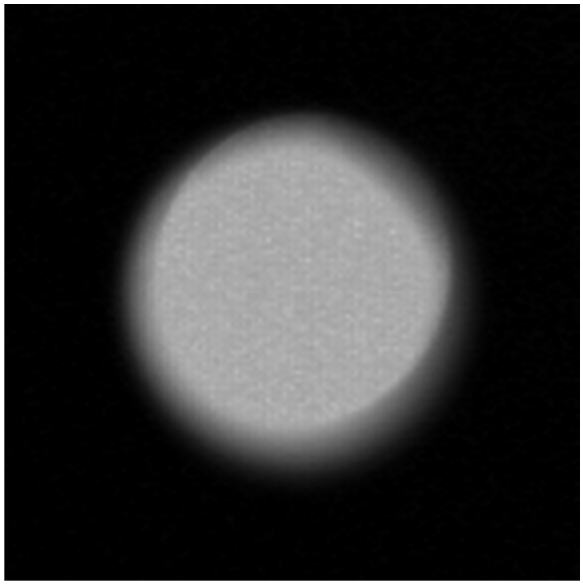


Figure 5-8: Image of projector pupil after ground-glass diffuser was installed. Illumination uniformity is greatly improved.

5.1.6 MEMs Deformable Mirrors

Two sets of MEMs deformable mirrors were acquired to be used in the test bed as the system under test. Both DMs are hexagonally segmented with piston, tip and tilt control on each segment. The first set of mirrors was manufactured by Sandia National Laboratories. Table 5-7 lists the relevant specifications of the device. While the actuators on the Sandia mirror have a large amount of stroke (27 μm), the mirror segments themselves are thin layers of silicon deposited directly over the actuator structure. This configuration causes the underlying structure to print through to the mirror surface, causing a significant amount of surface roughness. Furthermore, the top layer of silicon is deposited in such a way that the mirror segments are pre-stressed to account for a gold coating process, causing a large

degree of cupping in uncoated mirrors. The Sandia mirrors and drive electronics are also uncalibrated, making it difficult to repeatedly and accurately position the mirror as desired. For example, flattening the mirror requires a complicated iterative process that is dependent on the order in which segments were moved due to inter-actuator influence.

The second set of mirrors was purchased from Iris AO, Inc. Table 5-7 lists the relevant specifications of the Iris AO mirrors. While the stroke of the Iris AO mirrors is not as large as that of the Sandia mirrors, the actuators and drive electronics are fully calibrated, allowing for segments to be more accurately positioned very repeatedly. Furthermore, Iris AO attaches thick mirror surfaces to each actuator, reducing print through and improving surface roughness and surface figure. Typical segment surface quality is less than 20 nm RMS. The Iris AO mirrors were used for all of the experiments presented in this thesis.

Table 5-7: Specifications for the Sandia National Laboratories and Iris AO, Inc. hexagonally-segmented MEMs DM.

	Sandia DM	**Iris AO DM**
Number of Segments:	61	37
Flat-to-flat Segment Size (μm):	497	606
Mirror Array Size (mm):	~4	~4
Actuator Stroke (μm):	27	5
Fill Factor:	99%	98.6%

5.2 Experimental Procedure

5.2.1 Mirror Actuation

Before collecting image frames for phase diversity, the mirror segments were actuated to introduce between 0.1 λ_o and 0.2 λ_o RMS of random pistons, tips, and tilts into the system.

5.2.2 Pupil Orientation and Image

A pupil image was collected using the projector as the illumination source. The pupil image was rotated, thresholded, and scaled to be used as the known aperture in the phase-diversity algorithm.

The WFS camera was translated inside of the nominal focus as far as possible (~2.5 λ_o PV defocus) where regions of the PSF could be mapped to regions of the pupil. By obstructing portions of the pupil and seeing where the PSF was shadowed, the gross orientation of the pupil with respect to the WFS camera was determined.

Diffraction from the segment array causes a very distinct spot pattern in the PSF. By measuring the angle of the spots in the PSF with respect to the WFS camera pixels, the fine rotation of the pupil with respect to the WFS camera was measured to be about 1.9º. In support of this measurement, several rotation angles between 0º and 5º were tested in the phase-diversity algorithm and 1.9º produced the best phase estimate. Figure 5-9 shows an example PSF of a point source and angle measurement.

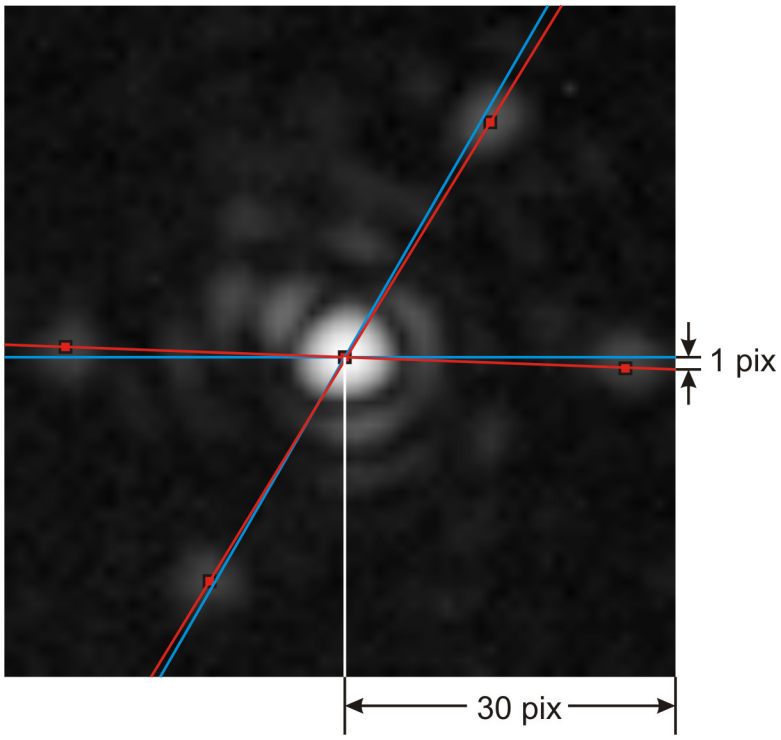


Figure 5-9: Example PSF of a point source used for determining the rotation angle of the pupil with respect to the WFS camera. Red squares identify the centers of the diffraction peaks, as determined by the maximum pixel value, and are connected by red lines. Blue lines are drawn along the expected angles of diffraction if the relative rotation between pupil and camera were 0º. Calculating the angle between the red and blue lines yields a relative rotation of 1.9º.

After rotation, the pupil image was thresholded to produce a binary mask of the iris in conjugate pupil plane P1. The binary mask was then zero-padded to the appropriate Q and resampled to match the array size of the images used in the phase diversity algorithm. Q is defined as

$$Q = \frac{\lambda F_{\#}}{\Delta_u} \tag{5.1}$$

where λ is the wavelength, Δ_u is the detector pixel size and the $F_{\#}$ is the image domain f-number. When Q = 2, the intensity in the image plane is Nyquist sampled.

For λ = 632.8 nm, $Q \approx 4$ for the test bed, and it depends on the diameter of the iris in the pupil plane P1.

The rotated, scaled, and thresholded binary image of the iris was then applied to the segmented array generated in the phase diversity algorithm to create the pupil amplitude. Figure 5-10 shows an example of the pupil amplitude. Notice the edges of the hexagonal array are shaded to better model the smooth edges of the mirrors.

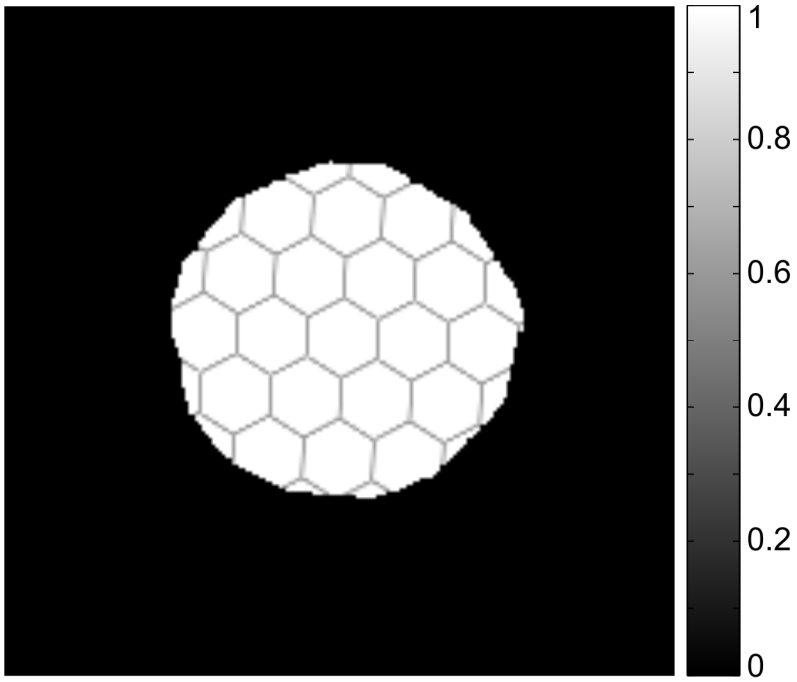


Figure 5-10: Pupil amplitude as constructed by the phase diversity algorithm. The irregular edge is due to the iris used as the limiting aperture in the conjugate pupil plane. The underlying hexagonal array uses shaded edges to approximate the smooth edges of the real mirror. The entire pupil is rotated by approximately 1.9º clockwise. Only the center 200 x 200 pixel portion of the array is shown.

5.2.3 Interferometry

Interferometry data was collected before and after the phase diversity image frames were collected. This allows any drift in the system during frame collection to be quantified. In experiments for which sub-aperture piston diversity was implemented, interferometry data was collected after each phase diversity frame in order to confirm the actual pistons that were introduced. As described in Section 5.1.2, thirteen phase-shifting frames were collected and processed by a Zygo® 13-step algorithm [1]. The resulting phases were then unwrapped as necessary on a segment-by-segment basis. Figure 5-11 shows a 0.008 λ_o RMS difference between interferometry data before and after phase diversity image collection. This places a limit on how well we can expect the phase estimated by phase diversity to agree with the phase measured by the interferometer.

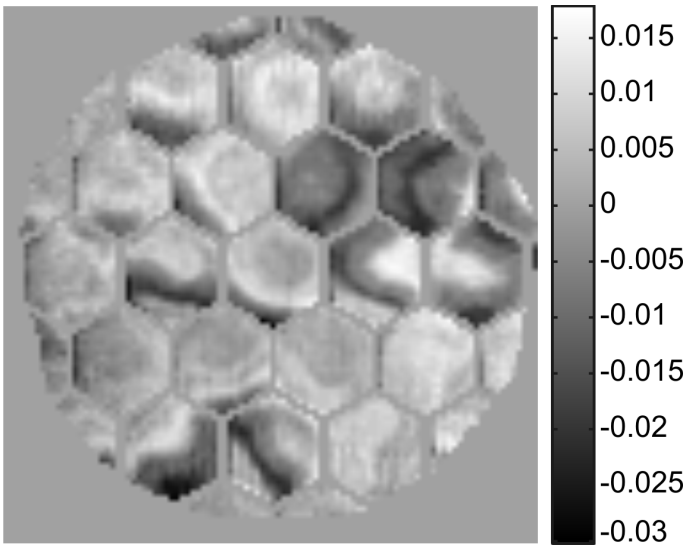


Figure 5-11: Difference between interferometry data taken before and after phase-diversity images. The scale is in units of waves and the RMS difference is 0.008 λ_o.

5.2.4 Phase Diversity Image Collection and Processing

Figure 5-12(a) shows the color image input to the projector for display. Figure 5-12(b) shows an example raw frame collected by the WFS camera. The rectangular halo surrounding the region of interest (ROI) is due to the fact that the DLP™ projector cannot produce a pure black background. To remove this halo, 10 "background frames" were collected where the projector displayed a black screen. The 10 frames were averaged and subtracted from the average of 10 image frames. Figure 5-13(a) shows an averaged background frame and Figure 5-13(b) shows an averaged, background-subtracted frame. Since the ROI occupied only a small portion of the CCD, the background-subtracted images were cropped to about a quarter of their size from the 1600 × 1200 array.

When the WFS camera is translated through focus, the image moves across the CCD since the axis of the translation stage is not perfectly aligned with the optical axis. To mitigate this effect, the diversity images were first registered using an efficient, sub-pixel registration algorithm [2]. Also, global tip and tilt phases for each diversity image were allowed to vary in the phase-diversity algorithm so that any residual translations of the images would be accounted for.

Exposure times for each image were set by integrating until the peak pixel in the image was equal to 80% of the 40,000 e^- well depth.

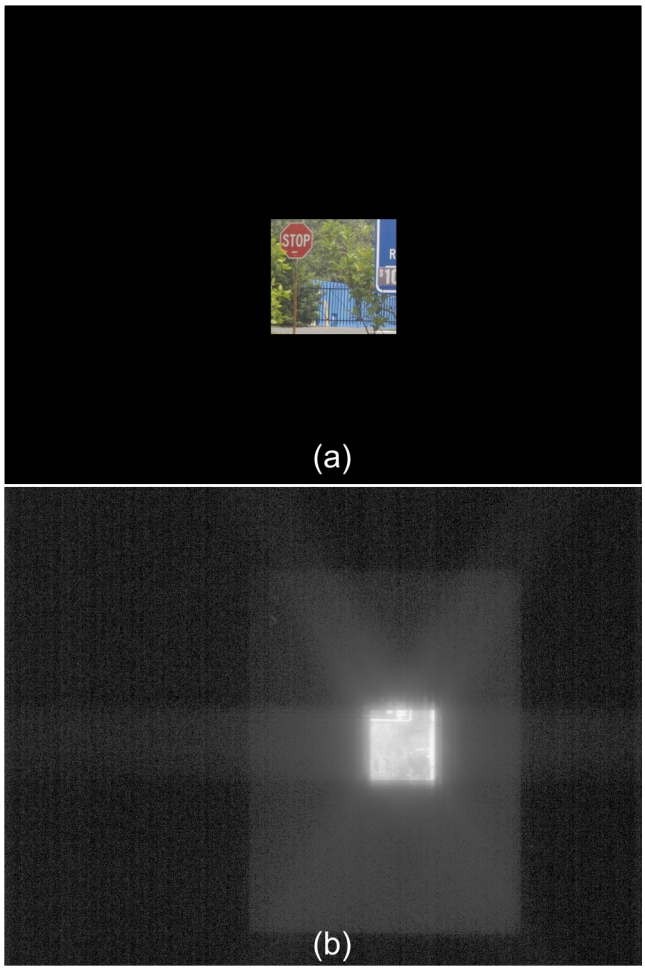


Figure 5-12: (a) True image input to projector. (b) Detected zero-diversity raw frame. The intensity in (b) is stretched to the 1/5 power to show diffraction streaks, background, and noise. The image in (b) appears rotated since the projector is mounted on its side.

5.2.5 Phase-Diversity Algorithm

For all of the experiments performed, the regularized reduced Gaussian (RRG) metric of Section 2.2.2 was used. Initially, only piston, tip, and tilt were estimated for each sub-aperture and only the second order Zernike polynomials were estimated across the global aperture (phase-diversity algorithms are insensitive to global piston, tip, and tilt). As the algorithm progressed, gradually, the number of parameters was

increased until the Zernike polynomials up to and including 5^{th} order were estimated on both the sub-apertures and the global aperture. Each time the number of parameters was increased, the object and noise power spectra were re-estimated using the new estimate to the system phase.

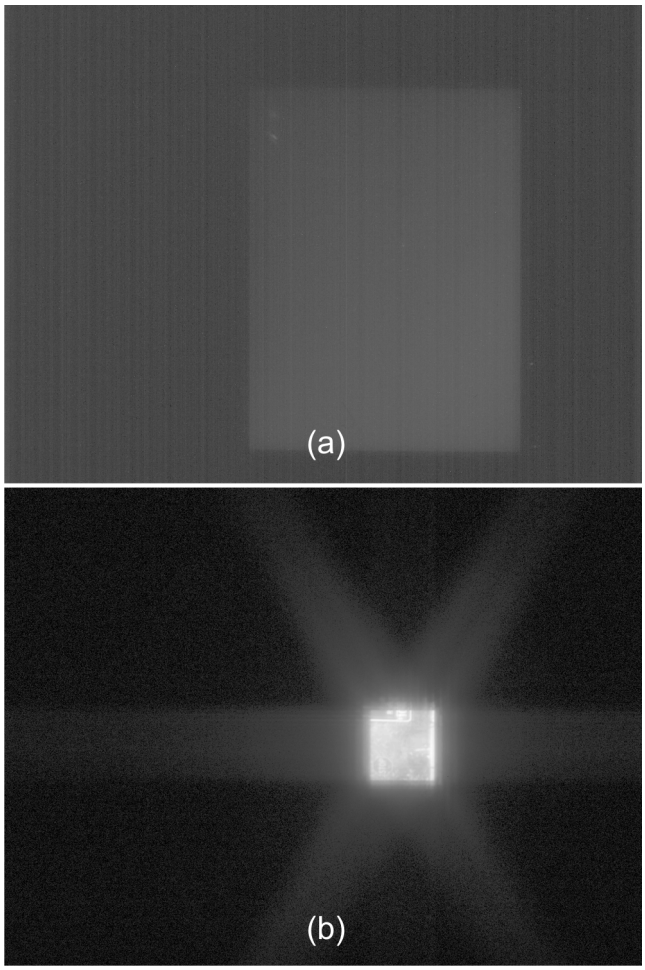


Figure 5-13: (a) Average of ten background frames. (b) Average of ten image frames minus the averaged background frame of (a). Both intensities are stretched to the 1/5 power.

When focus diversity was used to gather the phase diversity images, a focus correction factor was estimated for each diversity plane. This accounts for the inherent uncertainty in the actual translation of the WFS camera.

5.3 Experimental Results

Four experiments were performed to validate monochromatic focus diversity, broadband focus diversity, monochromatic sub-aperture piston phase diversity (SAPPD) and broadband SAPPD.

5.3.1 Monochromatic Focus Diversity

The image of Figure 5-12(a) was projected with the 632.8 nm laser-line filter in place. Five diversity images were captured by translating the WFS camera from the nominal focus position by [–15, –10, 0, 10, 15] mm. These focus positions correspond to [–1.315, –0.877, 0, 0.877, 1.315] λ_o peak-to-valley (PV) of quadratic phase, where the relationship between z-translation and PV quadratic phase in waves is given by

$$\phi_{PV}^{foc} = \frac{\Delta z}{8\lambda_o F_\#^2} \tag{5.2}$$

where Δz is the distance translated from the nominal focus, λ_o is the center wavelength and $F_\#$ is the image space f-number, which was 47.5 for these experiments. Figure 5-14 shows the five diversity images after background subtraction, registration and cropping. Each image is 400 × 400 pixels.

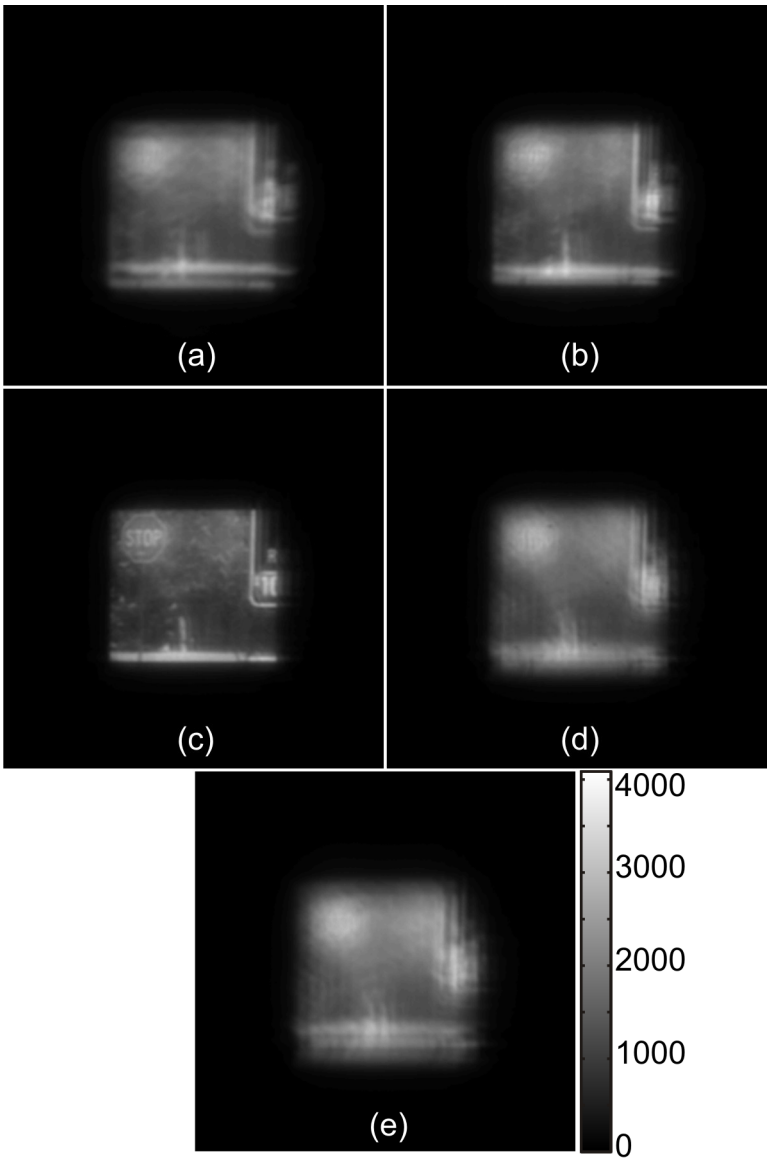


Figure 5-14: The five diversity images input to the phase diversity algorithm. (a) –1.315 λ_o PV defocus, (b) –0.877 λ_o, (c) 0 λ_o, (d) 0.877 λ_o, (e) 1.315 λ_o. Scale has units of bits.

Figure 5-15(a) shows the unwrapped, measured interferometry phase minus global piston, tip, tilt, and focus. Figure 5-15(b) shows the phase estimated by the phase-diversity algorithm minus global piston, tip, tilt, and focus. Figure 5-15(c)

shows the difference phase. The estimated and measured phases differ by 0.06 λ_o RMS.

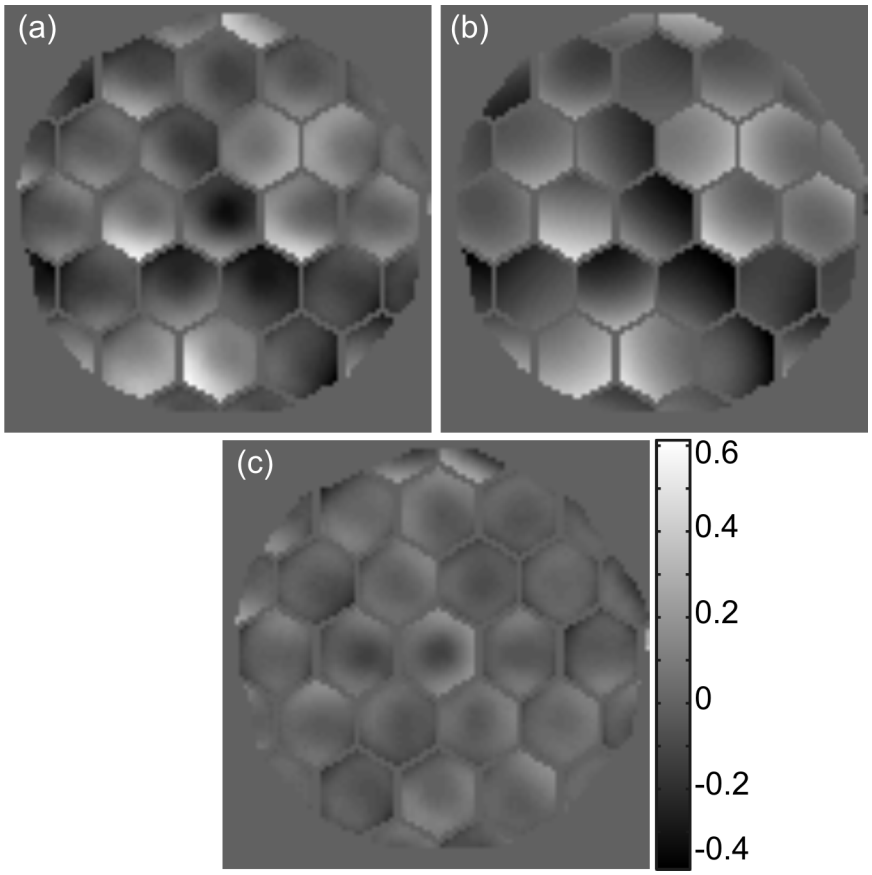


Figure 5-15: Phase estimation results for monochromatic focus diversity. (a) Phase measured by interferometry, 0.15 λ_o RMS. (b) Phase estimated by focus diversity, 0.16 λ_o RMS. (c) Difference phase, 0.06 λ_o RMS. Scale has units of waves.

Figure 5-16(a) shows the image input to the DLP™ projector. Figure 5-16(b) shows the in-focus detected image. Figure 5-16(c) shows the image restored by a multi-frame Wiener-filter using Eq. (2.16), using the phase estimated by the phase-diversity algorithm. Figure 5-16(d) shows the image restored using the phase measured by phase-shifting interferometry and using only the in-focus detected

image. Only a single frame reconstruction is used for the interferometry phase because the global tip, tilt and focus terms with respect to the interferometry camera are unknown for the out-of-focus images. The phase-diversity algorithm estimates these quantities with respect to the wavefront sensing camera during the optimization so they can be included in the multi-frame reconstruction. The Wiener-filter tuning constant, *c*, was equal to 1. The image reconstructed using the estimated phase has similar quality to the image reconstructed from the interferometry phase.

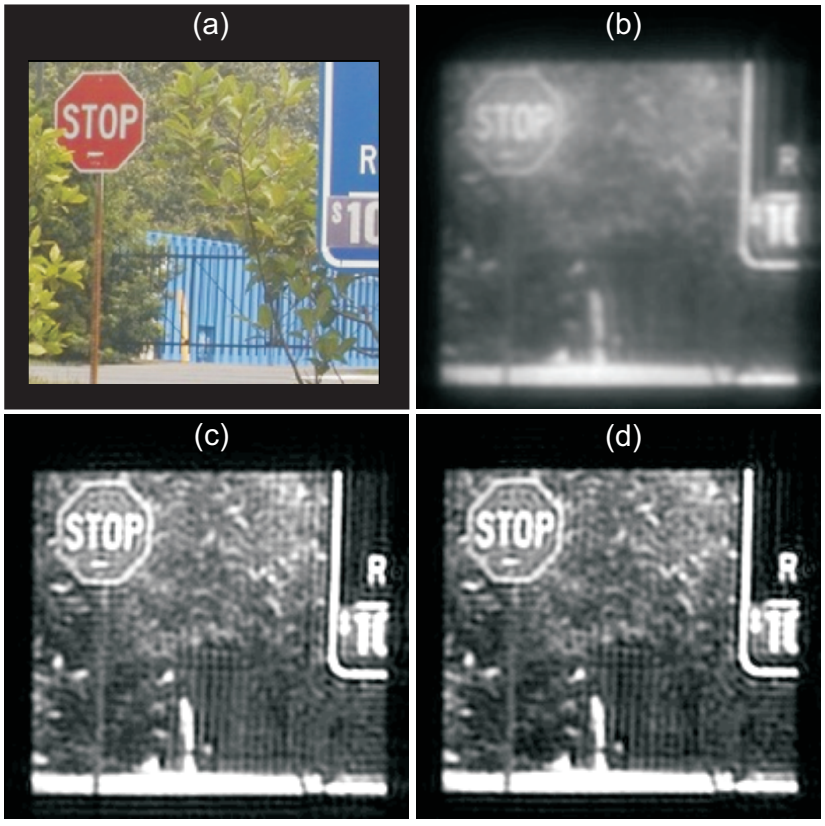


Figure 5-16: Monochromatic focus diversity images. (a) Image input to projector. (b) Zero-diversity image detected through 632.8 nm laser-line filter. (c) Monochromatic, multi-frame reconstructed image using estimated phase. (d) Monochromatic, single-frame reconstructed image using interferometry phase.

5.3.2 Broadband Focus Diversity

The image of Figure 5-12(a) was projected with the 500 nm longpass and 650 nm shortpass filters combined to create a 150 nm bandwidth centered on 575 nm. Five diversity images were captured by translating the WFS camera from the nominal focus position by [–15, –10, 0, 10, 15] mm. These focus positions correspond to [–1.484, –0.989, 0, 0.989, 1.484] λ_o peak-to-valley (PV) of quadratic phase using Eq. (5.2) and an image space f-number of 46.9. For this experiment, five gray-world spectral coefficients were used to sample the spectrum and were assumed to be uniform.

Figure 5-17(a) shows the unwrapped, measured interferometry phase minus global piston, tip, tilt, and focus. Figure 5-17(b) shows the phase estimated by the phase diversity algorithm minus global piston, tip, tilt, and focus. Figure 5-17(c) shows the difference phase. The estimated and measured phases differ by 0.06 λ_o RMS.

Figure 5-18 (a) shows the image input to the DLP™ projector. Figure 5-18(b) shows the in-focus detected image. Figure 5-18(c) shows the image restored by a multi-frame Wiener-filter using Eq. (2.16), using the phase estimated by the phase-diversity algorithm. Figure 5-18(d) shows the image restored using a single-frame Wiener filter and phase measured by phase-shifting interferometry. The Wiener-filter tuning constant, c, was equal to 1. Again, the image reconstructed from the estimated phase is of similar quality to the image reconstructed from the phase measured by interferometry. Notice the slats of the fence are clearly visible in the reconstructed

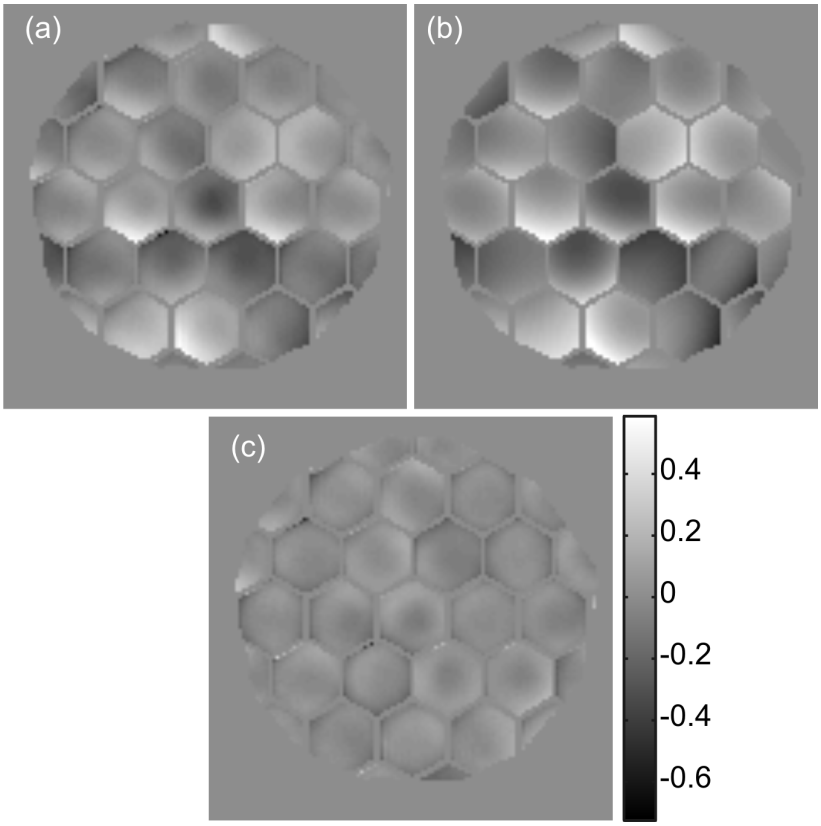


Figure 5-17: Phase estimation results for broadband focus diversity ($\Delta\lambda/\lambda_o$ = 26.1%). (a) Phase measured by interferometry, 0.15 λ_o RMS. (b) Phase estimated by focus diversity, 0.18 λ_o RMS. (c) Difference phase, 0.06 λ_o RMS. Scale has units of waves.

images, but not in the measured, in-focus image. We believe that slats are visible in Figure 5-18(c) and Figure 5-18(d) and not in the monochromatic reconstructed image of Figure 5-16(c) and Figure 5-16(d), despite having a similar degree of phase correction, because the narrow-band filter in the monochromatic experiment blocked much of the blue light coming from the container behind the fence. The black fence on a nearly black background is therefore a very low contrast region of the scene, in both the detected and reconstructed images. However, in the broadband experiment,

much of the blue light from the container passes through the spectral filter and the black fence on a bright blue background creates a high contrast region that is easily visible in the reconstructed image.

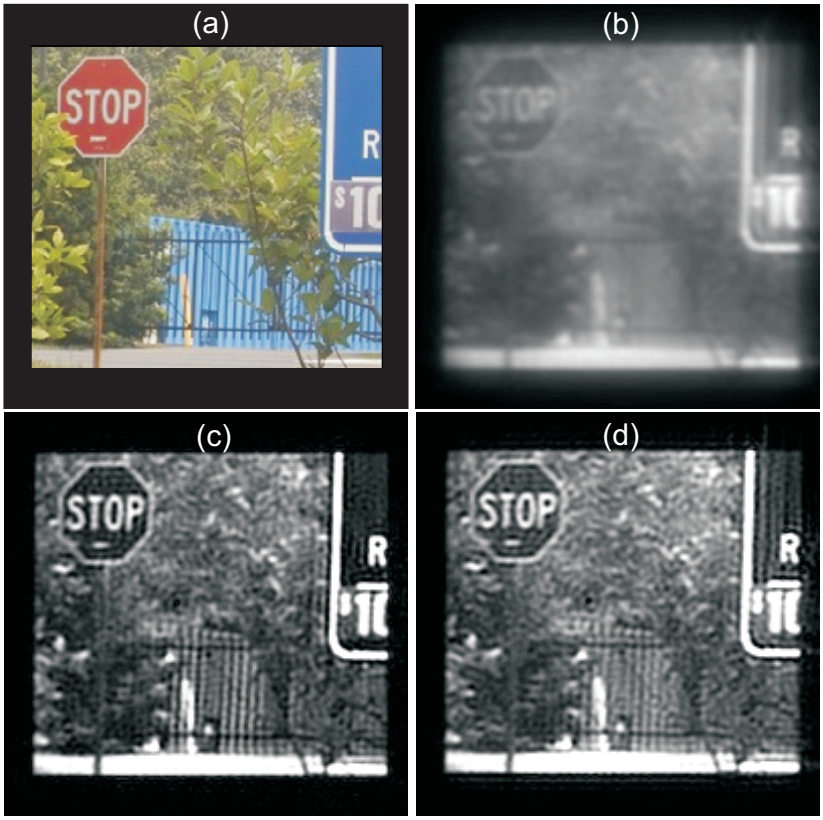


Figure 5-18: Broadband focus diversity images. (a) Image input to projector. (b) Zero-diversity image detected through a 150 nm bandpass filter. (c) Panchromatic, gray, multi-frame image reconstructed using estimated phase. (d) Panchromatic, gray, single-frame image reconstructed using interferometry phase. Note the slats of the fence are clearly visible in (c) and (d) but are washed out in the detected aberrated image (b).

5.3.3 Monochromatic Sub-aperture Piston Phase Diversity

The image of Figure 5-12(a) was projected through the 632.8 nm laser-line filter. Five diversity images were captured using sub-aperture piston phase diversity. The

implementation used for these experiments is shown in Figure 5-19. Triplets of segments were pistoned by [0.19, 0.09, 0, –0.09, –0.19] μm, resulting in twice that amount of piston OPD, [0.38, 0.19, 0, –0.19, –0.38] μm. With the center wavelength of 632.8 nm, the pistons are [0.6, 0.3, 0, –0.3, –0.6] λ_o.

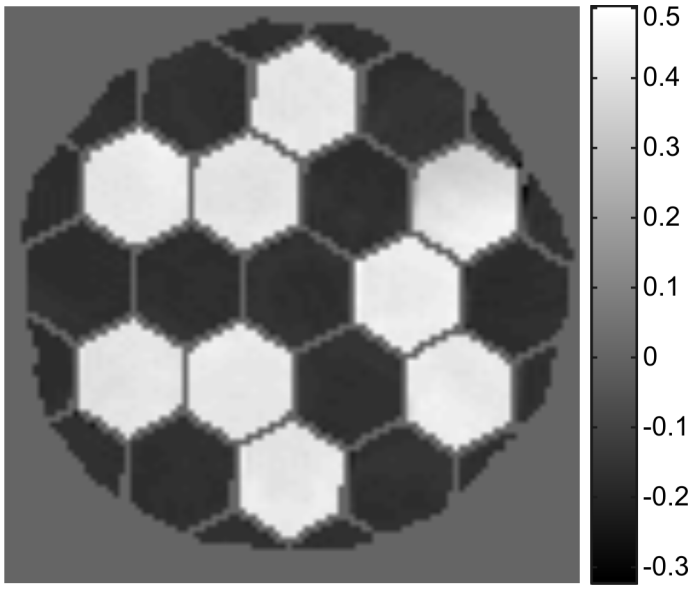


Figure 5-19: Example of sub-aperture piston diversity implementation. Shown is the difference between the 0.6 λ_o and 0 λ_o unwrapped interferometry frames. The scale is in units of waves.

Figure 5-20(a) shows the unwrapped, measured interferometry phase minus global piston, tip, tilt, and focus. Figure 5-20(b) shows the phase estimated by the phase-diversity algorithm minus global piston, tip, tilt, and focus. The difference phase, shown in Figure 5-20(c), reveals two segments that are in error by a wavelength. For monochromatic illumination, piston errors of integer wavelengths are invisible to both the phase-diversity algorithm and single-wavelength interferometry. Adding a single wave of piston to the appropriate segments of the

estimated phase reduces the error between the estimation and interferometry to 0.08 λ_o RMS, and is shown in Figure 5-20(d).

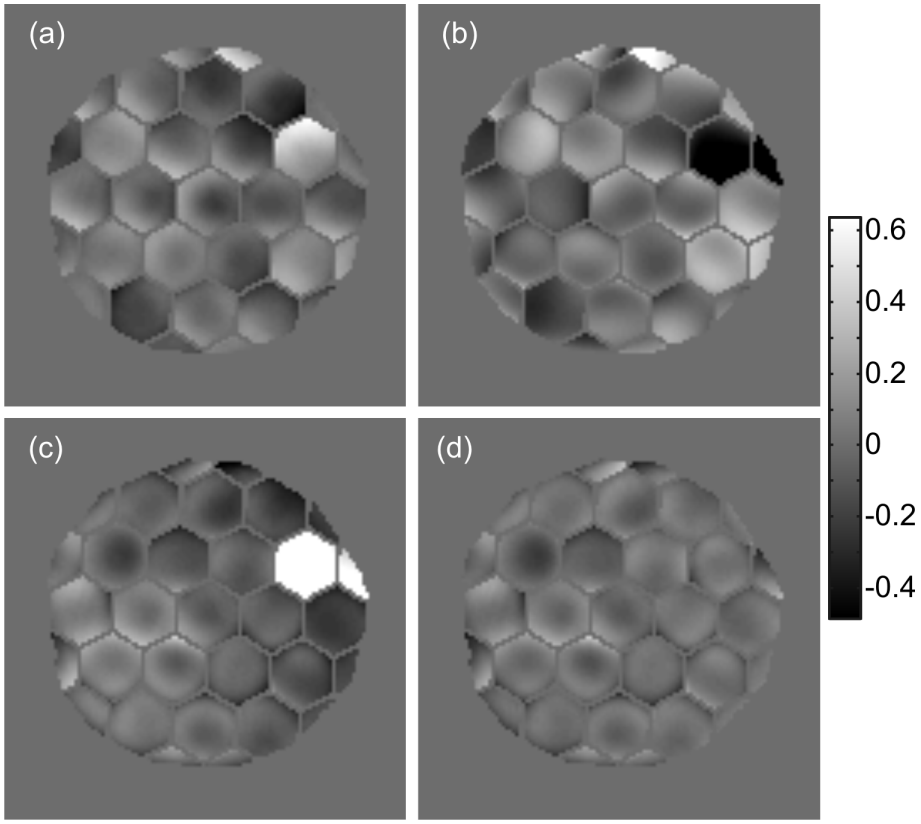


Figure 5-20: Phase estimation results for monochromatic sub-aperture piston phase diversity (λ_o = 632.8 nm). (a) Phase measured by interferometry, 0.15 λ_o RMS. (b) Phase estimated by focus diversity, 0.22 λ_o RMS. (c) Difference phase, 0.23 λ_o RMS. Visible in the difference phase are two segments that are in error by a single wavelength. (d) Difference between measured interferometry phase and estimated phase for monochromatic SAPPD after the piston errors in (c) were corrected. The RMS error is 0.08 λ_o. Scale has units of waves.

Figure 5-21(a) shows the image input to the DLP™ projector. Figure 5-21(b) shows the in-focus detected image. Figure 5-21(c) shows the image restored by a multi-frame Wiener-filter using Eq. (2.16) using the phase estimated by the phase diversity

algorithm. Figure 5-21(d) shows the image restored using only the zero-diversity image and the phase measured by the interferometer. The Wiener-filter tuning constant, c, was equal to 1. Here, it appears that the image reconstructed using the interferometry phase is slightly sharper than the image reconstructed using the estimated phase, but is also degraded more by noise artifacts. The difference in the image reconstructions may be due to the larger error in the phase estimation of 0.08 λ_o RMS.

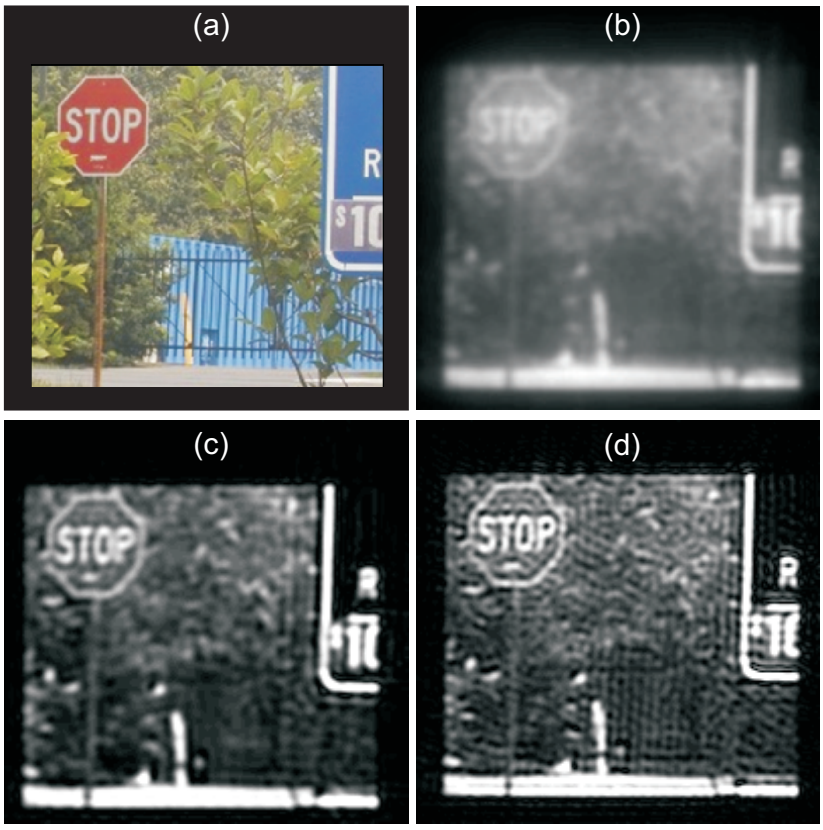


Figure 5-21: Monochromatic SAPPD images. (a) Image input to projector. (b) Zero-diversity image detected through 632.8nm laser-line filter. (c) Monochromatic, multi-frame reconstructed image using estimated phase. (d) Monochromatic, single-frame reconstructed image using interferometry phase.

5.3.4 Broadband Sub-aperture Piston Phase Diversity

The image of Figure 5-12(a) was again projected with the 500 nm longpass and 650 nm shortpass filters combined to create a 150 nm bandwidth centered on 575 nm. Five diversity images were captured using sub-aperture piston phase diversity. The implementation used for these experiments is shown in Figure 5-19. Triplets of segments were pistoned by [0.173, 0.086, 0, –0.086, –0.173] μm, resulting in twice that amount of piston OPD, [0.345, 0.173, 0, –0.173, –0.345] μm. With the center wavelength of 575 nm, the pistons are [0.6, 0.3, 0, –0.3, –0.6] λ_o. Five gray-world spectral coefficients were used to sample the spectrum and were assumed to be uniform.

Figure 5-22(a) shows the unwrapped, measured interferometry phase minus global piston, tip, tilt, and focus. Figure 5-22(b) shows the phase estimated by the phase-diversity algorithm minus global piston, tip, tilt, and focus. Figure 5-22(c) shows the difference phase. The estimated and measured phases differ by 0.06 λ_o RMS.

Figure 5-23(a) shows the image input to the DLP™ projector. Figure 5-23(b) shows the in-focus detected image. Figure 5-23(c) shows the image restored by a multi-frame Wiener-filter using Eq. (2.16), using the phase estimated by the phase-diversity algorithm. Figure 5-23(d) shows the image restored from the phase measured by interferometry using only the zero-diversity image. The Wiener-filter tuning constant, c, was equal to 1. Again, both reconstructed images appear to be of

similar quality. The slats of the fence are once again sharply defined in the reconstructed image where they are not visible at all in the detected, zero-diversity image. Similar to the focus diversity experiment, the slats of the fence are much more visible in the broadband reconstructed image in Figure 5-23(c) than the monochromatic reconstructed image in Figure 5-21(c). Again, this is due to the broadband spectral filter allowing the blue light of the container to pass and creating a high contrast feature with the fence.

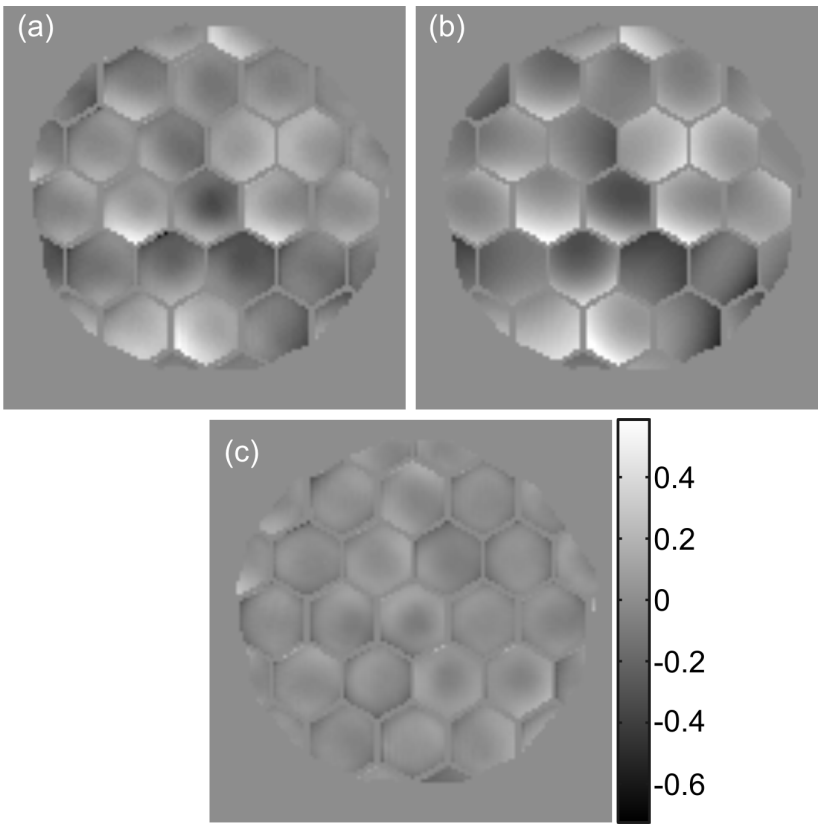


Figure 5-22: Phase estimation results for broadband sub-aperture piston diversity ($\Delta\lambda/\lambda_o$ = 26.1%). (a) Phase measured by interferometry, 0.15 λ_o RMS. (b) Phase estimated by sub-aperture piston diversity, 0.18 λ_o RMS. (c) Difference phase, 0.06 λ_o RMS. Scale has units of waves.

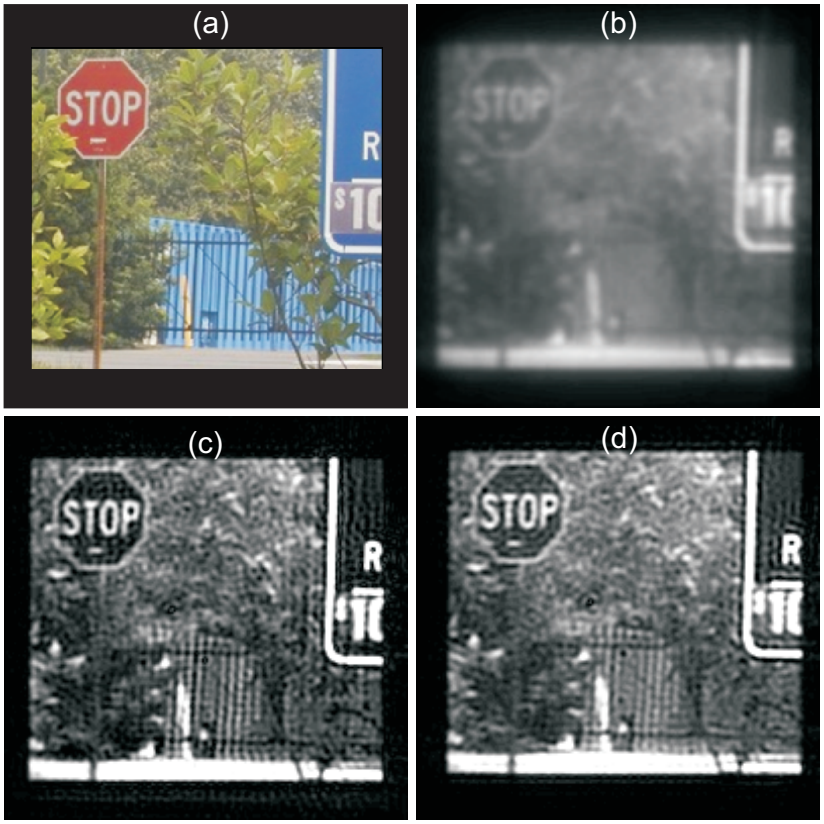


Figure 5-23: Broadband SAPPD images. (a) Image input to projector. (b) Zero-diversity image detected through a 150 nm bandpass filter. (c) Panchromatic, gray, multi-frame reconstructed image using the estimated phase. (d)Panchromatic, gray, single-frame reconstructed image using the phase measured by inteferometry. Note the slats of the fence are clearly visible in (c) and are washed out in the detected aberrated image (b).

Figure 5-24 shows how image reconstruction depends on the Wiener-filter tuning constant, *c*. In general, lower values of *c* emphasize sharpening of the image, but introduce noise artifacts. Conversely, large values of *c* emphasize noise suppression but result in blurred imagery.

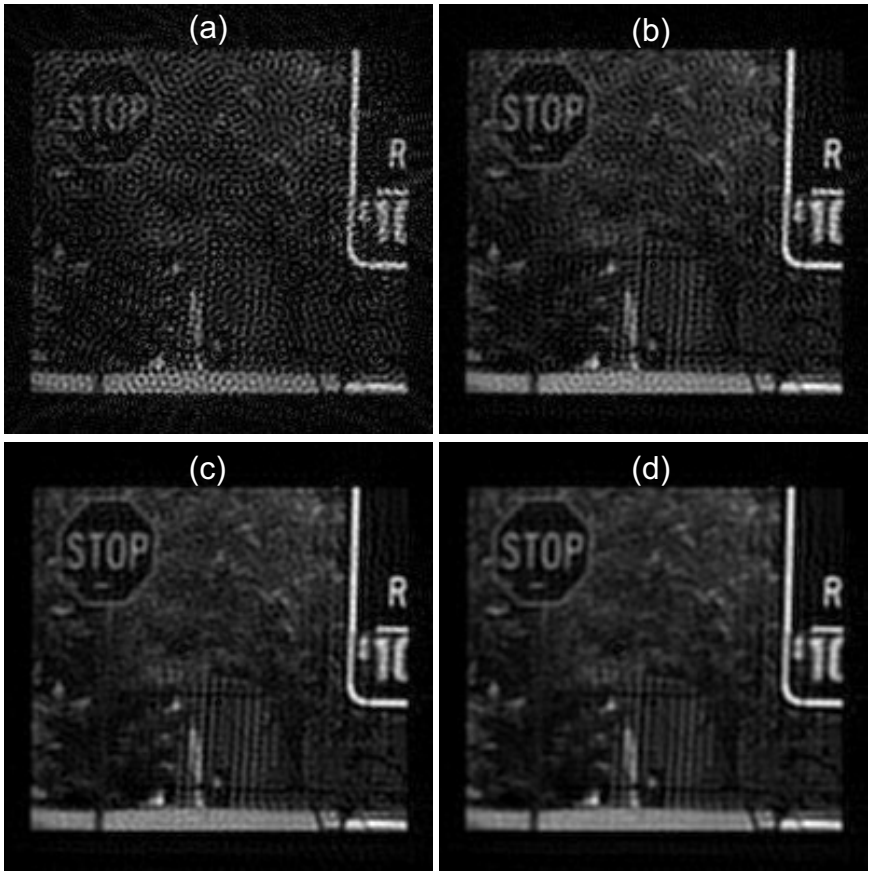


Figure 5-24: Broadband SAPPD image reconstructions for various values of Wiener-filter tuning constant, *c*: (a) *c* = 0.001, (b) *c* = 0.1, (c) *c* = 1, (d) *c* = 5.

5.4 Summary

An experimental test bed was constructed capable of sequentially capturing phase diversity images and phase-shifting interferometry data. The interferometry data was used to corroborate the phase estimated by the phase-diversity algorithm. A hexagonally segmented MEMs DM was used to introduce calibrated segment-level piston, tip, and tilt errors into the system. A DLP™ projector was modified and used to project a known color scene into the system.

We successfully demonstrated monochromatic and broadband phase diversity, implementing both focus diversity and sub-aperture piston phase diversity (SAPPD). Focus diversity was implemented by translating the wavefront sensing camera through focus. SAPPD was implemented by pistoning triplets of segments on the DM. In each case the estimated phase agreed with the interferometry to 0.06 λ_o RMS, except for monochromatic SAPPD which agreed to 0.08 λ_o RMS.

5.5 References

1. P. de Groot, "Phase shifting interferometer and method for surface topography measurement," (US Patent 5,473,434, 1995).

2. M. Guizar-Sicairos, S. T. Thurman, and J. R. Fienup, "Efficient subpixel image registration algorithms," Opt. Lett. **33**, 156-158 (2008).

6 Multi-field Wavefront Sensing

6.1 Introduction

It is not uncommon for multiple scientific instruments to share the focal plane of an optical system, particularly in the case of large, astronomical telescopes. For example, the James Webb Space Telescope (JWST) will have four scientific instruments populating the field of view. When wavefront correction is performed, it is necessary to sense and correct the field-dependent aberration terms so that the images detected by off-axis instruments are not degraded by an unacceptable wavefront error.

For deployable systems that allow active control of the primary mirror, such as the JWST, it is important to characterize the off-axis wavefront aberrations to accurately know the position of the deployed optics. For example, if the secondary mirror of a three mirror anastigmat is deployed such that it is decentered and tilted with respect to the nominal position, astigmatism that is linearly dependent on the field height and coma that is independent of field height will be introduced [1]. If wavefront sensing is performed only for a single field point, the primary mirror can be actuated to correct both of these aberrations. However, instruments viewing a different portion of the field may still be highly aberrated, as illustrated in Figure 6-1. Multi-field wavefront sensing can be performed to distinguish between field-dependent aberrations that are caused by a deployment error and field-dependent aberrations that are inherently part of the optical design, allowing the former to be corrected.

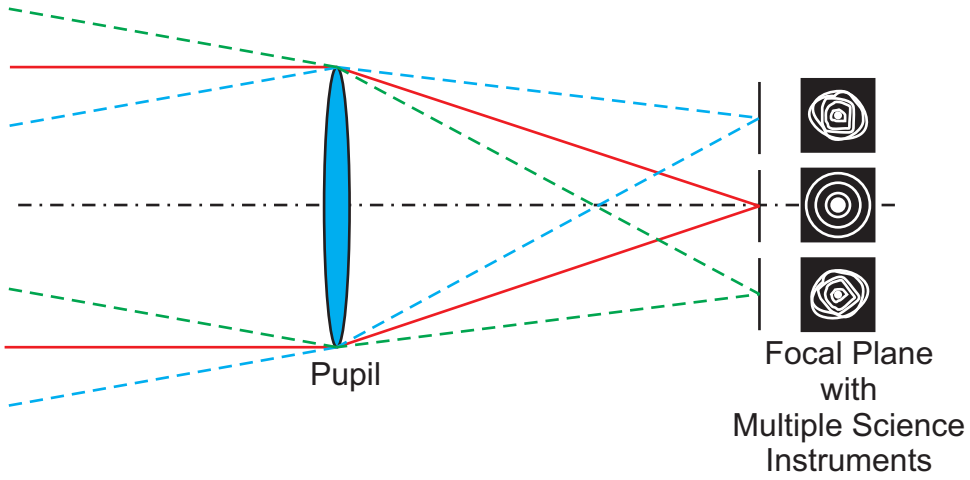


Figure 6-1: Schematic of a system with multiple science instruments in the focal plane. Even though the wavefront may be well corrected for the on-axis science instrument, other instruments in the FOV may be unacceptably aberrated.

Current techniques of phase-diverse phase retrieval (PDPR) use images from a single point source to determine the wavefront error for that point source location only. By performing PDPR separately on point sources across the field of view, one can eventually build a model for the field dependent aberrations in the system.

We propose a method of multi-field wavefront sensing in which the images from multiple point sources are used simultaneously in the PDPR algorithm to estimate the field-dependent aberrations. This method uses a technique of parameterizing the phase, proposed by Kwee and Braat, which uses a double Zernike expansion to introduce field dependence into the wavefront [2].

In this section we compare two methods of performing multi-field wavefront sensing: directly estimating the field-dependent wavefront for all point-sources using PDPR, or indirectly constructing a field-dependent model from a number of field-

independent PDPR estimates. Information-theoretic Cramer-Rao bounds are calculated to determine the theoretical limit of each technique. Digital simulations are also performed to compare each technique in practice.

In Section 6.2 we present a model for incorporating the field-dependence of the wavefront into the phase parameterization and use this model in a PDPR algorithm discussed in Section 6.3.

In Section 6.4, we introduce the information theoretic Cramer-Rao bounds and compute for the two methods of estimation.

In Section 6.5, we describe digital simulations that were run to compare PDPR estimation to predictions made by the CRBs. Results are presented in Section 6.6 and a summary is presented in Section 6.7.

6.2 Field-dependent Wavefront Model

For any single point in the field of view (FOV), the wavefront in the pupil can be written as an expansion of a basis set,

$$W(x,y) = \sum_{j=1}^{J} \alpha_j Z_j(x,y) \tag{6.1}$$

where (x, y) are the pupil coordinates, α_j is the coefficient of the j^{th} basis function, Z_j is the j^{th} member of the basis set and J is the total number of terms in the basis set. Typically, Zernikes are used for the polynomials because individual terms represent balanced optical aberrations. Since Eq. (6.1) does not depend on the field position of the point source, then a different set of coefficients is needed to characterize the

wavefront at a different field point. Therefore, a field-dependent wavefront can be written as

$$W\left(\xi_f,\eta_f;x,y\right)=\sum_{j=1}^{J}\alpha_{j,f}Z_j\left(x,y\right) \tag{6.2}$$

where (ξ_f, η_f) are the coordinates of a single discrete field point. The set of coefficients $\boldsymbol{\alpha} = \{\boldsymbol{\alpha}_f\}$ represents the wavefront at $\{(\xi_f, \eta_f)\}$, the set of discrete points in the FOV. Using a sufficient number of field points, a model of the wavefront at all field points can be extrapolated or interpolated.

Kwee and Braat [2] derived a set of field dependent polynomials in the form of a double Zernike expansion, such that the field dependent wavefront can be written in the form

$$W\left(\xi,\eta;x,y\right)=\sum_{\ell=1}^{L}\sum_{j=1}^{J}\beta_{j,\ell}Z_\ell\left(\xi,\eta\right)Z_j\left(x,y\right) \tag{6.3}$$

where L is the number of terms in the Zernike expansion over the field coordinates and the coefficients $\boldsymbol{\beta}$ now characterize the wavefront for every point in the field.

A relationship between $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ can be found by setting Eqs. (6.2) and (6.3) equal to each other:,

$$\sum_j\alpha_{j,f}Z_j\left(x,y\right)=\sum_{\ell=1}^{L}\sum_{j=1}^{J}\beta_{j,\ell}Z_\ell\left(\xi_f,\eta_f\right)Z_j\left(x,y\right) \tag{6.4}$$

$$\alpha_{j,f}=\sum_{\ell=1}^{L}\beta_{j,\ell}Z_\ell\left(\xi_f,\eta_f\right). \tag{6.5}$$

Equation (6.4) can be written in matrix notation as

$$\boldsymbol{\alpha}_{J\times F} = \boldsymbol{\beta}_{J\times L}\mathbf{M}_{JL\times JF} \tag{6.6}$$

where $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ are vectors of length $J \times F$ and $J \times L$, respectively, and $\mathbf{M}$ is a matrix of dimension $JL \times JF$, where J is the number of polynomials in the pupil expansion, L is the number of polynomials in the field expansion and F is the number of discrete field points. Conversely, it is possible to compute the β coefficients from the α coefficients by inverting Eq. (6.6) as

$$\boldsymbol{\alpha}_{J\times F}\mathbf{M}^{-1}_{JF\times JL} = \boldsymbol{\beta}_{J\times L}\,. \tag{6.7}$$

If the transformation matrix, $\mathbf{M}$, is not square, it can be inverted using the Moore-Penrose [3] pseudo-inverse.

6.3 Multi-field Phase-diverse Phase Retrieval

It is possible to estimate the coefficients, $\boldsymbol{\beta}$, that characterize the wavefront for every point in the field in two ways. First, one can estimate a set of $\boldsymbol{\alpha}$ coefficients for a number of different field positions and then obtain $\boldsymbol{\beta}$ with Eq. (6.7). Alternatively, one can use Eq. (6.3) in the forward model of the PDPR algorithm and use the set of off-axis PSFs jointly to directly estimate the $\boldsymbol{\beta}$ coefficients.

Using Eq. (6.2) or Eq. (6.3), the pupil function corresponding to field point (ξ_f, η_f) can be written as

$$P\left(x, y; \xi_f, \eta_f\right) = \left|A\left(x, y\right)\right| \exp\left[i\frac{2\pi}{\lambda} W\left(x, y; \xi_f, \eta_f\right)\right] \tag{6.8}$$

where $|A(x,y)|$ is the pupil amplitude which we assume does not change with field position. The pupil function is propagated to the k^{th} image plane by a Fresnel-like transform given by

$$h_k\left(u,v;\xi_f,\eta_f\right)=\exp\left[i\frac{\pi D_k}{\lambda B_k}\left(u^2+v^2\right)\right]\iint\left\{P\left(x,y;\xi_f,\eta_f\right)\right.$$
$$\left.\times\exp\left[i\frac{\pi A_k}{\lambda B_k}\left(x^2+y^2\right)\right]\exp\left[-i\frac{2\pi}{\lambda B_k}\left(xu+yv\right)\right]dxdy\right\} \tag{6.9}$$

where $h_k(u, v;\xi_f, \eta_f)$ is the k^{th} coherent impulse response function of the f^{th} field point, (u, v) are the image plane coordinates, A_k, B_k and D_k are the elements of the ray-transfer matrix that relates the pupil plane to the k^{th} image plane [4], and λ is the wavelength. The point spread functions (PSFs) are then given by

$$s_k\left(u,v;\xi_f,\eta_f\right)=\left|h_k\left(u,v;\xi_f,\eta_f\right)\right|^2. \tag{6.10}$$

As described in Section 1.3.2, one implementation of PDPR is to minimize an error metric using a nonlinear optimization routine. For this study we chose to minimize a squared difference of amplitudes metric, given by

$$E\left(\boldsymbol{\alpha}\right)=\sum_{u,v}\sum_{f=1}^{F}\sum_{k=1}^{K}\left[\sqrt{d_k\left(u,v;\xi_f,\eta_f\right)}-\left|h_k\left(u,v;\xi_f,\eta_f;\boldsymbol{\alpha}\right)\right|\right]^2 \tag{6.11}$$

or

$$E\left(\boldsymbol{\beta}\right)=\sum_{u,v}\sum_{f=1}^{F}\sum_{k=1}^{K}\left[\sqrt{d_k\left(u,v;\xi_f,\eta_f\right)}-\left|h_k\left(u,v;\xi_f,\eta_f;\boldsymbol{\beta}\right)\right|\right]^2 \tag{6.12}$$

where $d_k(u, v;\xi_f, \eta_f)$ is the k^{th} detected intensity PSF for the f^{th} field point and $h_k(u, v;\xi_f, \eta_f)$ is the estimated coherent impulse response function for the k^{th} image

plane that depends on either the field-dependent $\boldsymbol{\alpha}$ coefficients or the field-independent $\boldsymbol{\beta}$ coefficients.

6.4 Cramer-Rao Bound Analysis

A useful method to compare the two techniques of multi-field wavefront sensing is to calculate the information-theoretic Cramer-Rao bounds (CRB). For an unbiased estimator, the CRBs provide a bound on the minimum variance associated with estimating a particular quantity in the presence of noise [5-9]. They are independent of the method of estimation and only depend on the system model and noise statistics. For an unbiased estimator, the CRBs are defined as

$$\sigma_m^2 \geq \left[\mathbf{F}\right]_{m,m}^{-1} \tag{6.13}$$

where σ_m^2 is the variance associated with the error in estimating the m^{th} quantity and $\mathbf{F}$ is the Fisher Information Matrix (FIM), the elements of which are given by

$$\left[\mathbf{F}\right]_{m,n} = -\left\langle \frac{\partial^2 \mathbf{L}}{\partial \omega_m \partial \omega_n} \right\rangle \tag{6.14}$$

where $\mathbf{L}$ is the log-likelihood function for the given noise model and ω_m and ω_n are the m^{th} and n^{th} members of $\boldsymbol{\omega}$, the set of quantities being estimated.

When the estimator is biased, it is still possible to calculate the CRBs if the bias, $b(\boldsymbol{\omega})$, is known and differentiable. When this is the case, the CRBs are given by

$$\sigma_m^2 \geq \left\{ \left[\mathbf{I} + \mathbf{J}_b\left(\boldsymbol{\omega}\right)\right] \mathbf{F}^{-1} \left[\mathbf{I} + \mathbf{J}_b\left(\boldsymbol{\omega}\right)\right]^T \right\}_{m,m} \tag{6.15}$$

Where $\mathbf{I}$ is the identity matrix, $\mathbf{J}_b(\boldsymbol{\omega})$ is the Jacobian of the bias, and the superscript T denotes a transpose operation. If the bias is not known, or not differentiable, then the CRBs cannot be computed. However, it may still be useful to evaluate Eq. (6.13) and use the computed variances to track the relative performance of an estimator as certain conditions, such as SNR, change.

We now derive the FIM for the case of the multi-field PDPR estimation using measured images with additive Gaussian noise. The noisy detected images are given by

$$d_k\left(u,v;\xi_f,\eta_f\right)=s_k\left(u,v;\xi_f,\eta_f\right)+n_{k,f}\left(u,v\right) \tag{6.16}$$

where $n_{k,f}(u, v)$ is zero-mean, additive Gaussian noise for the k^{th} diversity PSF of the f^{th} field point. The log-likelihood function is given by

$$\mathbf{L}\left(\boldsymbol{\omega}\right)=-\frac{1}{2\sigma_n^2}\sum_{u,v}\sum_{f=1}^{F}\sum_{k=1}^{K}\left[d_k\left(u,v;\xi_f,\eta_f\right)-s_k\left(u,v;\xi_f,\eta_f;\boldsymbol{\omega}\right)\right]^2 \tag{6.17}$$

where σ_n^2 is the variance of the noise and $\boldsymbol{\omega}$ represents either the $\boldsymbol{\alpha}$ or $\boldsymbol{\beta}$ coefficients. Evaluating the first derivative of Eq. (6.17) gives

$$\begin{aligned}\frac{\partial\mathbf{L}\left(\boldsymbol{\omega}\right)}{\partial\omega_m}=\frac{1}{\sigma_n^2}\sum_{u,v}\sum_{f=1}^{F}\sum_{k=1}^{K}&\left[d_k\left(u,v;\xi_f,\eta_f\right)-s_k\left(u,v;\xi_f,\eta_f;\boldsymbol{\omega}\right)\right]\\&\times\frac{\partial s_k\left(u,v;\xi_f,\eta_f;\boldsymbol{\omega}\right)}{\partial\omega_m}.\end{aligned} \tag{6.18}$$

Evaluating the second derivative gives

$$\frac{\partial^2 \mathbf{L}(\boldsymbol{\omega})}{\partial \omega_n \partial \omega_m} = \frac{1}{\sigma_n^2} \sum_{u,v} \sum_{f=1}^{F} \sum_{k=1}^{K} \left\{ \left[d_k\left(u,v;\xi_f,\eta_f\right) - s_k\left(u,v;\xi_f,\eta_f;\boldsymbol{\omega}\right) \right] \times \frac{\partial^2 s_k\left(u,v;\xi_f,\eta_f;\boldsymbol{\omega}\right)}{\partial \omega_n \partial \omega_m} - \frac{\partial s_k\left(u,v;\xi_f,\eta_f;\boldsymbol{\omega}\right)}{\partial \omega_m} \frac{\partial s_k\left(u,v;\xi_f,\eta_f;\boldsymbol{\omega}\right)}{\partial \omega_n} \right\}. \tag{6.19}$$

Finally, noting that

$$\left\langle d_k\left(u,v;\xi_f,\eta_f\right) \right\rangle = s_k\left(u,v;\xi_f,\eta_f;\boldsymbol{\omega}\right) \tag{6.20}$$

and taking the negative of the expectation value of Eq. (6.19) to obtain the FIM yields

$$\left[\mathbf{F}\right]_{m,n} = \frac{1}{\sigma_n^2} \sum_{u,v} \sum_{f=1}^{F} \sum_{k=1}^{K} \frac{\partial s_k\left(u,v;\xi_f,\eta_f;\boldsymbol{\omega}\right)}{\partial \omega_m} \frac{\partial s_k\left(u,v;\xi_f,\eta_f;\boldsymbol{\omega}\right)}{\partial \omega_n}. \tag{6.21}$$

Using Eqs. (6.8) through (6.10) the derivatives can be computed, giving

$$\left[\mathbf{F}\right]_{m,n}^{\boldsymbol{\alpha}} = -\frac{16\pi^2}{\lambda^2 \sigma_n^2} \mathrm{Re}\left\{ \sum_{u,v} \sum_{f=1}^{F} \sum_{k=1}^{K} \left[h_k^*\left(u,v;\xi_f,\eta_f;\boldsymbol{\alpha}\right) \right]^2 \times \mathbf{P}\left[Z_m\left(x,y\right) P\left(x,y;\xi_f,\eta_f;\boldsymbol{\alpha}\right) \right] \times \mathbf{P}\left[Z_n\left(x,y\right) P\left(x,y;\xi_f,\eta_f;\boldsymbol{\alpha}\right) \right] \right\} \tag{6.22}$$

and

$$\left[\mathbf{F}\right]_{m,n}^{\boldsymbol{\beta}} = -\frac{16\pi^2}{\lambda^2 \sigma_n^2} \mathrm{Re}\left\{ \sum_{u,v} \sum_{f=1}^{F} \sum_{k=1}^{K} \left[h_k^*\left(u,v;\xi_f,\eta_f;\boldsymbol{\beta}\right) \right]^2 \times \mathbf{P}\left[Z_m\left(x,y;\xi_f,\eta_f\right) P\left(x,y;\xi_f,\eta_f;\boldsymbol{\beta}\right) \right] \times \mathbf{P}\left[Z_n\left(x,y;\xi_f,\eta_f\right) P\left(x,y;\xi_f,\eta_f;\boldsymbol{\beta}\right) \right] \right\} \tag{6.23}$$

where **P**[...] is the Fresnel-like propagation of Eq. (6.9). We note from Eqs. (6.22) and (6.23) that the FIM, and therefore the CRBs, depends on the specific realization of the wavefront parameterized by $\boldsymbol{\alpha}$ or $\boldsymbol{\beta}$. Therefore, for this study we calculate the

CRBs for several wavefront realizations and average them to obtain an average minimum variance of the estimation error. While this removes some of the novelty of calculating the CRBs it will still provide a relative comparison between the two methods of multi-field wavefront estimation.

6.5 Digital Simulations

PSFs were simulated for five different field points of a monolithic circular aperture. Figure 6-1 shows a schematic of the locations of the field points within the system's FOV. For each field point, two diversity PSFs were simulated with +4 λ and –4 λ of focus diversity. Five noise levels were tested by varying the amount of read noise in the image from 5 e^- to 25 e^-. PDPR was used to estimate either the set of the β coefficients directly, or estimate them indirectly by first estimating the α_f coefficients for each field point and using the transformation of Eq. (6.7). The double Zernike expansion included terms up to and including 4^{th} order in the pupil coordinates and 1^{st} order in the field coordinates. Five wavefront realizations were generated, each with 10 independent noise realizations, for a total of 50 independent trials. The CRB calculations are averaged over the 5 wavefront realizations since they do not depend on individual noise realizations. The PDPR results are averaged over all 50 trials. Figure 6-3 shows an example of a set of simulated wavefronts and Figure 6-4 shows the corresponding PSFs on one side of focus.

For each set of PSFs, the FIM was calculated using Eq. (6.22) or Eq. (6.23). When Eq. (6.23) was used, the FIM was inverted to calculate the CRBs

corresponding to when the $\boldsymbol{\beta}$ coefficients are directly estimated. When the FIM was calculated for the $\boldsymbol{\alpha}$ coefficients, the CRBs of indirectly estimating the $\boldsymbol{\beta}$ coefficients were calculated using the transformation

$$\sigma_m^2 \geq \left[\mathbf{J}_{\boldsymbol{\beta}}(\boldsymbol{\alpha}) \mathbf{F}(\boldsymbol{\alpha})^{-1} \mathbf{J}_{\boldsymbol{\beta}}(\boldsymbol{\alpha})^T \right]_{m,m}, \tag{6.24}$$

where $\mathbf{J}_{\boldsymbol{\beta}}(\boldsymbol{\alpha})$ is the Jacobian of $\boldsymbol{\beta}$ with respect to $\boldsymbol{\alpha}$, which was calculated by differentiating Eq. (6.7). Since we are using $F = 5$ discrete field points and using only $L = 3$ terms in the polynomial expansion over the field coordinates, the matrix $\mathbf{M}$ in Eq. (6.6) is not square, and therefore not directly invertible and we use the Moore-Penrose pseudo-inverse [3].

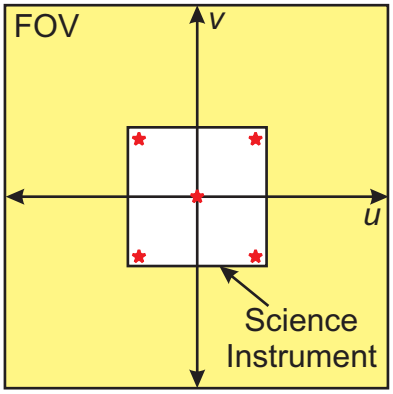


Figure 6-2: Schematic of the positions of the five field points on the focal plane. A single science instrument is used to perform multi-field wavefront sensing and the wavefront is then extrapolated over the entire FOV.

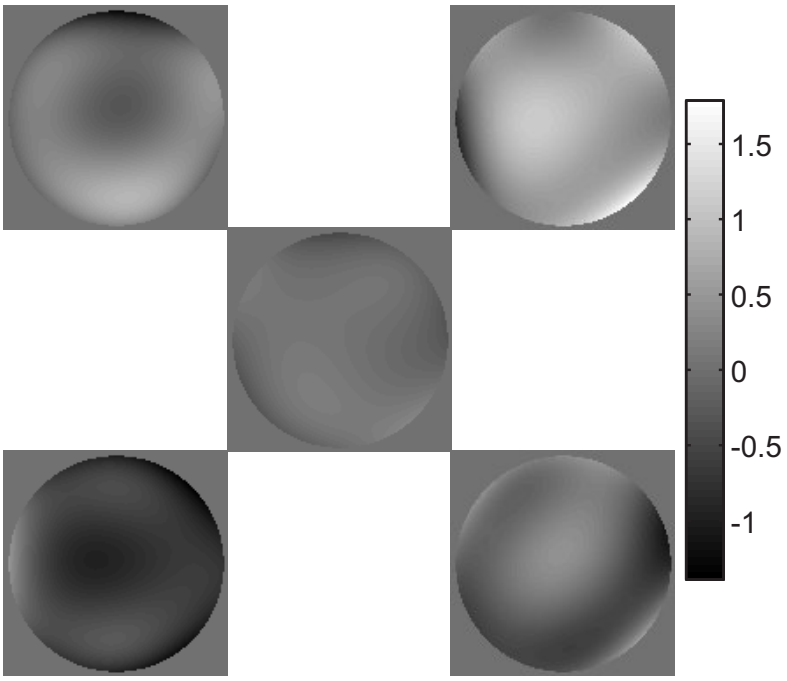


Figure 6-3: One realization of simulated wavefronts corresponding to the five field points shown in Figure 6-2. Scale has units of waves.

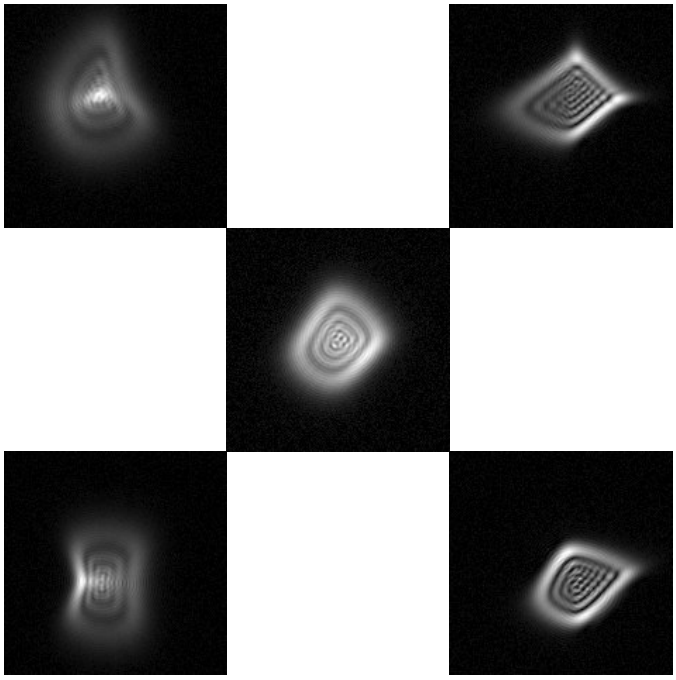


Figure 6-4: Simulated PSFs at –4 λ inside of focus corresponding to the wavefronts shown in Figure 6-3. Intensity raised to the ½ power is shown..

6.6 Results

Figure 6-5 shows the square root of the average CRBs for the high SNR case when the standard deviation of the read noise was 5 e^-. That is, each point is given by

$$y = \sqrt{\overline{\sigma_m^2}} \tag{6.25}$$

and the limits of the error bars are given by

$$\begin{aligned} \text{Upper Limit} &= \sqrt{\overline{\sigma_m^2} + std\left(\sigma_m^2\right)} \\ \text{Lower Limit} &= \sqrt{\overline{\sigma_m^2} - std\left(\sigma_m^2\right)} \end{aligned} \tag{6.26}$$

where $\sigma_m{}^2$ is the CRB for the m^{th} coefficient and *std*(...) denotes the standard deviation of the argument. The pairs of points along the bottom of the chart correspond to the coefficients of polynomial terms that are linear in the field coordinates, while the other points correspond to coefficients of polynomial terms where the field dependency is constant. For example, index 9 corresponds to the polynomial

$$U_9\left(\xi,\eta;x,y\right) = \left(r\cos\theta\right)\left(\rho^2\sin 2\phi\right) \tag{6.27}$$

where

$$\begin{aligned} r &= \sqrt{\xi^2+\eta^2}, \quad \theta = \tan^{-1}\left(\frac{\eta}{\xi}\right), \\ \rho &= \sqrt{x^2+y^2}, \quad \phi = \tan^{-1}\left(\frac{y}{x}\right), \end{aligned} \tag{6.28}$$

and is astigmatism that varies linearly with field position. Index 7 corresponds to the polynomial

$$U_7(x,y) = \rho^2 \sin 2\phi \tag{6.29}$$

which is astigmatism that is independent of field position. The CRBs indicate that there is a larger error associated with estimating the coefficients of the field-independent polynomials. Furthermore, the CRBs show that theoretically there is a modest benefit to estimating the β coefficients directly, using the field-dependent polynomial expansion of Eq. (6.3), rather than first estimating the α coefficients and performing the calculation of Eq. (6.7).

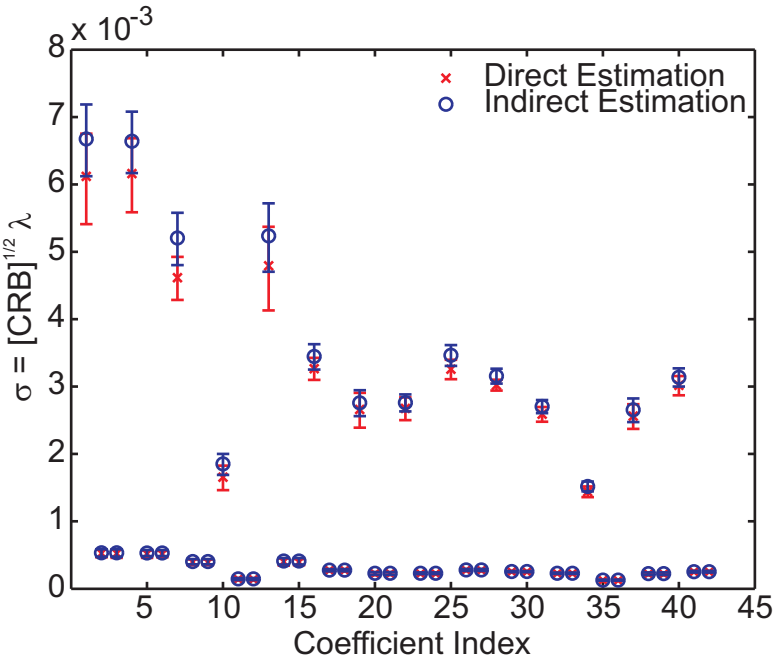


Figure 6-5: Square root of the Cramer-Rao bounds for each coefficient when the read noise standard deviation was 5 e^-. The pairs of points along the bottom correspond to polynomial terms that are linear in the field coordinates, while the other points correspond to polynomial terms for which the field dependence is constant.

Figure 6-6 shows the same results for the low SNR case when the standard deviation of the read noise was 25 e^-. The same trends are seen, except the magnitude of the CRBs has increased, indicating a larger estimation error with lower SNR. From the inverse relationship between the CRBs and the FIM, as calculated in Eq. (6.22) or Eq. (6.23), we expect the CRBs to vary linearly with the variance of read noise. This expectation is confirmed in Figure 6-7, where we plot the average CRB for coefficient #7 against the standard deviation of the read noise. It is also clear from Figure 6-7 that in the limit of high SNR, the two methods of estimating the field-dependence are nearly equal and for low SNR, the method of directly estimating the β coefficients has a clear advantage.

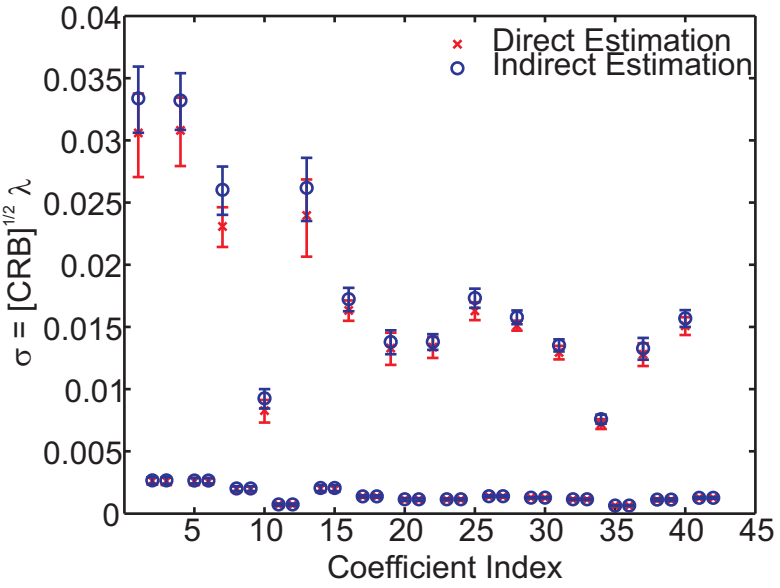


Figure 6-6: Square root of the Cramer-Rao bounds for each coefficient when the read noise standard deviation was 25 e^-. There is an advantage to using the direct estimation method, specifically in the estimation error of the field-independent polynomials.

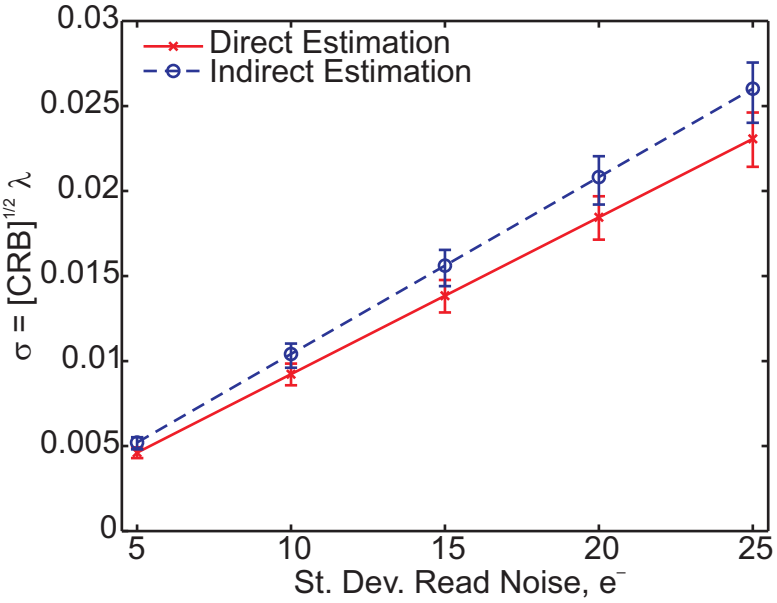


Figure 6-7: Square root of the CRB for coefficient #7 vs. the standard deviation of the read noise. As expected, the CRB is linear with the standard deviation. Coefficient #7 corresponds to field-independent astigmatism. The advantage to using the direct estimation method increases with the amount of noise.

Figure 6-8 shows the PDPR root-mean-squared-error (RMSE) for each coefficient, in good qualitative agreement with the CRBs. The PDPR results show that there is a larger error in estimating the coefficients of the field-independent polynomials than estimating the coefficients of the field-dependent polynomials. The PDPR results also show a benefit to the direct estimation method vs. the indirect estimation method. Figure 6-9 shows the RMSE for coefficient #7 vs. the standard deviation of the read noise, again in good qualitative agreement with the CRBs. In both Figures the direct estimate CRBs are also plotted for comparison. PDPR did not come close the CRBs, although it did much better for the linearly field-dependent polynomial terms than the terms that were constant with respect to the field

coordinate. This indicates that the PDPR algorithm used here is not the optimal estimator and that further improvement could be made.

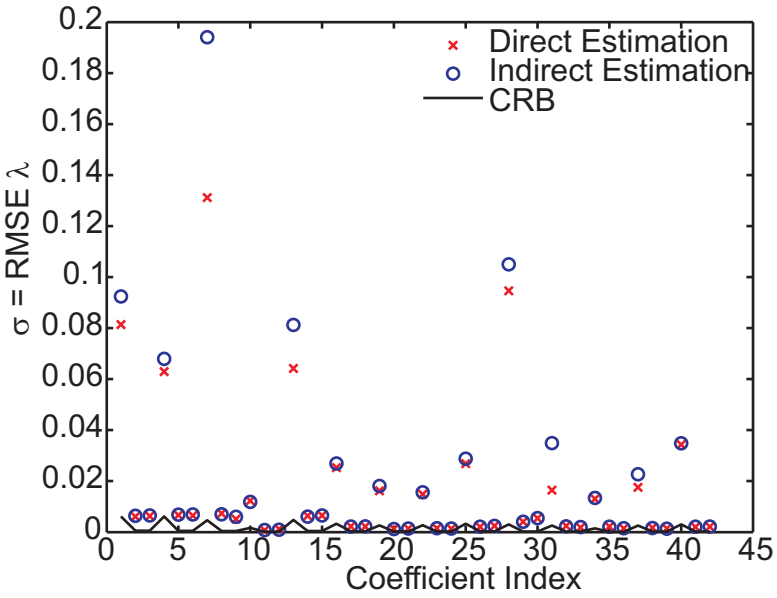


Figure 6-8: RMS error in PDPR estimation of the coefficients. The square root of the CRB for the direct estimation method is also plotted for comparison.

6.7 Summary

We presented two methods of performing multi-field wavefront sensing with phase-diverse phase retrieval. In the first method, sets of polynomial coefficients are estimated for discrete field points and a field-dependent model of the wavefront is constructed after estimation. In the second method, the field-dependent model of the wavefront is included in the PDPR algorithm and coefficients that characterize the wavefront for any field point are directly estimated.

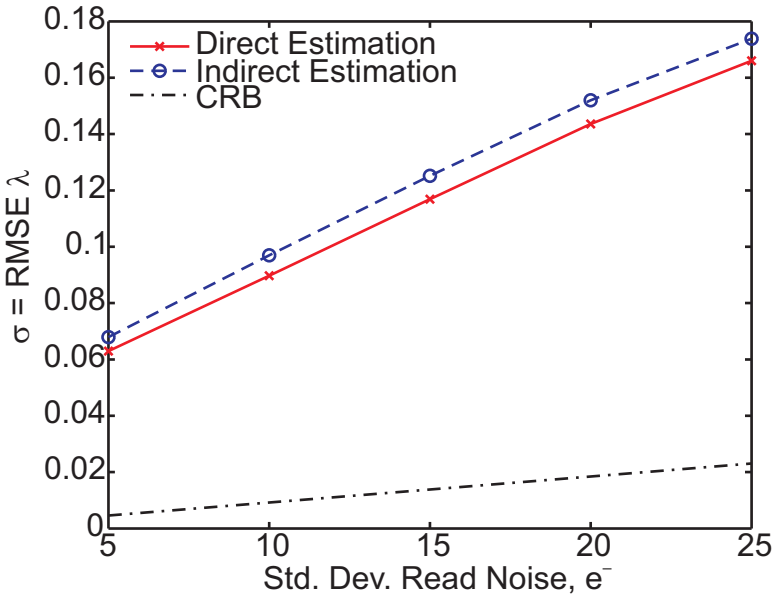


Figure 6-9: RMS error in PDPR estimation of coefficient #7. The square root of the CRB for the direct estimation method is also plotted for comparison. As expected, the error increases linearly with the standard deviation of the read noise.

Information-theoretic Cramer-Rao bounds were calculated to compare these two methods, and show that for low SNR, there is an advantage to the direct estimation method. For situations of high SNR, the two methods of estimation yield similar results.

PDPR was performed on digitally simulated data. The average error in estimating the coefficients was in good qualitative agreement with the CRBs. Again, an advantage was seen for the direct estimation method in situations of lower SNR.

6.8 References

1. D. J. Schroeder, *Astronomical Optics, 2 Ed.* (Academic Press, San Diego, CA, 2000).

2. I. Kwee, and J. Braat, "Double Zernike expansion of the optical aberration function," Pure and Applied Optics: Journal of the European Optical Society Part A **2**, 21-32 (1993).

3. R. Penrose, "A generalized inverse for matrices," in *Proc. of Cambridge Philos. Soc.*, **51** (1955), p. C655.

4. A. E. Siegman, *Lasers* (University Science Books, Sausalito, CA, 1986).

5. J. Cederquist, S. Robinson, D. Kryskowski, J. R. Fienup, and C. C. Wackerman, "Cramer-Rao lower bound on wavefront sensor error," Optical Engineering **25**, 586-592 (1986).

6. J. Cederquist, J. Fienup, C. Wackerman, S. Robinson, and D. Kryskowski, "Wave-front phase estimation from Fourier intensity measurements," J. Opt. Soc. Am. A **6**, 1020-1026 (1989).

7. S. M. Kay, *Fundamentals of Statistical Signal Processing, Vol. 1: Estimation Theory* (Prentice-Hall, Upper Saddle River, NJ, 1993).

8. H. L. Van Trees, *Detection, Estimation, and Modulation Theory, Part 1* (John Wiley & Sons, Inc., New York, NY, 2001).

9. B. R. Frieden, *Science from Fisher Information: A Unification* (Cambridge University Press, Cambridge, UK, 2004).

7 Summary & Conclusion

7.1 Summary

The goal of this work was to investigate the use of phase diversity for segmented and multiple aperture telescopes, which is important because such systems are becoming more common as telescopes become larger. We first made improvements to the phase-diversity algorithm implementing a broadband version of the algorithm and exploring different techniques of regularizing the metrics used. We then presented a new technique of implementing phase diversity on segmented and multi-aperture systems that makes use of individual segment or sub-aperture actuation to introduce the diversity function. Using this new technique we have shown it is possible to estimate not only the system phase and object being observed, but also the gray-world (or average) object spectrum. These developed techniques were then successfully demonstrated on an experimental test bed using a segmented MEMs deformable mirror.

In addition to work on phase diversity, we present supporting work using phase-diverse phase retrieval (PDPR). A method of implementing PDPR using a field-dependent model of the wavefront was introduced and evaluated using information-theoretic Cramer-Rao bounds and digital simulations.

In Chapter 2, we described a method using a gray-world object model to implement a broadband phase-diversity algorithm. This technique was first introduced by Seldin *et al* in 2000 and demonstrated in simulation [1]. Here, we provide an in depth comparison between using the broadband algorithm and using the

monochromatic algorithm when a broadband object is present. We further explore the performance of the algorithm when the gray-world spectral coefficients are assumed to be uniform. It was shown that for bandwidths ($\Delta\lambda/\lambda_o$) of 100%, a broadband phase-diversity algorithm, using known gray-world spectral coefficients, could achieve a residual Strehl ratio of 0.80 which corresponds to an RMS wavefront error of 0.075 λ_o. The conventional monochromatic algorithm obtained a residual Strehl ratio of less than 0.50 for the same bandwidth. For a bandwidth of 50%, the broadband algorithm achieved a residual Strehl ratio of 0.95, corresponding to an RMS wavefront error of 0.036 λ_o, while the monochromatic algorithm achieved a Strehl of 0.85. If the gray-world spectral coefficients were assumed to be uniform, the broadband algorithm achieved a residual Strehl ratio of 0.80 for bandwidths of 90% and a Strehl ratio of 0.95 for bandwidths of 40%.

Chapter 2 also presented results of a study comparing three techniques of implementing regularization in a phase diversity algorithm. These techniques were compared to the conventional reduced Gaussian (RG) metric developed by Gonsalves and Chidlaw in 1979 [2]. The first technique used a multi-frame, Wiener-Helstrom filtered estimate of the object to obtain the regularized reduced Gaussian (RRG) metric. The second and third techniques were developed for the case of two diversity images by Blanc *et al* [3], and involved using a Bayesian maximum *a posteriori* derivation to obtain the joint maximum *a posteriori* (JMAP) and marginal *a posteriori* (mAP) metrics. In this work, we extend their derivations for an arbitrary number of diversity images. Through digital simulation, it was shown that the RRG

and JMAP metrics yield lower phase estimation error for situations of low signal-to-noise ratio (SNR), achieving a maximum residual Strehl ratio of 0.85 where the RG metric only achieved a maximum residual Strehl ratio of 0.50. For high SNR, the RG and mAP metrics performed the best, achieving a maximum residual Strehl ratio of 0.9985 where the RRG metric achieved maximum residual Strehl ratio of 0.9980.

In the third Chapter, we presented a new implementation of phase diversity specific to segmented and multi-aperture systems. While conventional implementations use a focus shift to introduce a known phase aberration, sub-aperture piston phase diversity (SAPPD) uses segment or sub-aperture actuation to introduce piston phases over individual segments or sub-apertures. This technique could be useful for systems on which the focus diversity mechanism becomes disabled. Later, it is also shown to be useful for estimating object spectra using phase diversity. Through digital simulation, SAPPD was compared to conventional focus diversity and shown to be better than focus diversity for situations of high SNR when the pistons introduced were equivalent to fractional values of the center wavelength. At integer and half-integer values, the performance of SAPPD decreased due to a lack of diversity in the images. For low SNR, the conventional focus diversity yielded better phase estimates.

In the fourth Chapter we presented work showing that SAPPD could be used to obtain estimates of the object gray-world spectral coefficients, in addition to the system phase and object. It was shown that the amount of diversity introduced determined the spectral resolution of the algorithm, indicating that SAPPD would be

better suited to spectral estimation than focus diversity. Furthermore, a constraint was placed on the optical path difference (OPD) step size between diversity images to prevent aliasing of the spectrum. Through digital simulation, it was shown that good qualitative agreement between the estimated spectrum and object average-pixel spectrum could be obtained regardless of the OPD step size. The best quantitative agreement was obtained when both the spectral resolution and OPD step conditions were met. Pistoning of sub-apertures by 40 μm of OPD allowed seven spectral coefficients (27.4 nm) to be estimated with a normalized root-mean-squared error of 0.15.

An experimental test bed was constructed using a MEMS deformable mirror as the system under test and a modified digital light processing (DLP™) projector as the scene projector. Phase-shifting interferometry was used to corroborate the phase-diversity results. Using this test bed, conventional focus diversity was demonstrated using both monochromatic and broadband implementations. For both cases, the phase was estimated to within 0.06 λ_o RMS of the phase measured by the interferometer. Monochromatic and broadband implementations of SAPPD were also demonstrated, resulting in 0.08 λ_o RMS and 0.06 λ_o RMS errors, respectively. In all cases, an object was reconstructed that was in good agreement with the projected scene.

Work involving phase-diverse phase retrieval (PDPR) was also presented. First, a method of incorporating the field-dependence of a wavefront into the forward model of a PDPR algorithm was introduced. This new method is in contrast to

traditional methods where a wavefront is estimated separately for each field point and a field-dependent model is extrapolated afterwards. These two methods were compared using information theoretic Cramer-Rao bounds and digital simulation. Both comparisons showed an advantage to using the field-dependent model in the PDPR algorithm, especially as the SNR decreases.

7.2 Future Work

While the work presented here has begun to explore the potential of phase diversity for segmented and multi-aperture systems, there is much more to be done. For example, it is well known that low contrast objects or regions of a scene yield poor estimates of the phase. Formally investigating the performance of phase diversity as a function of object contrast and spatial frequency content requires further study. Furthermore, object spectral content will probably have an effect on broadband phase-diversity algorithms, especially when a gray-world spectrum is being estimated. We have collaborated with the Digital Imaging and Remote Sensing (DIRS) Laboratory at the Rochester Institute of Technology to investigate this [4], however further study is required. It is also necessary to test phase diversity with astronomical objects as the objects, where spatial and spectral statistics may be much different than terrestrial scenery. Finally, the relationship between phase estimation and spatial frequency content of the phase being estimated is not well understood, though it is expected that more complex phases may be harder to estimate.

Perhaps one of the biggest tasks is that of the full, joint estimation of both the object and phase using the Poisson or Gaussian log-likelihood functions of Eq. (1.12) or Eq. (1.15). While some work has been done to investigate this [5], a full study is needed, exploring the trade-offs between the phase and object when both are explicitly being estimated, and perhaps including either a gray-world model or full spectral model of the object. A comparison is also needed between an object explicitly estimated by the algorithm, and one reconstructed using the phase estimated by a reduced metric.

The combination of phase diversity with the multi-field wavefront sensing technique presented in this thesis would be a good project for future investigation. Instead of using individual field points to estimate a field-dependent model of the wavefront, one could use an extended object or scene that exceeds the isoplanatic patch to estimate wavefront information for the entire field of view.

One problem that has not been discussed in detail in this work is an inherent ambiguity of segmented and multi-aperture systems when sub-apertures or segments are pistoned with respect to each other by exactly an integer number of wavelengths. With a monochromatic phase-diversity algorithm it is impossible to detect such a misalignment. Even broadband phase-diversity algorithms have a stagnation mode in which it would require an increase of the error metric in order to bring a segment or sub-aperture within a single wavelength of piston. Therefore, if the phase-diversity result was used to actuate the segments or sub-apertures, it could be possible that they would be driven further away from alignment. Some techniques have been developed

to mitigate this problem [1, 6]. An automated method, implemented as part of the phase-diversity algorithm, would be extremely useful. One possible technique would be to sequentially add an integer wavelength of piston to the estimated phase of each segment or sub-aperture and calculate the change in the error metric. Noise would hinder this technique and would need to be addressed.

Another opportunity for improvement of phase-diversity algorithms would be to investigate performance under adverse conditions. Some work has been done here to test phase diversity when SNR is low, but effects such as jitter, background, charge diffusion, finite pixel size, detector gain and bias, etc. have not been included in the simulated images. Phase diversity should be tested when all of these effects are present to varying degrees. Furthermore, incorporating these effects into the forward model or even allowing phase diversity to estimate some of them may be very beneficial.

Finally, an experimental demonstration of the spectral estimation technique presented in Chapter 4 is required. A well-calibrated mirror with large amounts of stroke is needed to achieve fine resolution of the bandwidths provided on our test bed. It may also be necessary to develop a different method of scene projection in order to produce objects with sufficiently complex spectra; the DLP projector's five segment color wheel may produce spectra that are too smooth to distinguish from the uniform starting guess of the algorithm. Also, an independent measurement of the spectrum is needed to compare with the estimated gray-world coefficients.

7.3 References

1. J. H. Seldin, R. G. Paxman, V. G. Zarifis, L. Benson, and R. E. Stone, "Closed-loop wavefront sensing for a sparse-aperture, phased-array telescope using broadband phase diversity," in *Imaging Technology and Telescopes, Proc. SPIE*, J. W. Bilbro, J. B. Breckinridge, R. A. Carreras, S. R. Czyzak, M. J. Eckart, R. D. Fiete, and P. S. Idell, eds., **4091** (2000), pp. 48–63.

2. R. Gonsalves, and R. Chidlaw, "Wavefront sensing by phase retrieval," in *Applications of Digital Image Processing III, Proc. SPIE*, **207** (1979).

3. A. Blanc, L. M. Mugnier, and J. Idier, "Marginal estimation of aberrations and image restoration by use of phase diversity," J. Opt. Soc. Am. A **20**, 1035-1045 (2003).

4. B. J. Daniel, M. R. Bolcar, J. R. Schott, and J. R. Fienup, "Phase retrieval in sparse aperture systems with phase diversity: a trade space study," in *Sensors and Systems for Space Applications II, Proc. SPIE*, R. T. Howard, and P. Motaghedi, eds., **6958** (2008), pp. 69580K-69512.

5. R. Paxman, B. Thelen, and J. Seldin, "Phase-diversity correction of turbulence-induced space-variant blur," Optics Letters **19**, 1231-1231 (1994).

6. R. G. Paxman, J. H. Seldin, and P. P. Sanchez, "Applied Phase Diveristy," (Environmental Institute of Michigan [ERIM], Ann Arbor, Michigan, 1992).