Smart optics: Wavefront sensor-less adaptive optics - Image correction through sharpness maximisation

by

Larry P. Murray, B.Sc, M.Sc.

Supervised by Prof. J. C. Dainty

A thesis submitted to the National University of Ireland, Galway, as partial fulfilment of the requirements for the Degree of Doctor of Philosophy.

Department of Physics

National University of Ireland, Galway.

October 2006

CONTENTS

LIST OF FIGURES

ACKNOWLEDGMENTS

Firstly, I would like to thank my supervisor, Prof. Chris Dainty, for all his help and guidance throughout the duration of my Ph.D. I would also like to thank the members of the Applied Optics Group, especially Liz, Sasha and Dave for assistance, encouragement and good cheer on many occasions. I want to thank comrades David, Denis, Alfredo and Andrew for their company and good humour over the past few years. Thanks also to Eugenie for her assistance and summer research students, Milan, Jean, Franz and Elaine for their input and collaboration.

I would like to thank HP Labs and Science Foundation Ireland for the all important funding of my Ph.D. research, and Larry Hubby and John Rudin of HP Labs for collaborating with me on my project.

I owe special thanks to my family, especially my parents who have given me evermore support and help throughout my Ph.D. Lastly, I'd like to thank Linda for all her love and patience.

ABSTRACT

The original proposal of wavefront-sensorless aberration correction was suggested by Muller and Buffington. In this technique, we attempt to correct for wavefront aberrations without the use of a conventional wavefront sensor. Commands are applied to a corrective element with adjustable segments in an attempt to maximise a "sharpness" metric which correlates to image quality. Search algorithms are employed to find the optimal combination of actuator voltages on a deformable mirror to maximise a certain sharpness metric. The "sharpness" is based on image intensity measurements taken with a Charged Coupled Device (CCD) camera. It is shown that sharpness maximisation, using the simplex algorithm, can minimise the aberrations and restore the Airy rings of an imaged point source. This technique is then applied to extended objects which have been aberrated using a Hamamatsu spatial light modulator to induce aberrations. The correction achieved using various search algorithms are evaluated and presented.

The work carried out in this research experientially validates the use of sharpness metrics to correct for point source and extended object images. The suitability of various metrics is shown, and by application of mirror modes and Zernike aberrations into the system, a greater understanding of the search space is gained. This enables insight into the choice of search algorithm and the degree of correction achievable in a future system. A stochastic parallel gradient descent algorithm is shown to offer a high degree of

correction in a low number of iterations, demonstrating that sharpness maximisation has potential to correct quickly for aberrated images when employed in a system with high speed components.

It is shown that sharpness maximisation has potential for correction of slowly varying aberrations and enables the important calibration for non-common path errors, a limitation in many existing adaptive optics systems. This is further encouragement for the development of transmissive corrector which has the potential to spread the usage and applications of adaptive optics much further, notably in industrial and medical applications.

1. INTRODUCTION

1.1 Adaptive Optics Background

"Smart optics" means much more than adaptive optics on telescopes and in vision science: new technologies are changing the way adaptive optics (AO) instruments are built and operated, bringing new technologies into everyday life in the form of cheaper, lighter, and more robust optical systems. Smart optics includes optical systems, subsystems or devices that adjust to enhance the performance of imaging and image processing systems. Smart optics encompasses technologies such as, adaptive optics, programmable diffractive optics, optically based control systems, and wavefront controllers.

This vast area now affected by, and dependent on, adaptive optics stems originally from an adaptive optics system proposed to correct for atmospheric distortion in stellar images. The genesis for adaptive optics was in 1953 when H. Babcock[1] first proposed the use of a spatial light modulator to correct for aberrations introduced into images produced by ground-based telescopes due to atmospheric turbulence. The implementation of adaptive optics in astronomy led recently to its use in vision science and these are the two main areas where adaptive optics is employed.

Adaptive optics correction is, in theory, simple, but due to some demanding technical requirements for the wavefront sensor (WFS) and deformable mirror, realisation of adap-

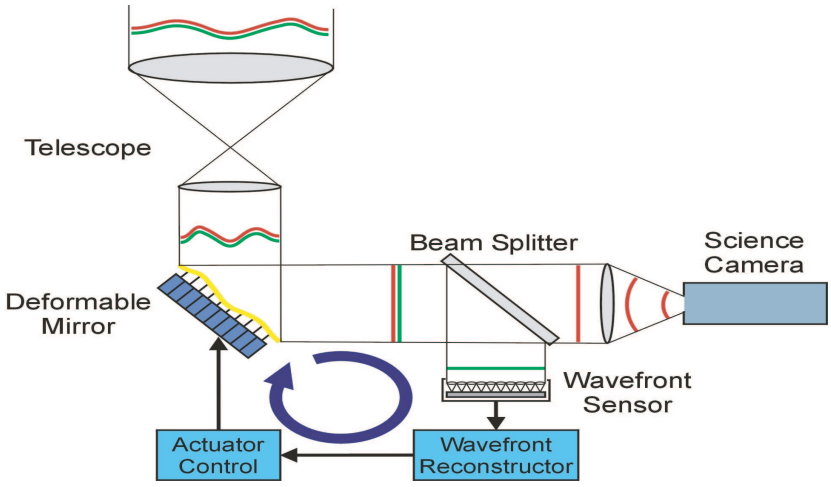


Fig. 1.1: A schematic of a conventional adaptive optics system showing key components such as a wavefront sensor and deformable mirror.

tive optics systems was delayed by several decades. In the USA, adaptive optics systems have been developed by the Department of Defense since 1970 in classified research[2]. In the civilian sector it took until the late 1980s before the COME-ON system of the European Southern Observatories was installed on the 3.6-m telescope in Chile[3]. A generic adaptive optics system for astronomy is presented in Figure 1.1, showing the key components.

As can be seen from Figure 1.1, the idea behind adaptive optics is, as previously mentioned, relatively simple. In most cases the aberrations of an incoming wavefront are determined by a wavefront senor and the necessary commands to be applied to the compensation device are determined by a control computer. However, the devil is in the detail. Performing the wavefront measurements, calculation of the control commands and, physical implementation at a rate high enough for real time correction, is the root of the complexity of these adaptive optics systems.

The technical requirements for adaptive optics systems include the frame rate and sensitivity of the wavefront sensor camera, and the maximum frequency that can be applied to the corrective element. Depending on the coherence time of the aberrating medium in a given adaptive optics system, corrections need to be made at a high rate, often KHz, or as a rule of thumb, measurements need to be made at a rate approximately 5-10 times faster than the coherence time of the aberration. In atmospheric adaptive optics, the required frame rate is determined by the rate of change of the atmospheric turbulence, whereas in vision science, the time constraints are largely imposed by the physical movements of the subject, both voluntary and involuntary, i.e., movements of the subject's eye due to pulse, thus each application has its individual requirements, which need to be met using an adaptive approach. A further difficulty of correction in adaptive optics, is that the modal structure of the wavefront sensor and wavefront corrector have to match that of the aberration being corrected to compensate properly for the wavefront aberration.

1.2 Image Sharpening

Prior to the implementation of modern wavefront sensors, and in the first practical demonstration of adaptive optics, Muller and Buffington[4], in 1974, proposed image sharpening as a method to correct for aberrated wavefronts. The basis, as outlined in Chapter 5, was to maximise an image metric which related to wavefront aberrations and, thus, image quality. However, with the advent of Shack-Hartmann and curvature wavefront sensors, which could be used to correct with high temporal frequency, the image sharpening approach was neglected because of the limited speed of correction that it could facilitate and problems related to correction algorithms. But, with the advancement of the component technol-

ogy required in image sharpness maximisation, this method has seen a re-emergence of interest. Key components, such as CCD cameras, are now able to operate in a much faster regime and computers have improved in speed.

The significant benefit that is offered by conventional wavefront sensing techniques, is the direct determination of necessary commands to be applied to the correcting device, providing close conjugation with the aberration. This, although a mathematically complex calculation, provides directly and quickly the commands to correct for the aberrations. The direct determination of the correcting device commands requires a reconstruction of the wavefront from the wavefront sensor measurements. The general problem is the determination of the wavefront phase from a map of its gradient or Laplacian. The reconstruction problem can be expressed in a matrix-algebra framework. The unknowns, a vector ϕ, of N commands, or of N phase values over a grid, must be calculated from the data, a measurement vector S, of M elements of slopes in two directions, or Laplacians and edge slopes. This is discussed in more detail in Chapter 2.

As sharpness maximisation does not directly measure the wavefront, but maximises the image sharpness metric through a series of measurements and trial corrections applied to the correcting device, it is a slower process. As such, sharpness maximisation may be better suited to slow-varying or static aberrations and might be useful in cases where a wavefront sensor cannot be relied upon. Central to the process of sharpness maximisation, and its speed, is the need for an intelligent and efficient search algorithm which, based on the previous sharpness calculation, determines the new set of voltages to be applied to the corrector. The aim is to drive the wavefront correction device to its optimal shape to minimise the wavefront aberrations.

1.3 Project Aims

The aim of this project is to develop a "smart" optics wavefront sensor-less device that can correct for optical aberrations. The major difference of this system, compared to most adaptive optics systems in use, is that the necessary commands to be applied to the wavefront correction device will be determined without direct knowledge of the wavefront aberrations, which are usually determined from wavefront sensor data. In this approach, image sharpness will be used as an image quality metric to be maximised, with image sharpness defined to be a maximum only in the presence of zero wavefront error[4]. The current status and components of conventional adaptive optics systems are presented in Chapter 2.

Initially, phase corrections will be implemented using a deformable mirror with the view to finally incorporating a transmissive correcting element. These corrections are determined without the use of a wavefront sensor, which is an integral part of most adaptive optics systems from astronomy to vision science. The inclusion of a wavefront sensor affords adaptive optics systems high speed correction and close conjugation of the deformable mirror to the phase aberrations. An expense of these benefits is the mathematical complexity involved in determination of the necessary commands to apply to the deformable mirror. These commands are mostly determined through a singular value deconvolution matrix operation[5, 6] which is presented in Chapter 2.4.

A distinction can be made between direct and indirect adaptive optics - presented in Chapter 3 - whereby, direct adaptive optics uses a wavefront sensor, be it a Shack-Hartmann or other sensor[7], to directly measure the phase distortion of the incoming wavefront. Indirect adaptive optics does not directly measure the wavefront but uses quan-

tities related to the wavefront to determine the correction. The range of possible correction devices are also discussed in detail in Chapter 3.

The quantity to be used in this project is the "sharpness" of the image and definitions of this will be set out later in Chapter 4. This sharpness quantity is determined directly from the intensities measured at the CCD camera so calibration of the CCD camera is important. The sharpness will be measured for an image recorded by the CCD camera, and the corrections determined by a search algorithm will be input by the deformable mirror. These three elements will be the limiting factors of the system in terms of speed. Chapter 4 also details the mathematical tools and computer programs used for calculation of the sharpness metrics. The various search algorithms used to determine the correction to be applied to the deformable mirror are considered in Chapter 5. A spatial light modulator (SLM) is used to introduce aberrations and its aberrating ability and calibration are discussed in Chapter 6. CCD calibration issues, experimental apparatus and the optical set-up employed are also discussed in Chapter 6. Correction results for various algorithms are displayed in Chapter 7 and conclusions from this research is presented in Chapter 8.

The correction determined via indirect adaptive optics may not be suitable for real-time correction for aberrations with a short coherence time, but may offer a system of reduced complexity. We aim to show that an image can be corrected by maximising the "sharpness" of the image blindly, that is, without knowing its aberrations.

1.4 Presentations and Publications

Papers and presentations on this research have been written and presented at various meetings and conferences which I shall detail below. Papers have been published in three Inter-

national Society for Optical Engineering (SPIE) journals following several meetings and an invited talk was given and an Optical Society of America Topical meeting.

In April 2005 a paper entitled "Image Correction through Sharpness Maximisation", was presented at a SPIE conference, "Opto-Ireland 2005: Imaging and Vision" in Dublin, Ireland. An invited paper, "Real-time Aberration Compensation without a Wavefront Sensor", was presented to the Optical Society of America at a meeting in North Carolina, USA, in June 2005. A presentation was given at an European Optical Society (EOS) conference in Munich, Germany, on "Industrial Imaging and Machine Vision" as part of the "Imaging: Methods, Sensors and Processing" section of the conference, also in June 2005. A second SPIE paper was presented at the "5th International Workshop on Adaptive Optics in Industry and Medicine", in Beijing, China, in September 2005. And a third presentation on "Image sharpening and MEMs Mirrors", was given at a SPIE conference on "MEMS/MOEMS and Their Applications III", at a Photonics West conference in San Jose, USA, in January 2006. Two presentations were given at an Optical Society of America Annual Meeting on "Frontiers in Optics" in Rochester, NY, USA in October 2006. D. R. McGaughey gave a presentation on work derived from a collaboration entitled, "Sharpness Metric Nonstationarity in Undersampled Systems"[8] and I co-authored a presentation on "Wavefront Sensor-less Adaptive Optics - Image Correction Through Sharpness Maximisation".

A collaboration has been on going with HP Labs Bristol in an attempt to further the development of a transmissive continuous liquid crystal correction device. As part of that collaboration I tested the initial devices measuring the phase change range of these devices, providing feed back to HP Labs.

My contribution in this field has been to develop a robust system which corrects for an aberrated image of an extended object, relying only on the image at the science camera. This research furthers the idea that wavefront sensor-less correction may be feasible for some applications. The presentation of this idea at conferences has helped publicise the idea and outline key principles of sharpness maximisation. Through the physical implementation of such a system I have encountered physical limitations of wavefront sensor-less correction which might not have be obvious had I used only numerical simulation. The examination of the search space has provided insight into the choice of algorithm.

2. ADAPTIVE OPTICS

2.1 Introduction

The image quality of ground based large telescopes suffers from atmospheric turbulence. Independent of the telescope size, the long-exposure angular resolution in the visible is equivalent to that of a telescope with a diameter of 10-20cm[7, 9]. This effect is caused by the turbulent mixing of air with different temperatures in the atmosphere. Thus, a perfectly plane wave from a star at infinity is aberrated before it enters the telescope. The same effect can be observed in vision science due to the imperfect optics of the human eye and in other imaging systems, where lens aberrations contribute to a deteriorated image.

It was the idea of H. Babcock in 1953 to correct atmospheric aberrations with a deformable mirror to obtain diffraction limited images[1]. The principle of an adaptive optics system is displayed in Figure 1.1. A deformable mirror, a wavefront sensor and a camera in the corrected focus, form the main elements. The wavefront sensor measures the aberrations with a high frame rate and sends the control signals to the deformable mirror in order to correct the aberrations. Then, the corrected focus can be recorded by a camera with an exposure time largely independent of atmospheric turbulence.

A wavefront can be seen as a surface over which an optical disturbance has a constant phase. An aberrated wavefront will not be plane, and as such, will have a distorted shape

whereby parts of the wavefront lead or lag others, as can be seen in the incoming wavefront of Figure 1.1.

The majority of conventional adaptive optics systems use direct wavefront sensing to measure the wavefront distortion. Examples of direct wavefront sensors are Shack-Hartmann, shearing interferometry and curvature sensing. These wavefront sensors are discussed in greater detail in Chapter 3.

The wavefront shape can be corrected by modifying the optical phase profile. The key principle here is that the correcting phase profile should be optically conjugated to the measured phase aberration. Correction with the conjugated phase profile results in an ideal flat wavefront. However, if the wavefront cannot be accurately determined or replicated fully by the relay optics, or if diffraction effects dominate, the phase conjugation cannot be accurately employed and the correction is subsequently limited[7].

The conjugation of the aberration is essential to correct the image - this is achieved using a wavefront corrector to alter the wavefront phase. Deformable mirrors are the most often used corrector, but liquid crystal correctors are becoming increasingly common. Operation of these phase correction devices is based on control of the optical path difference (OPD) shown in Figure 2.1, which can be written as, $OPD = n \triangle z$, where n is the refractive index, and $\triangle z$ is the physical distance traveled by the wave. OPD is related to the phase ϕ, as $\phi = 2\pi \cdot OPD/\lambda$. Deformable mirrors, which are the primary technology for wavefront correctors, modulate $\triangle z$, operating in the reflective mode. Liquid crystal (LC) phase modulators represent a low-cost alternative to mechanically driven mirrors; their operation is based on modulation of n of the LC layer, under the applied electric field in transparent or reflective mode[10].

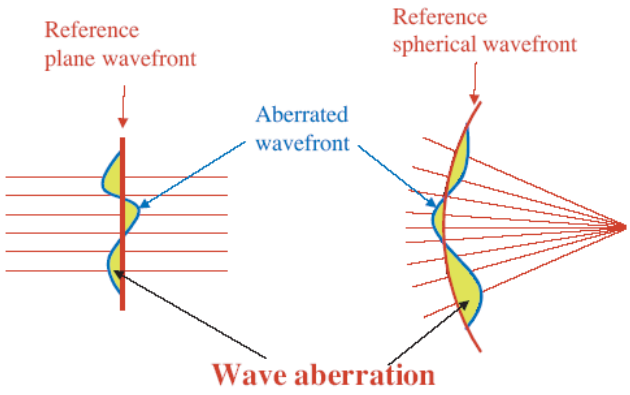


Fig. 2.1: Optical deviation of an aberrated wavefront from a reference wavefront.

Wavefront correctors are traditionally subdivided into two classes according to the implemented compensation technique - zonal and modal. Zonal correctors such as segmented piston and tip-tilt mirrors allow individual control of a phase over a set of subapertures providing step-wise phase compensation, whereas modal correctors use a set of smooth functions (modes, or influence functions) to approximate the required phase function[10].

2.2 Optical Aberrations

The basic principle of conventional adaptive optics is to measure the aberrations of an incoming wavefront and apply compensating aberrations in real time. To do this it is important to understand the nature and origin of the aberrations. In atmospheric adaptive optics, the source of the aberrations is known to arise from the turbulence in the atmosphere. These aberrations can be characterised, and Kolomgorov sought to present a statistical analysis of the aberrations[11]. A whole field of adaptive optics research is driven to bet-

ter understand the profile of atmospheric turbulence with techniques such as SCIDAR[12] and SLODAR[13]. In other imaging systems the aberrations arise from the physical components, be they the optical lenses of the human eye, or the optical lenses of a microscope system. In either case the aberrations can be characterised in a known way. The two main methods used to describe optical aberrations in adaptive optics are Karhunen-Loeve[14] and Zernike polynomials[15, 16], the latter being discussed in greater detail in Chapter 2.3. Firstly, the fundamentals of optical aberrations need to be examined.

2.2.1 Scalar Diffraction Theory

There are a number of different theories of diffraction and these can be divided into two classes, vector and scalar. Scalar treatments for describing diffraction include Huygens' principle[17], Rayleigh-Sommerfeld theory[17, 18], the Kirchhoff formulation[17, 18], and a model referred to as the angular spectrum of plane waves[17, 19]. The well-known Fraunhofer and Fresnel approximations can be derived from the any of the models, which are equivalent in this limit[17]. Vector theories include a treatment by Stratton and Chu[20] and a rigorous electromagnetic boundary value model which appears in Jackson's text[21]. Each of these theoretical models have strengths and weaknesses, and each can be satisfactorily employed for some range of problems. The choice of an appropriate model is based on what is known about a specific problem.

Scalar diffraction theory can be regarded as a first approximation for optical diffraction which can lead to a description of optical aberrations. In this treatment of diffraction, Huygens' principle is used as a starting point. Huygens wrote a treatise called *Traite de la Lumiere* on the theory of light, and his work stated that the wavefront of a propagating wave of light at any instant conforms to the envelope of spherical wavelets emanating

from every point on the wavefront at the prior instant[22]. From this principle Huygens was able to derive the laws of reflection and refraction, but the principle is deficient in that it fails to account for the directionality of the wave propagation in time. Also, the principle stated by Huygens does not account for diffraction. Subsequently, Fresnel elaborated on Huygens principle by stating that the amplitude of the wave at any given point equals the superposition of the amplitudes of all the secondary wavelets at that point. The Huygens-Fresnel principle is adequate to account for a wide range of optical phenomena and it was later shown by Kirchhoff how this principle can be deduced from Maxwell's equations[23]. The Fresnel-Kirchhoff theory can be used as a starting point to describe optical aberrations and is discussed in the following Section.

2.2.2 Fresnel-Kirchhoff Diffraction

Light is an electromagnetic wave with coupled electric and magnetic fields traveling through space. Disregarding polarization, the field can be described by a scalar function $\mathcal{U}(x;t)$ representing either the electric or the magnetic field amplitude. The time dependence of the field is harmonic and can be explicitly written as,

$$\mathcal{U}(x;t) = Re\left(U(x)\exp^{(-jwt)}\right). \tag{2.1}$$

Here w, denotes the angular frequency of the light, and the complex-valued amplitude $U(x)$, depends on the spatial coordinates only. If both quantities $\mathcal{U}(x;t)$ and $U(x)$ represent an optical wave, they must satisfy the wave equation and the Helmholtz equation,

$$\left[\bigtriangledown^2 - \frac{1}{c^2}\frac{\partial^2}{\partial t^2}\right]U(x;t) = 0, \quad and \quad \left[\bigtriangledown^2 + k^2\right]U(x) = 0 \tag{2.2}$$

where $\bigtriangledown^2$ is the Laplacian operator.

The Helmholtz equation directly follows from the Maxwell equations under the condition of a homogeneous medium, and in the absence of sources. The quantity k, is termed the wave number, or propagation constant, of the medium and is related to the light velocity c, the angular frequency w, and the vacuum wavelength λ, by,

$$k = \frac{w}{c} = \frac{2\pi}{\lambda} \quad and \quad c = \frac{c_0}{n} \tag{2.3}$$

where n, is the refractive index of the medium. The complex disturbance $U(x)$, at any observation point x_0, in space can be computed from the Helmholtz equation (eq: 2.2) with the help of Green's theorem. If a unit-amplitude spherical wave, expanding about the observation point x_0, is chosen as Green's function, i.e.,

$$G(x) = \frac{\exp^{jkr_0}}{r_0} \qquad with \qquad r_0 = \|x_0 - x\|, \tag{2.4}$$

the so-called integral theorem of Helmholtz and Kirchhoff is obtained,

$$U(x_0) = -\frac{1}{4\pi}\int_S\int \left[\frac{\partial U(x)}{\partial n}G(x) - U(x)\frac{\partial G(x)}{\partial n}\right] ds. \tag{2.5}$$

This relation plays a central role in the derivation of the scalar theory of diffraction as it allows the field $U(x)$, at an observation point x_0, to be expressed in terms of the "boundary values" on any closed surface S, surrounding x_0.

The integral theorem of (eq: 2.5) can readily be used to study diffraction effects occurring at a plane object. Therefore the integration surface S, is segmented into three disjoint parts, i.e., $S = A, B, C$. The boundary A, is chosen across the object location in the

screen plane, B, is the opaque part of the screen, and C, is a half-sphere containing the observation point x_0. A graphical illustration of this specific choice is given in Figure 2.2.

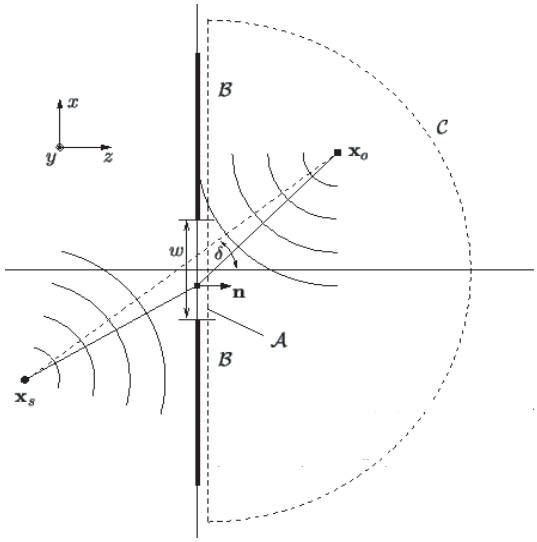


Fig. 2.2: Range of validity of scalar diffraction theories.

The solution of the diffraction problem is found by specifying Kirchhoff boundary conditions on A and B, and the Sommerfeld radiation condition on C :

$$on\ A: \quad U(x) = U_s, \quad \frac{\partial U(x)}{\partial n} = \frac{\partial U_s(x)}{\partial n},$$

$$on\ B: \qquad U(x) = 0, \quad \frac{\partial U(x)}{\partial n} = 0, \tag{2.6}$$

$$on\ C: \quad \lim_{\|x\|\to\infty} \|x\| \left(\frac{\partial U(x)}{\partial n} - jkU(x)) = 0 \right),$$

whereby $U_s(x)$, is the field incident on the screen. The Kirchhoff boundary conditions are an idealization of the real field distribution on the screen by assuming that the field and its derivative across A, is exactly the same as they would be in the absence of the screen. In the geometrical shadow across B, the field is simply set to zero. Although these assumptions are reasonable, they fail in the immediate neighbourhood of the rim of the opening. Thus, the range of validity is related to the wavelength and the geometrical dimension, w, of the opening (cf. Figure 2.2). The Sommerfeld radiation condition on C is satisfied, if the disturbance, $U(x)$, vanishes at least as fast as a spherical wave. Since the illuminating light, $U_s(x)$, invariably consists of a linear combination of spherical waves, this requirement is always fulfilled.

Inserting the boundary conditions of (eq: 2.6) into the integral theorem (eq: 2.5) shows that only the illuminating light, $U_s(x)$, across the opening A, contributes to the integral, i.e.,

$$U(x_0) = -\frac{1}{4\pi}\int\int_A \left[\frac{\partial U_s(x)}{\partial n}G(x) - U_s(x)\frac{\partial G(x)}{\partial n}\right] ds. \tag{2.7}$$

Its further assumed that the aperture is illuminated with a single expanding spherical wave arising from a point source located in x_s, a distance r_s, away from the screen,

$$U_s(x) = A_s\frac{\exp^{jk_0 r_s}}{r_s} \qquad with \qquad r_s = \|x_s - x\|. \tag{2.8}$$

Hence, both Green's function in (eq: 2.4) and the incident disturbance in (eq: 2.8) have the shape of spherical waves. A further simplification is obtained by noting that the distances from the screen to the observation point and the location of the point source, r_0 and

r_s, respectively, are many optical wavelengths. Therefore the following approximations hold for $\lambda \ll r_0, r_s$.

$$\frac{\partial U_s(x)}{\partial n} \cong jk_0 A_s \cos(n, x_s - x) \frac{\exp^{jk_0 r_s}}{r_s} \quad and \quad \frac{\partial U_0(x)}{\partial n} \cong jk_0 \cos(n, x_0 - x) \frac{\exp^{jk_0 r_0}}{r_0}. \tag{2.9}$$

Insertion of these approximations into (eq: 2.7) yields together with (eq: 2.4) and (eq: 2.8) the Fresnel-Kirchhoff diffraction formula,

$$U(x_0) = \frac{jA_s}{\lambda} \int\int_A \frac{\exp^{jk_0(r_s + r_0)}}{r_s r_0} \left[\frac{\cos(n, x_0 - x) \cos(n, x_s - x)}{2} \right] ds, \tag{2.10}$$

where the term in square brackets is called obliquity or inclination factor. This factor was introduced by Fresnel because he believed that the effect of the element ds would be greater in the forward direction than in an inclined direction. The Fresnel-Kirchhoff formula forms a starting point that is used to characterise optical aberrations and is also used by Muller and Buffington to prove their sharpness metrics shown in Chapter 4.

2.3 Representing the Wavefront - Zernike Polynomials and Seidel Aberrations

In the theory of optical aberrations, Zernike polynomials are very often used to describe the aberrations[15, 16]. They were introduced in 1934 by F. Zernike who deduced them from the Jacobi polynomials and slightly modified them for applications in optics. Zernike polynomials have the advantage that they are mathematically well defined, and that low

order terms are related to the classical aberrations like astigmatism, coma and spherical aberration.

Although interferometrists often like to represent wavefronts by Zernike polynomials, optical designers are more accustomed to Seidel aberration polynomials. The Seidel aberrations were developed in the mid-19th century to account for the monochromatic geometrical aberrations of centered optical systems, i.e. defects from perfect imagery in optical systems that have an optical axis. These types of aberrations can be developed from considerations of symmetry, and are given names such as spherical aberration, coma, astigmatism, field curvature, and distortion[17]. In addition to their type, aberrations are usually specified according to their order (third-order, fifth-order, etc.), although sometimes they are called primary, secondary, etc.

Seidel aberrations are generally used for, and can be derived by, ray tracing data, where as Zernike's are used to describe the aberrations over a pupil. It is generally possible to compute a Zernike polynomial from Seidel coefficients. Zernike aberrations can be thought of as a summation of individual Seidel aberrations, and so in this way, Zernike's$\equiv \sum$Seidel's, therefore both can be computed from ray tracing data. Because of their relationship Zernike polynomials can be computed from Seidel aberrations, and vice versa.

The Zernike coefficients are ordinarily found using a least-squares fit to a grid of exact ray data, while the Seidel coefficients can be computed from paraxial ray data. A more detailed analysis on Seidel aberrations and their distinction from Zernike aberrations can be found in Born and Wolf[17]. As Zernike polynomials are the the primary classification in adaptive optics an explanation of their basis is presented below.

It is convenient to express the wavefront aberrations in terms of Zernike polynomials since the Zernike polynomials are defined on the unit circle. Noll[15] introduced a normalisation for the polynomials that is particularly suited for application to wavefront aberrations. In this normalisation, the rms value of each polynomial over the circle is set equal to one. The Zernike polynomials form a set of orthogonal polynomials and it is convenient to write them as a function of ρ and θ[24]:

$$Z_{jeven} = \sqrt{n+1}R_n^m(\rho)\sqrt{2}\cos(m,\theta),\ for\ m \neq 0,$$

$$Z_{jodd} = \sqrt{n+1}R_n^m(\rho)\sqrt{2}\sin(m,\theta),\ for\ m \neq 0,$$

$$Z_j = \sqrt{n+1}R_n^0(\rho),\ for\ m = 0,$$

where,

$$R_n^m(\rho) = \sum_{s=0}^{\frac{n-m}{2}} \frac{(-1)^s(n-s)!}{s!(\frac{n+m}{2}-s)!(\frac{n-m}{2}-s)!}\rho^{n-2s}.$$

Representation of some of the lower order Zernike aberrations can be seen in Figure 2.3. As shown above, the Zernike polynomials are commonly characterised by radial order n, and an azimuthal order m. Frequently, a continuous numeration with single index, j, is used, instead of the two indices, n and m. For a given radial order N, there are a total of $(N+1)(N-1)/2$ Zernike polynomials. The power of Zernike modes comes from the fact that they are orthonormal; the scalar product $Z_i.Z_j$ is equal to 1 if $i = j$ and zero otherwise.

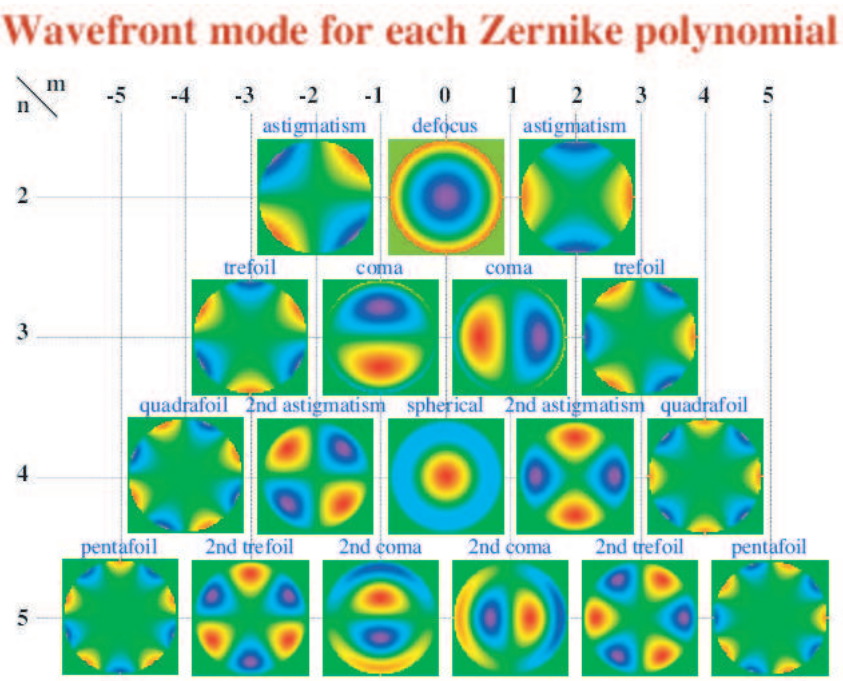


Fig. 2.3: Wavefront modes for first 5 orders of Zernike polynomials.

In many adaptive optics systems the wavefront is estimated by the coefficients of a Zernike expansion, which can then be used to drive the adaptive optics system, primarily the deformable mirror. However, whilst Zernike modes are a useful tool in AO, they have no fundamental importance; the modes of fundamental importance are those of the deformable mirror and wavefront sensor.

2.4 Wavefront Reconstruction

In this section the problem of computing the wavefront shape from the data provided by a wavefront sensor is addressed in a general way. The measurements (wavefront sensor data) can be represented by a vector S (its length is twice the number of sub-apertures, N,

for a Shack-Hartmann wavefront sensor, because slopes in two directions are measured, and equal to N for curvature sensing). The unknowns (wavefront) is a vector ϕ, which can be specified as phase values on a grid, or, more frequently, as Zernike coefficients. It is supposed that the relation between the measurements and unknowns is linear, at least in the first approximation. The most general form of a linear relation is given by matrix multiplication,

$$S = A\phi,$$

where the matrix, A, is called interaction matrix. In real adaptive optics systems the interaction matrix is determined experimentally: all possible signals (e.g. Zernike modes) are applied to a deformable mirror, and the wavefront sensor reaction to these signals is recorded. A reconstructor matrix, B, performs the inverse operation, retrieving wavefront vector from the measurements:

$$\phi = BS.$$

The number of measurements is typically more than the number of unknowns, so a least-squares solution is useful. In the least-squares approach a phase vector ϕ, that would best match the data is sought. The resulting reconstructor is,

$$B = \left(A^T A\right)^{-1} A^T.$$

Here the superscript, T, means a matrix transpose, and superscript, -1, means an inverse matrix. Matrix operations are very frequently encountered in adaptive optics. In almost all cases the matrix inversion presents problems because the matrix $A^T A$ is singular, which means that some parameters (or combinations of parameters) are not constrained by the data. For example, the first Zernike mode (piston) can not be determined from the slope measurements (but fortunately this is not important). In practice the matrix inversion is done by removing the undetermined (or poorly determined) parameters with the help of a singular value decomposition (SVD) algorithm[5, 6].

Wavefront sensors can be insensitive to certain kinds of wavefront errors. In Shack-Hartmann systems with a square geometry, poorly determined modes typically include "waffle" (quasi-periodic deformation with actuator-grid frequency). Waffle is a chess board-like pattern of phase error. This zero mean slope phase error is low over one set of wavefront subapertures arranged as the black squares of a chess board, and high over the white squares. These waffle modes can be present in the reconstructed wavefront unless the wavefront reconstructor explicitly removes or attenuates these modes which can corrupt the corrected image. Clearly, the wavefront sensor should be designed so that all mirror modes can be sensed with a good signal-to-noise ratio, to avoid "waffle" and similar effects.

The least-squares reconstructor is not the best one. There are two basic approaches to derive the control matrix which are distinguished from each other by the quantity that is minimised in the solution process. In the first approach the control matrix is derived by minimising the difference between the measured wavefront slopes and the slopes associated with the deformable mirror surface. This minimisation falls within the formalism of "maximum a posteri estimation" and "least squares" estimation[25]. In the second

approach statistical knowledge of both the wavefront statistics and the wavefront sensor noise characteristics are used to derive a control matrix that minimizes the aperture average mean square residual error. Using *a priori* information on the signal properties means a better reconstruction can be achieved. In the adaptive optics context it means that both compensation order and servo bandwidth are reduced when there is not enough photons. These two approaches are further developed by Roggemann and greater detail can be found in[26].

There are two approaches to wavefront reconstruction, one for modal correction and one for zonal correction. In modal reconstruction, for example, the modes of a polynomial, e.g. Zernike modes are reconstructed from the information obtained by the wavefront sensor. In the zonal approach, for example, the error in each subaperture of a Shack-Hartmann sensor is minimised by tilting the wavefront in the subaperture. In a curvature sensor system this approach is even more intuitive in combination with a bimorph as is shown in Chapter 3.

In both cases the local piston of the wavefront elements in each subaperture have to be treated separately in order to smoothly model the wavefront. This requires some difficult reconstruction techniques[27]. Together with the high accuracy that is required in the opto-mechanical alignment to ensure a precise correspondence between the wavefront sensor elements and the deformable mirror actuators, the zonal approach becomes less attractive than the modal approach.

2.5 Measures of Optical Quality

For extended objects it is difficult to quantify the correction achieved in a given image. Visually, there should obviously be an improvement, but its much harder to measure a related quality. Power spectrum analysis of corrected and uncorrected images provides some indication of increased image quality. For correction of USAF target images, it's clearer to see improvement in resolution due to the sections of decreasing spatial frequency. They also allow for the profiling of individual bar sections, from which, improvement in contrast can be measured. The point spread function (PSF) of an image, its optical transfer function (OTF), modulation transfer function (MTF) and Strehl ratio (SR) describe the effects and degradation of an object being imaged, as can be seen in Figure 2.4. These object-image relationships are discussed in the following section.

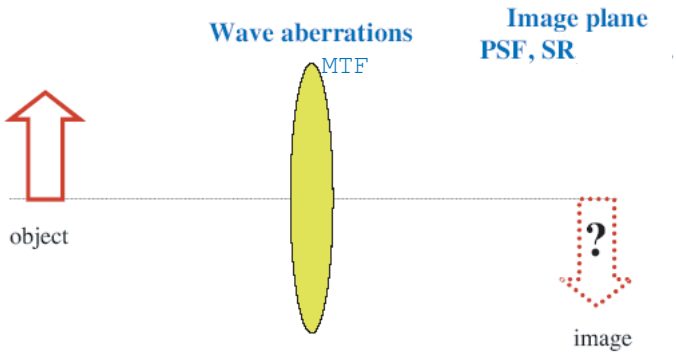


Fig. 2.4: Image quality relationships between object and image.

2.5.1 Point Spread Function

The PSF describes the two-dimensional distribution of light at the focal plane of an imaging system for a point source. It can be represented by $PSF = \left|\left(\tilde{P}(x,y)\right)\right|^2$, where $P(x,y)$

is the pupil function and $\sim$ denotes a Fourier transform. The PSF for a perfect optical system, based on circular elements, would be an "Airy pattern", which is derived from Fraunhofer diffraction theory. This is used to determine the Rayleigh resolution for two point sources. If the two point sources of light overlap such that the centres of the images are closer than the radius of the Airy disk, the images are considered to be unresolvable. The Rayleigh resolution limit is given by,

$$r = \frac{1.22\lambda f}{D},$$

where λ, is the wavelength of the light, f is the focal length and D is the aperture size.

2.5.2 Optical Transfer Function

The optical transfer function (OTF) can be represented by ,

$$\tilde{I}(\overrightarrow{f}) = \tilde{O}(\overrightarrow{f}) \cdot \tilde{P}(\overrightarrow{f})$$

Here $(\overrightarrow{f})$ is the spatial frequency, I is the image and O represents the object. $\tilde{P}(\overrightarrow{f})$ is called the optical transfer function. It describes the change of the modulus and phase of the object Fourier transform (FT) in the imaging process. The modulus of the OTF is called the MTF. For incoherent imaging, $\left|\tilde{P}(\overrightarrow{f})\right| \leq 1$. Typically, the MTF decreases with increasing frequencies, hence the small (high-frequency) details in the image are weakened and eventually lost.

It is known for any optical system $\left|\tilde{P}(\overrightarrow{f})\right| = 0$ for $\left|(\overrightarrow{f})\right| \geq f_c$, where $f_c = \frac{D}{\lambda}$ is called the cutoff frequency, and D is the maximal size of the aperture.

The relation between the PSF and OTF is a Fourier transform, as can be seen in Figure 2.5, so if you know one you have the other. From Fourier transfer properties it follows that $P(0) = \int_{-\infty}^{\infty} P(\tilde{u})du$ and that the Strehl ratio (described in Section 2.5.4) is proportional to the integral of the OTF over all frequencies. Figure 2.5 also presents the relationship between other image quality relations.

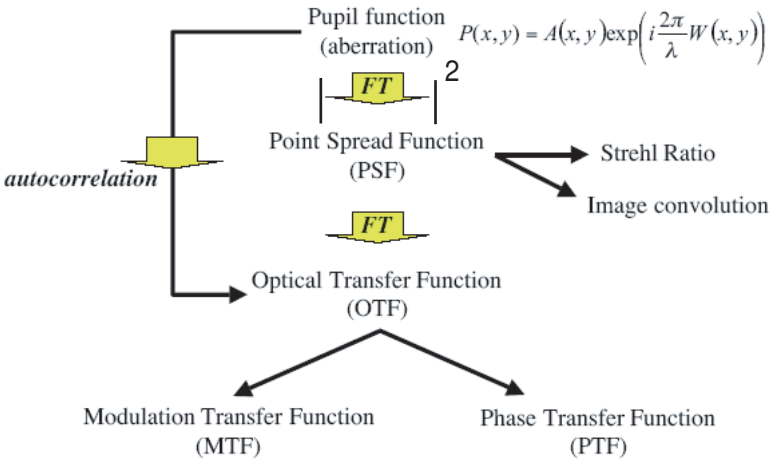


Fig. 2.5: Aberration vs image quality for incoherent imaging - sets out relationship between pupil function and other image quality relations.

2.5.3 Wavefront Variance

Before discussing the Strehl ratio, it is important to describe wavefront variance, presented below.

The normalised intensity at a Gaussian focus for an aberration free lens can be represented by:

$$i(P) = \frac{1}{\pi^2}\left|\int_0^1 \int_0^{2\pi} \exp\left(ikW\left(\rho,\theta\right)\right)\rho d\rho d\theta\right|^2. \tag{2.11}$$

For small aberrations this can be written as,

$$\exp\left(ikW\left(\rho,\theta\right)\right) \approx 1 + ikW\left(\rho,\theta\right) + \frac{1}{2}ikW\left(\rho,\theta\right)^2 + \ldots$$

$\left\langle W^2 \right\rangle$ denotes the average value of $W^n\left(\rho,\theta\right)$ over the pupil, i.e.,

$$\left\langle W^2 \right\rangle = \frac{\int_0^1 \int_0^{2\pi} W^n \rho d\rho d\theta}{\int_0^1 \int_0^{2\pi} \rho d\rho d\theta},$$

$$= \frac{1}{\pi}\int_0^1 \int_0^{2\pi} W^n \rho d\rho d\theta.$$

Equation 2.11 now becomes,

$$i(P) \approx \left| 1 + ik\left\langle W \right\rangle + \frac{1}{2}ik^2\left\langle W \right\rangle^2 \right|^2,$$

$$\approx 1 - k^2\left(\left\langle W^2 \right\rangle - \left\langle W \right\rangle^2\right),$$

where the term $\left(\left\langle W^2 \right\rangle - \left\langle W \right\rangle^2\right)$ is the variance σ_w^2, of the wavefront aberration over the pupil.

2.5.4 Strehl Ratio & Maréchal Criterion

The relationship between the phase aberration and image quality is quantified, in a high quality imaging system, by the Strehl ratio, defined as the ratio of the central intensities of the aberrated point spread function and the diffraction-limited point spread function:

$$S = \frac{I(0,0)_{aberrated}}{I(0,0)_{unaberrated}},$$

where $I(\xi,\eta)$ is the intensity point spread function, and (ξ,η) are image plane coordinates.

For small, arbitrary aberrations, $\sigma_\phi^2 \ll 1$, the Strehl ratio is related to the variance, σ_ϕ^2, of the phase aberration by:

$$S \simeq 1 - \sigma_\phi^2,$$

where σ_ϕ^2 is the phase variance of the wavefront aberration over the pupil. A system is "well-corrected" when $S \geq 0.8$, the equality being called the Maréchal criterion. At the Maréchal limit,

$$\sigma_\phi^2 = \left(\frac{2\pi}{\lambda}\right)^2 \sigma_W^2 = 0.2\,rad^2,$$

corresponding to an rms wavefront aberration of $\sigma_W \simeq \lambda/14$, or a wavefront variance, $\sigma_W^2 \simeq \lambda^2/200$.

2.5.5 Task-Based Assessment

Image quality metrics are methods by which image correction is often evaluated. The Strehl ratio is an often used criterion used for point source images in astronomy, which gives a clear and intuitive assessment of image quality which may correspond to perceived visual improvement. The modulation transfer function and optical transfer function are measures of how faithfully an object has been imaged, and as described previously, the variance of wavefront phase is linked to the image quality. These qualities all have a mathematical basis and can be readily understood.

They are specific in their judgement of whether an image is optically improved. They do not however, provide any latitude for the ability of the observation method to discern specific improvement. Reasons include the limited dynamic range of display and the extent to which human vision encodes such departures from perfect imaging. Regarding human observation, human eyes will be insensitive to marginal increases in image quality, and as such, what is referred to as task-based assessment should be considered for some imaging systems. The actual image improvement relative to the system components and task needs to be considered.

Barrett et al[28] assert that "scientific and medical images are acquired for specific purposes, and the quality of an imaging system is ultimately determined by how well the images fulfill those purposes". The task of the imaging system is to learn something about the object that produced the image. As such, tasks can be divided into two classes, classification and estimation. Classification labels the object or determines which class it belongs to. Estimation tasks involve the extraction of numerical information from the images. Barrett and Myers examine how well the task can be performed, which depends not only on

the task and imaging system, but also in the means by which the task is performed, or the observer. Some classification tasks will have a human observer and, alternatively, images can be classified by computer algorithms or mathematical models. Estimation tasks are less often performed by humans and mostly are analyzed via computer algorithm.

Task-based assessment is something that is required in adaptive optics[28]. As an example, consider the problem of comparing the suitability of different deformable mirrors in and adaptive optics system. Intuitively one might assume that a mirror with a higher number of actuators would provide the best correction. This is not necessarily the case in practical experiments as the stroke of each actuator may be too small, or the influence functions too wide. This is demonstrated in a comparative analysis of deformable mirrors for ocular adaptive optics by Dalimier and Dainty[29]. Here 3 types of deformable mirrors are evaluated in the correction of modeled ocular aberrations. It was shown that a piezo-electric deformable mirror with 19 actuators performs better than a bimorph mirror with 37 actuators. Also a mirror with a high stroke may not necessarily correct efficiently if the actuators are spaced too far apart as this can result in broad influence functions.

For applications where an image is being corrected for human visual perception it may not be critical to completely minimise the aberrations, as at a certain limit, further correction will be imperceptible. As an example, an image may be greatly and sufficiently improved for correction of lower order modes and further correction time not be required. In this sense the task for which correction is being employed can dictate the level of correction as opposed to metrics such as Strehl ratio.

3. WAVEFRONT SENSING AND CORRECTION DEVICES

Introduction

The various types of wavefront sensor can largely be divided into two generic classes: image-plane sensors and aperture-plane sensors. The aperture-plane sensors measure the phase gradient or wavefront slope directly, while the image-plane sensors generally measure the intensity distribution and deduce from this the wavefront aberration. Aperture plane sensing can be viewed as ***direct*** wavefront sensing, and image-plane sensing usually as ***indirect*** wavefront sensing.

Although most wavefront correction techniques fall into these two categories, some techniques do not strictly belong to either classification. Phase diversity is a technique which measures the phase based on image-plane, not aperture-plane, information. In this way it varies from true indirect methods as the wavefront phase is sought as in direct methods, but based upon image-plane measurements. For our purposes phase diversity is described as part of indirect wavefront sensing in Chapter 3.2.

3.1 Direct Wavefront Sensing

In direct wavefront sensing, a device is required to sense the wavefront with high spatial resolution and speed to apply realtime correction[27]. Direct methods provide informa-

tion about wavefront phase and this information is used to drive a wavefront corrector as described in Chapter 2.4. Generally direct methods are employed in atmospheric and vision science where correction needs to be made at a higher rate, and is the most common form of correction in adaptive optics. Further discussion on direct wavefront sensors is presented below; three different types of higher order wavefront sensor (direct sensing) are described in this Section.

3.1.1 Shack-Hartmann

The Shack-Hartmann[30] sensor, Figure 3.1, divides the telescope aperture into an array of smaller subapertures, and a lenslet array is used to produce multiple images. When an incoming wavefront is plane, all images are located on a regular grid defined by the lenslet geometry. A distorted wavefront which is incident on the lenslet array will cause the the images to be displaced from the normal positions. Displacements of image centroids in two orthogonal directions x, y are proportional to the average wavefront slopes x, y over the subapertures.

Thus, a Shack-Hartmann wavefront sensor measures averaged wavefront slopes over each lenslet. If $\phi(\overrightarrow{r})$ is the wavefront phase, the x-slope measured by a Shack-Hartmann wavefront sensor is,

$$x = \frac{\lambda}{2\pi S} \int_{sub-aperture} \frac{\partial \phi(\overrightarrow{r})}{\partial r_x} d\overrightarrow{r},$$

where S is the area of sub-aperture. The x-slopes are typically estimated from the displacements of the image centroid,

$$x = \frac{\sum_{i,j} x_{i,j} I_{i,j}}{\sum_{i,j} I_{i,j}},$$

where $I_{i,j}$, are intensities of light on the detector pixels, with a similar expression for the y-slopes. The centroid displacement of each subimage gives an estimate of the average wavefront gradient over the subaperture. The measurement error of the image centroid is due to photon noise and read noise in the detector. In practice Shack-Hartmann wavefront sensors used to image extended objects often use a correlation method to determine the centroids and hence the wavefront slopes.

The wavefront itself is reconstructed from the arrays of measured slopes. The Shack-Hartmann sensor is achromatic - the image movement is independent of wavelength and extended sources of limited extent can be used as long as they fit into the subimage boundary.

In practice, a Shack-Hartmann sensor is built by putting a lenslet array in the re-imaged pupil plane. The subimages from each subaperture are re-imaged (usually de-magnified) onto a CCD camera by a lenslet array.

3.1.2 Curvature Sensor

The idea of curvature sensing was first discussed by Roddier[31, 32]. It relies on measuring the intensity distribution in two different planes on either side of focus, using the normalized difference between the distributions. This difference is a measure of the curvature of the wavefront in the entrance pupil of the optical system and of the wavefront tilt at the pupil edge. The principle is outlined in Figure 3.2. The central lines show the

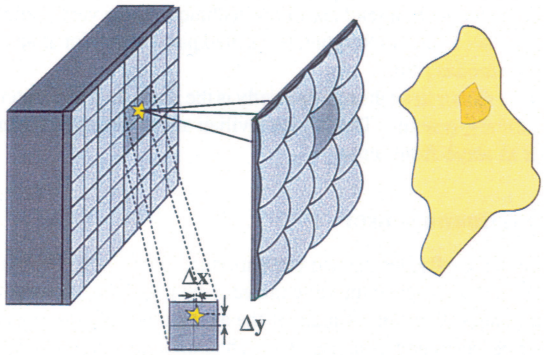


Fig. 3.1: Measurement principle of a Shack-Hartmann sensor. Image from each lenslet is imaged onto detector and displacement of aberrated image points from plane wave are measured.

rays from a curved part of the wavefront form a focus before the focal plane, leading to a local increase in intensity plane P_1 and a decrease in P_2. The two intensity distributions are recorded in the two planes, P_1 and P_2, each a distance δ, from the focal plane.

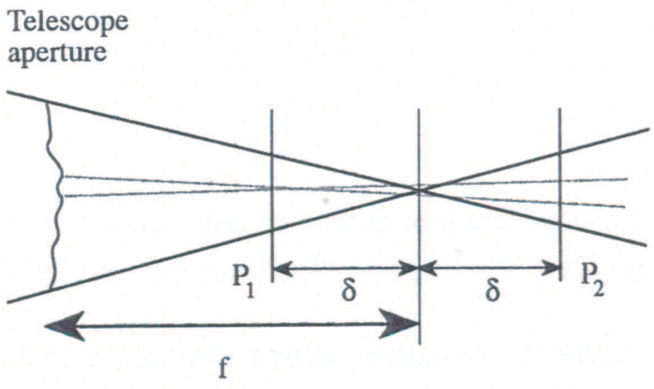


Fig. 3.2: Principle of curvature sensing. Curvature of aberrated wavefront causes an increase in intensity at P1 and a decrease in intensity at P2.

For extended sources and different degrees of correction the situation becomes more complicated. The local resolution in the wavefront measurement - given by the size of the subapertures in a Shack-Hartmann sensor - is determined by the size of the blur that

is caused, *e.g.* by the small piece of curved wavefront in Figure 3.2. If the separation δ, is too small, this blur is too small to be measured. In addition, if the detector pixels are too large this blur cannot be resolved and the mode corresponding to aberrations of this size cannot be measured. Thus, the separation, δ, and the pixel size have to be adjusted according to the degree of sensitivity required.

The normalized difference between the two required intensity distributions is [31]:

$$c(x,y) = \frac{I_+(x,y) - I_-(x,y)}{I_+(x,y) + I_-(x,y)} \propto \left[\frac{\partial}{\partial \rho} \phi(\rho,\theta) \Psi - \bigtriangledown^2 \phi(\rho,\theta) \right],$$

where $\bigtriangledown^2 = \frac{1}{\rho} \frac{\partial}{\partial \rho} \left(\rho \frac{\partial}{\partial \rho} \right) + \frac{1}{\rho^2} \frac{\partial^2}{\partial \theta^2}$, is the Laplacian operator representing the curvature of the wavefront. This equation is derived using the so-called intensity transport theory, which is essentially a geometrical optics approximation to wave propagation. The wavefront radial tilt $\frac{\partial \phi}{\partial \rho}$, has to be weighted by an impulse distribution, Ψ, around the pupil edge. Curvature sensing has the advantage that its output can be directly coupled to some correction devices, such as bimorph deformable mirrors, which are discussed later[31][35], without any intervening computer (at least in principle).

3.1.3 Lateral Shearing Interferometer

The two wavefront sensing methods discussed so far rely on geometrical optics. A third method, the shearing interferometer, uses wave optics. Here, the wavefront interferes with a shifted version of itself in the pupil. The interference pattern, the fringes, are then used to determine the local slope of the wavefront. Figure 3.3 displays the basic principle. The form of the fringes in the interferogram are determined by the shape of the wavefront.

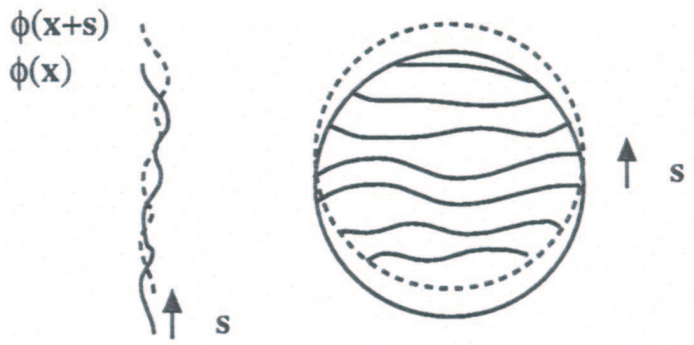


Fig. 3.3: Principle of the shearing interferometer. Aberrated wavefront interferes with a shifted version of itself causing an interference pattern which is the used to determine the local slope of the wavefront.

The signal that is measured in a shearing interferometer can be expressed quite simply as the sum of complex amplitudes,

$$I(\overrightarrow{x}) = \frac{1}{2}\left|\exp\left(i\phi(\overrightarrow{x})\right) + \exp\left(i\phi\left(\overrightarrow{x}+\overrightarrow{s}\right)\right)\right|^2,$$

$$= 1 + \cos\left(\phi\left(\overrightarrow{x}\right) - \phi\left(\overrightarrow{x}+\overrightarrow{s}\right)\right),$$

where ϕ, is the phase of the wavefront in *rad*, and the signal is wavelength independent.

It becomes independent of the wavelength if the shift is proportional to the wavelength. Grating interferometers have been used to achieve achromatic performance[33]. For small shift $|\overrightarrow{s}|$, the phase difference $\phi(\overrightarrow{x}) - \phi(\overrightarrow{x}+\overrightarrow{s})$, is approximately equal to the slope of the wavefront in the direction of the shear vector, $\overrightarrow{s}$. Using a Taylor approximation for small shear, s_x, along the x-axis gives,

$$\phi(\overrightarrow{x}) - \phi(\overrightarrow{x} + (s_x, 0)) = s_x \frac{\partial \phi(\overrightarrow{x})}{\partial x} + \varepsilon(s_x).$$

In order to reconstruct the wavefront, two interferograms with orthogonal shear are required. Extended sources reduce the fringe contrast since the quality that is being measured is the coherence function of $|\overrightarrow{s}|$. Like the curvature sensor, the wavefront sensor geometry has to be adjusted according to the object size. The subapertures that determine the spatial resolution of the slope measurements are represented by the size of the detector pixels.

Another type of wavefront sensor that is being increasingly used is a pyramid senor[34]. The basic idea is that a refractive pyramid is placed in the focal plane and dissects an image into four parts. Each beam is deflected, these beams form four images on the CCD detector and is similar in principle to the Foucault knife edge test[9].

3.2 Indirect Wavefront Sensing

Indirect techniques do not directly measure the wavefront, but use information related to the wavefront to provide the signal for the corrective element without reconstruction. In this sense they are not wavefront sensors per se but are an element of a closed loop system which operates to improve image quality and through this correct for wavefront aberrations. Indirect techniques include, image sharpness, phase diversity and phase retrieval and these techniques are outlined in in the following sections. Indirect methods are used more often in industrial and medical applications. One of the main indirect techniques

is phase diversity. Other techniques include image sharpness and various post processing correction methods. e.g. PSF deconvolution[78, 79, 80, 81]. Image sharpening and phase diversity are discussed in the following Section.

3.2.1 Phase Diversity

Phase retrieval[36] has been proposed as an alternative to conventional wavefront sensing methods for control of an adaptive optics system. It uses the principle output of the optical system to estimate the wavefront. A modification, called phase diversity, uses two or more images captured simultaneously. A general representation of the phase diversity technique can be seen in Figure 3.4.

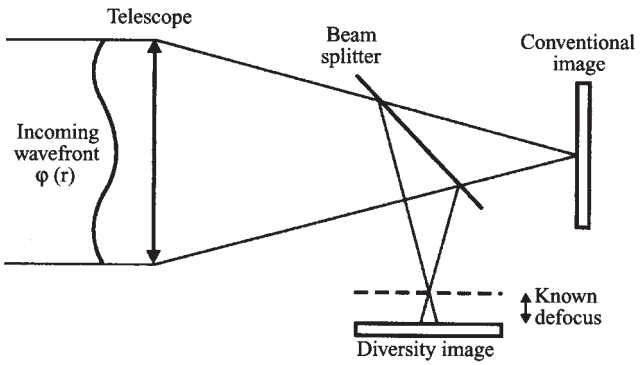


Fig. 3.4: Principle of the phase diversity technique. Two images are taken of an image, one of which is defocused by a known distance to introduce a quadratic diversity, which can be used to determine the phase of the wavefront.

Phase diversity[37] allows joint estimation of both the wavefront and an extended source. The diversity is introduced most simply by measuring the image in several focal planes. Phase diversity uses two or more images of an extended object to make joint estimation of the object and distorting wavefront. Typically, phase diversity is introduced

by changing the focal plane, a procedure that introduces quadratic phase diversity. The wavefront might be estimated by the coefficients of a Zernike expansion, which could then be used to drive the adaptive optics system. The literature on phase diversity is extensive and a review was recently presented by Gonsalves[38].

The phase diversity method offers several advantages over other aberration sensing methods[39]. The optical hardware is modest as requirements are essentially a beam splitter and second detector. The technique is less susceptible to systematic errors introduced by the hardware as it relies on an external reference: the object being imaged. The technique also works well for extended objects and finally, each photon is used for both imaging and aberration estimation. This is preferable to diverting, what can sometimes be valuable, photons from the imagery to a separate wavefront sensor.

3.2.2 Image Sharpening

This area of research is based on a "sharpness" criterion, which is used as an image metric to measure the degree of correction of the wavefront phase. The principle of image sharpening can be explained by Figure 3.5, a schematic diagram of image sharpening methods.

The image sharpness metric is a measure of the image quality and in general the higher this metric the better the image quality. In sharpness maximisation, a trial phase correction is applied to the image, via a corrector, and the effect on the sharpness metric is noted. Using a suitable sharpness metric, and a search algorithm which determines the trial phase to be applied to the corrector, the system is driven to maximise the sharpness metric and minimise the aberration.

The basic principle of indirect wavefront measurement systems is to make trial adjustments of phase in the optical aperture; determine the effect of each trial using an easily

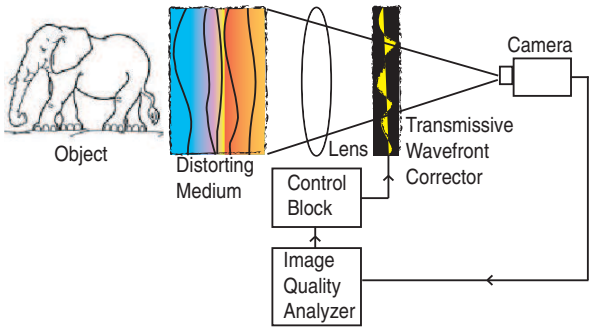


Fig. 3.5: Schematic of image sharpness correction system(after Vorontsov[73]) which shows mode of image correction without a wavefront sensor.

measurable quantity sensitive to wavefront phase, such as the intensity of a focused spot or "sharpness"; and then compute the correction required to maximise the measured quantity. The trial adjustments and phase corrections must be repeated at a rate high enough to track real-time changes occurring in the wavefront.

Many variations of this basic technique have been reported in the literature, both for the transmitted wave and received wave systems[41]. Two methods of aperture correction have been employed: frequency-division ("multi-dither")[42, 43] and time-division ("sequential"). In each case the trial perturbations in the aperture may be made zonally or modally. In this context, "zonally" implies independent perturbations of separate sections of aperture, whereas "modally" implies systematic perturbations of whole aperture using, for example, Zernike polynomials. Image sharpening based on sharpness maximisation is explained further in Chapter 4, where some of the available sharpness metrics are given.

3.2.3 Wavefront Coding

Wavefront coding (CDM Optics Inc) [44] is not a form of wavefront sensing as such, but is a "smart" optics device which can extend the depth of focus obtained from an image.

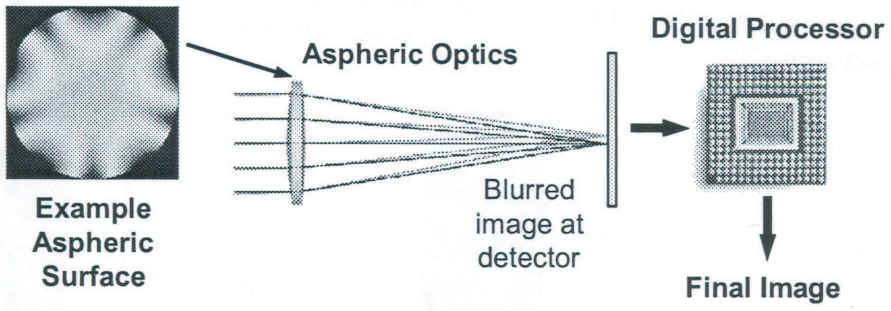


Fig. 3.6: Schematic wavefront coded imaging system. A special purpose blur is introduced into an image which causes invariance to various aberrations. Signal processing is then used to provide an image with an increased depth of focus.

Wavefront Coding increases the depth of field that can be seen in images through a manipulation of the imaging systems aberration. Wavefront coded[45] imaging systems differ from traditional imaging systems in that they use aspheric optics to form images with a special-purpose blur. This causes invariance to many optical aberrations including: spherical aberration, field curvature, astigmatism, chromatic aberration, defocus, and alignment related to defocus. Signal processing is used to remove the blur. The general system is shown in Figure 3.6. The aspheric optics can be a separate element of the imaging system, or can be integrated onto one or more optical elements. The signal processing is independent of the object being imaged and, in general, it depends only on the imaging optics and detector.

Joint design of the optics and signal processing is used to ensure that the amount and form of the blur is best suited to the amount and form of signal processing for minimization of noise effects. Signal processing to remove the image blur amplifies and changes the phase of the spatial frequencies. This amplification not only amplifies the spatial frequency content of the ideal image, but also the noise in the image. In practice it is this

noise amplification that typically sets the limit on the amount of benefit achieved from wavefront coding in a particular system configuration. The benefits of the wavefront coding methodology are balanced with a moderate decrease in the system signal-to-noise ratio(SNR)[46].

3.3 Wavefront Correction Devices

Introduction

Wavefront correctors are traditionally subdivided into two classes according to the implemented compensation technique - zonal and modal. Zonal correctors such as segmented piston and tip-tilt mirrors allow individual control of a phase over a set of subapertures providing step-wise phase compensation, whereas modal ones such as bimorph or membrane deformable mirrors use a set of smooth functions (modes) to approximate the required phase function[10]. There are two main types of wavefront correction devices - deformable mirrors and LCD spatial light modulators - and these are discussed in the following section.

Deformable Mirrors

Deformable mirrors are currently essential to the vast majority of adaptive optics systems and formed the basis of the initial proposal by Babcock in 1953[1]. A main advantage of these devices is their reflective nature which permits low light losses, an important aspect in energy-starved systems in astronomy and, to a lesser extent, vision science, where there are limited numbers of photons.

Deformable mirrors typically require a stroke of a few microns in astronomy and up to 10 microns for vision science, and must respond within the coherence time of the aberrations. In the past there have been three main classes of deformable mirrors. Segmented mirrors with single tip-tilt elements, and two mirror types with a continuous surface[47], bimorph mirrors[35, 48, 49] where the curvature of the mirror surface is affected by applying a force parallel to the surface, and piezo-electric mirrors[1], where a stack of actuators behind the mirror surface pushes and pulls the mirror by applying a force perpendicular to the surface. A more recent type of deformable mirror - membrane mirror, often referred as MEMs (Micro-Electro-Mechanical system)[50, 51] are also described.

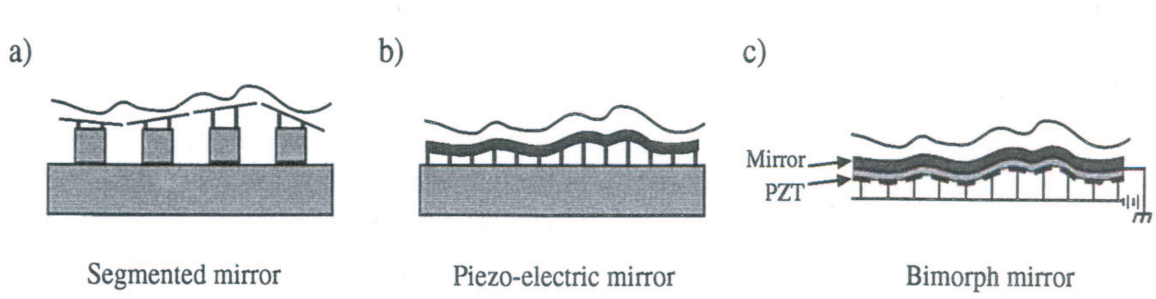


Fig. 3.7: Three classes of deformable mirrors.

Three classes of deformable mirrors can be seen in Figure 3.7. where 3.7(a), is a segmented mirror with tip-tilt actuators behind each segment, 3.7(b) and 3.7(c) are two types of continuous mirrors with piston actuation in the piezo-electric mirror 3.7(b), and bending actuation in the bimorph mirror 3.7(c). A short review of deformable mirrors is presented below outlining their mode of operation and advantages & disadvantages. The choice of deformable mirror can depend upon the chosen task and the quoted physical

[1] Bimorph mirrors are also composed of piezo-electric materials, but as a distinction, they are most often referred to solely as bimorph mirrors.

qualities of deformable mirrors such as actuator stroke and rms error need to be considered, depending on the system and purpose that the deformable mirror is chosen for. Dalimier and Dainty conducted a comparative study on 3 types of deformable mirror for the correction of ocular aberrations demonstrating this aspect of deformable mirrors[29]

3.3.1 Segmented Mirrors

Segmented mirrors (Figure 3.7a) have a number of advantages over continuous mirrors: the segments can be moved independent of each other, can be easily replaced, and the single segments can be combined to form large mirrors. Each subaperture of a Shack-Hartmann sensor can be associated with a segment of the mirror so that the high order adaptive optics system basically consists of many tip-tilt systems that can be run in parallel. Just as a bimorph (curvature) mirror is well suited to work with a curvature wavefront sensor, a segmented mirror is well suited to a Shack-Hartmann wavefront sensor. However, the discontinuous nature of segmented mirrors means one also has to avoid errors due to residual piston. The disadvantages include problems with diffraction effects from individual segments and the inter-segment alignment.

3.3.2 Continuous Facesheet Mirrors

Piezo-electric mirrors (Figure 3.7b) (continuous mirrors with an array of piezoelectric actuators expanding perpendicular to the mirror surface) are in widespread use now. They are available with >100s actuators, and the technology is well tested and very reliable. The typical voltage that is required to move the actuators is below $100V$, the bandwidth is in the KHz range, and the typical stroke is in the $5\mu m$ range. Some manufacturers use

electrostrictive material like a lead-magnesium-niobate (PMN) crystal that is similar to the PZT ceramics, but displays a smaller hysteresis and better motion control.

3.3.3 Bimorph Mirrors

Bimorph mirrors (Figure 3.7c) are constructed from a thin piezoelectric (PZT) material bonded to a thin mirror membrane. When a voltage is applied to a PZT material it expands in area, similar to a bimetallic strip. The curvature of the surface is proportional to the applied voltage. The application of a curvature to the wavefront makes bimorph mirrors the natural counterpart of the curvature sensor. As such, the signal from the curvature sensor can be fed directly to the bimorph mirror[31, 32, 35]. Bimorph mirrors are commercially available only at a lower number of actuators and can be expensive if custom built.

3.3.4 MEMs Devices

A more recent development in terms of deformable mirror has been the MEMs (Micro-Electro-Mechanical system) mirror. These mirrors consist of a thin membrane surface which can be manipulated when a voltage is applied. The advantage of these deformable mirrors is that they can be produced much more inexpensively than their predecessors, which often cost in the region of $1000 per actuator. For adaptive optics in vision science an actuator stoke of upwards of $10\mu m$, is required and this level of stroke can be provided by MEMs technology [50]. OKO Technologies[51, 52] produce a range of membrane mirrors for a relatively low cost and these are helping to extend the applications of deformable mirrors and increase their accessibility. This research uses a $15mm$, 37 channel actuator OKO Technologies membrane deformable mirror, which will be described later.

There are a range of commercially available MEMS mirrors, examples of which are Boston Micromachines[53] and IrisAO[54].

3.4 LCD Spatial Light Modulators

Spatial light modulator devices have the advantage that they are compact, lightweight and less expensive that deformable mirrors. They can also be used in transmission which is desirable for some applications. These attributes make them useful in many applications including those not traditionally associated with conventional adaptive optics[55, 56]. There are a number of liquid crystal (LC) correctors with pixelated structure of actuators whose operation is similar to that of piston-type segmented mirrors and also several configurations of LC devices with modal-type operation similar to that of deformable mirrors with a continuous face plate.

Traditionally, there are two major types of LC spatial light modulator:

1. Optically addressed spatial light modulators (OA-spatial light modulators), in which a two-dimensional image controls an output intensity, phase or polarization profile; and,

2. Electrically addressed spatial light modulators (EA-spatial light modulators), where the modulation is controlled by electrical signals[10].

3.4.1 Optically Addressed Spatial Light Modulators

A typical OA-spatial light modulator consists of a photoconductor and a thin LC layer situated between a pair of transparent electrodes. In the absence of recording light the photoconductor exhibits very high resistivity limiting the voltage across the LC layer. Sufficient

intensity of light lowers the resistivity of the photoconductor resulting in adequate voltage being dropped across the LC. Thus, local properties of the LC layer at a given point of the aperture depend on the intensity of the recording light. Due to the very high resolution of OA-spatial light modulators, the intensity of the recording beam is directly converted to the output intensity, phase or polarization profile.

Current OA-spatial light modulators are characterized by high diffraction efficiency and good temporal performance[57]. This allows the compensation of hundreds of waves of aberration caused by poor-quality primary mirrors in telescopes. Other applications of OA-spatial light modulators in adaptive optics are their use for sensing and correction of small phase distortions. Wavefront correction using a phase conjugation scheme with an OA-spatial light modulator placed in the Fourier plane has been demonstrated[58].

3.4.2 Electrically Addressed Spatial Light Modulators

The first electrically-addressed LC spatial light modulators with pure phase modulation were used as low-order wavefront correction devices such as adaptive LC lenses but their use has broadened[59]. Currently, Meadowlark Optics (USA)[60] produces transparent LC modulators with direct addressing of pixels, whose optical performance is similar to that of piston-type-segmented mirrors. Holoeye (Germany)[61] makes phase modulators with 1920x1200 pixels. Hamamatsu Photonics produces programmable phase modulators consisting of an EA-spatial light modulator with amplitude modulation coupled to an OA-spatial light modulator with phase modulation. Their recent model X8267, which is used in this project to create controlled aberrations, allows addressing of 768x768 pixels with a 100% fill factor[62].

One of the known drawbacks of LC phase modulators is their slow speed. For nematic LC modulators, the switching speed is about $100ms$ for 1 wave modulation. In its present state, this technology is not suitable for use in astronomy to correct for atmospheric turbulence, but there are a number of applications where fixed or slowly varying aberrations need to be compensated. For example, correction of static aberrations and heat effects in laser systems[63] and telescopes[64], adaptive focusing for machine and human vision, correction of the human eye aberrations[65], wide field-of-view imaging[66], and compression and shaping of femtosecond laser pulses[67].

4. IMAGE SHARPENING

4.1 Introduction

The concept of correcting for aberrations in an optics system, based on image metrics, was first proposed by Muller and Buffington in 1974[4]. This method uses a definition of sharpness of image to minimize aberrations in conjunction with a wavefront correcting medium. This technique was superceded with the advent of modern day wavefront sensors, such as the Shack-Hartmann, due to the speed of correction that they facilitated. Wavefront sensors measure quantities directly related to the wavefront which mean that the wavefront phase can be reconstructed. Therefore, deformable mirror commands to correct for the wavefront can be calculated and applied in realtime, offering a much faster speed of correction, whereas image sharpening relies on a search algorithm to find the solution in a search space which can often have a large number of degrees of freedom. However, with increasing CCD camera speeds and computation powers of modern day computers, this method can potentially be used for slowly varying or static aberrations, provided convergence to the minimum residual phase error can be achieved.

An image sharpness metric is a measure of image quality and in general the higher the metric the better the image quality. In sharpness maximisation, a trial phase correction is applied via a corrector and the effect on sharpness is noted. Using a suitable sharpness

metric and a search algorithm, which determines a trial phase to be applied to the wavefront corrector, the system is driven to maximise the sharpness metric and minimise the aberration. The two central issues are therefore;

1. choice of metric, and,

2. choice of search algorithm.

The sharpness maximisation relies on how the sharpness is defined. The problem arises when one quantitatively tries to define the word "sharp". Muller and Buffington define the sharpness such that its value for an aberration-degraded image is always less than that of the true image. In their paper Muller and Buffington set out eight sharpness metrics, some of which they showed were maximised for zero aberration. One such set of sharpness definitions are:

$$S_1 = \int I^n(x,y)dxdy \;\; n = 2 or 3 or 4, \; etc \tag{4.1}$$

where x, y denote coordinates in the image plane and $I(x,y)$ is the image irradiance. These sharpness definitions are maximised when there is zero wavefront error, in the presence of irregular object radiance distribution. This amplitude sensitivity makes this method useful for large extended objects. Other sharpness functions, such as higher-order moments of distribution, or entropy minimization functions,

$$S_2 = -\int I(x,y) ln[I(x,y)]dxdy, \tag{4.2}$$

have been examined by Muller and Buffington[4] and are shown to relate to low wavefront error. A review and analysis of sharpness was carried out by Doble[68]. Some met-

rics which have been developed are object dependant and the metric is chosen based on the scene[41, 71, 72, 73, 74, 75, 76].

4.2 Principle of Image Sharpening

Muller and Buffington[4] base their definition of sharpness by saying the sharpness S, of an image is one for which S reaches a maximum only for a true (undistorted) image. They consider only aperture plane distortion for the sharpness definition. Let $W(u,v)$ represent the complex amplitude at the telescope aperture without aberrations. If the distorted complex amplitude $W^{'}(u,v)$ is given by;

$$W^{'}(u,v) = W(u,v) \cdot \triangle(u,v),$$

then we have an aperture plane distortion $\triangle(u,v)$. Here $\triangle(u,v)$ is any complex function of the position (u,v) in the plane of the aperture. A system with a correcting element will allow the addition of extra phase shifts to $\triangle(u,v)$, for reduction or elimination of the distortion. Muller and Buffington sought to define the sharpness S, in such a way that any $\triangle(u,v)$ (which includes the effects of correction as well as the original distortion) other than a simple translation of the image, or a constant complex amplitude [$\triangle(u,v) = \exp^{ik(a+bu+cv)}$, where a, b and c are constants], will reduce the value of S.

Many definitions of image sharpness S, satisfy this criterion as shown in Chapter 4.3 An example is:

$$S = \int dxdy I^2(x,y),$$

where $I(x,y)$ is the irradiance at a point (x,y) in the image plane of the telescope system.

Muller and Buffington proved that for monochromatic light, S reaches its maximum only when $\triangle(u,v)$ is reduced to no more than the image translation described. This is proved using a version of the Fresnel-Kirchhoff equation (outlined in Chapter 2.2.3) to calculate the irradiance[4]. S is maximised for zero distortion regardless of the object-radiance distribution. This indicated that sharpness metrics could be used to provide feedback to the correcting element in an optical system to remove distortion, even for complex objects. There is no way which S can be increased beyond the value of an unaberrated image, regardless of the aperture distortion.

4.3 Image-plane Sharpness Functions

The metric used for image correction in practice could be any quantity that indicates system "quality" as affected by wavefront distortion and that can be calculated very quickly (in "real-time"). Depending on the type of adaptive optics system, the performance metric might be intensity of radiation at the focus, image sharpness[4], or scattered field statistical moments[73].

Feinup and Miller[72] explored variations of the sharpness metrics proposed by Muller and Buffington and found that different metrics worked better depending on the type of scene e.g. its spatial frequency content. Power law metrics with larger powers tend to perform better with scenes having prominent scatters, whereas power-law metrics with smaller powers perform better with scenes having no prominent scatters. The original power-law metrics set out by Muller and Buffington are of the form, $S = \int I^n(x,y)dx,dy$

where $n = 2,3,4$. Feinup and Miller extended their form to include the second derivative of the nonlinear point transformation.

This behaviour is due to the noise levels present in the images. An image with prominent scatters will be less susceptible to noise because the image will be more readily corrected as the power law metric will amplify the prominent scatter values to a greater extent than the noise. Conversely, an image with no prominent scatters will not benefit as the noise will be on a similar level as the signal and an amplification will not enhance the signal level with respect to the noise.

They found that the behaviour of a metric is determined by the second derivative of its point nonlinearity as a function of the image intensity. Its shown both theoretically and experimentally that metrics having similar second derivatives gave similar results when used in an image sharpening algorithm. For, example negative Shannon entropy acts in a similar way to a power law near unity. Metrics whose second derivative increases with increasing intensity emphasise bright points, whereas, metrics whose second derivative decreases with decreasing intensity emphasize making shadows and low return areas darker. Given these trends it is then possible to specify optimal metrics for particular types of images.

Vorontsov and Cohen[74] have applied image sharpening techniques to imaging extended objects. In one technique a coherent optical processor is used in which the signal from the imaging camera is used to control a spatial light modulator illuminated by a coherent wavefront. The output from the spatial light modulator will then be spatially modulated depending on the camera signal. If this is focused onto a rotating frosted glass plate a characteristic speckle pattern will be observed. The smaller the speckle scale, the

higher the spatial frequency content of the image. The system is then driven to minimise the speckle size[73].

Several of the sharpness metrics outlined by Muller and Buffington are listed below. We start with the S_1 metric shown below,

$$S_1 = \int I^2(x,y)dxdy, \tag{4.3}$$

Muller and Buffington proved this metric for monochromatic light and is maximised when there is zero aperture plane distortion. Although proven only for monochromatic light Muller and Buffington suggested that the metric should hold valid, at least for quasi-chromatic illumination.

Another sharpness metric is:

$$S_2 = I(x_0,y_0), \tag{4.4}$$

where the irradiance of the image at an particular point (x_0,y_0) in the image plane. Muller and Buffington suggest that objects that have a brightest spot S_2 will be maximised when the image of that bright spot is shifted over the point (x_0,y_0). This is essentially maximising the peak intensity of an image of a point source, taking that the maximum intensity of a point image will be at the central maximum.

A sharpness metric used in a system described by O' Meara[42, 43] is:

$$S_3 = \int |I(x,y)M(x,y)|\,dxdy, \tag{4.5}$$

where $M(x,y)$ is a mask function. If M is an accurate replica of the true undistorted image, then S_3 reduces to S_1 when distortion is eliminated.

As described earlier, the power-law metrics can be used as a measure of image quality:

$$S_4 = \int I^n(x,y)dxdy, \tag{4.6}$$

where $n = 2,3,4$.

A commonly employed sharpness metric, often known as "power in the bucket" is given by:

$$S_5 = \int_A I\, dxdy, \tag{4.7}$$

which is an integral of the intensity in some region A, of the image space - often a central disc or the maximisation of the amount of light passing through a pinhole.

The choice of the sharpness metric depends upon the object. S_3 has been shown to work well for point objects even for cases where the mask, M, only approximately matches the undisturbed image. For example M, can be a pinhole which has a diameter equal to the diffraction limited image. However this metric would not work as well for extended objects as *a priori* information about the object would be required, whereas, S_1 is expected to work better for extended objects[68].

4.4 Non-common Path Errors

As image sharpness maximisation determines corrections from the science camera, all the aberrations in the system can in principle be corrected. This is in contrast to direct wave-

front sensing, where only aberrations up until the beamsplitter which feeds the wavefront sensor are corrected for. Inherent aberrations in the imaging system after the wavefront sensor beamsplitter degrade image quality but are not seen by the wavefront sensor. Using image sharpening it is possible to correct for the whole imaging pass in an adaptive optics set up. This could, for example, be achieved by imaging a USAF target image through the system. A search algorithm can then be used to maximise a sharpness metric, cancelling the inherent aberrations of the optics system, by driving the deformable mirror to its optimal shape. The deformable mirror can be set to its "nominally" focused position - even before adaptive correction begins.

Using an indirect wavefront sensing technique also offers the advantage that corrections to be applied are determined from the science camera as opposed to in direct wavefront sensing where a portion of the light needs to be directed to the wavefront sensor. This is important for imaging in low light levels and preserves the photon flux being imaged.

Transmissive correctors will be of particular benefit to such regimes and will add simplicity to the optical system. Such correctors could be placed in front of the science camera in existing adaptive optics systems and used to calibrate for non-common path errors. The correction system does not need a reference as it uses light from the object under observation in order to determine the necessary correction commands.

4.5 Implementation of Sharpness Maximisation

4.5.1 Sharpness Calculation

The sharpness calculation is generally a straightforward calculation of the intensities measured at the CCD camera. These values are read directly from the imaging camera buffer

and a sharpness calculation is made, which the algorithm then uses to determine the next search space step.

The search and control algorithms to perform these operations were written in C++. Functions were also written to control the application of voltages to the deformable mirror for the results described in this thesis and were integrated into the main optimisation loop. An open source function called "Image Basic"[77] was used to read the CCD camera TIFF files and process the image. Image Basic gives access to a host of values from the CCD camera and provides a range of processing options.

For the S_1 metric, $\sum I^2(x,y)/(\sum I(x,y))^2$, the program calculates the squared intensity of all the pixel values and sums these (top line of metric), then the denominator, that is, the squared sum of all the pixels, is a normalisation factor inline with conservation of energy. This is to ensure that the sharpness calculation is calibrated for each image as the amount of light from one image to the next may vary due to defocus modes on the mirror changing the focus of the image.

The Image Basic function is used to measure the maximum intensity of the TIFF file read from the camera buffer before the optimisation algorithm is run. This is used to determine the maximum exposure of an unaberrated image. The exposure level is set by varying the image exposure time. It is set at a value just below the maximum exposure level of the CCD camera, therefore an aberrated image will have a lower maximum exposure. This is done to use the maximum dynamic range of the CCD camera and means the optimisation algorithm will be less sensitive to noise.

A study was done to examine the shape of the search space for undersampled images[8]. It was found that undersampling by increasing factors increased the non-stationarity of the sharpness metrics with respect to sub-pixel shifts in the image. This simulation showed,

as expected, that sampling at the Nyquist sampling rate, or above, avoided shift variance of the image. Undersampling was also shown experimentally to increase the noise levels for a maximised image and this is shown in Chapter 7.

4.5.2 Comparison of (Power-law) Sharpness Metrics

Certain Muller and Buffington metrics have been proven to be maximised for extended objects whereas others are proven only for point sources, and others still are unproven to guarantee image improvement. The metrics defined by Feinup[72] use higher order metrics and are scene dependant as are many of the other image metrics mentioned in Chapter 4.3. The metrics suggested by Muller and Buffington are robust and general extended object metrics, and as such, are easy to calculate and a good choice of metric to proceed with. The primary objective is a "proof-of-principle", that in a robust and simplistic physical system, wavefront aberrations can be compensated for by maximising a sharpness metric using a wavefront sensor-less technique, and so the S_1 sharpness metrics are used as a general metric for extended objects.

4.5.3 Determination of Image Quality

It is shown in Chapter 7.1, from experiments made on point source images in this thesis, that an increased sharpness value corresponds to an increased Strehl ratio. For correction of a USAF target image, an increased resolution for a corrected image can be identified as corresponding to greater clarity in the sections of decreasing spatial frequency. It can be measured physically by profiling sections of the bar target for aberrated and unaberrated images. An increased contrast should be observed for corrected images which can be related to an increased MTF.

5. CORRECTION ALGORITHMS

Introduction

A substantial portion of both past and present research in the field of image processing has been dedicated to reconstructing or estimating the photometric parameters of an object from the observed data. These methods are *post*-processing techniques and correct an image already acquired to give estimates of the unaberrated image. Deconvolution of observed data is widely used as a post-processing technique[78, 79, 80, 81]. The Gerchberg-Saxton algorithm is commonly employed for correction in phase diversity techniques[37] and Doble presents an analysis of various search algorithms for sharpness metrics[68].

In this project sharpness maximisation requires an algorithm that provides updates on a "live" image. The image is updated by applying a new set of parameters to the search space. It is the task of the search algorithm to find the optimum set of parameters to maximise the sharpness in the search space. As the search space is large, in our case 255^{37} degrees of freedom, mainly non-systematic search routines will be employed. The search space is too large for a 37 actuator deformable mirror with 8-bit control, imaging onto a 12-bit camera, to be tested systematically (except for modal searching as it reduces the degrees of freedom), where systematically means trying every possible combination

of 255^{37} degrees of freedom, so a suitable multi-dimensional search algorithm needs to be used.

Generally, for complicated optimisation problems, non-systematic search routines, such as, simulated annealing, the Nelder-Mead simplex, stochastic gradient descent (SGD) or stochastic parallel gradient descent (SPGD), will give the best solution[68].

5.1 Algorithm Requirements

In the feedback system, the "new" corrective voltages to be applied to the wavefront corrector are derived by a search algorithm. As opposed to direct wavefront sensing where the voltages are determined in one step[1], image sharpness requires a suitable search algorithm to drive the corrector to its optimal shape. Ideally when using a search algorithm the global minimum/maximum is the final state i.e. the true minimum or maximum of the system. This is very difficult to find for systems having a large number of degrees of freedom, due to the probability of getting stuck at a local minimum (or maximum). Currently there are a wide range of search algorithms which can be classified as either systematic or non-systematic, a brief overview of which is presented below.

Systematic search algorithms, as the name would suggest, are thorough methods which test every possible combination or solution to the problem. They are, as a consequence, slow but effective and are better suited for smaller search spaces and problems with low degrees of freedom. For example, if a system had only a single aberration (such as de-

[1] In fact, to ensure stability of closed-loop control, the full correction is never applied in a single step. In practical control systems, the closed-loop bandwidth is typically 5-10 times smaller than the sampling speed of the wavefront sensor.

focus) it would be highly efficient to systematically vary the aberration (focus) until the global maximum of sharpness was found.

For more complex search spaces it is often better to use non-systematic search routines. These techniques search in a quasi-random nature. They generally apply a random solution and use information gained from this to guess the next possible iterative attempt. Of the various non-systematic search routines, some, such as the Nelder-Mead simplex algorithm, are not proven to determine the global maximum or minimum of a particular problem.

Many search problems do not contain a unique solution, that is a global maximum or minimum, but contain many local minima/maxima and some search algorithms can get stuck in one of the local maxima/minima. Algorithms such as simulated annealing are proven to determine the global maximum, but the speed of determination is compromised and so the choice of algorithm will be different depending on the confines of the solution requirements. In the following section I shall outline the operation of four search algorithms which I used to correct for induced aberrations.

An extremum (maximum or minimum point) can be either *global* (truly the highest or lowest function value) or *local* (the highest or lowest in a finite neighborhood and not on the boundary of that neighborhood). As can be seen from Figure 5.1, a region may contain local maxima/minima as well as a global maximum or minimum, and this is a difficulty search algorithms must overcome.

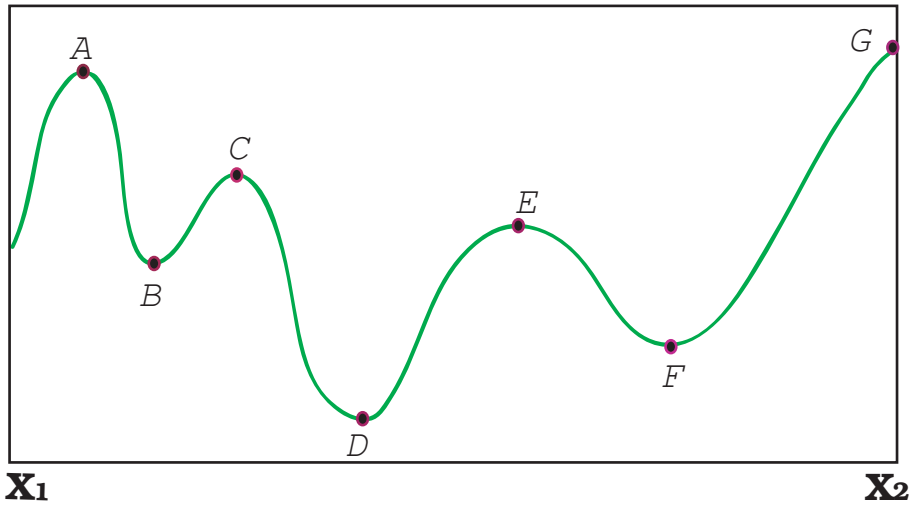


Fig. 5.1: Schematic of 2-D search space with local (A,B,C,E,F) and global (D,G) maxima and minima.

The search space may contain many local (or false) maxima, and ideally a search algorithm will find the global solution. In Figure 5.1, Points A, C and E are local but not global maxima, which lies at point G. Similarly points B and F are local but not global minima, which lies at D[82].

The graphical representation of Figure 5.1 is for a one-dimensional search space and its obvious that for a 37 actuator deformable mirror the search space will be much more complicated. In Chapter 7.2.3 the search space is examined in greater detail by applying mirror modes to an unaberrated image, with the view of proving further insight to the nature of the search space for various aberrations.

As maxima and minima are essentially the same thing, differing only in the sign of their definition, I shall refer to maxima from this point onwards, since the sharpness metric is maximised in the examples given.

5.2 Development of Correction Algorithms

An integrated system was necessary which evaluated sharpness and then controlled deformable mirror actuation in a closed-loop. Also, an interface was needed to QCapture, the Retiga CCD camera's software, in order to evaluate the "sharpness" of each image. The cameras exposure level needed to be checked before the search algorithm was run. This was necessary because, as an aberrated image is corrected, its expected that the intensity level of some pixels will increase as the image sharpens. If the initial maximum exposure level was set too close to the saturation point of a camera, the exposure level may saturate before the image is fully corrected, giving a "false" sharpness value and limiting the degree of correction achieved. To avoid this the exposure level of the camera was set just below the threshold value for an unaberrated image. Therefore, any aberrated image would have a lower intensity, meaning that the full dynamic range of the CCD camera could be utilised, reducing the effects of the read noise on the sharpness maximisation.

5.3 The Nelder-Mead Simplex Algorithm

As mentioned above, the simplex algorithm is not a systematic search algorithm in the sense that is does not search every possible node in the search space, but it does move in a methodical way through the search space. The simplex search method proposed by Nelder and Mead[83] is an algorithm that tries to minimise a scalar-valued nonlinear function of n real variables using only function values.

A simplex is a geometrical figure consisting of $n+1$ vertices, where n is the number of degrees of freedom. In two dimensions a simplex would be a triangle, in three, a tetrahedron and in this research, with a 37 actuator deformable mirror, a complex figure

with 38 vertices. Each vertex represents a set of mirror voltages. Initially, 38 $(n+1)$ random sets of voltages are generated and the corresponding sharpness value is measured for each vertex. Based on the initial evaluations the simplex attempts to adapt to the local landscape, with the aim of contracting to the global minimum. This is done by continuously replacing the simplex vertex with the highest (worst) function value by a new one, using the operations reflection, expansion and contraction as shown in Figure 5.2. If none of these steps are helpful, all vertices are contracted toward the point with the lowest (best) function value. This operation is called shrink.

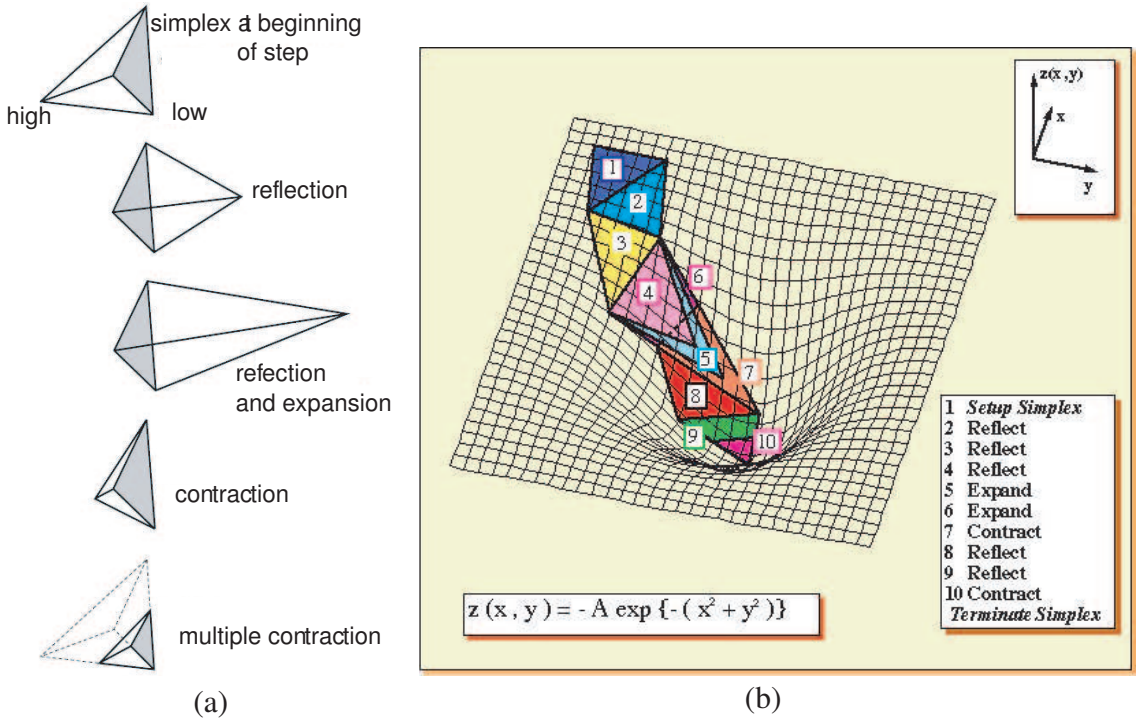


Fig. 5.2: Operations through which a simplex moves through search space.

For a system with three degrees of freedom the simplex is a tetrahedron, as seen in Figure 5.2a, and for two degrees of freedom a triangle, Figure 5.2b. The simplex opera-

tions through which a simplex moves through the search space can be seen in Figure 5.2a, and an example of a simplex moving to a minimum is shown in Figure 5.2b.

Four scalar constants have to be specified to define a Nelder-Mead method: coefficients of reflection (ρ), expansion (χ), contraction (γ), and shrinkage (σ). These parameters should satisfy,

$$\rho > 1, \quad \chi > 1, \quad 0 < \gamma < 1, \quad 0 < \sigma < 1.$$

The most common choices are,

$$\rho = 1, \ \chi = 2, \ \gamma = 0.5 \ \textit{and} \ \sigma = 0.5.$$

The simplex can be set to run for a certain number of iterations, or to stop when a termination criteria is met.

5.3.1 Algorithm Implementation

Firstly the function value is evaluated at the vertices of an initial simplex, which is randomly generated. Then the simplex enters the optimisation routine, where it runs through the following steps:

1. Order the vertices such that,

$$f(x_1) \leq f(x_2) \leq ... \leq f(x_{n+1}),$$

where $x_1, ..., x_{n+1}$ denote the coordinate vectors. The ordering of the vertices is repeated with each new iteration of the algorithm.

2. Compute the centroid of the n best points with,

$$\bar{x} = \sum_{i=1}^{n} x_i / n,$$

3. Reflect the "worst" point (which is x_{n+1}) through the centroid by,

$$x = \bar{x} + \rho\,(\bar{x} - x_{n+1})\,;$$

if $f(x_1) \leq f(x_r) < f(x_n)$, accept the reflected point and terminate the iteration.

4. If $f(x_r) < f(x_1)$, calculate the expansion point by,

$$x_e = \bar{x} + \chi\,(x_r - \bar{x})\,;$$

if $f(x_e) < f(x_r)$ accept x_e and terminate the iteration. Otherwise, if $f(x_e) \geq f(x_r)$, accept x_r and terminate.

5. If $f(x_n) \leq f(x_r) < f(x_{n+1})$, perform and outside contraction by,

$$x_{oc} = \bar{x} + \gamma(x_r - \bar{x})\,;$$

accept x_{oc} if $f(x_{oc}) \leq f(x_r)$, otherwise go to step 7 (perform shrink step).

6. If $f(x_r) \geq f(x_{n+1})$, perform and outside contraction; calculate

$$x_{ic} = \bar{x} - \gamma(\bar{x} - x_{n+1})\,;$$

accept this point if $f(x_{oc}) \leq f(x_{n+1})$, otherwise go to step 7 (perform shrink step).

7. Shrink the simplex toward the "best" point by,

$$v_i = x_1 + \sigma(x_i - x_1), \qquad i = 2, ..., n+1$$

and return to step 1.

The criteria for halting the procedure proposed by Nelder and Mead is,

$$\sqrt{(\sum_{i=1}^{n+1}(f(x_i) - \bar{f})^2 / n)\delta}\,; \qquad \overline{f} = \frac{1}{n+1}\sum_{i=1}^{n+1} f(x_i),$$

where δ is a small positive scalar.

Although Nelder and Mead published their paper in 1965, no theoretical results regarding convergence properties of the Nelder-Mead method in higher dimensions have yet been proven. Even finding any function in R^2 for which the algorithm would always converge to a minimum still remains an open problem[84]. Therefore, one cannot say with certainty that the maximum sharpness value reached by the simplex is the global maximum. However, I have found it produces good results in practice for the images produced in this thesis.

5.3.2 Simplex Algorithm Constraints

The original Nelder-Mead simplex algorithm was intended for minimization of functions with unbounded parameters and it had to be adapted for this project. The conditions for the execution of the simplex operations were modified for the maximisation of the sharpness

function. As the mirror operates over a discrete range of voltages, limits were set on the voltage values that could be applied.

The termination criteria proposed by Nelder and Mead were used to restart the optimisation procedure with a new simplex, where all vertices's except for the "best" point are randomly re-generated. This method is intended to prevent the simplex from shrinking to a false minimum.

5.4 Stochastic Gradient Descent Algorithms

Stochastic is synonymous with "random"; the word is of Greek origin and means "pertaining to chance" [85]. It is used to indicate that a particular subject is seen from point of view of randomness. Stochastic is often used as the complement of the word "deterministic", which means that random phenomena are not involved. Therefore, stochastic models are based on random trials, while deterministic models always produce the same output for a given starting condition. As such, *stochastic gradient descent* is an iterative multivariate optimisation search method. Stochastic gradient descent methods were first implemented on an adaptive optics system by Vorontsov[86]. Stochastic methods add randomness which help the algorithm to converge to the global maximum.

The basic premise of gradient descent algorithms is that, at each iteration of the algorithm, a new approximation is obtained by modifying the old one in the direction of the gradient of the discrepancy function. The standard gradient descent algorithm uses the true gradient to update the actuator voltages. The true gradient is usually the sum of the gradients caused by each individual trial perturbation. The voltages are adjusted by

the negative of the true gradient multiplied by a step size. Therefore, standard gradient descent requires one sweep through the training set before any voltages can be changed.

In stochastic parallel gradient descent the true gradient is approximated by the gradient of the cost function only evaluated on a single trial perturbation. The voltages are then adjusted by an amount proportional to this approximate gradient. Therefore, the parameters of the algorithm are updated after each perturbation. For a large number of degrees of freedom stochastic parallel gradient descent is much faster than general gradient descent and is more likely to find the global maximum.

Note that (stochastic) gradient descent methods work in spaces of any number of dimensions, even in infinite-dimensional ones. Two weaknesses of gradient descent methods are:

1. The algorithm can take many iterations to converge toward a maximum, if the curvature in different directions is very different.

2. Finding the optimal γ per step can be time-consuming. Conversely, using a fixed γ, can yield poor results. Conjugate gradient is often a better alternative and is described in Chapter 5.4.3.

The procedure used to implement each of the gradient descent algorithms is presented in the following Section.

5.4.1 Stochastic Gradient Descent Method

A single iteration cycle of the control voltages updated at the nth iterative steps consists of the following steps :

1. Initially a set of sharpness values $S^{(n)}$ depending on the application of a set of voltages $(u_1^{(n)}, u_2^n ..., u_{j,}^n ..., u_{37}^n)$ are measured. Each sharpness value corresponds to the application of a voltage to individual actuators.

2. Then random voltages, δu_j, of equal amplitude $\left|\delta u_j\right| = \alpha$, with a Bernoulli probability distribution $\rho\left[\delta u_j = \alpha\right] = \rho\left[\delta u_j = -\alpha\right] = 0.5$ are applied to each actuator (j^{th}) individually and the sharpness variation is measured:

$$\delta S = S\left(u_1, u_2, ...u_j + \delta u_j, ...u_{37}\right) - S\left(u_1, u_2, ...u_j, ...u_{37}\right).$$

The new voltages to be applied are then,

$$u_j^{(n+1)} = u_j^{(n)} + \gamma^{(n)}.\delta S_j.\delta u_j,$$

where, $\gamma^{(n)}$, is the gain, which can be either adapted through the iterations, or constant.

5.4.2 Stochastic Parallel Gradient Descent Method

One of the best ways to improve iteration speed is to decrease the number of sharpness calculations per iterations. Instead of applying random perturbations separately on each actuator and then measuring the corresponding sharpness variation, one can introduce a set of random voltages on all the actuators, and then measure the global sharpness variation. This method was first tested by Vorontsov[86, 87].

The principle of this method is similar to the non-parallel method and consists of the following steps :

1. First the sharpness value, $S(n)$, which corresponds to a set of voltages,

$$\left(u_1^{(n)}, u_2^{(n)}, ...u_j^{(n)}, ...u_{37}^{(n)}\right),$$

is measured.

2. Then, random voltages δu_j, of equal amplitudes $\left|\delta u_j\right| = \alpha$, and a Bernoulli probability distribution $\rho\left|\delta u_j = \alpha\right| = \rho\left|\delta u_j = -\alpha\right| = 0.5$ are applied simultaneously on the j^{th}actuator. These perturbations, applied in parallel, lead to a new set of voltages:

$$\left(u_1 + \delta u_1, u_2 + \delta u_2, ...u_j + \delta u_j..., u_{37} + \delta u_{37}\right).$$

3. Sharpness variation corresponding to the parallel perturbation is then calculated:

$$\delta S = S\left(u_1 + \delta u_1, u_2 + \delta u_2, ...u_j + \delta u_j..., u_{37} + \delta u_{37}\right) - S\left(u_1, u_2, ...u_j..., u_{37}\right).$$

4. The new voltages are then:

$$u_j^{(n+1)} = u_j^{(n)} + \gamma^{(n)}.\delta S.\delta u_j$$

where $\gamma^{(n)}$ is the gain, which can either be adapted through the iterations, or constant.

5.4.3 Conjugate Stochastic Parallel Gradient Descent Method

The conjugate gradient method is an algorithm for finding the nearest local minimum of a function of variables which presupposes that the gradient of the function can be computed. It uses conjugate directions instead of the local gradient for moving downhill, that is, to

a minimum. If the vicinity of the minimum has the shape of a long, narrow valley, the minimum is reached in far fewer steps than would be the case using the method of steepest descent[88, 89].

The conjugate gradient method is an effective method for symmetric positive definite systems. It is the oldest and best known non-stationary method. The method proceeds by generating vector sequences of iterates (i.e., successive approximations to the solution), residuals corresponding to the iterates, and search directions used in updating the iterates and residuals. Although the length of these sequences can become large, only a small number of vectors needs to be kept in memory. In every iteration of the method, two inner products are performed in order to compute update scalars that are defined to make the sequences satisfy certain orthogonality conditions. On a symmetric positive definite linear system these conditions imply that the distance to the true solution is minimized in some norm.

The principle of this method is again similar to that for the stochastic parallel gradient descent method and the procedure is the same for steps 1-3. The only difference is in the determination of the next voltage step, the formula of which is shown below. Conjugate gradient methods continue with the following step :

4. The new voltages are then given by:

$$u_j^{(n+1)} = \frac{\left\| \bigtriangledown S_j^{(n+1)} \right\|^2}{\left\| \bigtriangledown S_j^{(n)} \right\|^2} . u_j^{(n)} - \bigtriangledown S_j^{(n+1)}$$

where $\bigtriangledown S_j^{(n)} = \delta S^{(n)} . \delta u_j^{(n)}$ is the gradient on the j^{th} actuator on the n^{th} iteration.

5.5 Mirror Modal Correction

The previous search algorithms are methods whereby the program tries to find a combination of voltages to minimise the aberration and maximise the sharpness. To do this the search algorithm will search a space with 255^{37} degrees of freedom (for our experimental conditions). Each search routine will have a different approach and path to find the optimum voltage setting for the deformable mirror. Due to noise factors it is likely that the maximum sharpness that can be determined is insensitive to small perturbations of a given voltage combination, and further still, to a small change in the voltage setting for an individual actuator. Thus, while a given search routine may have approached the region of the maximum there may be a number of voltage combinations which provide the same or similar sharpness values. Therefore the search algorithm may spend excess time trying to improve the sharpness value and mirror shape unfruitfully.

The deformable mirror is of course limited by the number of physical shapes it can make. An algorithm may determine that in a particular step the voltage of one actuator may be $0V$ and the adjacent actuator be at $255V$. While commands will be sent to the mirror to emulate this shape there will not be a pure step between these two actuators. So, as the algorithm predicts that this should be the next step its not feasible for the mirror to accurately take this form. As the mirror is limited by the number of physical shapes it can make, a possible method of correcting for aberrations might be to cycle through the *modes* of the mirror, mirror modes often being mirror shapes close to the Zernike modes which model Zernike aberrations. In this way it will be possible to cycle through the mirror modes of the low order aberrations which have the biggest affect on image degradation.

In reality the number of viable mirror shapes is lower than 255^{37} degrees of freedom, as the eigenvalues associated with higher order modes are very small.

The mirror modes for the OKO deformable mirror to be used in this experiment have been measured and the necessary voltages required to apply a certain mode or given magnitude of it are known. The mirror modes were measured using a Fisba interferometer to determine the response of each actuator individually. Two measurements are required for each actuator - two voltage values are applied to the actuator and the response is used to determine the influence function of that actuator. This process is repeated for all actuators and the mirror modes derived. The mirror modes for the OKO deformable mirror can be seen in Figure 5.3[2]. The low order modes have some resemblance to Zernike modes but this is not the case as the mode number increases.

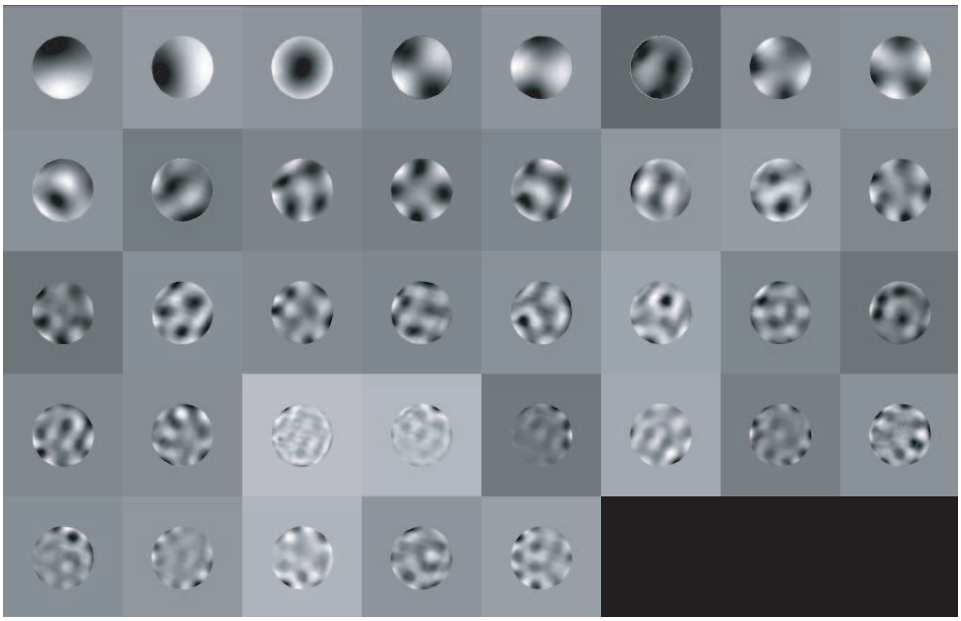


Fig. 5.3: Mirror modes of 15*mm* OKO membrane DM.

[2] Measurement of mirror modes carried out by Eugenie Dalimier in the Applied Optics Group, National University of Ireland, Galway, Ireland.

By searching the mirror mode space it is hoped the number of correcting combinations will be reduced. It may not be necessary to correct for higher order aberrations as the sharpness may be insensitive to these. Also it is important to note that the application of higher order modes can often have adverse effects on the correction being applied. This means that a limited number of modes should be used in correction using a modal approach. This further reduces the number of possible corrections, in effect reducing the search space. As can be seen from Figure 5.4, modes higher than mode 26 (approximately) have lower singular values and will contribute little to correction and so can be ignored.

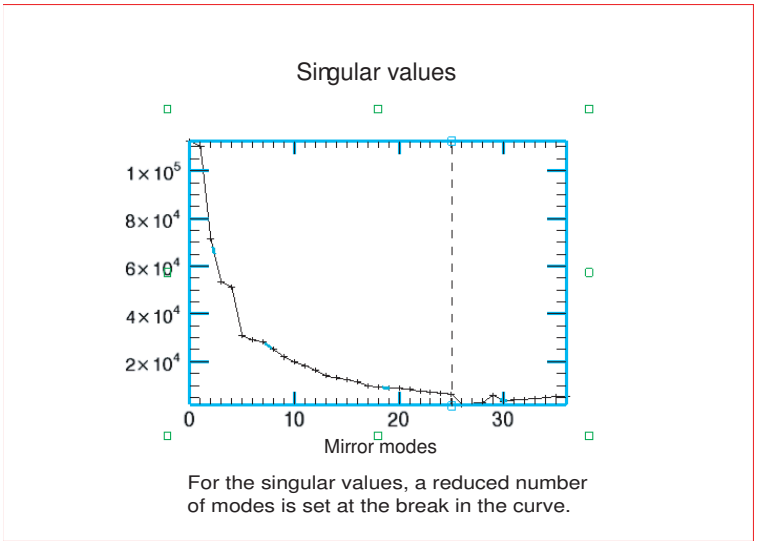


Fig. 5.4: Singular values for first 37 modes of OKO DM.

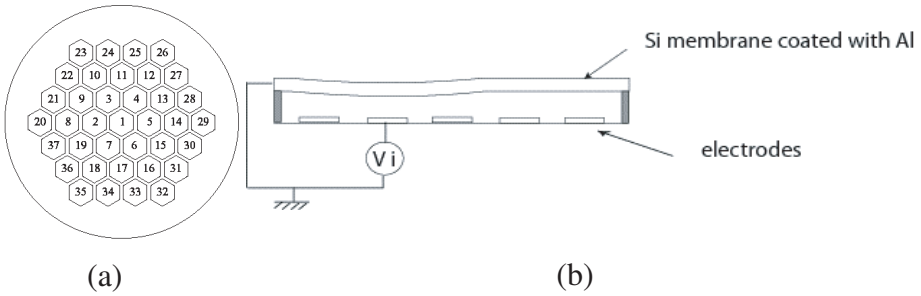
6. EXPERIMENTAL APPARATUS

Introduction

The aim of the experiments in the Chapters that follow is to experientially show that an aberrated image can be corrected by maximising its sharpness value. Initially an aberrated point source image is corrected using sharpness metrics and then the technique is applied to extended object images. In this Chapter the experimental conditions used to correct for an aberrated point source and extended object images are presented. A description is given of the key elements of the experimental apparatus such as the deformable mirror, CCD camera and the generation of Zernike aberrations using the Hamamatsu spatial light modulator. The operation of devices that are used is explained and other considerations such as the rate of sampling detailed.

6.1 Deformable Mirror Control

The OKO deformable mirror is a silicon micromachined membrane mirror, the shape of which is controlled by a set of 37 voltages, between 0 and 255 volts, through the generation of electrostatic forces. A voltage V applied on an actuator (or electrode) attracts the membrane of the mirror, proportional to V^2. The maximum stroke of the OKO deformable mirror used in this mode is $1.5\mu m$, over a $9.6mm$ pupil. This is a rather small stroke and



(a) (b)

Fig. 6.1: OKO DM actuator structure, Figure 6.1a, and operation, Figure 6.1b.

limits our ability to compensate for aberrations. The voltage was set at $230V$ as the maximum voltage, to prevent damage of the OKO deformable mirror. A housing unit was made to mount the deformable mirror in to guard against accidental contact and contamination with dust particles. The actuator structure and operation of an OKO deformable mirror can be seen in Figure 6.1.

The OKO deformable mirror is controlled using drivers developed in C++[1]. Commands are applied to the deformable mirror via a high voltage amplifier board which amplifies the voltage output of the PC to the necessary voltages in the range of 0-$255V$, to drive the mirror. It's recommended by the manufacturer not to drive the deformable mirror at its full voltage and so the maximum voltage was set to be $230V$. When aligning the optical system the deformable mirror is set in its bias position, which for a maximum voltage of $230V$, is $165V$. The bias is taken as the position in which the deformable mirror has equal stroke for the maximum and minimum voltages away from the bias. The value of $165V$ as the bias voltage comes from the squared nature of the deformable mirrors response to application of voltages, e.g. $230 \simeq \sqrt{165^2 + 165^2}$. The deformable mirrors squared response to voltages can be seen below,

[1] Written by Dr. Liz Daly of the Applied Optics Group, at the National University of Ireland, Galway.

$$ActuatorVoltages = \sqrt{biasV^2 + \triangle V^2},$$

where a change in voltage, $\triangle V^2$ of ± 165 will result in $0V$ or $230V$ being applied to the actuators.

With the bias voltage applied to the deformable mirror the CCD camera was manually translated to the position for which the extended image was focused. The sharpness of the image should be a maximum, and the focus position could be checked by running a program which applied voltage increments to all the actuators, from 0 to $230V$, and measured the corresponding sharpness. If the image is focused correctly for the bias position, the sharpness should peak at the bias voltage. If the CCD camera was not at the focus this was reflected in the corresponding position of the sharpness peak and the CCD camera could be repositioned. Results for alignment of the deformable mirror are presented in Chapter 7.2.

6.2 CCD Camera Noise Characteristics and Calibration

An integral component of any adaptive optics system is the science camera. Before a CCD camera is assembled, the manufacturer can use the nominal tolerances of the electronic components to estimate the gain to within some level of uncertainty. This calculation is based on resistor values used in the gain stage of the CCD readout electronics. However, since the actual resistance is subject to component tolerances, the gain of the assembled

camera may be quite different from this estimate. The actual gain can only be determined by actual performance in a gain calibration test.

For this reason the Retiga 1300 CCD camera to be used as the imaging camera was calibrated. The Retiga 1300 has 1280 x 1024 pixels with pixel size of $6.7\mu m$. It was calibrated using a Bentham Instruments Integrating Sphere, type ILlOENC. The image files were captured using QCapture, which is the software supplied with the camera. The calibration files are then processed in MATLAB. The Retiga 1300 camera was found to have a gain of 2.9 ± 0.1, where the gain is the analogue-to-digital conversion factor (ADU - analogue to digital unit). This gives the number of electrons required to register 1 digital unit. The gain value determined for the camera is a vital parameter as the resultant calculation of dark current, read-noise, and thus, dynamic range are based on this value.

A linearity test of the camera showed that it is linear until near saturation point. The linearity of the camera only deviates from the manufacturer's quoted linearity of gamma, approaching saturation. This is significant in that the maximum intensity value in the image should not approach the saturation point as it will affect the fidelity of the sharpness calculation. The signal-to-noise ratio of a CCD camera relates to its dynamic range and is given by the ratio of the maximum usable signal ($12^2 = 0 - 4095$ digital counts) to the read noise of the CCD, which is was measured to be $\simeq 5$ digital counts (or 15 e^-). The ratio is usually represented using a decibel scale (dB), or as a power of 2 ("bits"). The signal-to-noise ratio in dB is given by

$$SNR = 20\log\frac{4095}{5} = 58dB.$$

Therefore the signal-to-noise ratio was found to be $58dB$ slightly less than the $60dB$ specified, and thus it operates effectively as a 10-bit camera. The dark count was measured and found to be $0.26e-/pix/sec$. This value is higher than specified by the manufacturer $(0.15e-/pix/sec)$ but, is still quite small - especially for *ms* exposure times - and is insignificant for most purposes when a dark frame is subtracted from the image. These calibration results will need to be considered when determining the intensity values for evaluation of the sharpness metric. The reduced signal to noise ratio may have an effect on the calculation of the sharpness metric approaching a maximum in the search space.

6.3 Sampling of Point Source and Extended Object Set-ups

The sampling theorem is considered to have been articulated by Nyquist in 1928 and mathematically proven by Shannon in 1949[90, 91]. Some books use the term "Nyquist Sampling Theorem", and others use "Shannon Sampling Theorem". They are in fact the same theorem.

The sampling theorem states that for a limited bandwidth (band-limited) signal with maximum frequency f_{max}, the equally spaced sampling frequency, f_s, must be greater than twice the maximum frequency f_{max}, i.e.,

$$f_s > 2 \cdot f_{max},$$

in order for the signal be uniquely reconstructed without aliasing. The frequency $2 \cdot f_{max}$ is called the Nyquist sampling rate.

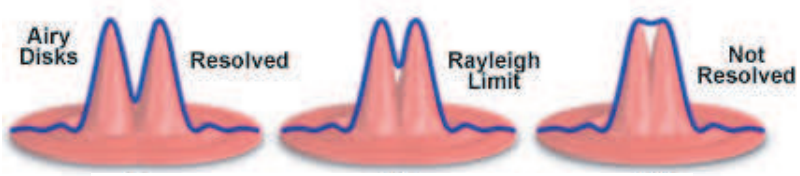


Fig. 6.2: Airy disk separation and Rayleigh criterion.

A function whose Fourier transform is zero for $|f| > f_{max}$ i.e. a bandlimited function, is fully specified by values spaced at intervals $\delta x \leq \frac{1}{2f_{max}}$, except for any harmonic term with zeros at the sampling points. In a diffraction-limited optical system, $f_{max} = \frac{1}{F\lambda}$ where $F = \frac{f}{d}$. This leads to, $\delta x \leq \frac{f\lambda}{2d}$. For the parameters used in this project gives $\delta x = 7.94\mu m$ and with $6.7\mu m$ CCD pixels its clear that the image is being sampled correctly.

In a two-dimensional image, two point sources are resolvable if their Airy disk diffraction patterns are distinct. According to the Rayleigh criterion, two closely spaced Airy disks are distinct if they are farther apart than the distance at which the principal maximum of one Airy disk coincides with the first minimum of the second Airy disk (as illustrated in Figure 6.2). If the point sources are of equal wavelength, then their Airy disks have the same diameter, and the Rayleigh criterion is then equal to the radius of one Airy disk, measured from its point of maximum intensity to the first ring of minimum intensity.

Experimentally, for CCD cameras, a sufficiently high pixel density (2X the highest spatial frequency in the image which is equivalent to at least 2 pixels across the full-width half maximum (FWHM) of a point source image) will guarantee a distortion-free representation of the image. For that reason, for optimum performance, it's required to

have at least two pixels across the core of a point source image, where the core is defined by the Full-Width Half Maximum (FWHM); i.e. the diameter at which the point source intensity falls to one-half its peak value. For practical purposes, the sampling really needs to be 2.5 to 3 pixels across the FWHM.

The extended source imaging system has the following parameters; a focal length (f) of 200mm, LED illumination light at $\lambda = 635nm$, an aperture diameter of $D = 9mm$ and 6.7μm CCD pixels. Using these values to determine the sampling rate through the radius of the Airy disk of a diffraction-limited lens of diameter D, and focal length f :

$$r = \frac{1.22f\lambda}{D},$$

gives the resolution r to be 17.2μm. This means that an image of a point source imaged by the extended object set-up will be sampled by > 2 pixels, i.e, $17.2/6.7 \simeq 2.6$ pixels, and thus is being sampled above the Nyquist rate.

The point source was imaged using a microscope objective and was greatly over sampled. For an aperture diameter, $D = 7mm$, an imaging lens with focal length, $f = 150mm$, and laser illumination light of $\lambda = 635nm$ being imaged onto the CCD camera results in the point source being imaged by >30 pixels.

6.4 Initial Point Source Optical Set-up

Initially sharpness maximisation was examined for the image of a point source object. The set-up used for the experiment is shown in Figure 6.3. The initial optical set-up design was made using Zemax[92], an optical design software package which allows optimisation and

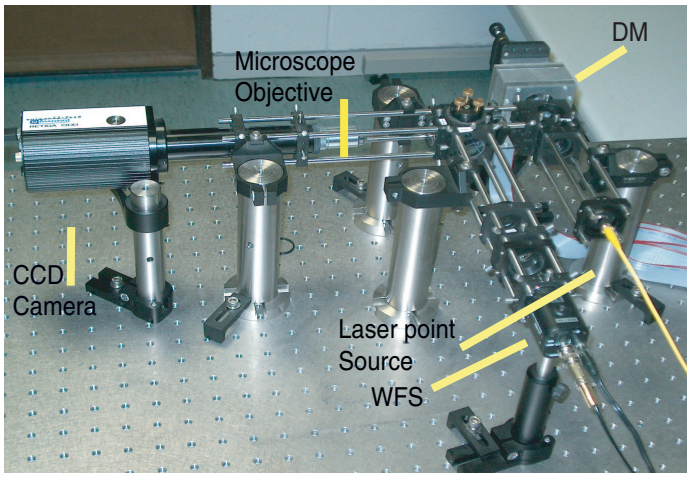


Fig. 6.3: Initial point source set-up.

evaluation of optical systems. The set-up is mounted in a Linos[93] microbench frame work, which aids alignment of the lens and optical components. Simplex algorithms were used initially to control a closed-loop operation, as good correction had been previously shown by Doble[68]. The sharpness value was re-evaluated and algorithm determined a new "correction" to be applied until a maximum sharpness value was determined.

A linearly polarised fibre optic He-Ne laser, wavelength $635nm$, is used to produce a point source to be imaged onto the CCD camera. The point source is placed at the focal length of the aperture lens $(f = 125mm)$ which collimates the beam. An aperture stop controls the input beam diameter and is set at $\sim 7mm$. This is done to operate the system with a high $F-number$ which helps reduce the aberrations of the system. It is important to note that while the OKO mirrors have a diameter of $15mm$, only the central 45-50 % of the mirror is useful for full stroke correction[2]. The collimated beam, produced by the first imaging lens, is then reflected off the OKO deformable mirror and back through a beam

[2] Based on simulation performed by Gordon Kennedy (Imperial College London).

splitter. The beam is then passed through a second beam splitter to produce two beams. One beam is imaged onto the OKO wavefront sensor and the other is focused onto the CCD camera.

The beam to be focused onto the CCD camera is brought to a focus and magnified with a Linos 20x microscope objective where it is finally imaged onto the Retiga CCD camera. The objective is used to increase the number of pixels which sample the Airy disk image. The other collimated beam is scaled to be re-imaged as a collimated beam onto the OKO wavefront sensor.

A neutral density filter is required to attenuate the intensity of the beam collimated onto the wavefront sensor. The intensity of the point source image on the CCD camera is controlled by the exposure time. For the wavefront sensor, saturation in the detector array will lead to problems detecting the centroids of the subapertures. Saturation of the point source image on the CCD will blur the point source image.

The wavefront sensor included is a Hartmann wavefront sensor which is essentially the same as a Shack-Hartmann, but has apertures instead of a lenslet array. The Hartmann mask is a hexagonal array of 91 subapertures with a mask-to-CCD distance of 12.16 mm and the detector CCD has 8.3 μm pixels. A lenslet array is useful when imaging in low light levels, but in the case of these initial experiments with a laser point source, this is not a factor. Its important to note that the wavefront sensor was not used for correction but was solely used as a check, to verify that increased sharpness did indeed relate to lower wavefront error.

6.4.1 Point Source Aberrating Phase Plates

Initially phase plates created by the National Centre for Laser Applications (NCLA) at NUI Galway[3], were used to aberrate the point source image. These consisted of phase structures etched onto Foturan[94], a photosensitive glass. The aberrations generated were weak static aberrations of approximately $1/4\lambda$ astigmatism and defocus. An interferometric image of the refractive index change for a defocus plate can be seen below in Figure 6.4.

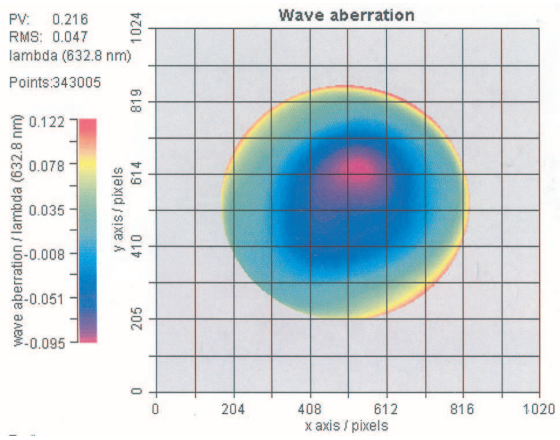


Fig. 6.4: Example of aberration induced by defocus plate.

6.4.2 Point Source Experimental Process

With the experimental set-up described above, the power law metrics described by Muller and Buffington were examined. Each metric was evaluated for 50 and 100 iterations of the simplex algorithm and their sharpness values and improvement in strehl ratio compared.

[3] Phase plates made by Dr Liz Daly, Applied Optics Group, National University of Ireland, Galway, Galway, Ireland.

6.5 Extended Object Apparatus

A narrow-band Luxeon Star LED, wavelength 635*nm*, was used to illuminate a 2-dimensional extended object. Transparent photographic slides taken of university buildings were used as extended objects, and to make any correction more readily discernible, a USAF target image was also used. The extended object is imaged onto a Retiga CCD camera via a Hamamatsu spatial light modulator (described in greater detail in the following section) and a deformable mirror. The spatial light modulator generates Zernike aberrations which are created in MATLAB. The aberrations are fed to the spatial light modulator via the green component of a RGB cable which is connected to a second graphics card port on the control computer.

The resolution of the Zernike images created in MATLAB are created to match the pixel dimensions of the spatial light modulator which has 1024 x 768 pixels. The spatial light modulator can create a 2π phase change and the phase can be wrapped in the MATLAB program to place stronger aberrations in the system.

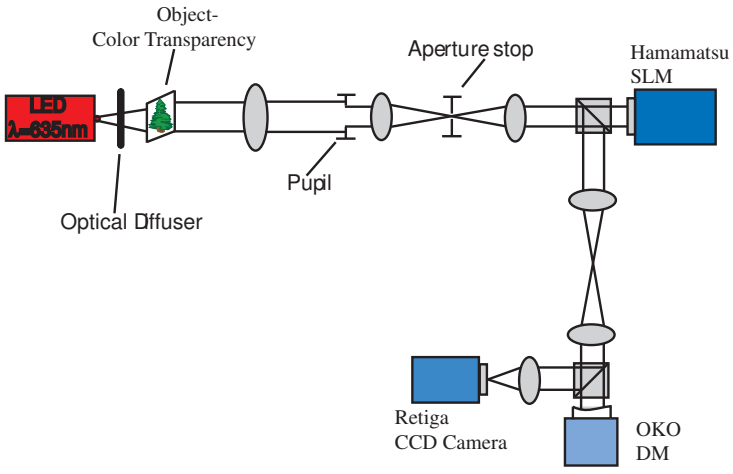


Fig. 6.5: Extended object experimental set-up.

The deformable mirror is a 37-channel OKO MEMS mirror[52] with a diameter of $15mm$ and has a frequency range of up to $1KHz$. The device can be used for fast dynamic correction of low-order optical aberrations such as defocus, astigmatism, coma, etc. The mirror is operated over a range of $0-220V$. The experimental set-up can be seen in Figure 6.5.

6.5.1 Liquid Crystal Phase Modulators

Introduction

The Hamamatsu X8267 Programmable Phase Modulator (PPM) consists of an optically addressed liquid crystal spatial light modulator optically coupled to an intensity light modulator[98]. The former is composed of a continuous layer of parallel-aligned liquid crystals, a light blocking layer and a photoresistive layer sandwiched between transparent electrodes. When an image of the desired phase map is displayed in the intensity light modulator, the spatial distribution of light over the photoresistive material causes the voltage between the electrodes to change locally, varying the effective refractive index and, consequently, inducing the associated wavefront distortion. Both elements are coupled through an optical system that removes pixelation. Therefore, the PPM is to a large extent free from the diffraction patterns associated with the pixel structure typical in electrically-addressed liquid crystal devices. For this particular model (X8267), phase maps can be generated with XGA resolution (1024×768 pixels) over the $20\times20mm$ liquid crystal active area that correspond to the central 768×768 pixels.

Hamamatsu spatial light modulator modulators have been shown previously to be suitable as an aberration generator. Artal has used the Hamamatsu spatial light modulator to

Fig. 6.6: Hamamatsu SLM used to generate Zernike aberrations.

create Zernike aberrations and subsequently compensate for aberrations in an artificial eye[62]. An image of the Hamamatsu spatial light modulator can be seen in Figure 6.6.

6.5.2 Application of Zernike Aberrations

Zernike aberrations are introduced into the system using a Hamamatsu spatial light modulator. Zernike aberrations are modeled in MATLAB and an image of the aberration is displayed on a second computer monitor. A dual graphics card installed in the computer allows control of two PC monitors from one desktop. This facilitates the operation of the search algorithm, MATLAB program, etc, on one monitor and the display of an image of a given Zernike aberration on the other monitor. The RGB signal from the dual graphics card that is normally connected to the monitor displaying the aberration is instead connected to the spatial light modulator. A pattern of this image is then placed on the spatial light modulator. The image of the Zernike aberration was linearised between 0-255, corresponding to a 8-bit signal. This corresponds to the range of values of the RBG signal.

The red and blue components of the image are zeroed as the spatial light modulator only uses the green component from the RGB cable to apply the phase pattern.

6.5.3 Measurement of Zernike Aberrations Generated by the SLM

A Fisba interferometer[99] was used to calibrate the Hamamatsu spatial light modulator, details of which are given below. Zernikes created in MATLAB were imposed on the spatial light modulator and the phase change created was measured. An example of Zernike images that are applied to the spatial light modulator in order to introduce the aberration can be seen in Figure 6.7.

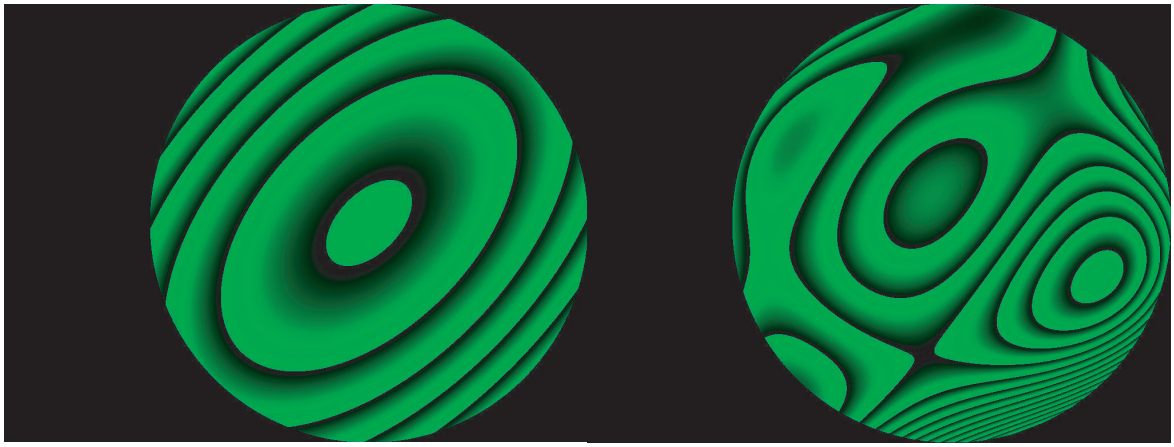


Fig. 6.7: Zernike images which are placed on SLM to generate aberrations.

The strength of each Zernike was increased and the corresponding phase change recorded. This produced a "look-up" table that indicates which wrapping factor was required to create a given magnitude of a Zernike aberration. The MATLAB images shown in Figure 6.7 are of a combination of 0.8λ P-V defocus and astigmatism in Figure 6.7a, and a random combination of several Zernikes in Figure 6.7b.

6.5.4 Calibration of Hamamatsu Spatial Light Modulator

To measure the amount of phase change introduced on the spatial light modulator the polarisation of the Fisba read laser needs to be aligned parallel to the director molecules of the spatial light modulator. For normal Fisba testing of optical surfaces, no polarisation sensitivity is present so this makes measurements a reasonably simple process. However, the spatial light modulator is highly polarisation sensitive for phase modulation. This meant that great care needed to be taken when aligning the Fisba read laser, parallel to the direction of the director molecules on the spatial light modulator. This was done in a simple set-up with a linear polariser and a power meter. The Fisba was mounted in a holder which facilitated the rotation of the Fisba laser to determine the angle it should be fixed at. With the Fisba in rotation a linear polariser was placed between the Fisba and a power meter. The polariser was set to pass vertically polarised light and the Fisba was rotated through 360^0 while the power meter was monitored. The Fisba interferometer was fixed at the position which produced the minimum intensity reading on the power meter. This position then corresponded to the horizontal components of the Fisba read laser being blocked by the linear polarisation pass of the fixed polariser. In this position the Fisba polarisation was aligned horizontally with the director of the spatial light modulator. This then produced a phase modulation of the read light when a phase pattern was applied to the spatial light modulator.

6.5.5 Aberration Generation

As mentioned above a MATLAB program was written which could create an image of Zernike aberrations, or any combination thereof. These aberrations could be “dialled

up" and placed onto the spatial light modulator. The phase change created by placing Zernike patterns on the Hamamatsu spatial light modulator were measured using a Fisba interferometer as described in Chapter 6.3.3, and two typical aberrations can be seen in Figure 6.8.

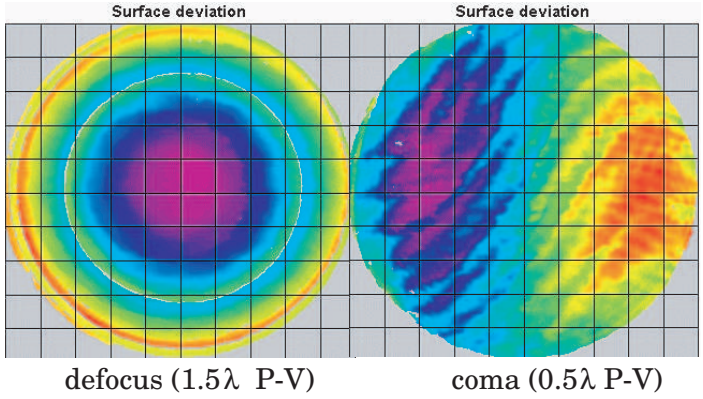


Fig. 6.8: Example of Zernike aberrations generated on SLM.

6.6 Extended Source Set-up

The main extended object used for sharpness maximisation was a USAF target, as it enabled easy detection of improvement in image quality. A Luxeon LED was used as the illuminating light because a purely coherent source such as a laser would introduce speckle. A problem encountered in using the LED as a source is that the LED contains some structure which was being superimposed onto the target image on the CCD camera. To avoid this superposition of the images a diffuser was placed 20*mm* after the LED in the imaging path. This meant that the LED needed to placed away from the focal length of the lens that imaged the extended object slide, so that LEDs structure would not be imaged onto the CCD camera.

The pupil of the system was conjugated with the spatial light modulator and the deformable mirror so that aberrations introduced by the spatial light modulator could be corrected for by the deformable mirror. The imaging beam was set to be $9mm$ because as mentioned earlier although the deformable mirror has a diameter of $15mm$ the stroke of the edge actuators are restricted. The Zernikes placed on the spatial light modulator where then created over the same beam size so the aberrations created by the SLM match the correcting area of the deformable mirror.

6.7 Experimental Process for Extended Objects

Various combinations of Zernike aberrations were generated and each algorithm was run to correct for the aberrations. The algorithms were tested to maximise the S_1 sharpness metric for a combination aberrations. The algorithms were also used to test the systems limit of correction. This was determined by placing an aberration in the system, running the algorithm to correct for the aberration and subsequently increasing the aberration until the deformable mirror/algorithm could not find any corrected solution. In this sense the limitation of correction could be a limitation of either the algorithm or the physical limits of the deformable mirror deformation.

6.7.1 Search Space Analysis - Cycling Through Mirror Modes

One approach to examine the search space - varying the aberrations in the system using the spatial light modulator - has already been mentioned. Another method is to vary the mirror shape for an originally focused and aberration free image. This can be done by placing on the mirror, individual mirror modes. By cycling through various modes, the

profile of the sharpness metric can be measured. Its expected that higher order modes will have less effect on the sharpness value of the focused image. In this way a profile can be built up of each of the mirror modes and their corresponding effect on the sharpness values, and as such, will give information on the search space for given modes.

6.7.2 Mirror Modal Search Method

Image correction using a modal approach reduces the number of search nodes in the search space and may offer a quicker form of correction. The mirror modes are applied sequentially, stepping through the full magnitude of each. The algorithm applied the full magnitude of the first mode and measured the corresponding sharpness then applied increments. The algorithm moved through each mode from the maximum magnitude of the mode, applying decreasing steps while measuring the corresponding sharpness values. The magnitude of each mode that gives the maximum sharpness value is retained before moving onto the application of the next mode. The algorithm cycles through all the modes determining what magnitude of each mode produced the highest sharpness value.

The algorithm then returns to mode #1 and applies the mode magnitude that produced the highest sharpness value, again records the sharpness value. Then the magnitude of mode #2 that produces the highest sharpness value is added and sharpness again calculated. If the new sharpness value is higher the addition of mode #2 is kept and the algorithm moves on to add mode #3, otherwise the addition of mode #2 is canceled and the addition of mode #3 is applied to mode #1. A flow chart representing this process is shown in Figure 6.9. As the modes are not orthogonal, the algorithm then cycles through all increments of each mode, only keeping the addition of mode increments which improve the

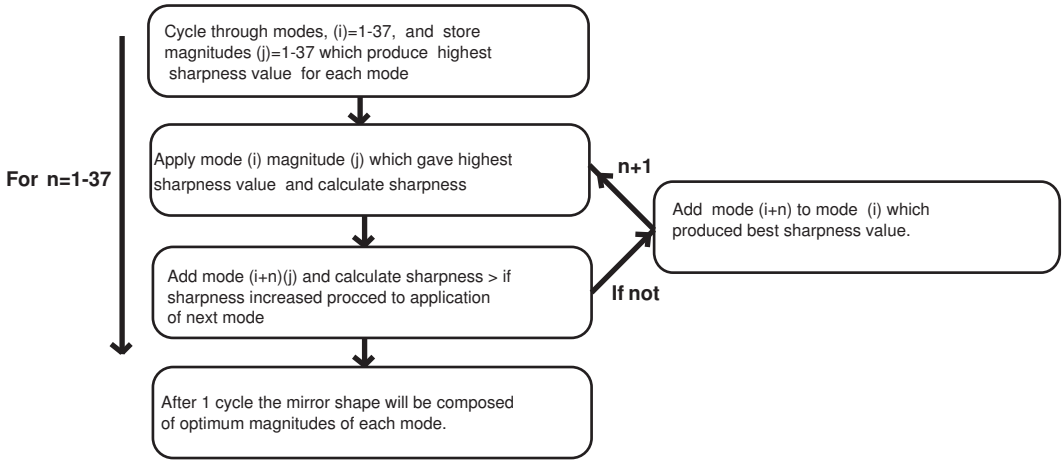


Fig. 6.9: Flow chart which shows approach used to cycle through the mirror modes.

sharpness value. The final set of mirror voltages are then applied and this is taken to be the best correction achieved by the modal search method.

6.7.3 Speed Limitations of Correction

The speed of correction achieved for both the point source and extended object was largely limited by the CCD camera's frame rate. The stated readout speed of the Retiga CCD camera was 11 full frames per second. For this experiment the key task was the "proof-of-principle", that experientially, an aberrated extended image could be corrected by maximising the sharpness value and so the emphasis was not optimising the speed of correction. It was found that for a minimum exposure time of 40μs the algorithm could capture and readout 7 full frames per second. Clearly then the speed of the CCD was a major limitation to the speed of correction that could be achieved. This is in stark comparison to other conventional CCD cameras which can operate at rates of 1000's of frames per second.

7. RESULTS

7.1 Results for a Point Source Object

The simplex algorithm was used to correct for aberrations introduced into the image of the point source in the experimental arrangement described in Section 6.4. The image was aberrated by phase plates with 1/4λ(P-V) defocus and astigmatism. The algorithm was run for I^2, I^3, I^4 for 50 and 100 iterations. For each power-law sharpness definition, 100 iterations of the algorithm were run. It was seen that for each case the algorithm converged to a maximum in less than 50 iterations so future corrections were limited to 50 iterations. An example of the convergence of I^3 in less than 50 iterations can be seen below in Figure 7.1.

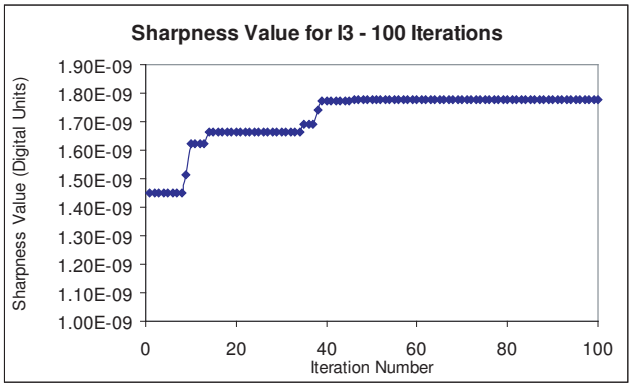


Fig. 7.1: Convergence of I^3 sharpness metric to a maximum.

Each power-law metric achieved a similar increase in the comparative Strehl ratio of the image. The comparative Strehl was taken to be factor by which the intensity density of the central maximum increased when the correction algorithm was applied. A mask was applied around the maximum intensity region and the average intensity was taken. This average for the corrected image was divided by the uncorrected average to get the factor of improvement.

The algorithm ran the fastest for the I^3 sharpness metric as the algorithm was less likely to perform shrink operations, whereby, each vertex of the simplex (38) is shrunk and re-evaluated within one iteration. Often 50 iterations would take tens of seconds for the I^3 sharpness metric, compared to tens to hundreds of seconds for I^2 and I^4. However, the limiting factor for the correction time for each algorithm was not connected to the nature of the sharpness metrics. The limiting factor was the processing of the image TIFF files and the camera readout speed as described in Chapter 6.

As mentioned before, the algorithm would reach a maximum sharpness value in less than 50 iterations and this can be seen (Figure 7.2) in a sample of some algorithm trials for I^2. Each iteration reaches a "maximum" value which corresponds to a restored image. The difference in the final sharpness value may be due to variations in the shot noise, resulting in a higher sharpness for a given CCD image or slight variations in the final deformable mirror voltage set.

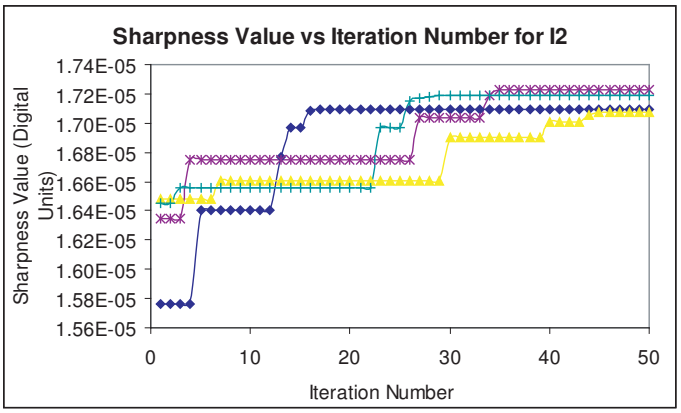


Fig. 7.2: Sharpness value vs Iteration number for I^2sharpness metric.

The increase in the Strehl of the aberrated point source achieved using I^4 as a sharpness definition and with 50 iterations can be seen below in Figure 7.3. As can be seen this metric improves the image intensity significantly.

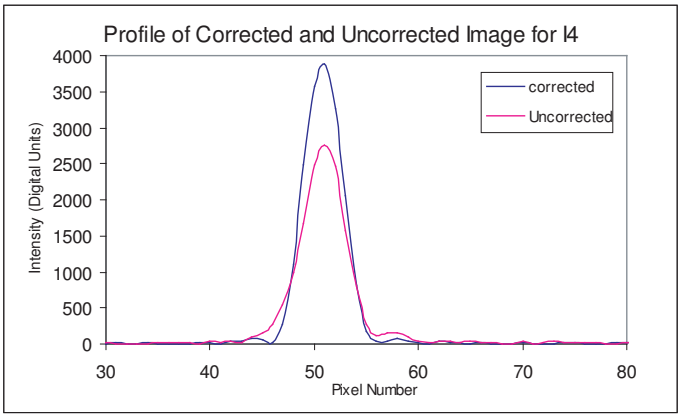


Fig. 7.3: Image profile of corrected and uncorrected image for I^4sharpness metric.

The correction achieved by the three power-law metrics is compared in Figure 7.4. It can be seen that each metric produces a similar result after 50 iterations.

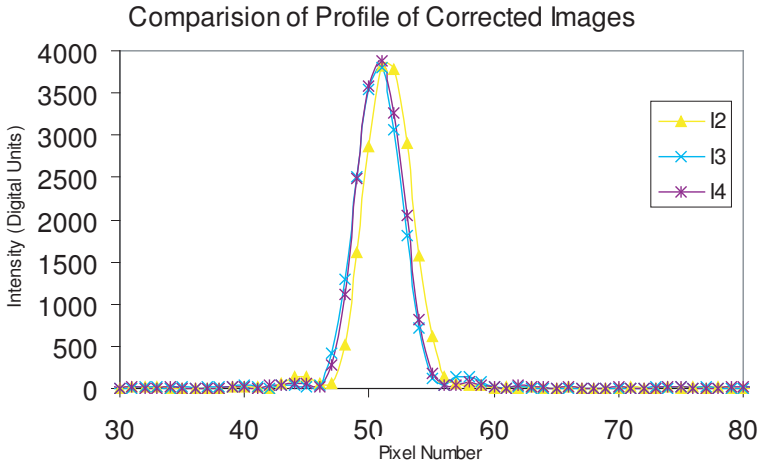
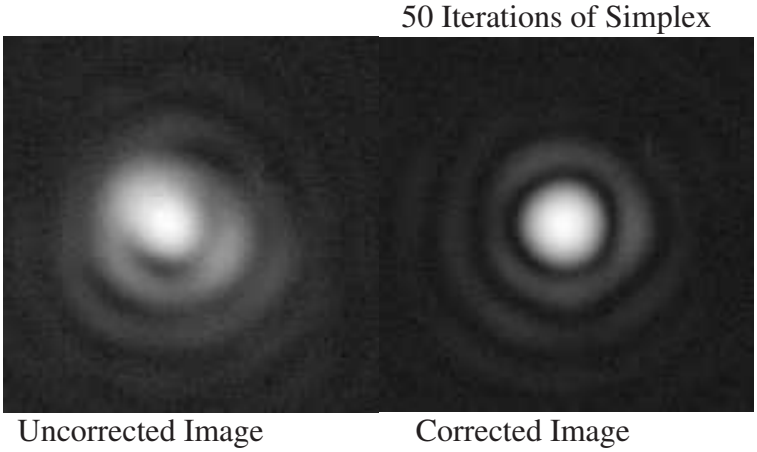


Fig. 7.4: Comparison of corrected image profile for I^2,$I^3$and I^4.

The maximum intensity of each image was increased for each image metric. The I^2 sharpness metric increased the central pixel intensity by a factor of 1.36 $\pm$ 0.07, for metric I^3 this factor was 1.39 $\pm$ 0.07 and for metric I^4, 1.30 $\pm$ 0.09, where the error values correspond to the range of the maximum intensity values found for each metric over ten correction trials.

A comparison can be seen in Figure 7.5 between an aberrated image and the correction achieved after 50 iterations of the simplex algorithm with the power-law I^2 as the sharpness metric.

It should be noted that although the amount of aberration in the image was small, the image is corrected for the induced aberration as well as the inherent aberrations. This situation is similar to that encountered in conventional adaptive optics for non-common path errors.



Results acheived after 50 iterations of the simplex algorithm. As can be seen from the images the Airy rings of the point source have been restored. Images are displayed on a log scale.

Fig. 7.5: Aberrated and corrected point source image.

7.2 Initial Experiments for Image of Extended Object

The alignment of the system was quite a sensitive process as the deformable mirror had such a limited stroke. This meant that the mirror had to be very well aligned to maximise the amount of usable stroke. The results of this can be seen in Chapter 7.2.1. The effect on the position of the peak sharpness value for an unaberrated image is discussed in Chapter 7.2.2 and the search space corresponding to the application of mirror modes can be seen in Chapter 7.2.3.

7.2.1 Bias Positioning of Deformable Mirror

As discussed in Chapter 6.1 its important to have the deformable mirror set at its bias position in order for the mirror to have equal stroke in each direction away from the bias position. This is necessary to enable the deformable mirror to faithfully produce large amplitude mirror modes such as astigmatism. The extended object is then focused

on the CCD camera for the deformable mirror's bias position. As explained in Chapter 6.1 the accuracy of this positioning is reflected by the position of the peak sharpness value when all of the deformable mirror actuators are changed incrementally from 0-250V, whilst measuring the sharpness for each voltage increment. It should be noted that, as the bias voltage is set a 165V, the corresponding maximum voltage should be $\simeq 230V\left(\sqrt{165^2+165^2} \simeq 230\right)$. The results for this can be seen in Figure 7.6

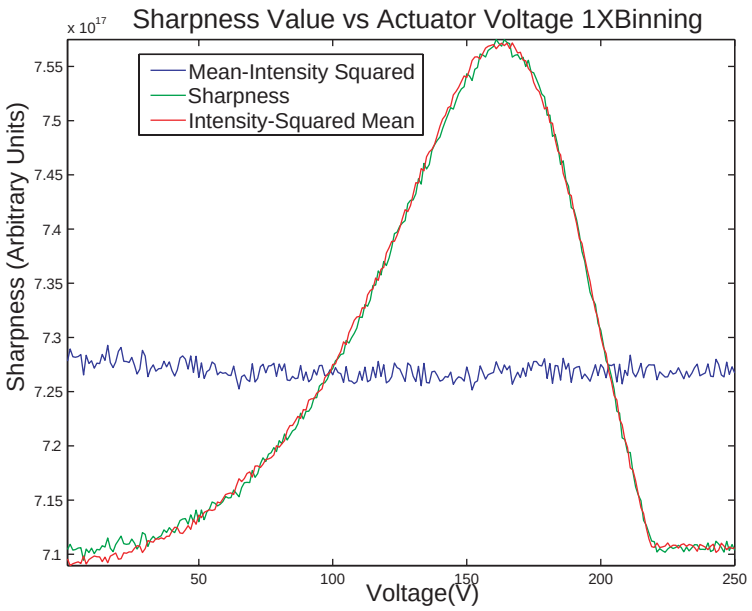


Fig. 7.6: Sharpness value as biased DM is moved in increments from 0-250 V, for 1x1 binning.

As can be seen from Figure 7.6 the sharpness metric is well defined when defocused through the application of the same voltage to all actuators sequentially. The sharpness is seen to be a maximum in the region of the bias voltage and fall sharply for voltages away from the bias value. It can be seen that the sharpness value corresponding to 0V and >220V on all actuators is the same. This indicates that the image is well focused for

the deformable mirror position as the image is equally defocused for the minimum and maximum voltage applied to the mirror. It can also be seen that for voltages higher than $220V$, the sharpness remains constant. This indicates that the mirror is limited and cannot deform at voltages higher that $220V$. Thus the search algorithm will be confined to apply voltages between 0 and $220V$ to reduce the search space.

The sharpness is normalised and plotted with the two component values of the sharpness calculation, the "mean of the intensity-squared", $\sum I^2(x,y)$, and the "mean-intensity squared", $(\sum I)^2$. It can be seen that the "mean-intensity squared" value is constant as expected, this is because it essentially represents the amount of energy in the image, which should not vary; only its distribution should. The "mean of the intensity-squared" value mirrors the sharpness value closely, as expected, and this indicated good alignment of the system, that is the system is telecentric. That the set up is telecentric is important as the energy in the image would vary and affect the fidelity of the sharpness value. This result was taken initially to ensure the mirror was focused correctly at its bias position before any further analysis was made.

The sharpness measurement in Figure 7.6 was made with the CCD camera operating in a $1x1$ binning regime, for which the image is sampled above the Nyquist frequency. To determine the effect undersampling of the image the same measurement was repeated with the CCD camera operating in a $4x4$ binning regime, which would then undersample the image, with < 1 pixels across the airy disk of a point source. The results can be seen in Figure 7.7. It can be seen that the sharpness curve is not as well defined. This would decrease the ability of a search algorithm to locate a maximum sharpness value when correcting for aberrated images, and indicates the sensitivity of the sharpness function to undersampling.

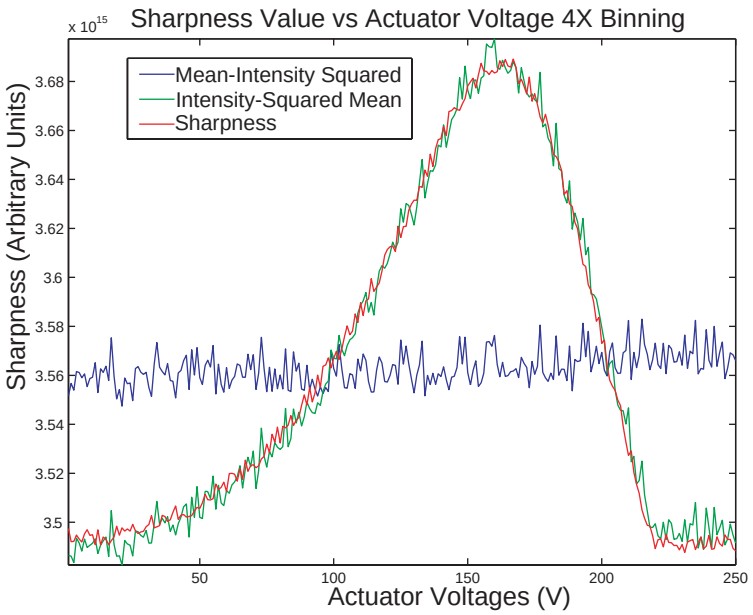


Fig. 7.7: Sharpness value vs. actuator voltage for 4*x*4 binning.

7.2.2 Application of Zernike Defocus Aberration to Extended Object

As shown previously the sharpness metric for an unaberrated extended object peaks at the bias value of the deformable mirror. This result is expected as the image is at its focus position on the CCD camera. The application of Zernike aberrations should therefore, change the position of the peak sharpness value. The target image was then aberrated with various amounts of Zernike defocus via the Hamamatsu SLM, to examine how the sharpness value varies. The results due to the addition of defocus aberration can be seen below in Figure 7.8.

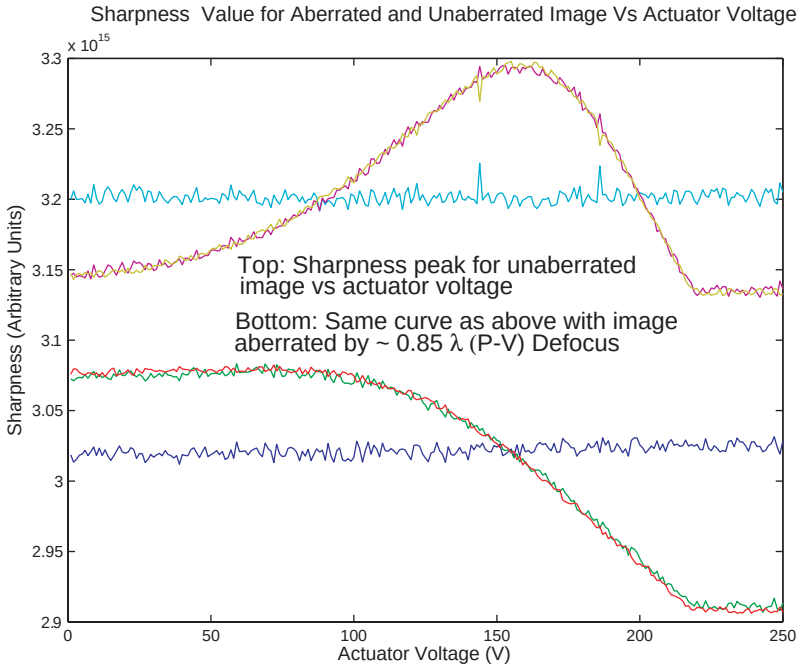


Fig. 7.8: Addition of defocus Zernike to unaberrated image focused at the DM bias position.

The sharpness value for the aberrated image is normalised and plotted with the sharpness evolution for an unaberrated image, as the deformable mirror moves from 0-250V. The mean intensity for the aberrated image is slightly decreased because, instead of acting purely as a reflector, the SLM slightly attenuates the light intensity in its phase modulation mode. Approximately $0.85\lambda(P-V)$ defocus aberration was placed on the spatial light modulator to aberrate the image. As can be seen the sharpness no longer peaks at the bias position. The position of the maximum sharpness has shifted towards $0V$ on all actuators, as the image is now in focus for this voltage. The image becomes increasingly defocused as the mirror moves from 0-250V, thus the sharpness decreases. Figure 7.9 displays the sharpness evolution for an image aberrated by 3.4 $\lambda(P-V)$ defocus. Here the sharpness falls more steeply as the mirror moves from 0-250V. The introduction of

defocus aberration has the effect of shifting the peak sharpness value as the image focal position changes.

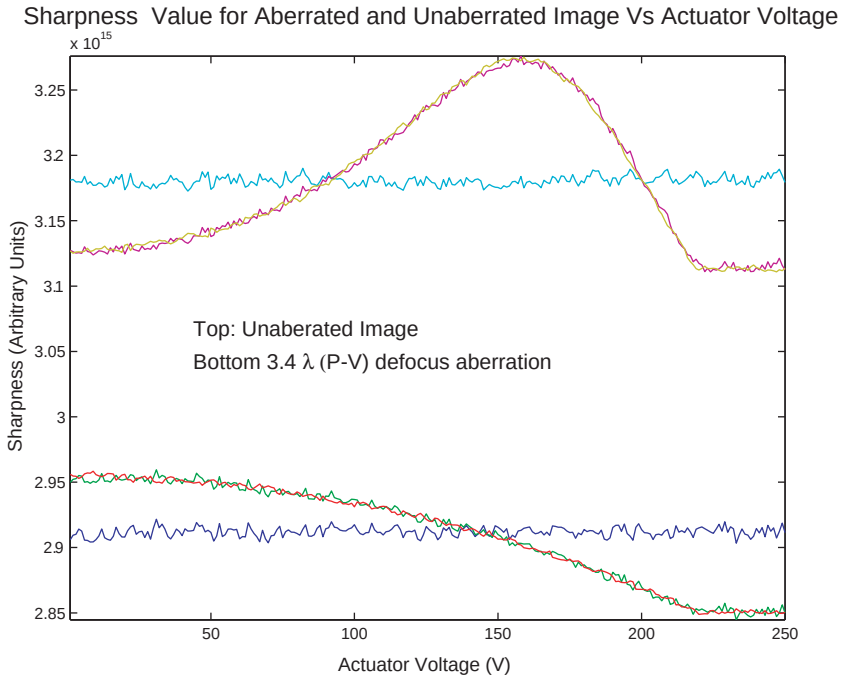


Fig. 7.9: Sharpness evolution of image aberrated by $3.4\lambda(P-V)$ defocus aberration.

It is clear from Figures 7.8 and 7.9 that the membrane deformable mirror used had a very limited stroke and can be only expected to correct relatively small aberrations. This was a major limitation in the project and means that we can simply obtain "proof-of-principle" results.

7.2.3 Search Space Analysis Through variation of the Mirror Modes

To gain insight into the search space and the response of the sharpness metric to the application of mirror modes, the sharpness of an unaberrated image is measured as mirror

modes are applied to the mirror. The application of mirror modes has the effect of introducing a mirror modal form of aberration. As the mirror modes have similarities to Zernike modes, at least for low orders, this indicates the nature of the search space with respect to various aberrations, and in return how the application of mirror modes may increase the sharpness for an aberrated image. The normalised sharpness metric's response, corresponding to the application of mirror modes, is also plotted with the sharpness values generated by driving the mirror from 0-250V, as described previously. This compares the degree to which the mirror modes can alter the sharpness values, with respect to the maximum change in the sharpness value generated by the application of the maximum and minimum voltage, i.e., the largest possible sharpness changes the mirror can produce.

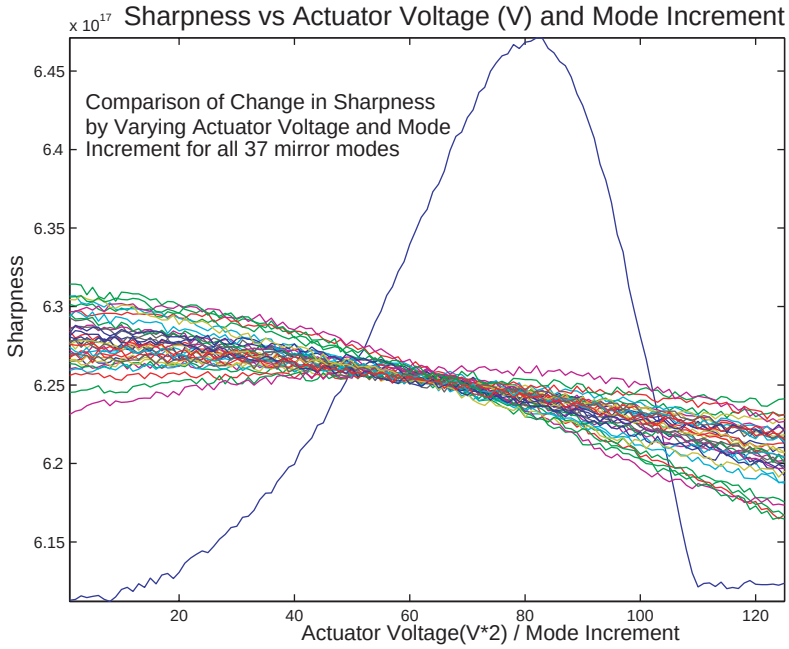


Fig. 7.10: Sharpness change induced by applying full mirror modes in 125 increments.

Figure 7.10 shows a plot of the mirror modes, each of which is incremented in steps of 1/125th of the full modal value, against the sharpness value measured by moving the mirror through minimum to maximum in voltage steps of $2V$. As can be seen, the sharpness variation of the full range of the mirror modes does not have as dramatic effect on the sharpness as that of driving all the actuators from the minimum to maximum voltage. This shows that the mirror modes have a lower influence on the sharpness value. This is to be expected as, to produce the mirror modes the deformable mirror actuators assume a variety of voltages and adjacent actuators can limit each others stroke, that is, the mirror actuators are not orthogonal.

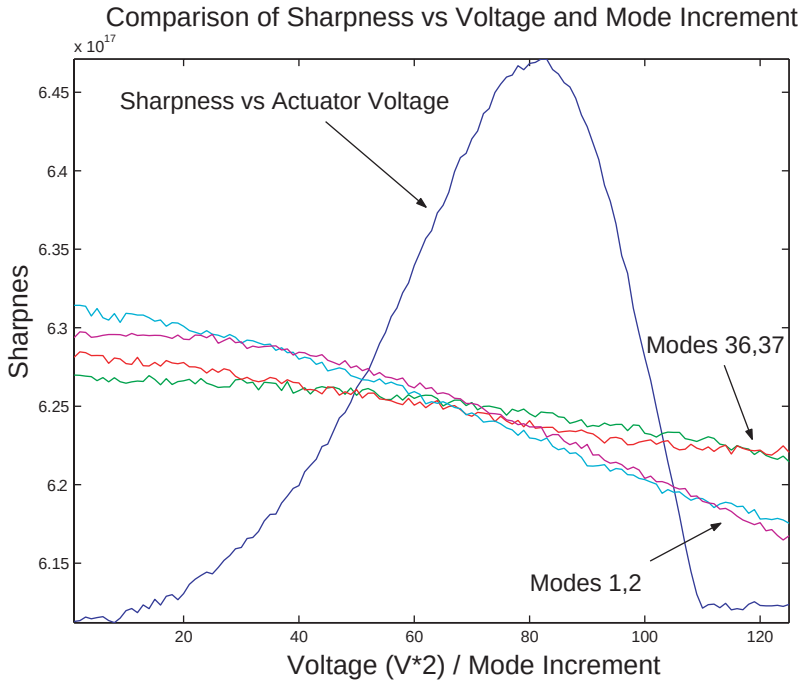


Fig. 7.11: Comparison of effect on sharpness value of high and low order mirror modes.

The sharpness corresponding to the application of low order mirror modes 1 and 2, and higher order modes 36 and 37 can be seen in Figure 7.11. As expected, it can be seen that the higher order modes have a lesser effect on the sharpness value. This can be seen from Figure 5.4 which shows that higher order modes have lower singular values.

The reduced effect on the sharpness of the higher order modes reflects that the Zernike aberrations, on which they are based, reduce image quality to a lesser extent. This suggested that for modal search routines, correction for low order modes will yield the biggest increase in sharpness of an aberrated image.

The sharpness change corresponding to the application of defocus and both astigmatism modes is shown below in Figure 7.12, and the sharpness change resulting from the introduction of Zernike aberrations produced by the spatial light modulator are shown in Figure 7.13.

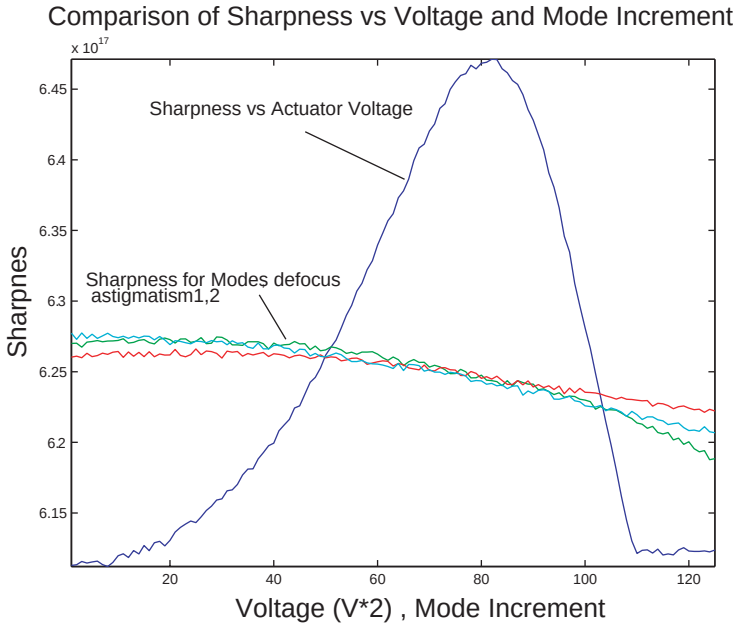


Fig. 7.12: Sharpness variation with increasing defocus and both astigmatism mirror modes.

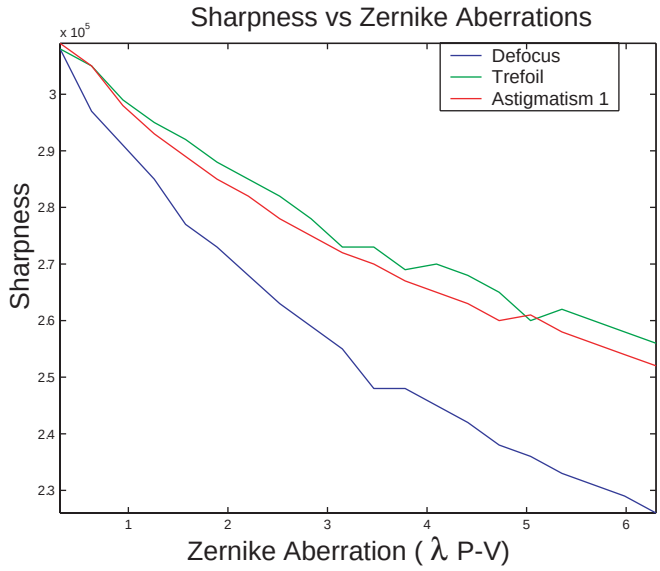


Fig. 7.13: Sharpness variation with increasing Zernike aberration applied via SLM.

It can be seen from Figure 7.12, as was seen in Figure 7.11 that the mirror modes have a limited effect on the sharpness compared to the application of voltages on all actuators. The sharpness for astigmatism 1,2 and defocus have a similar response and decrease the sharpness in smooth manner. This suggests that the sharpness might be a smooth function for extended objects, with a global maximum. The variance of the sharpness with respect to the application of mirror modes may differ from that due to the application of Zernike aberrations so the sharpness was measured while increasing magnitudes of Zernike aberrations were applied using the SLM. As can be seen from Figure 7.13 the sharpness decreases smoothly as the Zernike aberration is increased. The Zernike aberrations were increased in steps of $0.32\lambda(P-V)$. The aberrations are increased from 0 to $6.3\lambda(P-V)$.

Aberrations created by the spatial light modulator can be wrapped to increase the amount of phase change generated, and as can be seen in Figure 7.13, the spatial light modulator can create aberrations greater than what can be corrected by the deformable mirror. As such the deformable mirror will limit the correction, aberrations created by the SLM should be matched to what the DM can correct.

7.3 Extended Object Correction Results

7.3.1 Performance of Search Algorithms

A USAF target image is used as the extended object to be imaged. This image was aberrated by placing various Zernike aberrations on the Hamamatsu spatial light modulator to distort the image. The correction algorithms used to correct the aberrations are as follows, simplex algorithm, stochastic gradient descent, stochastic parallel gradient descent and a modal algorithm which attempts to correct for the aberrations by applying varying amounts of deformable mirror modes. Simulated annealing and conjugate gradient descent algorithms were also tested but their use was not pursued as no meaningful correction results were achieved.

The correction achieved by each algorithm for various Zernike aberrations are presented below. I shall present the results of correction of the algorithms beginning with the slowest - the Nelder-Mead simplex.

7.3.2 Nelder-Mead Simplex Algorithm Results

The Nelder-Mead simplex algorithm was shown in Chapter 7.1 to provide good correction to an aberrated point source image and was the first algorithm to be used to correct for images of extended objects. Figure 7.14 shows the correction achieved for an image aberrated by ~$1.3\lambda(P-V)$ defocus aberration. Although the original unaberrated image is well focused, it is not, of course, aberration free and so the algorithm is not purely correcting for defocus, but also for the inherent aberrations of the system. The algorithm will, in effect, try to correct for all the aberrations of the system.

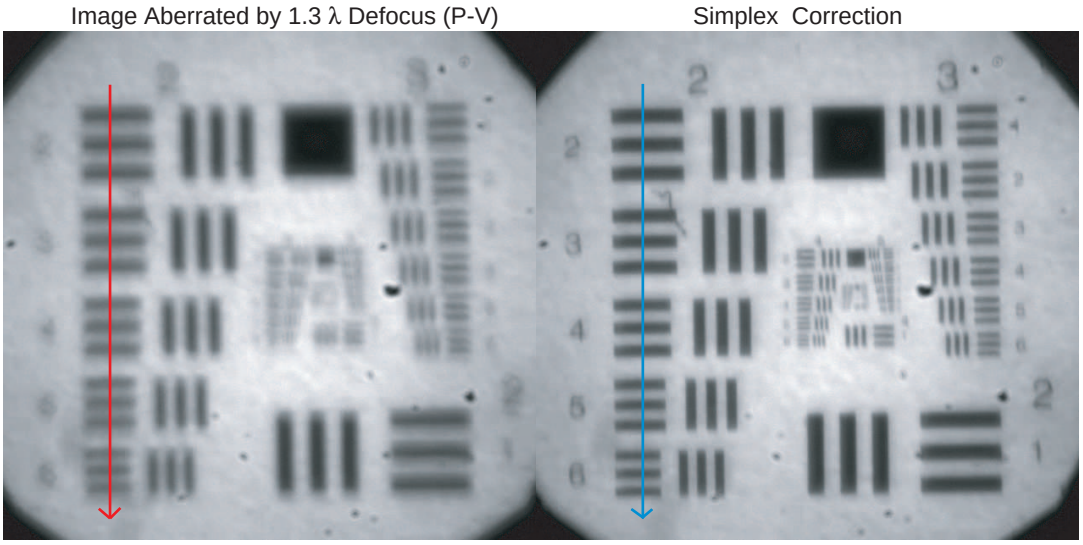


Fig. 7.14: Correction of defocus aberration by simplex Algorithm.

Figure 7.14 shows the image before correction on the left and the correction achieved after 500 iterations on the simplex. Figure 7.15 shows line profiles taken through one of the bar-target sections, as highlighted in the image, for the corrected and uncorrected images. In Figure 7.14 a red arrow is superimposed on the area which a profile is taken

through for the aberrated image, and similarly, a blue arrow through the corrected image. The profiles through corrected and uncorrected images will be subsequently represented by these colours in line profile plots. The profiles through the two images show the increase in contrast with correction. The resolution of the image has also increased which can be seen by the increased clarity of the central bar-targets.

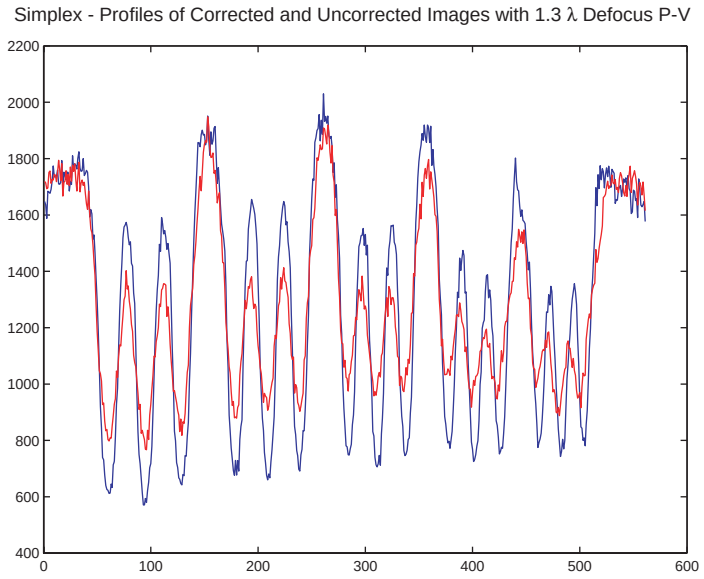


Fig. 7.15: Profiles taken through corrected and uncorrected images for simplex algorithm.

While the simplex can corrected well for aberrated images, it is a slow algorithm. The typical time required to perform 500 iterations is ~ 40 minutes. This is because the simplex performs a high number of shrink operations for complicated optimisation problems. Each shrink operation involves 38 sharpness measurements and the application of 38 set of mirror voltages, so in this sense it does not strictly perform 500 iterations, but many more depending on the search problem. The simplex algorithm was also used

to optimise small amounts of astigmatism, and combinations of astigmatism and defocus, and corrected the image to a similar degree as above.

7.3.3 Stochastic Gradient Descent

The operation of the Stochastic Gradient Descent algorithm is, of course, similar to that of the Stochastic *Parallel* Gradient Descent algorithm, but where as the stochastic parallel gradient descent algorithm performs two sharpness measurements to update the voltages, the stochastic gradient descent algorithm performs 37 sharpness measurements, and is as a result, slower. The stochastic gradient descent algorithm was shown to correct well for a degree of aberrations, and results for the correction of $1.3\lambda(P-V)$ defocus aberration, and the corresponding line profiles taken through the aberrated and corrected images are shown in Figures 7.16 and 7.17 respectively. Figure 7.18 shows the correction achieved for a combination of ~ $1\lambda(P-V)$ astigmatism and ~$1.3\lambda(P-V)$ defocus aberration.

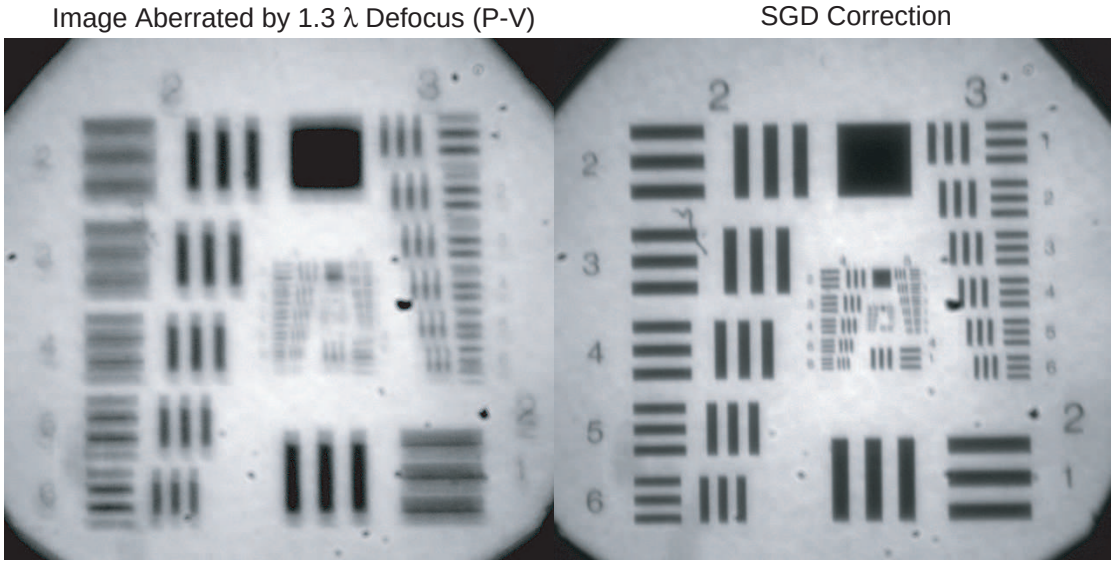


Fig. 7.16: SGD correction of defocus aberration.

Fig. 7.17: Profile of corrected and uncorrected images for SGD algorithm

The stochastic gradient descent algorithm is shown to correct well for images aberrated by the introduction of defocus - and other aberrations present in small amounts, as mentioned previously. The central target image is much more visible due to the increase resolution of the system with the reduction of aberrations. The profiles through the aberrated and corrected image show a large increase in the image contrast. In Figure 7.18 the image correction is not as good as that of the image aberrated by defocus, but given that the mirror has a maximum stroke of $1.5\mu m$ over the image, this is probably a limitation of the deformable mirror. The magnitude of the combined defocus and astigmatism aberration appears to be the limit of aberration that the mirror can correct for. Also, as can be seen from Figure 7.12, the application of the full mirror mode for astigmatism has a relatively small effect on the sharpness, as compared to the variation of all actuators through

the full range of voltages. So, while the maximum stroke of the mirror is $1.5\mu m$, its unlikely that mirror will be able to correct fully for ~$1\lambda(P-V)$ astigmatism alone, besides the combination with $1.3\lambda(P-V)$ defocus.

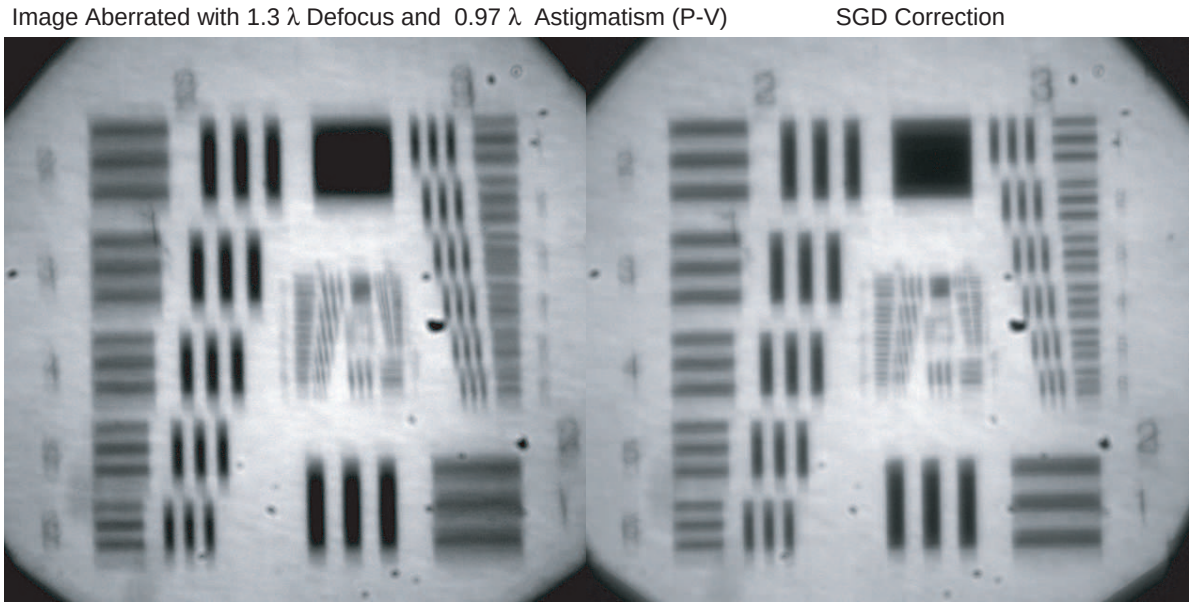


Fig. 7.18: SGD correction for combination of astigmatism and defocus aberration.

The stochastic gradient descent algorithm typically takes ~300 seconds to correct for aberrations, yielding results on a par, or slightly better than the simplex algorithm. The algorithm time of 300 seconds corresponds to approximately 55 iterations, where each iteration requires 38 sharpness measurement which are taken at a rate of approximately 7 per second.

7.3.4 Mirror Modal Correction

The mirror modal approach takes a fixed correction time - depending on the number of increments of each mode it applies - as it moves through a set number of iterations. The correction achieved using different numbers of mode increments was examined and set at 20 increments per mode. If applied to a mirror with a larger stroke, more increments of each mode would enhance the sensitivity of correction, but no significant improvement in correction was found for mode increments greater than 20. The correction for an image aberrated by ~$1.3\lambda(P-V)$ is shown below in Figure 7.19 and the profiles taken through the corrected and uncorrected images can be seen in Figure 7.20.

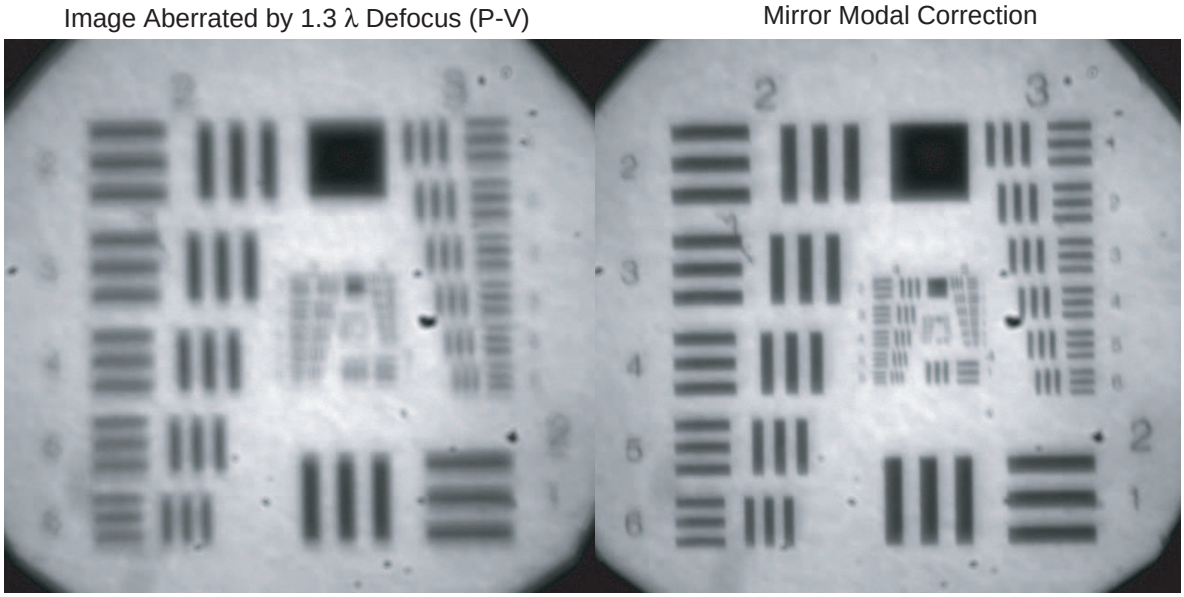


Fig. 7.19: Mirror modal correction for defocused image.

Figure 7.19 shows that the modal search method provides results similar to the previous methods with increased resolution in the central target area. Profiles through the

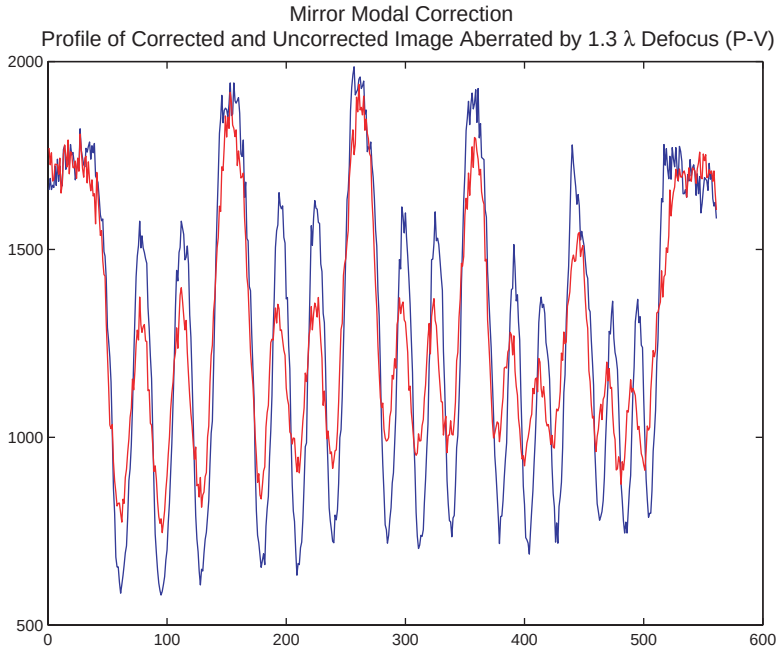


Fig. 7.20: Profiles of corrected and uncorrected images for modal algorithm.

corrected and uncorrected images again, show a large increase in image contrast after correction. As a more quantitative analysis of the increase in image quality, the modulation of the corrected and uncorrected images were measured and it was found that, over the range of frequencies in Figure 7.20, the modulation has been increased by $> 55\%$ on average. The modal search algorithm takes a set time of 220 seconds for each correction which requires 1517 sharpness measurements as the modes are cycled through.

7.3.5 Stochastic Parallel Gradient Descent

For the previous search algorithms the corrected images shown were mostly for defocus and astigmatism aberrations and each algorithm performs similarly, only varying in the speed of correction. The stochastic parallel gradient descent algorithm was used to correct for random combinations of Zernikes for $j = 4 - 18$, as well as for defocus and astigmatism combinations.

Figure 7.21 shows aberrated and corrected images for increasing magnitudes of defocus aberration. The defocus aberration applied, ranges from ~ 0.8 to 2.7 $\lambda(P-V)$. In the top two images, the correction determined by the stochastic parallel gradient descent algorithm is very good and corrected well for the defocus aberration. In the image below, as defocus aberration increases, we see that the algorithm corrected the aberration, but to a lesser extent. It can be seen in the bottom two images, where the aberration applied is approximately $2.7\lambda(P-V)$, the aberration is too strong for the deformable mirror to fully correct. The image is improved but only to the limit of the deformable mirrors actuator stroke. In this sense, the deformable mirror has run out of stroke and cannot correct the image further.

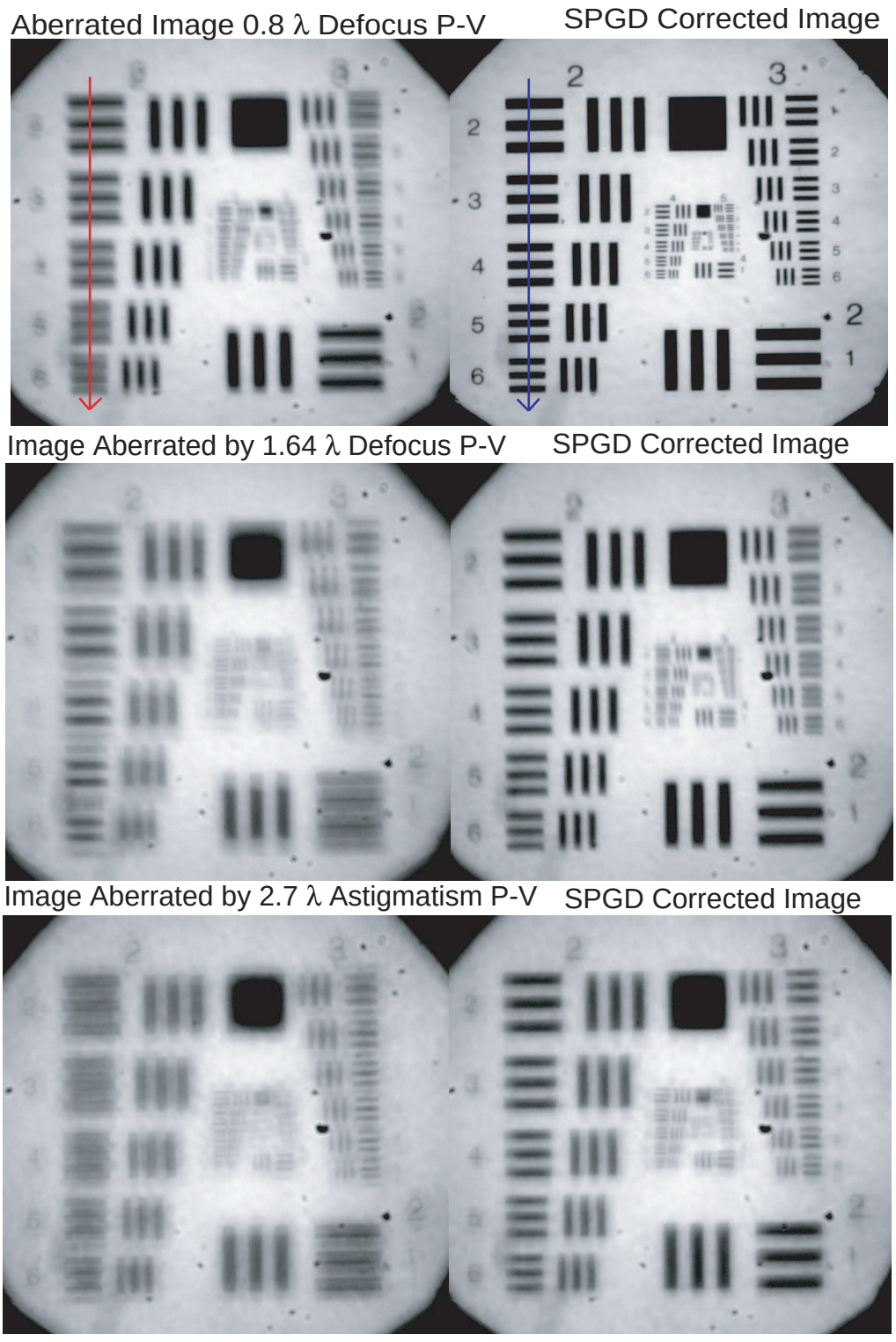


Fig. 7.21: Correction results of SPGD algorithm as aberration increases.

As mentioned earlier the original image is at its best focal position on the mirror, but its likely the image suffers due to aberrations from the systems optics and limits of the optical set-up. The stochastic parallel gradient descent algorithm was used to try and correct for the small amount of inherent aberration in the system. The results of this correction can be seen in Figure 7.22.

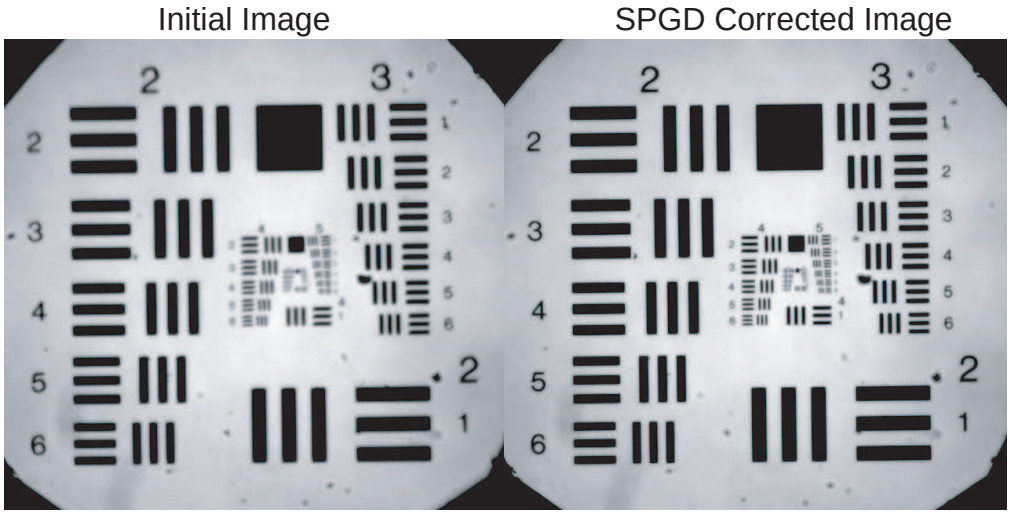


Fig. 7.22: SPGD correction for system aberrations.

Such is the sensitivity of the sharpness metric and stochastic parallel gradient descent algorithm, that small amounts of inherent aberration can be measured and compensated. In this way the stochastic parallel gradient descent algorithm may be used remove inherent optical aberrations from an existing experimental set-up. The deformable mirror voltages applied to cancel the system aberrations could then be used as the bias values for further correction.

Stochastic parallel gradient descent correction for images aberrated by ~1 and 2.7λ$(P-V)$ astigmatism can be seen in Figure 7.23 and the correction for a combination of defocus and astigmatism can be seen in Figure 7.24

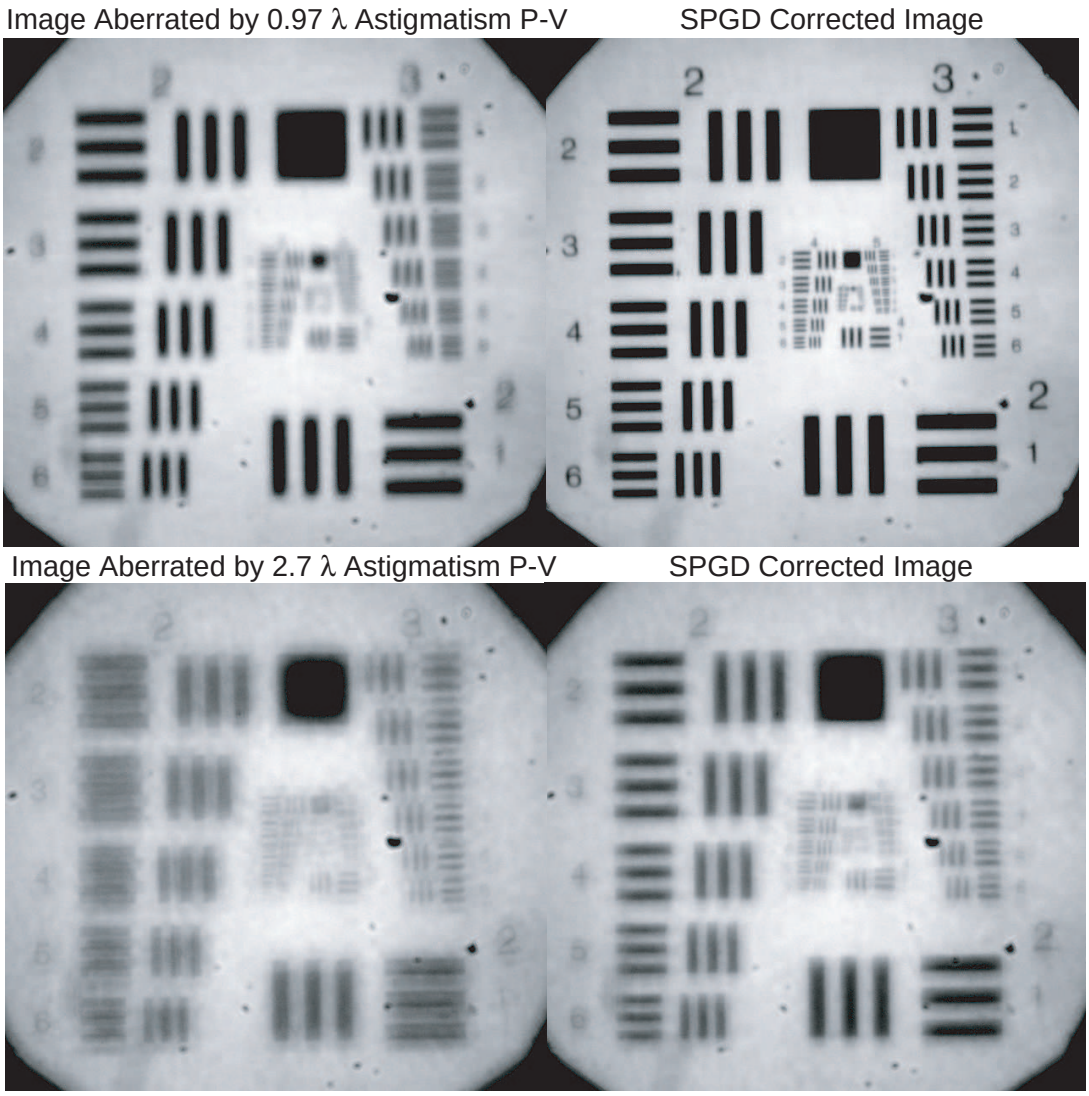


Fig. 7.23: SPGD correction of images aberrated by increasing magnitudes of astigmatism.

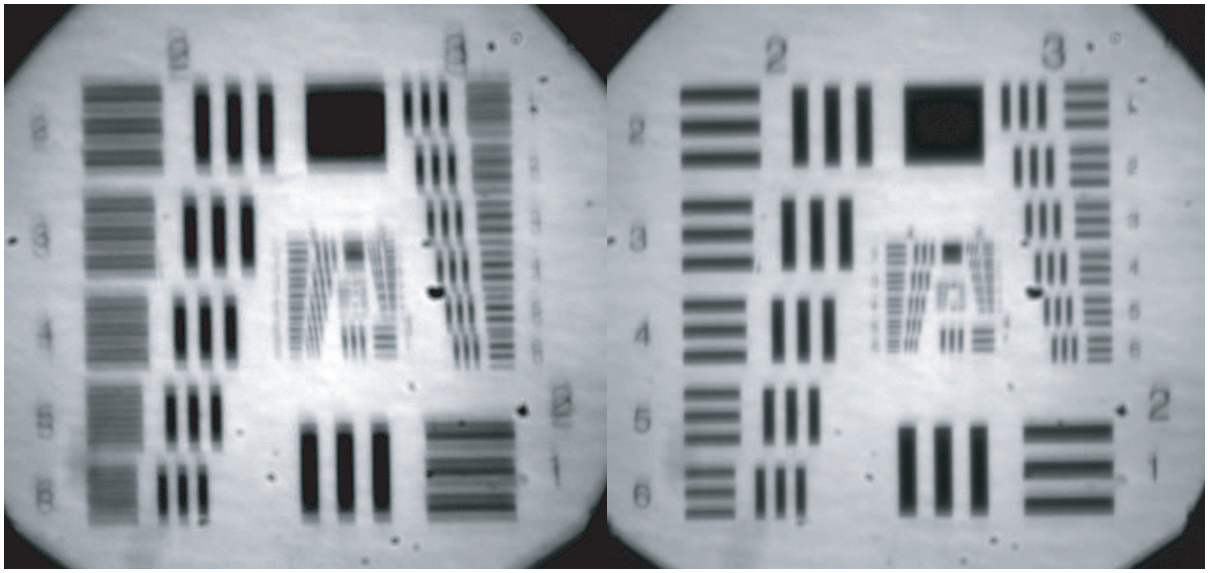


Fig. 7.24: SPGD correction for combination of defocus and astigmatism.

Figure 7.23 shows a similar result for the correction of increasing amounts of astigmatism as to that for increasing amounts of defocus. The deformable mirror can correct for lower amounts of astigmatism, but as expected can only partially correct for aberrations outside the range of stroke of the deformable mirror. The stochastic parallel gradient descent algorithm achieves quite a good correction for a combination of astigmatism and defocus as seen in Figure 7.24.

The ability of the stochastic parallel gradient descent algorithm to correct for multiple orders of Zernike aberrations is shown in Figure 7.25 The aberration applied is ~ 0.8 $\lambda(P-V)$ for each of the Zernikes, $j = 4-18$. It can be seen that the stochastic parallel gradient descent algorithm increases the contrast and resolution of the central bar target area.

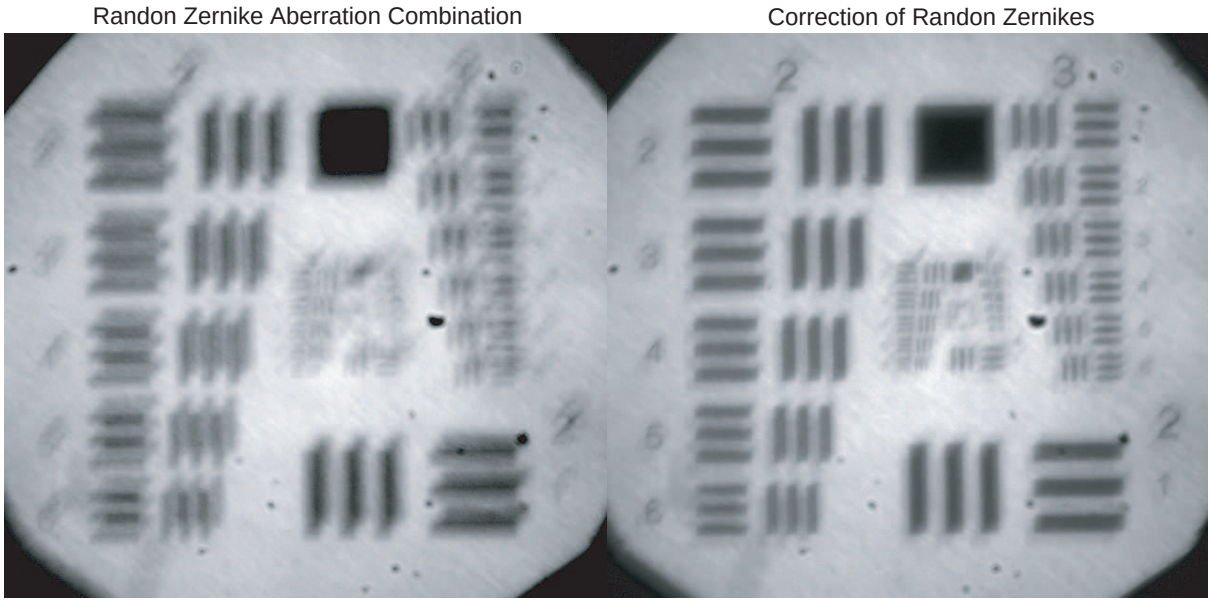


Fig. 7.25: SPGD correction of random combination of Zernike aberrations.

The corrections shown thus far for each of the algorithms have been performed on an aberrated USAF target image. Figure 7.26 shows correction as applied to a real image, taken of a University building. This image is aberrated by ~ $1.3\lambda(P-V)$. The edges of the buildings windows are enhanced, becoming clearer.

The SPGD algorithm correction time varied depending on the strength of the aberration to be corrected, but typically corrections took < 60 seconds making it the fastest algorithm, while also producing the best correction results. The correction time corresponds to approximately 200 iterations where each iteration requires 2 sharpness measurements.

7.3.6 Performance Of Search Algorithms

Four search algorithms were used to correct for aberrated extended object images and their ability to correct, both in terms of speed and degree of correction can be compared.

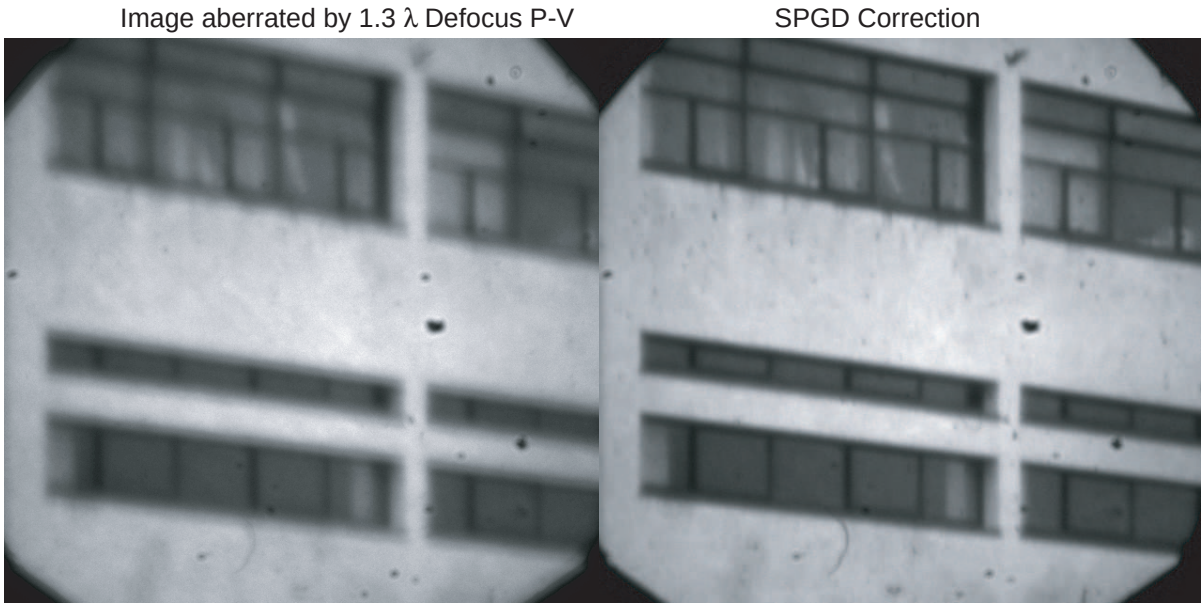


Fig. 7.26: SPGD correction of real image.

The results of the algorithms were discussed in order from the slowest to the fastest correction provided. The simplex algorithm took in the region of 40 minutes to correct an aberrated image, requiring 500 iterations. The stochastic gradient descent algorithm took ~ 300 seconds to perform approximately 55 iterations, offering a similar degree of correction. The speed of correction of the stochastic gradient descent algorithm was found to be similar to the modal algorithm implemented, which took ~220 seconds to perform 1517 iterations, yielding results on a par with the simplex algorithm and the modal algorithm. It can be seen from Figure 7.14, Figure 7.16 and Figure 7.19, which correspond to the simplex, stochastic gradient descent and the modal algorithm, respectively, that each algorithm provided similar correction results for an extended image aberrated by the same magnitude of aberration.

The algorithm that performed the best both in terms of speed and degree of correction was the stochastic parallel gradient descent algorithm. It was found to provide good correction of aberrated images in < 60 seconds, for approximately 200 iterations. These results mean that the stochastic parallel gradient descent algorithm corrects, approximately, a factor of 40 times faster that the simplex algorithm, 5 times faster than the stochastic gradient descent and a factor of 4 times faster than the modal algorithm. The stochastic parallel gradient descent algorithm was shown to correct better that the other algorithms for higher amounts of aberration and was also found to be the algorithm which was most sensitive to small amounts of aberration. It was able to correct for small amounts of inherent aberration present in the extended object image as can be seen in Figure 7.22. The other three algorithms were unable to provide any meaningful correction to the inherent aberrations of the system.

7.3.7 Noise Limitation of Sharpness Metrics/Search Algorithms

Each correction algorithm was shown to be well corrected for lower order Zernike aberrations applied to the image of a USAF target image. The target image enabled easy detection of correction for the images due to the areas of decreasing resolution as well as profiles through sections of the bar target to show the increased contrast achieved upon correction.

Although Figures 7.6 and 7.7 show that 4x4 binning of CCD images increases the noise in the sharpness metric, as 4x4 binning effectively undersamples the image, the correction results shown previously for each algorithm were obtained for 4x4 binning of the CCD camera. 1x1 binning for each algorithm was also carried out but yielded similar results. This may be because image frames contained more noise partially due to the

increasing effect of read noise for longer exposure times. The illumination of the LED used to image the objects had a limited magnitude, and this coupled with a relatively long imaging pass, imaging through a linear polariser with a high extinction ratio which greatly attenuated the intensity of the light, meant that long image exposures were needed.

When the camera was operating in a 4x4 binning mode, the exposure time required to utilise most of the dynamic range of the camera was 17*ms* compared to 280*ms* when the camera was operating in a 1x1 binning mode. This greatly increased the iteration time for the search algorithms, without yielding better correction results. The fact that the search algorithms still corrected for aberrated target image indicates that the search space is sufficiently well defined and that the maximum was unaffected by the noise in the sharpness value resulting from undersampling the image, and the lower signal-to-noise ratio due to 4x4 binning.

When attempting to correct for extending images of real objects the exposure times increased further still due to the reduced transparency of the image slides. The images camera exposure time required for 4x4 binning increased to 70*ms* and for 1x1 binning to 1.2*s*. Figure 7.26 shows correction achieved with 4x4 binning which has a similar signal-noise-ratio as that of the USAF target images. The correction achieved is of limited extent for 4x4 binning, probably because the image has lower contrast and is more sensitive to undersampling. For 1x1 binning any decrease in the noise of the sharpness metric when the image is sampled correctly, is outweighed by the large increase in the read noise of the CCD camera which decreases the signal-to-noise ratio.

8. CONCLUSIONS

8.1 Correction of Image of a Point Source

The power-law metrics set out by Muller and Buffington have been shown to produce good correction for a point source image. The Strehl factor improvement achieved by image sharpness metrics I^2, I^3 and I^4 is comparable. For each metric the maximum sharpness value corresponded to an increase in the images intensity in the central core. This is evidence that the image quality can be improved by maximising its sharpness. The simplex algorithm has been shown to determine a maximum sharpness value for each image metric. The simplex algorithm ran fastest for I^3, often achieving good correction in tens of seconds, whereas the I^2 and I^4 algorithms often took longer to converge.

The simplex algorithm preforms many more time consuming "shrink" operations for sharpness metrics $I^2$and I^4. This may be due to the nature of the search space for these metrics and the effect of noise for I^2,I^4. Nevertheless, each metric achieved a significant increase in the image quality. Each metric was seen to reach a plateau in terms of sharpness value and would not increase, no matter how many iterations were performed. This is a measure of the level of correction achieved as, from inspection, the maximum sharpness was effectively reached. Further correction was limited by the low stroke of

the deformable mirror. Each final image corresponding to 50 iterations and an maximum sharpness was a clear point source with its Airy rings restored.

It has been shown that the power-law metrics can be used to "blindly" correct for small aberrations in an indirect wavefront sensing system, or, act as a wavefront-sensor less correction. The sharpness metrics provided good correction for an aberrated point source image in relatively few iterations. A benefit of sharpness maximisation is that the corrections are determined from the science camera and thus this method may be useful to correct for non-common path errors in astronomy using a reference star. Sharpness maximisation could be used to maximise the sharpness value of a reference star, calibrating for the whole image path up to the science camera, where as wavefront sensors can only correct the optics system partially.

8.2 Modal Analysis of Search Space

In Chapter 7.2.3 mirror modes are applied to an unaberrated image in attempt to gain insight on how modal aberrations effect the sharpness value. Varying magnitudes of the mirror modes, which are closely modelled on Zernike aberrations, indicate that the sharpness is a smooth function which decreases steadily for increasing mirror modal aberrations. This also reflects the influence of the application of mirror modes on the sharpness value, and further shows the magnitude of aberrations that mirror modes are capable of correcting. When compared to the normalised sharpness value generated by moving the mirror from 0-250V, this shows that the application of mirror modes has a much reduced effect on the sharpness value. Generally modal aberrations based upon lower Zernike aberrations had the greatest effect on the sharpness value. Although the sharpness value for these

modal aberrations was measured for the variation of single aberrations at a time, its likely that the sharpness response for multiple aberrations is similar and that the sharpness space is well defined - at least for the extended object used in this experiment.

Recently Booth[100] has suggested a search method with relies upon sphere packing of the search space. These spheres are generated in such a way as to cover the whole search space using the least number of spheres. The size of the spheres are chosen based upon the desired accuracy in finding the global maximum and clearly the greater the accuracy the more complicated the search method becomes. Booth's paper assumes that the starting point of the search space is asymptotically close to the solution. In such a case its argued that the global maximum can be determined in N+1 measurements. This method is an example of a search algorithm being developed depending upon the topology of the search space.

8.3 Correction of Extended Objects Images

Four search algorithms were used to correct for aberrated images of extended objects: a simplex algorithm, stochastic gradient descent, stochastic parallel gradient descent and a modal algorithm. Each algorithm was tested to correct for defocus and astigmatism aberrations, as well as aberrations from the systems optics, present in small amounts. It was shown that the simplex algorithm, stochastic gradient descent and modal algorithm all achieved similar results offering good correction for aberrated images.

A large discrepancy in the performance of the search algorithms lay in the correction time. While the stochastic gradient descent algorithm and modal algorithm corrected to a similar degree and in a similar time (~200-300s), the simplex algorithm took much longer

to correct for the applied aberrations. While the simplex algorithm corrected aberrations to a good degree, it took on average ~ 40 minutes for each correction. This result is not altogether unexpected as the simplex is, by its nature a slow but effective search algorithm.

The stochastic parallel gradient descent algorithm yielded the best correction results, both in terms of speed and degree of correction. It was shown to have the required sensitivity and ability to correct for small amounts of inherent aberrations, as well as large amounts of random combinations of Zernike aberrations. The iteration time required for the stochastic parallel gradient descent algorithm was typically $< 60s$, and often as short as $30s$. These results demonstrated that the stochastic parallel gradient descent is an attractive algorithm for the application of sharpness maximisation on extended objects.

This correction time for each algorithm varied due to the nature of the algorithms but it is something that can be approved upon significantly by optimising the performance of key components such as the CCD camera. As explained earlier the exposure time for the USAF target image was $17ms$ for 4x4 binning, $280ms$ for 1x1 binning and the CCD camera has an operational maximum readout rate of 7 full frames per second. Coupled to this is the fact that the sharpness calculation and control algorithm were not optimised. The primary purpose of this research was to show that aberrations of an extended object image can be corrected by maximising the sharpness of the image. To further, and enhance, the correction achieved through sharpness maximisation an optimised control algorithm and faster CCD camera would allow increased speed of correction. For example, using a CCD camera operating at 500 frames per second, it should be possible to obtain correction using the stochastic parallel gradient descent method in less than a second, and whilst this is inadequate for the correction of atmospheric turbulence it may be sufficient in other applications of adaptive optics, for example, in vision science and in lasers.

The deformable mirror employed to correct for the aberrated images was an OKO membrane mirror. The deformable mirror was measured to have ~1.5μm stroke over the central 9.6mm of its diameter. The limited actuator stroke of the deformable mirror restricted the amount of aberration the mirror could correct. The use of a deformable mirror with greater actuator stroke would allow for correction of stronger optical aberrations. An in the case of the modal search method might improve the degree of correction as more mode increments could be applied giving greater sensitivity to the sharpness measurement.

An important benefit that sharpness maximisation offers to many adaptive optics systems was shown in Figure 7.22 where the inherent optical aberrations are compensated using the stochastic parallel gradient descent algorithm to maximise the sharpness. In this way the aberration of the whole imaging pass can be calibrated for which is not the case in adaptive optics systems which use a wavefront sensor to correct for aberrations. Image correction based upon the science camera also has the benefit that none of the imaging light is "lost" to a wavefront sensor.

8.4 Future Work - Analysis of the Search Space through Pair-wise Variation of Aberrations

Further research on sharpness maximisation could include the analysis of the search space corresponding to various Zernike aberrations. This would provide some insight into the nature of the search space. Presently the shape of the search space of sharpness metrics applied to aberrated images is unknown and the issue to address here is whether the search space is "more like Mount Fuji or more like the Alps", i.e. its complexity. This

might help determine the choice of algorithm depending on the degree of aberration and correction required. As the search space has a large number of degrees of freedom - too many to search systematically - non-systematic search routines have been applied to locate the global maximum of the system in this research. These metrics are chosen on past results for search problems elsewhere. These search routines are not tailored to image degradation based on specific aberrations, that is, the shape of the search space is not reflected in the choice of metric. In an attempt to gain insight into the nature of the search space corresponding to various aberrations, the sharpness could be measured whilst varying the magnitude of pairs or combinations of aberrations. A comprehensive analysis of the search space corresponding to combinations of Zernike aberrations might allow search algorithms to be tailored to images aberrated by specific aberrations, as although the stochastic parallel gradient descent algorithm is shown to provide good correction, some systems containing only low order aberrations may be optimised within a few iterations, if the corresponding search space is known.

This might be achieved by placing one aberration of a given fixed magnitude and measuring the sharpness value as a second aberration is varied. This process could then reversed. Various combinations of low order aberrations could be tested and the effect on the sharpness value noted.

Discussion

While image sharpness maximisation may not be able to compete with direct wavefront sensor methods, especially as regards speed of implementation or guaranteed correction, its advantage lies in its simplicity and as such may be of use in industrial or medical ap-

plication for the correction of relatively slow-varying or even static aberrations. A further benefit of image sharpness maximisation is that the system uses the science camera to determine the corrections as opposed to using a wavefront sensor where a section of the beam is diverted to a wavefront sensor. This reduces the amount of light reaching the science camera, and for applications where light levels are critical, can be a significant drawback. For example, in vision science where only 1 photon in every 10,000 sent into the eye is returned. In astronomical observations great lengths and expense are diverted to purchase high spec CCD cameras with exceptionally high quantum efficiency due to limited number of photons from faint astronomical objects. This results in the need for longer integration time and thus the effects of turbulence become more pronounced.

The stochastic parallel gradient descent algorithm provided good correction and required a relatively short correction time. An optimised sharpness maximisation technique based upon the stochastic parallel gradient descent, operating with a fast CCD camera and a deformable mirror with greater actuator stroke might have the potential to corrected for large degrees of aberration in a short time. This form of wavefront sensor-less adaptive optics may be of use for systems having slowly varying, or static, aberrations which need to be corrected or calibrated.

BIBLIOGRAPHY

[1] H. W. Babcock, "*The possibility of compensating astronomical seeing*", Publ. Astro. Soc. Pac., **65**, 229-236 (1953).

[2] J. W. Hardy, J. E. Lefevre, and C. L. Koliopoulos, "*Real-time atmospheric compensation*", J. Opt. Soc Am., **67**, 360-369 (1977).

[3] F. Merkle, P. Kern, P. Lena, F. Rigaut, J. C. Fontanella, G. Rousset, C. Boyer, J. P. Gaffard, and P. Jagourel, "*Successful test of adaptive optics*", ESO Messenger, **58**, 1-4 (1989).

[4] R. A. Muller and A. Buffington, "*Real-time correction of atmospherically degraded telescope images through image sharpening*", J. Opt. Soc. Am., **67**, 1200-1210 (1974).

[5] Singular Value Decomposition `http://www.kwon3d.com/theory/jkinem/svd.html`

[6] Introduction to Singular Value Decomposition `http://www.uwlax.edu/faculty/will/svd/index/html`

[7] R. K. Tyson, "*Principles of Adaptive Optics*", Academic Press, 2nd Edition (1998).

[8] D. R. McGaughey, L. P. Murray, J. C. Dainty, "*Sharpness Metric Nonstationarity in Undersampled Systems*", Royal Military College of Canada, Kingston, Ontario, Canada.

[9] J. W. Hardy, "*Adaptive Optics for Astronomical Telescopes*", Oxford University Press (1998).

[10] M. Loktev, " *Modal wavefront correctors based on nematic liquid crystals*", Thesis submitted to Technische Universiteit Delft (2005).

[11] A. Kolomgorov, "*Dissipation of energy in locally isotropic turbulence*", in : S. Friedlander, E. Topper (Eds.), "*Turbulence, Classical Papers on Statistical Theory*", Wiley-Interscience, New York (1961).

[12] V. A. Kluckers, N. J. Wooder, T. W. Nicholls, M. J. Adcock, I. Munro, and J. C. Dainty, "*Profiling of atmospheric turbulence strength and velocity using a generalised scidar technique*", Astronomy and Astrophysics supplement series, **130**, 141-155 (1998).

[13] R. W. Wilson, "*Slodar: Measuring optical turbulence altitude with a Shack-Hartmann wavefront sensor*", MNRAS **337,** 103 (2002).

[14] Karhunen-Loeve Decomposition: Literature Review `www.nd.edu/~sgordeye/Project1/project1.html`

[15] R. Noll, "*Zernike polynomials and atmospheric turbulence*", J. Opt. Soc. Am., **66**, 207-211 (1976).

[16] Zernike Polynomials www.ctio.noao.edu/~atokvin/tutorial/part1/turb.html#SEC1.9

[17] M. Born, E. Wolf, "*Principles of Optics*", 6th Edition, Cambridge University Press, Cambridge (1997).

[18] J. W. Goodmann, "*Introduction to Fourier Optics*", McGraw-Hill, New York (1941).

[19] P. Clemmow, "*The Plane Wave Spectrum Representation of Electromagnetic Fields*", Pergammon Press, London (1966).

[20] J. A. Stratton, "*Electromagnetic Theory*", 1st Edition, McGraw-Hill, New York (1992).

[21] J. D. Jackson, "*Classical Electrodynamics*", Wiley, New York 1962.

[22] Hugens Principle http://www.mathpages.com/home.math242/math242.htm

[23] Maxwell's Equations http://scienceworld.wolfram.com/physics/MaxwellEquations.html

[24] J. Y. Wang and D. E. Silva, "*Wave-front interpretation with Zernike polynomials*", Appl. Opt., **19**, 1510-1518 (1980).

[25] J. L. Melsa and D. L. Cohn, "*Decision and Estimation Theory*," New York: Mc Graw-hill Book Company (1978).

[26] M. C. Roggemann and B. M. Welsh, "*Imaging through Turbulence*", CRC Press (1996).

[27] G. Rousset, Nato Advanced Study Institute Series, Volume: "*Adaptive Optics for Astronomy*", Cargese (1993).

[28] H. H. Barrett, K. J. Myers, N. Devaney and J. C. Dainty, "*Objective Assessment of Image Quality: IV. Application to Adaptive Optics*", J. Opt. Soc. Am. A, (2006).

[29] E. Dalimier and C. J. Dainty, "*Comparative analysis of deformable mirrors for ocular adaptive optics*", Opt. Exp., **13**, 11 (2005).

[30] J. Primot, G. Rousset and J. C. Fontella, "*Deconvolution from wavefront sensing: a new technique for compensating turbulence-degraded images*", J. Opt. Soc. Am. A. 1598-1608 (1990).

[31] F. Roddier, "*Curvature sensing and compensation: a new concept in adaptive optics*", Appl. Opt., **27**, 123-125 (1988).

[32] F. Roddier, M. Northcott, J. E. Graves, "*A simple low-order adaptive optics system for near-infrared applications*", Astron. Soc. Pac., **103**, 131-149 (1991).

[33] J. C. Wyant, "*White light extended source shearing interferometer*", Appl. Opt., **13**, 200-202 (1974).

[34] R. Ragazzoni, "*Pupil plane wavefront sensing with an oscillating prism*", J. Modern Opt., **43,** 289-293 (1996).

[35] C. Schwartz, E. Ribak, and S. G. Lipson, "*Bimorph adaptive mirrors and curvature sensing*", J. Opt. Soc. Am. A, **11**, 895-902 (1994).

[36] R. A. Gonsalves, "*Phase retrieval from modulus data*" Opt. Soc. Am., **66**, 961-964 (1976).

[37] R. A. Gonsalves, "*Phase retrieval and diversity in adaptive optics*", Opt. Eng., **21**, 829-832 (1982).

[38] R. A. Gonsalves, "*Perspectives on phase diversity*" Conference on wavefront sensing controls, Hawaii, Nov. (2000).

[39] R. G. Paxmann, T. J. Schulz and J. R. Feinup, "*Joint estimation of object and aberration by using phase diversity*", Opt. Soc. Am. A, 1072-1085 (1992).

[40] M. A. Vorontsov, A. V. Koriabin and V. I. Shmalausen, "*Controlling Optical Systems*", Nauka, Moscow (1998).

[41] J. Hardy, "*Active Optics: A new technology for the control of light*", IEEE Proc., **66**, 651-697 (1978).

[42] T. R. O Meara, "*The multi-dither principle in adaptive optics*" Opt. Soc. Am., **67**, 306-315 (1977).

[43] T. R. O Meara, "*The theory of multi-dither adaptive optical systems with zonal control of deformable mirrors*", Opt. Soc. Am., **67**, 318-325 (1977).

[44] E. Dowski and W. T. Cathey, "*Extended depth of field through wavefront coding*", Appl. Opt., **34** (1995).

[45] W. T. Cathey and E. Dowski, "*New paradigm for imaging systems*", Appl. Opt., **44**, 29 (2002).

[46] K. Kubala, E. Dowski and W. T. Cathey E, "*Reducing the complexity in computational imaging systems*", Opt, Express, **11**, 8 (2003).

[47] R. H. Freeman and J. E. Pearson, "*Deformable mirrors for all seasons and reasons*", Appl. Opt. **21**, 580-588 (1982).

[48] E. Steinhaus and S. G. Lipson, "B*imorph piezoelectric flexible mirror*" J. Opt. Soc. Am., **69**, 478-481 (1979).

[49] S. A. Kokorowsky, "*Analysis of adaptive optical elements made from piezoelectric bimorphs*", Opt. Soc. Am., **69**, 181-187 (1979).

[50] MEMs http://www.cvs.rochester.edu/williamslab/research/option13.html

[51] G. Vdovin and P. M. Sarro, "*Flexible mirror micromachined in silicon*", Appl. Opt., **34**, 2968-2972 (1995).

[52] OKO Technologies http://www.okotech.com

[53] Boston Micromachines Mirrors http://www.bostonmicromachnes.com

[54] Iris AO, Inc. Mirrors http://www.irisao.com

[55] G. D. Love, "*Wavefront correction and production of Zernike modes with a liquid crystal spatial light modulator*", Appl. Opt., **36**, 1517-1524 (1997).

[56] J. Gourlay, G. D. Love, P. M. Birch, R. M. Sharples, A Purvis, "*A real-time closed-loop crystal adaptive optics system: first results*", Opt. Commun., **137**, 17-21 (1997).

[57] D. V. Wick, T. Martinez, et al, "*Deformed-helix ferroelectric liquid-crystal spatial light modulator that demonstrates high diffraction efficiency and 370-line pairs/mm resolution*", Appl. Opt., **38**, 3798-3803 (1999).

[58] S. R. Dale, G. D. Love, R. M. Myers, A. F. Naumov, "*Wavefront correction using a self-referencing phase conjugation system based on a Zernike cell*", Opt. Comm., **191**, 31-38 (2001).

[59] S. T. Kowel, D. S. Cleverly, in: "*Proceedings of NASA conference Optical Information Processing for Aerospace Applications*", 329-340 (1981).

[60] Meadowlark Optics, http://www.meadowlark.com

[61] Holoeye Photonics AG, http://www.holoeye.com

[62] P. M. Prieto, E. J. Fernandez, S. Manzanera, P. Artal, "*Adaptive optics with a programmable phase modulator: applications in the human eye*", Opt. Exp., **12**, 4059-4071 (2004).

[63] X. H. Wang, B. Wang, J. Pouch, F. Miranda, J. E. Anderson, P. J. Bos, "*Performance evaluation of a liquid-crystal-on-silicon spatial light modulator*", Opt. Eng., **43**, 2769-2774 (2004).

[64] M. T. Gruneisen, R. C. Dymale, J. R. Rotge, L. F. DeSandre, D. L. Lubin, "*Compensated telescope system with programmable diffractive optics*", Opt. Eng., **44**, 023201 (2005).

[65] P. M. Prieto, E. J. Fernandez, S. Manzanera, P. Artal, "*Adaptive optics with a programmable phase modulator: applications in the human eye*", Opt. Exp., **12**, 4059-4071 (2004).

[66] D. V. Wick, T. Martinez, S. R. Restaino, B. R. Stone, "*Foveated imaging demonstration*" Opt. Exp., **10**, 60-65 (2002).

[67] A. Monmayrant, B. Chatel, "*New phase and amplitude high resolution pulse shaper*", Review of Scientific Instruments, **75**, 2668-2671 (2004).

[68] N. Doble, "*Image Sharpening Metrics and Search Strategies for Indirect Adaptive Optics*", Thesis submitted to the University of Durham (2000).

[69] L. P Murray, J. C. Dainty and E. Daly, "*Wavefront correction through image sharpness maximisation*" Proc. SPIE, **5823**, 40-47 (2005).

[70] E. N. Ribak, E Gershnik, and M.Cheselka, "*Stellar scintillations as a remote atmospheric wavefront-sensor*", Opt. Lett., **21**, 435-437 (1996).

[71] V. I. Polejaev and M. A. Vorontsov, "*Adaptive active imaging system based on radiation focusing for extended targets*", SPIE Proc., **3126**, (1997).

[72] J. R. Feinup, J. J. Miller, "*Aberration correction by maximising generalised sharpness metrics*", Opt. Soc. Am., **20**, 4, 609-620 (2003).

[73] M. A. Vorontsov, "*Image quality criteria for an adaptive imaging system based on statistical analysis of the speckle field*", Opt. Soc. Am., **13**, 7, 1456-1466 (1996).

[74] M. Cohen, G. Cauwenbergs and M. A. Vorontsov, "*Image sharpness and beam focus VLSI sensors for adaptive optics*", IEEE Sensors, **2**, 6, (2002).

[75] R. T. Brigantic, M. C. Roggemann, K. W. Bauer and M. Welsh, "*Image-quality metrics for characterizing adaptive optics systems performance*", Appl. Opt., **36**, 6583-6593 (1997).

[76] M. A. Vorontsov et al, "*Adaptive imaging system for phase-distorted extended source and multiple-distance objects*", Appl. Opt., **36**, 15, 3319-3328 (1997).

[77] Image Basic C++ Code `http://www.kwasan.kyoto-u.ac.jp/~kamio/archive/image/`

[78] J. L. Starck, E. Pantin, F. Murtagh, "*Deconvolution in Astronomy: A Review*", Astron. Soc. Pacific, **114**, 1051 (2002).

[79] J. M. Conan, T. Fusco, L.M. Mugnier, F. Marchis, C. Roddier, F. Roddier, "*Deconvolution of adaptive optics images: from theory to practice*", SPIE, 4007, 913 (2000).

[80] J. C. Christou, E.K. Hege, S.M. Jefferies, "*Speckle Deconvolution Imaging Using An Iterative Algorithm*", SPIE, **2566**, 134 (1995).

[81] J.C. Christou, D. Bonaccini, N. Ageorges, F. Marchis, "*Myopic Deconvolution of Adaptive Optics Images*", ESO Messenger, **97**, 14 (1999).

[82] W. H. Press, S. A. Teukolsky, W. T. Vetterling and B. P. Flannery. "*Numerical Recipes in C*" Second Edition. Cambridge University Press, Cambridge, (1992).

[83] J. A. Nelder and R. Mead, "*A simplex method for function minimisation*", Computer Journal, **7**, 308-313 (1965).

[84] J. C. Lagarias, J. A. Reeds, M. H. Wright and P. E. Wright, "*Convergence properties of the Nelder-Mead Simplex Method in low dimensions*," SIAM J. Optim., **9**, 1, 112-147 (1998).

[85] E. Parzen, "*Stochastic Processes*," Oakland CA: Holden Day p. 7, (1962).

[86] M. A. Vorontsov, G. W. Carhart, J. C. Ricklin, "*Adaptive phase-distortion correction based on parallel gradient-descent optimization*", Opt. Lett., **22**, 907-909 (1997).

[87] T. Weyrauch, M. A. Vorontsov, T. G. Bifano, "*Microscale adaptive optics : wavefront control with a μ-mirror array and a VLSI stochastic gradient descent controller*", Appl. Opt., **40**, 4243-4253 (2001).

[88] Conjugate Gradient Method whttp://www.mathworld.wolfram.com/conjugateGradeintMethod.html

[89] H. H. Barrett, K J. Myers, "*Foundations of Image Science*", Wiley-Interscience (2004).

[90] Nyquist Sampling http://mathworld.com/NyquistFrequency.html

[91] H. Nyquist, "*Certain topics in telegraph transmission theory*", Trans. AIEE, **47**, 617-644 (1928).

[92] Zeemax Optical Design http://www.zeemax.com

[93] Linos Photonics Alignment http://www.linos-photonics.com

[94] Foturan Glass http://mikroglas.com/foturane.htm

[95] B. Y. Zeldovich, N. F. Pilipetski, V. V. Shkunov, "*Principles of phase conjugation*", Springer Series in Optical Sciences, **42**, Springer, Berlin (1985).

[96] M. A. Vorontsov, A. V. Koryabin, V. I. Shmalgausen, "*Controllable optical systems*", Nauka, Moscow (1988).

[97] S.Kirkpatrick and C. D. Gelatt and M. P. Vecchi, "*Optimization by Simulated Annealing*", Science, **220**, 4598, 671-680 (1983).

[98] Hamamatsu Photonics SLM `http://www.hamamatsu.com/`

[99] Fisba Interferometer `http://www.fisbaoptik.com`

[100] M. J. Booth, "*Wave front sensor-less adaptive optics: a sphere model-based approach using sphere packings*", Opt. Exp. **14**, 4, 1339-1352, (2006).