43

FERMILAB
OCT 20 1999
LIBRARY

PREPRINTS

of the

STEWARD OBSERVATORY

THE UNIVERSITY OF ARIZONA
TUCSON, ARIZONA 85721-0065 U.S.A.

STEWARD-1469

No. 1469

Fermilab Library
0 1160 0067835 3

ADAPTIVE OPTICS: NEURAL NETWORK WAVEFRONT SENSING, RECONSTRUCTION AND PREDICTION

PATRICK C. MCGUIRE[1], DAVID G. SANDLER[1,2], MICHAEL LLOYD-HART[1], AND TROY A. RHOADARMER[1]

[1]*The University of Arizona, Steward Observatory*
Center for Astronomical Adaptive Optics, Tucson, AZ 85721, U.S.A.
[2]*ThermoTrex Corporation, 10455 Pacific Center Court, San Diego, CA 92121, U.S.A.*

NAME	I.D No.	M.S.

RETURN TO FERMILAB LIBRARY

TO APPEAR IN SCIENTIFIC APPLICATIONS OF NEURAL NETS, LECTURE NOTES IN PHYSICS (SPRINGER, HEIDELBERG, 1999), IN PRESS.

Adaptive Optics: Neural Network Wavefront Sensing, Reconstruction and Prediction

Patrick C. McGuire[1], David G. Sandler[1,2], Michael Lloyd-Hart[1], and Troy A. Rhoadarmer[1]

[1] The University of Arizona, Steward Observatory,
Center for Astronomical Adaptive Optics,
Tucson, AZ 85721, U.S.A.
[2] ThermoTrex Corporation
10455 Pacific Center Court
San Diego, CA 92121, U.S.A.

Abstract. We introduce adaptive optics as a technique to improve images taken by ground-based telescopes through a turbulent blurring atmosphere. Adaptive optics rapidly senses the wavefront distortion referenced to either a natural or laser guidestar, and then applies an equal but opposite profile to an adaptive mirror. In this paper, we summarize the application of neural networks in adaptive optics. First, we report previous work on employing multi-layer perceptron neural networks and back-propagation to learn how to sense and reconstruct the wavefront. Second, we show how neural networks can be used to predict the wavefront, and compare the neural networks' predictive power in the presence of noise to that of linear networks also trained with back-propagation. In our simulations, we find that the linear network predictors train faster, they have lower residual phase variance, and they are much more tolerant to noise than the non-linear neural network predictors, though both offer improvement over no prediction. We conclude with comments on how neural networks may evolve over the next few years as adaptive optics becomes a more routine tool on the new large astronomical telescopes.

1 Principles of Adaptive Optics

Astronomy has the purpose of exploring the heavens, and in so doing, requires observations be made with ever-increasing clarity, usually meaning that the images be as sharp as possible (blurriness is to be avoided), and that the images have the highest possible contrast (so that dim features can be studied). Military reconnaissance has similar requirements of the images taken of earth-orbitting satellites. The first requirement, called the 'high-resolution imaging' requirement, has repercussions on the second requirement, called the 'faint imaging' requirement – usually when the image is made sharper, then it is possible to see fainter features of the image. Adaptive optics has the main purpose of providing much higher-resolution images for ground-based optical/infrared telescopes, that would otherwise be severely limited by the turbulent blurring of the atmosphere.

There are many motivating factors and many challenges associated with observing astronomical objects or man-made satellites from the surface of the Earth at near-visible (i.e. visible, near- to mid-infrared, and ultra-violet) wavelengths. The motivating factors include:

1) Discovery and imaging of planets and brown dwarfs orbitting nearby stars;
2) Study of star-formation in nearby star-forming regions;
3) Searching for dark matter in our Galactic halo in the form of stellar remnants (i.e. neutron stars) or substellar objects (brown dwarfs) by their gravitational micro-lensing of background field stars;
4) Verification, discovery, and study of gravitationally-lensed Active Galactic Nuclei (AGNs) in order to constrain the Hubble Constant;
5) Mapping of the inner realms of star-forming galaxies ("starbursts") at moderate red-shift ($z \equiv \frac{\Delta\lambda}{\lambda} = 1$) and possible associated AGNs to constrain galaxy evolution models;
6) High-resolution mapping of planetary bodies like Jupiter's moon Io to determine the spatiotemporal properties of its volcanoes, as well as the asteroid Vesta to study its mineralogy;
7) High-accuracy determination of the time-dependence of the two- and three- dimensional shape of the Sun in order to determine solar structure and to constrain theories of gravitation;
8) Spectroscopic studies of the rotation curves of nearby galaxies in order to infer dark matter content;
9) Spectroscopic studies of distant galaxies and protogalaxies which measure redshifts to infer distances and also to determine primordial element/isotope/molecular abundances in order to better understand big-bang nucleosynthesis;
10) Direct imaging of satellites orbitting the Earth in order to infer their capabilities.

In order to achieve the above goals for high-resolution imaging with ground-based telescopes, there are many technical challenges. These include the following:

1) The long-exposure blurring of images by the turbulent atmosphere at *all* near-visible wavelengths, producing seeing-limited images with resolution (1 arcsecond), which are not limited by the diffraction due to the aperture size;
2) The near- to mid-infrared thermal background of the atmosphere and the telescope;
3) Light pollution at visible wavelengths due to encroaching cities;
4) Wind-buffetting of the telescope;
5) Selective atmospheric absorption at various infrared and near-UV wavelengths;

6) The motion or apparent motion of the science object due to Earth rotation or satellite revolution-about-the-Earth;
7) The diurnal rising of the Sun and Moon;
8) Weather shutdowns (humidity, thick clouds, snowstorms) or weather annoyances (high-altitude cirrus);
9) The construction of telescopes of ever-increasing size in order to observe dimmer objects and also to attain higher spatial resolution;
10) The construction of CCD and Infrared Detectors of ever-increasing pixel number (e.g. 4096×4096), focal-plane area, quantum efficiency, electronic speed, spectral coverage, and ever-decreasing read-noise and thermal 'dark-current' noise.

A great deal of progress, perhaps a 'quantum leap', can be made on all of the motivating science factors if practical solutions can be found to Challenge 1, which is highly-linked to the effective usage of large ground-based telescopes (Challenge 9).

A much more expensive solution to all the Challenges facing ground-based astronomical and military observing is to place telescopes in orbit around the Earth. The absence of atmospheric blurring and other ground-based problems makes images with resolution better than 0.1 arcsecond available to astronomers (for telescope diameters greater than 2.4 m diameter at a wavelength of 1μm). For example, the Hubble Space Telescope (HST) serves *many* astronomers in their near-visible research; therefore observing time on the HST is relatively hard to obtain. Astronomers would be well-suited to have 5-10 different HST's, but at a price tag of over US$2 billion/HST, this will not happen. The current plan of US space-based near-visible astronomy (the Next Generation Space Telescope (NGST)) is to put into orbit by rocket by 2007 a single large telescope (8-meter class) at less cost than the HST. A ground-based 8-meter class telescope (e.g. Gemini, LBT, Keck) costs an order of magnitude less than the HST – hence, at the cost of 10-20 large ground-based telescopes, a single large space-based telescope can be put into orbit to produce images with resolution of 0.025 arcsecond (at 1 micron). Is such a space-based effort worth the high cost? Can we make ground-based competitive with the HST and NGST at a much lower cost?

One relatively inexpensive solution to (time-dependent) atmospheric blurring (Challenge 1) is first to record a series of very short exposures (1-20ms) of the science object or of a bright 'guidestar' near the science object in order to 'freeze' the turbulence. Each of these short-exposure images will predominantly have a high-resolution bright spot, called a 'speckle', somewhere in the image, and the resolution of this bright speckle is limited only by the telescope size, and not by the atmospheric turbulence. With the long series of short exposure images, during offline processing, the observer's computer program would first shift each image so that the dominant speckle of that image is coincident with all the other images' dominant speckle, after which the computer program would proceed to co-add all the images. This 'Shift-

and-Add' technique and its application to speckle images was developed by astronomers in the 1980's and achieved near diffraction-limited imagery (0.1 arcsecond resolution for a 2.4 m diameter telescope at a wavelength of 1 μm), but required very high signal-to-ratios. Hence, speckle-imaging is limited to relatively bright stars and their small neighborhoods. Unfortunately, only a limited amount of science can be produced by imaging near the bright stars that are suitable for speckle imaging.

Following up on military research that was declassified in the early 1990's, most major observatories around the world are developing a technique called 'adaptive optics' (AO). An AO-equipped telescope continually takes short-exposure measurements of the turbulence-distorted wavefronts of light coming from a guidestar above the large telescope, and then rapidly applies the conjugate of the turbulence's phase-map to a deformable mirror (DM) in either the AO instrument or telescope in order to cancel the blurring effect of the turbulence on the image of the science object. Long-exposures of the science object can easily be obtained just by integrating the image continuously during the application of the ever-changing shape of the DM (the AO system operates in 'closed-loop' servo). An AO system costs much less than the large telescope itself and costs significantly more than speckle-imaging offline processing. With the diffraction-limited performance obtained with large ground-based telescopes equipped with AO at only a small fraction of the cost of similar performance from a large space-based telescope, one begins to understand the allure and ongoing renaissance of ground-based telescope construction. In the remainder of this Section, we present a beginner's introduction to adaptive optics; we defer the astronomers of the audience to the final two Sections where we discuss old and new results pertaining to the use of neural networks in adaptive optics.

1.1 Large Ground-Based Telescopes

Since Newton and Galileo, astronomers and telescope builders have been entranced by the large-telescope mystique for two compelling reasons: more photons and higher spatial resolution. The rate of photon collection is proportional to the area of the primary mirror and hence the square of the telescope primary mirror diameter. Therefore, the 10-meter reflective telescopes of today afford a factor of 100 higher sensitivity to dim stars and dim galaxies than the 1-meter refractive telescopes of the nineteenth century.

Due to the diffraction theory of wave-like propagation of light developed by Frauenhofer and Airy, perfect 'stigmatic' focussing of light by a mirror or lens of finite size was shown to be impossible (see Meyer-Arendt (1984)), as was naively predicted by geometric optics. The half-width of the perfectly diffraction-limited spot was shown to be very close to the ratio $1.22\lambda/D$, where λ is the wavelength of light and D is the diameter of the telescope aperture. For example, for a wavelength of $\lambda = 0.5$ μm (blue-green light), and a lens of aperture $D = 1$ m, we obtain a diffraction-limited resolution of

0.61 μRadian $\equiv$ 0.12 arcseconds. From the inverse-proportional dependence of resolution upon aperture size, astronomers and opticians like Herschel and Frauenhofer immediately realized that by increasing the aperture size, one enhances the ultimate resolution of the optic. Unfortunately, the art of that period was refracting telescopes instead of reflecting telescopes, so chromatic aberration of refractive elements prevented diffraction-limited performance of these telescopes. It wasn't until the early twentieth century that the trend shifted, and the world's cutting edge 2 to 5 meter telescopes were constructed from reflective components.

Fifty to seventy years ago a few 5-meter-class reflecting telescopes were constructed (e.g. Hale). But until the 1980's, construction of such large telescopes had stagnated, with much of the telescope building effort going towards 2 to 3 meter telescopes. One reason for this was that light-weight mirror-making technology had not yet blossomed, and 5-meter telescopes meant very heavy, very high thermal time-constant primary mirrors. The high thermal time-constants meant that the telescopes did not cool to ambient temperatures during the night, causing some dome turbulence, but more destructively, adding significant thermal emission and making infrared astronomy much more difficult. Another important reason for the delay of the large primary mirror renaissance until the 1990's, was that large, shallow mirrors demanded long telescopes, which are difficult to construct and stabilize from a mechanical engineering standpoint. In the 1980's and 1990's, methods were perfected using spin-cast techniques (Hill & Angel (1992)), stressed-lap computer-controlled polishing (Martin *et al.* (1997)), and also actively-controlled segmented primary mirrors (Cohen, Mast & Nelson (1994)), to construct large, deep mirrors with small focal ratios ($f/D \sim 1.1$), which allowed compact telescope structures.

But the primary reason for the stagnation was that without adaptive optics technology to correct for atmospheric blurring, the payoff in spatial resolution from building larger telescopes was small. In the 1970's the Multiple Mirror Telescope (MMT) (a co-mounted array of six 1.8-meter diameter telescopes, spanning 6.9 meters) was constructed on Mt. Hopkins in Arizona (see Figure 1). Not equipped with adaptive optics until the 1990's, the main advantage of this telescope was the simultaneous pointing of 6 moderate-sized telescopes with a common focal-plane at the science object of interest, thus increasing the number of photons collected by a factor of 6 over 'easily' constructed telescopes of the day. The main disadvantage of the MMT is that when compared to other telescopes of similar size, the 41% filling factor means a factor of 2.5 less light collected than the state-of-the-art 6.5 meter telescope which is replacing the old MMT in the winter of 1998 (see Figure 2).

In Table 1, we list the large telescopes in order of size being built around the world that will be or already are incorporating adaptive optics systems. The 6.5 meter MMT upgrade $f/D = 1.25$ primary mirror was spin-cast in a

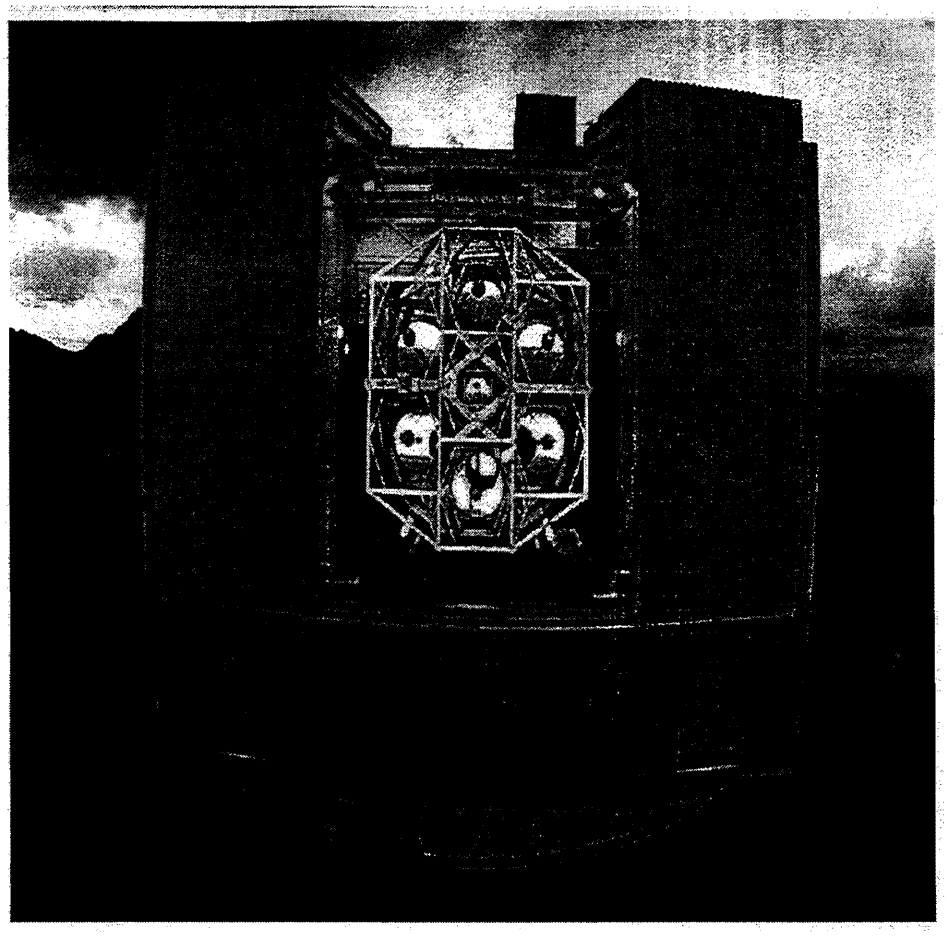


Fig. 1. The Multiple Mirror Telescope (MMT) (before March 1998) atop Mt. Hopkins in Arizona (after Serge Brunier of *Ciel et Espace*).

huge rotating oven at the University of Arizona Steward Observatory Mirror Laboratory in the early 1990's, and was also polished at the Mirror Lab in the mid 1990's, and will be taken to the telescope site on Mt. Hopkins in November 1998. A twin of the 6.5 meter MMT primary mirror is now being polished at the Mirror Lab for installation at the Magellan telescope in Chile. The largest telescope mirror in the world (one of the Large Binocular Telescope (LBT) 8.4 meter $f/D = 1.14$ primaries destined for Mt. Graham in Arizona) was successfully spun-cast at the Mirror Lab in 1997. The Keck I and II 10 meter $f/D = 1.75$ telescopes, built and maintained by the University of Califonia, have been installed and are operating on Mauna Kea in Hawaii, and each consist of 36 separately controlled close-packed segments. The Gemini

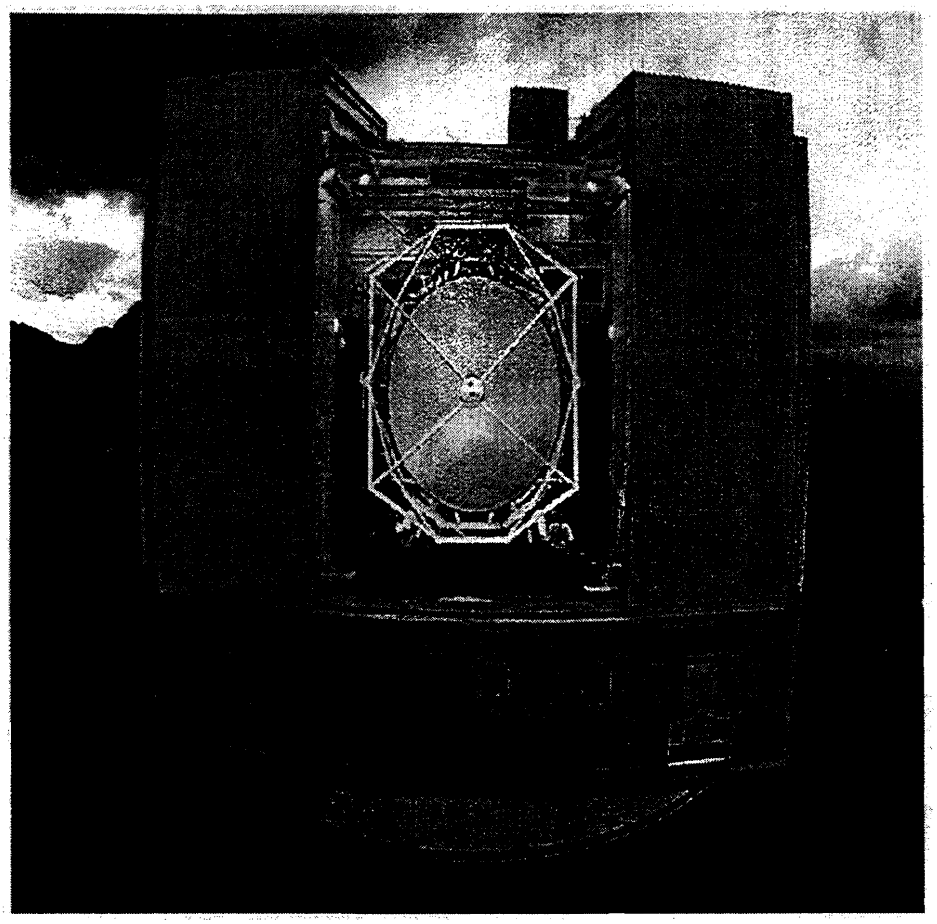


Fig. 2. The MMT-upgrade (6.5 meters diameter), as it will appear in January 1999 (after Serge Brunier of *Ciel et Espace*).

I and II 8 meter $f/D = 1.5$ telescopes will operate independently on Mauna Kea and in Chile, and are operated by a British/US/Canadian consortium. The Very Large Telescope Array, operated by the European Southern Observatory, consists of four 8.2 meter telescopes which can operate together in interferometric mode, of which one telescope became operational in 1998 in Cerro Paranal in northern Chile. Starfire Optical Range, run by the US Air Force Research Lab, in New Mexico has a 1.5 meter and a new 3.5 meter telescope, the primary mirrors being made at the University of Arizona. The CFHT (Canada-France-Hawaii telescope) is a 3.6 meter telescope operating on Mauna Kea, and the SUBARU telescope is a Japanese 8.3 meter telescope soon to be operating on Mauna Kea. And last but not least among this non-

TELESCOPE	DIAMETER	LOCATION
SOR I	1.5 m	Starfire Optical Range, New Mexico
MW 100"	2.5 m	Mt. Wilson, California
Shane	3.0 m	Mt. Hamilton, California
SOR II	3.5 m	Starfire Optical Range, New Mexico
Calar Alto	3.5 m	Calar Alto, Spain
Apache Point	3.5 m	Sacramento Peak, New Mexico
La Silla	3.6 m	La Silla, Chile
CFHT	3.6 m	Mauna Kea, Hawaii
WHT	4.5 m	La Palma, Canary Islands
MMT upgrade	6.5 m	Mt. Hopkins, Arizona
Magellan	6.5 m	Las Campanas, Chile
Gemini I	8.0 m	Mauna Kea, Hawaii
Gemini II	8.0 m	Cerro Pachon, Chile
VLT I-IV	8.2 m	Cerro Paranal, Chile
Subaru	8.3 m	Mauna Kea, Hawaii
LBT I & II	8.4 m	Mt. Graham, Arizona
Keck I & II	10.0 m	Mauna Kea, Hawaii

Table 1. List of telescopes that are using or will use adaptive optics

exhaustive list of large telescopes, the 3 meter Shane telescope is run by the University of Califonia's Lick Observatory on Mt. Hamilton in California. All of these large telescopes have been or will be outfitted with adaptive optics systems soon after commissioning.

1.2 The Blurring Effects of the Atmosphere

Newton (1730) recognized early on that the atmosphere was a major limiting factor to high resolution imaging through the atmosphere, and he recommended telescopic observations in the 'rarefied' air found on high mountaintops. Kolmogorov (1961) and Tatarskii (1961) developed a statistical model for the phase variations of light passing through turbulent air caused by refractive index variations. From his turbulence theory, Kolmogorov derived the following structure function:

$$D(r) \equiv E\left([\phi(\mathbf{r}+\mathbf{r}_0) - \phi(\mathbf{r}_0)]^2\right) = 6.88\left(\frac{r}{r_0}\right)^{\frac{5}{3}}, \tag{1}$$

where $D(r)$ represents the spatial structure function for phase variations as a function of the spatial separation r and parametrized by the Fried parameter r_0 (see Fried (1965)). The Fried parameter is calculated as an average over altitude of turbulence strength, and depends on zenith angle and wavelength

(Beckers (1993)), and is the maximum aperture through which diffraction-limited resolution can be obtained. This structure function equation is valid for distances between an empirically-determined inner scale and outer scale ($l < r < L$). The structure function is the mean-square phase error between two points in the aperture, and leads to a form for the residual RMS phase over an aperture which behaves as $(D/r_0)^{5/6}$. The Kolmogorov structure function states that for spatial separations $r < r_0$, there is relatively little phase variation, or in other words, there is spatial coherence of the wavefront of light passing through a patch of atmosphere of size r_0, whereas for separations of $r > r_0$, the wavefront of light is uncorrelated and hence the structure function is large, but limited by the outer scale. The Kolmogorov theory predicts (Noll (1976)) that there will be large-scale components which grow with aperture, including large components of tilt and focus. For very rapid images of a bright star taken with a telescope of diameter D, greater than the Fried coherence length r_0, there are typically D^2/r_0^2 bright speckles in the image, caused by the small coherence length. The Fried coherence length varies with time for a given site, but above Mt. Hopkins, the median r_0 at visible wavelengths (0.5 μm) is $\sim$ 0.15 meters. Due to the $\lambda^{6/5}$ dependence of r_0 on wavelength λ, this translates to a median r_0 in the near-infrared (2.2 μm) of $\sim$ 0.9 meters. Hence, the number of speckles (in the image plane) or coherence cells (in the pupil plane) for a short-exposure image taken with a large telescope (e.g. D = 6.5 meters) in the near-infrared (52 speckles at 2.2 μm) is 36 times less than in the visible (1900 speckles at 0.5 μm), thus making adaptive optics possible in the infrared and next-to-impossible in the visible.

The atmosphere is often composed of either 2 or 3 discrete layers of turbulence (Hufnagel (1974)) (see Figure 3). Each of these layers can in crude approximation be considered 'frozen' on timescales much longer than the time it takes the turbulence to be blown across the telescope aperture. This 'Taylor picture' (Taylor (1935)) of winds blowing a fixed pattern of turbulence past the observer is complemented by the prescription of the coherence timescale τ or 'Greenwood frequency' f_0 (Greenwood (1977), Fried (1994)). For temporal sensing frequencies below f_0, there is constant and large correlation of the wavefront of light impinging on a telescope (rigorously, the total phase error equals 1 radian at the Greenwood frequency, and the closed-loop error is 0.3 radians and one needs to sample about 10 times more frequently, so the sampling rate could be 30 times higher than the Greenwood frequency). For frequencies above f_0, the correlation plummets with a $f^{-8/3}$ power law dependence. The Greenwood frequency is given as:

$$f_0 \approx \frac{0.4v}{\sqrt{\lambda z}} = \frac{0.134}{\tau}, \tag{2}$$

where v is the bulk velocity of the wind in the turbulent layer of air and z is the altutide of the layer. For v = 40 m/s, λ = 2.2 μm, and z = 10000

meters, the Greenwood frequency is 100 Hz. Hence, for imaging in the near-infrared, adaptive optics system bandwidths need to be from several hundred Hz to 1kHz in order to effectively correct for the varying turbulence. This high bandwidth practically defines adaptive optics; lower bandwidth systems are called 'active optics' systems. The high bandwidth demands that only sufficiently bright adaptive optics guidestars be used; otherwise there will be insufficient photons to accurately measure or sense the shape of the distorted wavefront of light.

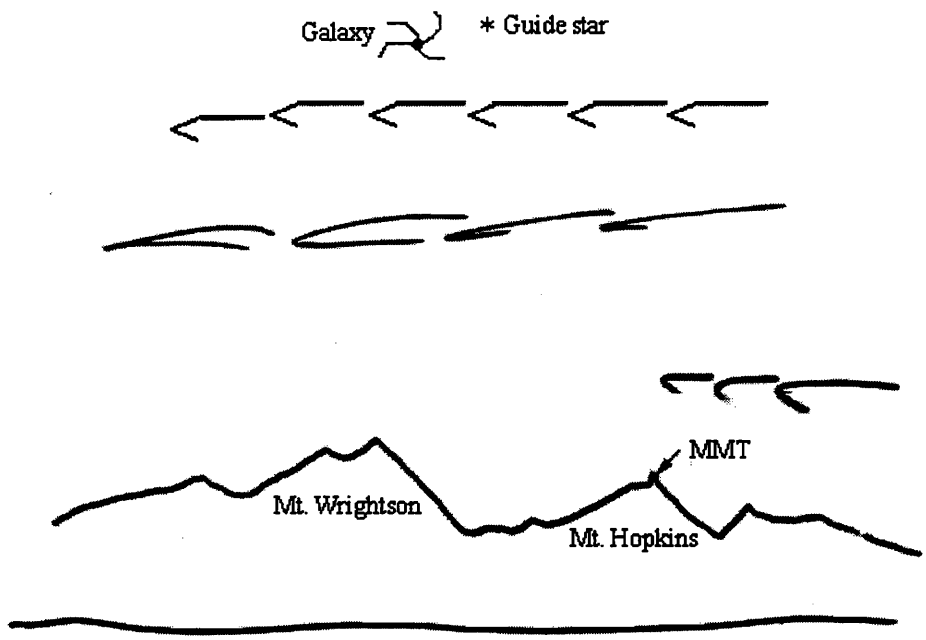


Fig. 3. A qualitative (not to scale) sketch of the main problem facing ground-based near-visible astronomy: atmospheric turbulence; shown here as three layers above Mt. Hopkins.

1.3 Resulting AO-Corrected Images: the Meaning of Strehl Ratios and Strehl Reduction by Non-Zero Wavefront Variance

The main purposes of any adaptive optics system are to improve resolution and concomitantly to increase the peak brightness of images of point-like sources taken beneath the blurring atmosphere. The fundamental limit to

both of these objectives is given by the diffraction due to the finite diameter of the primary mirror in the telescope (see section 1.1). Adaptive optics merit-functions are therefore derived with respect to this diffraction limit. The diffraction-limited (DL'ed) on-axis intensity is given by the brightness of the emitting source, the wavelength of the light, the diameter of the aperture, and the focal length of the telescope. For a fixed source and telescope, this DL on-axis intensity is *fixed* at I_0, and all aberrated performance is measured relative to the DL (see Figure 4), giving a relative on-axis intensity of:

$$\mathcal{S} \equiv \frac{I}{I_0}, \tag{3}$$

which is called the Strehl ratio. In the pupil plane, the atmosphere and optics will aberrate the phase-map of light coming from a point source, and the variance of this distortion (summed over the entire pupil), σ^2 (in rad^2), quantifies the disturbance. If the variance is given in terms of distance (i.e. $\mu\mathrm{m}^2$), the conversion to radians is accomplished by multiplying by $(2\pi/\lambda)^2$. Born & Wolf (1975) show that for moderate aberrations ($\sigma^2 < 1\,\mathrm{rad}^2$), the Strehl ratio is:

$$\mathcal{S} \approx \exp\left(-\sigma^2\right) \approx 1 - \sigma^2. \tag{4}$$

The DL resolution is given by Rayleigh's criterion ($0.251\lambda/D$ arcseconds, where λ is given in μm and D is given in meters), and seeing-limited resolution is often given in multiples of the DL resolution. However, for weak aberrations, the resolution may be very close to the DL, but the Strehl ratio may be as low as 0.6 or 0.7. Therefore, the Strehl ratio is sometimes a better performance criterion than the resolution. In the best of worlds, a full spatial transfer function (OTF/MTF) study should characterize an AO system, but we do not include such sophistication in the neural net prediction analysis reported here (Section 3). There are several contributions to the phase-variance σ^2 of wavefronts of light incident upon and going through an adaptive optics system (see Tyson (1991), Sandler *et al.* (1994b), Fried (1994)). These include fitting error, wavefront sensor error (due to finite number of photons and CCD read noise), servo lag error (discussed in detail in section 3), tilt anisoplanatism error, focus anisoplanatism (for laser guidestars), and reconstruction error.

1.4 Wavefront Sensing

The most common wavefront sensing technique used in adaptive optics systems was put forward by Shack & Platt (1971) and Hartmann (1900). Alternatives include the curvature sensor of Roddier (1988). A Shack-Hartmann sensor consists of an array of lenslets that is put at a pupil in the adaptive optics beam train and a camera put at the focus of the lenslet array. The most common lenslet and (CCD) camera geometry is called the Fried geometry, and consists of a square array of square lenslets, with each lenslet

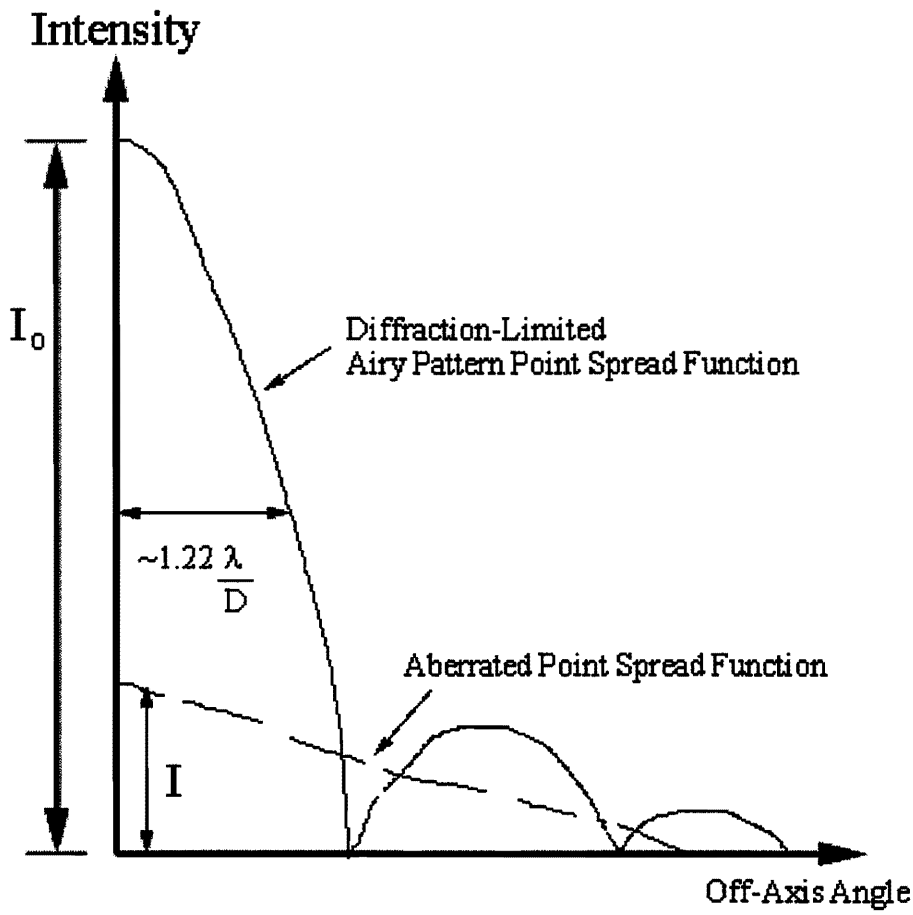


Fig. 4. A qualitative sketch showing the unaberrated and aberrated point spread functions for imaging. The ratio of the aberrated peak intensity to the unaberrated peak intensity is called the Strehl ratio, which is a key indicator of image quality.

focussing the unperturbed wavefront to a spot that is centered on the intersection of four CCD pixels, called a quad-cell (see Figure 5). The centers of the quad-cells in the lenslet-array pupil are made to be coincident with the corresponding positions of the actuators in the deformable mirror pupil. If the atmospherically-perturbed wavefront has tilt over a particular subaperture's lenslet in the x-direction, then there will be an imbalance of the signal registered by the CCD quad-cell in the x-direction (see Figure 6). Likewise, a tip in the y-direction will register a y-direction quad-cell imbalance. The array of tips and tilts over the pupil are often called the slopes, and can be

represented as both a matrix or a vector. Read-out noise and dark-current in the CCD detector (represented by n = RMS number of equivalent electrons of noise per pixel) and photon-counting noise (finite number of photons, N, caused by finite subaperture size, the required high temporal bandwidth, and the use of dim guidestars) will only allow imprecise determinations of the slopes. The RMS wavefront sensor centroid error is $N^{-1/2}$ waves, for pure Poisson noise – thus for 100 photons, there is 0.1 waves of centroid error contribution to the error in wavefront determination. For subapertures (on the sky) of 0.5×0.5 m^2, CCD integration times of 5 milliseconds, CCD read noise of $n = 3$ electrons, guidestars brighter than a visual magnitude of ~ 10 are considered bright and offer good adaptive correction (with an average of better than 3000 photons per subeparture per exposure and high signal to noise ratio), and guidestars dimmer than a visual magnitude of ~ 15 are considered dim (with less than 30 photons per subaperture per exposure and a signal to noise ratio of less than one).

1.5 Laser Guidestars

For purposes of accurate wavefront sensing and subsequent adaptive correction of the wavefront, guidestars of visual magnitude of at least 10 are required. Also, the guidestar is required to be within 20-40 arcseconds of the desired science object in order for the measured wavefront distortions in front of the guidestar to correspond closely to the wavefront distortions in front of the science object. Otherwise, due to the non-proximal guidestar, there is 'tilt anisoplanatism', and adaptive correction suffers. Rarely is there a guidestar of magnitude 10 or brighter within 20-40 arcseconds of the science object. In Figure 7, Quirrenbach and collaborators have found statistically that only 0.15% of the the sky is within 30 arcseconds of a magnitude 10 guidestar (near the Galactic pole), and even when observing in the plane of the Milky Way, only 0.6% of the sky is 'covered'. The scarcity of bright natural guidestars severely limits the quantity and perhaps the quality of astronomical studies which can be conducted with the improvements in resolution and brightness-sensitivity offered by adaptive optics.

Therefore, within the secret auspices of U.S. military research in the early 1980's, Happer (1982) proposed a key extension to the then nascent military artificial guidestar program, namely resonant back-scattering of a finely-tuned laser beam (projected from the ground) off a 10 km thick layer of naturally occurring atomic sodium at an altitude of 90 km and using the excitation of sodium from the hyperfine-split ground state to the first excited state as the resonance (Kibblewhite (1997)). They found that there is sufficient sodium column density and sufficient sodium absorption cross-section of yellow D2 light (589 nm), that a projected laser of moderate power (1-10 Watts) could produce a guidestar of magnitude greater than 10. Therefore, with such a laser co-projected from the observing telescope, a bright

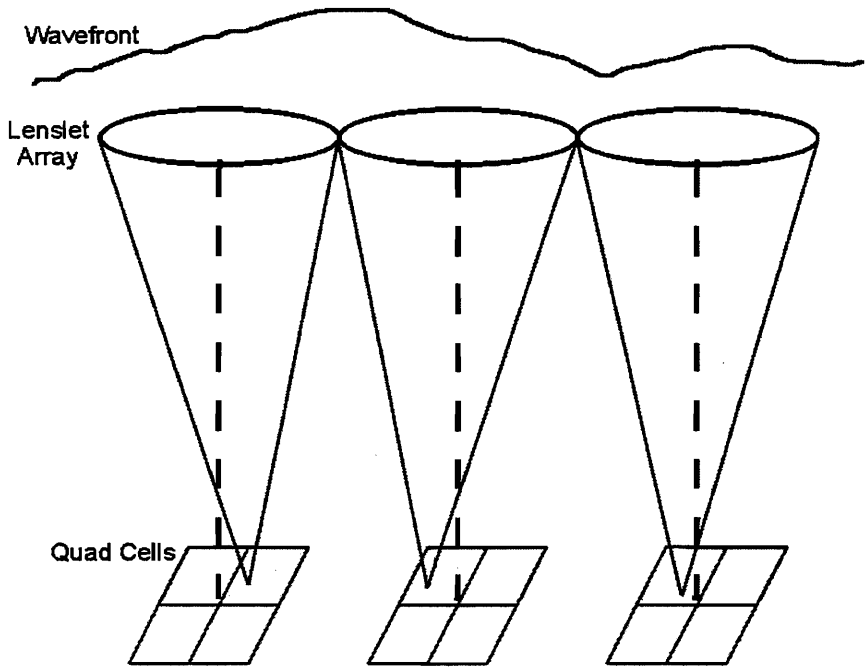


Fig. 5. A qualitative sketch showing the principle of Shack-Hartmann wavefront-sensing with a toy 1 × 3 lenslet array and 3 quad-cells. the wavefront slope is measured by the imbalance in the signals in the 4 CCD pixels that comprise a quad-cell.

guidestar could be projected anywhere on the sky, thus allowing adaptive-optics-fostered science to be pursued on any astronomical or man-made object of interest. These sodium guidestars have been developed and used by several research groups (Jacobsen (1997), Carter *et al.* (1994), Martinez (1998), Roberts *et al.* (1998), Shi *et al.* (1996), Friedman *et al.* (1995), Friedman *et al.* (1997), Avicola *et al.* (1994)); science has been produced with sodium guidestar correction (Max *et al.* (1997), Lloyd-Hart *et al.* (1998a)); and most of the world's AO groups will be outfitting their respective large telescopes with sodium laser guidestar systems in the coming years. The technology of tuning a laser to precisely the right wavelength (589.0 nm) is blossoming, with the past use of dye lasers (Martinez (1998)) and the present development of solid state Raman-shifted lasers (Roberts *et al.* (1998)).

Previous to the invention and development of the sodium laser guidestar, the US Air Force at Starfire Optical Range in New Mexico successfully devel-

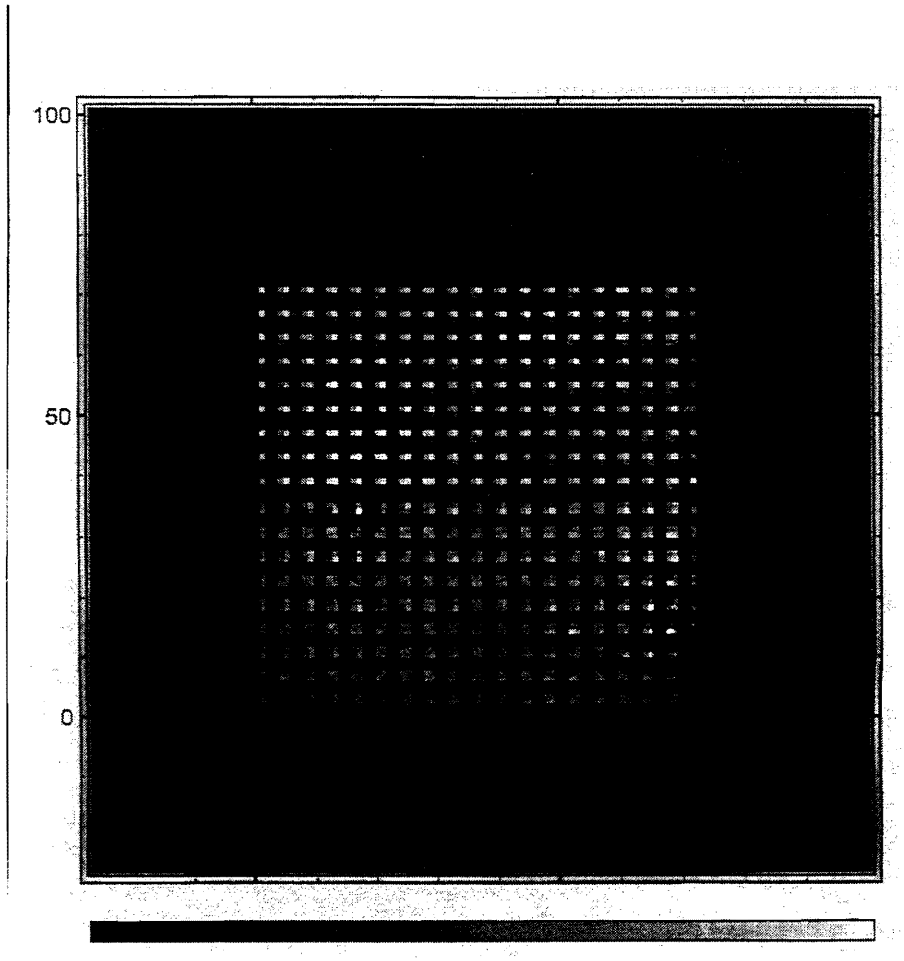


Fig. 6. An image of the CCD data output from the 13 × 13 subaperture Shack-Hartmann wavefront sensor camera being prepared for use with the MMT-upgrade AO system. The upper half of the array was not being clocked properly, so its signal was not centered on quad-cells, as in the lower half of the array (this problem has been fixed).

oped and is still primarily using another type of laser guidestar: the Rayleigh back-scattering guidestar (Fugate *et al.* (1991), Primmerman *et al.* (1991), Sandler *et al.* (1994a)). Such guidestars rely on the λ^{-4} Rayleigh scattering off of aerosols and dust in the atmosphere at altitudes of 10-30 km, and is pulsed and subsequently range-gated to select a sufficiently high-altitude to form the guidestar. The beauty of Rayleigh guidestars is that they can be made very

Sky coverage or probability of finding a natural guide star with a given magnitude within 30 and 10 arcsec from a random point in the galactic plane and at the pole.

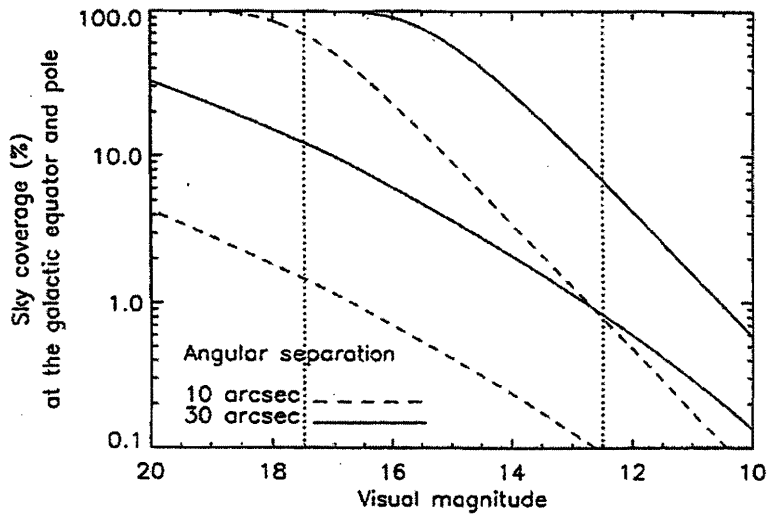


Fig. 7. A graph (after Quirrenbach (1997)) showing the fraction of the sky that has a star brighter than the given magnitude within 10 or 30 arcseconds. The upper solid and dashed lines are for directions towards the Galactic plane, and the lower solid and dashed lines are for directions towards the Galactic poles.

bright (V magnitude < 5) with copper vapor lasers due to the non-necessity of fine-tuning of the laser wavelength; such bright guidestars are needed to ensure sufficient resolution in the corrected visible images of the satellites in question. The main drawback of Rayleigh guidestars is that they are low altitude, and hence for two reasons do not sample the same turbulence as seen by the light emanating from the science object, but the US Air Force uses these low-altitude beacons because of the unavailability of magnitude 5 sodium laser guidestars. The first reason is that for laser guidestar altitudes of less than 20 km there can be unsensed turbulence *above* the Rayleigh guidestar. The second reason is that even if there was no atmosphere above the Rayleigh guidestar, 'focus anisoplanatism' or the 'cone effect' allows proper sampling of the turbulence within a cone emanating from the laser guidestar to the

telescope aperture, but does not allow *any* sampling of the turbulence outside this cone, but within the wider cone (or practically a cylinder) of light subtended by the science object at the apex and the telescope as the base. This cone effect also applies to the higher altitude sodium guidestars, but to a much lesser extent.

One problem does remain, even with the development of laser guidestars. This problem arises because the laser guidestar is projected from the ground, usually by a projection telescope attached to the large observing telescope. This means that the upgoing laser samples the same turbulence as the downgoing resonant backscatter laser light. Therefore, when the telescope aperture is divided into subapertures for wavefront sensing, each subaperture will only be able to measure the tilt *relative* to the upgoing projected laser beam, which is itself experiencing a *global* deflection or tilt relative to a 'fixed' direction towards the science object. This global tilt would go unsensed unless auxiliary countermeasures are taken. The countermeasure that is most frequently applied is using a third camera (the first being the science camera, the second being the wavefront sensor camera) to measure the global tilt, using light from a natural guidestar collected by the *whole* telescope aperture instead of the fine spatial subdivision of the light as done by the wavefront sensor. This still relies on using a natural guidestar, but allows the use of much dimmer natural guidestars (perhaps 300 times dimmer, or magnitude 16), which are *much* more plentiful and allow full sky coverage. Such coverage would not be possible without the laser guidestars used for differential tilt measurements.

1.6 Control of the Adaptive Mirror

The standard method of adaptive correction relies on inserting a reflective element in the adaptive optical system, prior to the science camera, and then from the signals derived from the wavefront sensor camera, deforming the mirror so as to null the wavefront sensor signals as accurately as possible. Of course, for accurate correction of the turbulence, three characteristics are required of the adaptive element:

1) sufficient stroke;
2) sufficient speed;
3) sufficient spatial resolution.

By the 1970's, low-order active control of optical systems had been demonstrated (Mikoshiba & Ahlborn (1973), Bridges *et al.* (1974), Bin-Nun & Dothan-Deutsch (1973), Feinleib, Lipson & Cone (1974), Hardy, Feinleib & Wyant (1974), Muller & Buffington (1974)), either correcting for tilt, focus or a small number of segments/zones; and high-order segmented mirrors and continuously-deformable mirrors (DM's) were in their nascency (Ealey (1989), Ealey & Wheeler (1990), Freeman *et al.* (1977)).

In the 1990's, after military declassification of adaptive optics technology, astronomers used this technology to develop AO systems for more heavenly

purposes. One such low-order system was developed for the Multiple Mirror Telescope (FASTTRAC2) (Lloyd-Hart *et al.* (1997)), and was innovative in the sense that the adaptive element replaced a mirror that was an essential part of the telescope without adaptive optics. In FASTTRAC2, the MMT's 6-facetted (static) beam combiner was replaced with a 6-facetted beam-combiner built by Thermotrex with rapid control of the tip/tilt of each of its facets. By making an essential telescope mirror the adaptive element, a performance-degrading beam train of perhaps 6 additional mirrors before the science camera was avoided. More standard AO systems which have these additional mirrors suffer considerably because of the additional reflection losses, and more importantly for infrared-imaging, the additional thermal emission from each of the uncooled mirrors. This concept of making an essential telescope mirror the adaptive one is currently being developed for the new single 6.5m mirror MMT (Foltz *et al.* (1998), West *et al.* (1996)): this rather bold program (Lloyd-Hart *et al.* (1998b)) replaces the solid $f/D = 15$ secondary with a 2 mm thick, 64 cm diameter thin-shell (high-order) adaptive secondary, with 336 voice-coil actuators behind the thin-shell (see Figure 8). The adaptive secondary is being built by the University of Arizona and an Italian consortium and should see first light in a complete AO system (being built by the University of Arizona and Thermotrex, Inc.) at the MMT at the end of 1999. The main problem that we have encountered in the development of the MMT adaptive secondary is that the deformable mirror is very 'floppy' and requires the damping effects of a thin air gap (40 microns) between the back of the secondary and a solid reference plate in order to eliminate resonances (see Figure 9, as calculated by T. Brinkley at Thermotrex). Another main problem that our Italian collaborators discovered in their prototype tests was that standard local Proportional-Integral-Derivative (PID) control of each actuator was much too slow to actuate high-stress high-order modes, sometimes taking 20-40 milliseconds to reach the commanded state. Therefore, a 'feed-forward' (not to be confused with feed-forward perceptrons) semi-global control algorithm was developed in which the required forces for the next commanded mirror-shape are calculated and applied to all the actuators within 1 millisecond, obviating the slow PID control.

1.7 Wavefront Reconstruction

With each measurement of the atmospherically-aberrated wavefront from a natural or laser guidestar, a high-speed computer needs to calculate the commands to send to the actuators that deform the mirror. The wavefront reconstruction is most standardly accomplished by a matrix multiplication:

$$\boldsymbol{\phi}(t+\tau) = \mathbf{R}\mathbf{s}(t), \tag{5}$$

where $\mathbf{s}(t)$ is a vector of all the wavefront sensor slopes (tip/tilts) (dimension $= S$) for an integration time of τ centered at the latest time t, $\boldsymbol{\phi}(t+\tau)$ is

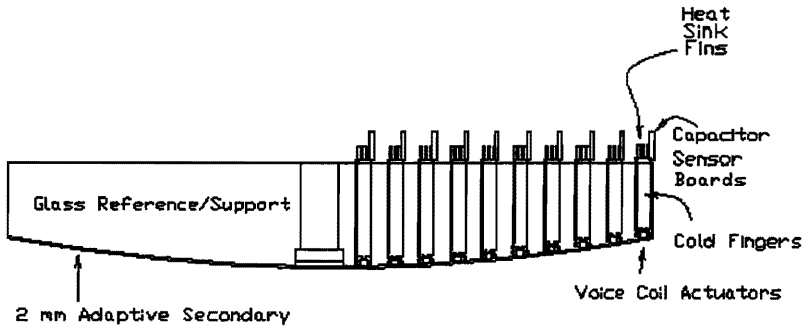


Fig. 8. A sketch showing the adaptive secondary for the MMT upgrade AO system, complete with a 2 mm deformable mirror made of Zerodur (low-expansion glass), voice coil actuators, and a glass reference plate.

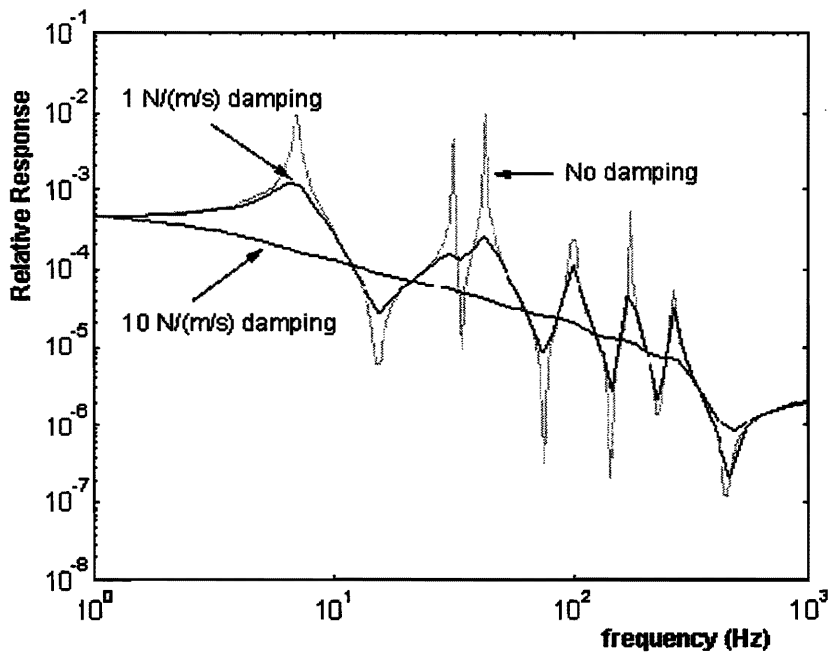


Fig. 9. A frequency response graph for the adaptive secondary shows the abundance of resonances at different sub-kiloHertz frequencies without damping, and how the resonances are eliminated with damping (after T. Brinkley, Thermotrex).

the vector of P actuator commands ('pokes') to be applied to the mirror for the next integration time centered at $t+\tau$, and $\mathbf{R}$ is the $P \times S$ reconstructor matrix that transforms from slope to phase. This reconstruction calculation for *all* the actuators has to be completed within a fraction of the wavefront sensor integration time. Since wavefront sensors for adaptive optics take exposures at a 100-1000 Hz rate, this demands wavefront reconstruction within 0.2-2 milliseconds. For high-order adaptive optics with hundreds of actuators over the aperture and an equal or greater number of wavefront sensor subapertures, such complex calculations continuously demanded within such a short time interval would tax even the fastest general purpose workstation computer. Therefore, special purpose computers with Digital Signal Processing (DSP) boards and matrix multiplication boards are employed to perform this dedicated reconstruction calculation.

In the standard AO scheme, the reconstructor matrix R is calculated prior to the astronomical or military observation and depends only on the subaperture/actuator geometry and a model of the deformable mirror. For the standard Fried geometry with DM actuators at the corners of all the wavefront sensor subapertures, a 'poke' matrix $\mathbf{M}$ is measured in which each matrix element M_{ij} represents the 'influence' of a unit poke by a single actuator j on each of the slope measurements i:

$$M_{ij} \equiv \frac{\delta s_i}{\delta \phi_j}. \tag{6}$$

Usually, these matrix elements are highly local in nature, with large numbers only for subapertures bordering an actuator. For $P \leq S$, the reconstructor matrix $\mathbf{R}$ is most often calculated by the Gaussian inverse of $\mathbf{M}$:

$$\mathbf{R} = \left(\mathbf{M}^T\mathbf{M}\right)^{-1}\mathbf{M}^T. \tag{7}$$

More sophisticated wavefront reconstruction algorithms have been developed which take into account the wind directions, atmospheric turbulence statistics, and measurement noise (Wallner (1983), Wild *et al.* (1995), Wild, Kibblewhite & Vuilleumier (1995), Angel (1994a)), and one of these algorithms employs neural networks (Angel (1994a)), which will be reviewed here. Another neural network algorithm, first developed by Lloyd-Hart (1991) & Lloyd-Hart *et al.* (1992), uses phase-diversity (two images one in-focus and one out-of-focus) as the input data and by training performs both the wavefront-sensing and wavefront-reconstruction in one-step, and will be a focus of this review.

2 Neural Network Wavefront Sensing and Reconstruction

2.1 Motivation

Astronomers currently measure atmospherically-perturbed wavefronts by two main techniques: Shack-Hartmann tip/tilt sensing and Roddier curvature sensing. Both of these techniques operate on the wavefront at a pupil of the telescope, thus demanding reimaging relay optics. A technique, described here, that operates on the *image*-plane data (i.e. the point-spread-function (PSF)) of the telescope would act directly on the quantity of interest to the observer. Additionally, Shack-Hartmann and Roddier wavefront sensing both require wavefront reconstruction from slopes or curvatures to phase. This requires integration over a connected pupil, which is not possible for array telescopes (i.e. the MMT), and therefore one begs for a more direct solution. Neural networks are offered as a solution, and due to their non-linear nature, these neural networks can effectively invert the non-linear transformations (PSF= $|FFT\left(e^{i\phi}\right)|^2$) from pupil-plane to image plane that are needed in phase-retrieval using phase-diversity. In the work described here, supervised-learning (back-propagation) was chosen because of its superiority to unsupervised/self-organizing networks in learning complex functional relationships.

2.2 Widrow-Hoff Wavefront Reconstruction

Angel (1994a) trained a network to transform simulated data from a 6 subaperture Shack-Hartmann quad-cell wavefront sensor (2 derivatives per quad-cell, which gives 12 inputs to the net) of circular/annular/segmented geometry (see Figure 10) for a 4 meter telescope in the infrared (2.2 microns, $r_0 = 1.18$meters) to reconstruct the amplitudes of the 6 pistons needed to minimize wavefront variance. Angel utilized the simplest of networks, without a hidden layer, using linear transfer functions without thresholds on the output layer, with Widrow-Hoff (WH) delta-rule training. When training with significant noise, the WH net's weights were found to be much smaller than without noise, thus allowing the conclusion that the WH net was able to learn that in high-noise situations, it is better to either do nothing or to simply average the data from all the inputs so as to reduce noise and determine global slope. In the absence of noise, Noll (1976) predicts that with 12 measurements and corresponding corrections of the wavefront that for this system, the idealized residual wavefront error would be $0.0339(D/r_0)^{5/3} = 0.26\,\mathrm{rad}^2$, and the WH neural net was found to allow an error of only $0.041(D/r_0)^{5/3} = 0.31\,\mathrm{rad}^2$, which is modest considering that the WH net includes both measurement error and fitting error.

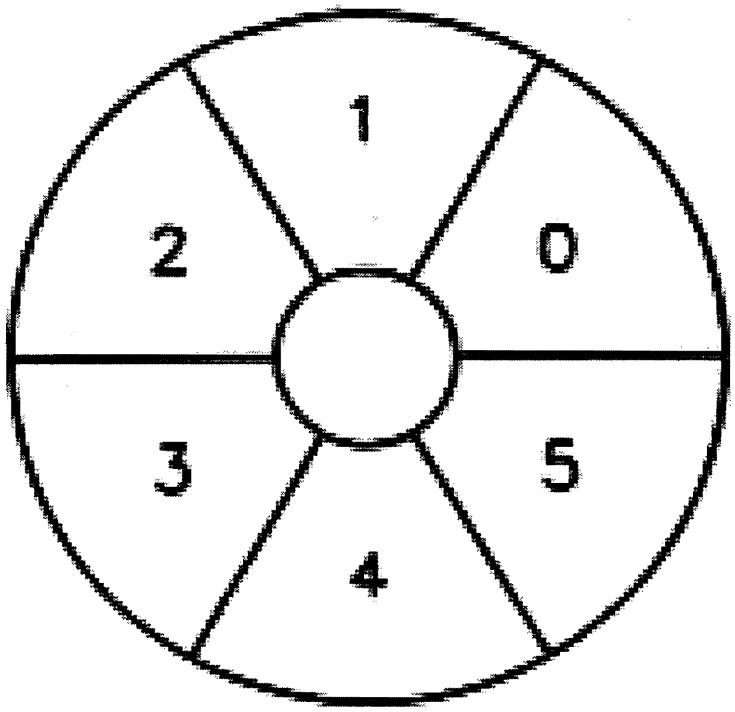


Fig. 10. The subaperture geometry for Angel's Widrow-Hoff linear net (no hidden layer) reconstructor.

2.3 Simulations for the Multiple Mirror Telescope

Angel *et al.* (1990) and Lloyd-Hart (1991) implemented a neural network phase-diversity wavefront sensor and reconstructor for low-order adaptive optics for a simulated Multiple Mirror Telescope. This input layer consisted of 338 input nodes, half being from an in-focus but aberrated 13×13 image of a star at a wavelength of 2.2 μm, and the other half being from a similar, simultaneous but deliberately out-of-focus image of the same star. The output layer of the multilayer perceptron neural network was composed of 18 linear nodes, each representing either tip, tilt or piston for each of the 6 MMT mirrors. And the hidden layer consisted of 150 sigmoidal nodes (see

Figure 11). The training data consisted of 2.5×10^5 image pairs calculated from Kolmogorov turbulence with $r_0 = 1$ m (see Figure 12c-d) on the input and their corresponding best-fit tip,tilt and pistons in the pupil plane for the outputs, and standard momentum-less back-propagation was employed. A comparison of the unmasked pupil-plane data in Figure 12a and the neural network output in Figure 12e shows the correspondence of the net's output to the 'real' world, and a comparison of the uncompensated image seen in Figure 12b with the neural network corrected image in Figure 12f shows the improvement in image resolution and Strehl ratio offered by a phase-diversity neural network. The long-exposure Strehl ratio, obtained by adding 500 simulated speckle images corrected by the trained neural network, was $\mathcal{S} = 0.66$, which is comparable to the Strehl predicted by tilt-corrected Kolmogorov turbulence ($\mathcal{S}_{\text{theory}} = 0.70$). This phase-diverse neural network wavefront sensor and reconstructor (which operates in the infrared (K band)) was found to handle photon-counting noise for stars as dim as 10th magnitude (assuming an infrared detector with 10 photo-electron read noise). Similar simulation results have been reported by Vdovin (1995).

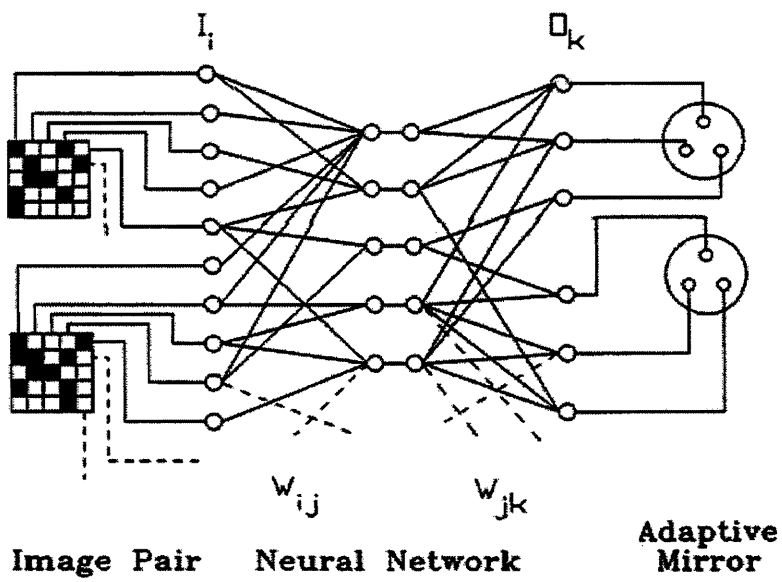
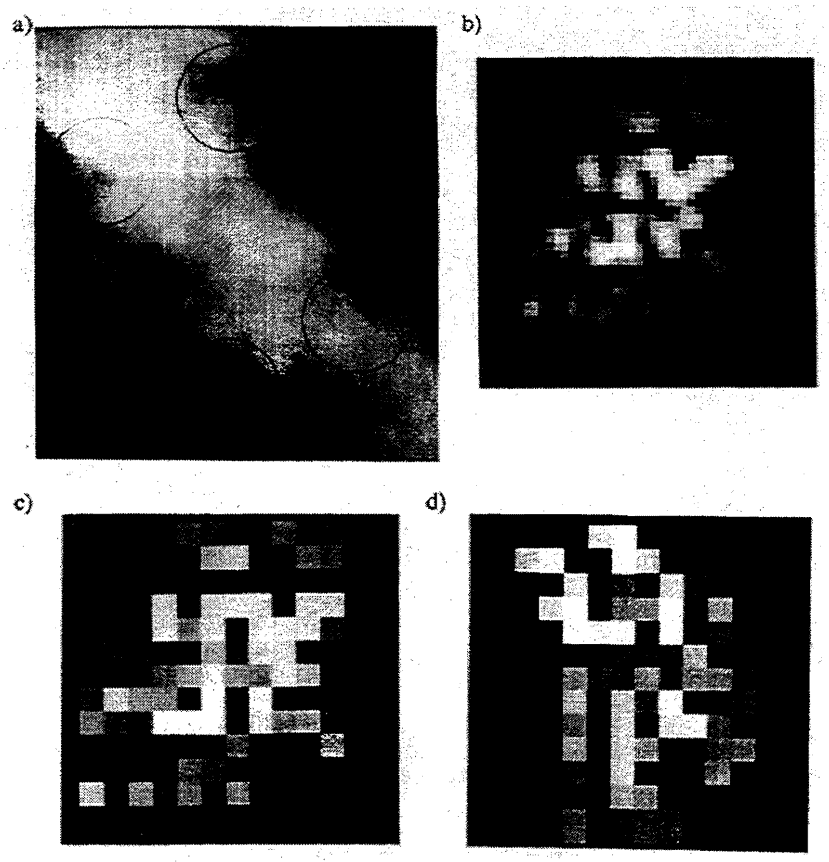


Fig. 11. The neural network used by Lloyd-Hart (1991) to sense and reconstruct the wavefront from phase-diverse MMT data.



2.4 Transputers at the Multiple Mirror Telescope

Lloyd-Hart (1991) and Lloyd-Hart *et al.* (1992) performed experiments that used a neural network with real starlight at both the Steward 2.3 meter telescope on Kitt Peak and at the old MMT on Mt. Hopkins. The experiments with the 2.3 meter continuous aperture telescope demanded the use of a pupil mask and a shift from K-band (2.2μm) to H-band ($1.6\,\mu$m) to simulate the bigger MMT and the MMT's better seeing conditions. Phase-diverse data was then taken at the 2.3 meter telescope, and then later in off-line mode, the pairs of images were shown to the neural net that was trained on simulated data for the MMT, and the net then outputted the wavefront (tips, tilts and pistons) for the six apertures. From this neural net wavefront, speckle images were computed, and found to be in decent agreement with the actual speckle phase-diverse images, with some discrepancies caused by the high sensitivity

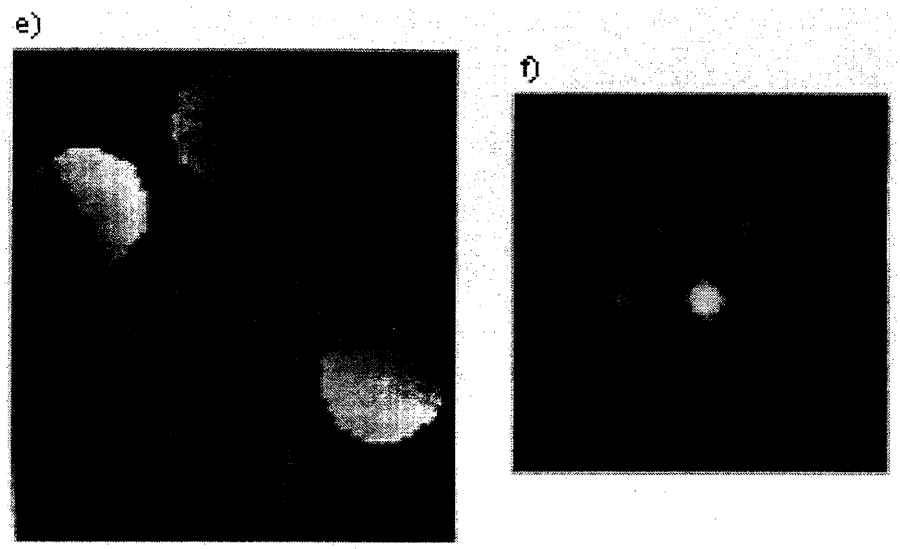


Fig. 12. (after Lloyd-Hart (1991)) Sub-figure a) shows the simulated phase wavefront that serves to distort the image formed by the MMT, as seen in b). Sub-figures c) and d) show the in- and out-of-focus downsampled images, respectively, that serve as the input to the phase-diversity neural net wavefront sensor. Sub-figure e) shows the graphically-depicted output of the neural net (tips,tilts & pistons: compare to the original wavefront in sub-figure a)), and sub-figure f) shows the resulting corrected image, which should be compared with sub-figure b).

of the image calculation to small changes in the nets output. By averaging 388 wavefronts derived by the net from the input image pairs, the mean wavefront was found not to be zero, but to have a significant (about half a wave) defocus aberration. A similar offline experiment was performed at the MMT, with non-phase-diverse infrared focal plane data taken with two of the six MMT mirrors at the same time as temporal visible centroid measurements were made of the images of each of the two mirrors separately. A neural net was then trained on simulated infrared images to output the relative tip/tilt and piston for the speckle image. With the trained neural net, the actual infrared data was passed through it to derive the tip output; when the net's infrared tip output was compared to the measured visible centroid tip, a significant correlation was determined but with 'a fair bit of scatter'. Prior to the completion of the MMT adaptive system, D. Wittman demonstrated, for the first time in the laboratory, closed-loop wavefront correction by a neural network with a HeNe laser and pin hole and a two-mirror aperture mask. Ten thousand phase-diverse random images were presented for training

to the net by randomly changing the piston, tip and tilt of the 2 element segmented mirror, and one thousand random images were used for testing. The control of the neural net effectively cancelled the random user-imposed mirror deformations, improving a very poor integrated image to one with a superb Strehl ratio of 0.98, or a wavefront variance of $0.02\,\mathrm{rad}^2$.

A six-element adaptive system was built for the MMT in 1991, complete with transputers to serve as the hardware for a neural net that sensed the wavefront from a pair of phase-diverse images and controlled the wavefront with the adaptive mirror (see Figure 13). The neural net was trained on 30,000 image pairs taken in the laboratory in Tucson with the adaptive instrument itself supplying known wavefront aberrations. Of course, the adaptive instrument was unable to aberrate the wavefront on a scale smaller than the subaperture size, so 10,000 computer simulated images with fine-scale turbulence *and* fixed trefoil MMT aberrations were used to further train the network. With 16 Inmos T800-25MHz transputer modules serving as the neural net, training was accomplished in 33 minutes. Since the neural net's feed-forward pass takes 160,000 floating point operations, and since the integration and read out time of the two 26×20 sub-arrays of the infrared detector is 9.4 milliseconds, a machine capable of sustaining computations at 17 Mflops was required; the transputer array was benchmarked at a sufficient 25 Mflops. Lloyd-Hart took 15×10 subframes that included most of the energy from the pair of 26×20 images from the infrared camera. Therefore the network had 300 inputs. After the wavefront is calculated by the transputer net, global tilt was subtracted from the wavefront to ensure that the image always stays centered, and the mean phase was subtracted so that the actuators stay centered in its range. Prior to the run, nets were trained on laboratory data, using either 36 or 54 hidden neurons, and 6 or 10 output neurons, depending on whether 2 or 3 telescope apertures were used. An extra output was found useful, so that instead of predicting phase which has a discontinuity at π, nets outputting the continuously-valued sine and cosine of the phase were trained. A three mirror net was trained with laboratory data to a phase error of $0.05\,\mathrm{rad}^2$ after 60,000 laboratory image pairs (50 minutes), but a net with 22 outputs for control of 6 mirrors did not achieve a low enough error, which was probably due to a local minimum. Neural network real-time adaptive control of *two* of the mirrors of the MMT was achieved in 1991, with training in the Tucson lab. For most of the 1991 observing run, the seeing was very poor (1 arcsecond at 2.2μm which is equivalent to an r_0 of 45cm, half the normal value), and since the net was trained for data of the normal character, sky correction with the neural network was impossible. When seeing improved later in the run, a two-mirror net achieved the first on-sky success in adaptive correction of atmospheric phase distortion on the image of the bright star ψ Pegasi (K magnitude of 0.0), see Figure 14, improving the seeing-limited resolution from 0.62 arcseconds to a diffraction-limited 0.1 arcseconds. By taking the summed product of the hidden and output weights

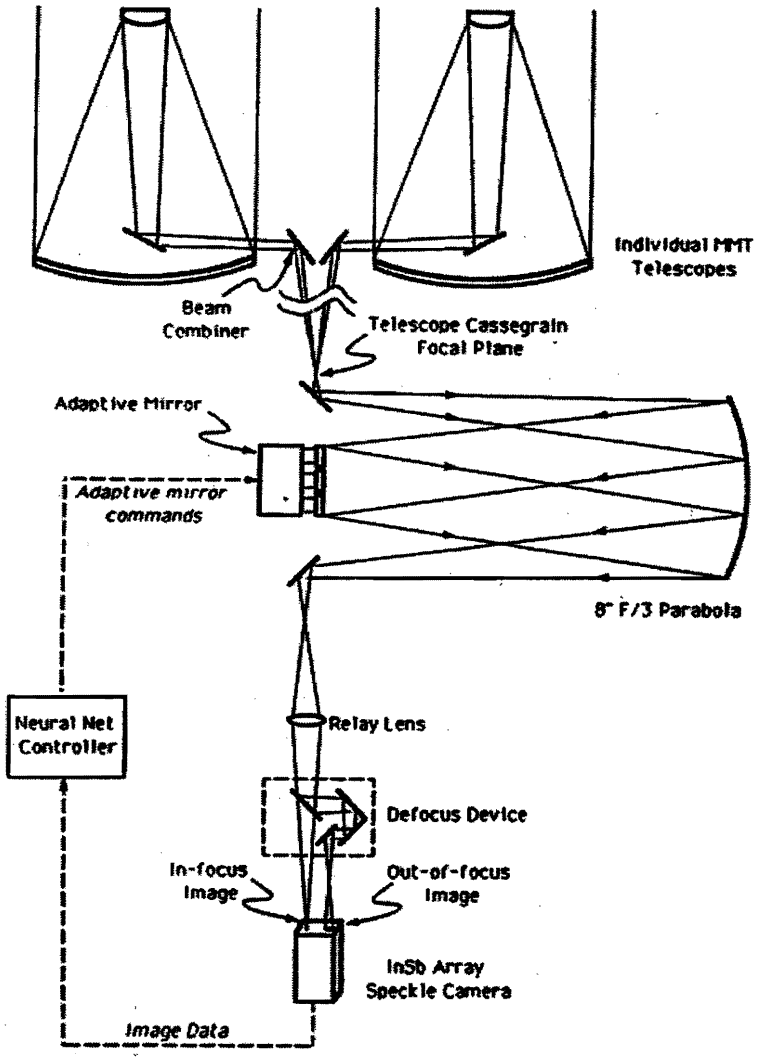


Fig. 13. Complete AO system (after Lloyd-Hart *et al.* (1992)) including the phase-diversity neural network wavefront sensor, as implemented at the MMT.

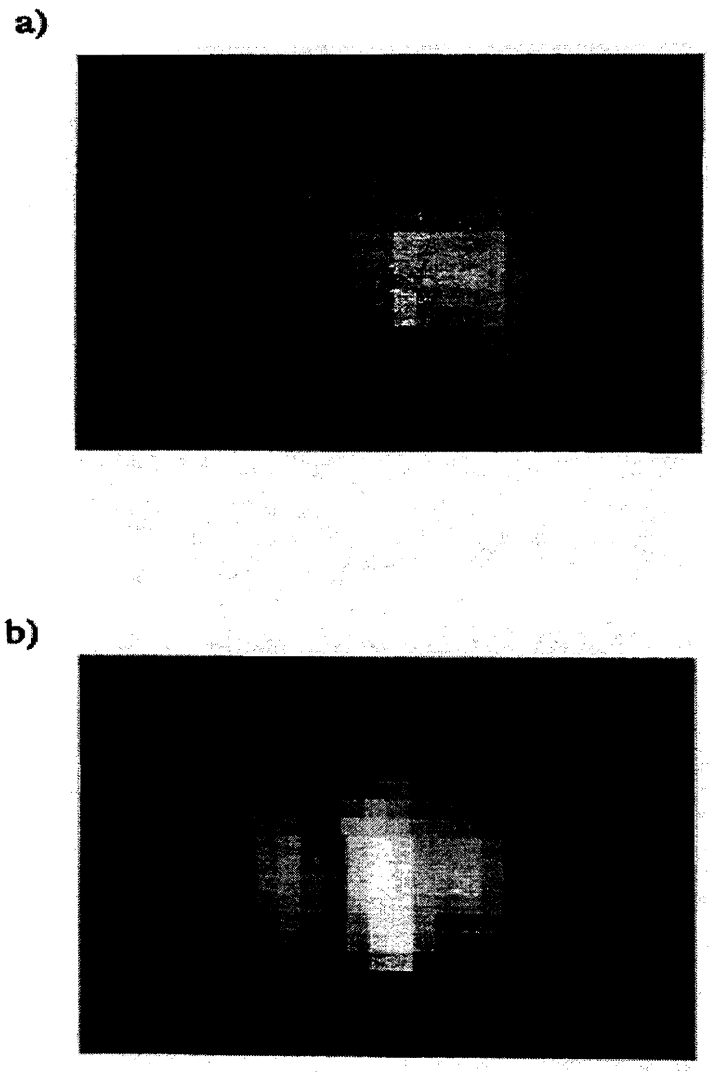


Fig. 14. (After Lloyd-Hart (1991)) With and without neural-network AO control of two MMT primary mirrors. The top figure shows the star ψ-Pegasi at 2.2μm wavelength without AO correction during a 10 second exposure (1000 10 millisecond images co-added) with only 2 of the 6 MMT primaries uncovered, giving a resolution of FWHM= 0.62 arseconds, and a Strehl ratio of 0.15. The bottom figure shows the same star with neural net phase-diversity 67 Hz closed-loop AO correction, giving a fringe resolution of 0.1 arcseconds, and a Strehl ratio of 0.27.

($V_{ij} = \Sigma_j W_{ij} W_{jk}$), Sandler (1991) claimed to see the Zernike modes reflected in this spatial product for his work on a continuous aperture neural network (Sandler *et al.* (1991)). Lloyd-Hart (1991) performed the same analysis and found encouraging results, shown in Figure 15. From the bottom to top of Figure 15, the panels represent V_{ij} for the tip and tilt for mirror A, tip and tilt for mirror F, and the sine and cosine of the phase; the images on the left side are for the in-focus image, and the images on the right side are for the out-of-focus image. The tip/tilt images of mirror A appear to be negative images of the tip/tilt images of mirror F, which is expected because positive tip on one mirror causes relative negative tip on the other mirror. These patterns, like the Zernike patterns in the weights seen for Sandler's continuous aperture net, are only expected in the case of a net with linear transfer functions on the hidden units. Since sigmoidal transfer functions were used, the appearance of this dominant structure in the weights is somewhat of a mystery. Perhaps the nets are not utilizing their non-linearities extensively.

2.5 Simulations for Single-Mirror High-Order Adaptive Optics Systems

Sandler *et al.* (1991), under the auspices of the Air Force Phillips Laboratory, at Starfire Optical Range, implemented a similar phase-diversity neural network wavefront sensor and reconstructor, but for a single-mirror (continuous aperture) and with data obtained from phase-diverse measurements of a *real* atmosphere (and a real star). Historically this work was performed prior to and was the inspiration for the work at the MMT outlined in the last two sections (despite similar publication dates). The 16×16 (I-band: $\lambda = 0.85\mu$m) images of Vega had 0.412-arcsecond pixels, an integration-time of 2 milliseconds, and were taken with the 1.5m Starfire Optical Range telescope in New Mexico; and the out-of-focus image was 1.1μm out-of-focus. A Shack-Hartmann wavefront sensor accumulated slope wavefront data simultaneously for later comparison to the wavefront derived off-line from the neural network which was trained with simulated data. The perceptron neural network consisted of 512 input neurons, a sigmoidal hidden layer, and 8 linear output neurons (see Figure 16). Each of the 8 outputs was trained via standard back-propagation to respond to the two input images (see the bottom pair of images in the inset of Figure 17) with the amplitude of a different Zernike mode's contribution to the atmospheric distortion. (Zernike polynomials (Born & Wolf (1975)) are like Hermite polynomials in that they are orthogonal, but they are applied to a unit disk, giving products of angular functions and radial polynomials (e.g., focus, astigmatism, coma, spherical aberration)). The neural-net reconstructed Zernike amplitudes were then used to derive a phase-map of the total atmospheric aberration (by summing all the Zernike modes, weighted by the NN amplitudes), and in Figure 17 for *one* atmospheric realization, the resulting contour-like NN phase-map ('interferogram') can be compared with the phase-map derived by the

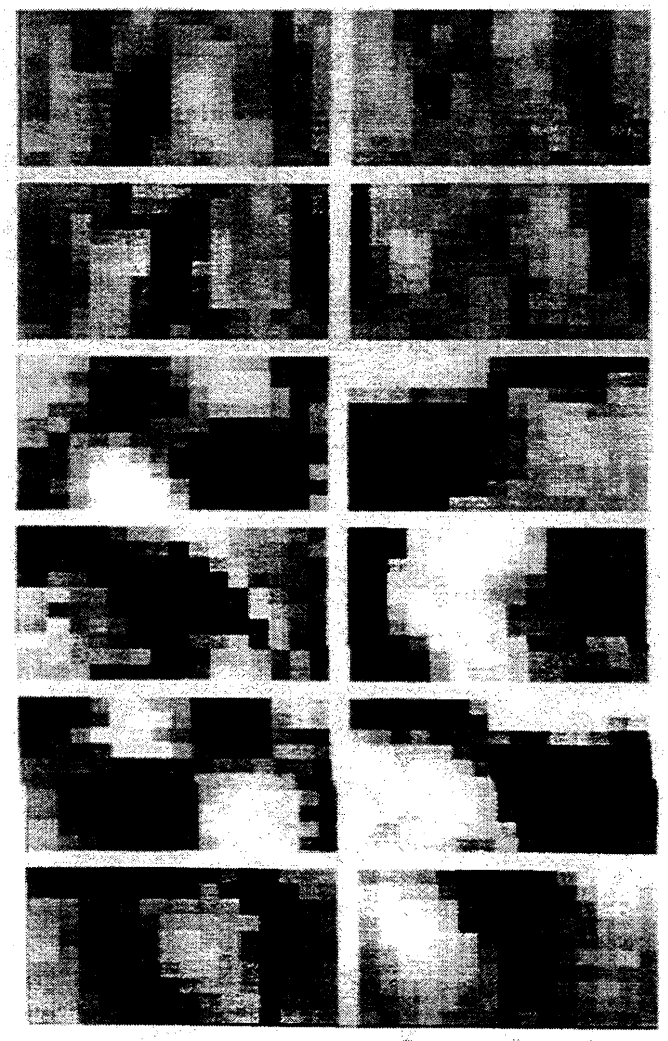


Fig. 15. The product of the hidden and output weights for the phase-diversity neural net for the MMT (after Lloyd-Hart *et al.* (1992)). The left column corresponds to the in-focus images, and the right column corrsponds to the out-of-focus images. From bottom to top, the six panels represent tip and tilt for mirror A, tip and tilt for mirror B, and the sine and cosine of the relative phase between the two mirrors. These images represent the filters which would be applied by the outputs to the input image pair if the net were linear.

Shack-Hartmann wavefront sensor and a standard reconstructor. Sandler *et al.* (1991) analyze 107 independent snapshots of the atmosphere with the off-line phase-constructing neural net, and find that the neural net would improve the wavefront error to 0.78 rad^2 (the uncorrected wavefront error was 1.77 rad^2), thus improving image resolution by a factor of 3. The use of neural networks for phase diversity wavefront sensing has been described in detail in Barrett & Sandler (1996). This same technique was applied by Barrett & Sandler (1993) to Hubble Space Telescope (HST) data before the Hubble optics were fixed. Using the real HST stellar images at different focal positions, the NN technique was compared to slow off-line Fourier based phase-retrieval methods, and the two methods were found to predict basically the same amount of spherical aberration.

3 Neural Network Wavefront Prediction

Wavefront sensor measurements of the atmospheric distortion need to be made with CCD exposures that have non-zero exposure time, in order to collect sufficient photons to characterize the wavefront. For typical laser guidestars of visual magnitude 10, exposure times of 1 millisecond will give about 600 photons per $0.5 \times 0.5\,\mathrm{m}^2$ subaperture, which gives high (but not infinite) signal-to-noise ratios. Laser guidestar exposure times much shorter than 1 millisecond would diminish performance. If a laser guidestar is not available, or if the researcher wishes to avoid laser guidestar focus anisoplanatism, then a natural guidestar may be used for the adaptive reference. In this case, the natural guidestar may not be optimally bright (dimmer than magnitude 10), so in order to maximize wavefront sensor signal-to-noise ratio, longer exposures are required. For a magnitude 12 natural guidestar, there would be about 475 photons per $0.5 \times 0.5\,\mathrm{m}^2$ subaperture for 5 millisecond exposures. With these non-zero exposure times and the additional time needed to reconstruct the wavefront and apply the actuator commands (0.5-2.5 milliseconds, using eight C40 digital signal processors), winds will have blown the 'old' measured turbulence some fraction of a subaperture across the pupil and new turbulence will take its place, either coming in from outside the aperture or from another subaperture. Concomitantly, the turbulent phase-screen may not remain static (as is assumed by Taylor's frozen flow hypothesis) due to the relative motion of turbulent eddies or due to the relative motion of the multiple turbulent layers of the atmosphere. For a single-layer atmosphere, blowing across at possible jet stream speed of 20m/s, the turbulent phase-screen will have shifted by 0.1 meters in a wavefront sensor CCD exposure time of 5 milliseconds, which is potentially a change of 20% in the turbulent phase-screen above a given $0.5 \times 0.5\,\mathrm{m}^2$ subaperture. For a dynamic atmosphere, the phase-variance between one wavefront and co-located wavefront t milliseconds later increases by $\sigma_t = (t/\tau)^{5/3}$, where $\tau = 0.31 r_0/v_w$, and v_w is the turbulence-weighted wind-velocity. In the infrared (e.g. $\lambda \sim 2.2\mu\mathrm{m}$),

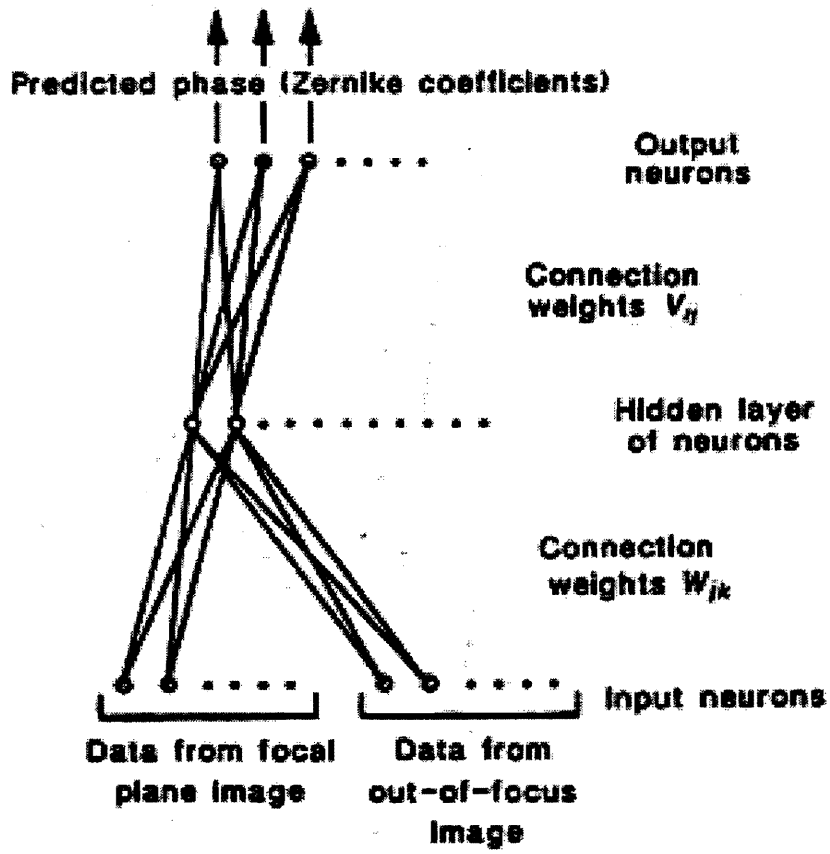


Fig. 16. A diagram showing the neural network used to recover phase for the continuous aperture phase-diversity telescope data (after Sandler *et al.* (1991)).

where the science is done, the Fried coherence length is $r_0 \sim 1\,\mathrm{m}$ and the wavefront decorrelation time is $\tau \sim$ 16milliseconds. Therefore, in the infrared, there would be an 'unavoidable' temporal variance of $\sim 0.14\,\mathrm{rad}^2$ for 5 millisecond exposures, unless auxiliary measures are employed. This decorrelation is a factor of 4 higher than for 1 millisecond exposures and is similar in magnitude to the fitting and reconstrution errors (Sandler *et al.* (1994b)),

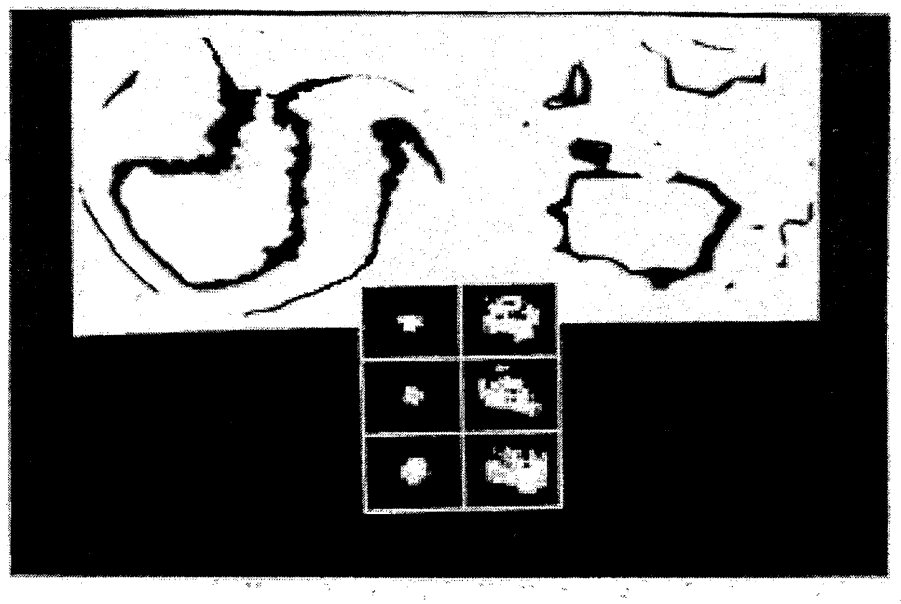


Fig. 17. An example of neural network performance (after Sandler *et al.* (1991)), showing three pairs of in- and out-of-focus far-field patterns, and two interferograms. The lowest pair of images is the actual camera data obtained at Starfire Optical Range and which are then input to the neural network. The middle set of images is produced by numerically simulating the in- and out-of-focus stellar images using the phase predicted by the neural network. The top set of images is produced by numerically simulating camera data corresponding to the phase reconstructed from the Shack-Hartmann sensor. The interferogram on the left represents the phase predicted by the neural network, and the interferogram on the right represents the phase measured by the Shack-Hartmann sensor.

so temporal decorrelation begins to become significant for these 5 millisecond exposures that might be demanded for AO with magnitude 12 natural guidestars.

Of course, the obvious solution to this temporal decorrelation and servo lag problem is the application of predictive technology to anticipate the wavefront changes that occur during the CCD exposures and actuator/mirror control. In order to apply prediction, the atmosphere needs to be predictable, not random in nature – indeed, Jorgenson, Aitken & Hege (1991) showed that centroids of a star's position followed a chaotic temporal trajectory with correlation dimension near 6. This implies that the atmosphere has a deter-

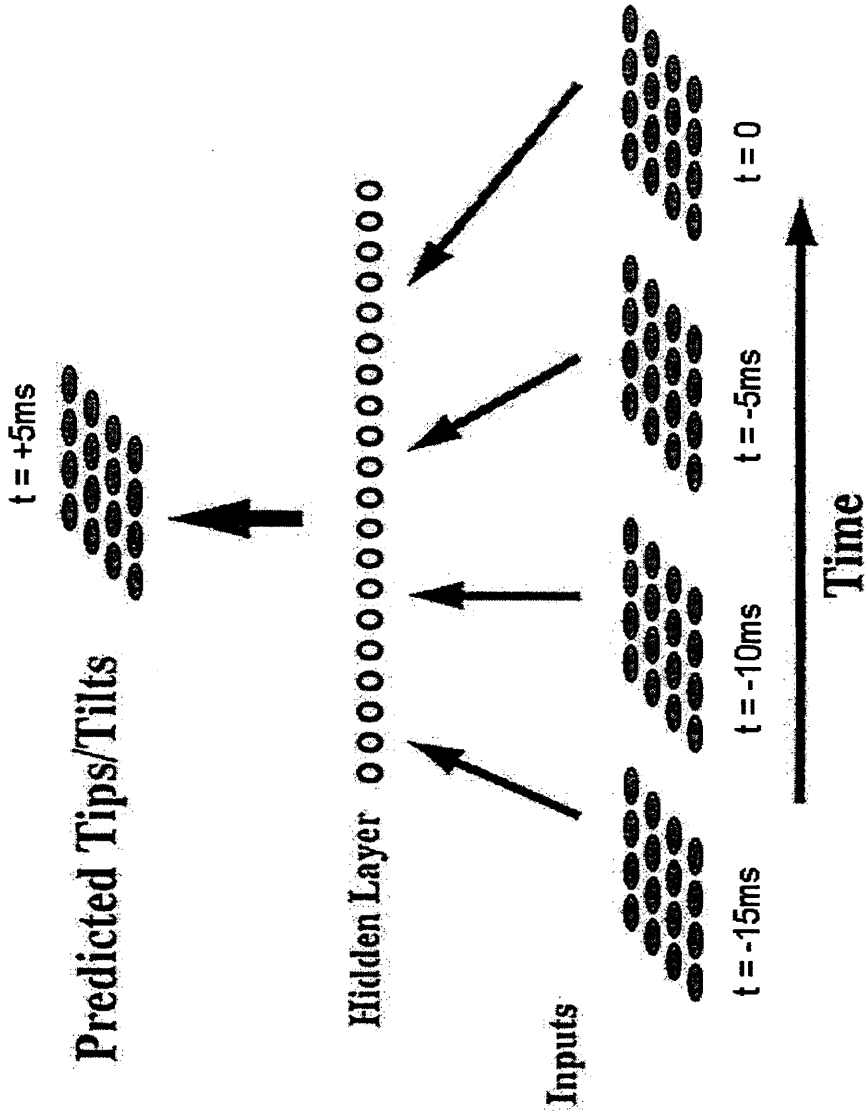


Fig. 18. A figure (adapted from Jorgenson & Aitken (1994)) showing the neural and linear network architecture used to predict wavefront slopes for a simulated 2 meter telescope into the future using the past arrays of slopes. For this case, there are 128 inputs, 60 hidden units, and 32 outputs (each wavefront sensor pixel corresponds to an x-tilt and a y-tip).

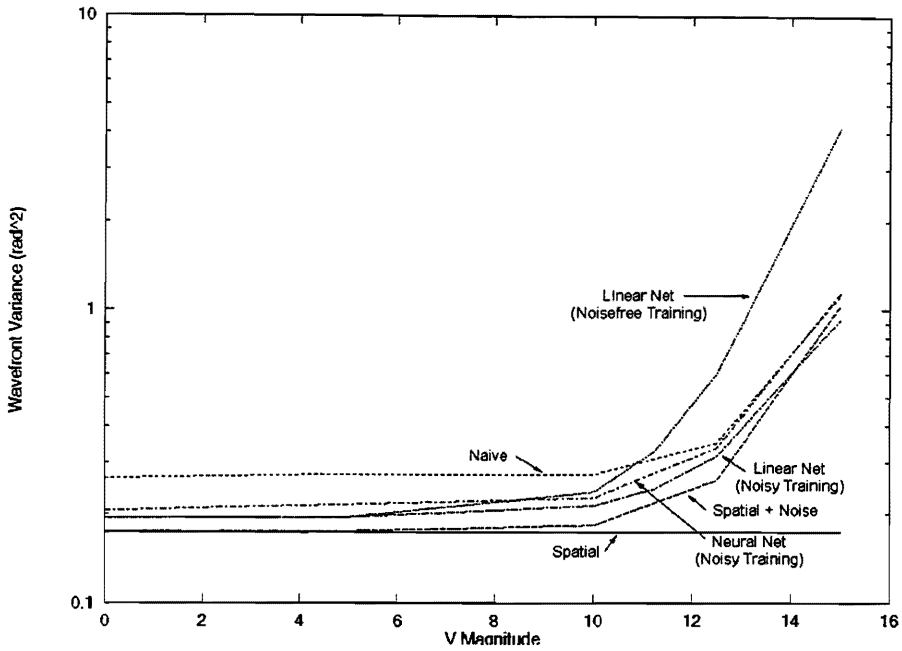


Fig. 19. A comparison of the prediction wavefront variance results of both the neural net and the linear net, after training for 178,000 5-millisecond Shack-Hartmann snapshots of a simulated $\tau = 14$ millisecond atmosphere. The bottom curve shows the simulated spatial (reconstructor and fitting) errors, the next curve adds photon-counting noise to the spatial error. The top curve shows the "naive" prediction error, computed by predicting that the slope will not change from one timestep to the next. And the remaining three curves are the predictor curves, discussed in the text. The performance is shown as a function of the guidestar brightness or magnitude, the dimmer guidestars being on the right.

ministic component that can be predicted, and thus some of the temporal decorrelation of the wavefront can be negated.

Several groups set out to use the past history of the measured wavefront slopes to predict the wavefront for the next WFS exposure. Jorgenson & Aitken (1994) used two-dimensional wavefront data from the COME-ON AO system in a neural network that incorporated both temporal and *spatial* information to predict the future array of slopes. In good seeing their wavefront variance with neural network prediction was a factor of two better than without prediction, with equivalent preformance for their linear predictors. For poorer seeing and longer-term predictions, the neural net predictor's wavefront variance was half that of the linear predictor and one-sixth that of no

prediction. Lloyd-Hart & McGuire (1996) also thoroughly investigated the efficacy of least-squares linear predictors for both simulated and real data as a function of lookback depth, signal-to-noise ratio and turbulence timescale, and they also did an initial foray into neural network prediction, capitalizing on McGuire's experience in currency futures market prediction with neural networks. The neural network was found to perform many times better than the linear predictors for low signal-to-noise experiments after the predictors had been trained without noise. Training without noise was the lore McGuire learned in his finance prediction studies, as well as the histogram equalization preprocessing used successfully in the neural network studies of Lloyd-Hart & McGuire (1996). The histogram equalization did *not* work well for the linear least-squares prediction. Other groups that have worked on spatiotemporal modelling and prediction of the atmosphere for AO include Schwartz, Baum & Ribak(1994), Bonaccini, Gallieni, & Giampieretti (1996), and Dessenne, Madec, & Rousset (1997).

Aitken & McGaughey (1996) have published significant work suggesting that the atmosphere is linear filtering process of Brownian motion with a Hurst exponent of 5/6, and that the predictability stems from the spatial averaging done by the wavefront sensor. They suggest that linear predictors "should be able to extract *all* of the available information of the linear process", despite evidence that non-linear predictors such as neural nets perform better than linear predictors. Part of the rationale for the new work that we present here is to study better the differences between linear and non-linear predictors, spurred by Aitken & McGaughey's result of linearity.

We follow Jorgenson & Aitken (1994) in the basic network architecture, but use only a single hidden layer. We use the past four $2 \times 4 \times 4$ arrays of open-loop slopes from a simulated 2 meter telescope and AO system to predict the next array of x & y slopes 5 milliseconds into the future (see Figure 18). For our neural network, we have linear transfer functions on the outputs and sigmoidal transfer functions on the hidden units; for our linear network, the hidden units are made linear. MinMax preprocessing is used for the input and output data, due to its linear nature and the linear output transfer functions (in Lloyd-Hart & McGuire (1996), sigmoids were used on the output layer, which allowed the non-linear preprocessing ("Uniform preprocessing") histogram equalization to work well). Standard back-propagation was the chosen training method for both the linear and non-linear nets, using 178,000 5 millisecond Shack-Hartmann wavefront sensor snapshots of a simulated atmosphere with a correlation time of $\tau = 14$ milliseconds, and guidestars of varying magnitudes. This training finally arrived at wavefront variance values which we depict in Figure 19 (for $\sigma^2 < 1$, the Strehl ratio is $\mathcal{S} \approx \exp(-\sigma^2)$). The bottom curve is the combination of spatial errors like reconstructor error and fitting error; when noise is added, we arrive at the next curve; and without any prediction we arrive at the curve that we call the "naive prediction" error, which is the error resulting from the servo lag that assumes

that the best prediction is 'the slopes will not change' that comes from assuming a random-walk type evolution. The intermediate curves are those with prediction. Noiseless training with noisy recall performed much worse for the linear predictor than noisy training with noisy recall, with errors that are almost twice the naive prediction error for a magnitude 15 guidestar. The linear net predictor with noisy training does better than the non-linear neural net predictor for all levels of noise (the higher the guidestar magnitude, the more photon counting noise). In high noise situations, the linear net predictor even does better than the spatial and noise errors combined (which might have been considered a floor in the possible perfomance). Both the linear net and the neural net predictors outperform the naive prediction at low noise levels, but at high noise levels only the linear net's performance is admirable, with the neural net's wavefront error approaching the naive prediction error. The sub-noise-level linear net performance means that the linear net is successfully temporally averaging out the noise. The outperformance of the neural net by the linear net may be caused by several things, first of which is the natural accordance of the linear net with McGaughey & Aitken's linear atmosphere, but possibly also the choice of linear output transfer functions and the decision not to use histogram equalization as in previous studies. We have also attempted recursive linear least squares (RLS) prediction (Strobach (1990)), but for this simulated data set, performance was poor compared to our linear net and our neural net. However, RLS prediction performance for a real data obtained from a high-order AO system on the 1.5 meter telescope at Starfire Optical Range Rhoadarmer (1996) was *much* better than our initial attempts at neural net prediction for this much higher dimensional real data set.

4 Future Work

Further network prediction studies, both linear and non-linear are warranted, and a careful comparison with other techniques (e.g., Wild (1996), Dessenne, Madec, & Rousset (1997), RLS and non-RLS) should be undertaken. The dimensionality curse needs solution, in order to extend the predictive networks' domain of excellence from low-order AO systems to high-order AO systems. This solution requires speeding up the training by 1-2 orders of magnitude, or it requires limiting the size of the networks by using only local, instead of global, neighborhoods of past influence. Once the dimensionality problem is solved, then the predictive networks may be of use in real-time observations at the telescope. Such real-time observations will require the development of these networks in our Adaptive Solutions VME computer or perhaps in a more sophisticated dedicated computer.

The phase-diversity wavefront-sensing neural networks are an inexpensive and rapid phase retrieval method that can be applied to both static optical deformation determination, as well as the rapid atmospheric deformation

sensing. It requires only two sets of images to determine the wavefront, and so it can be applied at almost any telescope with little difficulty. With a deformable mirror in the system, the phase-diversity neural network will form a complete, inexpensive adaptive optics system perhaps capable of the ultra-high-order adaptive correction (that may allow extra-solar planet imaging) currently in adaptive opticians' plans and dreams (Angel (1994b), Stahl & Sandler (1995)).

4.1 Acknowledgements

The new work outlined in Section 3 was supported by the US Air Force Office of Scientific Research under grants #F49620-94-1-0437 and #F49620-96-1-0366. The authors thank G. Angeli, J.R.P. Angel, C. Shelton, and A. Krogh for helpful discussions, R.Q. Fugate for collaboration at SOR, and Angeli and Angel for reviewing this paper.

References

G.J.M. Aitken & D. McGaughey (1996), "Predictability of Atmospherically-distorted Stellar Wavefronts", *Proc. European Southern Observatory Conf. on Adaptive Optics*, **54** Ed. M. Cullum, Garching, Germany, p. 89.

J.R.P. Angel, P. Wizinowich, M. Lloyd-Hart, & D.G. Sandler (1990), "Adaptive Optics for Array Telescopes using Neural network Techniques", *Nature*, **348**, 221.

J.R.P. Angel (1994a), "Wavefront Reconstruction by Machine Learning Using the Delta Rule", *Proc. SPIE Conf. on Adaptive Optics in Astronomy*, **2201**, Kona, Hawaii, 629.

J.R.P. Angel (1994b), "Ground Based Imaging of Extrasolar Planets Using Adaptive Optics", *Nature*, **368**, 203.

K. Avicola, J.M. Brase, J.R. Morris, H.D. Bissinger, H.W. Friedman, D.T. Gavel, R. Kiefer, C.E. Max, S.S. Olivier, D.A. Rapp, J.T. Salmon, D.A. Smanley, & K.E. Waltjen (1994), "Sodium Laser Guide Star System at Lawrence Livermore National Laboratory: System Description and Experimental Results" *Proc. SPIE Conf. on Adaptive Optics in Astronomy*, Eds. M.A. Ealey, F. Merkle **2201**, 326.

T.K. Barrett & D.G. Sandler (1993), "Artificial Neural Network for the Determination of the Hubble Space Telescope Aberration from Stellar Images", *Appl. Optics*, **32**, 1720.

T.K. Barrett & D.G. Sandler (1996), "Neural Networks for Control of Telescope Adaptive Optics", in *Handbook of Neural Computation*, Eds. E. Fiesler & R. Beale, Oxford University Press, p. G3.1:1.

J.M. Beckers (1993), "Adaptive Optics for Astronomy: Principles, Performance, and Applications", *Annu. Rev. Astron. Astrophys.*, **31**, 13.

E. Bin-Nun & F. Dothan-Deutsch (1973), *Rev. Sci Instrum.*, **44**, 512.

D. Bonaccini, D. Gallieni, & R. Giampieretti (1996), "Prediction of Star Wander Path with Linear and Non-Linear Methods", *Proc. European Southern Observatory Conf. on Adaptive Optics*, **54** Ed. M. Cullum, Garching, Germany, p. 103.

M. Born & E. Wolf (1975), *Principles of Optics*, 5th edition, Pergamon Press, Oxford.

W.D. Bridges, P.T. Brunner, S.P. Lazzara, T.A. Nussmeier, T.R. O'Meara, J.A. Sanguinet, & W.P. Brown (1974), *J. Appl. Opt.*, **13**, 2.

B.J. Carter, E.J. Kibblewhite, W.J. Wild, J.R.P. Angel, M. Lloyd-Hart, B.P. Jacobsen, D.M. Wittman, & J.W. Beletic (1994), "Sodium Beacon used for Adaptive Optics on the Multiple Mirror Telescope", *Proc. NATO/ASI School on Adaptive Optics for Astronomy*, Eds., D.M.Alloin, J.-M. Mariotti, Cargese, Corsica.

R.W. Cohen, T.S. Mast, & J.E. Nelson (1994), "Performance of the W.M. Keck Telescope Active Mirror Control System", *Proc. SPIE Conference on Advanced Technology Optical Telescopes*, Ed. L.M.Stepp, **2199**, 105.

C. Dessenne, P.Y. Madec, G. Rousset (1997) "Modal Prediction for Closed-Loop Adaptive Optics", *Optics Letters*, **22**, 1535.

M.A. Ealey (1989), Proc. Soc. Photo-opt. Instrum. Eng., **1167**.

M.A. Ealey & C.E. Wheeler (1990), *Proc. 3rd Int. Cong. Opt. Sci. and Eng.*, **1271**, Paper 23.

J. Feinleib, S.G. Lipson, & P.F. Cone (1974), *Appl. Phys. Lett.*, **25**, 311.

C. Foltz *et al.* (1998), http://sculptor.as.arizona.edu/foltz/www/mmt.html

R.H. Freeman, H.R. Garcia, J. DiTucci, & G.R. Wisner (1977), *High-Speed Deformable Mirror System*, U.S. Air Force Weapons Laboratory Report AFWL-TR-TR-76-146.

D.L. Fried (1965),"The Effect of Wavefront Distortion on the Performance of an Ideal Optical Heterodyne Receiver and an Ideal Camera", *Conf. on Atmospheric Limitations to Optical Propagation*, U.S. Nat. Bur. Stds. CRPL.

D.L. Fried (1994), "Atmospheric Turbulence Optical Effects: Understanding the Adaptive-Optics Implications", *Proc. NATO/ASI School on Adaptive Optics for Astronomy*, Eds. D.M. Alloin & J.M. Mariotti, Cargese, Corsica, p.25.

H.W. Friedman, G.V. Erbert, T.C. Kuklo, J.T. Salmon, D.A. Smauley, G.R. Thompson, J.G. Malik, N.J. Wong, K. Kanz, & K. Neeb (1995), "Sodium beacon laser system for the Lick Observatory" *Proc. SPIE Conf. on Adaptive Optical Systems and Applications*, Eds. R.K. Tyson & R.Q. Fugate, **2534**, 150.

H. Friedman, D. Gavel, C. Max, G. Erbert, G. Thompson, T. Kuklo, N. Wong, M. Feldman, B. Beeman, & H. Jones (1997), *Laser Guidestar System: Diagnostics and Launch Telescope Critical Design Review Report.*

R.Q. Fugate, L.M. Wopat, D.L. Fried, G.A. Ameer, S.L. Browne P.H. Roberts, G.A. Tyler, B.R. Boeke, & R.E. Ruane (1991), "Measurement of Atmospheric Wavefront Distortion using Scattered Light from a Laser Guide-star", *Nature*, **353**, 144.

J. Ge (1998), "Sodium Laser Guide Star Technique, Spectroscopy and Imaging with Adaptive Optics", Ph.D. dissertation, University of Arizona.

D.P. Greenwood (1977), "Bandwidth Specifications for Adaptive Optics Systems", *J. Opt. Soc. Am.*, **67**, 390.

J.W. Hardy, J. Feinleib, & J.C. Wyant (1974), *Real-Time Correction of Optical Imaging Systems*, OSA Meeting on Optical Propagation through Turbulence, Boulder, CO.

W. Happer (1982), Princeton U. & member of JASON's advisory group, private report; W. Happer, G.J. MacDonald, C.E. Max, & F.J. Dyson (1994), *J. Opt. Soc. Am.*, **A11**, 263.

J. Hartmann (1900), *Zt. Instrumentenkd.*, **24**, 1.

J.M. Hill & J.R.P. Angel (1992), "The Casting of the 6.5m Borosilicate Mirror for the MMT Conversion", http://medusa.as.arizona.edu/lbtwww/tech/mirror.htm, *Proceedings of the ESO Conference on Progress in Telescope and Instrumentation Technologies*, Garching, Germany, ed. M-H. Ulrich, 57.

R.E. Hufnagel (1974), *Proc. Topical Mtg. on Opt. Propagation through Turbulence*, Boulder, CO.

B.P. Jacobsen (1997), "Sodium Laser Guide Star Projection for Adaptive Optics", Ph.D. dissertation, University of Arizona.

M.B. Jorgenson, G.J.M. Aitken, & E.K. Hege (1991), "Evidence of a Chaotic Attractor in Star-wander Data", *Optics Letters*, **16**, 2.

M.B. Jorgenson & G.J.M. Aitken (1994), "Wavefront Prediction for Adaptive Optics", *European Southern Observatory Conf. on Active and Adaptive Optics*, Ed. F. Merkle, Garching, Germany, p. 143.

E. Kibblewhite (1997), "Physics of the Sodium Layer", *Proc. of the NATO/ASI School on Laser Guide Star Adaptive Optics for Astronomy*, Ed. J.C.Dainty, Cargese, Corsica.

A. Kolmogorov (1961), in *Turbulence, Classic Papers on Statistical Theory*, Eds. S.K.Friedlander & L. Topper, Interscience, New York.

M. Lloyd-Hart (1991), "Novel Techniques of Wavefront Sensing for Adaptive Optics with Array Telescopes Using and Artificial Neural Network", Ph.D. Dissertation, University of Arizona.

M. Lloyd-Hart, P. Wizinowich, B. McLeod, D. Wittman, D. Colucci, R. Dekany, D. McCarthy, J.R.P. Angel, & D. Sandler (1992), "First Results of an On-line Adaptive Optics System with Atmospheric Wavefront Sensing by an Artificial Neural Network", *Astrophysical Journal*, 390, L41-44.

M. Lloyd-Hart & P.C. McGuire (1996), "Spatio-Temporal Prediction for Adaptive Optics Wavefront Reconstructors", *Proc. European Southern Observatory Conf. on Adaptive Optics*, **54** Ed. M. Cullum, Garching, Germany, p. 95.

M. Lloyd-Hart, R. Angel, T. Groesbeck, P. McGuire, D. Sandler, D. McCarthy, T. Martinez, B. Jacobsen, T. Roberts, P. Hinz, J. Ge, B. McLeod, G. Brusa, K. Hege, & E. Hooper (1997), "Final Review of Adaptive Optics Results from the Pre-conversion MMT", *Proc. SPIE conference on Adaptive Optics and Applications* **3126**, San Diego.

M. Lloyd-Hart, J.R.P. Angel, T.D. Groesbeck, T. Martinez, B.P. Jacobsen, B.A. McLeod, D.W. McCarthy, E.J. Hooper, E.K. Hege, & D.G. Sandler (1998a), "First Astronomical Images Sharpened with Adaptive Optics Using a Sodium Laser Guide Star" *Astrophys. J.* **493**, 950; http://athene.as.arizona.edu:8000/caao/m13.html

M. Lloyd-Hart, R. Angel, D. Sandler, T. Barrett, P. McGuire, T. Rhoadarmer, D. Bruns, S. Miller, D. McCarthy, & M. Cheselka(1998b), "Infrared adaptive optics system for the 6.5 m MMT: system status and prototype results," *Proc. SPIE Conference on Adaptive Optical System Technologies*, ed. D. Bonaccini & R.K. Tyson, **3353**, Kona, Hawaii.

H.M. Martin, J.H. Burge, D.A. Ketelson, & S.C. West (1997), "Fabrication of the 6.5-m primary mirror for the Multiple Mirror Telescope Conversion", *Proc. SPIE Conference on Optical Telescopes of Today and Tomorrow*, Ed. A.L. Ardeberg, **2871**, 399.

T. Martinez (1998), "Continuous Wave Dye Laser for Use in Astronomical Adaptive Optics", Ph.D. dissertation, University of Arizona.

C.E. Max, S.S. Olivier, H.W. Friedman, J. An, K. Avicola, B.V. Beeman, H.D. Bissinger, J.M. Brase, G.V. Erbert, D.T. Gavel, K. Kanz, B. Macintosh, K.P. Neeb, & K.E. Waltjen (1997), "Image Improvement from a Sodium-layer Laser Guide Star Adaptive Optics System", *Science*, **277**, 1649;
http://ep.llnl.gov/urp/science/lgs_www/lgs_lick.html

J.R. Meyer-Arendt (1984): *Introduction to Classical and Modern Optics*. 2nd Edition (Englewood Cliffs, New Jersey: Prentice-Hall), 214–223.

S. Mikoshiba & B. Ahlborn (1973), *Rev. Sci. Instrum.*, **44**, 508.

R.A. Muller & A. Buffington (1974), *J. Opt. Soc. Am.*, **64**,9.

I. Newton (1730), *Opticks*, 4th edition, Dover, New York (1979).

R.J. Noll (1976), "Zernike Polynomials and Atmospheric Turbulence", *J. Opt. Soc. Am.*, **66**, 207.

C.A. Primmerman, D.V. Murphy, D.A. Page, B.G. Zollars, & H.T. Barclay (1991), "Compensation of Atmospheric Optical Distortion Using a Synthetic Beacon", *Nature*, **353**, 141.

A. Quirrenbach (1997), "Laser Guidestars", *Proc. of the NATO/ASI School on Laser Guide Star Adaptive Optics for Astronomy*, Ed. J.C. Dainty, Cargese, Corsica.

W.T. Roberts Jr., J.T. Murray, W.L. Austin, R.C. Powell,& J.R.P. Angel (1998), "Solid-State Raman Laser for MMT Sodium Guide Star", *Proc. SPIE Conf. on Adaptive Optical System Technologies*, **3353**, Kona, Hawaii.

T.A. Rhoadarmer (1996), unpublished.

F. Roddier (1988), *Appl. Opt.*, **27**, 1223.

D.G. Sandler, T.K. Barrett, D.A. Palmer, R.Q. Fugate & W.J. Wild (1991), "Use of a Neural Network to Control an Adaptive Optics System for an Astronomical Telescope", *Nature*, **351**, 300.

D.G. Sandler (1991), unpublished.

D.G. Sandler, L. Cuellar, M. Lefebvre, T. Barrett, R. Arnold, P. Johnson, A. Rego, G. Smith, G. Taylor, & B. Spivey (1994), "Shearing Interferometry for Laser-guide-star Atmospheric Correction at Large D/r_0", *J. Opt. Soc. Am.*, **A11**, 858.

D.G. Sandler, S. Stahl, J.R.P. Angel, M. Lloyd-Hart, & D. McCarthy (1994), "Adaptive Optics for Diffraction-Limited Infrared Imaging with 8-m Telescopes", *J. Opt. Soc. Am.*, **A11**, 925.

C. Schwartz, G. Baum, & E.N. Ribak (1994), "Turbulence-degraded Wavefronts as Fractal Surfaces", *J. Opt. Soc. Am.*, **A11**, 444.

R.B. Shack & B.C. Platt (1971), abstr. *J. Opt. Soc. Am.*, **61**, 656.

F. Shi, M.F. Smutko, M. Chun, J. Larkin, V. Scor, W. Wild, & E. Kibblewhite(1996), poster paper at January 1996 American Astronomical Society meeting, http://astro.uchicago.edu/chaos/poster3/poster3.html

S.M. Stahl & D.G. Sandler (1995), "Optimization and Performance of Adaptive Optics for Imaging Extrasolar Planets", Astrophys. Journal, **L454**, 153.

P. Strobach (1990), *Linear Prediction Theory: a Mathematical Basis for Adaptive Systems*, Springer-Verlag, Berlin.

V.I. Tatarskii (1961), *Wave Propagation in a Turbulent Medium*, McGraw-Hill, New York.

G.I. Taylor (1935), "Statistical Theory of Turbulence", *Proc. R. Soc. Lond.*, **A151**, 421.

R.K. Tyson (1991), *Principles of Adaptive Optics*, Academic Press, San Deigo.

G. Vdovin (1995), "Model of an Adaptive Optical System Controlled by a Neural Network", *Optical Engineering*, **34**, 3249.

E.P. Wallner (1983), "Optimal Wavefront Correction Using Slope Measurements", *J. Opt. Soc. Am.*, **73**, 1771.

S.C. West, S. Callahan, F.H. Chaffee, W. Davison, S. DeRigne, D. Fabricant, C.B. Foltz, J.M. Hill, R.H. Nagel, A. Poyner, & J.T. Williams (1996), "Toward first light for the 6.5-m MMT telescope", *Proc. of SPIE conference on Optical Telescopes of Today and Tomorrow*, **2871**, http://medusa.as.arizona.edu/lbtwww/tech/light.htm

W.J. Wild, E.J. Kibblewhite, R. Vuilleumier, V. Scor, F. Shi, & N. Farmiga (1995), "Investigation of Wavefront Estimators Using the Wavefront Control Experiment at Yerkes Observatory", *Proc. SPIE Conf. on Adaptive Optical Systems and Applications*, Eds. R.K. Tyson & R.Q. Fugate, **2534**, 194.

W.J. Wild, E.J. Kibblewhite, & R. Vuilleumier (1995), "Sparse Matrix Wavefront Estimators for Adaptive-Optics Systems for Large Ground-based Telescopes", *Optics Letters*, **20**, 955.

W.J. Wild (1996), "Predictive optimal estimators for adaptive-optics systems", *Optics Letters*, **21**, 1433.