Adaptive Optics for Vision Science and Astronomy
*ASP Conference Series, Vol. **VOLUME**, **PUBLICATION YEAR***
A. Quirrenbach

Deconvolution of Adaptive Optics Image Data

David W. Tyler

The University of New Mexico,Albuquerque High Performance Computing Center, 1601 Central Ave., NE, Albuquerque, NM 87131

Abstract. Adaptive optics compensation of turbulence effects in image data is seldom perfect. To achieve diffraction-limited resolution, especially with astronomical systems designed to guide on very dim natural objects, post-detection processing (deconvolution) must be accomplished. I discuss the history of deconvolution schemes in astronomical imaging, the limits imposed by data noise, and details of some algorithms. I also discuss observing techniques and instrumentation to optimize angular resolution and minimize artifacts in processed imagery.

1. Introduction

Since the original natural guide star (NGS) adaptive optics (AO) observations by the French ONERA group (Rousset, et al. 1990) and laser guide star (LGS) results from declassified US Air Force programs (Fugate, et al. 1991; Primmerman, et al. 1991), the promise of AO technology in astronomy has been realized to an astonishing degree. To quantify this statement, one could do no better than to cite the recent review by Close (2000) of published AO science results. Close cites over 60 refereed papers on solar system and galactic AO astronomy published between 1993 and 2000, and the trend in terms of AO papers published per year is clearly increasing.

New discoveries await the routine use of AO spectroscopy (Fugate 1995; Tyler & Ellerbroek 1998), but at the time of this writing, nearly all AO science results derive from image data (Close 2000). Compensation of turbulence effects that degrade image resolution is seldom perfect; this is especially true at astronomical observatories, where the objects to be observed are nearly always dim and diffuse. Systems capable of using these objects as NGS wavefront sensor references typically undersample the turbulent atmosphere in the near-IR both spatially and temporally, with the resulting image data showing significant residual degradation due to deformable mirror (DM) fitting error as well as servo lag. Even much more powerful systems fail to deliver diffraction-limited resolution on a routine basis due to system noise; it is a fact that a single AO system cannot be all things to all observers. As a result, research in post-detection processing (deconvolution) techniques to enhance the resolution of AO imagery began shortly after new groups in the astronomy community began work to build AO systems (e.g., Roggemann 1991; Roggemann & Matson 1992; Roggemann, Tyler, & Bilmont 1992; Bilmont, et al. 1992; Tyler & Roggemann 1992).

While laser guide star (LGS) technology (still somewhat immature) may provide sodium-layer reference sources with enough brightness to drive high-bandwidth, high-actuator-density systems on the present generation of 3-m class to 10-m telescopes, one may still expect deconvolution techniques to be required for the foreseeable future: First, only two such high-bandwidth, high-density systems currently exist, both at military facilities and both on relatively small telescopes (3.5-m and 3.63-m); building such a system for an 8- or 10-m telescope would be a daunting task. Second, LGS systems carry with them their own set of artifact-generating complications (e.g., Hippler, et al. 2000); and finally, it seems that as fast as developments in laser physics occur, the human imagination creates the need for even more power for such extreme projects as the 100-m OverWhelmingly Large (OWL) telescope!

It is my goal in this paper to provide an understanding of the principles of deconvolution, including practical issues, observing techniques, and the limits on what can and can't be done regardless of the algorithm used. In the next section, I will present the integral equation (the convolution integral) relevant to imaging through the Earth's atmosphere and discuss a simple technique for its solution. I will then discuss how data error is amplified in *any* attempt to solve the convolution integral, and show that noise amplification imposes a firm limit on what can ever be known about the object, given only the image data. Finally, I will discuss the history of deconvolution efforts in astronomical imaging, pointing out why most initial efforts were coolly received and why there has been a recent resurgence of interest in the field. In Section 3., I will discuss linear, or "direct," and nonlinear, iterative algorithms. Here, I will discuss the inadequacy of the maximum-likelihood criterion (the "stopping problem") as a springboard to a discussion of prior probability distributions as regularizing metrics; I will also discuss the use of ensemble statistics to regularize direct methods. In Section 4., I will discuss various observing and instrumentation techniques. I present a brief discussion of anisoplanatic PSFs and possible techniques for wide-field inversion in Section 5.; a summary is the final section. To save space, I have omitted sample deconvolved images and refer the interested reader to the references provided.

2. Deconvolution in astronomical imaging

2.1. The convolution integral

Scalar diffraction theory predicts the relation between the intensity distribution $i(x, y)$ at the focal plane of some physically realizable optical system and the image predicted by geometrical optics (the object intensity $o(x, y)$, scaled by some magnification factor) is given, under certain conditions, by the *convolution* of the object distribution with the squared-magnitude of the Fourier transform of the optical system's *pupil function* (Goodman 1968):

$$i(x, y) \;=\; \int dx \int dy \; |p(x - x', y - y')|^2 \; o(x, y), \tag{1}$$

where the function $p(x, y)$ is the Fourier transform of the pupil function. In the *diffraction-limited* case, the pupil function $P(\xi, \eta)$ simply defines the physical

extent of the pupil of the optical system; for the simplest example of a single circular lens, $P(\xi, \eta) = 0$ outside the radius defined by the lens. The squared-modulus of the pupil function's Fourier transform is central enough to the theory of image formation that it has its own name and notation:

$$|\mathcal{F}\{P(\xi, \eta)\}|^2 \;=\; |p(x, y)|^2 \;=\; h(x, y) \;= h(\mathbf{x}), \tag{2}$$

where $h(\mathbf{x})$ is called the *point-spread function* (PSF) of the optical system. Using this notation and vector notation, the convolution equation is written

$$i(\mathbf{x}) \;= \int d\mathbf{x}\; h(\mathbf{x} - \mathbf{x}')o(\mathbf{x}). \tag{3}$$

From this equation, with a circular aperture used to define $P(\xi, \eta)$, the well-known Airy pattern can be derived.

As one might infer from the form of (3), the PSF is the image formed when the object consists of a single unresolved point source; that is, when $o(\mathbf{x}) = K\delta(\mathbf{x})$, where K is the intensity in photons. The integral states that the image of any *extended object* can be calculated as the incoherent sum of PSFs from many such point sources. Among the conditions required for (3) to hold are that the PSF be *invariant* over the angular field subtended by the object. In all discussions in this paper regarding deconvolution, it will be assumed that this condition is met, with the exception of Section 5., where exceptions occurring in AO imaging will be presented briefly.

2.2. The optical transfer function

Consider now the Fourier transform of (3). Taking the Fourier transform of both sides admits the simpler relation between the spatial spectra of the object and image:

$$I(\mathbf{u}) \;=\; H(\mathbf{u})O(\mathbf{u}), \tag{4}$$

where $I(\mathbf{u})$, $H(\mathbf{u})$, and $O(\mathbf{u})$ are the Fourier transforms of $i(\mathbf{x})$, $h(\mathbf{x})$, and $o(\mathbf{x})$, respectively. In this form, one may see more directly why the image and object intensities do not have a simple magnification relation: Specifically, the function $H(\mathbf{u})$ attenuates, or "filters" (attenuates) the object spatial spectrum.

The Fourier transform of the PSF, $H(\mathbf{u})$, also has a special name; it is called the *optical transfer function* (OTF). Diffraction (or the presence of turbulence near the telescope's pupil plane) does not reduce the intensity of the incident light, so $H(\mathbf{u})$ takes unit magnitude at the spatial frequency origin. Also, diffraction will not change a real intensity distribution to a complex one, so the phase of $H(\mathbf{u})$ is zero at the origin. Finally, since the PSF is the squared-modulus of $p(x, y)$, the OTF is the autocorrelation of the bounded function $P(\xi, \eta)$. This means the bandwidth over which the OTF has non-zero magnitude is also bounded and the largest magnitude occurs at the origin. For our purposes, then, the OTF is a function that leaves total intensity unaffected and attenuates high spatial frequencies components (that is, fine structure in the image) relative to low spatial frequencies (gross image morphology). The OTF for a simple circular pupil with a central obstruction is shown in Fig. 1. The circular boundary outside of which all Fourier components are attenuated to zero is called the *diffraction cutoff*.

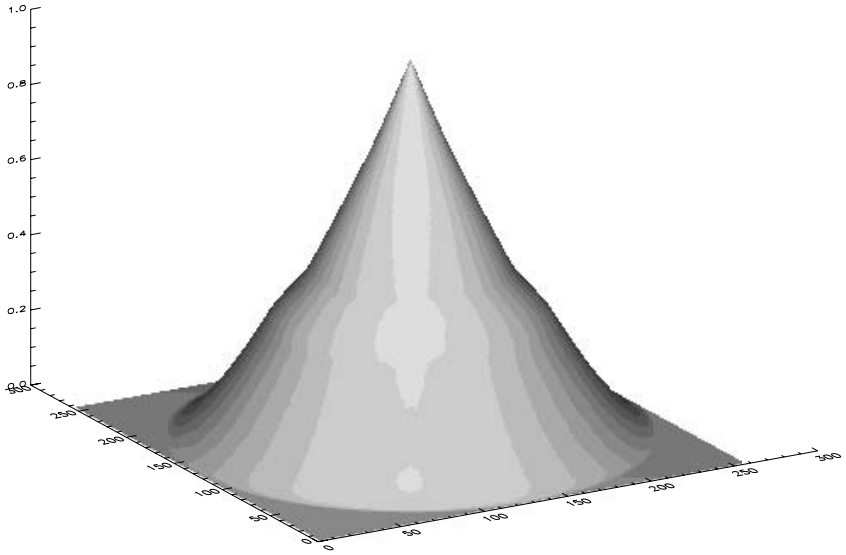

Figure 1. Surface rendering of the OTF corresponding to a circular pupil with a central obstruction

The effect of several optical phenomena can usually be accounted for simply by multiplying together the individual OTFs corresponding to those phenomena. As a relevant example, the combined effect of atmospheric turbulence and telescope diffraction can be treated as the product of two separate OTFs (Goodman 1985) under certain realistic conditions. Often, this multiplicative approximation is still valid when turbulence has been partially compensated by AO (Ellerbroek & Tyler 1999). Without AO compensation, the total OTF is given by the product

$$\begin{aligned} H(\mathbf{u}) &= H_d(\mathbf{u})H_a(\mathbf{u}) \\ &= H_d(\mathbf{u})\ \exp\{-3.44\left(\frac{D}{r_o}\ u\right)^{5/3}\}, \end{aligned} \tag{5}$$

where $H_{dl}(\mathbf{u})$ and $H_a(\mathbf{u})$ are the diffraction and atmospheric OTFs, respectively; $u = |\mathbf{u}|$; D is the telescope diameter; and r_o is the well-known Fried seeing parameter.

The effects of spatial frequency filtering by both atmospheric turbulence and diffraction are shown in Figs. 2 and 3. In Fig. 2, a Gaussian curve (solid line) is used to model the cross-section of a sample diffraction-limited image spectrum $H_d(\mathbf{u})O(\mathbf{u})$. The dashed line shows the effect of multiplying the diffraction-limited spectrum by $H_a(\mathbf{u})$ with $D/r_o = 10$. Note that the high-frequency components of the original diffraction-limited spectrum have been sharply attenuated by the additional filtering from atmospheric turbulence. Note also the

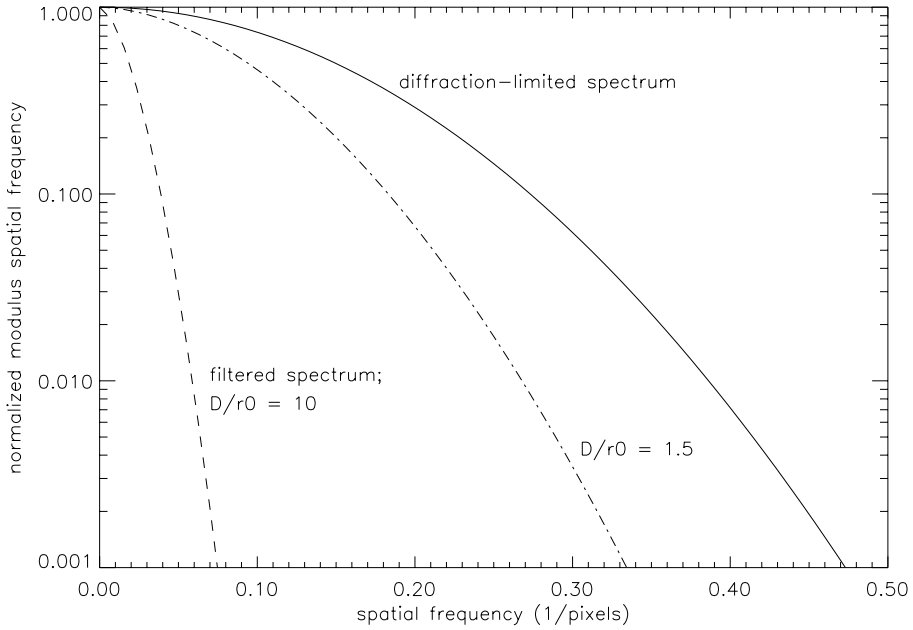

Figure 2. Cross-sections of modulus spectra filtered by (solid line) telescope diffraction alone and telescope diffraction with atmospheric turbulence characterized by $D/r_o = 10$ (dashed line) and $D/r_o = 5$ (dot-dashed line).

condition $D/r_o = 10$ may not be "severe" turbulence in terms of the Fried parameter if the telescope is large; for an 8-m telescope, this condition represents $r_o = 80$ cm. Figure 3 shows the effect of the turbulence filtering on resolution and image quality. The solid-line plot in Fig. 3 is the image corresponding to the diffraction-limited Gaussian spectrum in Fig. 2; the dashed line is the image corresponding to the $D/r_o = 10$ spectrum. Notice the loss of the high-frequency portion of the spectrum has caused considerable loss of image contrast, with significant intensity now spread beyond the angular extent of the object diffraction-limited image. This spreading means fine structure in any image so filtered is likewise smeared out and hard to detect. Faint companions are similarly hard to see both because of the large wings of the primary and the identical spreading out of the companion's light. Note the OTF is generally a complex number, so for short-exposure imaging, where the PSF is asymmetric and "speckled," or for certain fixed optical aberrations, the filtering of any object spectral component includes multiplication by the *phase* of the OTF at that frequency.

2.3. Deconvolution and the limit to image resolution

Attempts to mathematically invert this filtering and restore spectral amplitudes to at least the diffraction-limited level are referred to as *deconvolution* processing. At first glance, successful deconvolution appears much simpler than it is in practice. One might, for example, simply divide both sides of (4) by $H(\mathbf{u})$. The

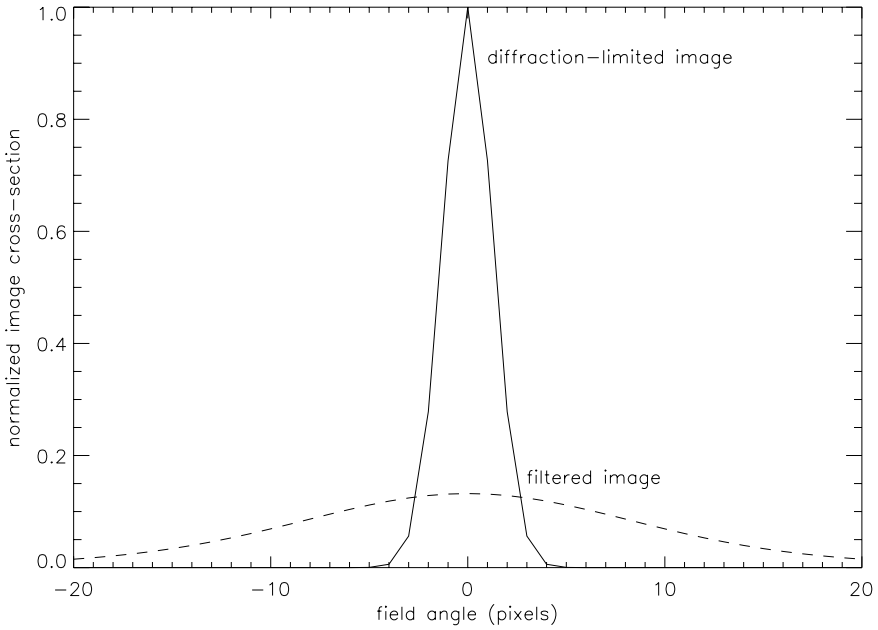

Figure 3. Slices across the images corresponding to the spectra in Fig. 2.

complications arise when one considers the unavoidable accumulation of noise when measuring and digitizing $I(\mathbf{u})$.
Using the notation $I_o(\mathbf{u})$ to indicate the Fourier transform of the image intensity in the absence of measurement noise, the object distribution $\tilde{o}(\mathbf{x})$ estimated by dividing out the OTF is written

$$\tilde{o}(\mathbf{x}) \;=\; \mathcal{F}^{-1}\left\{\frac{I_o(\mathbf{u})+N(\mathbf{u})}{H(\mathbf{u})}\right\}, \tag{6}$$

where $N(\mathbf{u})$ is the Fourier transform of the noise in the measured image. Now consider the magnitude of the error in the estimated spectrum $\tilde{O}(\mathbf{u})$:

$$|\delta\tilde{O}(\mathbf{u})| \;=\; \frac{|N(\mathbf{u})|}{|H(\mathbf{u})|}. \tag{7}$$

As I have already pointed out, $|H(\mathbf{u})|$ goes to zero for large u, so the only way the total error in (7) can be bounded is if $|N(\mathbf{u})|$ also approaches zero for large values of u, and approaches zero *faster* than $|H(\mathbf{u})|$ (Craig & Brown 1986). Unfortunately, realistic detector noises do neither; for example, both CCD amplifier "readout" noise and Poisson ("shot") noise have essentially flat spectra. *This fact establishes the limit of what can be recovered by* any *deconvolution algorithm* (Tyler, et al. 1998)
Formally, the signal-to-noise ratio of the data spectrum,

$$\mathrm{SNR}(\mathbf{u}) \;=\; \frac{<|I(\mathbf{u})|>}{\sqrt{\mathrm{var}\{I(\mathbf{u})\}}}, \tag{8}$$

defines what can ever be known about the object, because once a noisy measurement is made, any further scaling of the corrupted spectral components amplifies both signal and noise by the same factor! In (8) above, $< * >$ is the expectation operator, and the variance of the complex quantity $I(\mathbf{u})$ is given by

$$\text{var}\{I(\mathbf{u})\} \; =< |I(\mathbf{u})|^2 > \; - \; | < I(\mathbf{u}) > |^2. \tag{9}$$

Where SNR($\mathbf{u}$) decreases below unit value, the uncertainty in the measured spectrum is as great as the measured mean. At this point, no information is contained in that part of the spectra; this boundary is the *SNR cutoff* of the data. It is the SNR cutoff that actually determines available resolution, with or without deconvolution, and the SNR cutoff of $I(\mathbf{u})$ is nearly always short of the diffraction cutoff.

Returning to the plots in Fig. 2, consider the use of (6) to deconvolve the atmospheric turbulence and diffraction filters in the presence of a flat noise spectrum with a relative magnitude of 0.01. Dividing the filtered noisy spectrum by $H(\mathbf{u})$ yields what is almost totally amplified noise for u greater than about 0.05, or about 1/10 of the diffraction-limited passband. In other words, the deconvolved spectrum is so dominated by noise, the corresponding image would be both closed to interpretation and unrepeatable using subsequent measurements. Note that because the atmospheric filter is so strongly attenuating, reducing the noise level by a full order of magnitude only increases the portion of the spectrum not dominated by noise a small fraction of the diffraction limit.

Noise amplification upon deconvolution leads to the requirement for *regularization*. Regularization is nothing more than modification of the problem at hand to achieve some compromise between enhancing resolution and amplifying noise. For example, one might regularize estimation of the object spectrum by using a second filter in the deconvolution:

$$\tilde{o}(\mathbf{x}) \; = \; \mathcal{F}^{-1} \left\{ \frac{I_o(\mathbf{u}) + N(\mathbf{u})}{H(\mathbf{u})} \; R(\mathbf{u}) \right\}, \tag{10}$$

where $R(\mathbf{u})$ attenuates those spectral components where $|N(\mathbf{u})| \gg |I_o(\mathbf{u})|$ and passes those frequencies where the opposite is true.

2.4. A brief history of deconvolution

From the $D/r_o = 10$ plot in Fig. 2, it is clear the use of regularized deconvolution in this case will provide only a negligible increase in resolution beyond that obtained in the original image if significant noise is present. In typical near-IR and visible seeing, most of the spatial spectrum of images taken from research-size telescopes are attenuated well below the noise threshold, and use of (10) results in all the components amplified by deconvolution being immediately reset to zero by regularization. Images thus "deconvolved" may be free of amplified noise, but their resolution will not be noticeably improved by the processing. This is a practical consequence of my previous statement that deconvolution cannot change the data SNR. Further, although (10) is a very simple approach to deconvolution, *no algorithm will do better* unless the algorithm allows prior knowledge about the object to be used in some way.

Nearly simultaneously with the arrival of modest computing power and scanning densitometers to digitize photographic plates, Richardson (1972) and Lucy

(1974) independently proposed their iterative solution for $\tilde{O}(\mathbf{u})$. In what is now known as the Richardson-Lucy (RL) algorithm, the first iterations estimate low-spatial frequency features of the object, and subsequent iterations progressively estimate higher frequency (lower SNR) components of the object (Carasso 1999). Some details of the RL algorithm will be discussed in Section 3., but here it suffices to point out that Lucy (1974) cautioned readers not to iterate the algorithm to convergence in the presence of noise. Put another way, he recognized the need for regularization, provided in this case by early termination of the algorithm. Due to the challenges posed by the strong spectral filtering of the atmosphere discussed in the last paragraph, the RL algorithm was not then of any practical value and sat unused for nearly twenty years.

The first successful attempts at high-resolution imaging via deconvolution came when Labeyrie's (1970) *speckle interferometry* method for amplitude spectrum estimation was used together with the Knox and Thompson (1974) phase retrieval technique. The reason these techniques were successful is that neither algorithm estimates $O(\mathbf{u})$; instead, Labeyrie's method estimates $|O(\mathbf{u})|^2$ using the deconvolution

$$|\tilde{O}(\mathbf{u})|^2 \; = \frac{< |I(\mathbf{u})|^2 >}{< |H(\mathbf{u})|^2 >} \, R(\mathbf{u}), \tag{11}$$

where $< |H(\mathbf{u})| >^2$ is called the *speckle transfer function* (STF) and $R(\mathbf{u})$ is again a suitably-chosen regularizing function. The Knox-Thompson method estimates the average image *cross-spectrum*, $< I(\mathbf{u})I^*(\mathbf{u}+\delta\mathbf{u}) >$, where $\delta\mathbf{u}$ is a small spatial frequency offset. The phase spectrum of the object, needed with the amplitude spectrum $|\tilde{O}(\mathbf{u})|$ to form an image, is extracted from the cross-spectrum. The joint use of speckle interferometry with either the Knox-Thompson method or the bispectrum (Section 3.) is called *speckle imaging.*

Speckle interferometry and speckle imaging are so called because, rather than a single long-exposure (minutes to hours) image, the input data consists of an ensemble of short-exposure (tens to hundreds of ms) images. These exposures are short enough to capture instantaneous realizations of the random atmospheric phase fluctuations, resulting in the appearance of coherent "speckles" in each image (Goodman 1985, Dainty 1984).

The SNR distribution of $< |I(\mathbf{u})|^2 >$, in contrast to the SNR of $< I(\mathbf{u}) >$, is typically significant to a large fraction of the diffraction limit; however, if the object is dim and D/r_o is large, hundreds or even thousands of frames may need to be averaged to obtain high SNR($\mathbf{u}$) over a large part of the spectrum. This led many to view speckle imaging as unsuitable for imaging of deep-sky objects. Further, the Knox-Thompson technique proved to be somewhat less than robust, and various bispectrum phase extraction algorithms, while much more robust, were viewed as too computationally intensive until recently.

From the early 1980s, there was little progress or innovation in deconvolution techniques until the French ONERA group, developing AO hardware independently of the American military, demonstrated another speckle technique in 1990. In this method, *deconvolution from wavefront sensing* (DWFS), an AO wavefront sensor (WFS) is used to estimate both short-exposure OTFs and the STF (Primot, Rousset, & Fontanella 1990), which are then deconvolved from an ensemble of image spectra to estimate $O(\mathbf{u})$. Unlike full AO, image reconstruction is done after observing, and no DM is used. No separate phase spectrum

calculation is needed with DWFS. Although the original DWFS estimator has been shown to be biased (Roggemann, Welsh, and Devey 1994), DWFS was used to produce nearly-diffraction-limited images of Capella using the 4-m William Herschel Telescope (Marais, et al. 1991). Unfortunately, DWFS also proved to be practical only with relatively bright objects.

In the spring of 1990, the *Hubble Space Telescope* was launched from the cargo bay of the space shuttle *Discovery*. When its optics were found to be misfocused, the RL algorithm was taken off the shelf and began its first productive labors. The relatively high spectral SNR of most *HST* data meant RL would not be hobbled by the strong filtering evident in ground observations without AO. Prior to the 1993 *HST* servicing mission, a year of rather intense and productive deconvolution research went into the Image Restoration Project (Hanisch & White 1993). Although the servicing mission corrected the *HST* optics to their original design specifications, deconvolution remains a part of the *HST* data pipeline to boost spectral components within the diffraction limit and remove artifacts and secondary mirror support diffraction spikes.

Simultaneously, a dawning realization that AO compensation is rarely perfect began the "post-modern" interest in deconvolution of data from Earth-bound telescopes. Like *HST* imagery, AO data is not strongly filtered by the atmospheric OTF. Return to Fig. 2 and consider the dot-dashed plot marked "$D/r_o = 1.5$:" This curve shows the solid-line (diffraction-limited) spectrum again multiplied by $H_a(\mathbf{u})$, but this time with $D/r_o = 1.5$ to represent partial AO compensation of the turbulence represented by the $D/r_o = 10$ plot. Again considering a relative noise level of 0.01, it is now apparent that spectral components at frequencies lower than about 0.25 (half the diffraction limit) may be amplified by deconvolution without excessive noise. Also, due to the relatively shallow slope of the $D/r_o = 1.5$ spectrum, reductions in the noise level now result in corresponding increases in the SNR cutoff resolution. Partial AO compensation has a similar effect on the robustness of speckle imaging, as I will discuss in Section 3.

With RL and other algorithms investigated rather thoroughly during the course of the Image Restoration Project and made part of data processing packages such as IRAF and IDL, astronomers became not only motivated to deconvolve partially-compensated AO imagery but comfortable with the idea. The first Canada-France-Hawaii Telescope AO imagery used in an astrophysical investigation (Roddier, et al. 1995) was deconvolved using the IRAF RL package. The practice of deconvolving AO images continues, with recent workers still using RL (e.g., Chesneau, et al. 2000).

From this short historical review, it should be apparent that the limits of deconvolution, like those of AO, must be understood for this important tool to be effective. Deconvolution, with the exception of speckle imaging, was for some time dismissed by astronomers as an academic exercise; however, the concept itself is not flawed. With partial AO compensation, the need for deconvolution remains; however, AO pre-conditions the the solution of (3), providing a large arena in which these techniques can be robustly used.

3. Some deconvolution algorithms

In this section, I will discuss the workhorse Richardson-Lucy algorithm, including the general notion of *maximum likelihood* deconvolution. I will also discuss what Carasso (1999) calls the "inadequacy of the maximum likelihood criterion," motivating a discussion of the use of *prior knowledge* in Bayesian deconvolution schemes. I will conclude with a discussion of speckle imaging, how its domain of use is enlarged by AO compensation, and emphasize how ensemble statistics can be used to quantify limiting resolution.

3.1. The Richardson-Lucy algorithm and maximum-likelihood estimation

Algorithm description and characteristics. The RL algorithm codifies Bayes's theorem on conditional probabilities:

$$P(A|B)P(B) \;=\; P(B|A)P(A), \tag{12}$$

where the P functions are probabilities, and A and B are viewed as outcomes of random experiments. The expression $P(A|B)$ should be read as "the probability of A given that B has occurred." In the present case, $P(B|A)$ is interpreted as the probability of measuring the image B given that the object is A. Using the notation o_j to represent the unblurred object distribution at the jth pixel and d_i to represent the blurred and noisy image data at the ith pixel, one may write d_i/N as the *probability* of a signal event at the ith pixel in the image array, where N is the total intensity of both the object and image arrays. Similarly, o_j/N is the probability of an "event" at the jth pixel in the object array. With this established, the conditional probability $P(d_i|o_j)$ is now h_{ij}, an element of the point spread function *matrix* mapping the vector $\mathbf{o}$ to $\mathbf{d}$ via the discrete convolution

$$d_i \;=\; \sum_j \; h_{ij} o_j. \tag{13}$$

Note that $\sum_j h_{ij} = 1$ for all j, preserving the 1-norm or total intensity of $\mathbf{o}$. Using this notation and some algebra, Bayes's theorem can be rewritten and solved for o_j:

$$o_j \;=\; o_j \; \sum_i \left(\frac{h_{ij} d_i}{\sum_k h_{jk} o_k} \right). \tag{14}$$

With o_j expressed in terms of itself, iterative substitution can be used to find a solution. Using operator notation to simplify the convolutions,

$$\mathbf{o}^{n+1} \;=\; \mathbf{o}^n \; \mathcal{H}^* \left(\frac{\mathbf{d}}{\mathcal{H}\mathbf{o}^n} \right), \tag{15}$$

where $\mathcal{H}$ is the PSF convolution operator and $\mathcal{H}^*$ is its adjoint. Note I have used the ratio $\mathbf{d}/\mathcal{H}\mathbf{o}^n$ to indicate a vector formed by dividing each component of $\mathbf{d}$ by the corresponding components of $\mathcal{H}\mathbf{o}^n$. Likewise, the product on the right-hand side is the vector formed by multiplying each component of $\mathbf{o}^n$ with the components of $\mathcal{H}^*\mathbf{d}/\mathcal{H}\mathbf{o}^n$. The initial conditions for the iteration are all elements of $\mathbf{o}^0$ equal to N/P, where P is the number of pixels. This gives $\sum_j o_j^0 = N$.

The use of linear algebra in the preceding discussion again implies an invariant PSF. Also, the PSF must be *known* (or realistically, estimated from modeling or measurement). The latter implies the RL method is unsuitable for deconvolution of images short enough that the effects of the atmosphere are not well-averaged over time; were this not so, both the calibration star image and the PSF corrupting the science object image would be independent realizations of a random process.

To understand RL intuitively, let $\mathbf{r}$ denote the vector $\mathbf{d}/\mathcal{H}\mathbf{o}$. This vector can always be written in the form $\mathbf{r} = \mathbf{u}+(\mathbf{r}-\mathbf{u})$, where $\mathbf{u}$ denotes a vector with each component having unit value. Since $\sum_j h_{ij} = \sum_i h_{ij} = 1$, $\mathcal{H}^*\mathbf{u} = \mathbf{u}$. Inserting this in (15) yields

$$\mathbf{o}^{n+1} \;=\; \mathbf{o}^n \;+\; \mathcal{H}^*(\mathbf{r}-\mathbf{u}). \tag{16}$$

This says the object estimate at the $(n+1)$-st iteration can be written as the object estimate at the nth iteration plus the projection onto object space of the vector formed by differencing the data vector and filtered object estimate at the nth iteration.

Finally, consider RL deconvolution of noise-free data. Now let n be the iteration where

$$\mathbf{r} = \;\mathbf{u} \;+\; \mathbf{v}, \tag{17}$$

where $\mathbf{v}$ denotes a vector with components such that when they are ordered properly in a two-dimensional image array, the spatial spectrum of the array has only very low amplitude at low spatial frequencies. Specifically, let the norm of that part of $\mathbf{v}$ intersecting the *range* of $\mathcal{H}$ be small relative to the norm of $\mathbf{v}$. Recall the transformation of $\mathbf{o}$ by $\mathcal{H}$ produces vectors with appreciable power only in some restricted passband defined by the severity of the atmospheric turbulence (Fig. 2); the set of correctly-normed vectors with P components and attenuated spatial spectra define the range of $\mathcal{H}$. Referring again to (16) with $\mathbf{r}-\mathbf{u} = \mathbf{v}$, the RL iteration is

$$\mathbf{o}^{n+1} \;=\; \mathbf{o}^n \;+\; \mathcal{H}^*\mathbf{v}. \tag{18}$$

But since the range of $\mathcal{H}$ and the *nullspace* of $\mathcal{H}^*$ are orthogonal complements in the linear space on which $\mathcal{H}^*$ can operate, our previous statement that the spatial spectrum of $\mathbf{v}$ consists almost entirely of components at frequencies beyond the SNR cutoff of the OTF implies

$$\mathbf{o}^{n+1} \;\approx \mathbf{o}^n. \tag{19}$$

This tells us the RL algorithm will reproduce in its early iterations those spectral components not strongly filtered by the OTF; further, spectral components that *are* strongly filtered by the OTF will take relatively many iterations to reconstruct because the algorithm is relatively unresponsive in these bands. These conclusions are consistent with observed behavior of RL. The implication for *noisy* data is this: If the algorithm has gone through sufficiently many iterations that $\mathbf{o}^{n+1}$ doesn't differ significantly from $\mathbf{o}^n$ (using, for example, the 2-norm) any differences between successive estimates *are likely due to noise amplification.*

Carasso (1999) points out that RL can also be derived directly from (3) without using Bayes's theorem. Also, while neither the Richardson or Lucy derivations

nor the derivation directly from (3) require any assumptions about noise in $\mathbf{i}$, one may derive an iterative algorithm *identical to RL* by assuming $\mathbf{d}$ is corrupted by Poisson noise and estimating a $\mathbf{o}$ vector to maximize the log of $P(\mathbf{d}|\mathbf{o})$ (Shepp & Vardi 1982). This derivation demonstrates the *expectation-maximization* (EM) method. Carasso goes on to state that RL and Shepp-Vardi, as well as any algorithm that maximizes the log of $P(\mathbf{d}|\mathbf{o})$ when the noise is assumed to be *Gaussian*, all minimize some norm of $\mathbf{d} - \mathcal{H}\mathbf{o}$ (recall (16)). Since minimizing this quantity is equivalent to maximizing $P(\mathbf{d}|\mathbf{o})$ for Gaussian or Poisson noise, any $\mathbf{o}$ chosen to satisfy $||\mathbf{d} - \mathcal{H}\mathbf{o}||$ is called a *maximum likelihood estimator* for $\mathbf{o}$.

RL stopping criteria. From the preceding discussion of the response of RL to strongly filtered, noisy data, it should be clear that maximum likelihood estimation will at some point amplify data noise. By implication, the main difficulty with RL and other maximum-likelihood algorithms is *knowing when to stop the iterations*; *viz.*, RL and similar schemes cannot be iterated to convergence if the data is at all noisy. This is entirely equivalent to my earlier statement that the direct (non-iterative) deconvolution (6) must be regularized. By extension, my previous conclusions regarding the limits imposed by the data SNR also apply.

Suggestions for regularizing RL range from the reasonable to the silly. In the latter category are suggestions for "smoothing" the deconvolved image after allowing RL to converge. Whether the smoothing is done in the spectral domain with a filter or in the image domain with a pixel-averaging "sieve" makes no difference; the obvious flaw with this idea is that the extent of the reliable (high SNR) region of the spectrum is *unknown*. One may filter the image until it "looks good" (or better, until the norm of the residual is consistent with the noise variance), but unless the object morphology is already known, one has no idea whether the image has been smoothed too much or not smoothed enough. Put another way, if one doesn't know when to stop the iterations, one doesn't know how to filter the converged result, either. In a research area where images are typically collected to discover something previously undetected, such practice is hazardous. The perils of such blind regularization are documented by Michard (1996), who describes "large and far extending artifacts" in images of E-S0 galaxies acquired by CFHT and the *HST* and deconvolved with RL.

More useful approaches to RL regularization involve comparison with other imagery or "test" deconvolutions. Roddier, et al. (1996) suggest collecting PSF data by observing at least two calibration stars (always good practice, since many "guide stars" turn out to be binaries under AO scrutiny) and deconvolving the PSFs with each other using the same number of iterations as with the science object. If features observed in the deconvolved science image are not present in the deconvolved point-source images, they are deemed "real." One important caveat here is that the *mean* spectral amplitudes for a point source will be higher than those for an extended object, so the spectral SNR will also be higher, likely leading to the use of too many iterations with the science object. This caution leads to the idea of using simulations or adding model objects (Lai, et al. 1999) to blank fields of the actual data. One could, for example, add to the image field a Gaussian object with angular extent and intensity similar to the science object and then deconvolve, terminating iterations when asymmetries or noise features begin to appear in the known Gaussian object. If data from sev-

eral epochs (and seeing conditions) are available, deconvolved results from these data may be compared, assuming that features common to deconvolved images across epochs are also real. Potter, et al. (2000) note that when deconvolving AO polarimetry image data, excessive RL iterations give non-physical polarization results. Among the more time-consuming notions is to compare images from iterative algorithms like RL, Maximum-Entropy (MEM), and CLEAN with a direct method such as the linear Weiner filter (Stolovy, Hayward, & Herter 1996; Chesneau, et al. 2000) again looking for common features [1].

3.2. MAP and MEM estimation

Regularization of maximum-likelihood. What Carasso (1999) refers to as the "inadequacy of maximum-likelihood" (that RL and similar algorithms cannot be iterated to convergence in the presence of noise) has led to a variety of attempts to establish a more robust criterion. Returning to Bayes's theorem, the *posterior probability* $P(\mathbf{o}|\mathbf{d})$ (read "the probability of $\mathbf{o}$ being the object given the measured image $\mathbf{d}$") can be written

$$P(\mathbf{o}|\mathbf{d}) \;=\; \frac{P(\mathbf{d}|\mathbf{o})P(\mathbf{o})}{P(\mathbf{d})}. \tag{20}$$

Given a PSF and a noise model, it's possible to write down an expression for $P(\mathbf{d}|\mathbf{o})$. For Poisson noise, for example,

$$P(\mathbf{d}|\mathbf{o}) \;=\; \prod \frac{\exp\{-\sum_j h_{kj} f_j\}\ (\sum_j h_{kj} f_j)^{g_k}}{g_k!} \tag{21}$$

Using (21) in (20) and differentiating the log with respect to $\mathbf{o}$ yields another iterative algorithm, the exact form of which depends on (21) and $P(\mathbf{o})$. This *maximum a-posteriori* (MAP) method thus allows the introduction of prior knowledge in the deconvolution via the *prior probability distribution* (or "prior") $P(\mathbf{o})$.
The prior in the MAP algorithm allows the user to bias, or constrain, the solution vector f in some preferred direction. The nature of a MAP estimator can be seen by writing out the log posterior probability to be maximized:

$$\log P(\mathbf{o}|\mathbf{d}) \;=\; \log P(\mathbf{d}|\mathbf{o}) \;+\; \log P(\mathbf{o}) \;-\; \log P(\mathbf{d}). \tag{22}$$

A little thought reveals that while $P(\mathbf{d})$ is not formally dependent on $\mathbf{o}$, it is not *statistically independent* of $\mathbf{o}$. However, $P(\mathbf{d})$ is treated as independent of $\mathbf{o}$ in the MAP formalism, so only the sum of the first two terms are to be maximized. From our discussion of RL, the vector f that maximizes the first term is simply the maximum-likelihood estimator; in this light, a MAP estimator can be seen as a modification of maximum-likelihood designed to allow the algorithm to converge by constraining the solution away from excessive noise amplification. Alternatively (but perhaps less accurately for most MAP schemes), one may

[1] The data presented in the first paper were collected at 8.7 μm with the Hale 5-m using tip-tilt correction only. While not actually AO data, the D/r_o at such a long wavelength (possibly < 2) is typical of AO compensated data at shorter (near-IR) wavelengths in that deconvolution is a viable way to increase resolution.

regard the prior as providing *additional information* beyond the measured data, allowing resolution to be increased beyond that dictated by the SNR cutoff. Put yet another way, one can argue that the ultimate resolution obtainable using deconvolution is set by the *total information* I_t in the processed image, which is bounded from above by the sum of the information in the measured data (I_d) and the information in any constraints (I_c):

$$I_t \;\leq\; I_d \;+ I_c. \tag{23}$$

The above is an equality when the deconvolution algorithm perfectly communicates all information in the data and constraints to the processed image. Generally, a MAP estimator mimimizes a new criterion equal to the maximum-likelihood criterion plus a regularizing term weighted by some coefficient.

Maximum entropy. As one can imagine, the derivation of new priors has been at times a rich source of new publications. The popular *maximum entropy method* (MEM) algorithm (reviewed in Narayan & Nityananda 1986) was an early example. In MEM, Maxwell-Boltzmann statistics are used (with, in my opinion, only the vaguest of foundations) to weight the distribution of $N/\Delta N$ photons into the detector pixels, where ΔN is the intensity quantum established by the dynamic range of the detector. MEM was derived as a way of introducing prior knowledge without requiring detailed knowledge of the object morphology, but Carasso (1999) argues that the trial-and-error adjustment of various coefficients implies just such knowledge. Further, the form of (23) implies a prior must contain detailed knowledge in some form if the resolution enhancement associated with its use is to be significant. There are a variety of MEM implementations and even several definitions of "entropy!"

Priors from measured data. Whether the prior constrains the Laplacian of $\mathbf{o}$, its derivatives, or its "entropy," regularization terms depending on *assumed* knowledge are not very satisfying. Recent work demonstrates the potential of the MAP formalism for inclusion of *measured data* in addition to the imagery. For example, Piña & Puetter (1992) discuss constraining the solution $\mathbf{o}$ to minimize the norm of the autocorrelation of $\mathcal{H}\mathbf{o} - \mathbf{d}$. If the dominant noise source is zero-mean and independent between pixels, as is CCD read noise, the noise vanishes from the autocorrelation except at zero shift; minimizing the autocorrelation of the residual is an attempt to minimize the total cost function at the iteration number where the residual has only noise components. This idea has also been exploited by Mu & Plemmons (2000). Note that with an ensemble of short-exposure images, a multi-frame deconvolution could be accomplished using the norm of

$$< |O(\mathbf{u})|^2 >< |H(\mathbf{u})|^2 > - < |I(\mathbf{u})|^2 > \tag{24}$$

as a constraint. Recalling from the introduction to speckle interferometry that $< |I(\mathbf{u})|^2 >$ can have significant SNR out to a large fraction of the diffraction limit even in strong turbulence, this constraint also has potential to avoid noise amplification.

Mugnier, et al. (2001) have also used the MAP formalism to incorporate ensembles of WFS measurements collected during AO operation in the joint estimation of short-exposure PSFs and the object from image ensembles. Joint estimation of PSFs and the object without observing a calibration star is referred to as "blind" deconvolution (Ayers & Dainty 1988; Schulz 1993a); to emphasize their

incorporation of WFS data in the algorithm (see also Schulz 1993b), Mugnier et al. refer to their scheme as "myopic deconvolution." A variety of speckle-ensemble blind deconvolution schemes have been developed and studied experimentally for use with and without AO (Jefferies & Christou 1993; Christou 1995; Schulz, Stribling, & Miller 1997; Jeffs & Christou 1998; Tyler, et al. 1999). The multiplicity of image frames acts as an implied constraint because the object is common to every image, while noise and turbulence fluctuations vary randomly between frames. The ensemble myopic scheme used by Mugnier, et al. also makes use of the multi-frame constraint, but in a coherent framework allowing the incorporation of prior knowledge from a variety of sources. The noteworthy feature of Mugnier, et al., is that *actual measurements* are being used to constrain the deconvolution rather than supposition. In the Piña & Puetter scheme, well-understood characteristics of detector noise are used.

3.3. Noise removal

Recent research (Tyler & Matson 1997; Tyler 2000a) has shown the potential to use "hard" prior knowledge, such as the angular extent of the image blur (a constraint on the object's angular *support*), to actually *remove noise* from image data, increasing the spectral SNR. In a convex projection algorithm fundamentally different from simple filtering or using such a constraint in MAP estimation, Fourier-domain noise correlations are introduced when all image pixels outside the support boundary are set to zero; this correlation results in *noise transport* between bands in frequency space. By looking for bands where noise has accumulated relative to the unprocessed data, one may achieve significant increases in the SNR cutoff.

3.4. Classical speckle imaging

The pairing of Labeyrie's technique, given in (11), with phase recovery from the bispectrum are generally thought of as "classical" speckle imaging (use of other algorithms using short-exposure image ensembles, such as DWFS or the Mugnier, et al. scheme can also be called "speckle imaging"). The image bispectrum $B(\mathbf{u},\mathbf{v})$ is defined as follows:

$$B(\mathbf{u},\mathbf{v}) \;=\; I(\mathbf{u})I(\mathbf{v})I^*(\mathbf{u}+\mathbf{v}), \tag{25}$$

where $I^*(\mathbf{u})$ is the complex conjugate of $I(\mathbf{u})$. The phase of the bispectrum is

$$\beta(\mathbf{u},\mathbf{v}) \;=\; \phi(\mathbf{u})+\phi(\mathbf{v})-\phi(\mathbf{u}+\mathbf{v}). \tag{26}$$

Because the vector sum $\mathbf{u}+\mathbf{v}-(\mathbf{u}+\mathbf{v})$ is zero, (26) is referred to as the "closure phase" (Jennison 1958; Lohmann, Weigelt, & Wirnitzer 1983). It is easy to show that the closure phase is invariant under perturbations by the atmospheric phase, yielding $\beta_O(\mathbf{u},\mathbf{v}) = \beta_I(\mathbf{u},\mathbf{v})$, where $\beta_I(\mathbf{u},\mathbf{v})$ is given in (26) and $\beta_O(\mathbf{u},\mathbf{v})$ is the phase of the object bispectrum. While several algorithms exist for extracting an estimate of the object phase from $\beta_I(\mathbf{u},\mathbf{v})$ (Northcott, Ayers, & Dainty 1988; Matson 1992), one of the simpler to understand is the *recursive algorithm* to recover the unit-amplitude phasor $U(\mathbf{u})$:

$$U(\mathbf{u}+\mathbf{v}) \;=\; \sum w(\mathbf{u},\mathbf{v}) \left[\frac{\hat{B}(\mathbf{u},\mathbf{v})}{U(\mathbf{u})U(\mathbf{v})}\right]^*, \tag{27}$$

where the summation is over all previously-estimated phasors for which $\mathbf{u}+\mathbf{v}$ is a vector constant, $w(\mathbf{u},\mathbf{v})$ is a SNR-dependent weight, $\hat{B}(\mathbf{u}) = B(\mathbf{u})/|B(\mathbf{u})|$ and $U(\mathbf{u}) = O(\mathbf{u})/|O(\mathbf{u})|$. The algorithm is started by assuming that, since the object is constrained to be real, the phase of $U(\mathbf{u}=0)$ is zero.

These methods, widely used in binary star studies, are still in use and surprisingly productive in the AO era. The MPI-R group headed by Weigelt has produced high-resolution images of NGC 1068 (Wittkowski, et al. 1998), carbon star (Weigelt, et al. 1998) and AGB star (Gauger, et al. 1999) dust shells, and the mass-loss envelope of VY CMa (Wittkowski, Lauger, & Weigelt 1998). The Mira star R Cas has been spatially resolved, as well, to study its pulsation modes (Hofman, et al. 2000).

The above-referenced studies were accomplished without the use of AO due to the prevailing paradigm that speckle imaging is what one does if one can't afford an AO system. Avoidance of speckle imaging stems from the awareness that the spectral SNR of the image energy spectrum is *bounded from above by unity even for bright objects* (Dainty 1984); hence, hundreds and sometimes thousands of frames must be averaged to increase SNR($\mathbf{u}$). For very dim, extended objects, the number of frames required to deconvolve a significant part of the diffraction-limited object spectrum can be truly staggering. These facts have resulted in both classical speckle and ensemble implementations of iterative algorithms being avoided in favor of long-exposure AO imaging and RL deconvolution. The statement regarding the upper bound of the energy spectrum SNR, however, is only true regarding speckle imaging *without AO*. With AO running, the number of frames required in a data ensemble can be dramatically reduced (Roggemann, et al. 1994; Roggemann & Welsh 1996). Further, theory (Fried & Hench 1989, Tyler & Matson 1993; Welsh 1995) and experiment without AO (Tyler, et al. 1994) have shown the use of exposure times much longer than those conventionally regarded as "short" can also boost the energy spectrum SNR.

Finally, a significant advantage to the use of classical speckle *with* AO is the ability to use ensemble statistics to estimate the SNR distribution of the energy spectrum. Since it has been observed that the spectral band over which the object phase can be extracted from the bispectrum coincides with the spectral band where SNR($\mathbf{u}$) for the energy spectrum is greater than unity (Matson & Roggemann 1992), estimation of SNR($\mathbf{u}$) provides a valuable quantitative indicator of how the deconvolution should be regularized. Specifically, a regularizing filter can be designed as a functional of SNR($\mathbf{u}$).

4. Observing techniques and instrumentation

In this section, I will discuss some observing techniques and instrumentation methods found to be of practical value. Specifically, I will present techniques for PSF estimation along with ideas and instrumentation to increase data SNR. I have chosen to emphasize PSF estimation rather than practical implementation of "blind" deconvolution methods for several reasons: First, while relieving the observer of the requirement for PSF observations, blind methods are not well-tested in the field (*viz.*, "practical implementation of blind deconvolution" is not a well-defined phrase at the moment); second, blind deconvolution methods are generally relatively time-consuming because the object and one or more PSFs

are jointly estimated; third, fixed aberrations common to each frame of a data ensemble look like the object to a speckle blind deconvolution scheme. These fixed aberrations include the annoying "waffle mode" artifacts present when using a Shack-Hartmann wavefront sensor with the Fried geometry.

4.1. Estimating the PSF for deconvolution

The simplest way to estimate the effective PSF for use in a deconvolution algorithm is to observe an isolated, single star. Since the effectiveness with which the AO actually corrects atmospheric turbulence depends on several factors, however, some effort must be made to ensure observing conditions are consistent between observations of the science object and the calibration star; otherwise, the PSF estimated from a star observation can differ significantly from the PSF actually blurring the associated science object images. Such *miscalibration* can result in unnecessary residual blur or artifacts in processed imagery.

For NGS AO, one should ensure the WFS subaperture intensity (or SNR) is the same between object and cal star observations; this helps keep the AO error from WFS noise at the same level. Note that unresolved stars with the same magnitude (in the WFS band!) as an extended, diffuse science object will yield a better WFS SNR because they are compact. For LGS observations, the tilt references for both observations should be roughly the same brightness. Also, the calibrator should be near the science object (within a few degrees), since wind buffeting of the telescope and enclosure, as well as turbulent flow around the dome, creates effective "seeing" conditions that depend on the telescope pointing angle. Also, large changes in pointing between the object and cal star can cause large enough rotations and flexures in the optical path that fixed aberrations unseen by the WFS can differ significantly between pointings.

Finally, the cal star should be close the science object so the star can be frequently observed while acquiring object data. This is because even if great care is taken selecting a calibrator, seeing transients during and between observations can cause miscalibration effects during deconvolution. If one is observing with long exposures, several excursions of the local seeing can be averaged into the exposure with no hope of detection. Unless identical fluctuations occur during observation of the cal star, miscalibration is the inevitable result. Interrupting science object image acquisition with frequent, short observations of the cal star can minimize this effect. Note that many AO systems include a set of ND filters for use when the only nearby calibrator candidate is excessively bright (Note also that the laws governing luck and stellar distribution rarely work to provide bright stars anywhere near an object of astrophysical interest).

While miscalibration may be of great or small consequence, depending on its severity, the desire to avoid it entirely has motivated attempts to synthesize an effective PSF from WFS data (Véran, et al. 1997). Residual errors in AO correction include aberrations observed by the WFS but not corrected by the DM due to fitting error or servo lag and two types of aberrations not observable by the WFS: atmospheric turbulence with spatial frequency higher than the WFS can measure and fixed aberrations unseen by the WFS because they are accumulated in the imaging camera beam path (*non-common path errors*; see Ellerbroek & Tyler 1999) . Of these three types of errors, WFS data can be used to correct the first two. Roggemann & Meinhardt (1993) use the DWFS

technique with AO-compensated data to reduce servo lag error. Véran, et al. use theoretical covariances to estimate residual turbulence errors unseen by the WFS from the statistics of the turbulence errors observable by the WFS. Note that while both techniques provide valuable information about PSF evolution, the average atmospheric PSF is symmetric, and details of its morphology are not as valuable for recovering very fine structure as information about SNR($\mathbf{u}$) or details of PSF asymmetry due to unsensed fixed aberrations.

4.2. Pupil masking

An interesting technique to increase SNR($\mathbf{u}$) over a reduced set of spatial frequencies is *pupil masking* (also called "aperture masking," "optical aperture synthesis," and "non-redundant masking"). With this method, a mask with several small openings (like WFS lenslet openings, these are called "subapertures"), each with diameter $\approx r_o$ when scaled to the size of the actual telescope pupil, is placed at a pupil image or in the collimated beam. The advantage to this technique is apparent when one considers image formation as an interferometric process. In the interferometric model, the pupil of an Earth-based telescope is modeled as an ensemble of r_o-size coherent regions, each pair of which form an interference pattern on the detector. The fringe patterns have a variety of orientations and modulation frequencies depending on the orientation and separation of these patches, and the incoherent sum of the fringes is the detected image. In an unobscured telescope pupil in the presence of atmospheric turbulence, there may be many r_o patches separated by identical distances and having the same orientation. This means the fringe pattern in a given direction on the detector and with a given frequency is actually the sum of many *redundant* fringe patterns. Due to turbulence, the phase difference between any two redundant fringe patterns is a random variable, reducing the visibility of the sum fringe pattern at any spatial frequency and limiting the resolution of the detected image. Adaptive optics increase resolution by reducing the relative shift of the individual fringe patterns and increasing the spatial coherence of the field incident at the telescope pupil, thus increasing the visibility of all patterns at a particular frequency. Masking, however, simply reduces or eliminates the redundancy (Tyler 2000b). While the phases of the resulting fringe patterns are still randomized by turbulence, the resulting increase in fringe *amplitude* SNR can be used to enhance phase recovery by such methods such as speckle imaging or self-calibration. Of course, masking the light-gathering area of the telescope decreases the total signal and passes a reduced set of spatial frequencies, creating a somewhat complicated optimization problem. However, in addition to increasing the visibility SNR of the spatial frequencies that *are* measured (under the right conditions), pupil masking renders the resulting image data less susceptible to temporal and spatial PSF miscalibration (Haniff & Buscher 1992). The first astronomical observations using pupil masking were made by Haniff, et al. (1987). Haniff and a variety of collaborators went on to use this method for several very interesting investigations in the 1990s (e.g., see Haniff, Scholz, & Tuthill 1995), and recent masking measurements have been on the Keck I (Monnier, et al. 1999; Tuthill, et al. 2000a, Tuthill, et al. 2000b). Although Haniff & Wilson (1994) explored the use of masking with AO using a simulation, the potential of marrying these two techniques has not yet been realized.

Typically, a nonlinear MAP algorithm such as MEM must be used to estimate spectral components in bands left unsampled by the mask.

4.3. Frame selection

Another advantage to short-exposure ensemble imaging is the option of using only some subset of the total collected data for deconvolution (Roggemann, Stoudt, & Welsh 1993). The *frame selection* technique can be implemented subjectively using a "chi-by-eye" approach, but since an optimization problem is apparent, a more quantitative approach is desirable. The competing effects in frame selection are the reduction in the number of frames and the increase in the ensemble quality of the remaining frames.

Without AO, one can define an "effective" seeing parameter using the ensemble statistics *of the selected data*. With AO, one can still define an equivalent "compensated r_o" (Cagigal & Canales 2000) and, again, an effective compensated seeing after frame selection. Let r_o here stand for the AO-compensated seeing parameter as estimated using the statistics of the image ensemble prior to selection, and let r_{os} stand for the same parameter calculated using the statistics of the selected data subset. As an image quality metric, I choose the ensemble SNR distribution for the image energy spectrum, since the energy spectrum is the squared amplitude visibility, or contrast, of the fringe pattern with vector frequency $\mathbf{u}$ contributing to the image. The energy spectrum SNR is

$$\mathrm{SNR}(\mathbf{u}) \;=\; \frac{<|I(\mathbf{u})|^2>}{(\mathrm{var}\{|I(\mathbf{u})|^2\})^{1/2}} \tag{28}$$

At frequencies beyond the "seeing spike," $<|I(\mathbf{u})|^2>$ can be written (Dainty 1984)

$$<|I(\mathbf{u})|^2> = \; 0.435 \left(\frac{r_o}{D}\right)^2 H_d(\mathbf{u}). \tag{29}$$

For simplicity, let's assume the variance of the energy spectrum remains unchanged by frame selection. This is the most pessimistic assumption regarding the effect of selection on the SNR. Now let S be the number of frames remaining in the ensemble after selection. Then, since we are assuming the variance is unchanged by selection,

$$\mathrm{SNR}(\mathbf{u}) \;\sim \sqrt{M}\; r_o^2, \tag{30}$$

where M is the number of frames in the ensemble prior to selection, and

$$\frac{\mathrm{SNR}_s(\mathbf{u})}{\mathrm{SNR}(\mathbf{u})} \;=\; \sqrt{\frac{S}{M}} \left(\frac{r_{os}}{r_o}\right)^2. \tag{31}$$

For the SNR to improve as a result of frame selection, then,

$$\frac{M}{S} \;<\; \left(\frac{r_{os}}{r_o}\right)^4. \tag{32}$$

Clearly, although reducing the number of frames reduces the ensemble SNR, the reduction is more than offset if frame selection results in even a small increase

in the effective seeing, even without considering a variance decrease. Green & Hunt (1999) demonstrate the use of several frame selection metrics, such as Fisher information, and Schöller, et al. (1996) use an informal frame selection for uncompensated classical speckle and to deconvolve AO imagery of the Herbig Ae/Be binary NX Puppis.

4.4. WFS control of a fast shutter

Finally, one may consider an automated, real-time data selection for use with long-exposure imaging. Bloemhof, et al. (2000) suggest using the WFS to control a fast camera shutter, closing the shutter when seeing (as measured by the WFS) degrades below some threshold condition. Development of an appropriate metric for optimization and sensitivity of the resulting image quality to the threshold are interesting topics awaiting investigation.

5. Anisoplanatism in AO imaging

Up to this point, I have treated deconvolution assuming the blur spot created by an unresolved point source is invariant over the imaging field. Unfortunately, this is only true over a limited field about the (laser or natural) guidestar; variation of the blur spot over the field is referred to as *anisoplanatism*.

The simplest treatment of the effects of turbulence on astronomical imaging assume the atmosphere can be modeled as a single "phase screen" in the same plane as the telescope pupil. While this approximation in accurate for many purposes, reality is a little more complicated: First, while turbulence effects are strongest at lower altitudes, they are distributed over tens of kilometers. Second, the *dominant* effects often occur some distance from the pupil; for example, the mean altitude of atmospheric turbulence at the Gemini site on Mauna Kea is 6.5 km (Ellerbroek & Tyler 1998). The second of these realities means that if, in going from zero-th order to first, one models the atmosphere again as a screen but now at some finite altitude h, a pair of light rays eminating from the object does not accumulate the same phase difference as a pair of rays eminating from an angularly-distant guidestar (Fig. 4).

This model accounts for the observed results of anisoplanatism. In AO images of crowded star fields wider than a few arcseconds, variation of the PSF is seen as blur spots with morphologies depending on the distance between a particular star and the brightest star in the field (Marshall, et al. 2000). Star blur spots will subtend a smaller angle if they are closer to the guidestar than if they are farther away. In this case, "deconvolution" actually ceases to have meaning in the strict mathematical sense of the word, since the form of the convolution integral (3), along with the techniques used in Equations (10), (11), and (15), requires a shift-invariant kernel. Indeed, the very nomenclature "PSF," implying as it does linear analysis, can only be used legitimately in a local sense and not in reference to the whole image.

Along with varying blur angle, another effect of anisoplanatism is an asymmetry or *elongation* of blur spots over the field. The blur spot from any particular star will be elongated in the direction of the axis separating the star and the AO guidestar. To see why this is so, let us consider Figs. 4 and 5 in some detail. Let $\mathbf{r}_1$ and $\mathbf{r}_2$ be coordinates in the plane of the telescope pupil separated by

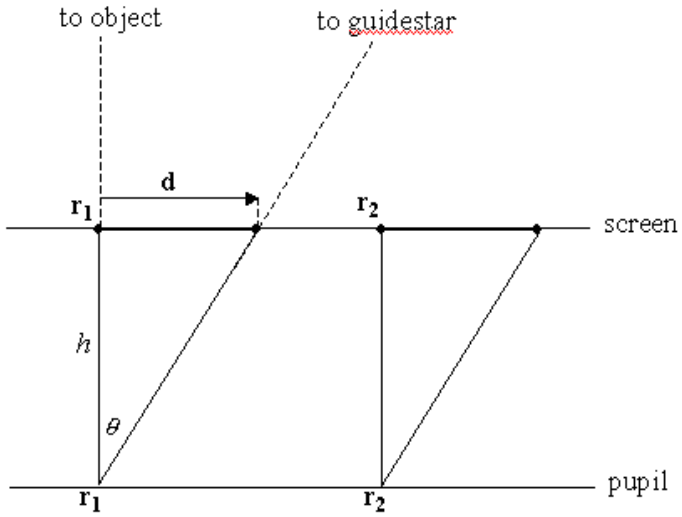

Figure 4. Rays eminating from object and guidestar with the separation axis parallel to the spatial frequency vector $\mathbf{u} = \mathbf{r}_1 - \mathbf{r}_2$.

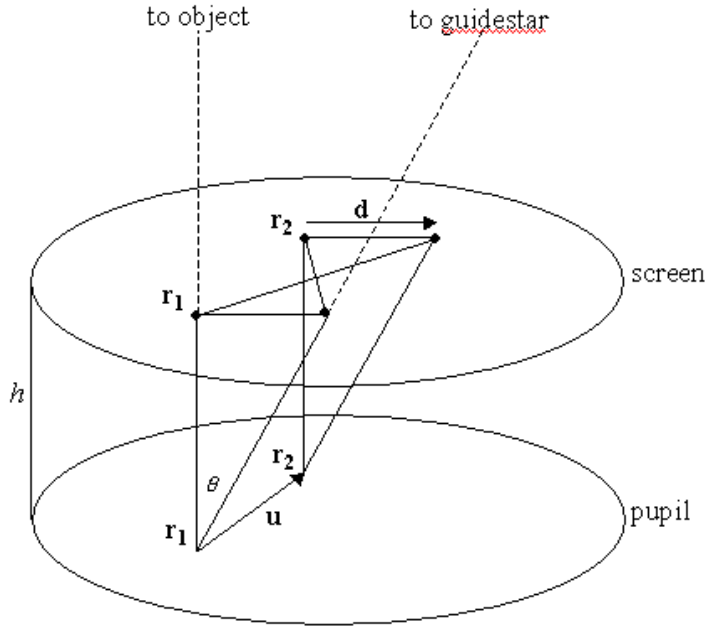

Figure 5. Rays eminating from object and guidestar with the separation axis normal to the spatial frequency vector $\mathbf{u}$.

u. Also, let the atmospheric turbulence at some instant in time be represented by a single, thin phase screen at at some non-zero alitude h above the pupil. The optical transfer function for a single random phase screen with Gaussian, wide-sense stationary statistics is given by

$$H(\mathbf{u}) \;=\; H_d(\mathbf{u})\exp\{-\frac{1}{2}D_\phi(u)\}, \tag{33}$$

where $H_d(\mathbf{u})$ is again the diffraction-limited OTF and $D_\phi(\mathbf{u})$ is the *phase structure function*; *viz.*, the variance of the difference between electric field phases separated in the pupil by **u**. Note the form of the structure function in terms of the pupil diameter and r_o was given in (5).

Projecting the pupil onto the screen identifies a coordinate system in the plane of the screen identical to the pupil plane coordinates. Now let two rays from the object intercept the screen and fall on the pupil at $\mathbf{r}_1$ and $\mathbf{r}_2$. With no loss of generality, one may choose the direction of the object rays to be normal to the planes of the pupil and screen, so the object rays also intercept the screen at $\mathbf{r}_1$ and $\mathbf{r}_2$. If the AO guidestar is very close to the object, then two rays from the guidestar will likewise fall on the pupil at $\mathbf{r}_1$ and $\mathbf{r}_2$ having intercepted the screen very near $\mathbf{r}_1$ and $\mathbf{r}_2$. In this case, the wavefront sensor will measure phases from the guidestar rays at $\mathbf{r}_1$ and $\mathbf{r}_2$ almost identical to the phases actually imposed on the two object rays by traversing the screen. The signals sent to the deformable mirror (DM) will then (neglecting several real-world effects) correctly conjugate the phases of the object rays. If, however, the object and guidestar are separated by some significant angle θ, then the two guidestar rays will intercept the screen at the points $\mathbf{r}_1 + \mathbf{d}$ and $\mathbf{r}_2 + \mathbf{d}$, where $|\mathbf{d}| = h\,\tan(\theta)$ (see Fig. 4).

One can write the electric field phase after AO correction at any point in the pupil as the difference between the phase $\phi_a(\mathbf{r})$ accumulated by the object ray after passing through the screen and the phase $\phi_c(\mathbf{r},\theta)$ imposed by the DM after the WFS measures the phase at $\mathbf{r}$ from a guidestar separated from the object by θ. The stucture function is then

$$D_\phi(\mathbf{u}) \;=< [\phi_a(\mathbf{r}_1) - \phi_c(\mathbf{r}_1+\mathbf{d}) - \phi_a(\mathbf{r}_2) + \phi_c(\mathbf{r}_2+\mathbf{d})]^2 > . \tag{34}$$

For the purposes of this discussion, the important terms in the expansion of the above are the covariances $\phi_a(\mathbf{r}_1)\phi_c(\mathbf{r}_2+\mathbf{d})$ and $\phi_a(\mathbf{r}_2)\phi_c(\mathbf{r}_1+\mathbf{d})$. Both terms are positively signed, so higher correlation between $\phi_a(\mathbf{r}_1)$ and $\phi_c(\mathbf{r}_2+\mathbf{d})$ or between $\phi_a(\mathbf{r}_2)$ and $\phi_c(\mathbf{r}_1+\mathbf{d})$ *increases* the numerical value of $D_\phi(\mathbf{u})$ and *decreases* $H(\mathbf{u})$. The difference in the effect of object-guidestar separation on spatial frequencies vectors **u** aligned parallel and perpendicular to **d** now becomes apparent. Figure 5 shows the relevant light rays and vectors for the case where **u** is normal to **d**: In this case, the distances $|\mathbf{r}_1 - (\mathbf{r}_2 + \mathbf{d})|$ and $|\mathbf{r}_2 - (\mathbf{r}_1 + \mathbf{d})|$ both increase with increasing θ. The covariances $\phi_a(\mathbf{r}_1)\phi_c(\mathbf{r}_2 + \mathbf{d})$ and $\phi_a(\mathbf{r}_2)\phi_c(\mathbf{r}_1 + \mathbf{d})$ thus both become smaller as θ and $|\mathbf{d}|$ increase, and their respective contributions to the magnitude of $D_\phi(\mathbf{u})$ both become smaller. In Fig. 4, however, **u** is parallel to **d**. In this case, it is apparent that while $|\mathbf{r}_1 - (\mathbf{r}_2 + \mathbf{d})|$ increases with $|\mathbf{d}|$, $|\mathbf{r}_2 - (\mathbf{r}_1 + \mathbf{d})|$ actually becomes *smaller*. While $\phi_a(\mathbf{r}_1)\phi_c(\mathbf{r}_2+\mathbf{d})$ decreases, then, $\phi_a(\mathbf{r}_2)\phi_c(\mathbf{r}_1 + \mathbf{d})$ increases, with the result that for spatial frequencies parallel to the vector separating the object and guidestart, $D_\phi(\mathbf{u})$ is larger than for spatial frequencies normal to the separation vector. Since the OTF is attenuated

by $\exp\{-1/2D_\phi(u)\}$, the associated PSF forms an ellipse with the major axis pointing along $\mathbf{d}$.
In principle, iterative methods can be used to invert a space-variant mapping from $o(\mathbf{x})$ to $i(\mathbf{x})$. The general form of (3) over a field much larger than the *isoplanatic angle* (or isoplanatic "patch") is

$$i(\mathbf{x}) \; = \int d\mathbf{x} \; h(\mathbf{x}, \mathbf{x}')o(\mathbf{x}). \tag{35}$$

The major difficulty here is the requirement for some knowledge about $h(\mathbf{x}, \mathbf{x}')$. The image of an unresolved source at a single location $\mathbf{x}$ in the image field only gives information about $h(\mathbf{x}, \mathbf{x}')$ within the isoplanatic patch about $\mathbf{x}$. One might consider a "mosaic" deconvolution of fields smaller than the isoplanatic patch, but crowded stellar fields are both a notoriously difficult environment for such a maneuver and the main science driver behind research into inversion methods for anisoplanatic fields. In such a case, one might also attempt to use the stars themselves as samples of $h(\mathbf{x}, \mathbf{x}')$ over the field, but this of course presumes that the morphology of each point source in the field is actually due to blur spot variation and not, for example, a close, faint companion. With measurement of $h(\mathbf{x}, \mathbf{x}')$ the main difficulty in solving the wide-field inversion problem, innovative use has been made of theory to develop an "anisoplanatic transfer function" formalism (Fusco, et al. 2000). Disregarding the minor rules infractions of using the linear systems phrases "transfer function" and "deconvolution" where they clearly do not apply, the authors demonstrate how knowledge of the turbulence profile and AO performance can be used to estimate and invert $h(\mathbf{x}, \mathbf{x}')$ given the OTF near the guidestar. One should note that the larger the number of turbulence modes compensated by AO, the smaller the isoplanatic patch (Roddier, et al. 1993); the fortuitous consequence is that if AO performance is relatively poor and the need for further resolution significant, conventional linear deconvolution may be used over a relatively larger field.

6. Summary

Deconvolution boosts spatial spectrum components that have been attenuated and distorted with spurious phase components. Once a measurement has been made, however, deconvolution amplifies *noise* as well as signal, so the limiting resolution for any image is determined by the *SNR cutoff* frequency. No algorithm that does not include prior knowledge in some way can overcome this limit.
Without AO compensation, Richardson-Lucy and similar algorithms that deconvolve the atmospheric OTF will not increase resolution of imagery collected at a telescope larger than a few tens of cm; however, RL can be very effective with AO compensation. Classical *speckle imaging* is effective without AO and can be much more robust *with* AO because the energy spectrum SNR is no longer bounded from above by unity. Prior knowledge can be used to modify RL and other *maximum-likelihood* algorithms in *maximum a posteriori* schemes.
Calibration star observations for PSF estimation should be carried out in such a way that observing conditions and AO performance are similar to those under which the science object was observed. *Pupil masking* and *frame selection* (or

WFS control of camera integration) are techniques to optimize SNR($\mathbf{u}$) for deconvolution. New methods are under study to invert the blurring of wide-field images where anisoplanatic effects are present.

Further reading. For a thorough treatment of most aspects of AO and deconvolution, see the excellent volume by Roggemann and Welsh (1996). Carasso's (1999) paper contains an informative review of several deconvolution methods widely used in imaging as well as an excellent discussion of some key principles. Goodman's volumes on Fourier optics (1968) and statistical optics (1985) are both informative and quite readable. Dainty's 1975 monograph with its recent update (Dainty 1984) provides additional information about speckle interferometry. For a mathematically formal treatment of image deconvolution, the reader is referred to the collection of monographs edited by Stark (1987). Finally, general strategies for inverse problems in many non-imaging astronomical applications can be found in Craig and Brown (1986).

Acknowledgments. I am grateful to Prof. Andreas Quirrenbach for the privilege of speaking at the Summer School. The Air Force Office of Scientific Research provided valuable support for the writing of this article. Finally, if I've managed to learn anything at all about AO and astronomical imaging, it is due entirely to the long-suffering patience and generosity of my colleagues, frequent co-authors, and good friends, Dr. Charles L. Matson, Air Force Research Laboratory; Prof. Michael C. Roggemann, Michigan Technological University; and Dr. Brent L. Ellerbroek, Gemini Telescopes Project. I'll always be struggling to catch up.

References

Ayers, G., & Dainty, J.C. 1988, Opt. Lett, 13, 547

Bilmont, M.J., Roggemann, M.C., Tyler, D.W., Von Bokern, M.A., Voelz, D.G., & Albetski, J.A. 1992, in Proc. SPIE 1688, Atmospheric Propagation and Remote Sensing, ed. A. Kohnle & W.B. Miller (Bellingham: SPIE), 489

Bloemhof, E., Marsh, K.A., Dekany, R.G., Troy, M., Marshall, J., Oppenheimer, B.R., Hayward, T., & Brandl, B. 2000, in SPIE Proc. 4007, Adaptive Optical Systems Technology, ed. P.L. Wizinowich (Bellingham: SPIE), 889

Cagigal, M.P., & Canales, V.F. 2000, J. Opt. Soc. Am. A, 17, 903

Carasso, A.S. 1999, SIAM J. Numer. Anal., 36, 1659

Chesneau, O., Roche, M., Boccaletti, A., Abe, L., Moutou, C., Charbonnier, F., Aime, C., Lanteri, H., & Vakili, F. 2000, A&A, 144, 523

Christou, J.C. 1995, in Proc. SPIE 2534, Adaptive Optical Systems and Applications, eds. R.K. Tyson & R.Q. Fugate (Bellingham: SPIE), 226

Close, L.M. 2000, in SPIE Proc. 4007, Adaptive Optical Systems Technology, ed. P.L. Wizinowich (Bellingham: SPIE), 758

Craig, I.J.D., & Brown, J.C. 1986, in Inverse Problems in Astronomy Bristol, U.K.: Adam Hilger), 49

Dainty, J.C. 1984 in Laser Speckle and Related Phenomena, 2 ed., ed. J.C. Dainty (Berlin: Springer-Verlag), 255

Ellerbroek, B.L, & Tyler, D.W. 1998, PASP, 110, 165

Ellerbroek, B.L, & Tyler, D.W. 1999, Appl. Opt., 38, 3857

Fried, D.L., & Hench, D.L. 1990, Tech. Report TR-1033 (Placentia: the Optical Sciences Company)

Fusco, T., Conan, J.-M., Mugnier, L.M., Michau, V., & Rousset, G. 2000, A&AS, 142, 149

Fugate, R.Q., Fried, D.L., Ameer, G.A., Boeke, B.R., Browne, S.L., Roberts, P.H., Ruane, R.E., & Wopat, L.M. 1991, Nature, 353,144

Fugate, R.Q. 1995, OSA Technical Digest 23 (Optical Society of America: Wash., D.C.), 170

Gauger, A., Balega, Y., Irrgang, P., Osterbart, R., and Weigelt, G. 1999, A&A, 346, 505

Goodman, J. 1968, Introduction to Fourier Optics (New York: McGraw-Hill), 101

Goodman, J. 1985, Statistical Optics (New York: Wiley & Sons), 361

Green, J.J., & Hunt, B.R. 1999, J. Opt. Soc. Am. A, 16, 2859

Hanisch, R.J., & White, R.L. 1993, Proc. STScI Worksop on The Restoration of HST Images and Spectra II (Baltimore: STScI)

Haniff, C.A., Mackay, C.D., Titterington, D.J., Sivia, D., Baldwin, J.E., & Warner, P.J. 1987, Nature, 328, 694

Haniff, C.A., & Buscher, D.F. 1992, J. Opt. Soc. Am. A, 9, 203

Haniff, C.A., & Wilson, R.W. 1994, PASP, 106, 1003

Haniff, C.A., Scholz, M., & Tuthill, P.G. 1995, MNRAS, 276, 640

Hippler, S., Kaspar, M., Feldt, M., Weiss, R., Looze, D.P., Montoya, L., Aceituno, J., Ott, T., & Davies, R., in SPIE Proc. 4007, Adaptive Optical Systems Technology, ed. P.L. Wizinowich (Bellingham: SPIE), 41

Hofmann, K.-H., Balega, Y., Scholz, M., & Weigelt, G. 2000, A&A, 353, 1016

Jefferies, S.M., and Christou, J.C. 1993, ApJ, 415, 862

Jeffs, B.D., & Christou, J.C. 1998, in Proc. SPIE 3353, Adaptive Optical System Technologies, eds. D. Bonaccini & R.K. Tyson (Bellingham: SPIE), 1006

Jennison, R.C. 1958, MNRAS, 118, 276

Knox, K.T., & Thompson, B.J. 1974, ApJ, 193, L45

Labeyrie, A. 1970, A&A, 6, 85

Lai, O., Rouan, D., Rigaut, F., Doyon, R., & Lacombe, F. 1999, A&A, 351, 834

Lohmann, A.W., Weigelt, G., & Wirnitzer, B. 1983, Appl. Opt. 22, 4028

Lucy, L.B. 1974, AJ, 79, 745

Marais, T., Michau, V., Fertin, G., Primot, J., & Fontanella, J.C. 1991, in ESO Symp. 39 on High Resolution Imaging, ed. J.M. Beckers & F. Merkle (Garching: ESO), 589

Marshall, J., Troy, M., Dekany, R., & Dekens, F. 2000, in SPIE Proc. 4007, Adaptive Optical Systems Technology, ed. P.L. Wizinowich (Bellingham: SPIE), 632

Matson, C.L. 1992, J. Opt. Soc. Am. A, 8, 1905

Michard, R. 1996, /aap, 117, 583

Monnier, J.D., Tuthill, P.G., Lopez, B., Cruzabeles, P., Danchi, W.C., Haniff, C.A. 1999, ApJ512, 351

Mu, Z., & Plemmons, R. 2000, in Proc. SPIE 4116, ed. F. Luk (Bellingham: SPIE), in press

Mugnier, L.M., Robert, C., Conan, J.-M., Michau, V., & Salem, S. 2001, J. Opt. Soc. Am. A, in press

Narayan, R., & Nityananda, R. 1986, ARA&A, 24, 127

Northcott, M.J., Ayers, G.R., & Dainty, J.C. 1988, J. Opt. Soc. Am. A, 5, 986

Piña, R.K., & Puetter, R.C. 1992, PASP, 104, 1096

Potter, D.E., Close, L.M., Roddier, F., Roddier, C., Graves, J.E., & Northcott, M.J. 2000, ApJ540, 422

Primmerman, C.A., Murphy, D.V., Page, D.A., Zollars, B.G., & Barclay, H.T. 1991, Nature, 353, 144

Primot, Rousset, & Fontanella 1990, J. Opt. Soc. Am. A, 7, 1589

Richardson, W.H. 1972, J. Opt. Soc. Am., 62, 55

Roddier, F., Roddier, N., Northcott, M., & Graves, J.E. 1993, J. Opt. Soc. Am. A, 10, 957

Roddier, F., Roddier, C., Graves, J.E., & Northcott, M.J. 1995, ApJ,443,249

Roddier, C., Roddier, F., Northcott, M.J., Graves, J.E., & Jim, K. 1996, ApJ, 463, 326

Roggemann, M.C. 1991, Appl. Opt., 30, 4227

Roggemann, M.C., & Matson, C.L. 1992, J. Opt. Soc. Am. A, 9, 1525

Roggemann, M.C., Tyler, D.W., & Bilmont, M.J. 1992, Appl. Opt., 35, 7429

Roggemann, M.C., Caudill, E.L., Tyler, D.W., Bilmont, M.J., Von Bokern, M.A., & Matson, C.L. 1994, Appl. Opt. 33, 3099

Roggemann, M.C., Stoudt, C.A., & Welsh 1993, Opt. Eng., 33, 3254

Roggemann, M.C., Welsh, B.M., and Devey, J. 1994, Appl. Opt., 33, 5754

Roggemann, M.C., & Meinhardt, J.A. 1993,, J. Opt. Soc. Am. A, 10, 1996

Roggemann, M.C., & Welsh, B.M. 1996, Imaging Through Turbulence (Boca Raton: CRC Press)

Rousset, G., Fontanella, J.K., Kern, P., Gigan, P., Rigaut, F., Lena, P., Boyer, C., Jagoural, P., Gaffard, J.P., & Merkle, F. 1990, A&A, 230, L29

Schöller, M., Brandner, W., Lehmann, T., Weigelt, G., & Zinnecker, H. 1996, A&A, 315, 445

Schulz, T.J. 1993a, J. Opt. Soc. Am. A, 10, 1064

Schulz, T.J. 1993b, in Proc. SPIE 2029, Digital Image Recovery and Synthesis II, ed. P.S. Idell (Bellingham: SPIE), 311

Schulz, T.J., Stribling, B.E., & Miller, J.J. 1997, Opt. Expr., 1, 355

Shepp, L.A., & Vardi, Y. 1982, IEEE Trans. Med. Img., 1, 113

Stark, H. 1987, Image Recovery: Theory and Applications (San Diego: Academic Press)

Stolovy, S.R., Hayward, T.L., & Herter, T. 1996, ApJ, 470, L45

Tuthill, P.G., Danchi, W.C., Hale, D.S., Monnier, J., & Townes, C.H. 2000a, ApJ, 534, 907

Tuthill, P.G., Monnier, J.D., Danchi, W.C., Wishnow, E.H., & Haniff, C.A. 2000b, PASP, 112, 555

Tyler, D.W., & Roggemann, M.C. 1992, in Proc. SPIE 1688, Atmospheric Propagation and Remote Sensing, ed. A. Kohnle & W.B. Miller (Bellingham: SPIE), 561

Tyler, D.W., & Matson, C.L. 1993, Opt. Eng. 32, 864

Tyler, D.W., Suzuki, A.H., von Bokern, M.A., Keating, D.D.B, & Roggemann, M.C. 1994, in SPIE Proc. 2198, Instrumentation in Astronomy VIII (Bellingham: SPIE), 1389

Tyler, D.W., and Matson, C.L. 1997, Opt. Expr., 1, 347

Tyler, D.W., & Ellerbroek, B.L. 1998, Appl. Opt., 37, 4569

Tyler, D.W., Ford, S.D., Hunt, B.R., Paxman, R.G., Roggemann, M.C., Rountree, J.C., Schulz, T.J., Schulze, K.J., Sheppard, D.G., Stribling, B.E., Van Kampen, W.C., and Welsh, B.M. 1998, in Proc. SPIE 3353, Adaptive Optical System Technologies, eds. D. Bonaccini & R.K. Tyson (Bellingham: SPIE), 165

Tyler, D.W., Seldin, J.H., Schulz, T.J., Schulze, K.J., Desonia, R.D., Miller, J.J., & Schroeder, A. 1999, in MHPCC Applications Briefs, ed. S. Swanson (Kihei: MHPCC), 7

Tyler, D.W. 2000a, Tech. Report 2000-006 (Albuquerque: AHPCC)

Tyler, D.W. 2000b, in SPIE Proc. 4007, Adaptive Optical Systems Technology, ed. P.L. Wizinowich (Bellingham: SPIE), 632

Véran, J.-P., Rigaut, F., Maître, H., & Rouan, D. 1997, J. Opt. Soc. Am. A, 14, 3057

Weigelt, G., Balega, Y., Blöcker, T., Fleischer, A.J., Osterbart, R., & Winters, J.M. 1998, A&A, 333, L51

Welsh, B.M. 1995, J. Opt. Soc. Am. A, 12, 1364

Wittkowski, M., Langer, N, & Weigelt, G. 1998, A&A, 329, L39

Wittkowski, M., Balega, Y., Beckert, T., Duschl, W.J., Hofmann, K.-H., & Weigelt, G. 1998, A&A, 329, L45