LAWRENCE
LIVERMORE
NATIONAL
LABORATORY

UCRL-TR-204793

Fast modal wave-front reconstruction

Lisa A. Poyneer

June 21, 2004

This work was performed under the auspices of the U.S. Department of Energy by University of California, Lawrence Livermore National Laboratory under Contract W-7405-Eng-48.

Fast modal wave-front reconstruction

Lisa A. Poyneer

1. Introduction

In this project I apply the concepts and techniques learned in Applied Computational Harmonic Analysis (ACHA) to my research. I am currently developing fast wave-front reconstruction techniques for high-order Adaptive Optics systems. The fast reconstruction technique which we plan on using is based on the DFT. Reconstruction and optional filtering occur in the frequency domain. This is a significantly different approach than current methods of wave-front reconstruction, which use a matrix multiply to do the reconstruction. The latest advances in these methods use modal control, which effectively seeks to control individual modes (or basis functions) as opposed to the signal on a point-by-point basis. These techniques can be adapted to incorporate statistical knowledge. My research task to is figure out how to do such a modal approach with my fast reconstructor.

This project report is divided into several sections. In the first section I summarize Adaptive Optics concepts and discuss the problem of wave-front reconstruction in greater detail. Following this is a brief summary of the most applicable concepts and techniques from ACHA that will be applied to the problem. The application of these techniques is then discussed.

2. Adaptive Optics and Wavefront Reconstruction

In many applications of optical systems, the observed field in the pupil plane has a non-uniform phase component. This deviation of the phase of the field from uniform is called a phase aberration. In imaging systems this aberration will degrade the quality of the images. In the case of a large astronomical telescope, random fluctuations in the atmosphere lead to significant distortion. These time-varying distortions can be corrected using an Adaptive Optics (AO) system, which is a real-time control system composed of optical, mechanical and computational parts. Adaptive optics is also applicable to problems in vision science, laser propagation and communication. For a high-level overview, consult this web site.[1] For an in-depth treatment of the astronomical case, consult these books.[2,3]

In a typical AO system the phase aberration is measured using a wave-front sensor (WFS). The sensor divides the pupil of the telescope into smaller sections called subapertures, by a small array of square lenslets. Each lenslet forms a spot image of the reference source on a small region of a CCD. As the phase aberration of the field in the subaperture changes, the spot moves. This motion is directly related to the spatial slope of the phase across the subaperture. By estimating this movement of the spot in both directions on the CCD, the x- and y-slopes of the wavefront in the subaperture can be estimated. These slopes are then sent to a wave-front reconstruction algorithm, which estimates the phase across the entire pupil based on the local slopes. This reconstructed phase is then applied to a flexible optic called the deformable mirror (DM), which conjugates the phase aberration of the field.

Current astronomical AO systems are of moderate order. The total number of estimated phase points and hence control points on the DM is less than 350. These control points are called actuators, and the total order of the system will be referred to as a. In these systems the wave-front reconstruction is done with a vector-matrix-multiply (VMM). The vector in this case consists of the WFS slope estimates for both x- and y-directions. The matrix converts the roughly $2a$ slopes to the a phase points. The matrix is calculated as the pseudo-inverse of a forward matrix. The forward matrix describes how the WFS measures the phase aberration described by the vector of phase points. Use of a matrix model allows the use of modes. A basic VMM reconstructor uses the SVD-based psuedoinverse of the forward matrix. More advanced schemes that use a better basis than the SVD basis can be developed. This will be discussed further below. The computational costs to invert the forward matrix is $\mathcal{O}(a^3)$. The cost to apply it is $\mathcal{O}(a^2)$ multiplications. Note that this can be highly parallelized, because each of the a phase points can be calculated independently of the others.

We are most concerned in this project with high-order (so-called 'Extreme' AO) systems that will be used for high contrast imaging. Using a large number of correction points, better compensation for atmospheric turbulence will be achieved, facilitating high-contrast imagery. A main effort of my current research is directed towards such a project,[4] now in the Conceptual Design stage with the Gemini Observatory as the Extreme Adaptive Optics Coronagraph (ExAOC) project. The end goal of such a system is to directly image an extra-solar planet in orbit around a star other than the sun.

For this system the number of phase points is around $a = 3500$. This is a factor of ten increase beyond present AO systems. In addition, to keep up with temporal changes in the atmosphere, the control system will run at 2.5 kHz. At this size and speed the costs of the VMM are prohibitively expensive. Initial computational costing indicates that a VMM would require substantially more computational power than we expect to have available. An asymptotically faster algorithm for wave-front reconstruction is necessary.

A. Fourier Transform wave-front reconstruction

Fourier transform wave-front reconstruction (FTR) is just such a method. Instead of describing the interactions of the phase points and the WFS measurements in a matrix, it uses a spatially-invariant model of the process of WFS measurement and inverts this process using a filter. The main hurdle to making this method work successfully is that the WFS measurements are only available inside a circular aperture. Fig. 1 shows an example setup. Given the circular aperture, the WFS slopes are available for only the valid subapertures (shown in blue). These are reconstructed to the phase points as shown by the red dots. The reconstruction must be done on a square grid at least two samples wider than the widest part of the aperture. The size of this square grid surrounding the aperture is $N \times N$.

If the slopes are not properly managed outside the valid aperture, the reconstruction will fail. With proper slope management, the true phase can be correctly reconstructed. After this slope management process (which requires relatively little computation) the x- and y-slopes signals are transformed with the FFT. The reconstruction filter is then applied in the frequency domain. The phase estimate is obtained with use of the inverse FFT. This method has been developed, fully

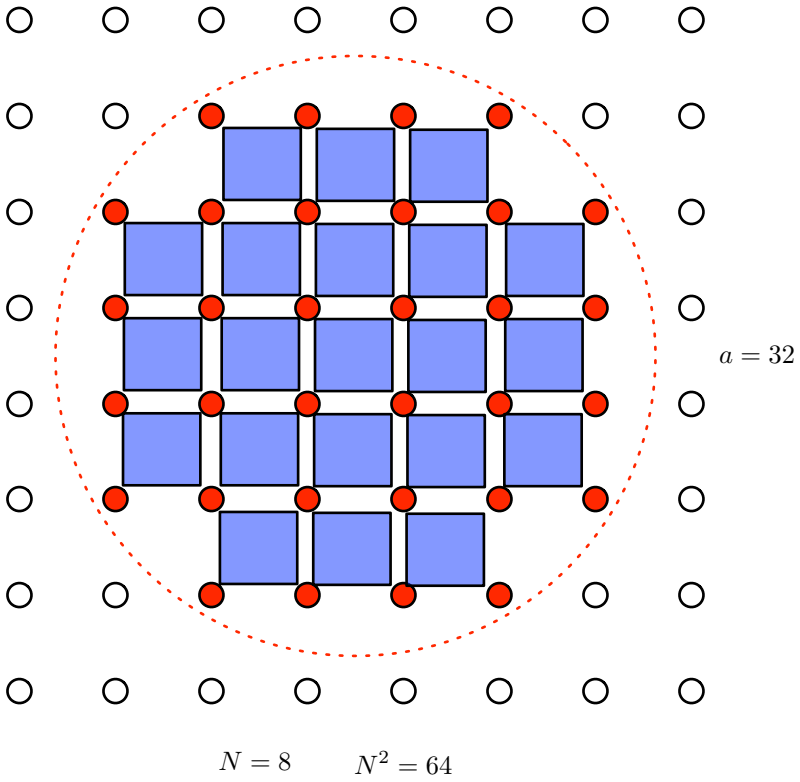

Fig. 1. The aperture is defined by the dashed red line. WFS slope measurements are available for the subapertures shown by blue squares. These are then reconstructed to the estimated wave-front phase at the phase points/actuators shown by the red dots.

analyzed[5] and experimentally validated at the Palomar observatory.[6]

Use of the FFT makes this method have computational costs of $\mathcal{O}(a \log a)$. (More precisely it is $\mathcal{O}(2N^2 \log N)$, but for systems of reasonable size $a \simeq .78N^2$.) When analyzed for the ExAOC system, this amount of computation is small enough that the whole FTR process will require just one of the four Xeons in the proposed quad-processor control computer. For comparison, a VMM for this system size would require 150 times more computation than FTR.

FTR has some additional benefits. The use of the filter allows easy modification of the control structure. Whereas modification of the matrix used in the VMM requires expensive off-line matrix inversion, the cost to update the filter is just the cost to write the N^2 coefficients. The filtering framework also fits in very nicely, with additional operations such as misalignment correction (implemented with sub-pixel shifting filters), noise reduction (implemented with low-pass filters) and DM pre-compensation simply added into the reconstruction filter. There are some weaknesses to FTR, however. It is in general not capable of implementing operations which are spatially variant. The method as it stands does not incorporate any knowledge of the atmosphere, which can be done to a VMM matrix with regularization or a better modal basis.

B. Optimal modal wave-front control

The basic VMM can be thought of as a method of zonal control. In this way individual phase points are estimated, and each actuator is controlled directly as a phase value. FTR as presently developed is a method of zonal control. An alternative method of control is called modal control. Instead of estimating and controlling the phase at specific points (which correspond to the pixel basis for an image) the phase is instead represented as coefficients to a modal basis set. As long as the modes which are reconstructed are orthogonal, they can be controlled individually. This allows the SNR of each mode to be estimated, and the control of each mode adjusted by a gain parameter to best compensate the atmosphere. Modes with high SNRs can be corrected for more aggressively, while modes with low SNRs can be corrected slightly or not at all. For more details on the theory, see this paper.[7] The term 'optimal' comes from the fact that the optimal gains are estimated during closed-loop control and the control matrix for the VMM is adapted. This process involves two parts: selection of the best modal basis to do the reconstruction and control in, and then implementation of the gain estimation in closed loop for that modal basis.

C. Research problem summary

For the ExAOC project I am collaborating with Jean-Pierre Veran, who developed an optimal modal control scheme for the Altair project[8] , which was also for the Gemini telescope. The Altair system runs at 2kHz, but the system is very small in comparison to ExAOC - it has only 136 actuators and modes. For that project Veran determined the best modal set and implemented the closed-loop gain estimation.

My research problem is how to adapt this modal technique to FTR for high-order systems. I cannot chose the basis, because the DFT has already set the basis. However, the DFT basis is clearly not an ONB within just the aperture. The first major task of my research is to determine what kind of modal set FTR uses and whether or not this can be adapted to modal control. The techniques of ACHA have enabled me to make significant progress on this problem.

3. Applied Computational Harmonic Analysis

Applied Computational Harmonic Analysis addresses the question "How can we analyze and synthesize signals of interest in an efficient and meaningful way?"

As a general methodology, ACHA choses a basis set of functions to represent a signal. Using the terminology common in AO, each member of the basis is a mode. The signal is projected onto this set to generate the coefficients. The coefficients are used to represent the signal. Many different bases are possible, from global (e.g.Fourier basis) to local (e.g. wavelet basis). It is desirable that the basis allow both analysis of the signal and synthesis. Analysis and synthesis with the basis set should also be computable quickly. It is also desirable that the basis be efficient and require as few coefficients as possible to well-approximate the signal. These same problems come up in wave-front reconstruction. As discussed above, multiple bases are possible for reconstructing the wave-front phase. Different methods have different computational complexities.

The rest of this section covers some basics of ACHA terminology that will be used in the remainder of the report. This material is primarily drawn from ACHA lectures[9] and the course handout on orthogonal sets.

The first important concept is that of the orthonormal basis (ONB). In this project we are concerned with two-dimensional (2D) signals for finite extent. These will be referenced as signals such as $x[m,n]$, $m,n \in \{0,...N-1\}^2$. We will assume that $x[m,n] \in L^2(\mathcal{R}^{N^2})$. Our signals and basis functions will be real, but we define the inner product of two functions as (for completeness)

$$< x[m,n], y[m,n] >= \sum_{m=0}^{N-1} \sum_{n=0}^{N-1} x[m,n]\overline{y[m,n]}, \tag{1}$$

where the bar represents complex conjugation. We will drop the double varialbe $[m,n]$ or switch to vector notation where convenient. An ONB for $L^2(\mathcal{R}^{N^2})$ consists of a set of N^2 functions such that

$$\{\phi_k\}_{k\in P} \text{ satisfies } < \phi_k, \phi_l >= \delta_{k,l}, \tag{2}$$

where P is the set of parameters the $\delta_{k,l}$ is the Kronecker delta function that is one when $k = l$ and zero else. Each mode has an inner product with itself of one and inner product with any other mode of zero.

ONBs have many desirable properties. Chief among them is the ease of analysis and synthesis. The modal coefficient for any mode ϕ_k given a signal x is just

$$c_k =< x, \phi_k > . \tag{3}$$

Given the complete set of the modal coefficients, the signal is exactly synthesized as

$$x = \sum_{k\in P} c_k \phi_k. \tag{4}$$

Another useful framework besides the ONB is the frame. A set $\{\phi_k\}_{k\in P} \subset \mathcal{H}$ is a frame of $\mathcal{H}$ if for all signals $x \in \mathcal{H}$,

$$\exists A, B > 0 \text{ s.t. } A||x||^2 \leq \sum_{k\in P} | < x, \phi_k > |^2 \leq B||x||^2, \tag{5}$$

where A, B are the frame bounds. A special type of frame is called the tight frame. This occurs when $A = B \geq 1$. Then the frame bounds represent the level of redundancy in the frame. Analysis with a frame is easy; the modal coefficients for a frame are calculated exactly as before with the inner product. Synthesis with a frame depends on the actual modes in the frame. Let the operator $\mathcal{U}$ be the operator that converts from the space of the signal $\mathcal{H}$ to the space of the frame coefficients $L^2(P)$. Explicitly,

$$(\mathcal{U}x)[k] =< x, \phi_k >,\ k \in P. \tag{6}$$

This operator then defines the dual modes of the frame as

$$\tilde{\phi}_k \equiv (\mathcal{U}^*\mathcal{U})^{-1} \phi_k. \tag{7}$$

Given these dual modes, any signal $x \in \mathcal{H}$ can be analyzed and synthesized as

$$x = \sum_{k\in P} < x, \phi_k > \tilde{\phi}_k, \tag{8}$$

or

$$x = \sum_{k \in P} < x, \tilde{\phi}_k > \phi_k. \tag{9}$$

This allows a frame to be used for analysis and synthesis, even if the frame does not form an ONB. The dual mode takes care of the linear dependences. ONBs and frames will be used in the following sections.

4. Deriving the FTR modal set

The modal basis of FTR is determined by the use of the $N \times N$ DFT. This basis forms on ONB for $L^2(\mathcal{R}^{N^2})$. However, we are only concerned with the phase estimate within a smaller region of interest in the aperture. The DFT set is not an ONB over this smaller region. Furthermore, due to computational constraints, we cannot substantially modify the basis. Instead we must determine what kind of set the DFT basis is for the aperture phase, how to determine the modal coefficients for this basis from the DFT, whether or not we can synthesize the signal from the coefficients, and finally if we can control each mode individually as in optimal modal control.

A. The DFT ONB

The method of FTR uses the two-dimensional DFT for its calculations. For a signal $x[m,n]$ of size $N \times N$, the 2D DFT of it is defined as

$$\tilde{X}[k,l] = \frac{1}{N} \sum_{m=0}^{N-1} \sum_{n=0}^{N-1} x[m,n] \mathrm{Exp}\left\{ \frac{-j2\pi(km+ln)}{N} \right\}. \tag{10}$$

Using the inner-product notation, we can alternatively express

$$\tilde{X}[k,l] =< x[m,n], \phi_{k,l}[m,n] >, \tag{11}$$

where the ONB is defined as

$$\{\phi_{k,l}[m,n]\}_{k,l \in \{0,\ldots,N-1\}^2} = \frac{1}{N} \mathrm{Exp}\left\{ \frac{j2\pi(km+ln)}{N} \right\} \tag{12}$$

When $x[m,n]$ is a real signal, the DFT $\tilde{X}[k,l]$ is Hermitian and we can define a entirely real ONB made up of pure sines and cosines. For the purposes of this project, we will define and number the modes as follows. Normalizing correctly, we obtain the $N^2 - 2$ sines

$$\sigma_{k,l}[m,n] = \frac{\sqrt{2}}{N} \sin\left(\frac{2\pi}{N}[km+ln] \right), \ [k,l] \in \mathcal{S} - \mathcal{S}_c, \tag{13}$$

and the $N^2 + 2$ cosines

$$\chi_{k,l}[m,n] = \frac{\sqrt{2}}{N} \cos\left(\frac{2\pi}{N}[km+ln] \right), \ [k,l] \in \mathcal{S} - \mathcal{S}_c, \tag{14}$$

$$\chi_{k,l}[m,n] = \frac{1}{N} \cos\left(\frac{2\pi}{N}[km+ln] \right), \ [k,l] \in \mathcal{S}_c, \tag{15}$$

where the sets of parameters $[k,l]$ are defined as

$$\mathcal{S} \equiv \{[k,l]|\,\{k=0, 0\le l\le N/2\}\cup\{0\le k\le N/2, l=0\}\cup \tag{16}$$

$$\{1\le k\le N-2, 1\le l\le N-1k\}\cup\{N/2\le k\le N-1, l=N-k,\}\cup\} \tag{17}$$

and

$$\mathcal{S}_c = \{[0,0],[0,N/2],[N/2,0],[N/2,N/2]\}\,. \tag{18}$$

This gives us a complete set that we can generate the DFT coefficients from using the Hermitian property. For $[k,l]\in\mathcal{S}_c$, we obtain the DFT coefficients by

$$\tilde{X}[k,l] =< x[m,n], \chi_{k,l}[m,n] > . \tag{19}$$

For $[k,l]\in\mathcal{S}-\mathcal{S}_c$, we obtain

$$\Re\left\{\tilde{X}[k,l]\right\} = \frac{1}{\sqrt{2}} < x[m,n], \chi_{k,l}[m,n] >, \tag{20}$$

$$\Im\left\{\tilde{X}[k,l]\right\} = \frac{1}{\sqrt{2}} < x[m,n], \sigma_{k,l}[m,n] > . \tag{21}$$

Fig. 2 shows the 2D grid of DFT coefficients indexed by k and l, and where the sine and cosine basis functions exist. For numbering convenience, the modes are indexed in ascending order along flipped diagonals, cosines first.

B. The DFT in a circular aperture

The real DFT ONB for $L^2(\mathcal{R}^{N^2})$ was presented above. However, in the problem of wavefront reconstruction the valid actuators only exist for a smaller number of points $a < N^2$ than the computational domain (see Fig. 1).

There are two possible ways around this problem. First, the estimated phase could somehow be managed (preferably at the slope management step) to make the phase valid over the entire square grid. Then the domain has been extended back to $L^2(\mathcal{R}^{N^2})$ and the DFT ONB can be used. This possibility I disregard, for a couple of good reasons. This type of management involves estimation of a large number, about 22%, of the points in the signal. To do so with any degree of accuracy would require detailed statistical knowledge of the signal within the aperture and the assumption that the signal is stationary. This is not necessarily a good assumption in closed-loop control and the computation necessary to estimate the process statistics in real time is most likely prohibitive.

This leaves us with the second option, which is to truncate the DFT basis to be valid only inside the aperture and zero outside. We can define a windowing function $w[m,n]$ for the aperture which is one within the aperture and zero outside. The truncated DFT modes are now $\breve{\chi}[m,n]_{k,l} = w[m,n]\chi[m,n]_{k,l}$ and $\breve{\sigma}[m,n]_{k,l} = w[m,n]\sigma[m,n]_{k,l}$. We can compute modal coefficients as $< x[m,n], \breve{\chi}[m,n]_{k,l} >$. Or we can truncate the signal and use the DFT as shown above to compute $< w[m,n]x[m,n], \chi[m,n]_{k,l} >$. This means we can compute modal coefficients with the DFT during the reconstruction step itself if the phase estimate is properly windowed.

This is made possible with a new method of slope management which has been developed which makes the reconstructed phase outside the aperture be very close to flat and zero. This is

	0	1	2	3	4	5	6	7
0	χ	χ/σ	χ/σ	χ/σ	χ			
1	χ/σ	χ/σ	χ/σ	χ/σ	χ/σ	χ/σ	χ/σ	χ/σ
2	χ/σ	χ/σ	χ/σ	χ/σ	χ/σ	χ/σ	χ/σ	
3	χ/σ	χ/σ	χ/σ	χ/σ	χ/σ	χ/σ		
4	χ	χ/σ	χ/σ	χ/σ	χ			
5		χ/σ	χ/σ					
6		χ/σ						
7								

	0	1	2	3	4	5	6	7
0	0	1, 2	5, 6	11, 12	19			
1	3, 4	7, 8	13, 14	20, 21	27, 28	35, 36	45, 46	57, 58
2	9, 10	15, 16	22, 23	29, 30	37, 38	47, 48	59, 60	
3	17, 18	24, 25	31, 32	39, 40	49, 50	61, 62		
4	26	33, 34	41, 42	51, 52	63			
5		43, 44	53, 54					
6		55, 56						
7								

Fig. 2. DFT frequency grid for 8×8 case. [Left]: Only consine modes exist at frequency pairs labelled χ. Both cosine and sine exist at pairs labeled χ/σ. [Right]: Modes are numbered along flipped diagonals, cosine mode comes first.

an improvement over the older method of slope management, as shown in Fig. 3. The successful operation of this method has be verified in simulations. For reconstruction of the DFT modes from the slopes for those modes, the new method has less than .0004 MSE for estimating the modal power for all the normalized modes. This is a high level of accuracy in coefficient estimation.

So now we have a truncated DFT basis composed of $\breve{\chi}[m,n]_{k,l}, \breve{\sigma}[m,n]_{k,l}$ modes. We can get the modal coefficients easily during the reconstruction step. However, it is clear that this is no longer an ONB. It has N^2 modes for the space $L^2(\mathcal{R}^a)$. The modes are no longer normalized and they are certainly not orthogonal. This set will work for analysis, but how will it work for synthesis?

C. Frame analysis

We can use frame theory to analyze the truncated DFT basis. This set is actually a tight frame on $L^2(\mathcal{R}^a)$. We term this frame the FTR Frame. Recall that for a tight frame, the limits A and B are equivalent and the dual modes are defined by the transform $\mathcal{U}$.

To show this, let us first examine the DFT ONB case. Converting the modes $\chi_{k,l}[m,n], \sigma_{k,l}[m,n]$ to column vectors $\chi_{k,l}, \sigma_{k,l}$, we can express the transform $\mathcal{U}$ as a $N^2 \times N^2$ matrix

$$\mathbf{U} = \left[\chi_{0,0}|\chi_{1,0}|\sigma_{1,0}|...|\chi_{N/2,N/2}\right]^T. \tag{22}$$

We can easily check that this is on ONB because $\mathbf{U}\mathbf{U}^* = \mathbf{I}_{N^2}$. The dual modes are the same as the original modes, so we know that $(\mathbf{U}^*\mathbf{U})^{-1} = \mathbf{I}_{N^2}$ and hence $\mathbf{U}^*\mathbf{U} = \mathbf{I}_{N^2}$.

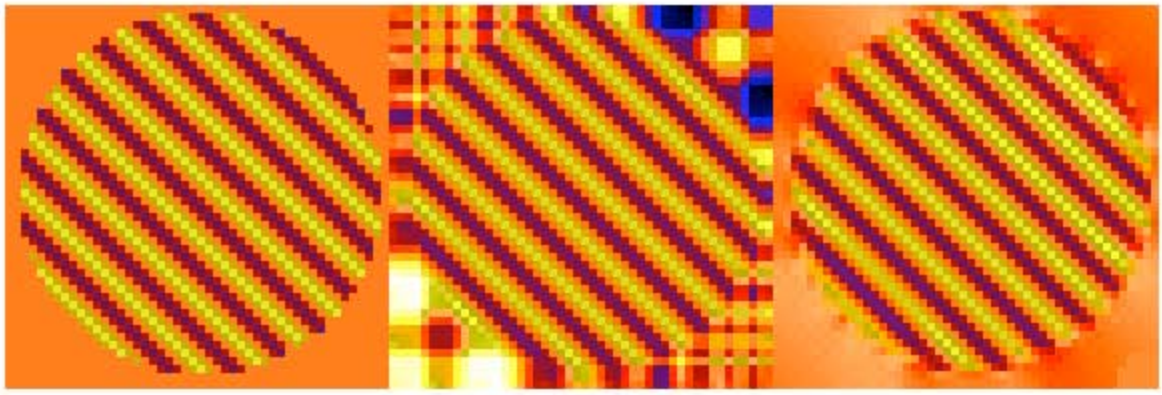

Fig. 3. [Left]: the true windowed phase. [Middle]: The reconstructed phase, unwindowed, with the old slope management method. [Right]: The reconstructed phase, unwindowed, with the new slope management method. Shown for an arbitrary DFT mode.

For the windowed DFT ONB, we use the truncated vectors as defined above $\breve{\chi}_{k,l}, \breve{\sigma}_{k,l}$. Now the matrix is N^2 rows by a columns,

$$\breve{\mathbf{U}} = \left[\breve{\chi}_{0,0}|\breve{\chi}_{1,0}|\breve{\sigma}_{1,0}|\cdots|\breve{\chi}_{N/2,N/2}\right]^T. \tag{23}$$

This new matrix $\breve{\mathbf{U}}$ is identical to $\mathbf{U}$, expect that it is missing $N^2 - a$ columns. When we examine the matrix for the dual frame $\breve{\mathbf{U}}^T\breve{\mathbf{U}}$, we see that in the matrix multiply the individual inner products that are calculated to generate the matrix coefficients are unchanged, though some are now missing. Those missing columns of $\breve{\mathbf{U}}$ have simply decimated the $\breve{\mathbf{U}}^T\breve{\mathbf{U}}$ matrix, leaving it $\mathbf{I}_a$. This means that the dual modes are exactly the original modes for the truncated set. This means that this truncated set, which we will call the FTR Frame, is a tight frame with bounds equal to 1. It is worth noting that this result holds regardless of the points over which the DFT ONB is truncated. The aperture can be of arbitrary shape.

Because the FTR frame is tight with bounds 1, we can analyze and synthesize in the same manner as with an ONB. For any signal $\mathbf{x} \in L^2(\mathcal{R}^a)$, we have

$$\mathbf{x} = \sum_{k \in P} < \mathbf{x}, \breve{\phi}_k > \breve{\phi}_k. \tag{24}$$

5. Efficiency of representation

The efficiency of a basis set is a measure of how well the power of the signal is concentrated in as few modal coefficients as possible. This is generally useful (e.g. data compression) but is also important for wave-front control. A highly efficient representation will allow better modal control, as power is concentrated in fewer modes, which presumably have higher SNR on measurement. In this section we examine the efficiency of the FTR Frame and the Altair basis in comparison to the optimal ONB.

First an explanation of the Altair basis is necessary. The basis is, to a first approximation, the discrete Zernike set. The Zernike polynomials form a radially symmetric, orthogonal set. They are

very commonly used in optics applications, since the low-order modes describe exactly common errors in optical systems such as focus and spherical aberration. Veran took the first 136 Zernikes and orthogonalized them for $L^2(\mathcal{R}^a)$ using Gram-Schmidt. The first 136 Zernikes, however, do not span the domain, so the orthogonalization produces only 91 modes. The remaining 45 modes are generated from the pixel basis for the set which has had the 91 modified Zernikes projected out. The eigenvectors of this subspace are then found and become the last 45 modes.[8]

In this section we assume the signals to be analyzed with the various basis sets are drawn from a stochastic process. If the first- and second-order statistics of the process are known (as embodied in the mean vector and the covariance matrix) the best basis is the Karhunen-Loeve (KL) basis. The KL basis is the best in the sense that it minimizes the mean-squared error when the signal is approximated by a finite number of basis functions. This concentrates the signal power in as few coefficients as possible. This means that the power of the signal is best concentrated in as few modes as possible. The KL basis is composed of the eigenvectors of the covariance matrix. The eigenvalues are the variances of the coefficients. Note that the covariance matrix will have all positive eigenvalues due to its structure.

The efficiency of representation of the signal is analyzed directly from autocovariance model. The random signal $\mathbf{x}$ is assumed to have mean $\mathbf{m}_x$ and covariance matrix $\mathbf{\Gamma}_x$. Any modal vector is denoted by ϕ_k. Using basic rules of probability, the coefficient c_k, which is the inner product $< \mathbf{x}, \phi_k >$, has mean

$$\mathrm{E}[c_k] = \phi_k^T \mathbf{m}_x, \tag{25}$$

and variance

$$\mathrm{Var}[c_k] = \phi_k^T \mathbf{\Gamma}_x \phi_k. \tag{26}$$

To analyze the efficiency of the FTR Frame and Altair basis in comparison to the optimal KL basis, models of the autocovariance matrix for reasonable situations need to be determined.

A. Open-loop atmosphere

The ExAOC system will be correcting the phase aberrations caused by atmospheric turbulence. The power spectrum of atmospheric phase aberrations can be described by a power law in spatial frequency. For the 2D case, the standard power law to use is[10]

$$P[k,l] = \left(\frac{2\pi}{D}\right)^2 0.00058 r_0^{-5/3} \left(\left(\frac{k}{D}\right)^2 + \left(\frac{l}{D}\right)^2\right)^{-11/6}, \tag{27}$$

where D is the width in meters of the signal and k, l go from 0 to $N-1$, where there are N samples across the width D, with $d = D/N$. The parameter r_0 is the coherence length of the turbulence and scales the power law. r_0 is normally on the order of 20 cm. Since it is just a multiplicative factor, we only need to calculate the variances of the modal coefficients once. The autocovariance matrix is calculated numerically from this PSD by proper sampling and use of the DFT to convert to an estimate of the autocovariance function. Then, based on distances between actuators, the autocovariance matrix is assembled. Note that this semi-analytic approach was taken because estimating the covariance matrix from Monte Carlo simulation was not exact enough. The

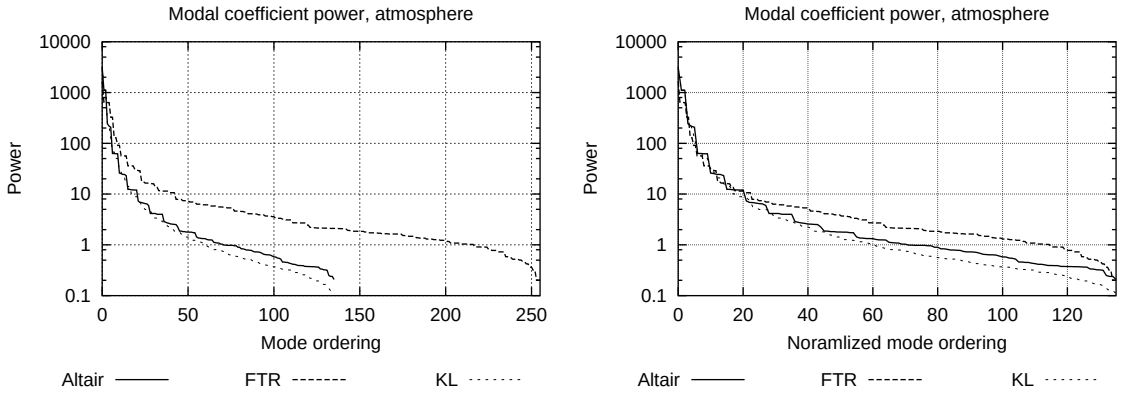

Fig. 4. The modal coefficient power, plotted in decreasing order. [Left]: versus mode rank. [Right]: versus normalized mode rank (to deal with the fact that FTR Frame has 256 modes and the other bases only 136.) The KL basis is the optimal basis. The Altair basis does very closely to the KL basis. The FTR frame decays more slowly.

sample covariance matrix converged on the theoretical value, but not quickly enough (even for large sample sizes) to ensure positive definiteness and a valid eigenvector decomposition.

Given this matrix $\mathbf{\Gamma}_a$ for the autocovariance of the atmosphere, the KL basis is computed (in Matlab with the `eig()` function). The modal coefficient variances for the FTR frame and the Altair basis are calculated as given in the above equations. Fig. 4 shows that variance of the modal coefficients for the 136 KL modes, the 136 Zernike-like Altair modes and the 256 FTR Frame modes, sorted in order of descending power. The KL basis has the best power concentration in the fewest modes. For the KL basis, 90% of the total is contained within the first 5 modes; 98% of the total power is contained with in the first 22 modes. The coefficient power drops off to 1% of the max after 10 modes. The Altair basis does nearly as well in concentrating the power in a small number of modes. For Altair the 90% mark is reached with 5 modes, and the 95% mark with 27 modes. The coefficient power drops off to to 1% of the max after 10 modes. This performance level is very close to the KL basis. The FTR Frame coefficient power decays more slowly. The 90% power mark is reached with 33 modes but the 95% power mark isn't reached until 142 modes are used. The coefficient power drops off to 1% of the max after 25 modes. The FTR Frame is not as efficient for this distribution of phase aberration, but it is still performs quite well. Note that we have plotted the coefficient power in descending order. The distribution of the power in spatial frequency space for the FTR Frame (see Fig. 2) is concentrated in a radially-symmetric manner around the lowest spatial frequencies.

Figs. 6 and 7 (at the end of the document) show specific modes from the sets that have the most and least power. The modes have been upsampled and low-pass filtered to give an estimate of what they would look like on the DM (and also for clarity given the signals are 16×16 samples.) The low-order KL modes look a lot like the modified Zernikes of the Altair basis. The analysis of this subsection should also be conducted for closed-loop phase aberrations, which will we a significantly different structure in terms of spatial frequency than the atmospheric case.

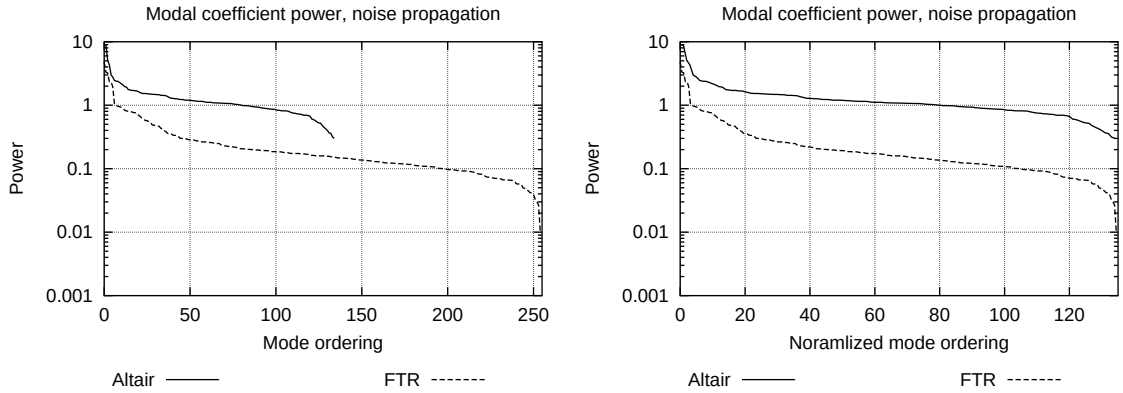

Fig. 5. Modal coefficient power due to noise propagation from WFS, plotted in decreasing order. [Left]: versus mode rank. [Right]: versus normalized mode rank (to deal with the fact that FTR Frame has 256 modes and the other bases only 136.) Though the two reconstructors produce the same amount of phase power for estimates, the noise propagates through the Altair reconstructor with 2.4 times the total power as for FTR.

B. Noise

The main source of noise in the wave-front reconstruction comes from noise in the slope estimates. It is important to study how this noise propagates through the reconstruction step and into the modes. A very reasonable model for the WFS noise is that it is white. To generate $\mathbf{\Gamma}_n$, Monte Carlo techniques are used. Many realizations of white noise are created and sent through the reconstructors (a specific matrix for the Altair basis and FTR for the FTR Frame.) These are used to estimate $\mathbf{\Gamma}_n$ using standard method of sample covariance. As above, the modal coefficient variances are calculated. These are shown in Fig. 5. The main result is that the Altair reconstruction matrix has 2.4 times the total noise power propagate through than FTR. The FTR Frame also concentrates the noise power in fewer modes. Only 30 modes in the FTR Frame have power greater than 10% of the maximum, while for Altair this number is 91. The Altair basis has more total noise and the noise is spread out less efficiently than the FTR Frame.

C. Meaningfullness

A final comment is worth making regarding efficiency and the various bases. The phase aberrations that the AO system is correcting are primarily due to the atmosphere. For both the Fourier basis and the Zernike basis, atmospheric power falls off with increasing spatial frequency or mode number. This makes both basis efficient. Recall that the power distribution of the Altair basis very closely resembled that of the KL basis. The FTR Frame has an advantage here, however, despite not closely resembling the KL basis. In closed-loop the residual phase has had most of its low-order power corrected. Most of what remains is higher order, and temporal lags can lead to errors along specific spatial frequencies. In addition, errors in the WFS measurements due to noise and errors in controlling the DM (e.g hysteresis) will also exist. For ExAOC project the proposed WFS array and the actuator array on the DM are both on a equi-spaced x-y grid. This makes the Fourier basis a natural basis for capturing these errors from these devices.

D. Summary

For a given reasonable model of atmospheric turbulence, we have compared the efficiency of the FTR Frame to the Altair basis and the optimal KL basis. The FTR is not quite as good as the best basis, but it does quite well. For all bases the coefficient power drops off very rapidly. In all cases fewer than 10% of the modes have coefficient power that is more than 0.01 times the maximum modal power. This bodes well for controlling most of the power of the wave-front in a small number of modes, as optimal modal control hopes to. The noise propagation in the reconstructors was also analyzed and the Altair matrix had 2.4 times the total noise power as FTR. The FTR Frame concentrates the noise better than the Altair basis.

The above analysis was in the case of a small system with only 136 modes. For the ExAOC case, the FTR Frame will have up to 4096 modes, corresponding to reconstruction on a grid up to the size of 64×64. Because of the power-law nature of the aberration, the efficiency results should still hold for large sizes.

6. Synthesis and modal control

As discussed early, the first part of the optimal control strategy is to find the basis to control in. The second part involves estimation of optimal gains and application of those gains to each mode. Because it is a tight frame, the FTR Frame can implement this gain adjustment directly in the frequency domain filter. Given the correspondence between FTR Frame modes and DFT coefficients (see Sec. 4.A), the modal gains can be directly multiplied into the filter at each frequency.

There is a slight snag, however, in this plan. Because the FTR Frame modes are not orthogonal, this method can actually be controlling the same signal component in multiple modes. The amount of this modal leakage is controlled by the shape of the windowing function $w[m, n]$ which was used to define the aperture. The Fourier transform of this window represents the corruption in the frequency space of the modal coefficients. Because the aperture is nearly the size of the grid, the width of its transform is small. In the Altair geometry case, the four nearest-neighbors in frequency space (recall Fig. 2) are affect by about 20%. The leakage into other modes is always under 10%, and mostly significantly below that. Since the distributions of atmospheric signal power and noise power in our model are also concentrated spatially, this should not pose a major problem. The leakage from the most powerful modes is contained nearby in modes that will most likely be controlled with a similar gain in the optimal control scheme. The most significant next step in this research project is to implement the optimal modal control and verify the above.

7. Conclusions

Using the techniques of ACHA, I have determined how reformulate the fast wave-front reconstruction method FTR into the framework of optimal modal control. The windowed DFT basis forms a tight frame for the phase signal, which allows both analysis and synthesis and as so facilitates easy gain adjustments for each mode. An analysis of efficiency for the case of atmospheric turbulence was conducted, comparing the FTR Frame and the optimal modal Altair basis to the KL basis. Though not quite as good as the KL basis, the FTR frame exhibits rapid decay of coefficient power, making it a good candidate for modal control.

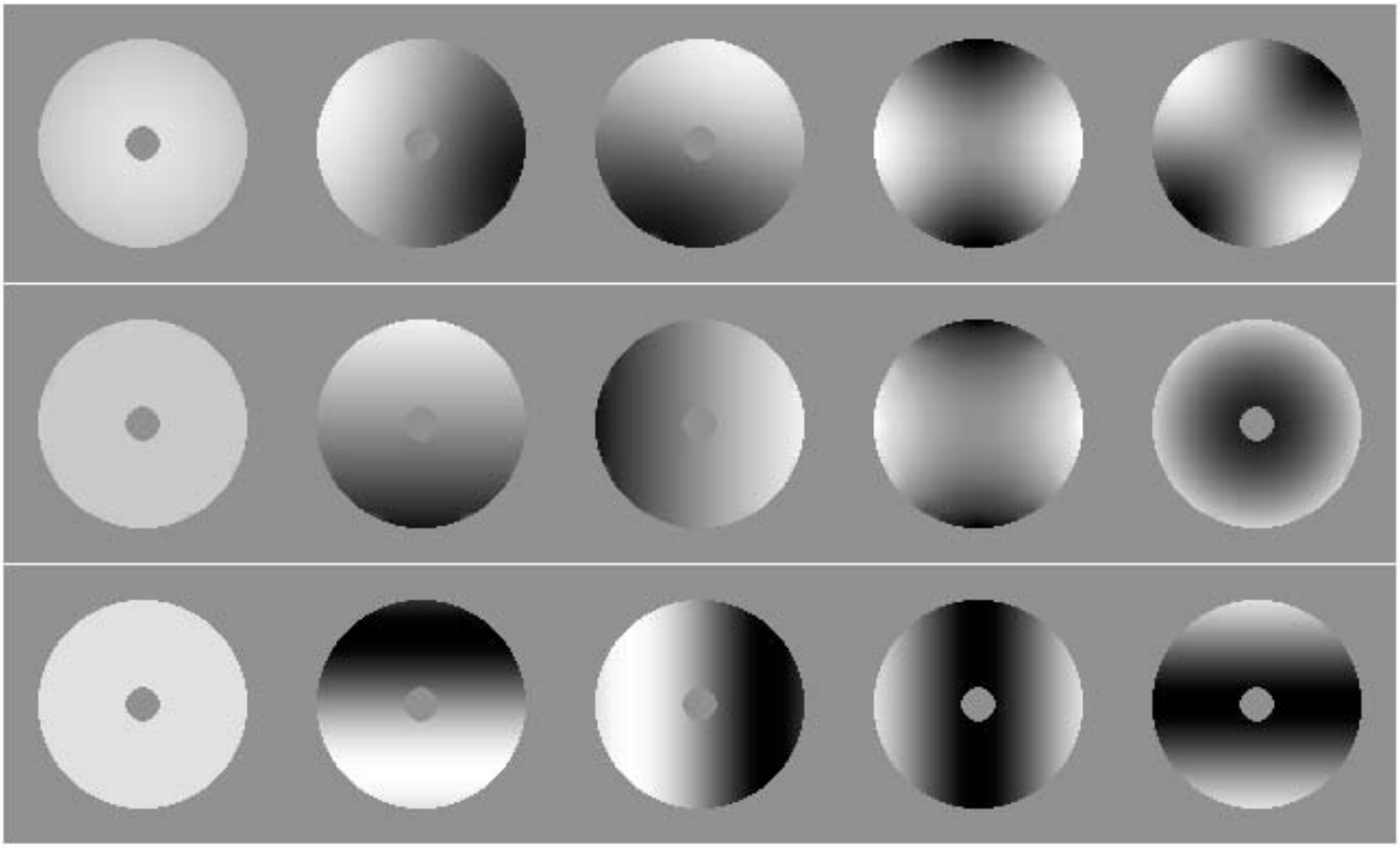

Fig. 6. The five modes with the most power under the atmospheric model, descending power left to right. Top: KL basis. Middle: Altair basis. Bottom: FTR Frame.

Acknowledgments

I would like to thank my collaborator Jean-Pierre Veran for contributions from the optimal-control side of the project.

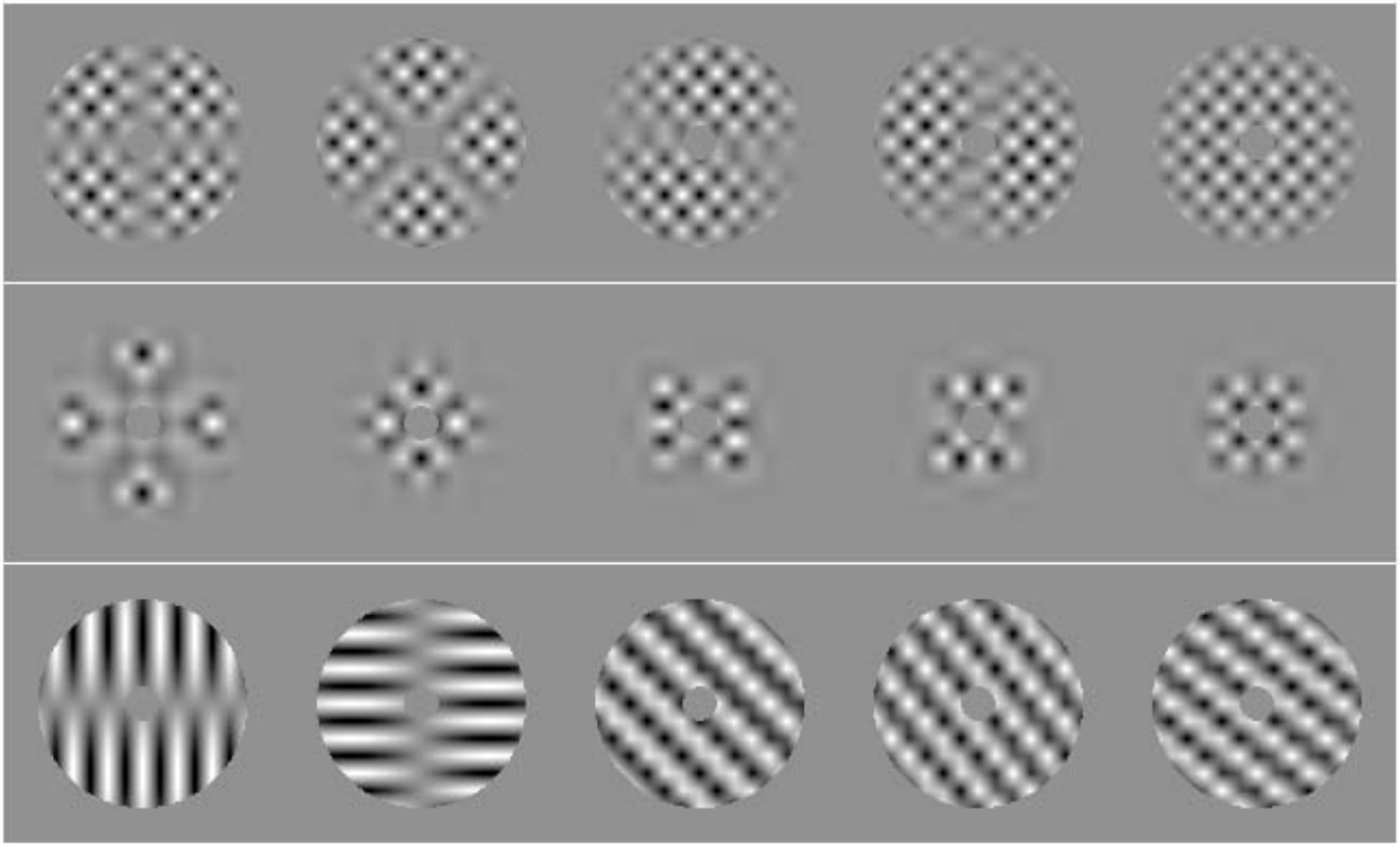

Fig. 7. The five modes with the least power under the atmospheric model, descending power left to right. Top: KL basis. Middle: Altair basis. Bottom: FTR Frame.

References

1. NSF Center for Adaptive Optics, `http://cfao.ucolick.org`

2. J. W. Hardy, *Adaptive Optics for Astronomical Telescopes,* (Oxford University Press, New York, 1998).

3. F. Roddier, Ed., *Adaptive Optics in Astronomy,* (Cambridge University Press, Cambridge, 1999).

4. B. Macintosh, et al, "Extreme Adaptive Optics Planet Imager: XAOPI," Proc. SPIE 5170, (in press 2003).

5. L.A. Poyneer, D.T. Gavel and J.M. Brase, "Fast wave-front reconstruction in large adaptive optics systems with use of the Fourier transform," *JOSA A* **19**, 2100-2111 (2002)

6. L.A. Poyneer, M. Troy, B. Macintosh and D.T. Gavel, "Experimental validation of Fourier transform wave-front reconstruction at the Palomar Observatory," *Opt. Let.* **28**, 798-800 (2003)

7. C. Dessenne, P.-Y. Madec and G. Rousset, "Optimization of a predictive controller for closed-loop adaptive optics," *App. Opt.* **37**, 4623-4633 (1998)

8. J.-P. Veran, "Altair's Optimser," Herzberg Institute for Astrophysics technical document(Victoria, Canada) (1998)

9. N. Saito, Notes from Applied Computational Harmonic Analysis , UC Davis course MAT280 (Spring 2004)

10. E. M. Johansson and D. T. Gavel, "Simulation of stellar speckle imaging," Proc. SPIE **2200**, 372–383, (1994)