Method of obtaining wavefront slope data from through-focus point spread function measurements

Samuel T. Thurman[1,2]

[1]*The Institute of Optics, University of Rochester, Rochester, New York 14627, USA*
[2]*Currently with Lockheed Martin Coherent Technologies, 135 South Taylor Avenue, Louisville, Colorado 80027, USA*
(sam.t.thurman@lmco.com)

Received June 7, 2010; revised October 26, 2010; accepted November 12, 2010;
posted November 15, 2010 (Doc. ID 129732); published December 14, 2010

A method is described for analyzing point source imagery collected with various amounts of defocus to obtain wavefront slope data at the exit pupil of an imaging system. Integration of this slope data yields the system wavefront aberration function. The method is based on a geometric optics interpretation of intensity point spread function measurements in the caustic region near focus. Algorithm performance is examined through Monte Carlo simulations. Application of the method to segmented-aperture systems is also explored. The proposed method is used to generate initial wavefront estimates for an iterative phase retrieval algorithm, significantly improving the capture range over a blind phase retrieval approach when segment tilt errors are large. © 2010 Optical Society of America

OCIS codes: 010.7350, 080.1010, 120.3940, 120.4820.

1. INTRODUCTION

There are a number of methods of characterizing the aberrations of an imaging system using a point source. Shack–Hartmann sensors reconstruct the wavefront aberration function $w(x,y)$ of a system based on pupil plane measurements of the slopes $\partial w(x,y)/\partial x$ and $\partial w(x,y)/\partial y$ [1–3]. In practice, this approach often requires additional optics to form a real image of the pupil, a lenslet array, and a dedicated focal plane array for making the slope measurements. Curvature sensing [4] involves measuring the optical intensity in two planes that are far enough from the image plane that the recorded intensity patterns look like blurred/distorted versions of the pupil. Subsequent analysis yields the wavefront curvatures $\partial^2 w(x,y)/\partial x^2$ and $\partial^2 w(x,y)/\partial y^2$ in the pupil plane. Focus-diverse phase retrieval [5,6] uses two or more intensity measurements from the region about the image plane and knowledge of the Fresnel/Fourier transformation relationship between the generalized pupil function and the field in each measurement plane to iteratively retrieve a $w(x,y)$ that is consistent with the intensity measurements. The intensity measurements often can be made with the image plane detector array while defocusing the system. This approach has minimal extra hardware requirements and the advantage of sensing the aberrations of the entire imaging system. Phase retrieval is based on Fourier or wave optics, while the other techniques described above are based on geometric or ray optics.

In this article, I describe a new wavefront sensing technique that is similar to focus-diverse phase retrieval in that it uses intensity measurements from the region about the image plane but differs from phase retrieval in that it is based on geometric optics and is noniterative. Analysis of the image domain intensity measurements yields pupil plane wavefront slope data like Shack–Hartmann wavefront sensing, but the method of obtaining the slope data is unique. Section 2 describes the wavefront sensing algorithm in detail. Algorithm performance is examined through Monte Carlo simulations in Section 3. An example application of the method for segmented-aperture systems is considered in Section 4. Section 5 is a summary.

2. ALGORITHM DEVELOPMENT

Suppose an optical system images an on-axis point source to the point P, as shown in Fig. 1. The exit pupil of the system is located in plane $z=0$, and P is in the plane $z=R$. The dashed curve in the exit pupil represents a reference sphere of radius R centered on P. The wavefront aberration function $w(x,y)$ represents the departure of the actual optical wavefront from this reference sphere. Consider ray m with coordinates (x_m, y_m) in the exit pupil. The dotted line in Fig. 1 indicates the trajectory of an ideal ray that passes through (x_m, y_m) and P, having paraxial ray heights as a function of z given by

$$r_{x,m}(z) = x_m[1-(z/R)] \qquad r_{y,m}(z) = y_m[1-(z/R)]. \tag{1}$$

The thin solid line in Fig. 1 represents the actual trajectory of ray m, which also passes through (x_m, y_m) but whose angular direction differs from that of the ideal ray by $s_{x,m}$ and $s_{y,m}$. The ray deviation angles $s_{x,m}$ and $s_{y,m}$ are equal to the slopes of the wavefront aberration function along the x and y directions, i.e.,

$$s_{x,m} = \frac{\partial w(x_m, y_m)}{\partial x} \qquad s_{y,m} = \frac{\partial w(x_m, y_m)}{\partial y}. \tag{2}$$

Also, the differences between the actual and ideal ray heights as a function of z are given by

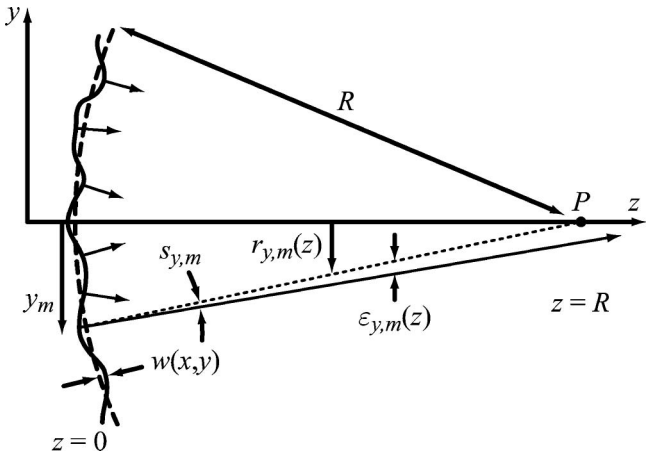

Fig. 1. Plane $z = 0$ corresponds to the exit pupil of an imaging system. Point $P = (0, 0, R)$ is the geometric image of an on-axis point source. The dashed curve in the exit pupil plane represents a reference sphere of radius R centered on P, while the solid curve represents the aberrated wavefront. The system aberrations are characterized by the wavefront aberration function $w(x, y)$. y_m indicates the height of ray m in the pupil plane. The dotted line represents the ideal ray trajectory that passes through P, with ray height $r_{y,m}(z)$. The solid line represents the actual ray trajectory, which deviates from the ideal ray height by $\varepsilon_{y,m}(z)$. The angular deviation $s_{y,m}$ of the actual ray from the ideal ray is equal to the slope of the wavefront aberration function $\partial w(x, y)/\partial y$.

$$\varepsilon_{x,m}(z) = s_{x,m} z \qquad \varepsilon_{y,m}(z) = s_{y,m} z. \tag{3}$$

Let $I_n(x, y)$ represent narrowband intensity distributions in planes $z = z_n$, for $n \in \{1, 2, \ldots, N\}$. Note that $I_n(x, y)$ represents the incoherent point spread function (PSF) of the optical system defocused by longitudinal distance $z_n - R$. In Fourier optics, $I_n(x, y)$ is related to the wavefront aberration function by

$$\begin{aligned} I_n(x, y) = \Bigg| \int\int A(x', y') \exp\left[i \frac{2\pi}{\lambda} w(x', y') \right] \\ \times \exp\left[i \frac{\pi}{\lambda} \left(\frac{1}{z_n} - \frac{1}{R} \right) (x'^2 + y'^2) \right] \\ \times \exp\left[-i \frac{2\pi}{\lambda z_n} (x x' + y y') \right] \mathrm{d}x' \mathrm{d}y' \Bigg|^2, \end{aligned} \tag{4}$$

where $A(x, y)$ is the amplitude of the optical field in the exit pupil. In geometrical optics, the intensity distribution $I_n(x, y)$ is proportional to the local density of rays in the plane $z = z_n$. When $s_{x,m}$ and $s_{y,m}$ are unknown, one could naively assume the probability that ray m passes through the point (x, y, z_n) is simply proportional to $I_n(x, y)$, i.e., the local ray density. Note that wavefront sensing is deterministic, and the association of probabilities with rays is purely heuristic. Specifically, the probability that ray m has angular deviation (s_x, s_y) based on measurement $I_n(x, y)$ is taken to be

$$\begin{aligned} p_{m,n}(s_x, s_y) = c_n I_n[z_n s_x + r_{x,m}(z_n), z_n s_y + r_{y,m}(z_n)] \\ * L(s_x, s_y), \end{aligned} \tag{5}$$

where c_n is a normalization constant, $*$ indicates a two-dimensional (2-D) convolution operation, and $L(s_x, s_y)$ is a low-pass filtering kernel used to reduce diffraction effects in $I_n(x, y)$. In cases where $I_n(x, y)$ are actually broadband intensity measurements, the visibility of interference features will be low and one would not necessarily low-pass filter the data with $L(s_x, s_y)$. Information from multiple measurements for $n \in \{1, 2, \ldots, N\}$ can be combined by assuming the $p_{m,n}(s_x, s_y)$ from different $z = z_n$ planes are statistically independent, such that the joint probability that ray m has angular deviation (s_x, s_y) is given by

$$p_{m,\mathrm{joint}}(s_x, s_y) = \prod_{n=1}^{N} p_{m,n}(s_x, s_y). \tag{6}$$

Estimates of the ray deviation angles $(s_{x,m}, s_{y,m})$ for ray m are taken as the most probable values of s_x and s_y, i.e.,

$$(s_{x,m}, s_{y,m}) = \max_{(s_x, s_y)} [p_{m,\mathrm{joint}}(s_x, s_y)]. \tag{7}$$

When $I_n(x, y)$ is measured for only $N = 1$ defocus plane, Eqs. (5)–(7) will yield wavefront slope estimates $(s_{x,m}, s_{y,m})$ consistent with all rays passing through the peak of $I_n(x, y)$, regardless of the actual spatial distribution of $I_n(x, y)$. For a modest number of measurements, however, Eq. (6) combines information from each defocus plane to yield useful slope estimates, as shown in the remainder of this section and in Section 3.

The wavefront aberration function $w(x, y)$ can be reconstructed from a set of slope estimates $(s_{x,m}, s_{y,m})$, for $m \in \{1, 2, \ldots, M\}$, using any one of several slope reconstruction algorithms [1–3]. The modal reconstruction algorithm of [7] is used here because it is noniterative. The wavefront aberration function is reconstructed as a polynomial expansion

$$w(x, y) = \sum_{k=1}^{K} c_k \psi_k(x, y), \tag{8}$$

where c_k are expansion coefficients and $\psi_k(x, y)$ are polynomial basis functions. The $\psi_k(x, y)$ obey the following orthonormality condition:

$$\begin{aligned} \sum_{m=1}^{M} \left[\frac{\partial \psi_j(x_m, y_m)}{\partial x} \frac{\partial \psi_k(x_m, y_m)}{\partial x} + \frac{\partial \psi_j(x_m, y_m)}{\partial y} \frac{\partial \psi_k(x_m, y_m)}{\partial y} \right] \\ = \delta_{j,k}, \end{aligned} \tag{9}$$

where $\delta_{j,k}$ is the Kronecker delta function, such that the expansion coefficients c_k can be directly computed from a set of slope data as

$$c_k = \sum_{m=1}^{M} \left[s_{x,m} \frac{\partial \psi_k(x_m, y_m)}{\partial x} + s_{y,m} \frac{\partial \psi_k(x_m, y_m)}{\partial y} \right]. \tag{10}$$

Figure 2 shows simulation data that is intended to provide additional insight into the method. Figure 2(a) shows an example wavefront aberration function $w(x, y)$ for a system with an exit pupil diameter $D = 3$ mm and a focal length of $f = 75$ mm. The corresponding PSFs $I_n(x, y)$, computed using Eq. (4) for $z_n = \{56, 65.5, 75, 84.5, 94\}$ mm, are shown in Figs. 2(e)–2(i). These z_n values are equivalent to $\{6, 3, 0, -3, -6\}$ wavelengths peak-to-valley of quadratic defocus aberration across the exit pupil at a wavelength of $\lambda = 632.8$ nm. The small square in Fig. 2(a) indicates the

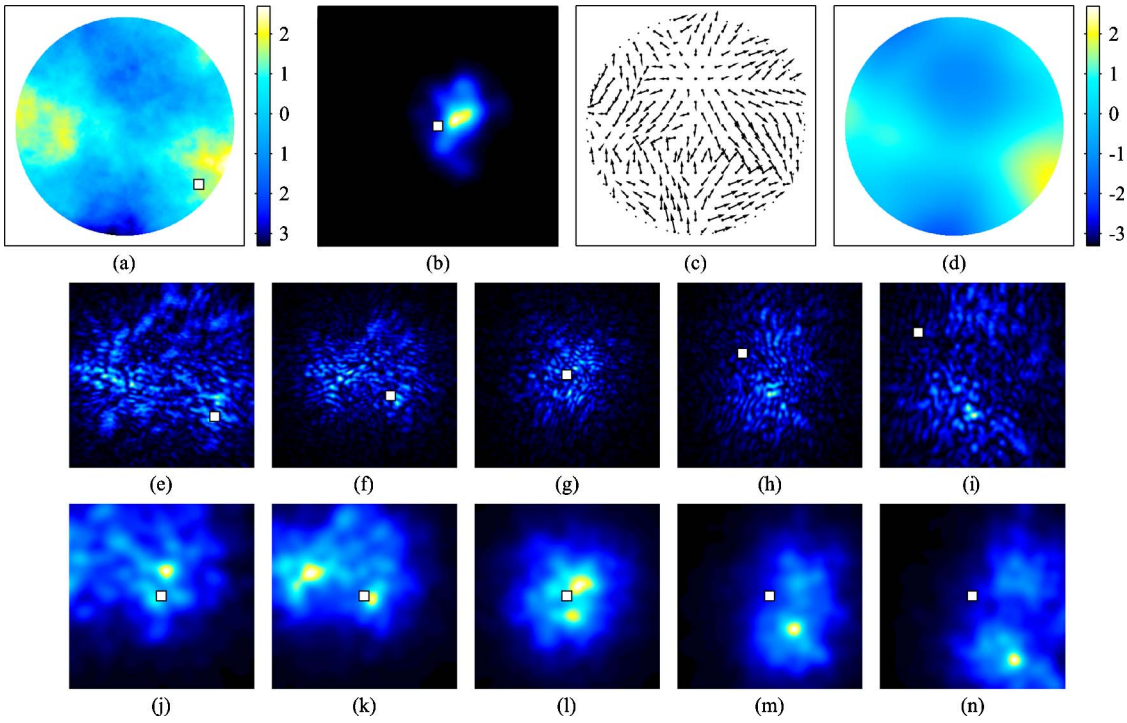

Fig. 2. (Color online) Simulation example: (a) wavefront aberration function $w(x,y)$ in units of wavelength, where the small white square indicates the pupil plane coordinates (x_m, y_m) of a particular ray m; (b) the corresponding joint probability distribution $p_{m,\text{joint}}(s_x, s_y)$ for the angular deviation of ray m computed using Eqs. (5) and (6); (c) wavefront slope data $(s_{x,m}, s_{y,m})$ at $M = 235$ points obtained through Eq. (7); (d) corresponding wavefront aberration function reconstructed using Eqs. (8)–(10) with $K = 91$ basis functions; (e)–(i) through-focus intensity measurements $I_n(x,y)$ of the system PSF, where the small white squares indicate the ideal image domain coordinates $[r_{x,m}(z_n), r_{y,m}(z_n)]$ for ray m shown in (a); and (j)–(n) the corresponding probability distributions $p_{m,n}(s_x, s_y)$ computed by low-pass filtering each $I_n(x,y)$, shifting by the ideal ray coordinates $[r_{x,m}(z_n), r_{y,m}(z_n)]$, and scaling to angular coordinates [see Eq. (5)]. The joint probability distribution $p_{m,\text{joint}}(s_x, s_y)$ shown in (b) is the product of the individual distributions $p_{m,n}(s_x, s_y)$ shown in (j)–(n) [see Eq. (6)].

location of a particular ray m in the pupil plane, with coordinates (x_m, y_m). The small squares in Fig. 2(e)–2(i) indicate the location of the ideal ray with coordinates $[r_{x,m}(z_n), r_{y,m}(z_n)]$ in each measurement plane. Figures 2(j)–2(n) show probability distributions $p_{m,n}(s_x, s_y)$ computed from each $I_n(x,y)$ via Eq. (5). These were computed using a 2-D Gaussian having a full width at half-maximum equal to $4\lambda/D$ for the low-pass kernel $L(s_x, s_y)$. This width was chosen since it is approximately twice the width of a diffraction-limited Airy spot. The corresponding joint probability distribution $p_{m,\text{joint}}(s_x, s_y)$ computed via Eq. (6) as the product of all the $p_{m,n}(s_x, s_y)$ is shown in Fig. 2(b). The location of the peak of this $p_{m,\text{joint}}(s_x, s_y)$ is used as an estimate for the wavefront slope data $(s_{x,m}, s_{y,m})$ at the location of the square in Fig. 2(a). The squares in Figs. 2(b) and 2(j)–2(n) are located at $(s_x, s_y) = (0,0)$. To obtain slope information at another point in the pupil, one traces the corresponding ideal ray through the measurement planes, computes new distributions $p_{m,n}(s_x, s_y)$, and locates the maximum of the resulting $p_{m,\text{joint}}(s_x, s_y)$. Figure 2(c) shows a set of slope data at $M = 235$ points across the pupil on a hexagonal lattice with spacing $D/16$. Figure 2(d) shows the corresponding estimate of the wavefront aberration function computed using Eqs. (8)–(10) for a set of $K = 91$ polynomial basis functions $\psi_k(x,y)$ similar in form to the Zernike polynomials [8]. This set of basis functions includes up to twelfth-order polynomial terms. The root-mean-square (RMS) difference (ignoring piston and tilt terms) between the actual and estimated wavefront aberration functions for this example was $\sigma = 0.37\lambda$. For comparison, the magnitude of the actual wavefront error shown in Fig. 2(a) was 1λ RMS. When the actual wavefront error is projected onto the first K Zernike polynomials (also including up to twelfth polynomial terms), the RMS difference is $\sigma_K = 0.33\lambda$. Theoretically, the modal reconstructor is capable of reconstructing accurately up to $K \approx M = 235$ Zernike terms (or up to twentieth-order polynomial terms). The fact that $\sigma_K (= 0.89\sigma)$ for $K = 91$ is a large fraction of the total error σ implies that the accuracy of the method is not limited by the modal reconstructor. Given the speckled appearance of the PSFs shown in Figs. 2(e)–2(i) and the *ad hoc* nature of the method, it is somewhat surprising that the estimated wavefront aberration function matches the actual $w(x,y)$ at all.

In summary, the basic steps of the method are

1. Choose a set of pupil plane coordinates (x_m, y_m) for $m \in \{1, 2, \ldots, M\}$;
2. Set $m = 1$;
3. Compute ideal ray coordinates $[r_{x,m}(z_n), r_{y,m}(z_n)]$ via Eq. (1) for each measurement plane $n \in \{1, 2, \ldots, N\}$;
4. Transform each PSF measurement $I_n(x,y)$ into a corresponding probability distribution $p_{m,n}(s_x, s_y)$ using Eq. (5);
5. Compute the joint probability distribution $p_{m,\text{joint}}(s_x, s_y)$ through Eq. (6);
6. Locate the maximum of $p_{m,\text{joint}}(s_x, s_y)$ to obtain wavefront slope data $(s_{x,m}, s_{y,m})$ at point (x_m, y_m) in the pupil plane;
7. Repeat steps 3–6 for $m \in \{2, 3, \ldots, M\}$;
8. Integrate the wavefront slope data to obtain the wavefront aberration function $w(x,y)$.

For results presented here, bilinear interpolation was used for step 4, the maxima in step 6 were determined by simply locating the largest sample value in each computed $p_{m,\text{joint}}(s_x, s_y)$ array, and the modal wavefront reconstructor described by Eqs. (8)–(10) was used for step 8. In some cases, a centroiding algorithm was used to locate the optic axis for each PSF measurement prior to performing the steps listed above. These details of the method, however, could vary in practice.

3. SIMULATION RESULTS

Monte Carlo simulations were used to explore the performance of the proposed method in various scenarios. Figure 3 shows results for the case when the actual wavefront aberrations $w(x, y)$ were generated by (i) synthesizing a random atmospheric phase screen [9], (ii) removing piston, tip, and tilt terms, and (iii) scaling this result to have a particular RMS value (plotted on the horizontal axis of Fig. 3). For each $w(x, y)$, five PSF measurements $I_n(x, y)$ were simulated via Eq. (4) and processed to yield an estimate of the wavefront aberration function. The RMS differences σ_K between the estimated and actual wavefront aberration functions are plotted on the vertical axis of Fig. 3. The plot indicates that the proposed wavefront sensing technique works reliably well ($\sigma_K \leq 0.1\lambda$) when the actual wavefront aberration is $\leq 0.4\lambda$ RMS. For larger wavefront errors, the results are variable and only moderately accurate at best.

Figure 4 shows Monte Carlo results for a similar scenario, except the atmospheric phase screens were projected onto the first $K = 91$ Zernike polynomials when generating $w(x, y)$ for the simulation. This projection operation eliminates high spatial-frequency aberrations, generally resulting in PSFs $I_n(x, y)$ that have a smaller speckle halo than those shown in Figs. 2(e)–2(i). Figure 4 indicates that the algorithm performs substantially better in this case. The error in the estimated wavefront is generally in the range of $\sigma_K = 0.1$–0.3λ for large amounts of actual wavefront error ($\sim 2\lambda$ RMS), and there is much less variability in performance compared to Fig. 3.

Additional Monte Carlo simulations were conducted to explore practical issues in implementing the proposed method. First, the impact of not knowing the location of the optic axis was considered. Figure 5 shows Monte Carlo results for the same scenario represented in Fig. 4, except a centroiding algorithm was used to determine the optic axis location for each PSF measurement prior to using the wavefront sensing algorithm. Comparing Figs. 4 and 5, the loss in performance due to lack of knowledge about the optic axis location is small and barely noticeable. When the actual wavefront error is $\sim 0.5\lambda$ RMS, the error in the estimated wavefront was typically in the range of $\sigma_K = 0.03$–0.09λ. Next, the impact of having fewer PSF measurements was considered. Figure 6 shows Monte Carlo results for the same scenario as Fig. 5, except only three PSF measurements were used instead of five. While the loss in performance with fewer measurements is noticeable, the algorithm still performed fairly well. For actual wavefront errors $\sim 0.5\lambda$ RMS, the error in the estimated wavefront is in the range $\sigma_K = 0.07$–0.2λ RMS.

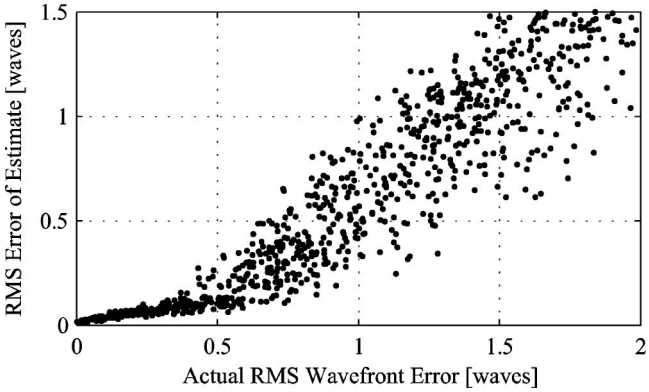

Fig. 3. Monte Carlo simulation results: plot of the RMS error in the estimated wavefront σ_K versus the RMS strength of the actual wavefront aberrations $w(x, y)$. These results correspond to the case where each $w(x, y)$ was an atmospheric phase screen with Kolmogorov statistics, and five PSF measurements were available.

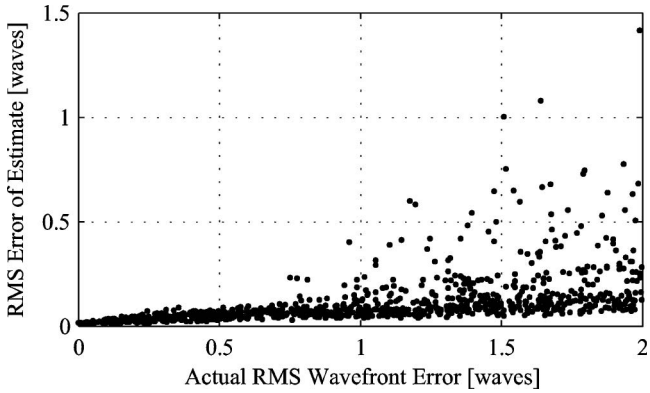

Fig. 4. Same as Fig. 3, but for the case where each $w(x, y)$ was an atmospheric phase screen with Kolmogorov statistics projected onto the first $K = 91$ Zernike polynomials, and five PSF measurements were available.

4. APPLICATION

Examples of segmented-aperture telescopes include the W. M. Keck Observatory [10], the Hobby–Eberly Telescope [11], the South African Large Telescope [12], and the James Webb Space Telescope (JWST) [13]. In each of these systems, hexagonal mirror segments are tiled together to form the primary of each telescope instead of using large monolithic mirrors. To achieve diffraction-limited performance, the mirror segments need to be accurately aligned to within a small fraction of a wavelength of light. Alignment of the Keck telescopes

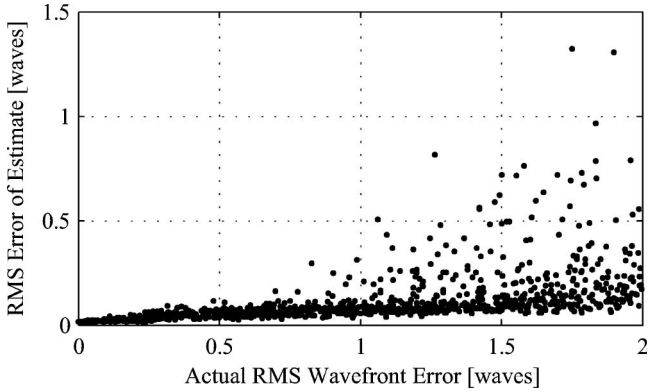

Fig. 5. Same as Fig. 3, but for the case where each $w(x, y)$ was an atmospheric phase screen with Kolmogorov statistics projected onto the first $K = 91$ Zernike polynomials, five PSF measurements were available, and a centroiding algorithm was used to determine the optic axis for each PSF measurement.

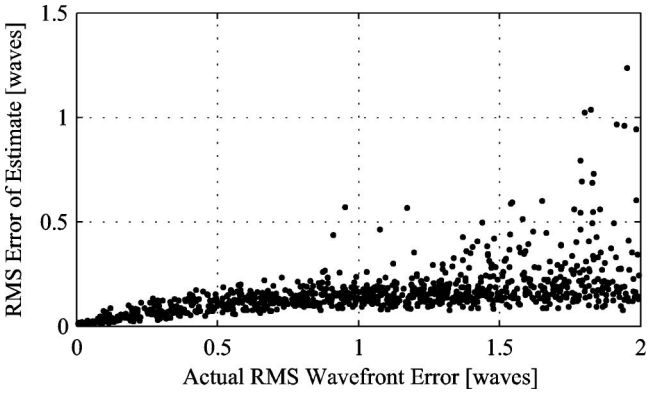

Fig. 6. Same as Fig. 3, but for the case where each $w(x,y)$ was an atmospheric phase screen with Kolmogorov statistics projected onto the first $K = 91$ Zernike polynomials, only three PSF measurements were available, and a centroiding algorithm was used to determine the optic axis for each PSF measurement.

is achieved using a modified Shack–Hartmann wavefront sensor with subapertures that span edges between adjacent mirror segments [14]. Coarse alignment is achieved using narrowband light ($\Delta\lambda/\lambda = 1\%$), in which case the sensor has a wide capture range ($\pm 67\lambda$) but coarse accuracy ($\sim 1.7\lambda$) [15]. As segments are brought into alignment, the optical bandwidth is increased to $\Delta\lambda/\lambda = 30\%$, improving accuracy ($\sim 0.07\lambda$) and reducing the capture range ($\pm 2.9\lambda$). Dispersed fringe sensing (DFS) [16] is planned for coarse alignment of JWST [17], while focus-diverse phase retrieval [18] will be used for fine alignment. DFS has a wide capture range ($\pm 37\lambda$) with good sensitivity (0.21–0.33λ RMS) but does not work well for small discontinuities ($<1\lambda$) between mirror segments [16]. Phase retrieval is more accurate, but has a limited capture range. In this section, I supplement phase retrieval with the method of obtaining wavefront slope information described in Section 2 to demonstrate a wavefront sensing approach that has the accuracy of phase retrieval with a significantly wider capture range. Specifically, the capture range was evaluated by numerical simulations for cases in which the surface figure of individual mirror segments is reasonably good, but there are varying amounts of segment piston and tilt misalignment.

The generalized pupil function $P(x,y)$ [19] of a segmented-aperture system can be expressed as

$$P(x,y) = \sum_{j=1}^{J} A_j(x,y) \exp\left[i2\pi \sum_{k=1}^{K} c_k \phi_k(x,y) + c_{j,k}\phi_{j,k}(x,y) \right], \tag{11}$$

where subscript $j \in \{1,2,\ldots,J\}$ indexes the various segments, $A_j(x,y)$ represents the amplitude transmittance of each segment, subscript $k \in \{1,2,\ldots,K\}$ indexes various wavefront error basis function terms $\phi_k(x,y)$ and $\phi_{j,k}(x,y)$, and coefficients c_k and $c_{j,k}$ are wavefront error expansion coefficients. $\phi_k(x,y)$ and $\phi_{j,k}(x,y)$ are based on Zernike polynomials [8] using the fringe ordering [20]. Table 1 indicates the aberrations represented by the first $K = 11$ basis functions. The $\phi_k(x,y)$ represent global aberrations across the entire pupil and are generated from the polynomial terms listed in Table 1 using Gram–Schmidt orthonormalization over the global pupil, i.e., the sum of all the $A_j(x,y)$'s. The $\phi_{j,k}(x,y)$ represent aberrations across the jth mirror segment and are orthonormalized over individual $A_j(x,y)$'s. Incoherent PSFs $I_n(x,y)$ of a segmented-aperture system can be computed with Eq. (4) by replacing $A(x',y')\exp[i2\pi w(x',y')/\lambda]$ with $P(x',y')$.

Table 1. Ordering and Form of Wavefront Basis Functions $\phi_{j,k}(x,y)$ and Statistics of Segment Wavefront Error Coefficients $c_{j,k}$ Used for the Capture Range Study

Index k	Aberration	Form [a]	Coefficient Max	Coefficient RMS
1	Piston	1	$0.079\kappa\lambda$	$0.046\kappa\lambda$
2	x Tilt	$\rho\cos\theta$	$0.18\kappa\lambda$	$0.10\kappa\lambda$
3	y Tilt	$\rho\sin\theta$	$0.18\kappa\lambda$	$0.10\kappa\lambda$
4	Power	ρ^2	0.063λ	0.036λ
5	Astigmatism	$\rho^2\cos 2\theta$	0.063λ	0.036λ
6	45° Astigmatism	$\rho^2\sin 2\theta$	0.063λ	0.036λ
7	x Coma	$\rho^3\cos\theta$	0.047λ	0.027λ
8	y Coma	$\rho^3\sin\theta$	0.047λ	0.027λ
9	Spherical	ρ^4	0.047λ	0.027λ
10	Trefoil	$\rho^3\cos 3\theta$	0.047λ	0.027λ
11	Trefoil	$\rho^3\sin 3\theta$	0.047λ	0.027λ

[a]The polynomial forms are given in polar coordinates (ρ,θ) instead of Cartesian coordinates (x,y).

A detailed process was used to generate wavefront realizations for the capture range study. Random global wavefront coefficients c_k for $k \in \{4,5,\ldots,11\}$ were generated and scaled to yield a combined RMS wavefront error of approximately 0.16λ across the global pupil. Segment figure errors were generated by selecting $c_{j,k}$ for $k \in \{4,5,\ldots,11\}$ values from a zero-mean uniform random distribution and scaled according to the maximum values listed in the fourth column of Table 1. For example, the maximum value of the segment power coefficients $c_{j,4}$ was 0.063λ. Thus, coefficients $c_{j,4}$ were drawn from the interval $[-0.063\lambda, 0.063\lambda]$. This process results in approximately 0.08λ RMS wavefront error associated with the segment figure. Segment piston and tilt coefficients were specified in terms of a wavefront error scaling parameter κ to simulate different levels of misalignment. The net effect of this process is to generate wavefront error realizations that correspond to a system with modest amounts of global ($\sim 0.16\lambda$) and segment ($\sim 0.08\lambda$) wavefront errors and scalable (depending on the value of κ) amounts of segment piston and tilt errors due to misalignment. For $\kappa > 1$, segment misalignment dominates the total wavefront error.

Figure 7(a) shows a particular wavefront error realization (in units of waves) for $\kappa = 5$ for a JWST-like system having $J = 18$ segments. For this example, the wavefront error associated with global coefficients c_k (for $k = 4$–11) is 0.13λ RMS, with segment coefficients $c_{j,k}$ (for $k = 4$–11) is 0.08λ RMS, with segment piston coefficients $c_{j,1}$ is 0.23λ RMS, and with segment tilt coefficients $c_{j,2}$ and $c_{j,3}$ is 1.1λ RMS. Adding these terms in quadrature results in a net wavefront error of 1.13λ RMS. This is a slightly pessimistic calculation of the actual RMS wavefront error, since the segment piston and tilt errors result in a small amount of global piston and tilt, which can be removed to obtain a net wavefront error of 0.94λ RMS. From here on, all RMS wavefront error values are reported after removing the global piston and tilt terms. Additionally, any segment piston errors outside of the interval $[-0.5\lambda, 0.5\lambda]$ are adjusted by an integer number of wavelengths to fall within this interval before calculating the RMS wavefront error. This was done because the simulations were performed

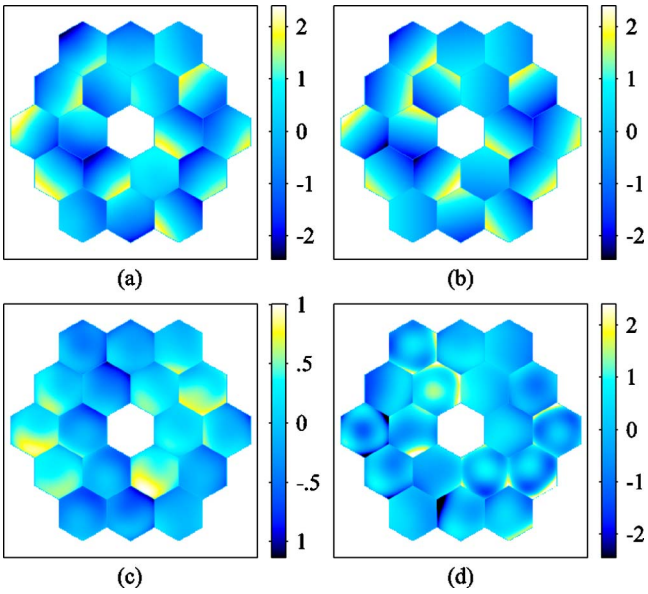

Fig. 7. (Color online) Segmented-aperture simulation example: (a) actual wavefront aberration function for a system with $J = 18$ segments (segment tilt errors dominate the wavefront error), (b) corresponding initial wavefront estimate of segment tilt errors obtained using the proposed wavefront slope estimation technique, (c) wavefront difference between (a) and (b), and (d) corresponding wavefront estimate obtained using the blind phase retrieval approach. All wavefronts are displayed in units of wavelength.

for monochromatic light, and, as such, there is an integer wavelength ambiguity in the estimated segment piston coefficients $c_{j,1}$.

For each wavefront realization, a set of five through-focus PSF measurements $I_n(x,y)$ were simulated with defocus amounts z_n equivalent to $\{8,6,0,-6,-8\}$ wavelengths peak-to-valley of quadratic defocus across the global pupil. The method of obtaining wavefront slope data described in Section 2 can be directly applied to segmented-aperture systems. The resulting slopes, however, cannot be integrated to yield $w(x,y)$ because of discontinuities between segments. Instead, initial estimates of the segment tilt errors were formed by (i) estimating wavefront slopes $(s_{x,j}, s_{y,j})$ for a set of $j \in \{1,2,\ldots,J\}$ rays that pass through the centers of each mirror segment and (ii) converting slope values $(s_{x,j}, s_{y,j})$ to segment tilt coefficients $c_{2,j}$ and $c_{3,j}$. The simulations assumed that accurate knowledge of the optic axis location was not available, and each PSF was recentered about its centroid prior to using the wavefront slope estimation method. Figure 7(b) shows the initial wavefront estimate corresponding to the example wavefront of Fig. 7(a). The difference between these wavefronts is shown in Fig. 7(c). Note that the segment tilt values for the initial estimate are reasonably accurate, but the remaining aberration terms have yet to be estimated.

Starting with an initial wavefront estimate, an iterative phase retrieval algorithm was used to obtain a final wavefront estimate, including all coefficients c_j and $c_{j,k}$. For this, a nonlinear optimization routine numerically searched for coefficients that produce model PSFs that best agree with the measured PSFs $I_n(x,y)$ in a mean square error sense. Specifically, the algorithm seeks the minimum of an objective function Φ given by Eq. (9) of Ref. [21]. Because of the nonlinear relationship between the wavefront coefficients and the PSF intensities, Φ is nonconvex, and the optimization routine can converge to a local minimum instead of the desired true solution. The likelihood of encountering a local minimum generally increases the further one starts from the true solution. This trend gives rise to the finite capture range of phase retrieval. In other words, the phase retrieval algorithm needs to start *sufficiently close* to the true solution in order to converge properly. Starting with the initial wavefront estimate shown in Fig. 7(b), phase retrieval was used to obtain a final wavefront estimate that agreed with the actual wavefront error shown in Fig. 7(a) to a level of 0.007λ RMS. For comparison, the result of a blind phase retrieval approach (starting with an initial estimate of zero wavefront error) is shown in Fig. 7(d). Note that this result bears little resemblance to the actual wavefront error because this example fell outside the capture range of blind phase retrieval. In simulations such as this, where the true wavefront error and noise statistics are known, the final value of Φ can be compared to its expected value for the true solution (based on noise statistics) to determine whether the phase retrieval algorithm converged to the true solution or not. For this example, the final value of Φ was within the expected range for the proposed (slope/tilt estimation plus phase retrieval) approach, but was larger than expected for the blind phase retrieval approach.

A series of simulations were performed using different values of the wavefront error scaling parameter κ to explore the capture range of both approaches. Figure 8 is a plot of the RMS error of estimated wavefronts versus the actual RMS wavefront errors for the proposed approach. For a number of trials (indicated by crosses), the algorithm failed to converge to a reasonable solution, as determined by the final value of Φ in comparison to the expected value. For a majority of the trials (indicated by circles), however, the algorithm did converge to a good wavefront estimate, with accuracies ranging from 0.0002λ to 0.02λ RMS. Many of the unconverged trials were examined to determine the cause of failure. In each case, there was a large error in one or more of the segment tilt coefficients $c_{j,2}$ or $c_{j,3}$ for the initial wavefront estimate. In these cases, the segment tilt estimation technique of Ref. [22] could be used to identify and possibly correct the one or two culprit segment tilts. With just one exception, the blind phase retrieval approach only achieved good results for $\kappa \le 1$, corresponding to actual wavefront errors $<0.3\lambda$ RMS. Figure 9 shows the fraction of trials that converged at each value of κ for

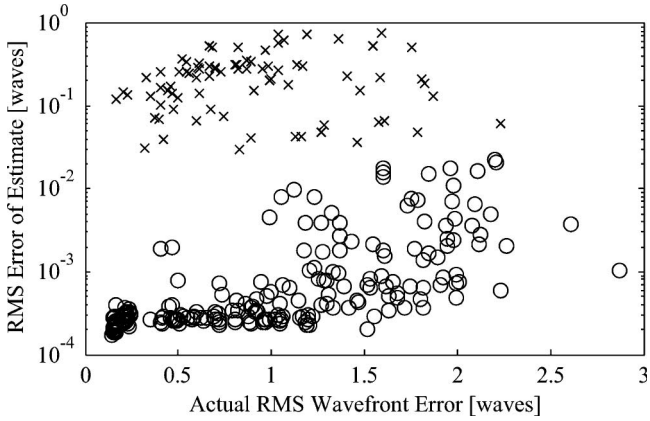

Fig. 8. Results of the capture range study using the proposed method: RMS error of estimated wavefronts versus the actual RMS wavefront errors, where the circles and crosses represent converged and unconverged trials, respectively.

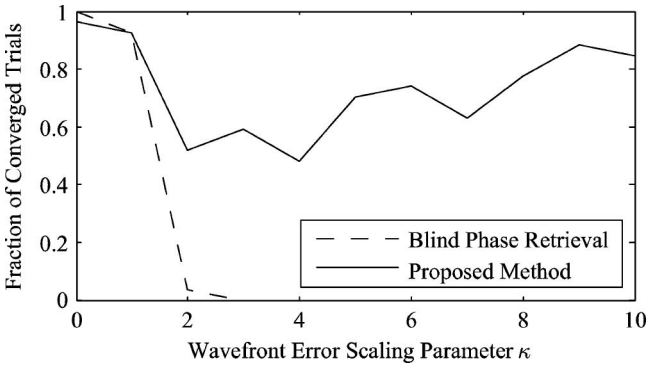

Fig. 9. Results of the capture range study: fraction of simulation trials that converged to the desired true solution versus the wavefront error scaling parameter κ. The dashed curve represents the blind phase retrieval results, while the solid curve represents results using the proposed method.

both the proposed and the blind phase retrieval approaches. The proposed method clearly has a wider capture range.

5. SUMMARY

A method was presented for obtaining pupil plane wavefront slope data from through-focus PSF measurements. This slope data can them be used to reconstruct the wavefront aberration function of the imaging system. The proposed technique differs from phase retrieval in that it is based on geometric optics instead of Fourier optics and differs from Shack–Hartmann and curvature wavefront sensing in that it uses measurements in the caustic region near focus instead of relying on data collected far from focus. Simulations suggest that the method works best for smooth wavefront aberrations. For smooth aberrations of the order of 1λ RMS, the error in the reconstructed wavefront in simulation was typically $<0.1\lambda$ RMS when using five PSF measurements. The loss in performance associated with not knowing the location of the optic axis was minimal. A small but noticeable loss in performance was observed when using just three PSF measurements instead of five. Other issues that may affect performance have yet to be investigated.

Image-based wavefront sensing is one method of phasing segmented-aperture systems. Phase retrieval alone, however, has a limited capture range. The proposed method here uses a wavefront slope estimation technique to generate an initial wavefront estimate for phase retrieval. Numerical simulations examined the case in which there was only a modest amount of wavefront error across individual segments but increasingly large segment tilt errors due to misalignment. In this scenario, the proposed algorithm was shown to have a much larger capture range than just starting the phase retrieval algorithm with a wavefront estimate of zero.

REFERENCES AND NOTES

1. D. L. Fried, "Least-square fitting a wave-front distortion estimate to an array of phase-difference measurements," J. Opt. Soc. Am. **67**, 370–375 (1977).
2. R. H. Hudgin, "Wave-front reconstruction for compensated imaging," J. Opt. Soc. Am. **67**, 375–378 (1977).
3. W. H. Southwell, "Wave-front estimation from wave-front slope measurements," J. Opt. Soc. Am. **70**, 998–1006 (1980).
4. F. Roddier, "Curvature sensing and compensation: a new concept in adaptive optics," Appl. Opt. **27**, 1223–1225 (1988).
5. A. J. Devaney, R. A. Gonsalves, and R. Chidlaw, "Application of phase retrieval techniques to adaptive imaging systems," J. Opt. Soc. Am. **67**, 1422 (abstract) (1977), http://www.opticsinfobase.org/josa/abstract.cfm?URI=josa-67-10-1361.
6. J. N. Cederquist, J. R. Fienup, C. C. Wackerman, S. R. Robinson, and D. Kryskowski, "Wave-front phase estimation from Fourier intensity measurements," J. Opt. Soc. Am. A **6**, 1020–1026 (1989).
7. E. Acosta, S. Bará, M. A. Rama, and S. Rios, "Determination of phase mode components in terms of local wave-front slopes: an analytical approach," Opt. Lett. **20**, 1083–1085 (1995).
8. R. J. Noll, "Zernike polynomials and atmospheric turbulence," J. Opt. Soc. Am. **66**, 207–211 (1976).
9. R. G. Lane, A. Glindemann, and J. C. Dainty, "Simulation of a Kolmorgorov phase screen," Waves Random Media **2**, 209–224 (1992).
10. URL:http://keckobservatory.org/.
11. URL:http://rhea.as.utexas.edu/.
12. URL:http://www.salt.ac.za/.
13. URL:http://www.jwst.nasa.gov/.
14. G. Chanan, M. Troy, F. Dekens, S. Michaels, J. Nelson, T. Mast, and D. Kirkman, "Phasing the mirror segments of the Keck telescopes: the broadband phasing algorithm," Appl. Opt. **37**, 140–155 (1998).
15. Performance is given here in terms of the wavefront error seen in reflection, which is a factor of 2 larger than the specifications reported in [14], which are in terms of surface error.
16. F. Shi, G. Chanan, C. Ohara, M. Troy, and D. C. Redding, "Experimental verification of dispersed fringe sensing as a segment phasing technique using the Keck telescope," Appl. Opt. **43** 4474–4481 (2004).
17. D. S. Acton, T. Towell, J. Schwenker, D. Shields, E. Sabatke, A. R. Contos, K. Hansen, F. Shi, B. Dean, and S. Smith, "End-to-end commissioning demonstration of the James Webb Space Telescope," Proc. SPIE **6687**, 668706 (2007).
18. L. D. Feinberg, B. H. Dean, D. L. Aronstein, C. W. Bowers, W. Hayden, R. G. Lyon, R. Shiri, J. S. Smith, D. S. Acton, L. Carey, A. Contos, E. Sabatke, J. Schwenker, D. Shields, T. Towell, F. Shi, and L. Meza, "TRL-6 for JWST wavefront sensing and control," Proc. SPIE **6687**, 668708 (2007).
19. J. Goodman, *Introduction to Fourier Optics*, 3rd ed. (Roberts, 2005).
20. "ZEMAX Optical Design Program, User's Guide" (ZEMAX Development Corporation, 2006).
21. S. T. Thurman and J. R. Fienup, "Phase retrieval with signal bias," J. Opt. Soc. Am. A **26**, 1008–1014 (2009).
22. T. Zielinski and J. R. Fienup, "Extending wavefront sensing capture range for segmented systems through tip and tilt estimation," in *Frontiers in Optics*, Tech. Dig. (CD) (Optical Society of America, 2006), paper FMF4.