

Phase diversity for three-dimensional imaging

Peter Kner

College of Engineering, University of Georgia, Athens, Georgia 30602, USA (kner@engr.uga.edu)

Received June 12, 2013; revised August 5, 2013; accepted August 18, 2013;
posted August 19, 2013 (Doc. ID 192104); published September 12, 2013

Phase diversity (PD) is a powerful technique for estimating wavefront aberrations from two-dimensional images of extended scenes. PD can work with extended incoherent images and, in an adaptive optics system, does not need extra hardware in addition to the deformable mirror. For these reasons, PD should be well suited to aberration measurement in microscopy applications. But, in biological widefield microscopy, the objects being imaged are frequently three-dimensional, and the images contain out-of-focus light. In this paper, we introduce multiplane PD and show that PD can be extended to widefield imaging of three-dimensional objects. This should be particularly useful in the field of biological fluorescence microscopy where the objects are very light sensitive and the aberrations cannot easily be determined. © 2013 Optical Society of America

OCIS codes: (180.6900) Three-dimensional microscopy; (180.2520) Fluorescence microscopy; (100.6890) Three-dimensional image processing; (100.5070) Phase retrieval.
<http://dx.doi.org/10.1364/JOSAA.30.001980>

1. INTRODUCTION

In adaptive optics (AO) systems, the optical wavefront is measured and then corrected in order to improve the performance of an optical instrument, resulting in images with better resolution and higher signal to noise ratios [1]. The use of AO is becoming widespread in astronomy [1,2], ophthalmoscopy [3], and, more recently, microscopy [4,5].

In an AO system, there are several choices for the wavefront sensing technique. The most common type of wavefront sensor for astronomical AO is the Shack–Hartmann wavefront sensor [1]. The Shack–Hartmann wavefront sensor is fast and accurate, but it requires complicated extra hardware and a bright point source. Shack–Hartmann sensors also do not provide any depth discrimination because of the low numerical aperture of the lenslets. For these reasons, Shack–Hartmann sensors have not been used widely in microscopy, where most of the aberrations are due to imaging through three-dimensional objects that are thicker than the depth of focus. Recently, the use of a Shack–Hartmann wavefront sensor with laser scanning confocal microscopy has been demonstrated [6,7]; in this case the depth-discrimination is provided by the confocal excitation spot.

A popular approach in microscopy applications has been to iteratively adjust the wavefront until some metric has been optimized. The metric is typically the signal intensity but can be any value related to the image quality such as the high-frequency content of the image. This is, in effect, a form of wavefront sensing because the optimized wavefront is a measure of the aberrations. Any algorithm for finding the maximum point of a multidimensional surface can be applied to the problem [8]. For example, genetic algorithms have been used to optimize the wavefront correction [9,10]. These approaches have the advantage that they are easy to implement but can require many exposures to optimize the metric. Minimizing the number of iterations in the optimization algorithm can be critically important to minimize the time and photobleaching due to the AO process. Algorithms have been proposed

which optimize a given number of orthogonal wavefront modes (i.e., Zernike modes) with the minimal number of measurements [11,12].

Phase retrieval is another approach to determining the wavefront without a sensor [13,14]. In phase retrieval, the wavefront is determined iteratively. Starting from an initial guess for the wavefront, the field is transformed between the pupil and focal planes. In the focal plane, the field must conform to the measured amplitude and in the pupil plane the field must conform to the pupil boundary. A coherent or quasi-coherent image is required for phase retrieval. Phase retrieval has been used in astronomy to determine the aberrations in the Hubble space telescope [15] and has been demonstrated in fluorescence microscopy using images of a subdiffraction fluorescent bead [16,17]. With phase retrieval, the resolution of the image is determined by the extent and the SNR of the images, so a high resolution wavefront can be determined from only a few images. The resolution in the back pupil plane is determined by the lateral extent of the image. Thus a $10\text{ }\mu\text{m} \times 10\text{ }\mu\text{m}$ image will have a resolution of $(\lambda/10\text{ NA})$ in units of NA/λ . For $\lambda = 0.5\text{ }\mu\text{m}$ and $\text{NA} = 1.2$, there would be ≈ 1750 pixels comprising the pupil.

As with the Shack–Hartmann sensor, phase retrieval requires, in practice, a point source that is not always available in microscopy. Phase diversity (PD) is a technique for inferring the wavefront aberrations and an improved estimate of the image, from multiple incoherent images [18,19]. Each image is taken with a different known phase function introduced into the back pupil plane of the imaging system. PD works with incoherent images of extended scenes so it is well suited to fluorescence microscopy in that sense, but PD so far has been developed only for two-dimensional imaging of two-dimensional objects. In widefield fluorescence microscopy, images are frequently of three-dimensional objects (objects whose axial extent is significantly greater than the depth of focus) in which out-of-focus light is present. The out-of-focus light affects the ability of PD to determine the wavefront and

the true object. PD can potentially determine the wavefront with high resolution from only two images. To characterize the wavefront with the first eight Zernike modes would require nine measurements using an efficient optimization scheme, whereas two PD measurements on an 128×128 pixel image would yield a wavefront described by over 1000 independent values although the signal to noise would determine the actual number of modes that could be measured accurately.

Multiframe myopic deconvolution is related to PD [20]. In this technique, one or a set of images is used to determine both the optimum object and point spread function (PSF) by minimizing the cost function. But in PD, the images are intentionally aberrated with known phase functions to determine the optimum object and PSF. This leads to a compact and efficient cost function whereas in myopic deconvolution the error function is constructed in a more ad-hoc fashion.

In this paper, we discuss various approaches to applying PD to three-dimensional imaging. We develop multiplane PD, an extension of PD for widefield images of three-dimensional objects, and show that, for Gaussian noise, an expression that eliminates the image degrees of freedom can be developed as with two-dimensional PD. Multiplane PD results in an improved estimate of the object in the focal plane and improved performance of the cost function for estimating the wavefront aberrations. Three-dimensional PD has been discussed in the context of retinal imaging, but in that work only one image stack was taken so no additional PD images were taken [21]. The use of phase functions other than defocus is discussed by Campbell *et al.* [22]. First we will review two-dimensional PD theory. Then we will discuss applying PD to three-dimensional widefield imaging and introduce multiplane PD. In the results section, we demonstrate the different approaches on simulated data. Finally, we discuss the performance of PD and future research plans.

2. TWO-DIMENSIONAL PHASE DIVERSITY

In PD, multiple images are taken with different known phase functions introduced into the pupil function of the system as described by Paxman *et al.* [19]. The pupil function for the k th diversity image is given by

$$H_k(u) = |P(u)| \exp(i\phi(u) + i\theta_k(u)). \quad (1)$$

$P(u)$ is the pupil function, which defines the size of the pupil and any apodization. $\theta_k(u)$ are the known phase functions, and $\phi(u)$ is the unknown phase aberration that is to be determined. u denotes the coordinates of the pupil plane; x and u form a Fourier transform pair. For a microscope, $P(u) = 0$ for $u > \text{NA}/\lambda$. The coherent PSF is then given by

$$h_k(x) = \mathcal{F}^{-1}\{H_k(u)\}, \quad (2)$$

where x denotes the coordinates in the image plane, and $\mathcal{F}\{\}$ denotes the Fourier transform operation. The incoherent PSF is then

$$s_k(x) = |h_k(x)|^2. \quad (3)$$

The set of PSFs, $s_k(x)$, and the set of images, $d_k(x)$, are then used to minimize the cost function, $L\{f(x), \phi(u)\}$. In the case of Gaussian noise, the cost function is the difference between

the measured images and the images expected from the object and the PSF:

$$L(f, \bar{\alpha}) = \sum_{k=1}^K \sum_x [d_k(x) - f(x) \otimes s_k(x)]^2. \quad (4)$$

$f(x)$ is the object, and $\bar{\alpha}$ denotes the parameterization of the phase degrees of freedom. If the phase is decomposed into basis functions $\phi_j(u)$, then

$$\phi(u) = \sum_{j=1}^J \alpha_j \phi_j(u) \quad (5)$$

and $\bar{\alpha} = \{\alpha_j\}$. The object, $f(x)$, can be eliminated from the expression for $L(f, \bar{\alpha})$ by requiring that $L(f, \bar{\alpha})$ be at an extremum with respect to $f(x)$. This is done in the Fourier domain where the cost function is

$$L(F, \bar{\alpha}) = \sum_{k=1}^K \sum_u |D_k(u) - F(u)S_k(u)|^2 \quad (6)$$

and we demand that

$$\frac{\partial L}{\partial F_{\text{re}}} = 0 \quad \text{and} \quad \frac{\partial L}{\partial F_{\text{im}}} = 0. \quad (7)$$

For $K = 2$, this leads to the well-known result of Gonsalves [18] for the estimate of the object Fourier transform, $F(u)$, and the cost function, $L(\bar{\alpha})$:

$$F(u) = \frac{D_1(u)S_1^*(u; \bar{\alpha}) + D_2(u)S_2^*(u; \bar{\alpha})}{|S_1(u; \bar{\alpha})|^2 + |S_2(u; \bar{\alpha})|^2} \quad (8)$$

$$L(\bar{\alpha}) = \sum_u \frac{|D_1(u)S_2(u; \bar{\alpha}) - D_2(u)S_1(u; \bar{\alpha})|^2}{|S_1(u; \bar{\alpha})|^2 + |S_2(u; \bar{\alpha})|^2}. \quad (9)$$

By minimizing $L(\bar{\alpha})$ over $\{\alpha_j\}$, the result is the best estimates for both the object, $f(x) = \mathcal{F}^{-1}\{F(u)\}$, and the aberrations parametrized by $\bar{\alpha}$. Eliminating $f(x)$ from the cost function is a substantial reduction in the degrees of freedom. While the aberrations may be adequately described in many cases by 8–33 Zernike modes, an $N \times N$ pixel image would contribute N^2 additional search degrees of freedom. Paxman *et al.* extended this result by considering the case of more than two diversity images. They also considered the case of Poisson noise in which case the object degrees of freedom cannot easily be eliminated. In many systems, PD is implemented by using defocus as the PD aberration, which can be accomplished by moving the detector.

3. WIDEFIELD IMAGING MODEL

A. Extension to Three Dimensions

In widefield fluorescence microscopy of three-dimensional objects, the image can be considered as the convolution with the three-dimensional PSF

$$d(x, y, z) = f(x, y, z) \otimes s(x, y, z). \quad (10)$$

This is true as long as the system is shift invariant, as for the two-dimensional case. For widefield imaging, $s(x, y, z)$ is not limited in its axial extent. The spot becomes larger, but the integrated intensity does not change. Thus, each lateral plane contains out-of-focus images from the rest of the object. The incoherent PSF is given by

$$s_k(x, y, z) = |h_k(x, y, z)|^2, \quad (11)$$

$$h_k(x, y, z) = \mathcal{F}^{-1}\{H_k(u) \exp(iz\gamma(u))\}, \quad (12)$$

$$\gamma(u) = \frac{2\pi n}{\lambda} \sqrt{1 - \left(\frac{\lambda u}{n}\right)^2}. \quad (13)$$

$\gamma(u)$ is the defocus function. The optical transfer function (OTF) is the Fourier transform of the PSF. $S(u) = \mathcal{F}\{s(x, y, z)\}$. With this approach, the two-dimensional PD equations can be directly extended to the three-dimensional case by extending the Fourier transforms to three-dimensions, using Eqs. (8) and (9) but with F and S now referring to the 3D object and OTF. In three-dimensional PD, defocus cannot be used as the phase diverse phase function because it provides no new information. In three-dimensional imaging, the PSF is a function of z and defocus just corresponds to a shift in the imaging plane:

$$\begin{aligned} s_k(\vec{x}) &= s(\vec{x} - d\hat{z}) \\ d_k(\vec{x}) &= s_k(\vec{x}) \otimes f(\vec{x}) \\ &= \mathcal{F}^{-1}\{e^{-j2\pi u d} S(\vec{u}) F(\vec{u})\} \\ &= d(\vec{x} - d\hat{z}). \end{aligned} \quad (14)$$

Because the shift translates directly to the image, the phase term, $e^{-j2\pi u d}$, will appear in both terms of $L(\vec{\alpha})$, and the cost function will be $L(\vec{\alpha}) \approx 0$. Thus, to use PD with a three-dimensional image, something other than defocus must be used for PD. In an AO system, the deformable mirror can be used to introduce aberrations; thus, it is straightforward to apply PD functions other than defocus as has been suggested in [22].

There are two problems with using three-dimensional images for PD. One problem is simply that the computational cost is very high. Computing the cost function requires computing the OTFs, $S_1(u)$ and $S_2(u)$. Computing each OTF requires the calculation of N_z 2D FFTs of size $N_x \times N_y$ and one 3D FFT of size $N_z N_x N_y$. Computing the 2D OTF requires only two 2D FFTs of size $N_x N_y$. The other, more serious difficulty is that the system is not shift invariant. Sample-induced aberrations will change throughout the sample so that if the image of the entire 3D object is used, it will not be clear what aberrations are being determined [23].

B. Multiplane Phase Diversity

Instead of using a full three-dimensional imaging model, we consider a model for the two-dimensional image that takes into account the out-of-focus planes:

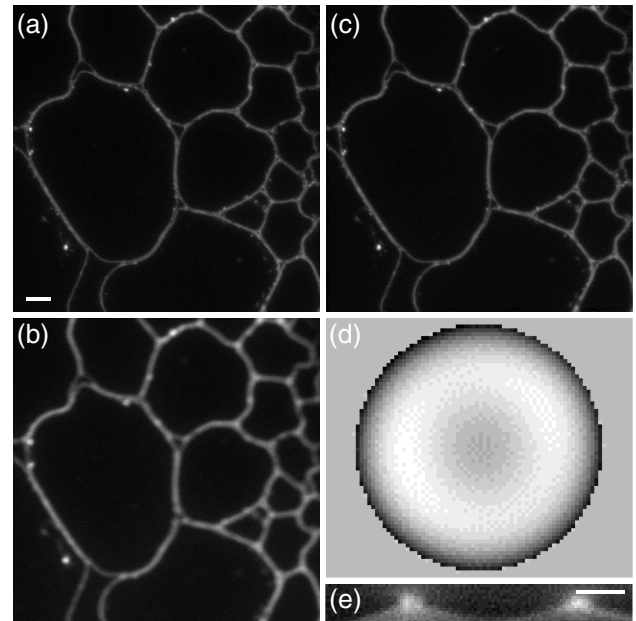


Fig. 1. Example of two-dimensional PD on a fluorescence image. (a) In focus image of GFP-labelled cell wall in *Arabidopsis thaliana*. (b) Image taken 1 μm out of focus. (c) Object estimated from PD calculation on images (a) and (b). (d) Wavefront estimated from PD calculation. (e) Axial cross section of the sample showing spherical aberration in agreement with the PD calculation and the wavefront in (d). Scale bar is 5 μm . These images were taken with a 40 \times air objective.

$$d_k(x, y) = \sum_{i=1}^P f(x, y, z_i) \otimes s_k(x, y; z_i - z_d). \quad (15)$$

The sum is over the P z -planes around the focal plane of the object, z_d , and $s_k(x, y; z)$ is the 2D PSF with the k th phase function with a defocus of z . Planes further away from the imaging plane will contribute less to the structure of the object so the

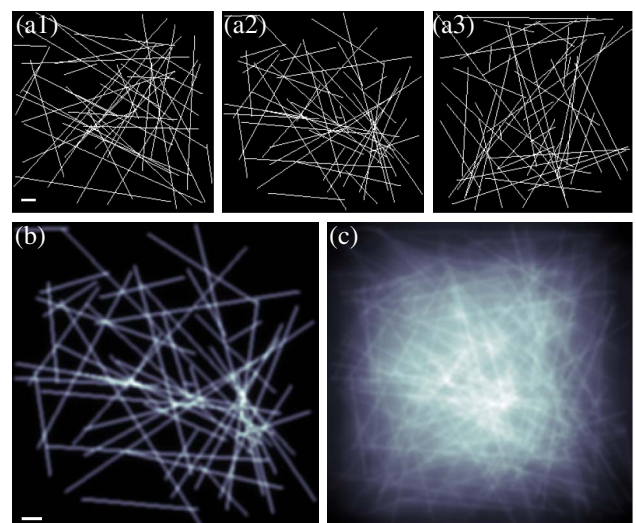


Fig. 2. Simulation using the imaging model from Eq. (15). (a) Three planes from simulated object composed of lateral lines. The planes are 0.25 μm apart, and the total image is 32 planes. (b) Image of the central plane. (c) Image of the central plane including all out-of-focus planes using Eq. (15). The out-of-focus OTF was generated using the Stokseth approximation [24]. Scale bar is 1 μm .

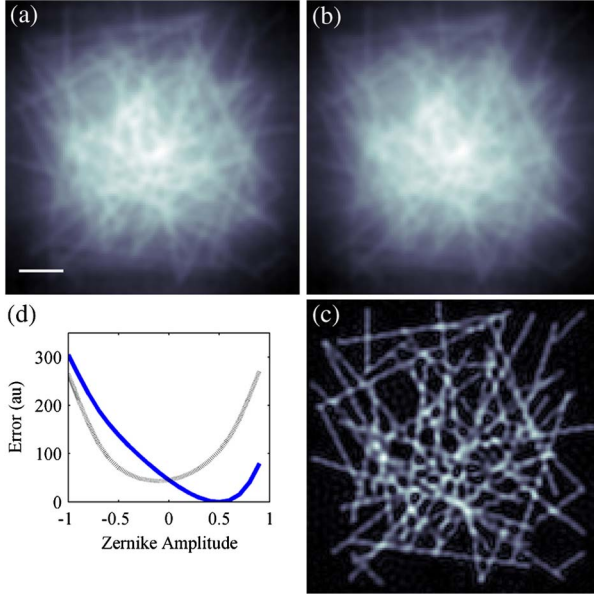


Fig. 3. Three-dimensional PD. (a) and (b) PD images of the simulated sample. These are the central planes from a $32 \times 128 \times 128$ stack with $dz = 0.5 \mu\text{m}$ and $dx = 50 \text{ nm}$. The aberration is $\phi(u) = 0.5Z_3^1(u)$. (a) Image with $\theta_1(u) = 0$. (b) Image with PD aberration $\theta_2(u) = 0.5Z_2^2(u)$. (c) Central plane of the recovered image from the PD calculation with $\mu = 0.001$. (d) The error metric, $L(\bar{\alpha})$, as the amplitude of aberration is varied. Blue solid curve: aberration is Z_3^1 and goes to 0 at the value of the actual aberration. Black dashed curve: varying another aberration, Z_3^{-1} , does not bring the error metric, L , to 0. Scale bar is $1 \mu\text{m}$.

level of approximation can be controlled by the number of image planes included in the sum as well as the spacing of the image planes. We can then insert this imaging model into the cost function, Eq. (11) from Paxman *et al.* [19]. The cost function then becomes

$$L(f, \bar{\alpha}) = \sum_{k=1}^K \sum_{x,y} \left(d_k(x, y) - \sum_{i=1}^P f(x, y, z_i) \otimes s_k(x, y, z_i) \right)^2, \quad (16)$$

$$L(F, \bar{\alpha}) = \sum_{k=1}^K \sum_u \left| D_k - \sum_{i=1}^P F_z(u) S_k(u; z_i) \right|^2, \quad (17)$$

where $F_z(u)$ and $S_k(u)$ are two-dimensional Fourier transforms over the x - y plane. As with 2D PD, the object degrees of freedom are eliminated by setting the derivatives of L with respect to F_z to 0. The result is a $P \times P$ matrix equation for $[F_z]$, the object Fourier transform for each z -plane included in the approximation. The result is

$$M_{z,z'} [F_z] = \begin{bmatrix} \dots \\ \sum_k D_k S_{k,z}^* \\ \dots \end{bmatrix}, \quad (18)$$

where

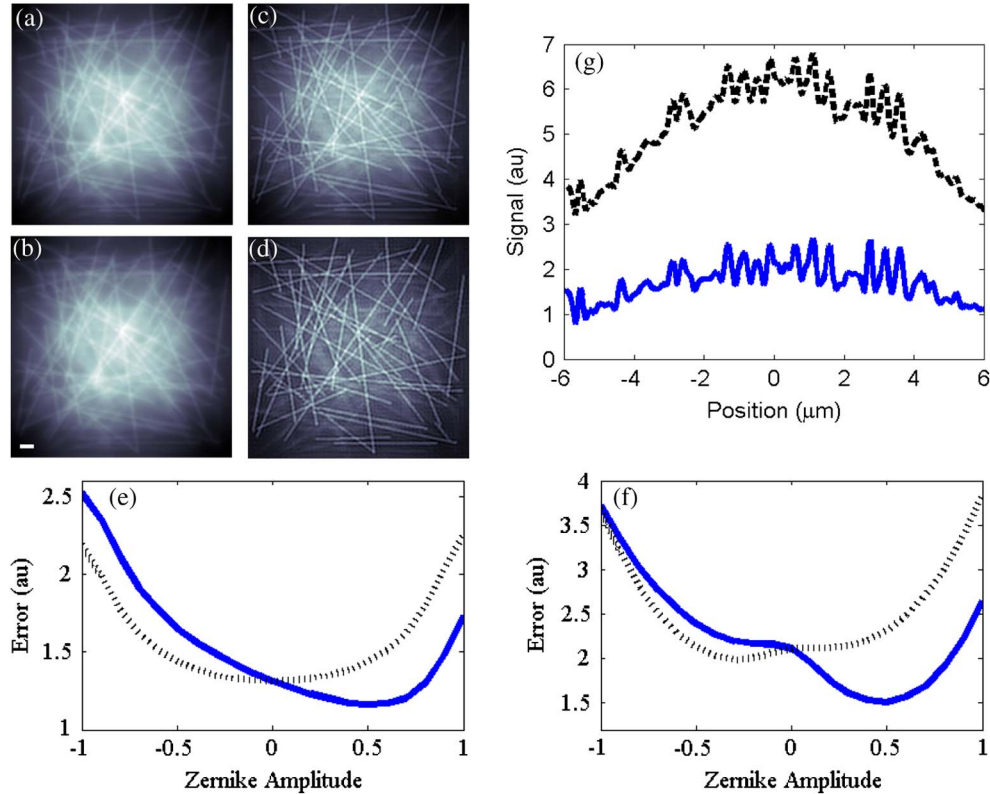


Fig. 4. Comparison of two-dimensional PD with three-plane PD on one plane of a three-dimensional stack. (a) PD image with $\theta_1(u) = 0$. (b) PD image with $\theta_2(u) = 0.5Z_2^2(u)$. The wavefront aberration is $\phi(u) = 0.5Z_3^1(u)$. (c) Optimal PD image from two-dimensional PD equation. (d) Optimal PD image from three-plane PD. (e) Error metric, $L(\bar{\alpha})$, versus Zernike amplitude for two-dimensional PD. Black dashed curve: error metric versus Z_3^{-1} blue solid curve: error metric versus Z_3^1 (f) error metric for three-plane PD. (g) Cross section across images (c) (black, dashed) and (d) (blue, solid). The three-plane PD results in lower background. $\mu = 0.1$. Simulated object has 32 planes spaced $0.5 \mu\text{m}$ apart. Scale bar is $1 \mu\text{m}$.

$$M_{z,z'} = \sum_k M_{z,z'}^k = \begin{bmatrix} \dots & \dots & \dots \\ \dots & \sum_k S_{k,z}^* S_{k,z'} & \dots \\ \dots & \dots & \dots \end{bmatrix}. \quad (19)$$

The solution of Eq. (18) for $[F_z]$ can then be inserted into Eq. (17), to get an expression for $L(F, \bar{\alpha})$ in which the object degrees of freedom are eliminated. It is easy to verify that for the case $K = 2, P = 1$, Eqs. (17) and (18) reduce to Paxman's result.

To invert Eq. (18), it is necessary that $K \geq P$ because each matrix $M_{z,z'}^k$ is singular. Physically this is a statement that we cannot estimate P images at different focal planes without at least P PD images; information cannot be created. To account for measurement noise and numerical artifacts in the results, we further introduce a Wiener parameter, μ . Then Eqs. (8) and (18) become

$$F(u) = \frac{D_1(u)S_1^*(u) + D_2(u)S_2^*(u)}{|S_1(u)|^2 + |S_2(u)|^2 + \mu^2} \quad (20)$$

and

$$\left(\begin{bmatrix} \dots & \dots & \dots \\ \dots & \sum_k S_{k,z}^* S_{k,z'} & \dots \\ \dots & \dots & \dots \end{bmatrix} + \mu^2 \mathcal{I} \right) \begin{bmatrix} \dots \\ F_z \\ \dots \end{bmatrix} = \begin{bmatrix} \dots \\ \sum_k D_k S_{k,z}^* \\ \dots \end{bmatrix}, \quad (21)$$

respectively, where \mathcal{I} is the identity matrix. Introducing the Wiener parameter results in an equation that is invertible even for $K < P$, although information would not be conserved.

4. RESULTS

While two-dimensional PD is not a new result, there are not many published results of 2D PD in biological microscopy, so we show an example of two-dimensional PD on a biological

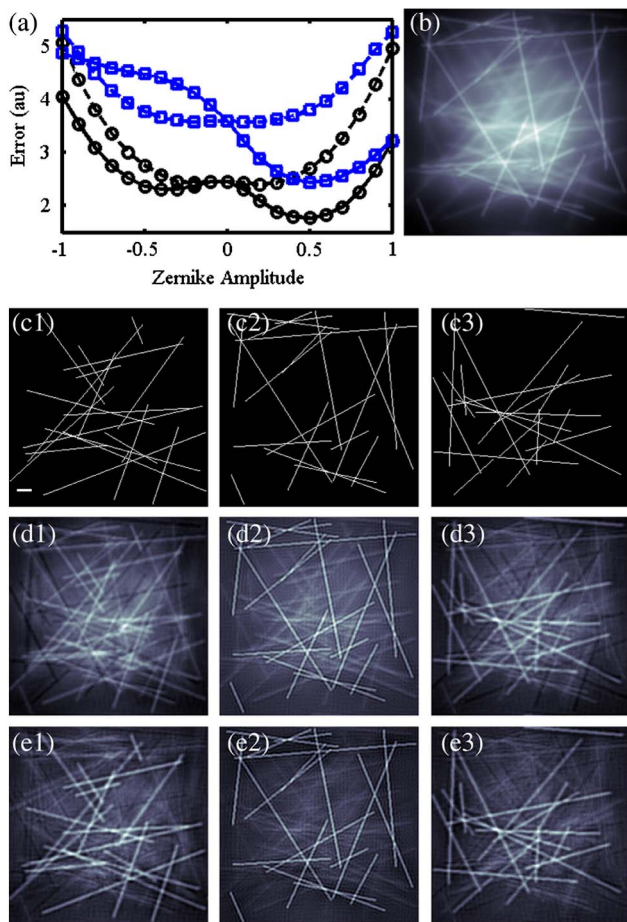


Fig. 5. Comparison of three-plane and five-plane PD. The aberration is $\phi(u) = 0.5Z_3^1(u)$. (a) Plot of the error function versus Zernike amplitude. The black curves with circles are the three-plane error curves and the blue curves with squares are the five-plane error curves. The solid curves show the error versus the amplitude of $Z_3^1(u)$. The dashed curves show the error versus the amplitude of $Z_3^{-1}(u)$. (b) PD image with no induced aberration. (c) Simulated object. (d) Result of three-plane PD calculation. (e) Result of five-plane PD calculation. (1) First column is $0.5 \mu\text{m}$ below focus. (2) Second column is focal plane. (3) Third column $0.5 \mu\text{m}$ above focus. $\mu = 0.1$. Simulated object has 32 planes spaced $0.5 \mu\text{m}$ apart. Scale bar is $1 \mu\text{m}$.

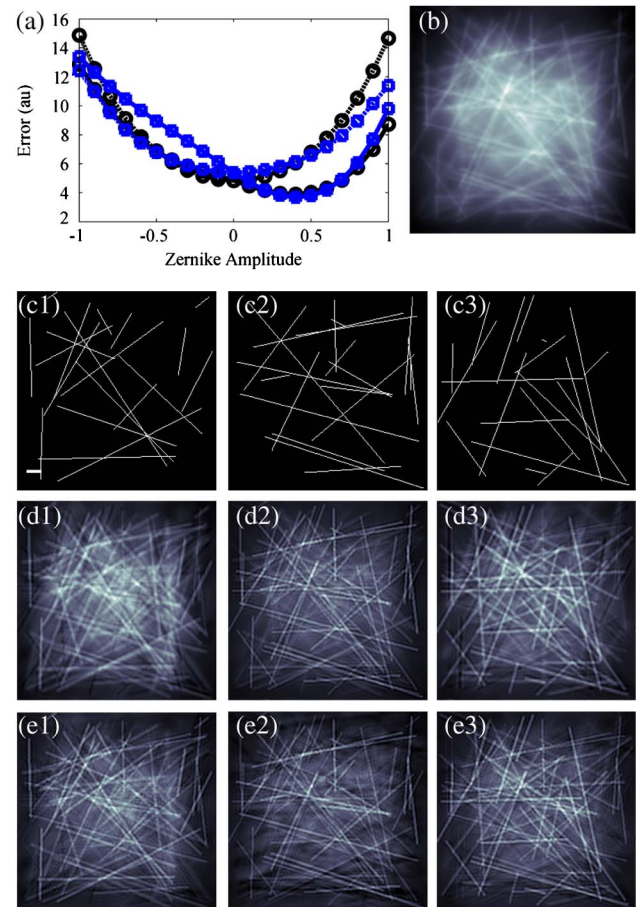


Fig. 6. Comparison of three-plane and five-plane PD for a simulated object with spacing of $dz = 0.2 \mu\text{m}$. The aberration is $\phi(u) = 0.5Z_3^1(u)$. (a) Plot of the error function versus Zernike amplitude. The black curves with circles are the three-plane PD error curves with $dz = 0.4 \mu\text{m}$, and the blue curves with squares are the five-plane PD error curves with $dz = 0.2 \mu\text{m}$. The solid curves show the error versus the amplitude of $Z_3^1(u)$. The dashed curves show the error versus the amplitude of $Z_3^{-1}(u)$. (b) PD image with no induced aberration. (c) Simulated object. (d) Result of three-plane PD calculation. (e) Result of five-plane PD calculation. (1) The first column is the plane below focus. (2) The second column is the focal plane. And (3) the third column is the plane above focus. $\mu = 0.1$. Scale bar is $1 \mu\text{m}$.

sample in Fig. 1. Two images were taken of a thin section of *Arabidopsis thaliana* with GFP labelled cell wall. One PD image is in-focus and the other image is $1.0\ \mu\text{m}$ away from focus. The minimization of the cost function results in the wavefront shown in Fig. 1(d) which shows strong spherical aberration. The accuracy of the result is confirmed by the axial cross-section of the image, Fig. 1(e), which shows a strong asymmetry above and below focus which is characteristic of spherical aberration. In this case the spherical aberration is due to the setting of the correction collar on the $40\times$ air objective.

To test the different three-dimensional PD calculations we used simulations of line objects. The simulated objects consisted of $32\ 256 \times 256$ pixel planes. Each plane contains between 20 and 50 line objects. Unless otherwise stated, the planes are separated by an axial distance, $dz = 0.5\ \mu\text{m}$. The lateral pixel dimension is $50\ \text{nm}$. The numerical aperture is 1.4, and the wavelength is $515\ \text{nm}$. Widefield microscopy images were created from the simulated object using Eq. (15). For each plane contributing to the image $\mathcal{F}\{f(x; z)\}$ was multiplied by the OTF, $\mathcal{F}\{s_k(x; z)\}$. Figure 2 shows an example of a simulated image with $dz = 0.25\ \mu\text{m}$. For the PD calculations on simulated objects, the phase is described by the first 15 Zernike

polynomials. In this paper, the aberration is set to be $\phi(u) = 0.5Z_3^1(u)$ in the back pupil plane. For the Zernike polynomials, $Z_n^m(u)$, n is the radial index and m is the azimuthal index where $m < 0$ for the odd phase term, and $\int Z_n^m(u)^2 d^2u = 1$.

An example of three-dimensional PD (Section 3.A) is shown in Fig. 3 on a simulated image. The PD aberrations are $\theta_1(u) = 0$ and $\theta_2(u) = 0.5Z_2^2(u)$. In this case, PD does an excellent job of estimating the image. Out-of-focus light is completely removed from the object estimate [Fig. 3(b)]. The cost function also is exactly zero at the correct value for the aberrations.

Next we look at multiplane PD. Figure 4 compares the results of two-dimensional PD with three-plane PD on a simulated object with planes spaced $0.5\ \mu\text{m}$ apart. The PD aberrations are $\theta_1(u) = 0$ and $\theta_2(u) = 0.5Z_2^2(u)$ for 2D PD. For three-plane PD, we add an extra image with $\theta_3(u) = 0.5Z_0^0(u)$. For both two-dimensional PD and three-plane PD, the cost function is a minimum at the correct value of the phase aberration, but in three-plane PD, the error function is lower. For three-plane PD, the central plane of the object estimate is significantly improved compared to 2D PD. The background is much lower and the image is sharper. The upper and lower planes of the three-plane PD object estimate

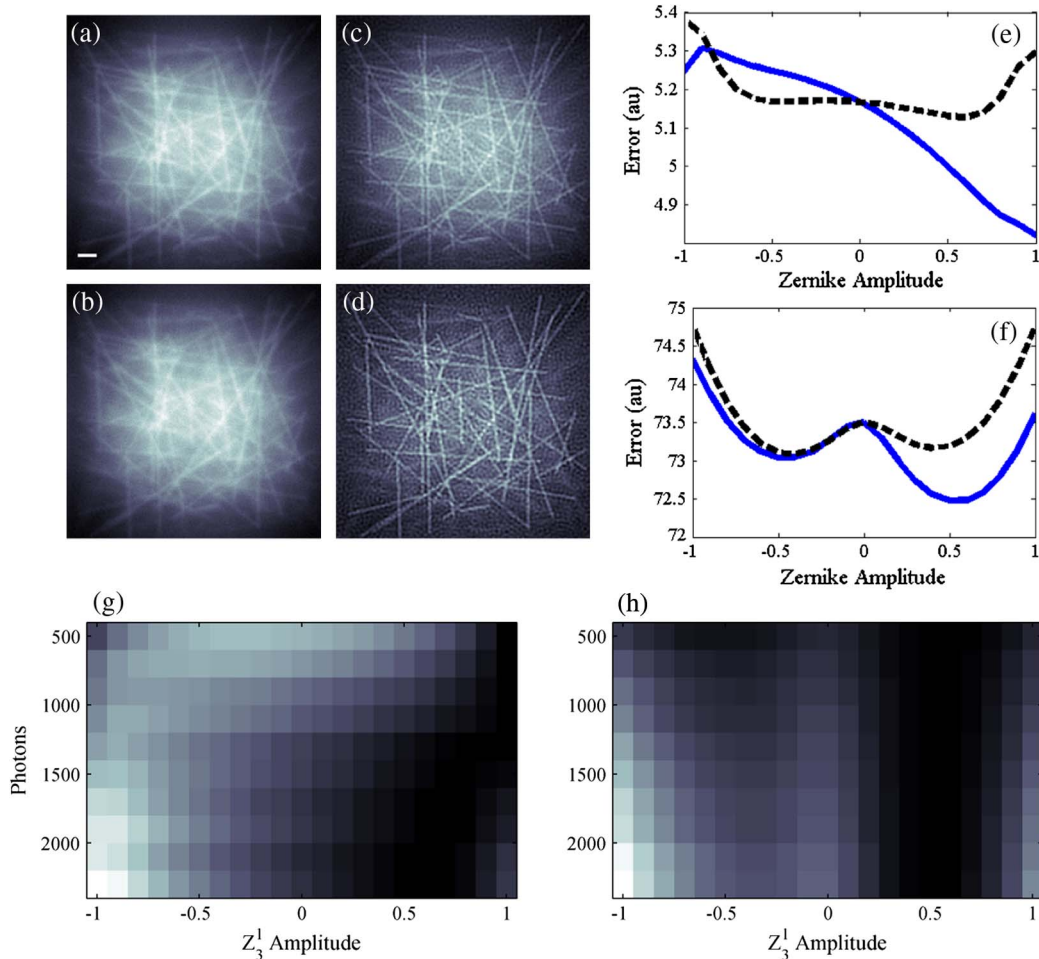


Fig. 7. Comparison of two-dimensional PD and three-plane PD with noise. Poisson noise is added to the image assuming the object emits 1000 photons per pixel. (a) PD image with $\theta_1(u) = 0$. (b) PD image with $\theta_2(u) = 0.5Z_2^2(u)$. The wavefront aberration is $\phi(u) = 0.5Z_3^1(u)$. (c) Optimal PD image from two-dimensional PD equation. (d) Optimal PD image from three-plane PD. (e) Error metric, $L(\bar{a})$ versus Zernike amplitude for two-dimensional PD. Black dashed curve: error metric versus Z_3^{-1} . Blue solid curve: error metric versus Z_3^1 . (f) Error metric for three-plane PD. (g) Intensity map of the 2D error metric. Black shows the minimum of the error metric, and 2D PD does not accurately find the correct amplitude. (h) Intensity map of the three-plane error metric. The minimum remains near $0.5Z_3^1$. Scale bar is $1\ \mu\text{m}$.

are poor—they seem to have absorbed the background that was correctly eliminated from the in-focus plane.

In Fig. 5, we compare three-plane PD to five-plane PD for a simulated object with 32 planes spaced $0.5\ \mu\text{m}$ apart. For both three-plane and five-plane PD, we use five PD images with $\theta(u) = 0, 0.5Z_2^0(u), 0.5Z_2^2(u), 0.5Z_2^{-2}(u), 0.5Z_3^1(u)$, respectively. Using five planes instead of three improves the sectioning of both the central (in focus) plane, the central column in Fig. 5, and the planes above and below focus. Both three-plane and five-plane PD assign objects to the correct plane, although in three-plane PD, the planes above and below focus contain more out-of-focus light and artifacts. To test multiplane PD further, we tested an object with planes spaced $0.2\ \mu\text{m}$ apart so that multiple planes are within the depth of focus. The results are shown in Fig. 6. Here we see that the object estimates of the planes above and below focus contain a mixture of information from the object's central plane and planes above and below focus. We cannot expect the object estimates to perfectly assign the information to the correct axial position to better accuracy than the depth of focus. Including more planes excludes more of the background out-of-focus light so that the adjacent planes have less blur in five-plane PD. The minimum of $L(\bar{a})$ is not exactly at the correct value, which is perhaps a consequence of there being uncertainty with many planes within the depth of focus.

To understand the effect of noise on the PD calculation, we scale each image by the number of photons emitted per pixel and add Poisson noise. Figures 7(a)–7(f) are the same comparison as in Fig. 4 but with noise corresponding to 1000 photons per pixel. We see that the images resulting from the PD calculation are essentially the same with the three-plane PD image, Fig. 7(d), having less background than the two-dimensional PD result [Fig. 7(c)]. On the other hand, the cost function shows a much different result. The cost function for

2D PD has a minimum at a value for the amplitude of the aberration that is much greater than the actual aberration, $\phi(u) = 0.5Z_3^1(u)$ [Fig. 7(e)]. In contrast, three-plane PD still estimates the aberration reasonably well. The minimum of the cost function is at $0.55Z_3^1(u)$, as shown in Fig. 7(f). Figures 7(g) and 7(h) show intensity maps of the cost function as a function of the number of photons and the aberration amplitude for 2D PD and three-plane PD, respectively. As can be seen, the minimum of the cost function wanders off much more rapidly for 2D PD.

If Eq. (20) is used with $K = 1$, we recover the formula for Wiener deconvolution. Applying Eq. (21) with $K = 1$ and $P > 1$ is not strictly valid because the result is undefined for $\mu \rightarrow 0$, but we show the results for $\mu = 0.1$. With $K = 1$, we cannot estimate the wavefront aberrations, \bar{a} , but estimates of the object, f , can be recovered. A comparison of Wiener deconvolution with three-plane deconvolution is made in Fig. 8. Without aberrations, or with knowledge of the actual aberrations, the deconvolution approach gives results comparable to PD. If the aberrations are unknown, the results of deconvolution will be degraded as shown in Fig. 8(d) where an aberration of $\phi(u) = 0.5Z_2^2(u)$ is introduced into the image. Figure 8(d) has higher background than Fig. 8(c). Three-plane deconvolution still does a better job of removing out-of-focus information from the in-focus plane [Fig. 8(b)].

5. DISCUSSION AND CONCLUSION

One concern is that PD assumes a laterally shift-invariant set of images because PD determines the aberrations in the back pupil plane. To handle variations in the aberrations across an image, the set of PD images can be tiled and PD can be applied to each set of tiled images. As the original images are tiled into smaller subimages, the resolution of the wavefront will decrease although speed will increase for each subimage. The degrees of freedom of the wavefront are $\propto l_x l_y$ where l_x and l_y are the extents of the sub-image.

Our future work will focus on improving and demonstrating the applicability of this approach to biological imaging. Depth aberrations occur frequently in microscopy due to a possible mismatch in refractive index between the sample and the immersion medium. Depth aberrations can be easily included in the imaging model, Eq. (15), so that the refractive index and depth will be variables in the cost function and can be determined from the PD calculation. Otherwise the depth aberration will appear as a shift-variant error source in the calculation.

Another important problem with PD is speed. Because PD requires assembling multiple pupil functions for each calculated aberration and calculating the OTFs, the calculation of the cost function $L(\bar{a})$ is slow; calculating the cost function can take several seconds and minimizing the cost function, which requires many evaluations, can take tens of minutes. To improve the speed of the calculation, the algorithm can be implemented in faster languages than Python, the interpreted language we used for these results, and a parallelized version of the algorithm could be implemented on graphical processing units, which would be significantly faster. Using the extended Nijboer–Zernike approach for calculating the PSF would also increase the speed of the algorithm [25]. The Nijboer–Zernike theory allows the calculation of the

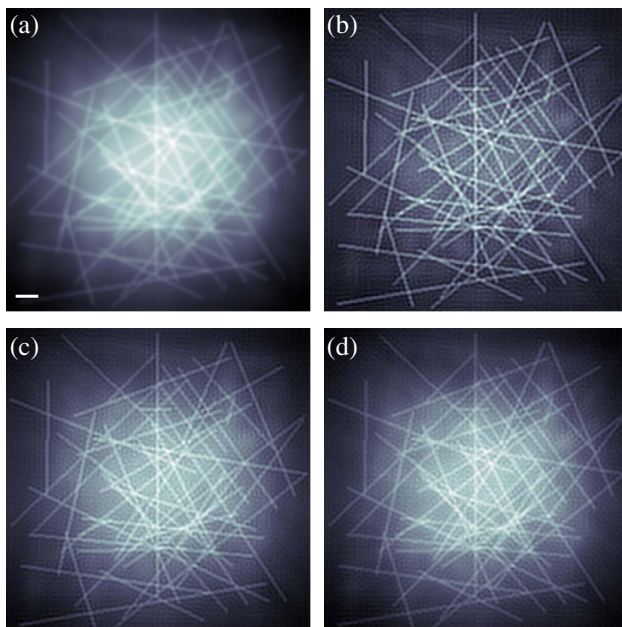


Fig. 8. Comparison of Wiener deconvolution and three-plane deconvolution. (a) Two-dimensional image from the center of a 32 plane stack with $0.5\ \mu\text{m}$ spacing. (b) Three-plane deconvolution. (c) Wiener deconvolution. (d) Wiener deconvolution with an aberrated image ($\theta(u) = 0.5Z_2^2(u)$). For (b)–(d), $\mu = 0.01$. Scale bar is $1\ \mu\text{m}$.

aberrated PSF quickly without the need for repeated Fourier transforms.

In conclusion, we have demonstrated that PD can be extended to three-dimensional wide-field imaging. For the case of a three-dimensionally shift-invariant system, the two-dimensional PD approach can be directly extended to three-dimensions. For the more general case of a system with aberrations that change along the axial dimension, we have developed a formalism for PD that takes into account the planes of the object that are out-of-focus. This multiplane PD approach can potentially determine the wavefront errors with significantly fewer measurements than the iterative optimization approaches currently in use. Another significant result is that the object estimate f returned by multiplane PD is a significant improvement over the original images and the object estimate returned by two-dimensional PD. f has significantly less background. These results should improve the performance of phase-diversity in biological imaging applications and possibly allow PD to be used as the wavefront measurement method in biological AO applications.

ACKNOWLEDGMENTS

This work was supported by the University of Georgia Research Foundation and by the National Science Foundation (under grant MCB1052672). We thank Utku Avci from the University of Georgia Complex Carbohydrate Research Center for supplying the *Arabidopsis thaliana* sample.

REFERENCES

1. J. W. Hardy, *Adaptive Optics for Astronomical Telescopes*, Oxford Series in Optical and Imaging Sciences; 16 (Oxford University, 1998).
2. R. K. Tyson, *Principles of Adaptive Optics*, 2nd ed. (Academic, 1997).
3. A. Roorda, F. Romero-Borja, W. Donnelly III, H. Queener, T. Hebert, and M. Campbell, "Adaptive optics scanning laser ophthalmoscopy," *Opt. Express* **10**, 405–412 (2002).
4. M. J. Booth, "Adaptive optics in microscopy," *Phil. Trans. R. Soc. A* **365**, 2829–2843 (2007).
5. J. A. Kubby, *Adaptive Optics for Biological Imaging* (CRC press, 2013).
6. X. Tao, B. Fernandez, O. Azucena, M. Fu, D. Garcia, Y. Zuo, D. C. Chen, and J. Kubby, "Adaptive optics confocal microscopy using direct wavefront sensing," *Opt. Lett.* **36**, 1062–1064 (2011).
7. X. Tao, O. Azucena, M. Fu, Y. Zuo, D. C. Chen, and J. Kubby, "Adaptive optics microscopy with direct wavefront sensing using fluorescent protein guide stars," *Opt. Lett.* **36**, 3389–3391 (2011).
8. A. J. Wright, D. Burns, B. A. Patterson, S. P. Poland, G. J. Valentine, and J. M. Girkin, "Exploration of the optimisation algorithms used in the implementation of adaptive optics in confocal and multiphoton microscopy," *Microsc. Res. Tech.* **67**, 36–44 (2005).
9. O. Albert, L. Sherman, G. Mourou, T. B. Norris, and G. Vdovin, "Smart microscope: an adaptive optics learning system for aberration correction in multiphoton confocal microscopy," *Opt. Lett.* **25**, 52–54 (2000).
10. S. P. Poland, A. J. Wright, and J. M. Girkin, "Evaluation of fitness parameters used in an iterative approach to aberration correction in optical sectioning microscopy," *Appl. Opt.* **47**, 731–736 (2008).
11. M. J. Booth, "Wave front sensor-less adaptive optics: a model-based approach using sphere packings," *Opt. Express* **14**, 1339–1352 (2006).
12. M. J. Booth, "Wavefront sensorless adaptive optics for large aberrations," *Opt. Lett.* **32**, 5–7 (2007).
13. R. W. Gerchberg and W. O. Saxton, "A practical algorithm for the determination of the phase from image and diffraction plane pictures," *Optik* **35**, 237–246 (1972).
14. J. R. Fienup, "Reconstruction of an object from the modulus of its Fourier transform," *Opt. Lett.* **3**, 27–29 (1978).
15. J. R. Fienup, J. C. Marron, T. J. Schulz, and J. H. Seldin, "Hubble space telescope characterized by using phase-retrieval algorithms," *Appl. Opt.* **32**, 1747–1767 (1993).
16. B. M. Hanser, M. G. Gustafsson, D. A. Agard, and J. W. Sedat, "Phase retrieval for high-numerical-aperture optical systems," *Opt. Lett.* **28**, 801–803 (2003).
17. P. Kner, L. Winoto, D. A. Agard, and J. W. Sedat, "Closed loop adaptive optics for microscopy without a wavefront sensor," *Proc. SPIE* **7570**, 757006 (2010).
18. R. A. Gonsalves, "Phase retrieval and diversity in adaptive optics," *Opt. Eng.* **21**, 215829 (1982).
19. R. G. Paxman, T. J. Schulz, and J. R. Fienup, "Joint estimation of object and aberrations by using phase diversity," *J. Opt. Soc. Am. A* **9**, 1072–1085 (1992).
20. E. F. Hom, F. Marchis, T. K. Lee, S. Haase, D. A. Agard, and J. W. Sedat, "AIDA: an adaptive image deconvolution algorithm with application to multi-frame and three-dimensional data," *J. Opt. Soc. Am. A* **24**, 1580–1600 (2007).
21. G. Chenegros, L. M. Mugnier, F. Lacombe, and M. Glanc, "3D phase diversity: a myopic deconvolution method for short-exposure images: application to retinal imaging," *J. Opt. Soc. Am. A* **24**, 1349–1357 (2007).
22. H. I. Campbell, S. Zhang, A. H. Greenaway, and S. Restaino, "Generalized phase diversity for wave-front sensing," *Opt. Lett.* **29**, 2707–2709 (2004).
23. Z. Kam, P. Kner, D. Agard, and J. W. Sedat, "Modelling the application of adaptive optics to wide-field microscope live imaging," *J. Microsc.* **226**, 33–42 (2007).
24. P. A. Stokseth, "Properties of a defocused optical system," *J. Opt. Soc. Am.* **59**, 1314–1321 (1969).
25. A. J. E. M. Janssen, "Extended Nijboer–Zernike approach for the computation of optical point-spread functions," *J. Opt. Soc. Am. A* **19**, 849–857 (2002).