Wave-front slope estimation

Marcos A. van Dam and Richard G. Lane

Department of Electrical and Electronic Engineering, University of Canterbury, Private Bag 4800, Christchurch, New Zealand

Received August 3, 1999; revised manuscript received March 23, 2000; accepted March 24, 2000

The conventional way of measuring the average slope of the phase of a wave front is from the centroid of the image formed at the focal plane. We show the limitations of using the centroid and present an optimal estimator along with the derivation of its lower error bound for a diffraction-limited image. The method is extended to slope estimation in the case of a random aberration introduced by atmospheric turbulence. It was found that the variance of the error of the slope estimator can be improved significantly at low turbulence levels by using the minimum mean-square-error estimator instead of the centroid. © 2000 Optical Society of America [S0740-3232(00)00407-5]

OCIS codes: 010.7350, 010.1080.

1. INTRODUCTION

An important problem in optics is to determine the least-mean-square (LMS) slope of the phase of an incoming wave front. The slope is conventionally estimated by using a centroid estimator that takes the center of gravity of the photons at the image plane. We show that this estimator is optimal for a Gaussian-distributed speckle image. However, for any finite-size aperture with diffraction-limited optics on an infinite plane, the variance of the centroid estimator is infinite.[1] In practice, the detection plane of all real sensors is a finite region, referred to in this paper as the truncated plane. Detection on this truncated plane has three effects. First, it causes the variance of the centroid estimator to become finite, as clearly the centroid estimate cannot lie outside the truncated plane. Second, it introduces a bias toward the center of the plane. Third, it results in the loss of information, as some photons are not detected.

A better estimator is the maximum-likelihood (ML) estimator, defined in Eq. (8). The Cramér–Rao bound is a lower bound on the variance of an unbiased estimator. It is shown that for estimates from a sufficient number of photons, the ML estimators for circular and rectangular apertures approach this bound. The estimate can be improved further by adding prior information about the distribution of the wave-front slopes. Conventionally, this is done by assuming a probability distribution for the slopes, but we also demonstrate that truncating the image plane is equivalent to adding prior information.

The ML estimator is then used to estimate the tip/tilt terms of a speckle image from a simulated phase screen with Kolmogorov statistics. This is useful in the Shack–Hartmann sensor, which divides the aperture into a two-dimensional array of lenslets. The average slope of the phase in the lenslets can then be used to reconstruct a more detailed estimate of the phase distortion, by either local or modal reconstruction.[2] The performance of this sensor is thus strongly dependent on the estimate of the slope in the individual apertures.

2. SLOPE ESTIMATION FROM SPECKLE IMAGES

Light from distant sources, such as astronomical objects, can be considered to consist of planar wave fronts with uniform amplitude before they enter the earth's atmosphere. The turbulence of the atmosphere causes a variation in refractive indices, which in turn changes the wave fronts' phase. The propagation of these aberrated wave fronts to ground level results in nonuniform wave-front amplitudes and phases at the aperture, with the latter effect by far the most significant. At good astronomical sites, the wave front is well modeled by an amplitude, $A(\mathbf{x})$, which is constant across the aperture and zero elsewhere, with an associated phase $\phi(\mathbf{x})$.[3]

The wave is focused onto the focal plane of a telescope. We adopt the following notation: $\mathbf{x} = (x, y)$ are the Cartesian coordinates in the aperture, with corresponding coordinates (ξ, η) in the focal plane, where the photons are detected. The relationship between the incident wave, $A(\mathbf{x})\exp[i\phi(\mathbf{x})]$, and the speckle image, $I(\xi, \eta)$, at the focal plane is given by[4]

$$I(\xi, \eta) = \frac{1}{\lambda^2 f^2}\left|\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} A(\mathbf{x})\exp[i\phi(\mathbf{x})] \times \exp\left[-i\frac{2\pi}{\lambda f}(x\xi + y\eta)\right]dx\,dy\right|^2, \quad (1)$$

where λ is the wavelength of the photons and f is the focal length of the lens or mirror. For convenience of notation, we will define $\mathbf{u} = (u, v) = (2\pi/\lambda f)(\xi, \eta)$, whereupon the complex amplitudes at the aperture and the image plane form a Fourier transform pair.

From a simple geometric argument, it can be seen that the slope, $\mathbf{s}$, is related to the displacement of the speckle image, $\mathbf{a} = (a_u, a_v)$, by $\mathbf{a} = \mathbf{s}$. The value of $\mathbf{a}$ corresponding to the LMS slope is referred to here as the center. This is different from the centroid, which represents an estimate of the center.

0740-3232/2000/071319-06$15.00

Consider a planar wave front from an incoherent source passing through an aberration-free circular lens of radius r. When the image is diffraction limited, an Airy disk pattern is observed.[4] The point-spread function (PSF), $h(\rho)$, of the Airy disk is the intensity pattern observed at the image plane. When normalized so that the area integral is unity, the PSF can also be thought of statistically as a probability density function (pdf), i.e., the probability of a photon landing at a certain region in the detection plane. The Airy disk corresponding to an aperture of radius $\lambda f/2\pi$ is given by

$$h(\rho) = \frac{J_1(\rho)^2}{\pi\rho^2}, \tag{2}$$

where $\rho = \sqrt{u^2 + v^2}$ is the radial distance from the center of the image and $J_1(\rho)$ is a Bessel function of order 1 of the first kind. For determining the location of the center of the PSF from a speckle image, the centroid estimator is usually used.[5] In Subsection 3.B we show that the centroid is an optimal estimator for a Gaussian PSF, not for an Airy disk.

3. MAXIMUM-LIKELIHOOD ESTIMATION

A. Maximum-Likelihood Estimator

The ML estimator yields the estimate for the parameter, $\mathbf{a}$, that maximizes the probability of obtaining the N received points, $\mathbf{u}_i = (u_i, v_i)$. This can be written as

$$\hat{\mathbf{a}} = \max_{\mathbf{a}} \prod_{i=0}^{N-1} h(\mathbf{u}_i|\mathbf{a}). \tag{3}$$

Here $\Pi_{i=0}^{N-1} h(\mathbf{u}_i|\mathbf{a})$ is the likelihood distribution and $\max_{\mathbf{a}}$ represents the value of $\mathbf{a}$ that maximizes it.

B. Slope Estimator for a Gaussian Point-Spread Function

If we consider a Gaussian PSF as the pdf,

$$g(\mathbf{u}) = \frac{1}{2\pi\sigma^2}\exp(-\mathbf{u}^2/2\sigma^2), \tag{4}$$

the likelihood equation is

$$\prod_{i=0}^{N-1} g(\mathbf{u}_i|\mathbf{a}) = \prod_{i=0}^{N-1} \frac{1}{2\pi\sigma^2}\exp\left[\frac{-(u_i - a_u)^2 - (v_i - a_v)^2}{2\sigma^2}\right]. \tag{5}$$

Taking the log of Eq. (5) gives

$$\sum_{i=0}^{N-1} \ln g(\mathbf{u}_i|\mathbf{a}) = N\ln\frac{1}{2\pi\sigma^2} - \sum_{i=0}^{N-1}\left[\frac{(u_i - a_u)^2 + (v_i - a_v)^2}{2\sigma^2}\right]. \tag{6}$$

Differentiating with respect to a_u and equating to zero to find the maximum gives

$$\frac{1}{\sigma^2}\sum_{i=0}^{N-1}(u_i - a_u) = 0, \tag{7}$$

which holds if $a_u = (1/N)\Sigma_{i=0}^{N-1} u_i$. Similarly, $a_v = (1/N)\Sigma_{i=0}^{N-1} v_i$. This is the conventional centroid estimator, and it is optimal for a Gaussian pdf.

C. Slope Estimator for an Airy Disk

We now consider the PSF to be an Airy disk. If a single photon is detected at $\mathbf{u}_0 = (u_0, v_0)$, then the ML estimator for the center of the disk is

$$\max_{\mathbf{a}} h(\mathbf{u}_0|\mathbf{a}) = \max_{\mathbf{a}} h\{[(u_0 - a_u)^2 + (v_0 - a_v)^2]^{1/2}\}. \tag{8}$$

Equation (8) has a solution at $\hat{\mathbf{a}} = (u_0, v_0)$. For one photon this produces the same estimate as the centroid. The expected value of the single-photon estimator on an infinite plane,

$$E[\hat{\mathbf{a}}] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \hat{\mathbf{a}} h(\mathbf{u}_0|\mathbf{a})\mathrm{d}u_0\mathrm{d}v_0, \tag{9}$$

is equal to $\mathbf{a}$, showing that it is an unbiased estimator. Its variance,

$$\begin{aligned} E[(\hat{\mathbf{a}} - \mathbf{a})^2] &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} (\hat{\mathbf{a}} - \mathbf{a})^2 h(\mathbf{u}_0|\mathbf{a})\mathrm{d}u_0\mathrm{d}v_0 \\ &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} (\mathbf{u}_0 - \mathbf{a})^2 \frac{J_1(\mathbf{u}_0 - \mathbf{a})^2}{\pi(\mathbf{u}_0 - \mathbf{a})^2}\mathrm{d}u_0\mathrm{d}v_0 \\ &= 2\pi\int_0^{\infty} \frac{J_1(\rho)^2}{\pi}\rho\mathrm{d}\rho, \end{aligned} \tag{10}$$

is infinite. Physically, the reason for this is that the probability of detecting photons very far away from the center of the Airy disk does not decay quickly enough.[1] The second moment of the PSF of any discontinuous aperture measured on any infinite plane, not just on the focal plane, is also infinite.[6] Since the photons are independent, increasing the number of photons does not make the variance of the centroid estimator finite. This shows that the centroid estimator on an infinite plane is not feasible. For N photons, however, the ML estimate is $\max_{\mathbf{a}} \Pi_{i=0}^{N-1} h\{[(u_i - a_u)^2 + (v_i - a_v)^2]^{1/2}\}$, a quantity that differs from the centroid. In Subsection 3.D the variance of the ML estimator of the Airy disk on an infinite plane is shown to decrease asymptotically to $1/N$ in each of the two dimensions.

Truncating the plane constrains the centroid estimate to a certain region, making the variance finite. The truncated plane is placed where the center is expected to be; thus the smaller the plane, the more prior information is assumed. Therefore by adding prior information, the truncated plane can improve the centroid estimate, even though some photons are lost.

There are two major differences between the Airy disk and Gaussian distributions and hence the ML and centroid estimators. The rate of decay of the PSF far away from the center is proportional to $\mathbf{u}^{-2}$ for the Airy disk and $\exp(-\mathbf{u}^2/2\sigma^2)$ for the Gaussian. The Airy disk PSF has oscillatory variations in its intensity, including nulls, where the probability of detecting a photon is zero. The

latter is significant because the position of the nulls is helpful in identifying the center, thus reducing the variance of the estimate. It is important to note that, unlike the centroid estimator, the ML estimator requires that the center be calculated in two dimensions simultaneously. In other words, the value of a_u affects a_v and vice versa. As a consequence, it is more computationally intensive.

D. Cramér–Rao Bound of the Airy Disk

The Cramér–Rao bound is a lower bound on the variance of an unbiased estimator given its conditional probability distribution, $h(\mathbf{u}|\mathbf{a})$.[7] The bound can be calculated provided that $[\partial(\mathbf{u}|\mathbf{a})]/\partial\mathbf{a}$ and $[\partial^2 h(\mathbf{u}|\mathbf{a})]/\partial\mathbf{a}^2$ exist and are absolutely integrable, which is the case for all the PSF's of interest. If $\hat{\mathbf{a}}$ is an unbiased estimator of $\mathbf{a}$, then we can define a lower bound

$$\mathrm{Var}(\hat{a}_i - a_i) \geqslant J^{ii}, \qquad i \in 1, 2, \tag{11}$$

where J^{ii} is the iith element of the 2×2 matrix J^{-1}. J is called the information matrix and is given by

$$J = -E\begin{bmatrix} \dfrac{\partial^2 \ln h(\mathbf{u}|\mathbf{a})}{\partial {u_u}^2} & \dfrac{\partial^2 \ln h(\mathbf{u}|\mathbf{a})}{\partial a_u \partial a_v} \\ \dfrac{\partial^2 \ln h(\mathbf{u}|\mathbf{a})}{\partial a_v \partial a_u} & \dfrac{\partial^2 \ln h(\mathbf{u}|\mathbf{a})}{\partial {a_v}^2} \end{bmatrix}. \tag{12}$$

We evaluate the leading diagonal elements of the matrix by letting $\rho = [(u - a_u)^2 + (v - a_v)^2]^{1/2}$. Then

$$\begin{aligned} \frac{\partial^2 \ln h(\mathbf{u}|\mathbf{a})}{\partial {a_u}^2} &= \frac{\partial^2 \ln h(\rho)}{\partial {a_u}^2} \\ &= \left(\frac{\partial \rho}{\partial a_u}\right)^2 \frac{\partial^2 \ln h(\rho)}{\partial \rho^2} + \frac{\partial^2 \rho}{\partial {a_u}^2} \frac{\partial \ln h(\rho)}{\partial \rho} \\ &= \frac{(u - a_u)^2}{\rho^2} \frac{\partial^2 \ln h(\rho)}{\partial \rho^2} \\ &\quad + \frac{(v - a_v)^2}{\rho^3} \frac{\partial \ln h(\rho)}{\partial \rho}. \end{aligned} \tag{13}$$

This gives

$$\begin{aligned} J_{11} &= -\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\left[\frac{(u - a_u)^2}{\rho^2} \frac{\partial^2 \ln h(\rho)}{\partial \rho^2}\right. \\ &\quad \left. + \frac{(v - a_v)^2}{\rho^3} \frac{\partial \ln h(\rho)}{\partial \rho}\right] h(\rho)\,\mathrm{d}u\,\mathrm{d}v. \end{aligned} \tag{14}$$

A similar approach for J_{22} can be used to show that

$$\begin{aligned} J_{11} + J_{22} &= -\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\left[\frac{(u - a_u)^2 + (v - a_v)^2}{\rho^2} \frac{\partial^2 \ln h(\rho)}{\partial \rho^2}\right. \\ &\quad \left. + \frac{(u - a_u)^2 + (v - a_v)^2}{\rho^3} \frac{\partial \ln h(\rho)}{\partial \rho}\right] \\ &\quad \times h(\rho)\,\mathrm{d}u\,\mathrm{d}v \\ &= -\int_0^{2\pi}\int_0^{\infty}\left[\frac{\partial^2 \ln h(\rho)}{\partial \rho^2} + \frac{1}{\rho}\frac{\partial \ln h(\rho)}{\partial \rho}\right] \\ &\quad \times h(\rho)\rho\,\mathrm{d}\rho\,\mathrm{d}\theta. \end{aligned} \tag{15}$$

Substituting Eq. (2) into Eq. (15) gives $J_{11} + J_{22} = 2$. By the symmetry of u and v, we can see that $J_{11} = J_{22} = 1$. Similarly,

$$\begin{aligned} J_{12} &= -\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(u - a_u)(v - a_v) \\ &\quad \times \left[\frac{1}{\rho^2}\frac{\partial^2 \ln h(\rho)}{\partial \rho^2} - \frac{1}{\rho^3}\frac{\partial \ln h(\rho)}{\partial \rho}\right] h(\rho)\,\mathrm{d}u\,\mathrm{d}v, \\ &= 0 \end{aligned} \tag{16}$$

since the integrand is antisymmetric:

$$J = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}. \tag{17}$$

Taking the inverse of J gives $J^{11} = J^{22} = 1$. This results in a lower bound of 1 for the variance in each of the axes. This bound is equivalent to the quantum mechanics limit imposed by the uncertainty principle, $\Delta x \Delta p \geqslant h/4\pi$, where h is Planck's constant and Δx and Δp are, respectively, the uncertainty of the position and the momentum of the photon.[8]

It can be shown that under certain general conditions[9] (which are satisfied in this case), the likelihood equation has a solution that converges to $\mathbf{a}$ as $N \to \infty$. It is also asymptotically efficient; i.e.,

$$\lim_{N\to\infty}\mathrm{Var}(\hat{a}_i - a_i) = J^{ii}, \qquad i \in 1, 2, \tag{18}$$

and asymptotically Gaussian. The ML estimator for N photons is $\max_{\mathbf{a}} \Pi_{i=0}^{N-1} h\{[(u_i - a_u)^2 + (v_i - a_v)^2]^{1/2}\}$, yielding a lower bound of $1/N$ for the variance. This is expected from the central-limit theorem, as the likelihood equation approaches a Gaussian distribution. Hence the variance of the ML estimator converges to $1/N$ as $N \to \infty$.

For completeness, the equivalent result for a diffraction-limited rectangular aperture is stated here. The PSF in one direction for a lens with its length equal to $\lambda f/\pi$ is given by

$$s(u) = \frac{\sin^2(u)}{\pi u^2}. \tag{19}$$

The information matrix is now a scalar:

$$\begin{aligned} J &= -\int_{-\infty}^{\infty}\frac{\mathrm{d}^2 \ln s(u)}{\mathrm{d}u^2} s(u)\,\mathrm{d}u \\ &= -\int_{-\infty}^{\infty}\left[\frac{2}{u^2} - \frac{2}{\sin^2(u)}\right]\frac{\sin^2(u)}{\pi u^2}\,\mathrm{d}u \\ &= 4/3. \end{aligned} \tag{20}$$

Hence the Cramér–Rao bound for the variance is 3/4 in each direction. This is slightly less than a circular aperture of radius $\lambda f/2\pi$, because the square lens has a larger area than the circular one.

4. PRIOR INFORMATION

A. Maximum *a posteriori* Estimator

The reason the centroid estimator works in practice is that the detector is a truncated plane. This introduces implicit prior information, as the detector is placed where the center of the image is expected to be. If prior information about the parameters to be estimated is known, then the maximum *a posteriori* (MAP) estimator is appropriate. From Bayes's rule, $f(\mathbf{u}, \mathbf{a}) = h(\mathbf{u}|\mathbf{a})p(\mathbf{a})$, where $f(\mathbf{u}, \mathbf{a})$ denotes the joint probability of $\mathbf{u}$ and $\mathbf{a}$ and $p(\mathbf{a})$ is the prior probability distribution of the centers. Hence the MAP estimate for one photon, obtained from the maximum of the a posteriori distribution is

$$\max_{\mathbf{a}} f(\mathbf{u}_0, \mathbf{a}) = \max_{\mathbf{a}} h\{[(u_0 - a_u)^2 + (v_0 - a_v)^2]^{1/2}\}p(\mathbf{a}). \tag{21}$$

Consider the case in which the center of the PSF is normally distributed around the origin with variance σ^2 in both u and v. Then

$$p(\mathbf{a}) = \frac{1}{2\pi\gamma^2}\exp(-\mathbf{a}^2/2\gamma^2). \tag{22}$$

The information matrix, J, now consists of the sum of two parts: the data, J_D, which is equal to J in Eq. (17), and the prior, J_P[7]:

$$J_P = -E\begin{bmatrix} \dfrac{\partial^2 \ln p(\mathbf{a})}{\partial a_u^2} & \dfrac{\partial^2 \ln p(\mathbf{a})}{\partial a_u \partial a_v} \\ \dfrac{\partial^2 \ln p(\mathbf{a})}{\partial a_v \partial a_u} & \dfrac{\partial^2 \ln p(\mathbf{a})}{\partial a_v^2} \end{bmatrix} = \begin{bmatrix} 1/\gamma^2 & 0 \\ 0 & 1/\gamma^2 \end{bmatrix}. \tag{23}$$

Hence the Cramér–Rao lower bound is given by

$$J = J_D + J_P = \begin{bmatrix} 1/\gamma^2 + N & 0 \\ 0 & 1/\gamma^2 + N \end{bmatrix}, \tag{24}$$

and the lower bound in each axis is $J^{11} = J^{22} = 1/(1/\gamma^2 + N)$.

B. Maximum *a posteriori* Simulations

The performance of the MAP estimator of the center of a translated Airy disk was simulated on a plane. The center of the Airy disk was assigned a displacement from the origin that was normally distributed with a variance of 1. The translated Airy disk was then used to define the pdf of photon arrival. Photons were assigned a radius and an angle at random with a distribution given by the pdf to ensure the extent of the detection plane and that the resolution available was limited only by the working precision of the computer. Using discrete pixels would increase the variance of the estimate, as some information about where the photon lands would be lost. The results of the MAP estimator are plotted along with the Cramér–Rao bound in Fig. 1. The Cramér–Rao bound is very tight, even at relatively small N.

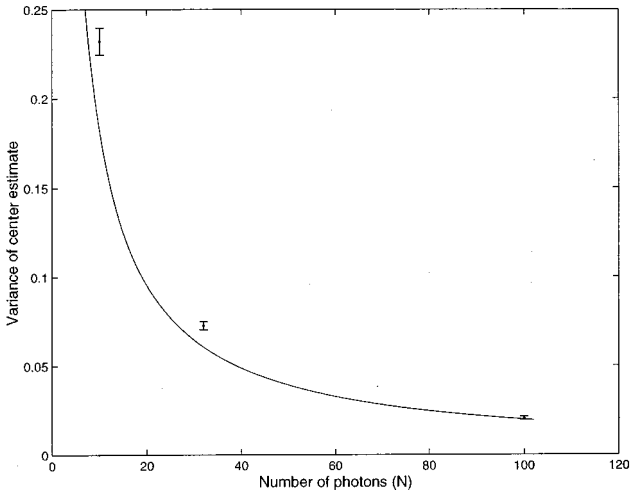

Fig. 1. Cramér–Rao bound for the Airy disk with N photons. The points, with 1σ error bars, represent the simulation results.

C. Effect of Truncated Plane

As mentioned in Subsection 3.B, the centroid is the optimal estimator for a Gaussian-distributed speckle image. Indeed, several authors explicitly assume a Gaussian distribution, as it is analytically more tractable.[10,11] The Airy disk is approximated by a two-dimensional Gaussian PSF, $g(\mathbf{u})$, with the e^{-1} point at a radial distance $\sqrt{2}\sigma$ from the peak, corresponding to $\sigma = 1.354$ for the PSF defined by Eq. (2).

This Gaussian model has proved successful in the design of adaptive optics systems, because when a truncated plane is used for detection, it introduces implicit prior information. The minimum mean square error (MMSE) is the mean of the a posteriori pdf and is given by

$$\hat{\mathbf{a}} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} a_u a_v p(\mathbf{a}) \times \prod_{i=0}^{N-1} h\{[(u_i - a_u)^2 + (v_i - a_v)^2]^{1/2}\}\mathrm{d}a_u\mathrm{d}a_v. \tag{25}$$

In the case of the centroid estimator, truncating the detection plane results in a trade-off between the amount of light collected and the implicit prior; the optimum detector size was previously investigated by Irwan and Lane.[1] Since each dimension can be treated separately, we consider the displacement in one dimension, u, on a square detector with length $2W$. The Gaussian approximation to the PSF is normally distributed with mean a and $\sigma = 1.354$. Using the centroid, we obtain the expected value of the estimator as

$$E[\hat{\mathbf{a}}] = \frac{\int_{-W}^{W} u g(u - a)\mathrm{d}u}{\int_{-W}^{W} g(u - a)\mathrm{d}u}, \tag{26}$$

where the denominator is a normalization constant such that the probability of receiving a photon on the truncated plane is unity. The estimator is biased toward the origin, as $E[\hat{a}] \leqslant a$. The MMSE estimate of a from an infinite detector is

$$E[\hat{a}] = \frac{\int_{-\infty}^{\infty} ug(u-a)p(a)\mathrm{d}u}{\int_{-\infty}^{\infty} p(a)\mathrm{d}a \int_{-\infty}^{\infty} g(u-a)\mathrm{d}u}$$

$$= \frac{p(a)\int_{-\infty}^{\infty} ug(u-a)\mathrm{d}u}{\int_{-\infty}^{\infty} p(a)\mathrm{d}a}$$

$$= \frac{p(a)a}{\int_{-\infty}^{\infty} p(a)\mathrm{d}a}, \tag{27}$$

where $p(a)$ is the pdf of the prior information. Equating the two estimates yields $p(a)$ to within a normalization constant:

$$p(a) = \frac{\int_{-W}^{W} ug(u-a)\mathrm{d}u}{a\int_{-W}^{W} g(u-a)\mathrm{d}u}. \tag{28}$$

This distribution, plotted in Fig. 2, must be normalized for the true pdf of the prior to be obtained. Note that this prior is associated with the detection of each photon, not with the estimation of the center. Hence for N photons the prior used should be $p(a)^N$.

From another viewpoint, truncating the plane is also equivalent to convolving the aperture with the scaled inverse Fourier transform of the detector.[6] The effective aperture is now continuous, so the variance of the PSF is finite. Hence the truncated plane can be thought of as an apodized aperture, where the variance is decreased at the expense of a reduction in photons.

D. Slope Estimation of Kolmogorov Turbulence

Atmospheric turbulence can be modeled as obeying Kolmogorov statistics.[3] This assumption results in randomly varying normally distributed wave-front slopes with zero mean and a variance in each direction of $7.17r_0^{-5/3}D^{-1/3}$, where r_0 is Fried's parameter and D is the diameter of the aperture.[12] In this case the PSF does not have an analytical description and must be computed numerically. The PSF can also be determined experimentally if the statistics of the turbulence are not known or there is significant noise in the detector.

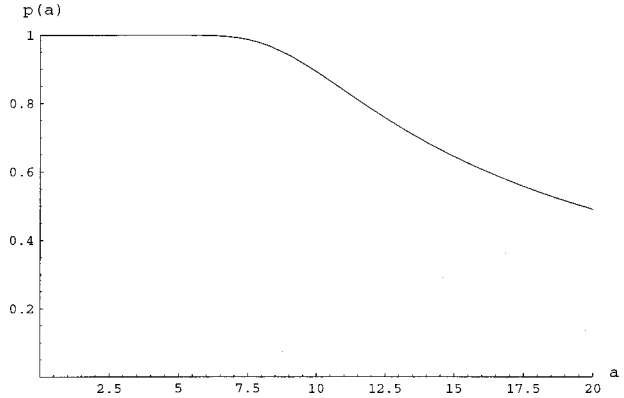

Fig. 2. Unnormalized prior distribution, $p(a)$, for $W = 10$.

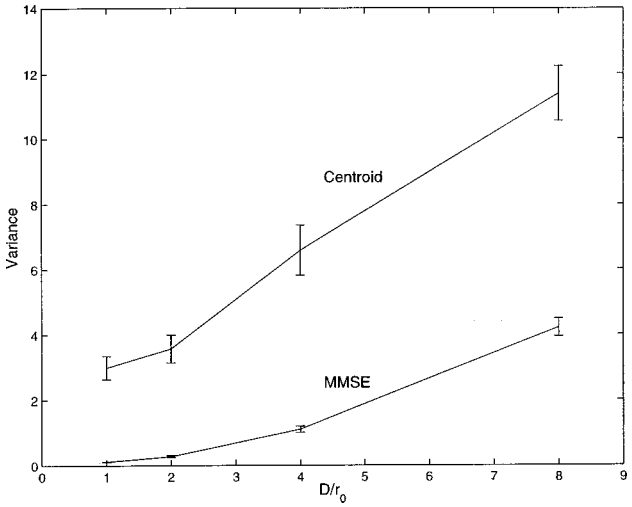

Fig. 3. Comparison of the variance of the error (pixels^2) of the MMSE and centroid estimators as a function of the level of turbulence on a truncated plane. The error bars are the 1σ uncertainty.

The MMSE is, by definition, the estimator with the smallest variance possible. Its main disadvantage is that it is computationally expensive, but for discrete data the integrations can be approximated by summations. In a CCD, the number of photons falling on each pixel is counted. For a $(2M+1)\times(2M+1)$ array, the MMSE can be approximated by a discrete version of Eq. (25):

$$\hat{\mathbf{a}} = \sum_{a_u=-M}^{M}\sum_{a_v=-M}^{M} a_u a_v p[(a_u^2 + a_v^2)^{1/2}] \times \prod_{i=0}^{N-1} h\{[(u_i - a_u)^2 + (v_i - a_v)^2]^{1/2}\}. \tag{29}$$

E. Kolmogorov Turbulence Simulations

The performance of the MMSE for Kolmogorov turbulence was simulated and compared with the centroid estimator. Random phase screens were generated with the method of Harding *et al.*[13] The LMS slope of the phase was computed directly from the phase screens for reference. The incoming wave fronts, assumed to be of unit amplitude, were 64 pixels in diameter inserted in an array of 128×128 zeros. The pdf at the image plane for that phase screen was obtained by using a discrete version of Eq. (1). From this pdf a speckle image was formed by use of Poisson statistics, with an average of ten photons, and 600 phase screens were simulated for each D/r_0 ratio. From the results plotted in Fig. 3, it can be seen that at low D/r_0 ratios the MMSE significantly outperforms the centroid estimator. As D/r_0 increases, the two estimates become more similar for several reasons. The definition of the diffraction nodes and antinodes becomes weaker, and these features are responsible for the low variance obtained in the Airy disk. The prior information becomes less useful as the variance of the distribution of the centers of the PSF increases. Also, as the

PSF gets broader, the PSF and the Gaussian become more similar inside the truncated plane.

A simulation was performed to see how sensitive the MMSE estimator is to the PSF and the prior information. In this simulation, the turbulence had a $D/r_0 = 2$ ratio but the estimator was fed the parameters corresponding to $D/r_0 = 4$. The MMSE produced an average error variance of 0.31, which is just outside the uncertainty bounds of the performance at $D/r_0 = 2(0.27 \pm 0.03)$. We conclude that the estimator is only weakly dependent on the turbulence statistics, provided that the roll-off of the intensity, $h(\mathbf{u})$, follows a $\mathbf{u}^{-2}$ law away from the center.

5. CONCLUSION

The optimal estimator for the LMS slope is the MMSE, which can include prior information about the PSF and the noise. When there is no prior information and the PSF is Gaussian, the optimal estimator is the conventional centroid. However, the centroid estimator of the speckle image of a sharp aperture on an infinite plane has infinite variance. By truncating the plane, one reduces the variance at the expense of using a biased estimate. This estimate is suboptimal also because of the loss of photons. Truncation introduces implicit prior information to the estimate, but this can also be introduced from a knowledge of the turbulence statistics.

At low photon counts, the MMSE estimator is much better than the centroid. For an average of ten photons for $D/r_0 = 1$, the difference was of the order of 30. As the level of turbulence per lenslet increases, the advantage obtained by using the MMSE is reduced.

The corresponding author, Richard G. Lane, can be reached by phone, 64-3-366-7001; fax, 64-3-364-2761; or e-mail, r.lane@elec.canterbury.ac.nz.

REFERENCES

1. R. Irwan and R. G. Lane, "Analysis of optimal centroid estimation applied to Shack–Hartmann sensing," Appl. Opt. **38**, 6737–6743 (1999).
2. W. H. Southwell, "Wave-front estimation from wave-front slope measurements," J. Opt. Soc. Am. **70**, 998–1006 (1980).
3. F. Roddier, "The effect of atmospheric turbulence in optical astronomy," *Progress in Optics*, E. Wolf, ed. (North-Holland, Amsterdam, 1981), pp. 283–376.
4. J. Goodman, *Introduction to Fourier Optics* (McGraw-Hill, San Francisco, Calif., 1968), pp. 59–65.
5. M. C. Roggemann and B. Welsh, *Imaging through Turbulence* (CRC Press, Boca Raton, Fla., 1996), pp. 182–183.
6. M. R. Teague, "Irradiance moments: their propagation and use for unique retrieval of phase," J. Opt. Soc. Am. **72**, 1199–1209 (1982).
7. H. L. Van Trees, *Detection, Estimation, and Modulation Theory*, Part 1 (Wiley, New York 1968), pp. 66–85.
8. S. Gasiorowicz, *Quantum Physics* (Wiley, New York, 1974), p. 55.
9. H. Cramér, *Mathematical Methods of Statistics* (Princeton U. Press, Princeton, N.J., 1946), pp. 498–505.
10. K. A. Winick, "Cramér–Rao lower bounds on the performance of charge-coupled-device optical position estimators," J. Opt. Soc. Am. A **3**, 1809–1815 (1986).
11. G. Cao and X. Yu, "Accuracy analysis of a Hartmann–Shack wavefront sensor operated with a faint object," Opt. Eng. **33**, 2331–2335 (1994).
12. R. J. Noll, "Zernike polynomials and atmospheric turbulence," J. Opt. Soc. Am. **66**, 207–211 (1976).
13. C. M. Harding, R. A. Johnston, and R. G. Lane, "Fast simulation of a Kolmogorov phase screen," Appl. Opt. **38**, 2161–2170 (1999).