

Joint estimation of Stokes images and aberrations from phase-diverse polarimetric measurements

John R. Valenzuela,^{1,*} Jeffrey A. Fessler,¹ and Richard G. Paxman¹

¹University of Michigan, 1301 Beal Avenue, Ann Arbor, Michigan 48109-2122, USA

²General Dynamics Advanced Information Systems, 1200 Joe Hall Drive, Ypsilanti, Michigan 48197, USA

*Corresponding author: jvalenz@umich.edu

Received November 24, 2009; revised March 22, 2010; accepted March 23, 2010;
posted March 23, 2010 (Doc. ID 120426); published April 30, 2010

The technique of phase diversity has been used in traditional incoherent imaging systems to jointly estimate an object and optical system aberrations. This paper extends the technique of phase diversity to polarimetric imaging systems. Specifically, we describe penalized-likelihood methods for jointly estimating Stokes images and optical system aberrations from measurements that contain phase diversity. Jointly estimating Stokes images and optical system aberrations involves a large parameter space. A closed-form expression for the estimate of the Stokes images in terms of the aberration parameters is derived and used in a formulation that reduces the dimensionality of the search space to the number of aberration parameters only. We compare the performance of the joint estimator under both quadratic and edge-preserving regularization; we also compare the performance of the reduced parameter search strategy to the full parameter search strategy under quadratic regularization. The joint estimator with edge-preserving regularization yields higher fidelity polarization estimates than with quadratic regularization. With the reduced parameter search strategy, accurate aberration estimates can be obtained without recourse to regularization “tuning.” © 2010 Optical Society of America

OCIS codes: 100.3020, 100.3190, 100.3010.

1. INTRODUCTION

Polarimetric imaging systems acquire data that can be used to infer the polarization state of an optical field [1,2]. The polarization state of an optical field across a scene contains information related to surface features such as shape and roughness [3]. Naturally occurring objects typically have a larger surface granularity than man-made objects, so polarimetry offers the potential for improved target detection and identification over other imaging modalities [4].

The polarization state of a transverse optical field can be specified by the Stokes vector $\mathbf{S}=(S_0, S_1, S_2, S_3)$ [5,6]. The elements of \mathbf{S} are functions of the optical intensity and defined in the following way: S_0 is the total optical intensity, S_1 is the difference between the optical intensity transmitted by a linear polarizer with pass axis oriented at 0° (reference) and one having pass axis oriented at 90° , S_2 is the difference between the optical intensity transmitted by a linear polarizer with pass axis oriented at 45° and one having pass axis oriented at 135° , and S_3 is the optical intensity transmitted by a right circular polarizer and a left circular polarizer. In the majority of remote-sensing applications the linear polarization state of the optical field is of interest and so the S_3 component is ignored. We adopt this usual simplification of considering only the first three components of the Stokes vector, though the method generalizes easily.

Polarimeters, like traditional incoherent imaging sensors, have resolution limits that depend on noise and system point-spread function. In remote-sensing applications, degradations in the point-spread function are often due to atmospheric turbulence, residual aberrations in

the optical system, or misalignment among components in the optical system. We previously developed a method for estimating Stokes images directly from polarimetric measurements [7]. That work assumed complete knowledge of the system point-spread function and was thus limited in its range of application. In this paper, we propose methods that overcome this limitation by introducing phase diversity into the polarimetric measurements. In traditional incoherent imaging the technique of phase diversity has been used to jointly estimate the object and optical aberrations in the presence of atmospheric turbulence [8]. Phase diversity requires the simultaneous collection of two or more images that are related via a deterministic phase perturbation. Typically, two images are collected: one is the conventional in-focus image and the second image is acquired on a separate focal plane that is translated along the optical axis thereby inducing a known defocus to the second image. Figure 1 shows a typical phase diversity configuration. A direct extension of the traditional phase diversity strategy to polarimetry would be to acquire two measurements per polarimetric channel; a four-channel polarimeter would be extended to an eight-channel polarimeter. Here we present two algorithms to jointly estimate the Stokes vectors and optical aberrations using a simpler four channel phase diverse polarimeter. The method could be adapted easily to eight-channel polarimeters and other variations, but a four-channel polarimeter configuration is particularly attractive in terms of cost and complexity of hardware.

One acquisition parameter that must be chosen is the amount of defocus in the diversity channel(s). Choosing the optimal amount of phase diversity for phase-diverse

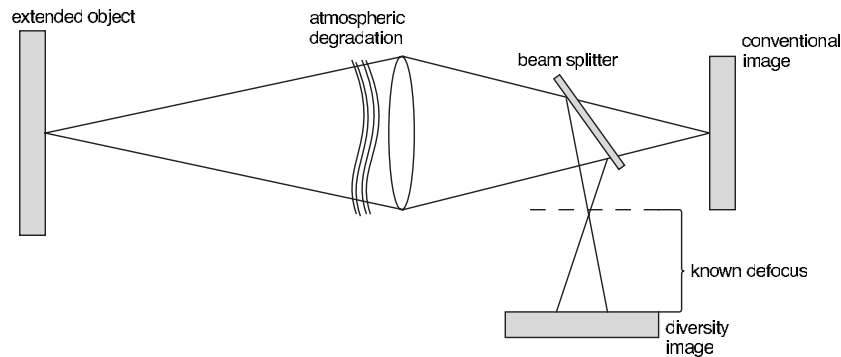


Fig. 1. Traditional phase-diversity imaging strategy.

phase-retrieval in a traditional incoherent imaging system was investigated in [9] using the Cramér–Rao lower bound. In this work we also use the Cramér–Rao lower bound for phase-diverse phase-retrieval as a guide in choosing the amount of defocus to introduce into the system.

For simplicity of presentation, all optical-system elements are assumed to be ideal and all polarimetric measurements are assumed to be perfectly registered.

The organization of this paper is as follows. Section 2 presents the mathematical framework of joint estimation of object and aberrations from polarimetric measurements. Section 3 formulates a reduced-parameter search strategy. Section 4 explores joint estimation numerically with both quadratic and edge-preserving regularization. Sections 5 and 6 give results and concluding remarks.

2. MATHEMATICAL FRAMEWORK

A. Stokes-Vector Imaging

The optical intensity, Γ , at a single point in an imaging system with a linear polarizer in the optical path having pass axis oriented at angle θ to the reference axis can be expressed in terms of the Stokes vector,

$$\Gamma(\theta) = \frac{1}{2}[S_0 + S_1 \cos(2\theta) + S_2 \sin(2\theta)]. \quad (1)$$

An imaging polarimeter has multiple channels, each with a different polarization angle. For J measurements (channels) at polarization angles $\theta_1, \dots, \theta_J$, Eq. (1) becomes a system of J equations. In matrix form the system is

$$\begin{bmatrix} \Gamma(\theta_1) \\ \vdots \\ \Gamma(\theta_J) \end{bmatrix} = \frac{1}{2} \begin{bmatrix} 1 & \cos(2\theta_1) & \sin(2\theta_1) \\ \vdots & \vdots & \vdots \\ 1 & \cos(2\theta_J) & \sin(2\theta_J) \end{bmatrix} \begin{bmatrix} S_0 \\ S_1 \\ S_2 \end{bmatrix}. \quad (2)$$

When $\Gamma(\theta_j)$, S_0 , S_1 , and S_2 are images each of size $N \times M$, Eq. (2) can be configured lexicographically to become

$$\mathbf{\Gamma} = (\mathbf{T}_{J \times 3} \otimes \mathbf{I}_{n_p}) \mathbf{S}, \quad n_p = NM, \quad (3)$$

where $\mathbf{S} = (S_0, S_1, S_2)$ is a $3n_p \times 1$ column vector, \mathbf{I}_{n_p} is the $n_p \times n_p$ identity matrix, \otimes is the Kronecker product, $\mathbf{T}_{J \times 3}$ is the matrix in Eq.(2), and $\mathbf{\Gamma}$ is a $Jn_p \times 1$ column vector. The conventional estimate of the Stokes images, $\hat{\mathbf{S}}_{\text{conv}}$, is formed by using the pseudo-inverse of $\mathbf{T}_{J \times 3}$ [5]:

$$\hat{\mathbf{S}}_{\text{conv}} = \{[(\mathbf{T}'_{J \times 3} \mathbf{T}_{J \times 3})^{-1} \mathbf{T}'_{J \times 3}] \otimes \mathbf{I}_{n_p}\} \mathbf{\Gamma}, \quad (4)$$

where “ $'$ ” denotes conjugate transpose. The matrix inverse in Eq.(4) is guaranteed to exist if $J \geq 3$ and the θ_j are chosen so that $\mathbf{T}_{J \times 3}$ has linearly independent columns. In words, in Eq. (4) the $J \times 3$ system of equations in matrix (2) is solved by least-squares at each voxel independently.

B. Forward-Imaging Model

The model (3) ignores measurement blur and noise. A more complete discrete–discrete forward model for an incoherent imaging system that accounts for space-invariant optical blur and additive noise can be represented by 2D discrete convolution:

$$y_j(n, m) = b_j(n, m) ** \Gamma_j(n, m) + \epsilon_j(n, m) \quad (5)$$

$$n = 1, \dots, N, m = 1, \dots, M,$$

where $y_j(n, m)$ are the data for the j th channel, $b_j(n, m)$ denotes the incoherent point-spread function (PSF) associated with the j th channel, $\Gamma_j(n, m)$ is the j th channel ideal intensity image, $**$ denotes 2D convolution, and $\epsilon_j(n, m)$ is additive noise. A matrix-vector representation of (5) is

$$\mathbf{y}_j = \mathbf{B}_j [(\mathbf{T}_{J \times 3})_j \otimes \mathbf{I}_{n_p}] \mathbf{S} + \boldsymbol{\epsilon}_j, \quad j = 1, \dots, J, \quad (6)$$

where \mathbf{B}_j denotes a $n_p \times n_p$ Toeplitz matrix whose entries depend on $b_j(n, m)$, $(\mathbf{T}_{J \times 3})_j$ denotes the j th row of $\mathbf{T}_{J \times 3}$, and $\boldsymbol{\epsilon}_j$ is an additive noise vector. Stacking J channels [each given by Eq.(6)] yields

$$\mathbf{y} = \mathbf{B} (\mathbf{T}_{J \times 3} \otimes \mathbf{I}_{n_p}) \mathbf{S} + \boldsymbol{\epsilon}, \quad (7)$$

where $\mathbf{y} \triangleq (\mathbf{y}_1, \dots, \mathbf{y}_J)$, $\mathbf{B} \triangleq \text{diag}\{\mathbf{B}_j\}$ is a block diagonal matrix with the single-channel blur matrices on the diagonal, and $\boldsymbol{\epsilon} \triangleq (\boldsymbol{\epsilon}_1, \dots, \boldsymbol{\epsilon}_J)$.

C. Point-Spread-Function Parameterization

Ideally the matrices \mathbf{B}_j (or equivalently the PSFs $b_j(n, m)$) would correspond to diffraction-limited PSFs. In practice the PSF is often degraded by known or unknown aberrations. In the presence of aberrations the generalized pupil function (which is also the coherent transfer function) for the system can be written

$$H(u, v) = A(u, v) \exp[iW(u, v)], \quad (8)$$

where $A(u, v)$ is a binary aperture function, $W(u, v)$ is a phase-aberration function, and (u, v) are frequency domain coordinates [10]. Aberrations in an optical system can be represented using a suitable basis set $\{\phi_k(u, v)\}$, such as Zernike polynomials [11]. Representing $W(u, v)$ in the basis $\{\phi_k(u, v)\}$ parameterizes the generalized pupil function:

$$H(u, v; \boldsymbol{\alpha}) = A(u, v) \exp \left[i \sum_{k=1}^K \alpha_k \phi_k(u, v) \right],$$

where $\boldsymbol{\alpha} = (\alpha_1, \dots, \alpha_K)$. (9)

Visible regime polarimeter configurations, such as division-of-focal-plane and division-of-amplitude, simultaneously acquire all of the polarimetric channels and so are exposed to identical optical aberrations, i.e., $W(u, v)$ is the same for each channel.

D. Phase Diversity

To aid in the estimation of aberrations, we propose to introduce phase diversity, typically by different amounts in the different polarimetric channels. Figures 2 and 3 show two possible polarimetric-phase-diverse imaging strategies. If the phase diversity function in channel j is denoted $\varphi_j(u, v)$, then the generalized pupil function for the j th channel can be written

$$H_j(u, v; \boldsymbol{\alpha}, \varphi_j) = A(u, v) \exp \left\{ i \left[\sum_{k=1}^K \alpha_k \phi_k(u, v) + \varphi_j(u, v) \right] \right\}. \quad (10)$$

The corresponding incoherent PSF, $h_j(x, y)$, and the optical transfer function, $\mathcal{H}_j(u, v)$, can be written in terms of the generalized pupil function:

$$h_j(x, y; \boldsymbol{\alpha}, \varphi_j) = c |\mathcal{F}^{-1}[H_j(u, v; \boldsymbol{\alpha}, \varphi_j)]|^2, \quad (11)$$

$$\mathcal{H}_j(u, v; \boldsymbol{\alpha}, \varphi_j) = c \mathcal{F} [|\mathcal{F}^{-1}[H_j(u, v; \boldsymbol{\alpha}, \varphi_j)]|^2], \quad (12)$$

where $\mathcal{F}[\cdot]$ is the Fourier transform and c is a constant that normalizes the PSF to unit volume [9]. The modeled system PSF, $b_j(n, m)$, and optical transfer function consist of samples of $h_j(x, y; \boldsymbol{\alpha}, \varphi_j)$ and $\mathcal{H}_j(u, v; \boldsymbol{\alpha}, \varphi_j)$ at the Nyquist sampling rate [10], respectively. Consequently, each blur matrix, \mathbf{B}_j , is parameterized by the vector $\boldsymbol{\alpha}$. For analysis and implementation we assume periodic boundary conditions on the object so that the blur matrices, $\mathbf{B}_j(\boldsymbol{\alpha})$, are circulant and thus diagonalized by a 2D discrete Fourier transform (DFT) matrix:

$$\mathbf{B}_j(\boldsymbol{\alpha}) = \mathbf{Q} \boldsymbol{\Omega}_j(\boldsymbol{\alpha}) \mathbf{Q}', \quad (13)$$

where \mathbf{Q} is a 2D unitary DFT matrix and $\boldsymbol{\Omega}_j(\boldsymbol{\alpha})$ is a diagonal matrix whose entries are the DFT coefficients of the first column of $\mathbf{B}_j(\boldsymbol{\alpha})$.

3. ALGORITHMS FOR JOINT ESTIMATION OF STOKES IMAGES AND ABERRATIONS

This section describes novel algorithms for estimating \mathbf{S} and $\boldsymbol{\alpha}$ jointly under the model (7). Under an additive Gaussian noise model $\boldsymbol{\varepsilon}_j \sim N(\mathbf{0}, \sigma^2 \mathbf{I}_{n_p})$ for $j=1, \dots, J$, the log-likelihood function for both the object \mathbf{S} and aberration parameters $\boldsymbol{\alpha}$ is

$$L(\mathbf{S}, \boldsymbol{\alpha}) = -\frac{1}{2\sigma^2} \|\mathbf{y} - \mathbf{B}(\boldsymbol{\alpha})(\mathbf{T}_{J \times 3} \otimes \mathbf{I}_{n_p})\mathbf{S}\|^2. \quad (14)$$

Conventional maximum-likelihood estimation is ineffective for this problem because $\mathbf{B}(\boldsymbol{\alpha})$ is ill-conditioned. Therefore we focus on penalized-likelihood estimators of the form

$$(\hat{\mathbf{S}}, \hat{\boldsymbol{\alpha}}) = \underset{(\mathbf{S}, \boldsymbol{\alpha})}{\operatorname{argmin}} \{-L(\mathbf{S}, \boldsymbol{\alpha}) + R(\mathbf{S})\} \triangleq \underset{(\mathbf{S}, \boldsymbol{\alpha})}{\operatorname{argmin}} \Psi(\mathbf{S}, \boldsymbol{\alpha}), \quad (15)$$

where $R(\mathbf{S})$ is a regularization term that penalizes an object, \mathbf{S} , according to how much it departs from our as-

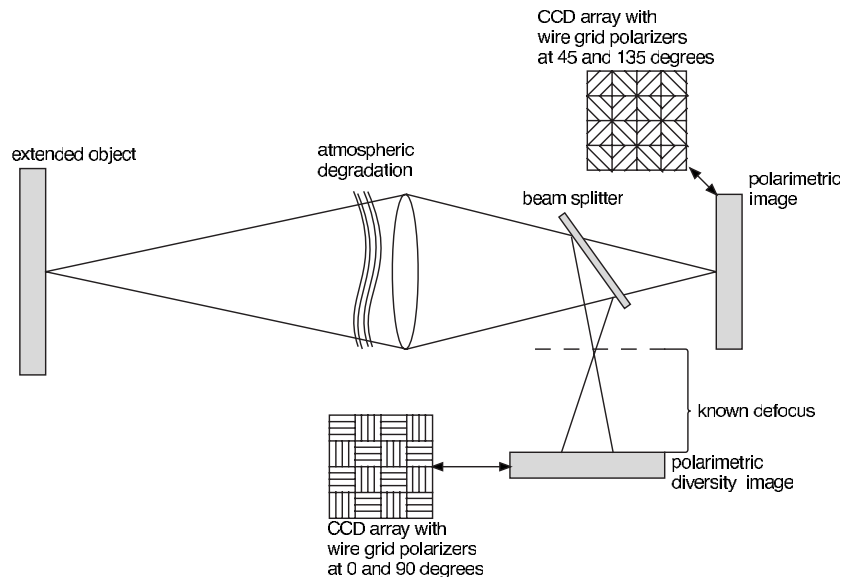


Fig. 2. Polarimetric phase-diversity strategy utilizing the division-of-focal-plane technique.

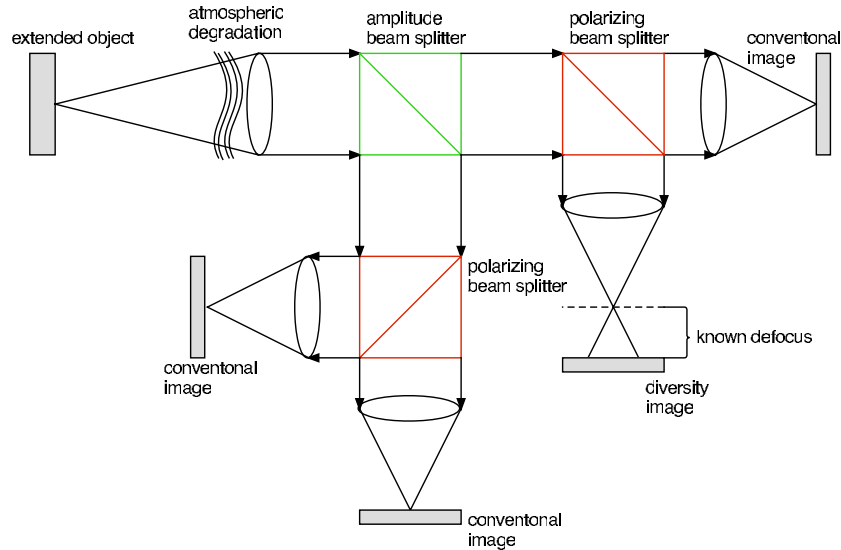


Fig. 3. (Color online) Polarimetric phase-diversity strategy utilizing the division-of-amplitude technique.

assumptions about the image properties [12]. In remote sensing α is often a nuisance parameter. However, in an adaptive-optics system with aberration correction capability, α is a parameter of interest. Depending on the task at hand, either α or \mathbf{S} or both can be parameters of interest. The choice of regularization penalty will in general depend on the task, i.e., which parameters are of interest and which are nuisance.

A. α as a Nuisance Parameter

When the Stokes images, \mathbf{S} , are of primary interest, then α is a nuisance parameter, and a regularization function that reflects *a priori* knowledge about the object should be chosen. Stokes images (S_1, S_2) typically have sharp edges due to man-made objects having stronger polarimetric signatures than naturally occurring objects. To recover as much polarization information as possible the regularization function, $R(\mathbf{S})$, should preserve edges. Since quadratic regularization tends to wash out edges and smooth noise we explore edge-preserving regularization using a hyperbolic potential function $\psi(t; \delta) = \delta^2(\sqrt{1 + (t/\delta)^2} - 1)$. For fixed δ this function is approximately quadratic for values of $t < \delta$ and approximately linear for $t > \delta$. This behavior will tend to smooth noise and preserve edges. Specifically, we choose $R(\mathbf{S})$ to be

$$R(\mathbf{S}) = \sum_{l=0}^2 \sum_{k=1}^{2n_p} \beta_l \psi([\mathbf{C}\mathbf{S}]_k; \delta_l), \tag{16}$$

where \mathbf{C} is a 2D finite-differencing matrix (horizontal and vertical differences). The estimator (15) is then

$$\begin{aligned} (\hat{\mathbf{S}}, \hat{\alpha}) = \operatorname{argmin}_{(\mathbf{S}, \alpha)} & \frac{1}{2\sigma^2} \|\mathbf{y} - \mathbf{B}(\alpha)(\mathbf{T}_{J \times 3} \otimes \mathbf{I}_{n_p})\mathbf{S}\|^2 \\ & + \sum_{l=0}^2 \sum_{k=1}^{2n_p} \beta_l \psi([\mathbf{C}\mathbf{S}]_k; \delta_l). \end{aligned} \tag{17}$$

B. α as a Parameter of Interest: Reduced Parameter Search Strategy

In [13] it was shown that, for a two channel phase-diversity system under an additive Gaussian noise model, the estimate of the object being imaged could be expressed in terms of the system aberration parameters. This result was generalized in [8] for phase-diverse imaging with an arbitrary number of channels. A similar procedure can be used to derive a closed-form expression for the Stokes images in terms of system aberrations for polarimetric phase-diverse imaging. Deriving this closed-form expression requires the use of a quadratic regularizer that can be diagonalized by the DFT as in Eq. (13). We focus on quadratic regularizers of the form

$$R(\mathbf{S}) = \frac{1}{2} \|(\sqrt{\boldsymbol{\beta}_3} \otimes \mathbf{C})\mathbf{S}\|^2, \tag{18}$$

where $\sqrt{\boldsymbol{\beta}_3} \triangleq \operatorname{diag}\{\sqrt{\beta_0}, \sqrt{\beta_1}, \sqrt{\beta_2}\}$, and $\beta_i > 0, i = 0, 1, 2$. Using this regularization function Eq. (15) becomes

$$\begin{aligned} (\hat{\mathbf{S}}, \hat{\alpha}) = \operatorname{argmin}_{(\mathbf{S}, \alpha)} & \frac{1}{2\sigma^2} \|\mathbf{y} - \mathbf{B}(\alpha)(\mathbf{T}_{J \times 3} \otimes \mathbf{I}_{n_p})\mathbf{S}\|^2 \\ & + \frac{1}{2} \|(\sqrt{\boldsymbol{\beta}_3} \otimes \mathbf{C})\mathbf{S}\|^2. \end{aligned} \tag{19}$$

For a fixed aberration vector, α , Eq. (19) is convex in \mathbf{S} and the column gradient of $\hat{\mathbf{S}}$ satisfies the stationary point condition $\nabla_{\mathbf{S}} \Psi(\mathbf{S}; \alpha) = \mathbf{0}$, where Ψ is defined in Eq. (15), which leads to

$$\begin{aligned} \mathbf{S}(\alpha) = & [(\mathbf{T}'_{J \times 3} \otimes \mathbf{I}_{n_p})\mathbf{B}(\alpha)' \mathbf{B}(\alpha)(\mathbf{T}_{J \times 3} \otimes \mathbf{I}_{n_p}) + \sigma^2 \boldsymbol{\beta}_3 \otimes \mathbf{C}'\mathbf{C}]^{-1} \\ & \times (\mathbf{T}'_{J \times 3} \otimes \mathbf{I}_{n_p})\mathbf{B}(\alpha)' \mathbf{y}. \end{aligned} \tag{20}$$

The matrix inverse in Eq. (20) is guaranteed to exist provided the intersection of the null spaces of the component matrices is the zero vector. Because \mathbf{C} is a first-order finite differencing matrix, the only nonzero vectors in its null space are of the form $\gamma \mathbf{1}$ where $\gamma \in \mathbb{R}$ and $\mathbf{1}$ is the

$n_p \times 1$ vector of ones. Therefore, nonzero vectors in the null space of $(\sqrt{\beta_3} \otimes \mathbf{C})$ are of the form $\mathbf{v} = (\gamma_1 \mathbf{1}, \gamma_2 \mathbf{1}, \gamma_3 \mathbf{1})$, where $\gamma_1, \gamma_2, \gamma_3$ are not all simultaneously zero. It remains to show that \mathbf{v} is not in the null space of $(\mathbf{T}'_{J \times 3} \otimes \mathbf{I}_{n_p}) \mathbf{B}(\boldsymbol{\alpha}) \mathbf{B}(\boldsymbol{\alpha}) (\mathbf{T}_{J \times 3} \otimes \mathbf{I}_{n_p})$. Now,

$$(\mathbf{T}'_{J \times 3} \otimes \mathbf{I}_{n_p}) \mathbf{B}(\boldsymbol{\alpha}) \mathbf{B}(\boldsymbol{\alpha}) (\mathbf{T}_{J \times 3} \otimes \mathbf{I}_{n_p}) \mathbf{v} = (\mathbf{T}'_{J \times 3} \otimes \mathbf{I}_{n_p}) \mathbf{Q} \Omega(\boldsymbol{\alpha})' \Omega(\boldsymbol{\alpha}) \underbrace{\mathbf{Q}' (\mathbf{T}_{J \times 3} \otimes \mathbf{I}_{n_p}) \mathbf{v}}_{\mathbf{u}},$$

where the circulant approximation has been used. Observe that \mathbf{u} is nonzero only in the DC components. Recall that the optical transfer function, $\Omega_j(\boldsymbol{\alpha})$, for space-invariant blur conserves energy and thus does not alter DC components. Thus,

$$(\mathbf{T}'_{J \times 3} \otimes \mathbf{I}_{n_p}) \mathbf{Q} \Omega(\boldsymbol{\alpha})' \Omega(\boldsymbol{\alpha}) \mathbf{u} = (\mathbf{T}'_{J \times 3} \otimes \mathbf{I}_{n_p}) \check{\mathbf{v}} \neq \mathbf{0},$$

since $\check{\mathbf{v}}$ is a nonzero constant vector.

Substitution of Eq. (20) into Eq. (19) yields an “aberration only” objective function,

$$\hat{\boldsymbol{\alpha}} = \underset{\boldsymbol{\alpha}}{\operatorname{argmin}} \left\{ \frac{1}{2\sigma^2} \|\mathbf{y} - \mathbf{B}(\boldsymbol{\alpha}) (\mathbf{T}_{J \times 3} \otimes \mathbf{I}_{n_p}) \mathbf{S}(\boldsymbol{\alpha})\|^2 + \frac{1}{2} \|(\sqrt{\beta_3} \otimes \mathbf{C}) \mathbf{S}(\boldsymbol{\alpha})\|^2 \right\}. \quad (21)$$

The estimate in Eq. (21) is a *joint* estimate of object and aberrations, that is, minimization over $\boldsymbol{\alpha}$ *implicitly* minimizes over \mathbf{S} . Once $\boldsymbol{\alpha}$ has been estimated the object estimate is given by Eq. (20).

4. SIMULATION EXPERIMENTS

We performed simulation experiments to evaluate joint estimation with both edge-preserving regularization (17) and quadratic regularization (21). Two situations were considered: (1) the object parameters are of interest and the aberration parameters are nuisance parameters, and (2) the aberration parameters are of interest and the object parameters are nuisance parameters. Because of the significant computational savings afforded by Eq. (21) we evaluated it with distinct regularization tuned for each object and aberrations. For comparison we also evaluated the conventional estimate (4) using the same data without phase diversity. For ground truth, we used polarimetric images collected using a division-focal-plane polarimeter by General Dynamics Advanced Information Systems, Ypsilanti, Michigan. The linear polarizer pass axes were oriented at $\{0^\circ, 45^\circ, 90^\circ, 135^\circ\}$, and the sub-sampled polarimetric image size was $[256 \times 256]$ (sub-sampled from a $[512 \times 512]$ micropolarizer array). The imagery was then corrupted by space-invariant optical blur and additive zero-mean Gaussian noise. The optical blur was constructed using an annular pupil with a phase distortion constructed from Zernike polynomials 4–19 as defined in [11]; the phase distortion had an RMS strength of 0.2 waves. The phase of the generalized pupil function is shown in Fig. 4. We define the SNR of an image to be $20 \log_{10}(\|\bar{y}\|/\|\bar{y} - y\|)$ dB, where \bar{y} and y are the noise-free and noisy images, respectively; the experiments were

done at two SNR levels: 45 dB and 25 dB. To emulate a traditional phase-diversity configuration the defocused channels were at angles $\{0^\circ, 90^\circ\}$. In this configuration the $\{45^\circ, 135^\circ\}$ channels sum to form the conventional in-focus channel and the $\{0^\circ, 90^\circ\}$ channels sum to form the defocus channel.

To aid in selecting the amount of defocus to use in the diversity channels we assumed complete knowledge of the object in Figs. 5 and 6 (phase-diverse phase-retrieval) and computed the Cramér–Rao bound for the aberration parameters over a range of defocus values. The Fisher-information matrix is computed from the log-likelihood function in Eq. (19):

$$\mathbf{F} = \frac{1}{\sigma^2} [\nabla_{\boldsymbol{\alpha}} \boldsymbol{\mu}(\boldsymbol{\alpha})] [\nabla_{\boldsymbol{\alpha}} \boldsymbol{\mu}(\boldsymbol{\alpha})]', \quad (22)$$

where ∇ denotes the column gradient and $\boldsymbol{\mu}(\boldsymbol{\alpha}) \triangleq \mathbf{B}(\boldsymbol{\alpha}) \times (\mathbf{T}_{4 \times 3} \otimes \mathbf{I}_{n_p}) \mathbf{S}$. The Fisher-information matrix was computed and inverted for various values of defocus. Since the Zernike polynomials are orthonormal, the mean of the diagonal elements corresponds to the minimum achievable mean-squared error, \hat{W}_{MIN} , of any unbiased estimator of the degrading wavefront $W(\boldsymbol{\alpha})$. In Fig. 7 \hat{W}_{MIN} is plotted against peak-to-valley defocus. The minimum occurs when the amount of defocus is 1.8 waves peak-to-valley; we used this amount of defocus in the simulations but we note that it is not necessarily the optimal choice for joint estimation of object and aberrations. The blurry and noisy data with and without phase diversity are shown in Figs. 8 and 9.

Numerical evaluation of Eqs. (17) and (21) requires the selection of regularization “tuning” parameters; for Eq. (17) six parameters must be chosen, $(\beta_0, \beta_1, \beta_2, \delta_0, \delta_1, \delta_2)$, and for Eq. (21) three parameters must be chosen $(\beta_0, \beta_1, \beta_2)$.

When there is no phase diversity the estimator PSF, $\mathbf{I}(\boldsymbol{\alpha})$, of Eq. (19) for a fixed aberration parameter, $\boldsymbol{\alpha}$, is given by [16]

$$\mathbf{I}_k(\boldsymbol{\alpha}) = [\mathbf{B}(\boldsymbol{\alpha})' \mathbf{B}(\boldsymbol{\alpha}) + \sigma^2 \beta_k \mathbf{C}' \mathbf{C}]^{-1} \mathbf{B}(\boldsymbol{\alpha})' \mathbf{B}(\boldsymbol{\alpha}) \mathbf{e}_k, \quad (23)$$

where $\mathbf{B}(\boldsymbol{\alpha})$ is the common blur across channels, \mathbf{e}_k is a Kronecker impulse, and $k = 0, 1, 2$ indicates the Stokes im-

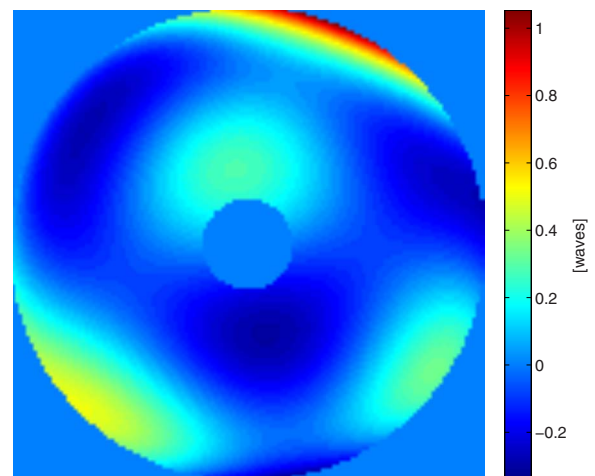


Fig. 4. (Color online) Phase of the generalized pupil function.



Fig. 5. (Color online) Image estimation results for SNR=45 dB. From left to right: object, estimate using edge-preserving regularizer, estimate using quadratic regularizer, and the conventional estimate.



Fig. 6. (Color online) Image estimation results for SNR=25 dB. From left to right: object, estimate using edge-preserving regularizer, estimate using quadratic regularizer, and the conventional estimate.

age; because the blur is space invariant Eq. (23) is independent of pixel location. The parameters $(\beta_0, \beta_1, \beta_2)$ in Eq.(17) were chosen so that in the limit that the hyperbolic potential is approximately quadratic, i.e., Eq.(17) \approx Eq.(19), the channel PSFs had full width at half maxima (FWHM) of $(r, 2r, 2r)$, where r is the FWHM of the diffraction-limited PSF in the absence of phase diversity; β_0 was calculated using

$$\hat{\beta}_0 = \underset{\beta_0}{\operatorname{argmin}} \|\operatorname{FWHM}[\mathbf{l}_0(\boldsymbol{\alpha})] - r\|^2, \quad (24)$$

and (β_1, β_2) were calculated similarly. Setting the PSF FWHM of \hat{S}_1 and \hat{S}_2 to twice that of \hat{S}_0 is reasonable because of the significantly lower SNR in the S_1 and S_2 images. We generated 20 realizations of $\boldsymbol{\alpha}$ each having RMS phase strengths of 0.2 waves over the pupil. For each aberration realization, Eq. (24) was solved numerically using a simplex search site [14], and then the final values for $(\beta_0, \beta_1, \beta_2)$ were determined by averaging over the ensemble. Once the β parameters were fixed the nonquadratic regularization parameters, $\{\delta_0, \delta_1, \delta_2\}$, were determined by a brute-force multidimensional search for the parameter combination that minimized the normalized RMS error between the Stokes object and the Stokes estimate.

For Eq. (21) there were three regularization parameters to set for each case. These parameters were determined by a brute-force multidimensional search for the parameter combination that minimized the normalized RMS error between (1) the Stokes object and the Stokes estimate, and (2) the true aberrations and the aberration estimate. The regularization parameters that were “tuned” for object estimation were 10 orders of magnitude larger than those for aberration estimation.

After the regularization parameters were set, the estimators were evaluated over a 20-realization noise ensemble for each of two SNR levels. The initial estimate in each case was formed using Eq. (4) with the phase-diverse data. Since closed form expressions for the minimizers of Eqs. (17) and (21) are not tractable they were minimized numerically. The optimization was done using the limited memory Broyden–Fletcher–Goldfarb–Shanno (L-BFGS) algorithm [15]. The minimization of Eq. (17) required preconditioning because of the different scales of the Stokes images and the aberration parameters. Samples of the Hessian matrix of Eq. (17) were calculated via finite differences and used in a diagonal preconditioner. The iterative search was stopped when the iteration, k , satisfied $(\Psi_{k+1} - \Psi_k) / \Psi_k < 10^{-10}$; this corresponded to ≈ 200 iterations for Eq. (17) and ≈ 30 iterations for Eq. (21).

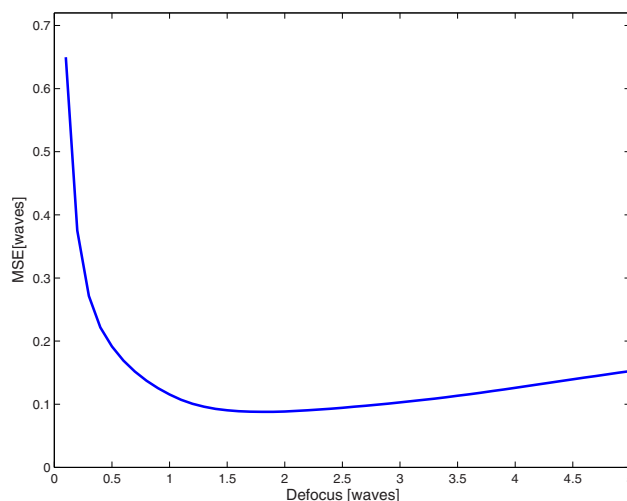


Fig. 7. (Color online) Minimum mean squared error as a function of defocus measured from peak to valley.



Fig. 8. Data for SNR=45 dB: from left to right: {0°, 45°, 90°, 135°}.

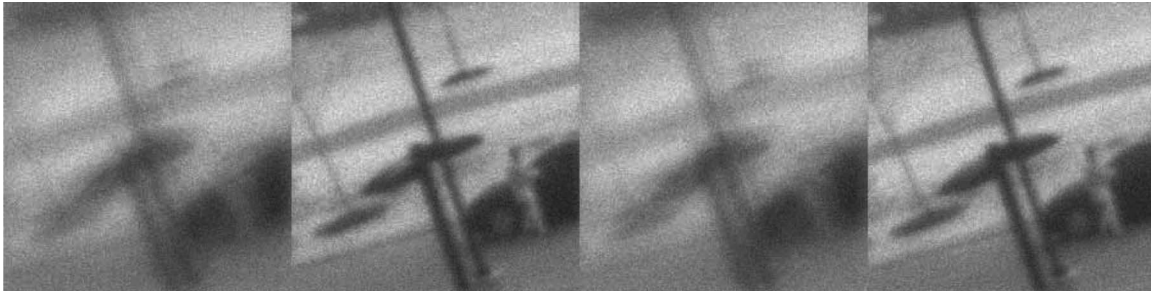


Fig. 9. Data for SNR=25 dB: from left to right: {0°, 45°, 90°, 135°}.

5. RESULTS

Tables 1 and 2 show normalized RMS estimation errors for each of Eqs. (17), (21), and (4). The reported errors are of the quantities S_0 , the total linear polarization (TPOL) $\sqrt{S_1^2 + S_2^2}$, and wavefront. There is no wavefront error to be reported for the conventional estimate, Eq. (4), so a value of N/A is listed. Also, the estimates of S_0 and TPOL are listed as N/A for Eq. (21) when the regularization was tuned for aberration estimation because the estimated images are unrecognizable. The poor object estimates in this case are due to the small values of the regularization parameter. Recall that the object estimate is given by Eq. (20), which approaches the inverse filter as $\beta \rightarrow 0$ and thus greatly amplifies noise. The aberration estimation errors for Eq. (21) when tuned for object estimation were reasonably good and are included for completeness.

Figures 8 and 9 show object estimates for SNR=45 dB and SNR=25 dB, respectively. Each estimate is displayed in red-green-blue (RGB) format with the RGB channels set as $[S_0 + 10\sqrt{S_1^2 + S_2^2}, S_0, S_0]$; in this display scheme the polarized elements of the scene are red while the unpolarized elements are in gray scale; the factor of 10 in the red channel was chosen for visual appeal. As expected, the estimates with data at a higher SNR have lower RMS errors and are more visually appealing. Figures 10 and 11 show cuts through TPOL reconstructions, at a column having an edge with large polarization content, for SNR = 45 dB and SNR=25 dB, respectively. The benefit of edge-preserving regularization is apparent in both cases but more pronounced at the 25 dB SNR level, as the quadratically regularized estimate shows significantly larger blurring across the edge.

Table 1. RMS Error Percentages for SNR=45 dB

Cost	Parameter of Interest	S_0	$\sqrt{S_1^2 + S_2^2}$	Wavefront
Edge-preserving	S	1.8% ± 0.01%	36% ± 0.3%	3.3% ± 0.2%
Quadratic	S	1.6% ± 0.01%	40% ± 0.2%	3.0% ± 0.2%
Quadratic	α	N/A	N/A	1.4% ± 0.2%
Conventional estimate	S	10% ± 0.0013%	60% ± 0.11%	N/A

Table 2. RMS Error Percentages for SNR=25 dB

Cost	Parameter of Interest	S_0	$\sqrt{S_1^2 + S_2^2}$	Wavefront
Edge-preserving	S	6.2% ± 0.2%	59% ± 1.4%	80% ± 4.6%
Quadratic	S	6.5% ± 0.02%	61% ± 1.0%	79% ± 0.36%
Quadratic	α	N/A	N/A	16% ± 7%
Conventional estimate	S	11% ± 0.011%	490% ± 1.7%	N/A

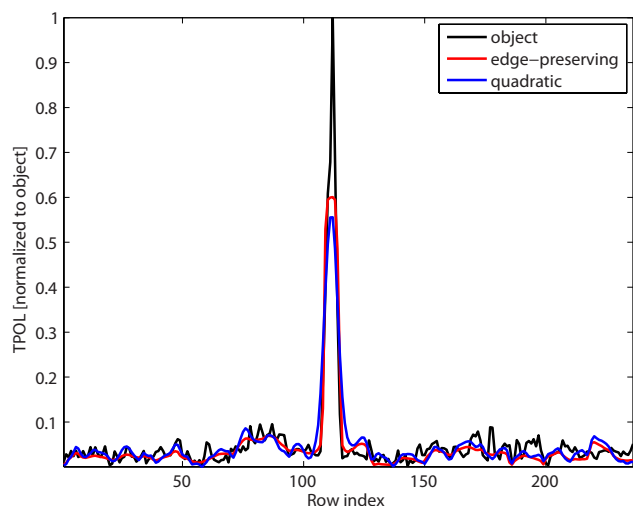


Fig. 10. (Color online) Cuts through a column of TPO_L for the object and reconstructions with edge-preserving and quadratic regularization at SNR=45 dB.

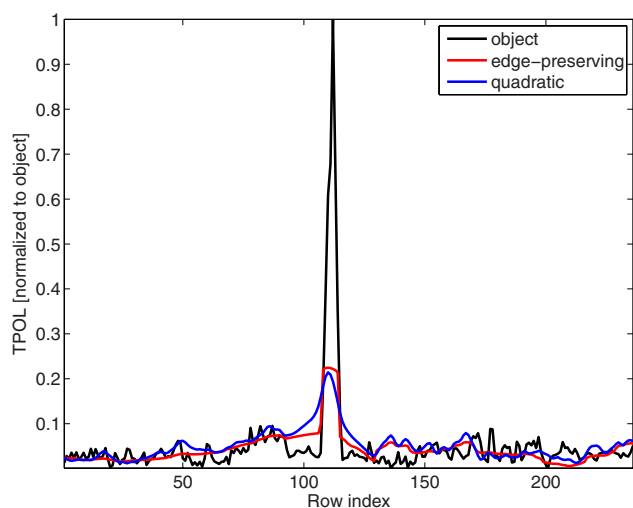


Fig. 11. (Color online) Cuts through a column of TPO_L for the object and reconstructions with edge-preserving and quadratic regularization at SNR=25 dB.

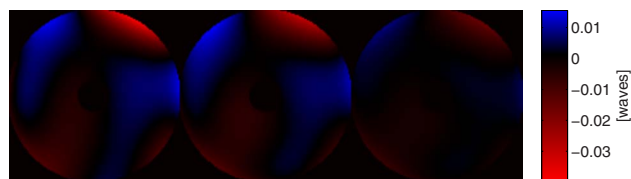


Fig. 12. (Color online) Residual wavefront errors for SNR=45 dB. From left to right: edge-preserving regularization, quadratic regularization tuned for object estimation, and quadratic regularization tuned for aberration estimation.

Figures 12 and 13 show the residual wavefronts, that is, the estimated wavefront less the true wavefront. The estimates in all cases have lower RMS errors with higher SNR data. At 45 dB SNR the wavefront estimation errors are all comparable. At 25 dB SNR the wavefront error in using Eq. (21) (when tuned for aberration estimation) is markedly lower than for Eqs. (17) and (21) (when tuned for object estimation). This significant reduction in esti-

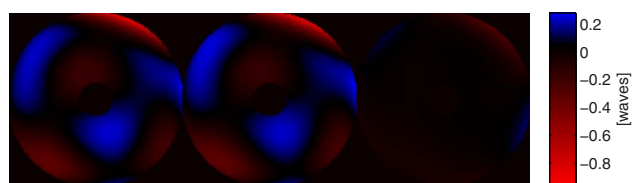


Fig. 13. (Color online) Residual wavefront errors for SNR=25 dB. From left to right: edge-preserving regularization, quadratic regularization tuned for object estimation, and quadratic regularization tuned for aberration estimation.

mation error can be attributed to the regularization being tuned for aberration estimation.

6. CONCLUSIONS AND FUTURE WORK

This paper has described two methods, Eqs. (17) and (21), for joint estimation of Stokes images and aberrations from polarimetric images with phase diversity. Estimation accuracy follows a task-based hierarchy, i.e., in a joint-estimation framework the choice of algorithm is task dependent. When the task is image restoration (aberrations are nuisance parameters) an algorithm that jointly estimates object and aberrations while incorporating *a priori* knowledge of the *object* is appropriate. However, if the aberration parameters are of interest and the object is a nuisance parameter then a reduced-parameter algorithm should be chosen.

Future work includes analyzing the bias and covariance of Eq. (19) and using those expressions to investigate how the choice of diversity channels affects estimation of Stokes images and aberrations.

REFERENCES

1. J. S. Tyo, D. L. Goldstein, D. B. Chenault, and J. A. Shaw, "Review of passive imaging polarimetry for remote sensing applications," *Appl. Opt.* **45**, 5453–5469 (2006).
2. H. Barrett and K. J. Myers, *Foundations of Image Science* (Wiley, 2004).
3. W. G. Egan, "Polarization and surface roughness," *Proc. SPIE* **3426**, 144–152 (1998).
4. G. C. Giakos, "Multifusion, multispectral, optical polarimetric imaging sensing principles," *IEEE Trans. Instrum. Meas.* **55**, 1628–1633 (2006).
5. R. A. Chipman, "Polarimetry," in *Handbook of Optics*, 2nd ed., M. Bass, ed. (McGraw-Hill, 1995), Vol. 2, Chap. 22.
6. E. Hecht, *Optics*, 3rd ed. (Addison-Wesley, 1998).
7. J. R. Valenzuela and J. A. Fessler, "Joint reconstruction of Stokes images from polarimetric measurements," *J. Opt. Soc. Am. A* **26**, 962–968 (2009).
8. R. Paxman, T. Schulz, and J. Fienup, "Joint estimation of object and aberrations by using phase diversity," *J. Opt. Soc. Am. A* **9**, 1072–1085 (1992).
9. D. J. Lee, M. C. Roggemann, and B. M. Welsh, "Cramer-Rao analysis of phase-diverse wave-front sensing," *J. Opt. Soc. Am. A* **16**, 1005–1015 (1999).
10. J. W. Goodman, *Fourier Optics*, 3rd ed. (Roberts, 2005).
11. R. J. Knoll, "Zernike polynomials and atmospheric turbulence," *J. Opt. Soc. Am.* **66**, 207–211 (1976).
12. H. W. Engl, M. Hanke, and A. Neubauer, *Regularization of Inverse Problems* (Kluwer Academic, 1996).
13. R. A. Gonsalves and R. Chidlaw, "Wavefront sensing by phase retrieval," *Proc. SPIE* **207**, 32–39 (1979).
14. J. A. Nelder and R. Mead, "A simplex method for function minimization," *Comput. J.* **7**, 308–313 (1965).

15. D. Liu and J. Nocedal, "On the limited memory BFGS method for large scale optimization," *Math. Program.* **45**, 503–528 (1989).
16. J. Fessler, Department of Electrical and Computer Engi-

neering, University of Michigan, 1301 Beal Avenue, Ann Arbor, Michigan 48109–2122, USA, is preparing a book to be called *Image Reconstruction: Algorithms and Analysis*.