in *Image Reconstruction and Restoration*, Proc. of the SPIE **2302-23**, San Diego, CA (July 1994)

Parameter dimension of turbulence-induced phase errors and its effects on estimation in phase diversity

Brian J. Thelen and Richard G. Paxman

Environmental Research Institute of Michigan
P.O. Box 134001, Ann Arbor, Michigan 48113-4001
(313) 994-1200 ext. 2363
internet email: thelen@erim.org

ABSTRACT

The method of phase diversity has been used in the context of incoherent imaging to estimate jointly an object that is being imaged and phase aberrations induced by atmospheric turbulence. The method requires a parametric model for the phase-aberration function. Typically, the parameters are coefficients to a finite set of basis functions. Care must be taken in selecting a parameterization that properly balances accuracy in the representation of the phase-aberration function with stability in the estimates. It is well known that over parameterization can result in unstable estimates. Thus a certain amount of model mismatch is often desirable. We derive expressions that quantify the bias and variance in object and aberration estimates as a function of parameter dimension.

1. INTRODUCTION

The method of phase diversity, first proposed by Gonsalves [1], requires the simultaneous collection of two images, as depicted in Figure 1. The first is the conventional incoherent focal-plane image that is degraded by system aberrations. System aberrations can arise from atmospheric turbulence, aberrated optical elements, or misalignments among optical elements. A simple beam splitter and a second detector array, translated along the optical axis, constitute a diversity channel that affords the collection of the second image, which is further degraded due to defocus. The defocus introduces a quadratic phase diversity. More general phase diversities can be introduced by reimaging the pupil and inserting a prescribed phase screen in the diversity channel conjugate to the pupil. The goal is to identify an object and phase aberrations that are consistent with both collected images, given the known phase diversity.

The method of phase diversity offers several advantages over other aberration-sensing methods. The optical hardware is modest. A simple beam splitter and a second detector array affords the simultaneous collection of the two images, as illustrated in Figure 1. In addition, the method relies on an external reference (the object being imaged) and is therefore less susceptible to systematic errors introduced by optical hardware. The technique also works well for extended objects or even scenes. The method of phase diversity should not be confused with curvature-sensing methods, which have been developed for point objects.

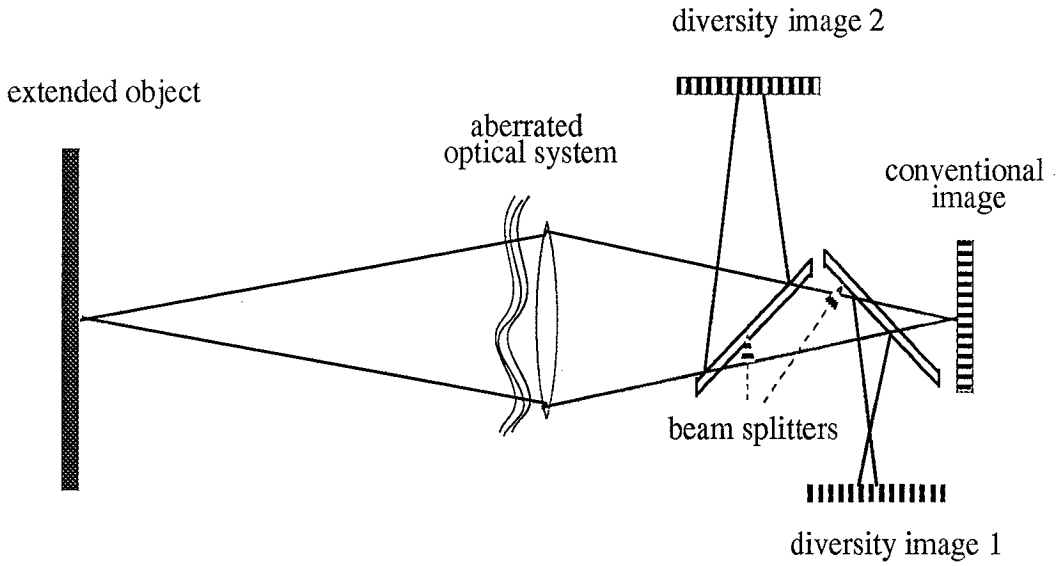

Figure 1: Optical layout for phase-diversity imaging.

Several authors have investigated the use of phase diversity in a variety of applications, including wavefront sensing, imaging with phased-array telescopes, and solar astronomy [1-9]. Researchers have also developed variations on the basic phase-diversity concept. One such variation, referred to as *phase-diverse speckle imaging*, requires the collection of a pair of short-exposure diversity images for each of several atmospheric realizations [10-13]. This novel imaging modality appears to be quite promising for use in ground-based astronomy, particularly solar astronomy. Another variation is the use of phase diversity to correct for space-variant blur [14-16].

Paxman, *et al.*, [6] developed the theory of phase diversity within an estimation-theoretic framework. In this context, the method of phase diversity is accomplished by jointly estimating the object and the system phase-aberration function. The method requires a parametric model for the phase-aberration function. Typically, the parameters are coefficients to a set of basis functions. Obviously, the number of aberration parameters to be estimated must be finite. It is natural to ask how many aberration parameters are appropriate. Care must be taken in selecting a parameterization that properly balances accuracy in the representation of the phase-aberration function with stability in the estimates. It is well known that over parameterization can result in unstable estimates (cf. [17], pg 88). Thus a certain amount of model mismatch is often desirable. In this paper we introduce an approach for assessing the effects of phase-aberration model mismatch on the joint estimation (of object and aberrations) in phase diversity. Specifically we give expressions for the expected value and covariance of the maximum-likelihood estimates as a function of the mismatch in the phase aberrations, and in turn, these expressions can be used to quantify the mean-squared error (MSE) of the object estimate, as a function of the mismatch in the phase-aberrations. We believe that such a quantification will provide valuable insight into phase-aberration model selection for the phase diversity estimation problem.

The organization of this paper is as follows. In the next section, we present a brief overview of the general model mismatch problem The general problem is that of parameter estimation from noisy data for which there is a model mismatch. Specifically, in this general problem, we suppose we have two models, I and II, which potentially describe the generation of the data, and we are using model I as the basis for the parameter estimation. However we are concerned that model II is actually more accurate and that there will be adverse effects on our estimation due to the model mismatch. As a concrete example within phase diversity, model II might correspond to a fully parameterized phase aberration model, while model I corresponds to a reduced parameterization of the phase aberrations. Assuming a maximum-likelihood estimation (MLE) framework, we then present an analytic characterization of the estimates (MLE) in the general mismatched problem. In two subsequent subsections, this is followed up with applications of the theory to two special cases which are of interest in phase diversity. In section 3, we apply the theory specifically to phase diversity to derive expressions for the expected value and covariance for the object estimate in the case of a mismatched phase aberration model and Gaussian noise. In the Gaussian noise case, these expressions can be numerically computed for moderate sized images. Covariance computations in the Poisson-noise case are more challenging. The final section summarizes the results of the paper and discusses areas for future research.

2. GENERAL THEORY

This section describes the general theory for MLE in the case of model mismatch. Specifically, we give analytic expressions for the approximate bias and covariance of the MLE in the case of model mismatch. We then apply these general results to two special cases. The first case, discussed in Subsection 2.1, is that of independent and identically distributed (iid) multivariate Gaussian random vectors with known covariance matrix, which is assumed to be a scalar multiple of the identity matrix, and a mean vector which is a function of the underlying (unknown) parameter. The second case, discussed in Subsection 2.2, is that of iid multivariate Poisson random vectors, where the components are independent with a mean vector which is a function of the underlying parameter.

For the general theory we assume we have noisy multivariate data which is modeled as iid random vectors $X_1, X_2, \ldots, X_n$. The random vectors are assumed to have a continuous distribution specified by a probability density function (pdf) or a discrete distribution specified by a probability mass function (pmf). For the sake of concreteness, we present the theory for the case of a continuous distribution (i.e., pdf's), but the theory carries over easily to the case of discrete distributions. We assume that we have two models: model I, the basis for MLE, and model II is the "true" model. Under model I, the pdf for X_1 is given by $f(x;\theta)$ where $\theta \in \Theta$ with Θ being a finite-dimensional parameter space. Under model II, X_1 has the true pdf $f(x;\theta_*)$ where $\theta_* \in \Theta_*$ is fixed, but need not be finite dimensional. For notational convenience, we use the same functional representation for all pdf's with the parameter notation dictating which is appropriate. Also, let E_θ, E_{θ_*} denote the expectation operators associated with the parameters θ and θ_* respectively.

In the matched model case, i.e., in the case were model I actually is true, the maximum likelihood estimator $\hat{\theta}_n$, defined by

$$\hat{\theta}_n = \text{argmax}_\theta \prod_{j=1}^{n} f(X_j;\theta), \tag{1}$$

has very well-known statistical properties provided that the pdf's are three-times differentiable with respect to the parameters and the partial derivatives satisfy certain technical conditions (cf. [18] or [19]). Specifically it is known to be an efficient estimator (smallest variance over the whole parameter space) which has a distribution well approximated by a multivariate Gaussian distribution with a mean vector of θ and a covariance equal to $\frac{1}{n}$ times the inverse of the Fisher information matrix. Here, the Fisher information matrix is given by

$$[J(\theta)]_{ij} = E_\theta \left(\frac{\partial L_1(\theta)}{\partial \theta_i} \frac{\partial L_1(\theta)}{\partial \theta_j} \right), \tag{2}$$

where $L_n(\theta)$ is the log-likelihood function, i.e.,

$$L_n(\theta) = \sum_{j=1}^{n} \log(f(X_j; \theta)). \tag{3}$$

As discussed in the Introduction, we are interested in the properties of the MLE $\hat{\theta}_n$ in the mismatched case, i.e., where model II is actually true. For this analysis, we assume that model I is used to generate the MLE's of θ and we assume that the true pdf is $f(\cdot; \theta_*)$ corresponding to model II. We also make the assumption that all pdf's have common support, i.e., the set where they are not equal to 0 is the same for all θ and θ_*. For fixed θ, and making mild assumptions on statistical model I, similar to those standardly assumed in MLE theory (cf. [19]), one can show that $\hat{\theta}_n$ converges in probability to $\theta_o \in \Theta$ as n gets large, where θ_o satisfies that

$$\theta_o = \text{argmax}_\theta E_{\theta_*}(\log(f(X_1; \theta)). \tag{4}$$

Here we are assuming that the expected value in the RHS of (4) is well-defined and finite for all θ. For the rest of this paper, this convergence is assumed to hold for the MLE. One can show that the maximization in the RHS of (4) is equivalent to determining the value of θ which minimizes the relative entropy (Kullback-Liebler distance) between the pdf $f(\cdot; \theta_*)$ and the pdf $f(\cdot; \theta)$, i.e., the pdf $f(\cdot; \theta_o)$ is that pdf in $\{f(\cdot; \theta) : \theta \in \Theta\}$ which is closest to $f(\cdot; \theta_*)$ in "relative entropy distance." Assuming that $\hat{\theta}_n$ converges in probability to θ_o, and assuming that the pdf's $f(x; \theta)$ are nice smooth functions of θ for fixed x, we have a generalization of the usual result for MLE's. Specifically it can be shown that $\hat{\theta}_n$ has a distribution well approximated by a multivariate Gaussian distribution with mean θ_o and a covariance matrix given by

$$\frac{1}{n}(J_*^2(\theta_o))^{-1} J_*^1(\theta_o)(J_*^2(\theta_o))^{-1} \tag{5}$$

where

$$\left[J_*^1(\theta)\right]_{ij} = E_{\theta_*} \left(\frac{\partial L_1(\theta)}{\partial \theta_i} \frac{\partial L_1(\theta)}{\partial \theta_j} \right) \tag{6}$$

and

$$\left[J_*^2(\theta)\right]_{ij} = -E_{\theta_*} \left(\frac{\partial^2 L_1(\theta)}{\partial \theta_i \, \partial \theta_j} \right). \tag{7}$$

Note that in the case of a matched model, the J_*^1 and J_*^2 matrices were actually equal and in fact were both equal to the Fisher information matrix. In the mismatched case, these matrices are in

general not equal, as we will see in the special cases. In the case of a 1-D parameter space Θ, the proof of the above is very similar to the result stated and proved in [19]. So far in this section, we have provided general expressions for the approximate expected value and covariance of the MLE. In the two subsequent subsections, we present expressions for the expected value and covariance in two special cases of interest for phase diversity (multivariate Gaussian and multivariate Poisson).

2.1 Application to a Gaussian model

In this subsection, we specialize the result to the case where $X_1, X_2, \ldots, X_n$ are iid M-dimensional multivariate Gaussian under both models I and II with a common covariance matrix of $\sigma^2 I$ where I is $M \times M$ identity matrix. In model I, we assume that the mean vector is a function of θ, i.e.,

$$\mu(\theta) = \begin{bmatrix} \mu_1(\theta) \\ \mu_2(\theta) \\ \vdots \\ \mu_M(\theta) \end{bmatrix}, \tag{8}$$

and under model II, we assume that the mean vector is a function of θ_*, i.e.,

$$\mu(\theta_*) = \begin{bmatrix} \mu_1(\theta_*) \\ \mu_2(\theta_*) \\ \vdots \\ \mu_M(\theta_*) \end{bmatrix}. \tag{9}$$

Again, though the functions $\mu_1, \ldots, \mu_M$ are actually different for each of the two models, for the sake of notational convenience we do not introduce extra notation to distinguish between the two. The log-likelihood function, for model I and one random vector X_1, is given by

$$L_1(\theta) = -\frac{1}{2\sigma^2} \sum_{m=1}^{M} (X_{1m} - \mu_m(\theta))^2. \tag{10}$$

By easy computations one can show that θ_o is the value of θ satisfying that it minimizes the Euclidean norm between the mean vectors $\mu(\theta)$ and $\mu(\theta_*)$, i.e.,

$$\theta_o = \text{argmin}_\theta \sum_{m=1}^{M} (\mu_m(\theta_*) - \mu_m(\theta))^2, \tag{11}$$

and $J_*^1(\theta_o)$ is given by

$$\begin{aligned} \left[J_*^1(\theta_o)\right]_{ij} &= \frac{1}{\sigma^4} \sum_{m,m'} (\mu_m(\theta_*) - \mu_m(\theta_o))(\mu_{m'}(\theta_*) - \mu_{m'}(\theta_o)) \frac{\partial \mu_m(\theta_o)}{\partial \theta_i} \frac{\partial \mu_{m'}(\theta_o)}{\partial \theta_j} \\ &\quad + \frac{1}{\sigma^2} \sum_{m=1}^{M} \frac{\partial \mu_m(\theta_o)}{\partial \theta_i} \frac{\partial \mu_m(\theta_o)}{\partial \theta_j}, \end{aligned} \tag{12}$$

and $J_*^2(\theta_o)$ is given by

$$\left[J_*^2(\theta)\right]_{ij} = \frac{1}{\sigma^2}\sum_{m=1}^{M}\left[-\frac{\partial^2\mu_m(\theta_o)}{\partial\theta_i\,\partial\theta_j}(\mu_m(\theta_*)-\mu_m(\theta_o)) + \frac{\partial\mu_m(\theta_o)}{\partial\theta_i}\frac{\partial\mu_m(\theta_o)}{\partial\theta_j}\right] \tag{13}$$

2.2 Application to a Poisson model

In this subsection, we specialize the result to the case where $X_1, X_2, \ldots, X_n$ are iid M-dimensional multivariate Poisson with independent components under both models I and II. Under model I, we assume that the mean vector is a function of θ, i.e.,

$$\lambda(\theta) = \begin{bmatrix} \lambda_1(\theta) \\ \lambda_2(\theta) \\ \vdots \\ \lambda_M(\theta) \end{bmatrix}, \tag{14}$$

and under model II, we assume the mean vector is a function of θ_*, i.e.,

$$\lambda(\theta_*) = \begin{bmatrix} \lambda_1(\theta_*) \\ \lambda_2(\theta_*) \\ \vdots \\ \lambda_M(\theta_*) \end{bmatrix}. \tag{15}$$

Again, though the functions $\lambda_1, \ldots, \lambda_M$ are actually different for each of the two models, for the sake of notational convenience we do not introduce extra notation to distinguish between the two. The log-likelihood for model I and one random vector X_1 is given by

$$L_1(\theta) = \sum_{m=1}^{M}\left(X_{1m}\log(\lambda_m(\theta)) - \lambda_m(\theta) - \log(X_{1m}!)\right) \tag{16}$$

Now one can show that θ_o is the value of θ satisfying that

$$\theta_o = \text{argmin}_\theta \sum_{m=1}^{M}\left(\lambda_m(\theta_*)\log(\lambda_m(\theta)) - \lambda_m(\theta)\right), \tag{17}$$

and $J_*^1(\theta_o)$ is given by

$$\begin{aligned}\left[J_*^1(\theta_o)\right]_{ij} &= \sum_{m,m'}\frac{\partial\lambda_m(\theta_o)}{\partial\theta_i}\frac{\partial\lambda_{m'}(\theta_o)}{\partial\theta_j}\left(\frac{\lambda_m(\theta_*)}{\lambda_m(\theta_o)}-1\right)\left(\frac{\lambda_{m'}(\theta_*)}{\lambda_{m'}(\theta_o)}-1\right) \\ &\quad + \sum_{m=1}^{M}\frac{\partial\lambda_m(\theta_o)}{\partial\theta_i}\frac{\partial\lambda_m(\theta_o)}{\partial\theta_j}\frac{\lambda_m(\theta_*)}{\lambda_m(\theta_o)}\end{aligned} \tag{18}$$

and $J_*^2(\theta_o)$ is given by

$$\left[J_*^2(\theta_o)\right]_{ij} = \sum_{m=1}^{M}\frac{\lambda_m(\theta_*)}{\lambda_m^2(\theta_o)}\frac{\partial\lambda_m(\theta_o)}{\partial\theta_i}\frac{\partial\lambda_m(\theta_o)}{\partial\theta_j} - \sum_{m=1}^{M}\left(\frac{\lambda_m(\theta_*)}{\lambda_m(\theta_o)}-1\right)\frac{\partial^2\lambda_m(\theta_o)}{\partial\theta_i\partial\theta_j} \tag{19}$$

3. APPLICATION TO PHASE DIVERSITY

In this section, we apply the general theory to the mismatched estimation problem within phase diversity. We take as our starting point the general phase diversity imaging model which is characterized by the following equation:

$$g_k(x) = \sum_{x' \in \mathcal{X}} f(x') s_k(x - x') \tag{20}$$

where f is the object array, s_k is the point-spread function (PSF) having diversity k, $g_k(x)$ is the expected k^{th} diversity image, x is a two-dimensional coordinate, and

$$\mathcal{X} = \{0, \ldots, N-1\} \times \{0, \ldots, N-1\}. \tag{21}$$

We treat the object, the PSF's, and the images as periodic arrays with period cell of size $N \times N$. The model for the k^{th} PSF is that

$$s_k(x) = |h_k(x)|^2 \tag{22}$$

where h_k is the inverse discrete Fourier transform of the generalized pupil function,

$$H_k(u) = A_k(u) \exp\{i(\phi(u) + \theta_k(u))\}, \tag{23}$$

where $A_k(u)$ is the binary aperture function, ϕ is the unknown phase-aberration function that we would like to estimate, and θ_k the known phase function associated with the k^{th} diversity image. It is typical to parameterize the phase-aberration function:

$$\phi(u) = \sum_{j=1}^{J_t} \alpha_j \phi_j(u) \tag{24}$$

where the J_t coefficients, $\alpha_1, \ldots, \alpha_{J_t}$, serve as parameters and $\{\phi_j\}_1^{J_t}$ is a convenient set of basis functions, such as discretized Zernike polynomials. With this parameterization of ϕ, the PSF depends on the parameter vector α and similarly the noiseless image values, $g_k(x)$ depend on f and α. Since we will be taking derivatives with respect to object and aberration parameters, we make this dependence explicit by writing $s_k(x;\alpha)$ and $g_k(x;f,\alpha)$. We now consider the case where the noise at each detector element is modeled as iid zero-mean Gaussian with a variance of σ^2. As discussed earlier, for a fixed finite set of basis functions, there is a question of how many to include in the model, i.e., how many aberration coefficients to estimate. One wants the dimensionality to be large enough to provide an accurate approximation, but not so many so as to cause instability of the aberration and/or object estimates. We can apply the theory developed in the previous subsection 2.1 to give analytic expressions for the effects of the mismatched phase-aberration model. Our true model in this setting is that the parameter $\theta_* \in \Theta_*$ consists of

$$\theta_* = \left(\{f_*(x) : \ x \in \mathcal{X}\}, \{\alpha_{*j}\}_1^{J_t}\right) \in \Theta_* \equiv \mathbf{R}_+^{\mathcal{X}} \times \mathbf{R}^{J_t} \tag{25}$$

and the mismatched model corresponds to

$$\theta = \left(\{f(x) : \ x \in \mathcal{X}\}, \{\alpha_j\}_1^{J}\right) \in \Theta \equiv \mathbf{R}_+^{\mathcal{X}} \times \mathbf{R}^{J_t} \tag{26}$$

where $J \leq J_t$, $\mathbf{R}$ denotes the real numbers, and $\mathbf{R}_+^{\mathcal{X}}$ denotes the space of all possible nonnegative objects. Notationally let the "true" object be denoted by f_* and the "true" aberrations be denoted by the J_t-dimensional vector α_*. For moderate to high signal-to-noise ratios, the current stochastic model is statistically equivalent to a model where one observes an iid sequence of n white noisy images, each which is multivariate Gaussian with the same mean and a variance of $n\sigma^2$. Hence in this case, our previous theory can be applied to this problem to assess effects of model mismatch on MLE.

The first step in this application is to determine the expected value for f and $\alpha^{(J)}$. Based on the results from subsection 2.1, it is easy to show that this corresponds to the values of f and $\alpha^{(J)} \equiv \{\alpha_j\}_1^J$ which minimize the sum of squared differences between the two sets of (noiseless) diversity images, i.e., the object f_o and phase parameters $\alpha_o^{(J)}$ satisfying

$$(f_o, \alpha_o^{(J)}) = \text{argmin}_{f,\alpha^{(J)}} \sum_{k=1}^{K} \sum_{x \in \mathcal{X}} |g_k(x; f_*, \alpha_*) - g_k(x; f, \alpha^{(J)})|^2. \tag{27}$$

Note that this is essentially the standard MLE estimation problem for f and aberrations $\alpha^{(J)}$, where one assumes that the data are the noiseless diversity images generated by f_* and α_* (cf. [6]). Therefore existing nonlinear optimization methods that are used for phase diversity estimates can be used here to find the expected values. From these expressions for the MLE expected value we can compute the bias for the mismatched MLE of the object. Specifically the bias is given by $f_o(u) - f_*(u)$, where f_o was determined from solving for the MLE.

The second step is to compute the covariance which depends on $f_o, \alpha_o^{(J)}$. To do this we need to compute the matrices $J_*^1(f_o, \alpha_o^{(J)})$ and $J_2^*(f_o, \alpha_o^{(J)})$ and apply the formula in (5). For notational convenience, we replace f_o by f and $\alpha_o^{(J)}$ by α for the rest of this section. The expressions for these matrices involves the first and second partial derivatives of the mean data vector, which in this case corresponds to the noiseless diversity images, with respect to the parameters. Using (20), it can be shown that

$$\begin{aligned}
&\left[J_*^1(f,\alpha)\right]_{f(u),f(u')} \\
&= \frac{1}{\sigma^4} \sum_{k,k'} \sum_{u'',u'''} [(g_k(u''; f_*, \alpha_*) - g_k(u''; f, \alpha))(g_{k'}(u'''; f_*, \alpha_*) - g_{k'}(u'''; f, \alpha)) \\
&\qquad \cdot s_k(u'' - u; \alpha) s_k(u''' - u'; \alpha)] \\
&\quad + \frac{1}{\sigma^2} \sum_{k=1}^{K} \sum_{u''} s_k(u'' - u; \alpha) s_{k'}(u'' - u'; \alpha).
\end{aligned} \tag{28}$$

$$\begin{aligned}
&\left[J_*^1(f,\alpha)\right]_{f(u),\alpha_j} \\
&= \frac{1}{\sigma^4} \sum_{k,k'} \sum_{u',u''} \Bigg[(g_k(u'; f_*, \alpha_*) - g_k(u'; f, \alpha))(g_{k'}(u''; f_*, \alpha_*) - g_{k'}(u''; f, \alpha)) \\
&\qquad \cdot s_k(u' - u; \alpha) \frac{\partial g_{k'}(u''; f, \alpha)}{\partial \alpha_j} \Bigg] \\
&\quad + \frac{1}{\sigma^2} \sum_{k=1}^{K} \sum_{u'} s_k(u' - u; \alpha) \frac{\partial g_k(u'; f, \alpha)}{\partial \alpha_j},
\end{aligned} \tag{29}$$

and

$$\begin{aligned}\left[J_*^1(f,\alpha)\right]_{\alpha_i,\alpha_j} &= \frac{1}{\sigma^4}\sum_{k,k'}\sum_{u,u'}(g_k(u;f_*,\alpha_*)-g_k(u;f,\alpha))(g_{k'}(u';f_*,\alpha_*)-g_{k'}(u';f,\alpha))\frac{\partial g_k(u;f,\alpha)}{\partial\alpha_i}\frac{\partial g_{k'}(u';f,\alpha)}{\partial\alpha_j} \\ &\quad + \frac{1}{\sigma^2}\sum_{k=1}^{K}\sum_{u}\frac{\partial g_k(u;f,\alpha)}{\partial\alpha_i}\frac{\partial g_k(u;f,\alpha)}{\partial\alpha_j} \end{aligned} \tag{30}$$

where

$$\frac{\partial g_k(u;f,\alpha)}{\partial\alpha_i} = \sum_{u'} f(u')\frac{\partial s_k(u-u';\alpha)}{\partial\alpha_i}. \tag{31}$$

Similarly the J_*^2 matrix is given by

$$\begin{aligned}\left[J_*^2(f,\alpha)\right]_{f(u),f(u')} &= \frac{1}{\sigma^2}\sum_{k=1}^{K}\sum_{u''}s_k(u''-u;\alpha)s_k(u''-u';\alpha) \\ &= \sum_{k=1}^{K}\sum_{u''}s_k(u''-(u-u');\alpha)s_k(u'';\alpha), \end{aligned} \tag{32}$$

$$\begin{aligned}\left[J_*^2(f,\alpha)\right]_{f(u),\alpha_j} &= \frac{1}{\sigma^2}\sum_{k=1}^{K}\sum_{u'}\left[-\frac{\partial s_k(u'-u;\alpha)}{\partial\alpha_j}(g_k(u';f_*,\alpha_*)-g_k(u';f,\alpha))\right. \\ &\qquad \left. + s_k(u'-u;\alpha)\frac{\partial g_k(u';\alpha,f)}{\partial\alpha_j}\right], \end{aligned} \tag{33}$$

and

$$\left[J_*^2(f,\alpha)\right]_{\alpha_i,\alpha_j} = \frac{1}{\sigma^2}\sum_{k=1}^{K}\sum_{u}\left[\frac{\partial^2 g_k(u;f,\alpha)}{\partial\alpha_i\partial\alpha_j}(g_k(u;f_*,\alpha_*)-g_k(u;f,\alpha)) + \frac{\partial g_k(u;\alpha,f)}{\partial\alpha_i}\frac{\partial g_k(u;\alpha,f)}{\partial\alpha_j}\right], \tag{34}$$

where

$$\frac{\partial^2 g_k(u;f,\alpha)}{\partial\alpha_i\partial\alpha_j} = \sum_{u'} f(u')\frac{\partial^2 s_k(u-u';\alpha)}{\partial\alpha_i\partial\alpha_j}. \tag{35}$$

The computation of the covariance matrix for the mismatched object MLE involves computing the upper (f,f) quadrant of the total covariance matrix

$$\Sigma \equiv (J_*^2(f_o,\alpha_o))^{-1}J_*^1(f_o,\theta_o)(J_*^2(f_o,\theta_o))^{-1}. \tag{36}$$

Now the matrices J_*^1 and J_*^2 are square matrices whose dimension is the sum of the object dimension plus the aberration dimension, which for typical real problems is quite large. This presents a potential problem in computing the covariance matrix in (36), due to the fact that we have to do

an inversion of the J_*^2 matrix. However if one carefully inspects the above expressions, especially (32), after lexicographic ordering the parameter $f(u)$ with respect to the two-dimensional variable u, the upper lefthand (f, f) quadrant of J_*^2 matrix is block circulant and so one can easily find the inverse of this quadrant using the FFT. If the dimension of the aberration parameterization is relatively small, one can use this and the formula for inverses of partitioned matrices (cf. pg 390-391 in [17]) to derive expressions for the inverse of J_*^2 matrix relatively easily. It only requires the multiplication of matrices and the inversion of matrices whose dimension is the dimension of the phase aberration parameterization. Thus in this case, the computation of the bias and covariance of the object MLE appears reasonable, and we are currently implementing this in software to do computations on simulated data.

Using the results in subsection 2.2, we have also derived analogous expressions for the (mismatched) MLE expected value and covariance in the phase diversity joint estimation problem for the case that Poisson noise is dominant. Again the expected value is computed by solving the MLE problem with the noiseless diversity images as data. However the nice block-circulant structure which was present in the J_*^2 matrix for the Gaussian noise case is no longer present in the Poisson noise case, so that the numerical computation of the covariance matrix appears to be considerably more difficult than was true in the Gaussian case. Alternative computational approaches in the Poisson case, which exploit other special structures of the matrices is the object of current research.

4. SUMMARY

In this paper, we have presented an approach for assessing the affects of model mismatch on joint estimation in phase diversity. We have given expressions for the expected value and the covariance of the MLE in the case of a mismatched phase aberration model and Gaussian noise. The expected value can be computed by solving a phase diversity estimation problem using noiseless diversity images as data. The covariance can be computed by inverting matrices whose dimension is that of the total dimension of all the parameters (object and aberration) and then doing matrix multiplication. In the case of moderate size objects, a brute force matrix inversion may be numerically impractical. However in the Gaussian case, invoking some special matrix structures which are present, we presented a reasonable numerical approach for computing the inverse and hence for computing the covariance matrix. Currently, we have no analogous simplification which holds in the Poisson case. This overall approach to assessing model mismatch within phase diversity carries over directly to the case of phase diverse speckle and we have worked out the details.

Areas for future research include:

(i) Application of the approach to simulated data for the Gaussian noise case. In particular, apply the methodology to quantify MLE performance as a function of different aberration parameterizations for a variety of SNR's, objects, and phase aberrations.

(ii) In the Poisson case, investigating the existence of special matrix structures (within J_*^2) which would make the numerical computation of the covariance matrix practical for moderate size objects.

ACKNOWLEDGEMENTS

This research was supported in part by the NASA Space Physics Program Office, NASA Grant No.: NAGW-4069.

REFERENCES

1. R.A. Gonsalves and R. Childlaw, "Wavefront sensing by phase retrieval," in *Applications of Digital Image Processing III*, A.G. Tescher, ed., Proc. Soc. Photo-Opt. Instrum. Eng. **207**, 32-39 (1979).
2. R.A. Gonsalves, "Phase retrieval and diversity in adaptive optics," Opt. Eng. **21**, 829-832 (1982).
3. R.G. Paxman and J.R. Fienup, "Optical misalignment sensing and image reconstruction using phase diversity," J. Opt. Soc. Am. A **5**, 914-923 (1988).
4. J.A. Högbom, "On the intensity distribution over the focal volume," in *High Spatial Resolution Solar Observations*, Proceedings of the 10th Sacramento Peak Summer Workshop, Sunspot, New Mexico (1988).
5. R.G. Paxman and S.L. Crippen, "Aberration correction for phased-array telescopes using phase diversity," in *Digital Image Synthesis and Inverse Optics*, Proc. Soc. Photo-Opt. Instrum. Eng. **1351**, 787-797 (1990).
6. 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).
7. S.R. Restaino, "Wavefront sensing and image deconvolution of solar data," Applied Optics **35**, 7442-7449 (1992).
8. M.G. Löfdahl and G.B. Scharmer, "Phase-diversity restoration of solar images," in *Real Time and Post-Facto Solar Image Correction*, 13th Sacramento Peak Summer Workshop, Sunspot NM, September 1992.
9. N. Miller and A. Ling, "Imaging with phase diversity: experimental results," in *Digital Image Recovery and Synthesis II*, Paul S. Idell, ed., Proc. Soc. Photo-Opt. Instrum. Eng. **2029-25** (1993).
10. R.G. Paxman, T.J. Schulz, and J.R. Fienup, "Phase-diverse speckle interferometry," in *Topical Meeting on Signal Recovery and Synthesis IV*, Technical Digest Series **11**, (Optical Society of America, Washington DC, 1992), New Orleans, LA, April 1992.
11. R.G. Paxman and J.H. Seldin,"Fine-resolution imaging of solar features using phase-diverse speckle imaging," in *Real Time and Post Facto Solar Image Correction*, 13th Sacramento Peak Summer Workshop, Sunspot NM, September 1992.
12. R.G. Paxman and J.H. Seldin,"Fine-resolution astronomical imaging using phase-diverse speckle," in *Workshop on Wavefront Supported Post-Facto Image Correction*, Torben Anderson, ed., Risø, Roskilde, Denmark, November 1992.

13. R.G. Paxman and J.H. Seldin, "'Fine-resolution astronomical imaging with phase-diverse speckle," in *Digital Image Recovery and Synthesis II*, Paul S. Idell, ed., Proc. Soc. Photo-Opt. Instrum. Eng. **2029-31** (1993).

14. R.G. Paxman, B.J. Thelen, and J.H. Selden, "Correction of Anisoplanatic Blur by Using Phase Diversity," in *Adaptive Optics in Astronomy* Proc. SPIE **2201-135**, Kona, Hawaii (March 1994).

15. R.A. Gonsalves, "Nonisolpanatic imaging by phase diversity," Optics Letters **19** 493 (1 April, 1994).

16. R.G. Paxman, B.J. Thelen, and J.H. Selden, "Phase-diversity correction of space-variant turbulence-induced blur," to be published in Optics Letters (15 August, 1994).

17. G. A. F. Seber, *Linear Regression Analysis*, Wiley, New York, 1977.

18. H. L. Van Trees, *Detection, Estimation, and Modulation Theory, Part I*, Wiley, New York, 1968.

19. E. Lehmann, *Theory of Point Estimation*, Wiley, New York, 1983.

20. T. M. Cover and J. A. Thomas, *Elements of Information Theory*, Wiley, New York, 1991.