Analytical solution to the phase-diversity problem for real-time wavefront sensing

Isabelle Mocœur,[1,2,*] Laurent M. Mugnier,[1] and Frédéric Cassaing[1]

[1]*Office National d'Etudes et de Recherche Aérospatiales, Optics Department, BP 72, 92322 Châtillon Cedex, France*
[2]*Centre National d'Etudes Spatiales, 18 Avenue Edouard Belin, 31401 Toulouse Cedex 4, France*
**Corresponding author: mocoeur_astro@orange.fr*

Received May 26, 2009; revised October 3, 2009; accepted October 8, 2009;
posted October 14, 2009 (Doc. ID 111870); published November 5, 2009

High-resolution optical systems require a very accurate control of the optical paths. For the measurement of aberrations on extended objects, several iterative phase-diversity algorithms have been developed, based on aberration estimation from focal-plane intensity measurements. Here we present an analytical estimator in the case of small aberrations. Under this assumption, a quadratic criterion is derived that allows us to express the solution (phase and object) under a simple analytical form. We also compare the performance of our algorithm with the iterative phase diversity, demonstrating that the analytic estimator is appropriate for closed-loop operation. © 2009 Optical Society of America

OCIS codes: 010.7350, 100.3190, 110.5100, 100.1830, 100.3020.

The image $\boldsymbol{i}$ recorded at the focal plane of an instrument is modeled by the discrete and noisy convolution of the point-spread function (PSF) $\boldsymbol{h}$ with the observed object $\boldsymbol{o}$, which defines the mathematical model of the data. Generally, the PSF is degraded by aberrations $\boldsymbol{a}=\{a_1,\cdots,a_k\}$ such as turbulence; therefore, these aberrations must be estimated in order to be corrected. The phase-diversity technique [1] uses simultaneous acquisition of the focal-plane image and at least a second image differing by a known set of aberrations $\boldsymbol{a_d}$, conventionally a small defocus. In this Letter, we consider two images, $\boldsymbol{i_1}$ and $\boldsymbol{i_2}$; the first one is acquired at the focal plane, whereas the second one is obtained in a plane defocused by a distance d. The corresponding PSFs are then given by $\boldsymbol{h_1}=\boldsymbol{h}(\boldsymbol{a})$ and $\boldsymbol{h_2}=\boldsymbol{h}(\boldsymbol{a}+\boldsymbol{a_d})$.

To retrieve both the aberrations and the object that are most compatible with the measurements, Gonsalves [1] first proposed to use a least-square approach; an extension of this method is the joint maximum *a posteriori* estimation of $(\hat{\boldsymbol{a}},\hat{\boldsymbol{o}})$ that are most compatible with the measurements by using statistical information on the data [2,3]. If the noise is assumed to be a stationary white Gaussian distribution with constant variance σ^2, the joint criterion J_{I} to be minimized can be written in the Fourier domain as

$$J_{\mathrm{I}}(\boldsymbol{o},\boldsymbol{a})=N_\nu \ln\sigma^2+\frac{1}{2\sigma^2}\sum_{\nu=1}^{N_\nu}\sum_{d=1}^{2}\left|\tilde{\boldsymbol{i}}_{\boldsymbol{d}}(\nu)-\tilde{\boldsymbol{h}}_{\boldsymbol{d}}(\boldsymbol{a},\nu)\tilde{\boldsymbol{o}}(\nu)\right|^2+\sum_{\nu=1}^{N_\nu}\frac{|\tilde{\boldsymbol{o}}(\nu)-\tilde{\boldsymbol{o}}_{\boldsymbol{m}}(\nu)|^2}{2S_o(\nu)}+\mathrm{R}(\boldsymbol{a}),\tag{1}$$

where $\tilde{}$ denotes Fourier transformation and where $\tilde{\boldsymbol{h}}$, which is the Fourier transform of $\boldsymbol{h}$, is the optical transfer function (OTF); N_ν is the number of pixels in the image; and d is the dth diversity plane (1 for focal plane, 2 for extrafocal); the two last terms can be used to introduce possible prior knowledge on the aberrations and/or on the object: S_o is the power spectral density model of $\boldsymbol{o}$, $\boldsymbol{o_m}$ is the mean object, and $\mathrm{R}(\boldsymbol{a})$ is the phase regularization term. Here we choose $\boldsymbol{o_m}=0$; we take $\mathrm{R}(\boldsymbol{a})=0$, which means that we do not regularize the estimation of the aberrations explicitly. An implicit regularization is achieved by expanding the phase on a finite, small set of Zernike polynomials and is enough to obtain good results for the considered noise levels.

Criterion J_{I} must be minimized with respect to both the object $\boldsymbol{o}$ and the aberrations $\boldsymbol{a}$. However, while the object is unknown, it can be estimated for given aberrations,

$$\hat{\tilde{\boldsymbol{o}}}(\boldsymbol{a},\nu)=\frac{\tilde{\boldsymbol{i}}_{\boldsymbol{1}}(\nu)\tilde{\boldsymbol{h}}_{\boldsymbol{1}}^{*}(\boldsymbol{a},\nu)+\tilde{\boldsymbol{i}}_{\boldsymbol{2}}(\nu)\tilde{\boldsymbol{h}}_{\boldsymbol{2}}^{*}(\boldsymbol{a},\nu)}{|\tilde{\boldsymbol{h}}_{\boldsymbol{1}}(\tilde{\boldsymbol{a}},\nu)|^2+|\tilde{\boldsymbol{h}}_{\boldsymbol{2}}(\boldsymbol{a},\nu)|^2+\dfrac{\sigma^2}{S_o(\nu)}},\tag{2}$$

where * denotes complex conjugate. We set $\sigma^2/S_o=\varepsilon$ close to 0, which means we under-regularize the inversion, because it has been shown [4] that doing so leads to a consistent estimator for the aberrations. Then, by introducing the estimated object of Eq. (2) into Eq. (1), we obtain a criterion that explicitly depends on the aberrations only,

$$J_{\mathrm{II}}(\boldsymbol{a})=\frac{1}{2\sigma^2}\sum_{\nu=1}^{N_\nu}\frac{|\tilde{\boldsymbol{i}}_{\boldsymbol{1}}(\nu)\tilde{\boldsymbol{h}}_{\boldsymbol{2}}(\boldsymbol{a},\nu)-\tilde{\boldsymbol{i}}_{\boldsymbol{2}}(\nu)\tilde{\boldsymbol{h}}_{\boldsymbol{1}}(\boldsymbol{a},\nu)|^2}{|\tilde{\boldsymbol{h}}_{\boldsymbol{1}}(\boldsymbol{a},\nu)|^2+|\tilde{\boldsymbol{h}}_{\boldsymbol{2}}(\boldsymbol{a},\nu)|^2+\varepsilon}+\mathrm{Cst}.\tag{3}$$

To derive the aberrations $\boldsymbol{a}$, J_{II} is usually minimized by an iterative gradient-based method. But although iterative estimators are optimal in terms of performance [5], they are time consuming, since each iteration costs $2N_d$ FFTs, where N_d is the number of diversity planes (here, $N_d=2$).

During the past 15 years, efforts have been made toward noniterative algorithms: first by proposing better numerical algorithms [6,7], then by modifying the criterion used to estimated the aberrations from the data [8,9]. However, none of these methods is

0146-9592/09/223487-3/$15.00

truly single iteration, since each of them requires at least two iterations to converge [8,10].

To derive an analytical solution for the aberrations, we use the small phase assumption in the two following ways [11]:

- We consider the denominator of Eq. (3) as a weighting term, at $\boldsymbol{a}=0$, for instance (or at the last estimate $\boldsymbol{a_0}$ for $\boldsymbol{a}$).
- We linearize the expression of the OTF in each diversity plane by a first-order Taylor expansion, obtaining an affine expression of $\tilde{\boldsymbol{h}}_{\boldsymbol{1}}$ and $\tilde{\boldsymbol{h}}_{\boldsymbol{2}}$ as a function of $\boldsymbol{a}$ for each frequency ν,

$$\tilde{h}_d(\boldsymbol{a},\nu)=\boldsymbol{\alpha}_d(\nu)\boldsymbol{a}+\beta_d(\nu), \tag{4}$$

where $\boldsymbol{\alpha}_d(\nu)$ is a row vector of size $k_{\max}$ and $\beta_d(\nu)$ is a scalar. The differentiation is done at $\boldsymbol{a}=0$ or at $\boldsymbol{a_0}$ for each of these OTFs. So $\beta_d(\nu)=\tilde{h}_d(\boldsymbol{a_0},\nu)$, $\boldsymbol{\alpha}_d(\nu)^T=\partial h_d/\partial a(\boldsymbol{a_0},\nu)$, where ^{T} denotes transposition.

We obtain a new expression of the criterion that is quadratic and can be written as

$$J_{\mathrm{II}}(\boldsymbol{a})=\frac{1}{2\sigma^2}\sum_{\nu=1}^{N_\nu}|\boldsymbol{A}(\nu)\boldsymbol{a}-B(\nu)|^2+\mathrm{Cst}, \tag{5}$$

with

$$\boldsymbol{A}(\nu)=\frac{\tilde{i}_2(\nu)\boldsymbol{\alpha}_1(\nu)-\tilde{i}_1(\nu)\boldsymbol{\alpha}_2(\nu)}{\sqrt{|\tilde{h}_1(0,\nu)|^2+|\tilde{h}_2(0,\nu)|^2+\varepsilon}}, \tag{6}$$

$$B(\nu)=\frac{-\tilde{i}_2(\nu)\beta_1(\nu)+\tilde{i}_1(\nu)\beta_2(\nu)}{\sqrt{|\tilde{h}_1(0,\nu)|^2+|\tilde{h}_2(0,\nu)|^2+\varepsilon}}. \tag{7}$$

We define the matrix $\boldsymbol{A}$ of size $N_\nu\times k_{\max}$ as the stack of all row vectors $\boldsymbol{A}(\nu)$ of Eq. (6). Similarly, we define vector $\boldsymbol{B}$ of size N_ν as the stack of the scalar $B(\nu)$ of Eq. (7). Equation (5) can be then rewritten as

$$J_{\mathrm{II}}(\boldsymbol{a})=\frac{1}{2\sigma^2}\|\boldsymbol{A}\boldsymbol{a}-\boldsymbol{B}\|^2+\mathrm{Cst}. \tag{8}$$

To minimize $J_{\mathrm{II}}(\boldsymbol{a})$, its gradient is derived with respect to the aberrations, which leads to a linear equation depending on $\boldsymbol{a}$. The resulting aberration vector is given by

$$\hat{\boldsymbol{a}}=[\Re(\boldsymbol{A}^H\boldsymbol{A})]^\dagger\Re(\boldsymbol{A}^H\boldsymbol{B}), \tag{9}$$

with $\Re$ the real part operator and $\dagger$ the generalized inverse of a matrix.

The resulting algorithm is much faster than the iterative one, requiring only $N_d=2$ FFTs. The inversion of the $\Re(\boldsymbol{A}^H\boldsymbol{A})$ matrix is not critical, since it is a square matrix of size $k_{\max}\ll N_\nu$.

Once the phase has been computed, the object can be then restored as well by introducing $\hat{\mathbf{a}}$ in Eq. (2).

To study the properties of the analytical estimator, we consider the specific low-order aberrations of a phased-array optical system, which are the positioning errors between the apertures, namely, the pistons and tip/tilts. To solve the inverse problem, we suppose that the instrument pupil is composed by N_T identical apertures. Each aperture n has a complex transmission $\boldsymbol{p_n}$, its phase being expanded on an orthonormalized basis, here a set of $k_{\max}$ scaled Zernike polynomials $\boldsymbol{Z_k}$,

$$\boldsymbol{p}_n(\boldsymbol{u})=\Pi(\boldsymbol{u})\exp\left[j\sum_{k=1}^{k_{\max}}a_{kn}\boldsymbol{Z_k}(\boldsymbol{u})\right], \tag{10}$$

where its modulus is described by the disk function Π,

$$\Pi(\boldsymbol{u})=\begin{cases}1 & \text{for } 0\leqslant|\boldsymbol{u}|\leqslant R_n\\ 0 & \text{elsewhere}\end{cases}. \tag{11}$$

In Eq. (10), $j^2=-1$ and a_{kn} is the rms amplitude of the kth mode over the nth subaperture. The corresponding aberration vector $\boldsymbol{a}$ is of size $N_Tk_{\max}$. As we deal only with piston ($k=1$) and tip/tilt ($k=2,3$), we will consider $k_{\max}=3$ in the following.

Two kinds of tests are reported here with two observation planes, one at $\boldsymbol{a_d}=0$ and the other with a 1 rad rms defocus. The object is an urban scene; monochromatic images of size $N_{\mathrm{pix}}=256\times256$ pixels are simulated with photon noise plus a read-out noise of 10 electrons per pixel and are sampled at the Shannon rate. The signal-to-noise ratio (SNR) is defined as the ratio between the mean value of the photon number and the noise standard deviation (per pixel).

Using three apertures in an equilateral configuration, we first consider a piston linearity test, applying at high flux (SNR=87) a 51-point ramp of $[-2\pi,+2\pi]$ on a given subaperture. One pair of the corresponding images is represented in Fig. 1.

The graph in Fig. 2 compares, for the aberrated subaperture, the piston estimated by our analytical method and the iterative algorithm with the introduced piston. First, we note that the reconstruction made with the analytic algorithm is excellent between $[-\pi/2;\pi/2]$ with an accuracy below $\lambda/60$ and a bias almost zero ($<10^{-3}$ rad) at the origin. In addition, the piston is reconstructed between ±1 rad with slope coefficients equal to 1 whatever the algorithm considered. Beyond 1 rad, where the small phase assumption is no longer valid, the bias increases rapidly. We also note that near $a=\pm\pi$, a wrapping occurs; since our imaging model is monochromatic, phase information is obtained only modulo 2π.

To test phase estimation in a case of small phase perturbation, we apply a set of random piston and tip/tilt listed in Table 1 to all subapertures (where to-

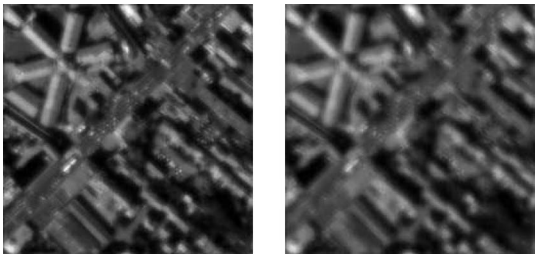

Fig. 1. Focal (left) and extrafocal (right) images obtained when a piston of 1 rad is applied on a subaperture.

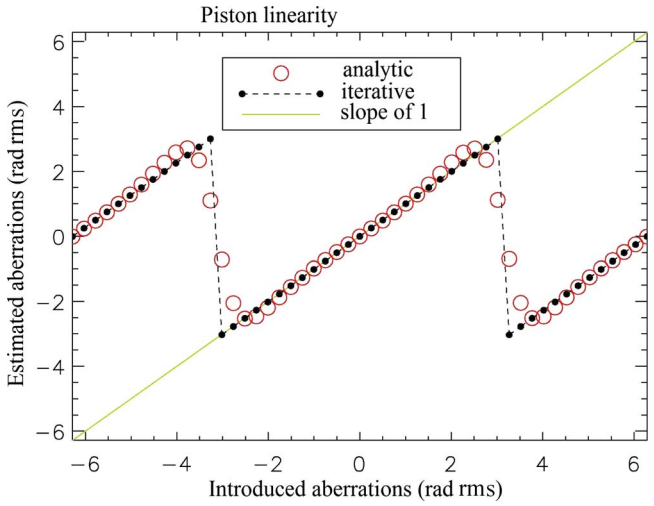

Fig. 2. (Color online) Piston linearity obtained when a piston ramp is applied on an subaperture.

tal absolute amplitude is equal to 0.53 rad rms, i.e., $\lambda/12$) for different levels of source brightness. For each data set, increasing fluxes are considered, ranging from $N_{\mathrm{bph}}=3\times10^6$ photoelectrons (phe$^-$) per image (SNR=4) to 3×10^{10} phe$^-$ (SNR=677). For each level, a data set of fifty images is simulated.

Figure 3 presents the total rms error obtained by both estimators versus the level of source brightness. Between 3×10^6 phe$^-$ and 1×10^9 phe$^-$, the two algorithms present the same behavior, following a law in $1/\sqrt{N_{\mathrm{bph}}}$. For higher flux, the error associated with the analytic algorithm remains constant around 2×10^{-3} rad. This observation is not surprising, since the affine approximation of the OTF is valid only for aberrations close to zero. The plateau we observe is then due to the inherent approximation of the linearization we made. However, the aberration estimations accuracy in the photon-noise regime remains better than 3×10^{-2} rad or $\lambda/200$ (equivalent to 10% of the total absolute amplitude introduced in piston), which is suitable for many applications.

Table 1. Aberrations Applied over the Configuration (rad rms)

	Piston	Tip	Tilt
Subaperture 1	−0.084	−0.017	0.019
Subaperture 2	0.161	−0.018	0.065
Subaperture 3	−0.055	0.019	0.092

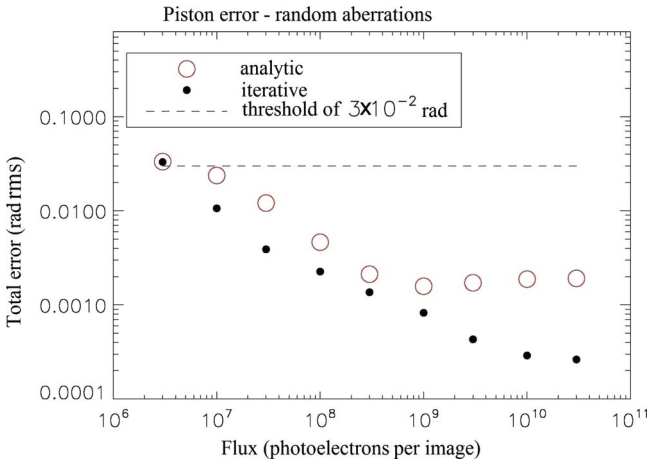

Fig. 3. (Color online) Error estimated over the subapertures when a set of random aberrations is applied.

As a conclusion, we have developed, under the small phase assumption, a noniterative focal plane algorithm for the phase diversity wavefront sensor that requires only 2 FFTs. This algorithm can typically estimate piston aberrations up to $|\pi/2|$ rad rms with an error of $\lambda/60$ and can be used for real-time correction of phase disturbances.

References

1. R. A. Gonsalves, Opt. Eng. **21**, 829 (1982).
2. O. M. Bucci, A. Capozzolo, and G. D'Elia, J. Opt. Soc. Am. A **16**, 1759 (1999).
3. A. Blanc, L. M. Mugnier, and J. Idier, J. Opt. Soc. Am. A **20**, 1035 (2003).
4. J. Idier, L. M. Mugnier, and A. Blanc, IEEE Trans. Image Process. **14**, 2107 (2005).
5. L. M. Mugnier, A. Blanc, and J. Idier, in *Advances in Imaging and Electron Physics*, P. Hawkes, ed. (Elsevier, 2006), Vol. 141, pp. 1–76.
6. M. G. Löfdahl and A. L. Duncan, Proc. SPIE **3353**, 952 (1998).
7. C. R. Vogel, T. Chan, and R. Plemmons, Proc. SPIE **3353**, 994 (1998).
8. G. B. Scharmer, in *High Resolution Solar Physics: Theory, Observations and Techniques*, T. R. Rimmele, K. S. Balasubramaniam, and R. R. Radick, eds., ASP Conferences Series (1999), Vol. 183, pp. 330–341.
9. M. G. Löfdahl and G. B. Scharmer, Proc. SPIE **4013**, 737 (2000).
10. J. Seldin, R. Paxman, V. Zarifis, L. Benson, and R. Stone, Proc. SPIE **401**, 48 (2000).
11. I. Mocœur, L. M. Mugnier, and F. Cassaing, in *21ième Colloque GRETSI* (2007).