A simple and efficient model for polychromatic focal plane wave-front sensor

S. Dandy^{*a*}, J.-F. Sauvage^{*a*}, T. Fusco^{*a*}, and L. Mugnier^{*a*} ^{*a*} ONERA, BP-72, 92322 Châtillon Cedex, France

ABSTRACT

Phase Diversity is a focal-plane technique which is chromatic by nature. The use of a monochromatic model on wide-band imaging results of an additional error function of the spectral range. We present here a second order modeling of the focal plane wave-front sensing error due to wide-band imaging and propose a first order correction by inverse problem and the first results of an end-to-end simulation for an iterative correction. Simulation results of 20 nm wave-front aberrations show that the reconstruction error decreases from 10 nm with classical focal-plane sensor to sub-nanometric error with optimal correction at $\Delta \lambda = 500$ nm.

Keywords: Focal plane sensor, Phase Diversity, inverse problem, chromatism

1. INTRODUCTION

Focal-plane sensor is widely used in astronomical observations systems nowadays. It allows accurate measurement of static or quasi-static aberrations of optical systems down to the scientific detector itself. In particular, such kind of techniques are essential for high contrast imaging applications in which nanometric accuracies are required. Phase Diversity [PD] is a powerful focal plane technique which uses several (at least 2) images linked together with known aberrations (most often defocus) and allows an unambiguous phase estimation.^{1,2} Such a technique has already been successfully applied for NAOS ^{3,4} and SPHERE.⁵ In addition to its good noise propagation properties and its ability to work on extended (not to say very extended) objects, one of the main interest of PD is its simplicity of optomechanical implementation. However, the use of wide-band imaging, in order to enhance SNR, leads to a dramatic increase of the algorithm complexity. Hence, in most cases, we use wide-band imaging with monochromatic image formation. This leads to approximations in PSF model which are known to be a major limitation as soon as the spectral resolution ($R = \lambda/\Delta\lambda$) becomes smaller than a few tens.

We propose in this paper a simple and efficient approach to overcome this limitation and to make monochromatic Phase Diversity work on wide spectral bandwidth images. The principle of Phase Diversity, its limitations and the quantification of the chromatic errors are detailed in Section 2. We propose in the third section a comprehensive modeling of the wide spectral bandwidth measurement errors using a second order expansion. Using this simplified approximation, we therefore propose a first order correction scheme which allows us to largely compensate for the chromatic error up to a sub-nanometric level even for very large bandwidths ($R \simeq 1$).

2. PHASE DIVERSITY

2.1 Principle and monochromatic model

Any focal-plane technique is based on the resolution of inverse problem which consist to estimate a corrugated phase from a focal-plane image. One of the major difficulties during phase estimation is the non-unicity of the solution.¹ This indetermination is due to the particular relation between the OTF (Optical Transfert Function) and the phase in the pupil. When we project the phase on the basis of Zernike polynomials, this indetermination leads to an ambiguity on the global sign of even radial order modes. In order to remove indetermination during phase estimation, phase diversity technique uses several images

Adaptive Optics Systems II, edited by Brent L. Ellerbroek, Michael Hart, Norbert Hubin, Peter L. Wizinowich, Proc. of SPIE Vol. 7736, 773629 · © 2010 SPIE · CCC code: 0277-786X/10/\$18 · doi: 10.1117/12.857226



FIG. 1. Phase Diversity principle

related by perfectly known phase relations. The first image i_f is acquired in the focal-plane of the optical system and the second image i_d in a slightly defocused plane with a known defocus distance d (see Figure 1).

In the following, all the images are simulated with a point source, without noise. In this particular case one may identify the image with the system PSF(point spread fonction) and are well described by the following relations :

$$h_f = |\mathbf{F}\mathbf{T}^{-1}(\mathcal{P}(r)e^{i\phi(r)})|^2 \tag{2.1}$$

$$h_d = |\mathrm{FT}^{-1}(\mathcal{P}(r)e^{i(\phi(r) + \phi_d(r))})|^2$$
(2.2)

where h_f and h_d are the respective PSF, r for the variable position in the pupil plane, $i = \sqrt{-1}$, $\phi(r)$ is the aberration function to estimated and ϕ_d is the known phase diversity. The inverse fourier transform, FT⁻¹ is defined by the relation :

$$\mathrm{FT}^{-1}\left(\mathcal{P}\right)\left(\nu\right) = \iint_{-\infty}^{-\infty} \mathcal{P}(r) e^{-2i\pi \frac{\nu r}{\lambda f}} \mathrm{d}r$$
(2.3)

In the following, the phase is well described by its decomposition on Zernike Basis.

$$\phi(r) = \sum_{k=1}^{N} a_k Z_k(r)$$
(2.4)

where $Z_k(r)$ are the Zernike polynomials with Z_1 being the piston, $Z_2 Z_3$ being the Tip / Tilt and a_k being the coefficient of the k^{th} mode.

a is the lexicographical vector gathering all the a_k coefficients. The phase estimation is then performed through the minimisation of the following Least Square Error :

$$J(\phi) = \left\| h_f - \left| \mathrm{FT}^{-1}(\mathcal{P}(r)e^{i\phi(r)}) \right|^2 \right\|^2 + \left\| h_d - \left| \mathrm{FT}^{-1}(\mathcal{P}(r)e^{i(\phi(r) + \phi_d(r))}) \right|^2 \right\|^2$$
(2.5)

using the previous image formation model. For sake of clarity, the phase estimation technique, in a monochromatic case, will be denoted "Classical Phase Diversity" [CPD], in the following.

2.2 Wide-band imaging

The appropriate formalism to describe wide-band imaging is given by :

$$I_{\Delta\lambda,\bar{\lambda}}(\nu) = \int_{\bar{\lambda}-\frac{\Delta\lambda}{2}}^{\bar{\lambda}+\frac{\Delta\lambda}{2}} \left[O_{\lambda} \star h_{\lambda}\right](\nu) \mathrm{d}\lambda$$
(2.6)

$$h(\nu;\lambda) = \left| \iint_{-\infty}^{-\infty} \mathcal{P}(r) e^{i\phi(r)} e^{-2i\pi \frac{\nu r}{\lambda f}} \mathrm{d}r \right|^2$$
(2.7)

email : sarah.dandy@onera.fr

Proc. of SPIE Vol. 7736 773629-2

with incoherent light, flat spectrum, and of bandwidth $\Delta \lambda$. We denote by $\overline{\lambda}$ the central wave-length of the imaging bandwidth. Ludovic Meynadier⁶ has shown that the optimal wavelenght for minimising reconstruction error on wide bande images is precisely $\overline{\lambda}$.

However, the CPD uses a monochromatic image formation model, in the most of the time. In the case of wide-band imaging largely used in astronomical contexts, the monochromatic image formation model used to perform the inversion is no more appropriated. Phase estimation using wide-band images therefore leads to corrugated measurements. In order to understand the impact of this limitation, the next section will expose the quantification of CPD chromatic error.

3. CLASSICAL PHASE DIVERSITY AND WIDE-BAND IMAGING

Differents solutions could be established to handle the chromatic effects. The first one is to use a polychromatic image formation model on Phase Diversity. However, this solution needs a complexe model with an arduous minimisation. That implies an important computation time. A second solution consists in adding an optical instrument, like a Wynne corrector, on the imaging way for correcting the chromatic beam. The introduction of additional optics necessarily implies additional aberrations. The third solution is to model wide-band measurement in order to correct the chromatism effects by a first order numerical correction, as explained in the following.

The different notations used in the following are gathered on Scheme 2:



FIG. 2. Scheme of our wide-band correction concept. a being the real aberrations. $\hat{a}^{\Delta\lambda}$ being the estimated aberrations by a Classical Phase Diversity

PSF h_f and h_d are computed thanks to Equation 2.7. a being the real aberrations. $\hat{a}^{\Delta\lambda}$ being the estimated aberrations by Phase Diversity.

3.1 Overall quantification of classical Phase Diversity chromatic error

The wider the band, the more different the PSf $h_{\Delta\lambda,\bar{\lambda}}(\nu)$ and $h(\nu;\lambda)$ are. Actually, wide-band imaging has the effect of spreading the PSF. Performing phase estimation on large band images with CPD leads to measurement errors. As shown in Figure 3, this error increases with bandwidth and with the number of estimated Zernike coefficients. The reconstructed error ε is defined by the following relation :

$$\varepsilon = \sqrt{\sum_{k=1}^{N} (a_k - \hat{a}_k^{\Delta\lambda})^2} = ||\boldsymbol{a} - \hat{\boldsymbol{a}}^{\Delta\lambda}||$$
(3.1)

with $a = \{a_k\}_{k=1...N}$ being the real aberrations of our corrugated phase and $\hat{a}^{\Delta\lambda} = \{\hat{a}_k^{\Delta\lambda}\}_{k=1...N}$ being the estimated aberrations by CPD.

For a 170 nm bandwidth around $\bar{\lambda} = 1050$ nm, an estimation of 100 Zernikes reaches 50% of error, while an estimation of only 20 Zernike reaches the same error for a larger bandwidth of 470 nm. In order to reach an interesting estimation (eg 10% error), the bandwidth has to be limited to 50 - 150 nm, depending on the number of estimated coefficient.

3.2 Detailed analyse of the chromatic error

A detailed study of the phase reconstruction error has been conducted in this section. The behaviour of each estimated coefficient $\hat{a}_k^{\Delta\lambda}$ has been studied wrt each introduced coefficient a_j . Figure 7 shows the coefficient $\hat{a}_k^{\Delta\lambda}$ estimated with a monochromatic model wrt true coefficient a_j for k = 11 and j = 11, and k = 17 and j = 6.



FIG. 3. Reconstruction error nm [rms] vs bandwidth. The phase error is 50 nm. The horizontal line show the relative error. Large band images are simulated with 12 images / $\Delta\lambda = 100$. Bandwidth is centered at 1050 nm, images are Shannon sampled at 950 nm



FIG. 4. Chromatic phase diversity 2^{nd} order error model, with 500 nm bandwidth. Bandwidth is centered at 1050 nm, images are Shannon sampled at 950 nm. Simulations and reconstructions are done on 20 Zernike.

Proc. of SPIE Vol. 7736 773629-4

The behaviour of the measured coefficient $\hat{a}_k^{\Delta\lambda}$ with respect to introduced coefficient a vary from a coefficient to another. On an amplitude range going up to 100 nm, this behaviour is well described by a second order model. We propose a second order model for the j^{th} wide-band estimated coefficient $\hat{a}_j^{\Delta\lambda}$:

$$\widehat{a}_{j}^{\Delta\lambda} = \underbrace{b_{j}^{\Delta\lambda}}_{\text{bias}} + \underbrace{\sum_{k} M_{jk}^{\Delta\lambda} a_{k}}_{\text{first order term}} + a^{t} \left\{ \begin{array}{c} \sum_{k} N_{j1k}^{\Delta\lambda} a_{k} \\ \vdots \\ \sum_{k} N_{j1k}^{\Delta\lambda} a_{k} \\ \vdots \\ \sum_{k} N_{jnk}^{\Delta\lambda} a_{k} \end{array} \right\} + o\left(||a||^{3}\right)$$
(3.2)

with $\hat{a}_{j}^{\Delta\lambda}$ the j^{th} wide-band estimated coefficient, $b_{j}^{\Delta\lambda}$ being the bias term on the j^{th} Zernike for bandwidth $\Delta\lambda$. $\sum_{k} M_{jk}^{\Delta\lambda} a_{k}$ being the first order term, a_{k} being the real aberrations of our corrugated phase. $M_{jk}^{\Delta\lambda}$ corresponds to the error due to the linear coupling between coefficients k and j at bandwidth $\Delta\lambda$. Finally, the second order term, means the quadratic part of error.

3.3 Bias vector

Bias vector is calibrated by simulation for each bandwidth $\Delta \lambda$. Bias term $b^{\Delta \lambda}$ is directly given by the measurement $a^{\Delta \lambda}$ when a = 0. An exemple of bias vector for bandwidth varying from 100 nm to 500 nm is shown on Figure 5. The only banwidth effect on CPD estimation is located on pure radial order Zernike polynomials (azimuthal order = 0).



FIG. 5. Bias vector on 20 Zernike for 100 nm to 500 nm bandwidth. Bandwidth is centered at 1050 nm, images are Shannon sampled at 950 nm.

The main effect of wide-band imaging with null phase is an enlargement of the PSF in the focal-plane. This is also true for the pure radial order polynomials. Hence, the PSF widennig due to the large bandwidth is interpreted by the CPD as the presence of pure radials orders abberations (defocus, spherical aberrations, ...).

3.4 First order term

The first order term $M^{\Delta\lambda}$ caracterises the gain coefficient and the coupling on phase coefficients induce by the CPD using wide-band imaging. The matrix $M^{\Delta\lambda}$ is computed for a given bandwidth $\Delta\lambda$ and could be calibrated by simulation by concatenating the Phase Diversity measurements $\hat{a}^{\Delta\lambda}$ when the true aberration is composed of the canonical basis ($a = [0 \dots a_0 \dots 0]$).

$$M_{j,*}^{\Delta\lambda} = \frac{\widehat{a}^{\Delta\lambda} - b^{\Delta\lambda}}{a_0}$$
(3.3)

The matrix $M^{\Delta\lambda}$ depends on the choice of a_0 which will be discussed further. This matrix is mainly composed of a diagonal, i.e. a measurement gain. Nevertheless, some non-diagonal terms are clearly visible, which introduces a non-negligible coupling between true and measured coefficients.



FIG. 6. Matrix $M^{\Delta\lambda}$ with 500 nm bandwidth. Bandwidth is centered at 1050 nm, images are Shannon sampled at 950 nm. The estimated coefficient values are coded in gray-scale. Max value is 1.2 and corresponds to white, min value is -0.4 and corresponds to black. Each introduced coefficient is explained in main text.

3.5 Second order term

Second order term caracterises the quadratic evolution of measurement wrt to the phase itself. The 3dimensional matrix $N^{\Delta\lambda}$ calibrated with simulation in the same way as $M^{\Delta\lambda}$. The matrix $N^{\Delta\lambda}$ includes second order coupling between estimated coefficients.

3.6 Results of modelisation

Figure 7 shows various scenarii of linear coupling, to refer to Figure 6 for the location of coupling. All next Figures are simulated with 20 Zernike starting from defocus with a bandwidth of 500 nm. As shown on the example of Figure 7, our second order model fits to phase measurement more precisely as a first order model. The first order is a good model for small aberrations, with amplitude up to $\simeq 10$ nm. The second order is a good model for a few tens of nanometers. In order to obtain an even better model, we need to use the Taylor series to the upper order. However, in our case, i.e. the XAO framework, the aberrations are less than a few tens of nanometers. The second order is therefore the model we need.

4. PSEUDO-LARGE BANDE CONCEPT FOR FIRST ORDER CORRECTION

We have shown in previous Section that chromatism effect is mainly present on pure radial order Zernike polynomials. The matrix $M^{\Delta\lambda}$ is mainly composed of a as shown in Figure 6, diagonal term. This term have an acceptable estimation by first order as we can see in Figure 7_a . The second order approximation is only required for large aberration amplitudes (typically larger than a few tens of nanometers). In a first step we are going to neglect it, in order to keep the solution simple. A more complex scheme will be presented in further work.

4.1 Inverse problem

We have validated in Section 3.2 our second order model for data formation. It is validated by simulation. This model is easy to inverse for the first order, and expresses itself in the following way :

$$\widehat{\boldsymbol{a}}^{\Delta\lambda} = \boldsymbol{b}^{\Delta\lambda} + \boldsymbol{M}^{\Delta\lambda}\boldsymbol{a} + o\left(||\boldsymbol{a}||^2\right) \tag{4.1}$$



FIG. 7. 2^{nd} order error model on various coefficients, with 500 nm bandwidth, centered at 1050 nm, images are Shannon sampled at 950 nm. Simulations and reconstructions are done on 20 Zernike modes. The matrix $M^{\Delta\lambda}$ is calibrated with a_0 equals to 10 nm and the matrix $N^{\Delta\lambda}$, with a_0 equals to 0.1 nm.



FIG. 8. Scheme of our wide-band correction concept. a being the real aberrations. $\hat{a}^{\Delta\lambda}$ being the estimated aberrations by Phase Diversity. \hat{a} being the estimated aberrations after the first order correction

It is easy to inverse beacause the matrix $M^{\Delta\lambda}$ is mainly diagonal. In our case, $M^{\Delta\lambda}$, here after called interaction matrix, is obtained by classical gaussian elimination method. Conditioning value is around 10. hence, we estimate the true aberrations by a direct matrix inversion.

$$\widehat{a} = M^{\Delta\lambda\dagger} (\widehat{a}^{\Delta\lambda} - b^{\Delta\lambda}) \tag{4.2}$$

4.2 Optimal calibration of the interaction matrix

When the inverse problem is solved, we obtain the estimated aberrations \hat{a} . However, this estimation depends on the pertinence of the interaction matrix $M^{\Delta\lambda}$, i.e. it depends on the way interaction matrix is calibrated and more precisely on the value of a_0 (amplitude for interaction matrix calibration). The next Figure 9 shows the optimized use for each interaction matrix. In Figure 9, we show the reconstruction



FIG. 9. Domain of validity of interaction matrix, $M^{\Delta\lambda}$. The Figure 9_a is simulated with a bandwidth of 200 nm and the Figure 9_b with a bandwidth of 500 nm. Bandwidth is centered at 1050 nm, images are Shannon sampled at 950 nm.

error of our pseudo-large band concept wrt the amplitude a_4 of true phase (being a pure defocus : $\phi = a_4 Z_4$). The matrix $M^{\Delta\lambda}$ used in our pseudo-large band concept is calibrated with different values of $a_0 = [5, 15, 30, 45, 60]$. Even if the true phase is composed of a pure defocus, 20 modes are estimated, and a 20X20 matrix $M^{\Delta\lambda}$ is considered.

As we can see, the reconstruction error presents a minimum when the calibration value of the interaction matrix gets close to the true phase. E.g. in the case of $\Delta \lambda = 500$ nm, and $a_0 = 60$ nm, the reconstruction error reaches it minimum for a true phase equals to 60 nm. Hence, the interaction matrix is optimal for a calibration at the true coefficient.

This is also visible on Figure 10, where the previous results are generalized to a 20 Zernike modes estimation. The wave-front error is 20 nm, the optimal value of a_0 is therefore $\frac{20 \text{ nm}}{\sqrt{20}} = 4.47$. In a realistic case, the true phase is of course unknown. The compensation of a by \hat{a} will therefore be non-optimal. In the next Section, we propose a way to mitigate this drawback.



FIG. 10. Gain of first order bandwidtqh correction to phase diversity estimation. Simulations and reconstructions are done on 20 Zernike. Bandwidth is centered at 1050 nm, images are Shannon sampled at 950 nm. Interaction matrix, $M^{\Delta\lambda}$, are calibrated on 20x20, with the optimal WFE.

4.3 Iterative "first order" correction

In order to account for non-optimal matrix $M^{\Delta\lambda}$, let us introduce the notion of iterative correction. As shown on Figure 8, the measurement is done iteratively after the phase correction. We have seen in precedent parts that a first order correction brings already an accurate estimation of the corrugated phase if we use an optimal interaction matrix. The optimal matrix is calibrated with a value of a_0 equal to the true coefficient. This value is unknown in a realistic case, we will therefore use a non optimal matrix $M^{\Delta\lambda}$, calibrated with an arbitrary value of a_0 . A good way to choose a_0 is to estimate it with the Strehl Ratio (SR) in the image :

$$a_0 = \sqrt{\frac{1 - SR}{N}} \tag{4.3}$$

with N being the number of estimated modes.

The Figure 11 shows the performance of iterative correction for optimal and non-optimal scheme. The aim is to quantify the impact of a non-optimal value a_0 on the overall performance. For this three Figures, we used the same 20 nm input phase. The first order iterative correction is done respectively with an 10 nm, 20 nm and 60 nm interaction matrix. The error at first iteration is given by the residual phase after the first correction, i.e. a CPD results.



FIG. 11. First order iterative correction. Comparison of performance with calibrated $a_0 = a)10 \text{ nm}, b)20 \text{ nm}, c)60 \text{ nm}$ between interaction matrix. Simulations and reconstructions are done on 20 Zernike for 20 nm of aberrations. Results are simulated at 100 nm, 200 nm and 500 nm bandwidth. Bandwidth is centered at 1050 nm, images are Shannon sampled at 950 nm.

Whatever the choice of a_0 , the convergence level is the same. This result is very encouraging as it shows that the effects on a non-optimal matrix $M^{\Delta\lambda}$ can be compensated by iterative correction. Meanwhile, the choice of a_0 impacts on convergence speed. The closer a_0 to true coefficient, the quicker the convergence is. The error given at first iteration therefore corresponds to the error plotted on Figure 10.

5. ACKNOWLEDGMENT

The work was partly funded by the European Comission under FP7 grant agreement N 226604 Optical Infrared Co-ordination Network for astronomy.

Proc. of SPIE Vol. 7736 773629-9

6. CONCLUSION

We studied in this paper the influence of large band imaging on focal plane aberration estimation. We proposed and validated a second order model for describing the wide band impact on phase measurement. We then validated here a first order iterative correction in order to overcome the chromatic error induced by focal plane sensing. This method is based on a first order model of chromatic focal plane measurement. The robustness of the method has been studed with respect to calibration of first order interaction matrix $M^{\Delta\lambda}$. The non-optimal correction is compensated by an iterative process. Encouraging results show a nanometric residual error even at very large bandwidth as $\Delta\lambda = 500$ nm. Next step consists in an experimental validation of this procedure. Such a validation requires an AO system capability in order to introduce calibrated aberrations, and iteratively compensate for them. This sensor will be helpful in the future development of WFS for faint natural guide star.

REFERENCES

- [1] R. A. Gonsalves, *Phase retrieval and diversity in adaptive optics*, Optical Engineering, 21 (5), pp. 829–832 (1982).
- [2] L. M. Mugnier, A. Blanc et J. Idier, *Phase Diversity : a Technique for Wave-Front Sensing and for Diffraction-Limited Imaging*, Dans Advances in Imaging and Electron Physics, sous la direction de P. Hawkes, vol. 141, chap. 1, pp. 1–76. Elsevier (2006).
- [3] A. Blanc, T. Fusco, M. Hartung, L. M. Mugnier et G. Rousset, *Calibration of NAOS and CONICA static aberrations. Application of the phase diversity technique*, Astron. Astrophys., 399, pp. 373–383 (2003).
- [4] M. Hartung, A. Blanc, T. Fusco, F. Lacombe, L. M. Mugnier, G. Rousset et R. Lenzen, *Calibration of CONICA static aberrations by phase diversity*, Dans *Instrumental Design and Performance for Optical/Infrared Ground-Based Telescopes*, sous la direction de M. Iye et A. F. M. Moorwood, vol. 4841, Bellingham, Washington, Proc. Soc. Photo-Opt. Instrum. Eng., SPIE (2002), Date conférence : août 2002, Waikoloa, Hawaii.
- [5] J.-F. Sauvage, T. Fusco, G. Rousset et C. Petit, *Calibration and Pre-Compensation of Non-Common Path Aberrations for eXtreme Adaptive Optics*, J. Opt. Soc. Am. A, 24 (8), pp. 2334–2346 (août 2007).
- [6] L. Meynadier, Analyse de surface d'onde pour le contrôle actif d'un télescope spatial, Thèse de doctorat, Université de Nice-Sophia Antipolis (1997).