

SCIENTIFIC REPORTS

OPEN

High-accuracy wavefront sensing by phase diversity technique with bisymmetric defocuses diversity phase

Peiguang Zhang¹, Chengliang Yang¹, Zihao Xu^{1,2}, Zhaoliang Cao¹, Quanquan Mu¹ & Li Xuan¹

We investigate a specific diversity phase for phase diversity (PD) phase retrieval, which possesses higher accuracy than common PD, especially for large-scale and high-frequency wavefront sensing. The commonly used PD algorithm employs the image intensities of the focused plane and one defocused plane to build the error metric. Unlike the commonly used PD, we explore a bisymmetric defocuses diversity phase, which employs the image intensities of two symmetrical defocused planes to build the error metric. This kind of diversity phase, named PD-BD (bisymmetric defocuses phase diversity), is analysed with the Cramer-Rao lower bound (CRLB). Statistically, PD-BD shows smaller CRLBs than the commonly used PD, which indicates stronger capacity of phase retrieval. Numerical simulations also verify that PD-BD has higher accuracy of phase retrieval than the commonly used PD when dealing with large-scale and high-frequency wavefront aberrations. To further affirm that PD-BD possesses higher accuracy of wavefront sensing than PD, we also perform a simple verification experiment.

Phase diversity (PD) algorithm was firstly introduced by Gonsalves and Chidlaw¹, which is a well-known algorithm of phase retrieval (PR) algorithms. This algorithm requires the image intensities of a simultaneous set of focused image and defocused image to build the error metric used in numerical optimization processing. The focused image is the target image degraded by system aberration, while the defocused image is the same blurred image, but with an additional known defocus aberration. The reconstruction of unknown wavefront aberration can be obtained by minimizing the error metric. PD algorithm could achieve near-diffraction-limited imaging and it is suitable for point target and extended target at the same time. PD algorithm is mainly used in high resolution imaging^{2,3}, and adaptive optics (AO) system^{4–10}. Other applications of PD algorithm also include measurement of laser beams^{11,12}, fresnel incoherent holography¹³ and biological microscopy imaging¹⁴.

The accuracy of phase reconstruction is closely linked with the optimization algorithm selected and the characteristics of PD physical model. Gradient-based optimization algorithms used in PD, such as conjugate-gradient algorithm (CG)^{15,16}, and Broyden-Fletcher-Goldfarb-Shanno (BFGS) algorithm^{17,18}, are easily trapped in local minimums near the initial positions. Based on a reliable global optimization algorithm, such as hybrid particle swarm global optimization algorithm¹⁹, the accuracy of PD is then determined by the behavior of PD physical model. Scholars have made some adjustments on the physical model of traditional PD to improve its accuracy, such as multiframe phase-diversity joint estimation²⁰ and phase diversity speckle (PDS) algorithm¹⁷. The commonly used diversity phase of traditional PD is a known defocus phase. A mixed diversity phase composed of astigmatism and focus has been empirically selected to improve the accuracy of PD²¹. Phase diversity technique has also been introduced to blind deconvolution image restoration^{22,23}.

The diversity phase is a phase relative to the measured system aberration, which provides a helpful constraint condition for the reconstruction of the measured system aberration. The commonly used diversity phase of PD employs the image intensities of the focused plane and one defocused plane to build the error metric. The measured system aberration within the focused plane can be seen as a zero-diversity phase. And the defocused plane contains a known defocus diversity phase relative to the measured system aberration. In this paper, we investigate

¹State Key Laboratory of Applied Optics, Changchun Institute of Optics, Fine Mechanics and Physics (CIOMP), Chinese Academy of Sciences, Changchun, Jilin, 130033, China. ²Graduate School of the Chinese Academy of Science, Beijing, 100039, China. Correspondence and requests for materials should be addressed to C.Y. (email: yclidahai@ciomp.ac.cn)

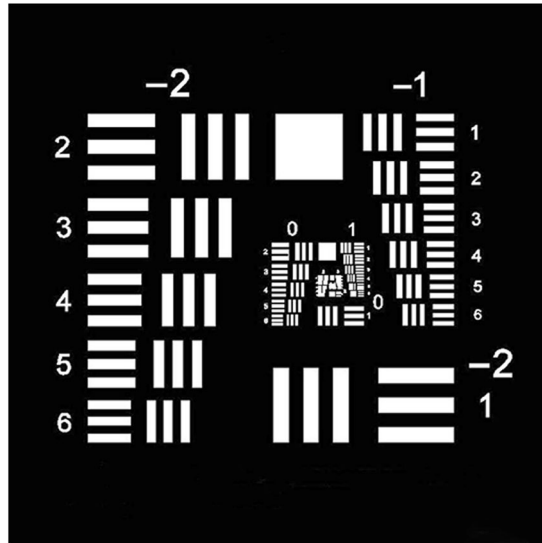


Figure 1. The pristine object used for CRLB analysis and numerical simulations.

a specific diversity phase for PD phase retrieval. The main idea is to remove the zero-diversity phase and to add a symmetric defocus diversity phase. The bisymmetric defocuses diversity phase, named PD-BD (bisymmetric defocuses phase diversity), employs the image intensities of two symmetrical defocused planes to build the error metric. Obviously, compared with traditional PD, PD-BD introduces no additional computational burden. PD-BD was first introduced by Lee *et al.* (1999)²⁴, which was analyzed by Cramer-Rao lower bound. Their work hinted that PD-BD might possess a better performance for wavefront sensing. In general, PD-BD could be seen as a special case of the multi-image PD algorithm presented by Paxman *et al.* (1992)²⁰, derived for an arbitrary number of diversity images. In our opinion, two bisymmetric defocus diversity phases can provide helpful constraint conditions for the reconstruction of the measured system aberration.

Results

CRLB analysis and numerical simulations. PD is a posteriori image-based aberration estimation technique. The smallest possible unbiased estimator variance for PD technique can be theoretically given by the Cramer-Rao lower bound (CRLB). CRLB could provide a lower bound on the mean-square error of an estimator, which can be used to theoretically evaluate the accuracy of wavefront sensing by phase diversity technique. The fundamental concept of CRLB is written symbolically as

$$\text{Var}(\tilde{a}_k) \geq [\mathbf{F}^{-1}(\mathbf{a})]_{kk}, \quad (1)$$

where \tilde{a}_k denotes an unbiased estimate of a_k . The $\text{Var}(\cdot)$ operator denotes the variance of the \tilde{a}_k . The symbol $[\mathbf{F}(\mathbf{a})]$ stands for the Fisher information matrix of \mathbf{a} . We use the same definition of CRLB as ref.²⁵, of which the CRLB is denoted as the corresponding square root that bounds the rms error of an estimator. So the CRLB of PD technique on a Zernike parameter a_k is expressed as

$$\text{CRLB}(a_k) = \sqrt{[\mathbf{F}^{-1}]_{kk}}. \quad (2)$$

The equality states that the estimated parameter vector \mathbf{a} is bounded from below by the corresponding square root of diagonal elements of a Fisher information matrix inverse. Detailed expression of the Fisher information matrix \mathbf{F} and descriptions for the CRLB of PD can be obtained in Refs^{24,25}. We suppose a high signal-to-noise ratio (SNR) for the computation of CRLB. In this paper, we will consider the CRLBs of PD and PD-BD for the simulated object under the assumption of Gaussian noise.

We will carry on the comparisons between PD and PD-BD by CRLB analysis and numerical simulations. The pristine object used for CRLB analysis and numerical simulations is shown in Fig. 1. The object here is a high signal-to-noise ratio (SNR) object which will help to approach theoretical CRLB. In fact, it is unimportant to know accurate SNR when we focus on the comparison of CRLB between PD and PD-BD. The size of the Nyquist sampled object is 200×200 pixels and the intensity levels are quantized to 16 bits. The size of simulated pupil is 100×100 pixels. For numerical simulations, the regularization parameter γ is set to be 1.0×10^{-6} . The optimization algorithm used here is a hybrid particle swarm optimization algorithm¹⁹, which can provide global search and avoid being trapped in local minimums. The amount of defocus used for PD is commonly set to be 0.5λ PV or 1.0λ PV (peak to valley). In the following, we will show that the amount of symmetrical defocuses used for PD-BD can be set as $\pm 0.5\lambda$ PV, $\pm 0.75\lambda$ PV and $\pm 1.0\lambda$ PV, which possess similar performance. We denote PD with 0.5λ PV and 1.0λ PV defocus as PD-0.5 and PD-1.0, respectively. Similarly, PD-BD with $\pm 0.5\lambda$ PV, $\pm 0.75\lambda$ PV and $\pm 1.0\lambda$ PV defocus are denoted as PD-0.5, PD-0.75 and PD-1.0, respectively. The measured wavefront aberration is denoted as ϕ_{set} . The rms value of ϕ_{set} is defined as follows:

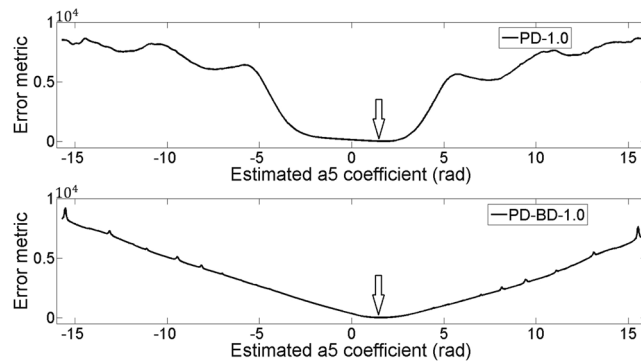


Figure 2. Error metrics plotting of PD-1.0 and PD-BD-1.0 with a astigmatism system aberration $a_5 = 1.5$ radian.

coe	Set	Reconstructed				
		PD-0.5	PD-1.0	PD-BD-0.5	PD-BD-0.75	PD-BD-1.0
a_4	0.1024	0.0299	-0.0416	0.0935	0.0936	0.0947
a_5	0.5409	0.6126	0.6801	0.5410	0.5437	0.5386
a_6	0.2027	0.2413	0.2778	0.2056	0.2058	0.2041
a_7	-0.1453	-0.1857	-0.2083	-0.1686	-0.1535	-0.1501
a_8	-0.0053	0.0195	-0.0146	-0.0216	-0.0117	-0.0104
a_9	-0.3886	-0.3858	-0.3847	-0.3717	-0.3920	-0.3917
a_{10}	0.1536	0.1674	0.1670	0.1704	0.1631	0.1617
a_{11}	0.0801	-0.1329	-0.2729	0.0485	0.0457	0.0532
a_{12}	-0.3007	-0.2697	-0.3247	-0.2938	-0.2988	-0.3024
a_{13}	-0.4227	-0.3354	-0.3456	-0.4238	-0.4184	-0.4242
a_{14}	-0.2725	-0.2514	-0.1163	-0.2672	-0.2632	-0.2587
a_{15}	0.1789	0.1818	0.1558	0.1717	0.1691	0.1725
RMS	0.0549 λ	0.0568 λ	0.0626 λ	0.0545 λ	0.0547 λ	0.0546 λ
RMS error		0.0178 λ	0.0309 λ	0.0032 λ	0.0027 λ	0.0022 λ

Table 1. Values(radians) of the coefficients used for CRLB analysis and numerical simulations.

$$RMS(\lambda) = \frac{1}{2\pi} \left[\sum_v (\phi_{set}(\mathbf{v}) - \phi_{aver}(\mathbf{v}))^2 / N_T \right]^{\frac{1}{2}}, \quad (3)$$

where v is the two-dimensional position coordinate in the pupil plane. ϕ_{aver} is the average value of ϕ_{set} , N_T is the total number of pixels within the pupil aperture. The accuracy of the reconstructed wavefront aberration ϕ is quantified with the rms phase error:

$$RMS \text{ error}(\lambda) = \frac{1}{2\pi} \left[\sum_v (\phi(\mathbf{v}) - \phi_{set}(\mathbf{v}))^2 / N_T \right]^{\frac{1}{2}}. \quad (4)$$

We first compare the error metrics of PD-1.0 and PD-BD-1.0 with a astigmatism system aberration $a_5 = 1.5$ radian. The error metrics of PD-1.0 and PD-BD-1.0 then reduce to functions of one single variable a_5 , which are explicitly plotted in Fig. 2. We can see that PD-BD-1.0 provides a steeper gradient than PD-1.0 near the minimum extreme point $a_5 = 1.5$. Steeper gradients near the global minimum position mean a easier search for the measured phase with a higher accuracy.

Moreover, we consider a randomly generated aberration listed in Table 1. We show the CRLBs plotting of PD-0.5, PD-1.0, PD-BD-0.5, PD-BD-0.75 and PD-BD-1.0 in Fig. 3. PD-BDs provides smaller CRLBs than PDs for all Zernike parameters, which indicates stronger capacity of phase retrieval. It is reasonably expected that PD-BD possesses a higher accuracy than PD. Corresponding simulation results for that aberration by PD and PD-BD are also shown in Table 1 and Fig. 4. The reconstruction accuracy of PD-0.5 is a little better than PD-1.0. Nevertheless, It can be seen that the reconstructed parameters by PD-BDs are apparently more accurate than that by PD-0.5. PD-BD-0.5, PD-BD-0.75 and PD-BD-1.0 provide similar high reconstruction accuracy. RMS errors obtained by PD-0.5 and PD-BD-0.5 are 0.0178 λ and 0.0032 λ , respectively. It is also displayed in Fig. 4 that the reconstructed phase by PD-BD-0.5 is obviously more accurate than that by PD-0.5.

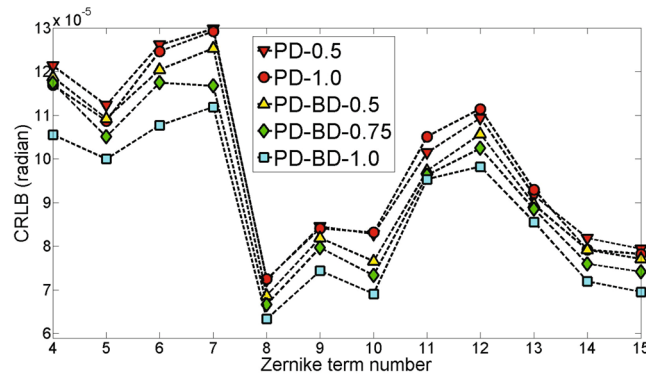


Figure 3. CRLBs plotting of PD and PD-BD on the aberration listed in Table 1.

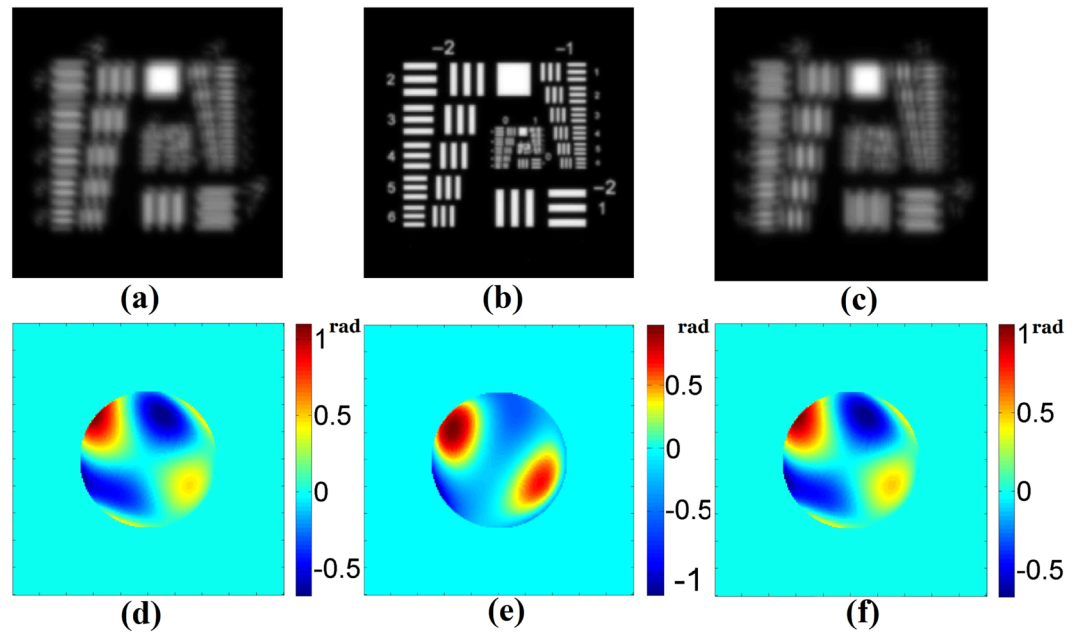


Figure 4. Simulation results by PD-0.5 and PD-BD-0.5 for the aberration listed in Table 1. (a), (b) and (c) are the minus focused image, the focal image and the plus focused image, respectively. (d) is the measured phase, (e) and (f) are the phases reconstructed by PD-0.5 and PD-BD-0.5, respectively.

To further testify the accuracy of PD-BD, we provide the statistical results based on 100 numerical cases. The measured 100 aberrations are randomly generated by setting $a_k \in [-1, 1]$, $k = 4, \dots, 15$ (radians). We define $\overline{CRLB}(i)$ and $RMSE(i)$ as the average CRLB and the RMS error for the i th numerical case, respectively. And the average CRLB and the average RMS error for the 100 numerical cases are defined as \overline{CRLB} and \overline{RMSE} , respectively.

$$\overline{CRLB}(i) = \frac{1}{12} \sum_{k=4}^{15} CRLB(a_k)(i), \quad i = 1, \dots, 100 \quad (5)$$

$$\overline{CRLB} = \frac{1}{100} \sum_{i=1}^{100} \overline{CRLB}(i), \quad (6)$$

$$\overline{RMSE} = \frac{1}{100} \sum_{i=1}^{100} RMSE(i). \quad (7)$$

We show \overline{CRLB} and \overline{RMSE} obtained by PD and PD-BD in Table 2. \overline{CRLB} obtained by PD-BD-0.75 and PD-BD-1.0 are smaller than that by PDs, which indicates a higher accuracy. Although, \overline{CRLB} obtained by PD-BD-0.5 is a

	Computed				
	PD-0.5	PD-1.0	PD-BD-0.5	PD-BD-0.75	PD-BD-1.0
\overline{CRLB}	2.4303E-5	2.3715E-5	2.4076E-5	2.3637E-5	2.3167E-5
\overline{RMSE}	0.0186 λ	0.0306 λ	0.0143 λ	0.0148 λ	0.0150 λ

Table 2. \overline{CRLB} (radians) and $\overline{RMSE}(\lambda)$ obtained by PD and PD-BD based on 100 numerical cases.

RMS (λ)	RMS errors				
	PD-0.5	PD-1.0	PD-BD-0.5	PD-BD-0.75	PD-BD-1.0
coe (0.0549)	0.0178 λ	0.0309 λ	0.0032 λ	0.0027 λ	0.0022 λ
2 \times coe (0.1098)	0.0222 λ	0.0302 λ	0.0067 λ	0.0069 λ	0.0054 λ
3 \times coe (0.1647)	0.0303 λ	0.0521 λ	0.0137 λ	0.0129 λ	0.0101 λ
4 \times coe (0.2196)	0.0375 λ	0.0615 λ	0.0194 λ	0.0207 λ	0.0176 λ
5 \times coe (0.2745)	0.0543 λ	0.0666 λ	0.0301 λ	0.0261 λ	0.0320 λ
6 \times coe (0.3294)	0.0761 λ	0.0820 λ	0.0772 λ	0.0401 λ	0.0508 λ

Table 3. RMS errors(λ) achieved by PD and PD-BD for wavefront aberrations of different scales.

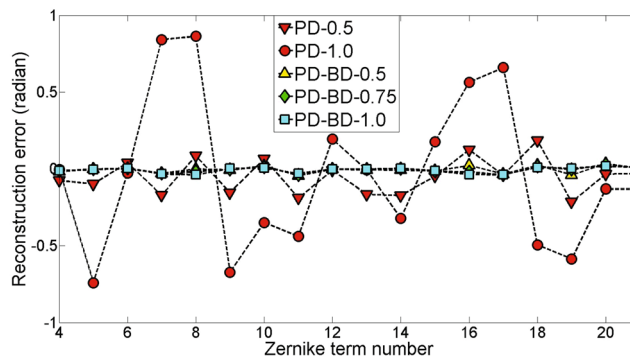


Figure 5. Reconstruction error of each Zernike coefficient by PD and PD-BD for the set coefficient vector $\{a_4, \dots, a_{21}\}$.

little greater than that by PD-1.0, \overline{RMSE} obtained by PD-BD-0.5 is 0.0143 λ , nearly two times smaller than that by PD-1.0. Overall, for the 100 numerical cases, PD-BD gains a higher accuracy than PD in 85% of all cases.

With the increasing of the scales of the measured wavefront aberrations, the accuracy of PD and PD-BD will drop gradually. Here, we will compare the performance of PD and PD-BD for wavefront aberrations of the same spatial frequency distribution, but with different scales. The coefficient vector listed in Table 1 is denoted as **coe**. The measured wavefront aberrations of different scales are generated from **coe**, of which each coefficient is multiplied by a same integer. Table 3 lists the RMS errors achieved by PD and PD-BD for wavefront aberrations of different scales. In general, it is can be seen that the accuracy of PD-BD is higher than PD for wavefront aberrations of different scales. It also seems that PD-BD with different amounts of defocus keep satisfying accuracy for wavefront aberrations of different scales.

At last, we will compare the performance of PD and PD-BD for high-frequency wavefront sensing. The measured high-frequency wavefront coefficient vector $\{a_4, \dots, a_{21}\}$ is generated by adding random high-frequency parts to the relative low-frequency wavefront aberration listed in Table 1. And the coefficient vector $\{a_4, \dots, a_{28}\}$ is also generated by adding random high-frequency parts to coefficient vector $\{a_4, \dots, a_{21}\}$. Reconstruction errors of each Zernike coefficient by PD and PD-BD for coefficient vectors $\{a_4, \dots, a_{21}\}$ and $\{a_4, \dots, a_{28}\}$ are plotted in Fig. 5 and Fig. 6, respectively.

PD-0.5 and PD-1.0 are obviously less accurate than that by PD-BDs. PD-BD-0.5, PD-BD-0.75 and PD-BD-1.0 provide similar high reconstruction accuracy, nearly five times higher than that by PD-0.5. With the addition of high-frequency parts of the measured wavefront aberration, the reconstruction accuracy of PD-0.5 drop gradually. For coefficient vectors $\{a_4, \dots, a_{15}\}$, $\{a_4, \dots, a_{21}\}$ and $\{a_4, \dots, a_{28}\}$, RMS errors obtained by PD-0.5 for different high-frequency wavefront aberrations are 0.0178 λ , 0.0282 λ and 0.0305 λ , respectively. By comparison, RMS errors obtained by PD-BD-0.5 for the same wavefront aberrations are 0.0032 λ , 0.0052 λ and 0.0054 λ , respectively. Compared with PD, it seems that the accuracy of wavefront sensing by PD-BD is less sensitive to the number of Zernike polynomials used and the diversity amplitude, which is a notable advantage.

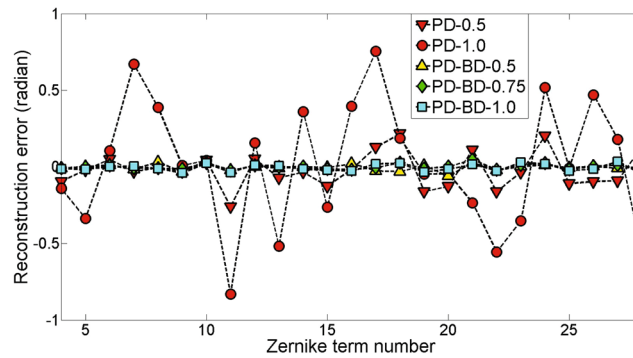


Figure 6. Reconstruction error of each Zernike coefficient by PD and PD-BD for the set coefficient vector $\{a_4, \dots, a_{28}\}$.

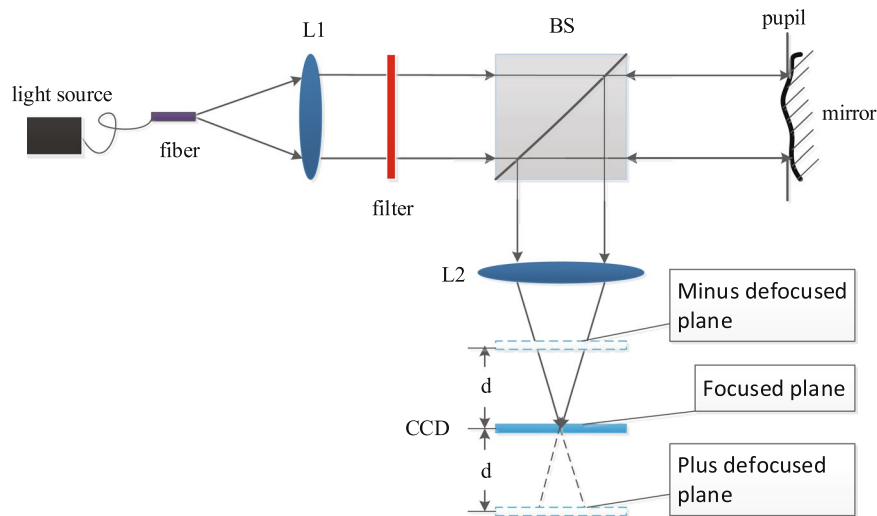


Figure 7. Simplified block diagram of the verification experiment system.

Experimental results. To further examine the accuracy of wavefront sensing with PD and PD-BD, we perform a simple verification experiment. The simplified block diagram of the optical system used here is shown in Fig. 7. The equivalent focal length F is 300 mm, the pupil diameter D is 10 mm, and the CCD pixel size is 5.2 μm . A common halogen lamp is selected as the light source. A fiber bundle is used as the extended object to be reconstructed. The measured wavefront aberration is randomly generated by a homemade deformed mirror with seven actuators. To obtain quasi-monochromatic images, a filter (632.8 nm) is located at the collimated beam. The regularization parameter γ is empirically set to be 5.0×10^{-3} .

We move the CCD by the distance $\pm d$ to get the plus defocused image and the minus defocused image. The corresponding peak-to-valley optical path difference Δ^{26} is equal to

$$\Delta = \frac{d}{8(F/D)^2}. \quad (8)$$

We choose d such that $\Delta = \pm 0.5\lambda$, so the distance d here is 2.279 mm. The measured object takes the area of 450×450 pixels and the exposure time is 20 ms. To get rid of the tilts aberrations induced by moving CCD, we match the plus defocused image and the minus defocused image by an image matching algorithm called speeded-up robust features (SURF)²⁷. So, similar as our simulations, we ignore piston and tilt aberrations $a_1 - a_3$. The computed parameter vector is set as $\mathbf{a} = [a_4, a_5, \dots, a_{28}]$.

To examine the accuracy of wavefront sensing with PD and PD-BD, we will compare the detailed clarity of the reconstructed object images. We denote the wavefront coefficients reconstructed by PD and PD-BD as **PD-coe** and **PD-BD-coe**, respectively. With the same input images, or the same formula for object reconstruction, we put **PD-coe** and **PD-BD-coe** into

$$O(\mathbf{u}) = \frac{H^*(\mathbf{u})I(\mathbf{u}) + H_{+d}^*(\mathbf{u})I_{+d}(\mathbf{u})}{|H(\mathbf{u})|^2 + |H_{+d}(\mathbf{u})|^2 + \gamma}, \quad (9)$$

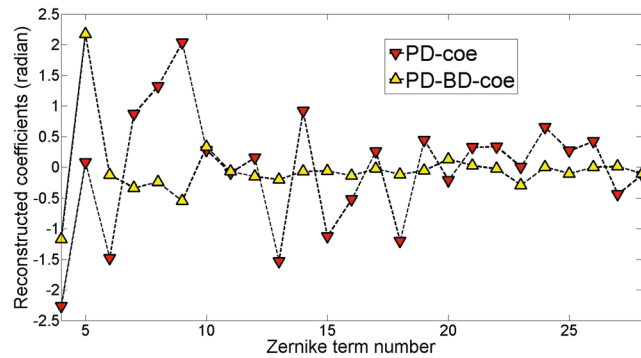


Figure 8. Reconstructed wavefront coefficients by PD and PD-BD for the verification experiment.

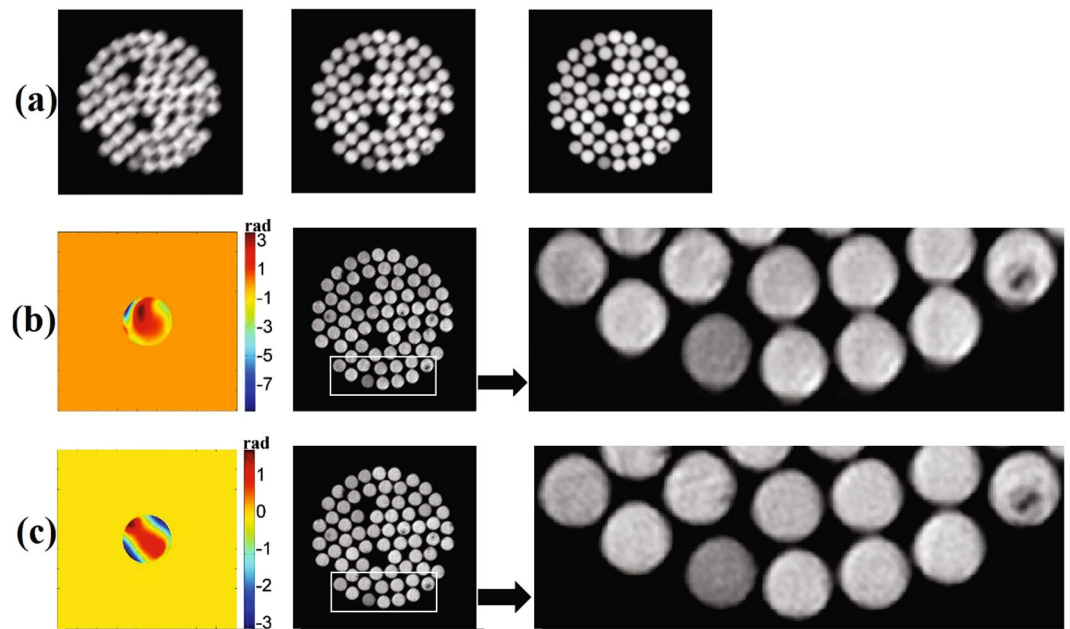


Figure 9. (a) contains the minus defocused image, the focused image and the plus defocused image. (b) and (c) contain reconstructed wavefront and object image by **PD-coe** and **PD-BD-coe**, respectively.

respectively. The reconstructed object images are then obtained by taking the inverse Fourier transform of $O(\mathbf{u})$. Thus, the only difference for the two reconstructed object images are the wavefront coefficients reconstructed by PD and PD-BD.

It can be seen in Fig. 8 that **PD-coe** and **PD-BD-coe** are quite different, especially for Zernike coefficients $a_5 - a_{15}$. The big difference probably means one of them is much more precise than the other one for wavefront sensing. We show from left to right the minus defocused image, the focused image and the plus defocused image in Fig. 9(a). Figure 9(b) shows the wavefront aberration reconstructed by PD and object image reconstructed by Eq. (9). Figure 9(c) shows the wavefront aberration reconstructed by PD-BD and the object image reconstructed by Eq. (9). From the partial enlarged view of Fig. 9(b), apparent distortions for each single fiber can be found, which means there are still non-negligible and uncorrected system aberrations. While the object image reconstructed by **PD-BD-coe** possesses greater clarity than that by **PD-coe** for each single fiber. We think our experiment results have proved indirectly that PD-BD possesses higher accuracy of wavefront sensing than PD.

Discussion

For the example of an astigmatism system aberration $a_5 = 1.5$ radian, PD-BD provides a steeper gradient near the minimum extreme point than PD. A steep gradient of PD-BD near the minimum extreme point can lessen calculative burden and improve the accuracy of wavefront sensing. For the simulations of wavefront sensing with increasing scales of wavefront aberrations, the accuracy of PD and PD-BD drop gradually. Taken together, for different scales of wavefront aberrations, the accuracy of PD-BD is mainly higher than that of PD. With the addition of high-frequency parts of the measured wavefront aberration, the reconstruction accuracy of PD drop gradually. PD-BD could maintain essential reconstruction accuracy for high-frequency wavefront sensing. Compared with

PD, it seems that the accuracy of wavefront sensing by PD-BD is less sensitive to the number of Zernike polynomials used and the diversity amplitude, which is a notable advantage for engineering applications. For the simple verification experiment, we have affirmed the fact that PD-BD possesses a higher accuracy of wavefront sensing than PD.

We have statistically compared the CRLB of PD and PD-BD, which indicates PD-BD has a stronger capacity of phase retrieval than PD. In our opinion, the in-depth reason that PD-BD possesses a higher accuracy than PD is due to the shape of the error metric which results in better convergence in the optimization process. For actual engineering application, PD-BD employs two symmetric images of close SNR near the focal plane. This characteristic of PD-BD is favourable to balance necessary amplitude of defocus diversity and necessary SNR of the defocused image. In other words, PD-BD could better ensure the maximum likelihood estimation, which is the foundation of phase diversity technique.

Methods

In this section, compared with common PD, we briefly describe the bisymmetric defocuses phase diversity. It is supposed that the object is illuminated with non-coherent quasi-monochromatic light, and the imaging system is a linear shift-invariant system²⁸. The intensity distribution of the image plane with Gaussian noise can be modeled by the following equation:

$$i(\mathbf{r}) = o(\mathbf{r}) \otimes h(\mathbf{r}) + n(\mathbf{r}), \quad (10)$$

where \otimes stands for the convolution operation, \mathbf{r} is a two-dimensional position vector in the image plane, $o(\mathbf{r})$ is the object to be found, $h(\mathbf{r})$ is the point-spread function (PSF) of the optical system, and $n(\mathbf{r})$ is Gaussian noise. The point-spread function associated with the focused image is given by

$$h(\mathbf{r}) = |\mathcal{F}^{-1}\{P(\mathbf{v})\exp[j\phi(\mathbf{v})]\}|^2, \quad (11)$$

where \mathbf{v} is a two-dimensional position vector in the pupil plane, $P(\mathbf{v})$ is the binary aperture function with values of 1 inside the pupil and 0 outside, and $\phi(\mathbf{v})$ is the unknown wavefront aberration. \mathcal{F} and \mathcal{F}^{-1} denote the Fourier transform and the inverse Fourier transform, respectively. Similarly, the plus and minus defocused images are given by

$$i_{\pm d}(\mathbf{r}) = o(\mathbf{r}) \otimes h_{\pm d}(\mathbf{r}) + n_d(\mathbf{r}), \quad (12)$$

where

$$h_{\pm d}(\mathbf{r}) = |\mathcal{F}^{-1}\{P(\mathbf{v})\exp[j(\phi(\mathbf{v}) \pm \phi_d(\mathbf{v}))]\}|^2, \quad (13)$$

and the subscript d indicates that a known phase diversity $\pm\phi_d(\mathbf{v})$ has been added to the optical system. The commonly used $\phi_d(\mathbf{v})$ is a known defocus aberration. $\phi_d(\mathbf{v})$ can be exactly determined by the location of the defocused image plane relative to the focused plane.

For common PD, the following error metric.

$$E_{PD} = \frac{1}{2} \sum_{\mathbf{u}} \{|I(\mathbf{u}) - O(\mathbf{u})H(\mathbf{u})|^2 + |I_{+d}(\mathbf{u}) - O(\mathbf{u})H_{+d}(\mathbf{u})|^2\} + \frac{\gamma}{2} \sum_{\mathbf{u}} \{|O(\mathbf{u})|^2\}, \quad (14)$$

is usually used to evaluate the mean-squared image intensity difference between the data predicted by optical system and the data actually collected. The second term $\frac{\gamma}{2} \sum_{\mathbf{u}} \{|O(\mathbf{u})|^2\}$, called the regularization function, which is usually used to prevent oscillations caused by noise amplification. The parameter γ is a small nonnegative constant, which is usually set manually. The major characteristic of PD-BD is that the measured system aberration within the focused plane is reconstructed by two symmetrical defocused planes. In detail, the error metric of PD-BD is

$$E = \frac{1}{2} \sum_{\mathbf{u}} \{|I_{-d}(\mathbf{u}) - O(\mathbf{u})H_{-d}(\mathbf{u})|^2 + |I_{+d}(\mathbf{u}) - O(\mathbf{u})H_{+d}(\mathbf{u})|^2\} + \frac{\gamma}{2} \sum_{\mathbf{u}} \{|O(\mathbf{u})|^2\}, \quad (15)$$

where u is a two-dimensional spatial frequency coordinate. We show the key difference between PD and PD-BD in Fig. 10. $I(\mathbf{u})$ is the image intensity distribution of the focal plane in frequency domain. $I_{+d}(\mathbf{u})$ and $I_{-d}(\mathbf{u})$ are the image intensity distributions of the plus defocused plane and the minus defocused plane in frequency domain, respectively. $H(\mathbf{u})$ is the optical transfer function (OTF) of the focused system. $H_{+d}(\mathbf{u})$ and $H_{-d}(\mathbf{u})$ are the OTFs of the plus defocused system and the minus defocused system:

$$H(\mathbf{u}) = \mathcal{F}\{|\mathcal{F}^{-1}\{P(\mathbf{v})\exp[j\phi(\mathbf{v})]\}|^2\}, \quad (16)$$

$$H_{\pm d}(\mathbf{u}) = \mathcal{F}\{|\mathcal{F}^{-1}\{P(\mathbf{v})\exp[j(\phi(\mathbf{v}) \pm \phi_d(\mathbf{v}))]\}|^2\}, \quad (17)$$

where \mathbf{v} is the two-dimensional position coordinate in the pupil plane, $P(\mathbf{v})$ is the binary aperture function with values of 1 inside the pupil and 0 outside. $\phi(\mathbf{v})$ is the measured system aberration, which is commonly approximated and parameterized by a finite set of Zernike polynomials²⁹:

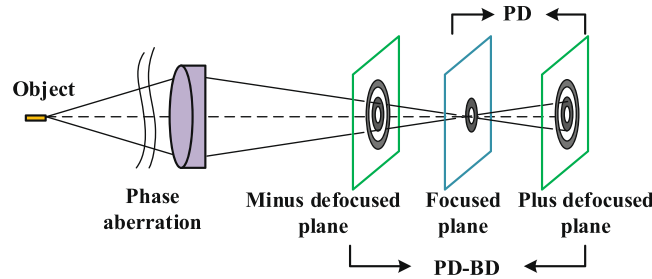


Figure 10. Intensity planes employed by PD and PD-BD.

$$\phi(\mathbf{v}) = \sum_{j=4}^K a_j Z_j(\mathbf{v}), \quad (18)$$

where K is the maximum number of Zernike polynomials. The coefficients $a_1 - a_3$ stand for piston and tilt of the wavefront aberration, which have no effect on the resolution of image. $\phi_d(\mathbf{v})$ is a known defocus diversity phase. Similar to traditional PD, the error metric of PD-BD is minimized by choice of $O(\mathbf{u})$ when

$$O(\mathbf{u}) = \frac{H_{-d}^*(\mathbf{u})I_{-d}(\mathbf{u}) + H_{+d}^*(\mathbf{u})I_{+d}(\mathbf{u})}{|H_{-d}(\mathbf{u})|^2 + |H_{+d}(\mathbf{u})|^2 + \gamma}, \quad (19)$$

where the subscript * represents conjugate operator. Substitution of $O(\mathbf{u})$ into E yields

$$E = \sum_{\mathbf{u}} \frac{|I_{-d}(\mathbf{u})H_{+d}(\mathbf{u}) - I_{+d}(\mathbf{u})H_{-d}(\mathbf{u})|^2}{|H_{-d}(\mathbf{u})|^2 + |H_{+d}(\mathbf{u})|^2 + \gamma}. \quad (20)$$

The error metric E is therefore independence of any object estimate. For a given Zernike parameters $\mathbf{a} = [a_4, a_5, \dots, a_K]$, the error metric $E(\mathbf{a})$ can be calculated. Optimization process of PD-BD is to find the optimal parameter set \mathbf{a} for which the error metric is globally minimum. The measured system aberration is then reconstructed by Eq. (18). And the reconstructed object image is obtained by taking the inverse Fourier transform of $O(\mathbf{u})$.

References

- Gonsalves, R. A. Phase retrieval and diversity in adaptive optics. *Optical Engineering* **21**, 215829–215829 (1982).
- Baba, N., Tomita, H. & Miura, N. Iterative reconstruction method in phase-diversity imaging. *Appl. Opt.* **33**, 4428–4433 (1994).
- Paxman, R. G., Thelen, B. J. & Seldin, J. H. Phase-diversity correction of turbulence-induced space-variant blur. *Opt. Lett.* **19**, 1231–1233 (1994).
- Lee, D. J., Welsh, B. M., Roggemann, M. C. & Ellerbroek, B. L. Diagnosing unknown aberrations in an adaptive optics system by use of phase diversity. *Opt. Lett.* **22**, 952–954 (1997).
- Lee, D. J., Roggemann, M. C., Welsh, B. M. & Crosby, E. R. Evaluation of least-squares phase-diversity technique for space telescope wave-front sensing. *Appl. Opt.* **36**, 9186–9197 (1997).
- Wirth, A., Gonsalves, R. & Jankevics, A. Adaptive optics enabled wavefront diversity sensing. In *Imaging and Applied Optics*, JTuC4 (Optical Society of America, 2011).
- Korkiakoski, V., Keller, C., Doelman, N., Fraanje, R. & Verhaegen, M. Joint-optimization of phase-diversity and adaptive optics. In *Imaging and Applied Optics*, JTuC5 (Optical Society of America, 2011).
- Dolne, J. J. *et al.* Real time phase diversity advanced image processing and wavefront sensing. In *Optical Engineering+Applications*, 67120G–67120 G (International Society for Optics and Photonics, 2007).
- Jean-François, S., Thierry, F., Gérard, R. & Cyril, P. Calibration and precompensation of noncommon path aberrations for extreme adaptive optics. *J. Opt. Soc. Am. A* **24**, 2334–2346 (2007).
- Clélia, R., Thierry, F., Jean-François, S. & Laurent, M. Improvement of phase diversity algorithm for non-common path calibration in extreme AO context. In *Adaptive Optics Systems*, 70156 A (Proceedings of SPIE, 2008).
- Védrenne, N., Mugnier, L. M., Michau, V., Velluet, M.-T. & Bient, R. Laser beam complex amplitude measurement by phase diversity. *Opt. Express* **22**, 4575–4589 (2014).
- Védrenne, N. *et al.* Design and performance of an integrated phase and amplitude diversity sensor. In *CLEO: 2015, STu2N.2* (Optical Society of America, 2015).
- Klapp, I. & Rosen, J. Phase diversity implementation in fresnel incoherent holography. In *Imaging and Applied Optics*, CTh3C.3 (Optical Society of America, 2013).
- Kner, P. Phase diversity for three-dimensional imaging. *J. Opt. Soc. Am. A* **30**, 1980–1987 (2013).
- Paxman, R. G. & Fienup, J. R. Optical misalignment sensing and image reconstruction using phase diversity. *J. Opt. Soc. Am. A* **5**, 914–923 (1988).
- Möller, M. F. A scaled conjugate gradient algorithm for fast supervised learning. *Neural networks* **6**, 525–533 (1993).
- Rondeau, X., Thiébaud, E., Tallon, M. & Foy, R. Phase retrieval from speckle images. *J. Opt. Soc. Am. A* **24**, 3354–3365 (2007).
- Johnson, P. M., Goda, M. E. & Gamiz, V. L. Multiframe phase-diversity algorithm for active imaging. *J. Opt. Soc. Am. A* **24**, 1894–1900 (2007).
- Zhang, P. G. *et al.* Hybrid particle swarm global optimization algorithm for phase diversity phase retrieval. *Opt. Express* **24**, 25704–25717 (2016).
- Paxman, R. G., Schulz, T. J. & Fienup, J. R. Joint estimation of object and aberrations by using phase diversity. *J. Opt. Soc. Am. A* **9**, 1072–1085 (1992).
- Sauvage, J.-F., Mugnier, L., Paul, B. & Villicroze, R. Coronagraphic phase diversity: a simple focal plane sensor for high-contrast imaging. *Opt. Lett.* **37**, 4808–4810 (2012).

22. Van Noort, M., Der Voort, L. R. V. & Löfdahl, M. G. Solar image restoration by use of multi-frame blind de-convolution with multiple objects and phase diversity. *Sol. Phys.* **228**, 191–215 (2005).
23. Scharmer, G. B., Löfdahl, M. G., van Werkhoven, T. & de la Cruz Rodriguez, J. High-order aberration compensation with multi-frame blind deconvolution and phase diversity image restoration techniques. *Astronomy & Astrophysics* **521**, A68 (2010).
24. Lee, D. J., Roggemann, M. C. & Welsh, B. M. Cramér-rao analysis of phase-diverse wave-front sensing. *J. Opt. Soc. Am. A* **16**, 1005–1015 (1999).
25. Dolne, J. J. & Schall, H. B. Cramer-rao bound and phase-diversity blind deconvolution performance versus diversity polynomials. *Appl. Opt.* **44**, 6220–6227 (2005).
26. Mugnier, L. M., Blanc, A. & Idier, J. Phase diversity: A technique for wave-front sensing and for diffraction-limited imaging. In Hawkes, P. (ed.) *Advances in Imaging and Electron Physics*, vol. 141, 1–76 (Elsevier, 2006).
27. Bay, H., Ess, A., Tuytelaars, T. & Gool, L. V. Speeded-up robust features (surf). *Comput. Vis. Image Underst.* **110**, 346–359 (2008).
28. Goodman, J. W. & Gustafson, S. C. Introduction to fourier optics. *Opt. Eng.* **35**, 1513–1513 (1996).
29. Noll, R. J. Zernike polynomials and atmospheric turbulence. *J. Opt. Soc. Am.* **66**, 207–211 (1976).

Acknowledgements

This work is supported by the National Natural Science Foundation of China (grant numbers 11774342, 61378075) and the Science Foundation of State Key Laboratory of Applied Optics, Changchun Institute of Optics, Fine Mechanics and Physics (CIOMP), Chinese Academy of Sciences.

Author Contributions

P.Z. wrote the main manuscript. C.Y. conceived the idea and designed the research. P.Z., C.Y. and Z.X. performed the simulations and analyzed the datas. P.Z., C.Y., Z.C. and Q.M. conducted the experiments. L.X. and Q.M. supervised the research. All authors discussed and reviewed the manuscript.

Additional Information

Competing Interests: The authors declare that they have no competing interests.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2017