Communication

Fast High-Resolution Phase Diversity Wavefront Sensing with L-BFGS Algorithm

Haoyuan Zhang [1,2,3], Guohao Ju [1,3,*], Liang Guo [1,2,3], Boqian Xu [1,3], Xiaoquan Bai [1,3], Fengyi Jiang [1,3] and Shuyan Xu [1,3]

[1] Changchun Institute of Optics, Fine Mechanics and Physics, Chinese Academy of Sciences, Changchun 130033, China; zhanghaoyuan21@mails.ucas.ac.cn (H.Z.); guoliang18@mails.ucas.ac.cn (L.G.); xuboqian@ciomp.ac.cn (B.X.); baixiaoquan@ciomp.ac.cn (X.B.); jiangfengyi@ciomp.ac.cn (F.J.)
[2] University of Chinese Academy of Sciences, Beijing 100049, China
[3] Chinese Academy of Sciences Key Laboratory of On-Orbit Manufacturing and Integration for Space Optics System, Changchun 130033, China
* Correspondence: juguohao@ciomp.ac.cn

Abstract: The presence of manufacture error in large mirrors introduces high-order aberrations, which can severely influence the intensity distribution of point spread function. Therefore, high-resolution phase diversity wavefront sensing is usually needed. However, high-resolution phase diversity wavefront sensing is restricted with the problem of low efficiency and stagnation. This paper proposes a fast high-resolution phase diversity method with limited memory Broyden–Fletcher–Goldfarb–Shanno (L-BFGS) algorithm, which can accurately detect aberrations in the presence of high-order aberrations. An analytical gradient of the objective function for phase-diversity is integrated into the framework of the L-BFGS nonlinear optimization algorithm. L-BFGS algorithm is specifically suitable for high-resolution wavefront sensing where a large phase matrix is optimized. The performance of phase diversity with L-BFGS is compared to other iterative method through simulations and a real experiment. This work contributes to fast high-resolution image-based wavefront sensing with a high robustness.

Keywords: active optics; phase diversity; L-BFGS

1. Introduction

During the long-term on-orbit observation and operation of space-based large-aperture astronomical telescopes, the influence of space temperature or micro-vibration will gradually cause mirror misalignment and deformation [1,2]. Mirror misalignments and deformations will introduce wavefront aberrations and de-grade the imaging quality of the system. The imaging performance of the space optical system can be maintained by using active optics technology [3,4]. It is necessary to obtain the wavefront information of the optical system, and then actively align the system and correct those aberrations with mirror actuators according to the obtained aberration information.

At present, the commonly used wavefront sensing methods are: Shack-Hartmann sensor [5], pyramid sensor [6], curvature sensing [7] and image-based wavefront sensing method. Image-based wavefront sensing -represents a class of methods that directly utilize image plane intensity measurements to recover the wavefront phase of the pupil plane of an optical system. This class of methods mainly include iterative-transform methods (developed from the Gerchberg-Saxton algorithm) [8–13], parametric methods (also known as model-based optimization algorithm or directly called phase diversity algorithm) [14–18], and deep learning methods [19–21]. This image-based wavefront sensing method does not require special hardware devices or complex calibration operations, so this type of method is particularly suitable for wavefront sensing in space telescopes [22].

The Phase Diversity (PD) algorithm is a well-known image-based wavefront sensing technique, which usually uses the known defocus aberration between two defocused images to obtain the unknown wavefront aberration of the telescope system [14–16]. The PD algorithm does not have high requirements on the hardware, and usually only two defocused images need to be collected when performing wavefront sensing on the optical system [23]. The addition of known diversity phase can improve the efficiency and robustness of the wavefront sensing process, and PD is applicable to extended scene. Since the birth of the PD algorithm, this technology has been widely used in many fields such as adaptive optics, active optics, biological microscopic imaging and quality control of laser beams [17,24,25].

The deviation between theoretical and actual acquired image intensities is evaluated by Fourier optics theory to establish the evaluation function. The key of PD algorithm is to find a suitable optimization algorithm to solve the global optimal value of the evaluation function. Many gradient-based nonlinear optimization algorithms, such as steepest descent (SD) algorithm [26,27], conjugate gradient (CG) algorithm [17] and the quasi-Newton algorithm [28,29], etc. have been applied. Among them, the BFGS algorithm proposed by Broyden et al. is a commonly used quasi-Newton method [30]. They used a matrix that does not contain the second derivative to approximate the Hessian matrix in the Newton method, and solved the problem of finding the second partial derivative in the Newton method. However, each iteration of the BFGS algorithm requires a large amount of storage space. When the optimization problem is large-scale, the storage and calculation of the matrix will be difficult. In response to this problem, Liu and Nocedal et al. proposed a limited-memory BFGS algorithm (L-BFGS), which replaces the previous Hessian matrix by storing a small amount of data from the previous m iterations [31].

Factors such as surface defects of the imaging system will produce high-order aberrations. These high-order aberrations can effectively influence the intensity distribution of the point spread function (PSF) which will decrease the accuracy of wavefront sensing. Therefore, high-resolution phase-diversity wavefront sensing is usually required to accurately detect aberrations in optical systems. However, high-resolution phase diversity wavefront sensing is restricted with the problem of low efficiency and stagnation, this paper proposes a fast high-resolution PD wavefront sensing with L-BFGS algorithm. The algorithm takes all pixels on the phase plane as unknown quantities to solve the problem, and uses the analytical gradient calculation method proposed by Fienup et al. to improve the computational efficiency of the algorithm [12].

The remainder of this paper is organized as follows. In Section 2, we review the classical PD algorithm and the L-BFGS algorithm, and propose the basic flow of the fast high-resolution PD algorithm. Section 3 is the simulation analysis of algorithm performance. Section 4 is the experimental verification part. And in Section 5, we conclude this paper.

2. Principle of Fast High-Resolution Phase Diversity Wavefront Sensing with L-BFGS Algorithm

2.1. Review of Phase Diversity Algorithm

In this section, we will give a brief description of the principle of the PD algorithm [14–16]. For a diffraction-limited incoherent imaging system, the image at the focal plane of the system is the convolution of the object with the system PSF:

$$d_1(\mathbf{r}) = o(\mathbf{r}) * h_1(\mathbf{r}) + n_1(\mathbf{r}), \tag{1}$$

where, $d_1(\mathbf{r})$ is the light intensity distribution at the focal plane of the system, $o(\mathbf{r})$ is the unknown target, $h_1(\mathbf{r})$ is the PSF at focal plane, $n_1(\mathbf{r})$ is the detector noise (Gaussian distribution), $*$ represents the convolution operation, $\mathbf{r}$ is the two-dimensional position vector of the image plane, and:

$$h_1(\mathbf{r}) = \left| \mathcal{F}\left\{ A(\boldsymbol{\rho}) \cdot e^{j\varphi(\boldsymbol{\rho})} \right\} \right|^2, \tag{2}$$

where, $A(\boldsymbol{\rho})$ is the pupil amplitude, $\varphi(\boldsymbol{\rho})$ is the unknown wave aberration of the system, $\boldsymbol{\rho}$ is the two-dimensional position vector of the pupil plane, $\mathcal{F}\{\cdot\}$ represents the Fourier transform operation.

The image and PSF of the defocused plane of the system can be expressed as:

$$d_2(\mathbf{r}) = o(\mathbf{r}) * h_2(\mathbf{r}) + n_2(\mathbf{r}), \tag{3}$$

$$h_2(\mathbf{r}) = \left| \mathcal{F}\left\{ A(\boldsymbol{\rho})e^{j[\varphi(\boldsymbol{\rho}) + \Delta(\boldsymbol{\rho})]} \right\} \right|^2, \tag{4}$$

where, $d_2(\mathbf{r})$ is the light intensity distribution at the defocus plane of the system, $h_2(\mathbf{r})$ is the PSF at defocus plane, $n_2(r)$ is the detector noise, and $\Delta(\boldsymbol{\rho})$ is the known defocus aberration.

We usually use Zernike polynomials to represent the wavefront aberrations of optical systems [32]:

$$\varphi(\boldsymbol{\rho}) = \sum_{i=4}^{N} C_i Z_i(\boldsymbol{\rho}), \tag{5}$$

where, $Z_i(\boldsymbol{\rho})$ is the i-th term of the Zernike polynomial, and C_i is the coefficient of the i-th term of the Zernike polynomial. Therefore, given a set of coefficients $a = [C_1, C_2, C_3, \ldots, C_N]$, the corresponding system wavefront aberration can be obtained.

Then, the evaluation function is constructed using the maximum likelihood estimation method:

$$E = \sum_{k=1}^{2} \sum_{\mathbf{r}} [d_k(\mathbf{r}) - o(\mathbf{r}) * h_k(\mathbf{r})]^2, \tag{6}$$

If the unknown system wavefront aberration is to be obtained, it is necessary to find the global optimal solution that makes the evaluation function shown in formula (6) the minimum value. At this time, the problem of using the PD algorithm to solve the system wave aberration is transformed into a nonlinear optimization problem: the unknown wavefront phase information can be obtained by selecting an appropriate optimization algorithm.

2.2. The Principle of L-BFGS Algorithm and the Application of Analytic Gradient in L-BFGS Algorithm

Compared with the BFGS algorithm, the L-BFGS algorithm reduces the requirement for storage capacity, it avoids the calculation of large-scale matrix and improves the calculation efficiency. The iterative formula of the L-BFGS algorithm is as follows:

$$v_{k+1} = v_k + \alpha_k H_k g_k, \tag{7}$$

where, v_k and v_{k+1} are the iteration results of the kth and k+1th iterations respectively, α_k is the step size of the kth iteration, g_k is the gradient of the kth iteration, and H_k is the Hessian matrix (a matrix of 2nd order partial derivatives) of the kth iteration. Define $V_k = I - \rho_k y_k s_k^T$, $\rho_k = \frac{1}{y_k^T s_k}$, $y_k = g_{k+1} - g_k$, $s_k = v_{k+1} - v_k$. Then the H_k can be expressed as:

$$H_{k+1} = V_k^T H_k V_k + \rho_k s_k s_k^T, \tag{8}$$

At the same time, H_k can be obtained by using the initial positive-definite matrix $H_0 = I$ and the information in the previous m steps. Therefore, Equation (8) can be expressed by the initial positive-definite matrix:

$$
\begin{aligned}
H_{k+1} = {} & V_k^T \ldots V_{k-m}^T H_0 \ldots V_k \\
& + \rho_{k-m}(V_k^T \ldots V_{k-m+1}^T)s_{k-m}s_{k-m}^T(V_{k-m+1}^T \ldots V_k^T) \\
& + \rho_{k-m+1}(V_k^T \ldots V_{k-m+2}^T)s_{k-m+1}s_{k-m+1}^T(V_{k-m+2}^T \ldots V_k^T) \\
& + \ldots \\
& + \rho_{k+1}V_k^T s_{k+1}s_{k+1}^T \\
& + \rho_k s_k s_k^T
\end{aligned}
\tag{9}
$$

In this paper, the whole phase plane is taken as the calculation target, and the gradient information of each pixel value on the phase plane is solved separately. Fienup et al. proposed an analytic gradient expression based on the Fourier transform [12], which can be used to obtain the derivative of Equation (6) for the phase plane:

$$\frac{\partial E}{\partial \varphi} = -2\text{Im}\left\{P(\varphi)\mathcal{F}^{-1}\left\{G^{w^*}(\varphi)\right\}\right\} \tag{10}$$

where, $G^{w^*}(\varphi) = 2 \cdot \left[|F(\varphi)|^2 - i\right] \cdot F(\varphi)^*$, $F(\varphi) = \mathcal{F}\left\{A(\boldsymbol{\rho})e^{j\varphi(\boldsymbol{\rho})}\right\}$, the superscript $*$ represents complex conjugate, $\mathcal{F}^{-1}\{\cdot\}$ represents the inverse Fourier transform, $\text{Im}\{\cdot\}$ represents the imaginary part of a complex number, and the generalized pupil function $P(\varphi) = A(\boldsymbol{\rho})e^{j\varphi(\boldsymbol{\rho})}$.

Equation (10) shows that only one inverse Fourier transform is needed to obtain the gradient information of $n \cdot n$ pixel values on the phase plane, and the complexity of calculating the gradient is reduced from $O(n \cdot n)$ to $O(1)$. And Equation (10) corresponds to g_k in Equation (7), the application of this analytical gradient expression in the L-BFGS algorithm can significantly improve the convergence efficiency of the algorithm.

2.3. Fast High-Resolution Phase Diversity Wavefront Sensing with L-BFGS Algorithm

Different from the traditional PD algorithm that takes the Zernike coefficient as the solution target, the fast high-resolution PD (hereinafter referred to as high-resolution PD) algorithm proposed in this paper uses all the pixels on the phase plane as the unknown quantity to solve the problem. Combining the analytical gradient calculation method shown in Formula (10) with the L-BFGS algorithm, the computational complexity of the Hessian matrix will change from $O(n \cdot n)$ to $O(n \cdot m)$ (usually m is much smaller than n), thus improving the computational efficiency. The algorithm process is as Figure 1:
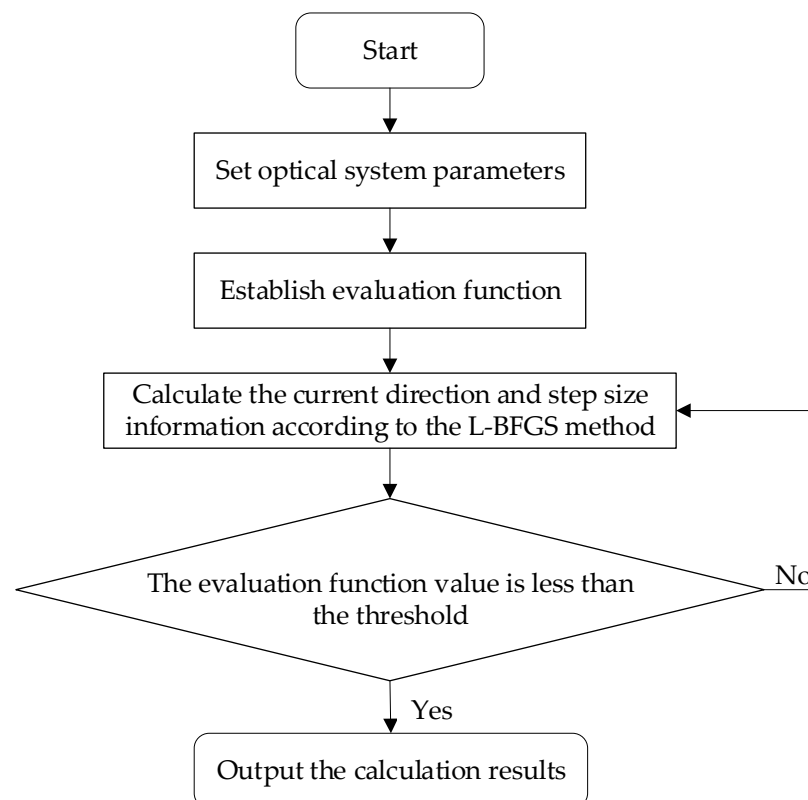

Figure 1. Algorithm flowchart.

3. Simulations

3.1. System Parameter Setting

We set the diameter of the primary mirror of the optical system to be 2 m, the focal length of the system to be 28 m, the observation wavelength to be 625 nm, the number of CCD samples to be 256 × 256 pixels, the pixel size of the CCD to be 5.5 μm, and the defocus distance between the two images to be 10 mm.

3.2. Solution Accuracy Analysis of High-Resolution PD Algorithm

In this paper, root mean square (RMS) is used to represent the magnitude of wavefront aberration, and use root-mean-square error (RMSE) to measure the accuracy of the solution. Three sets of Zernike polynomial coefficients with different RMS are used to simulate the wavefront aberration, and the generated PSF image is substituted into the high-resolution PD algorithm to solve the wavefront aberration. Then, high-order aberrations are added for simulation to analyze the ability of the high-resolution PD algorithm to solve wavefront aberrations when high-order aberrations are included. The calculation result is shown in Figure 2:
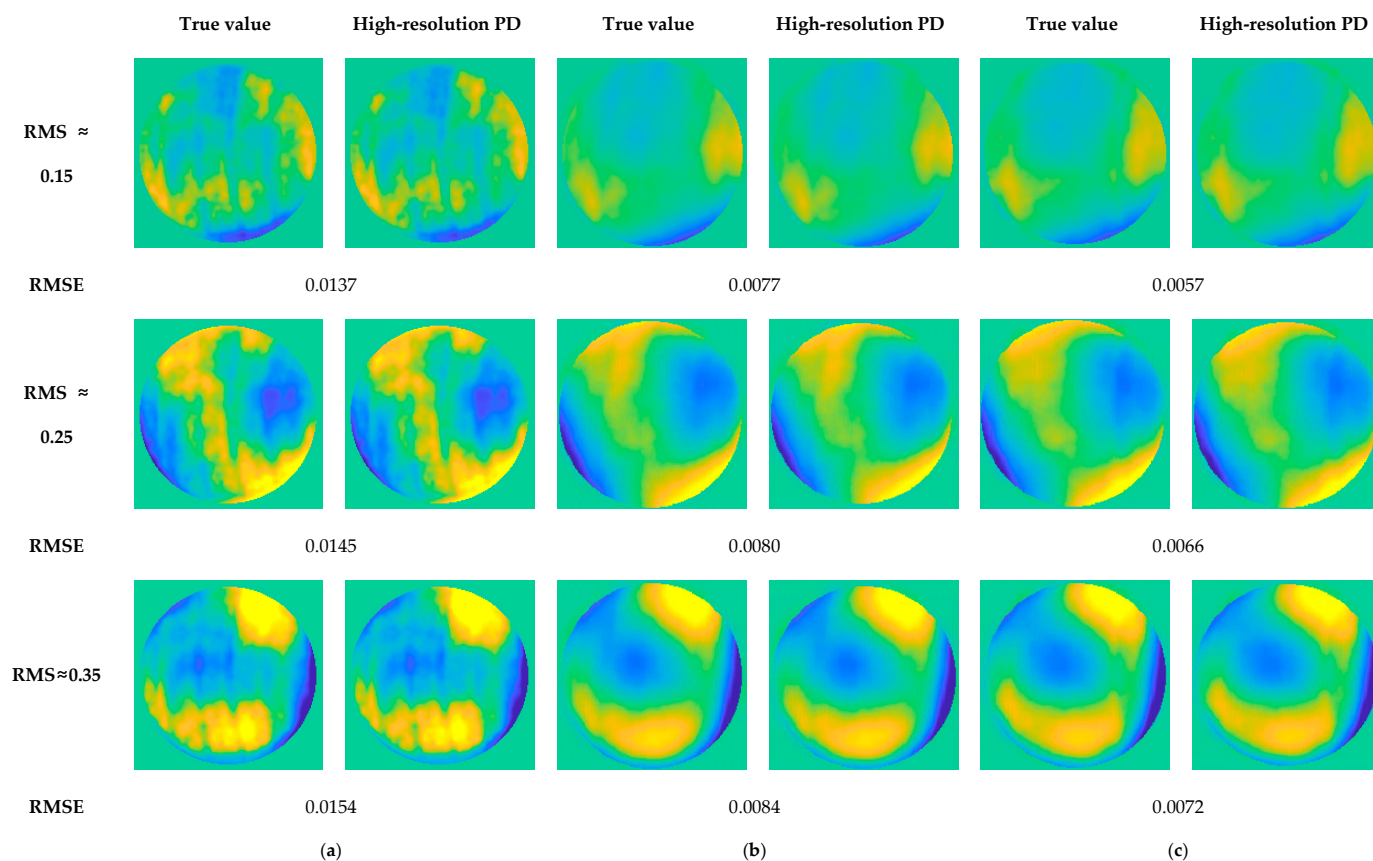

Figure 2. The simulation analysis results of solution accuracy of high-resolution PD algorithm. (**a**–**c**) are in three different high-order aberration situations, respectively. In each case, the simulation analysis is carried out by adding different magnitudes of low-order wavefront aberrations.

The following conclusions can be drawn from Figure 2:
- The high-resolution PD algorithm achieves convergence when introducing different high-order aberrations;
- The calculation accuracy of wavefront aberrations is affected by higher order aberrations;
- The larger the wavefront aberration, the lower the solution accuracy.

3.3. Comparative Analysis

3.3.1. Comparative Analysis of Convergence Efficiency

　　When high-order aberrations are involved, the high-resolution PD algorithm, the traditional PD algorithm for solving Zernike coefficients, and the GS algorithm are used to solve the aberrations to compare the convergence efficiency of the three algorithms. The results are shown in Figure 3:
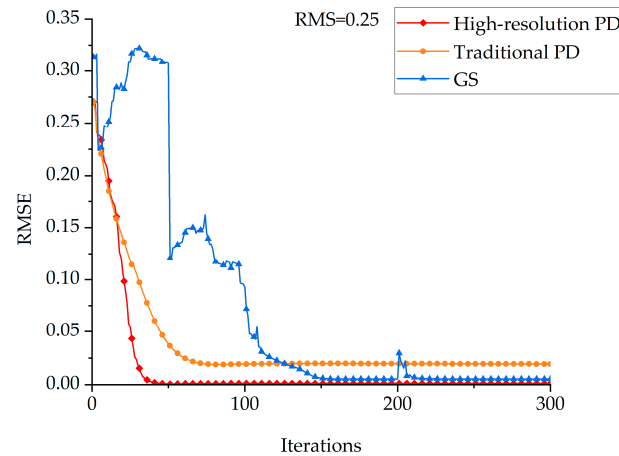

Figure 3. The solution results of the three algorithms for wavefront aberrations when the wavefront aberration size is RMS = 0.25 and with high-order aberrations.

　　The following conclusions can be drawn from Figure 3:

- The high-resolution PD algorithm converges faster and has higher solution accuracy than the traditional PD algorithm;
- The GS algorithm guarantees the solution accuracy through multiple cross iterations but greatly sacrifices the convergence efficiency;
- The high-resolution PD algorithm can also quickly converge while ensuring the solution accuracy. We can see the proposed algorithm converges 2 times faster than the GS algorithm, which is commonly used now.

3.3.2. Comparing Analysis of Solution Accuracy

　　Three sets of Zernike coefficients are selected to simulate wavefront aberrations with higher order aberrations. The results of the three algorithms are shown in Figure 4:
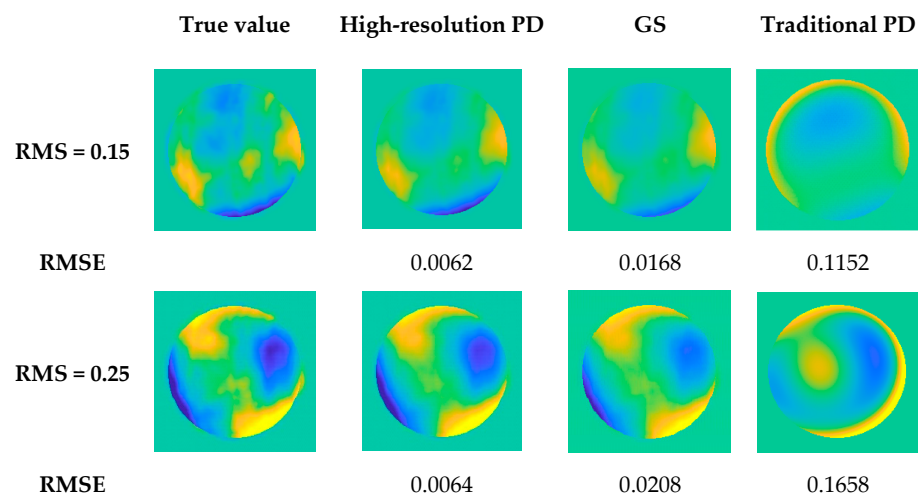

Figure 4. *Cont.*

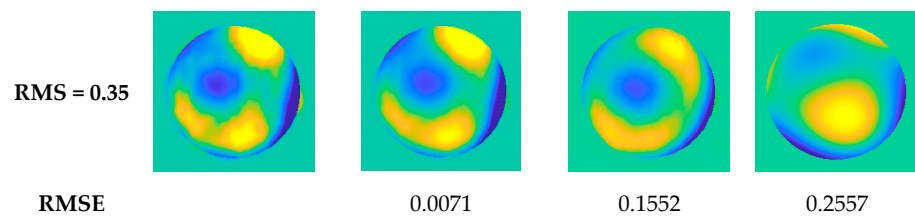| RMS = 0.35				
RMSE		0.0071	0.1552	0.2557

Figure 4. The result of the solution of the three algorithms.

The following conclusions can be drawn from Figure 4:

- In the case of high-order aberrations in the system, the accuracy of the high-resolution PD algorithm for solving aberrations is better than the other two algorithms;
- When the wavefront aberration is small, the GS algorithm can correctly solve the aberration. However, when the wavefront aberration is large, the GS algorithm falls into the local extremum, and the correct aberration information cannot be obtained;
- In the case of high-order aberrations in the system, the traditional PD algorithm cannot solve the aberrations;
- The solution accuracy decreases with the increase of wavefront aberration.

3.3.3. Comparing Analysis of Robustness Analysis

In order to be more realistic, the noise related to the intensity and satisfying the Gaussian distribution is added to the simulated PSF image. Use Peak Signal-to-Noise Ratio (PSNR) as a criterion for evaluating noise magnitude:

$$PSNR = 20\log_{10}\left(\frac{S_{peak}}{\sqrt{S_{peak} + \sigma^2_{read} + \sigma^2_{dark}}}\right) \qquad (11)$$

where, S_{peak} is the maximum value of the intensity in the noise-free image, σ^2_{read} and σ^2_{dark} are the variance of read noise and dark current noise, respectively.

In order to further verify the effectiveness of the proposed method, 100 groups of Monte Carlo tests are carried out and the results are presented in Figure 5. In the range of $[-0.3\lambda, 0.3\lambda]$, 100 sets of Zernike coefficients are randomly generated to verify the robustness of the algorithm under different SNR conditions:
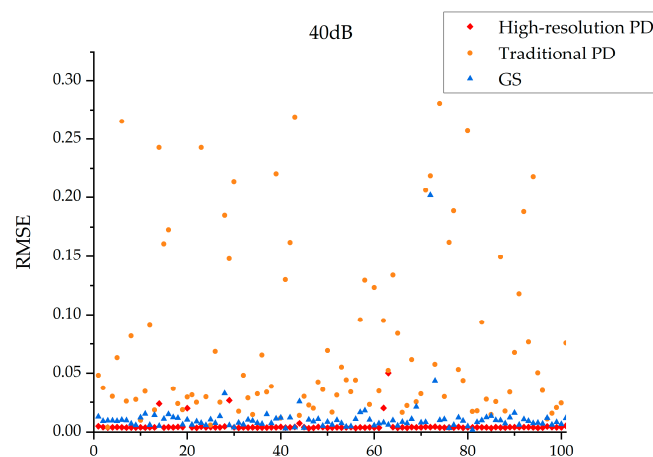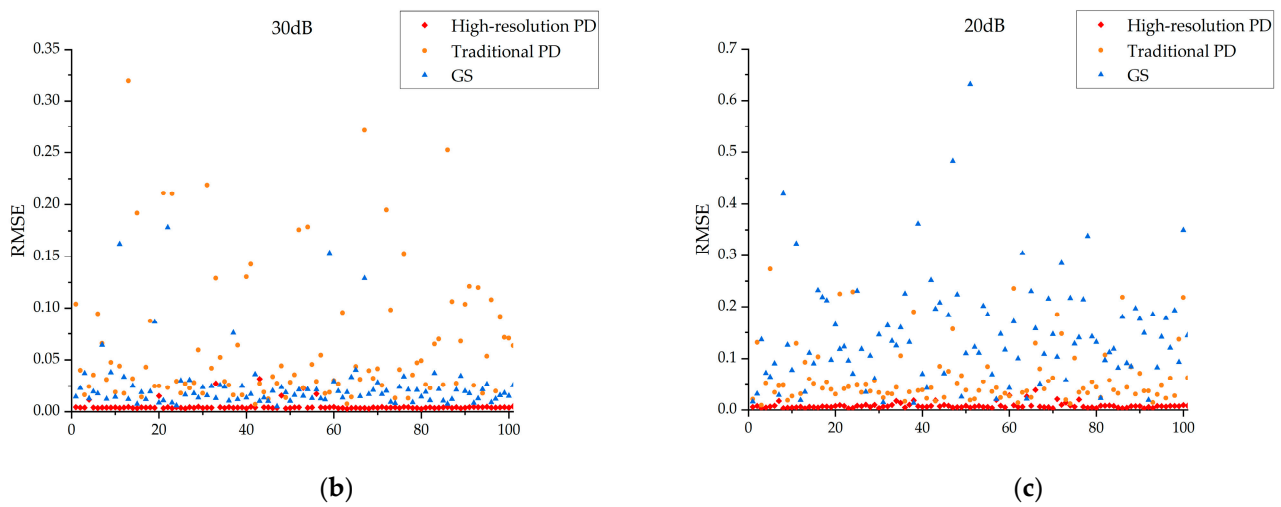

(**a**)

Figure 5. *Cont*.

Figure 5. The algorithm robustness analysis under different SNR conditions. The SNRs of (**a**–**c**) are 40 dB, 30 dB and 20 dB respectively.

The following conclusions can be drawn from Figure 5:

- In the case of low signal-to-noise ratio, the high-resolution PD algorithm and GS algorithm are relatively stable, while the traditional PD algorithm is more likely to fall into local extremum;
- When the noise becomes larger, the high-resolution PD algorithm can still maintain stability, while the stability of the GS algorithm will gradually decrease;

4. Experimental

The experimental verification optical path shown in Figure 6 was designed and built by using the existing equipment in the laboratory. Which is mainly composed of a telescope system, an interferometer, a CCD camera, and a beam splitter prism. And a flat mirror is placed in front of the main mirror of the telescope to form a self-collimating optical path. In order to reduce the influence of airflow disturbance, an interferometer and a CCD camera are used to collect data simultaneously.
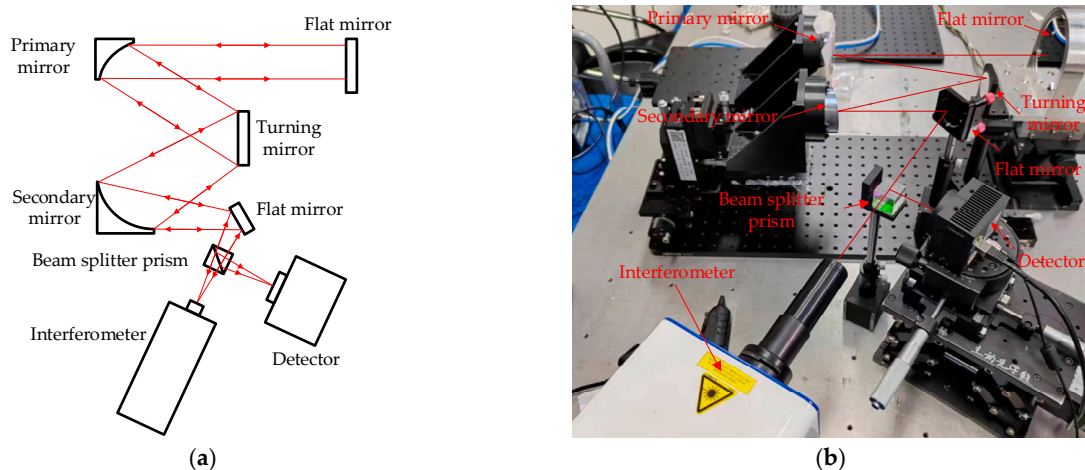

Figure 6. High-resolution PD algorithm experimental optical path diagram, (**a**) is the schematic diagram, and (**b**) is the experimental optical path diagram. The light emitted by the interferometer is irradiated to the flat mirror after passing through the secondary mirror, the turning mirror and the primary mirror of the telescope, and then returns according to the original path. The returned light is divided into two paths by the beam splitter prism, one path of light returns to the interferometer and the other path of light converges on the focal plane of the detector.

A set of images containing known defocus aberrations is collected by moving the CCD camera along the optical axis, and then the high-resolution PD algorithm is used to solve the wavefront aberrations in the optical system. At the same time, the high-resolution algorithm is verified with the data collected by the interferometer. The result is shown in Figure 7:
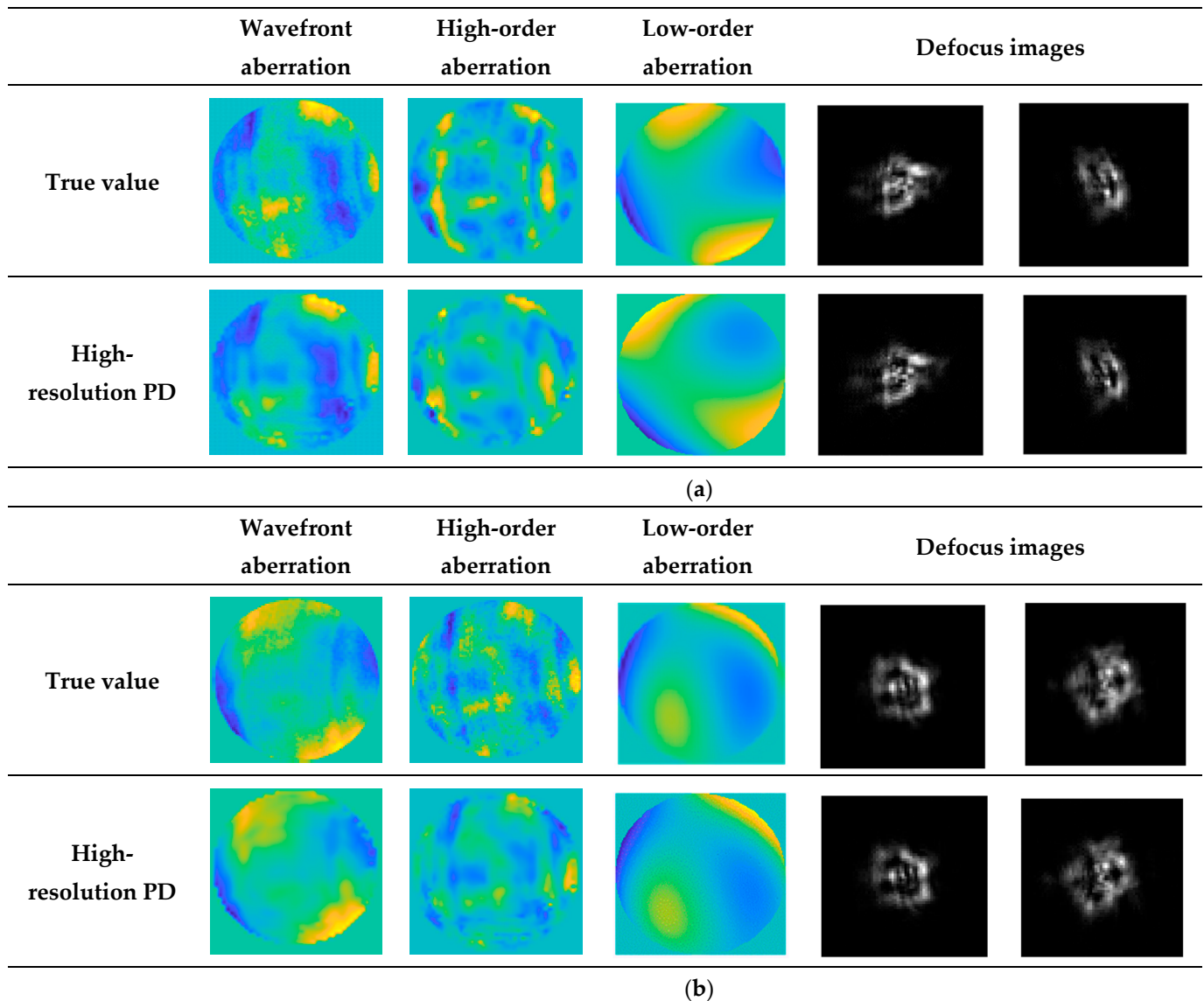

Figure 7. Experimental results. (**a**,**b**) are the experimental results conducted in two states of the optical system, respectively This figure shows the results of the high-resolution PD algorithm restoration of wavefront aberrations, higher-order aberrations, lower-order aberrations, and defocused images. It can be seen that the wavefront aberrations of the optical system are successfully solved by the high-resolution PD algorithm.

In order to avoid the influence of air flow disturbance on the experimental results, four sets of images were collected in each of the two states. Subsequently, the solution results are expressed by Zernike polynomials, and the solution results of the 4th to 7th items of the Zernike polynomials are shown in Table 1:

Table 1. Experimental verification results of high-resolution PD algorithm. (a) and (b) represent the two states of the optical system, respectively. A–D represents four sets of data collected in the same state.

		C4	C5	C6	C7	RMSE
(a)	True value	−0.0090	−0.1847	0.0539	0.0699	
	A	0.0069	−0.1792	0.0463	0.0902	0.0103
	B	0.0204	−0.1675	0.0490	0.0726	0.0140
	C	0.0184	−0.1706	0.0560	0.0687	0.0126
	D	0.0165	−0.2076	0.0423	0.0792	0.0149
		C4	**C5**	**C6**	**C7**	**RMSE**
(b)	True value	−0.0526	0.0636	0.0837	0.0306	
	A	−0.0406	0.0785	0.0775	0.0285	0.0081
	B	−0.0438	0.0787	0.0553	0.0199	0.0129
	C	−0.0453	0.0814	0.0787	0.0200	0.0089
	D	−0.0688	0.0841	0.0558	0.0215	0.0149

It can be seen from Figure 7 and Table 1 that under different conditions, the high-resolution PD algorithm can successfully solve the wavefront aberration of the optical system, and the result is relatively stable.

5. Conclusions

High-resolution phase diversity wavefront sensing is of great importance in the area of optics. However, high-resolution phase diversity wavefront sensing is restricted with the problem of low efficiency and stagnation. We propose a fast and high-resolution PD wavefront sensing method based on the L-BFGS algorithm to address this problem. The algorithm takes all pixels on the phase plane as unknown quantities, and uses the analytical gradient shown in Equation (10) to improve the computational efficiency of the algorithm.

Simulations are performed to demonstrate the accuracy and convergency of the proposed algorithm. On one hand, it is shown that the proposed algorithm can accurately recover high-resolution wavefront phase over a large range of wavefront error (i.e., this algorithm is robust to stagnation problem). On the other hand, the results also show that this algorithm is superior over other algorithms in accuracy and convergence efficiency.

Real experiments are performed to further validate the effectiveness of the proposed method. It is shown that the proposed method can accurately recover high-resolution wavefront phase when two defocused images are available.

This work provides a feasible solution to the problem of low efficiency and stagnation in high-resolution phase diversity wavefront sensing.

References

1. Noethe, L. Active optics in modern large optical telescopes. *Prog. Opt.* **2002**, *43*, 3–4.
2. Dolkens, D.; Van Marrewijk, G.; Kuiper, H. Active correction system of a deployable telescope for Earth observation. In Proceedings of the International Conference on Space Optics—ICSO 2018, Crete, Greece, 9−12 October 2018; pp. 111–121.
3. Tarenghi, M.; Wilson, R. The ESO NTT (New Technology Telescope): The first active optics telescope. In Proceedings of the Active telescope systems, Orlando, FL, USA, 28−31 March 1989; pp. 302–313.
4. Poberezhskiy, I.; Luchik, T.; Zhao, F.; Frerking, M.; Basinger, S.; Cady, E.; Colavita, M.M.; Creager, B.; Fathpour, N.; Goullioud, R. Roman space telescope coronagraph: Engineering design and operating concept. In Proceedings of the Space Telescopes and Instrumentation 2020: Optical, Infrared, and Millimeter Wave, Online, 14−18 December 2020; pp. 314–335.
5. Ko, J.; Davis, C.C. Comparison of the plenoptic sensor and the Shack–Hartmann sensor. *Appl. Opt.* **2017**, *56*, 3689–3698. [CrossRef] [PubMed]
6. Plantet, C.; Meimon, S.; Conan, J.-M.; Fusco, T. Revisiting the comparison between the Shack-Hartmann and the pyramid wavefront sensors via the Fisher information matrix. *Opt. Express* **2015**, *23*, 28619–28633. [CrossRef]
7. An, Q.; Wu, X.; Lin, X.; Ming, M.; Wang, J.; Chen, T.; Zhang, J.; Li, H.; Chen, L.; Tang, J. Large segmented sparse aperture collimation by curvature sensing. *Opt. Express* **2020**, *28*, 40176–40187. [CrossRef] [PubMed]
8. Fienup, J.R. Phase retrieval algorithms: A personal tour. *Appl. Opt.* **2013**, *52*, 45–56. [CrossRef] [PubMed]
9. Zhao, L.; Bai, J.; Hao, Y.; Jing, H.; Wang, C.; Lu, B.; Liang, Y.; Wang, K. Modal-based nonlinear optimization algorithm for wavefront measurement with under-sampled data. *Opt. Lett.* **2020**, *45*, 5456–5459. [CrossRef]
10. Zhao, L.; Yan, H.; Bai, J.; Hou, J.; He, Y.; Zhou, X.; Wang, K. Simultaneous reconstruction of phase and amplitude for wavefront measurements based on nonlinear optimization algorithms. *Opt. Express* **2020**, *28*, 19726–19739. [CrossRef]
11. Gerhberg, R.; Saxton, W. A practical algorithm for the determination of phase from image and diffraction plane picture. *Optik* **1972**, *35*, 237–246.
12. Fienup, J.R. Phase-retrieval algorithms for a complicated optical system. *Appl. Opt.* **1993**, *32*, 1737–1746. [CrossRef]
13. van Kooten, M.A.; Ragland, S.; Jensen-Clem, R.; Xin, Y.; Delorme, J.-R.; Wallace, J.K. On-sky Reconstruction of Keck Primary Mirror Piston Offsets Using a Zernike Wavefront Sensor. *Astrophys. J.* **2022**, *932*, 109. [CrossRef]
14. Gonsalves, R.A.; Chidlaw, R. Wavefront sensing by phase retrieval. In Proceedings of the Applications of Digital Image Processing III, San Diego, CA, USA, 27−30 August 1979; pp. 32–39.
15. Gonsalves, R.A. Phase retrieval and diversity in adaptive optics. *Opt. Eng.* **1982**, *21*, 829–832. [CrossRef]
16. Paxman, R.G.; Schulz, T.J.; Fienup, J.R. Joint estimation of object and aberrations by using phase diversity. *JOSA A* **1992**, *9*, 1072–1085. [CrossRef]
17. Paxman, R.G.; Fienup, J. Optical misalignment sensing and image reconstruction using phase diversity. *JOSA A* **1988**, *5*, 914–923. [CrossRef]
18. Qi, X.; Ju, G.; Xu, S. Efficient solution to the stagnation problem of the particle swarm optimization algorithm for phase diversity. *Appl. Opt.* **2018**, *57*, 2747–2757. [CrossRef]
19. Ju, G.; Qi, X.; Ma, H.; Yan, C. Feature-based phase retrieval wavefront sensing approach using machine learning. *Opt. Express* **2018**, *26*, 31767–31783. [CrossRef]
20. Möckl, L.; Petrov, P.N.; Moerner, W. Accurate phase retrieval of complex 3D point spread functions with deep residual neural networks. *Appl. Phys. Lett.* **2019**, *115*, 251106. [CrossRef] [PubMed]
21. Dumont, M.; Correia, C.; Sauvage, J.-F.; Schwartz, N.; Gray, M.; Beltramo-Martin, O.; Cardoso, J. Deep learning for space-borne focal-plane wavefront sensing. In Proceedings of the Space Telescopes and Instrumentation 2022: Optical, Infrared, and Millimeter Wave, Montréal, QC, Canada, 17−23 July 2022; pp. 1107–1114.
22. Acton, D.S.; Atcheson, P.D.; Cermak, M.; Kingsbury, L.K.; Shi, F.; Redding, D.C. James Webb Space Telescope wavefront sensing and control algorithms. In PProceedings of the Optical, Infrared, and Millimeter Space Telescopes, Glasgow, UK, 21−25 June 2004; pp. 887–896.
23. Bailén, F.J.; Suárez, D.O.; Rodríguez, J.B.; Del Toro Iniesta, J. Optimal Defocus for Phase Diversity Wave Front Retrieval. *Astrophys. J. Suppl. Ser.* **2022**, *263*, 8. [CrossRef]
24. Paxman, R.G.; Thelen, B.J.; Seldin, J.H. Phase-diversity correction of turbulence-induced space-variant blur. *Opt. Lett.* **1994**, *19*, 1231–1233. [CrossRef] [PubMed]
25. Bolcar, M.R.; Fienup, J.R. Sub-aperture piston phase diversity for segmented and multi-aperture systems. *Appl. Opt.* **2009**, *48*, A5–A12. [CrossRef]
26. Fienup, J.R. Phase retrieval algorithms: A comparison. *Appl. Opt.* **1982**, *21*, 2758–2769. [CrossRef]
27. Patwardhan, V. Solution of a NxN System of Linear algebraic Equations: 1—The Steepest Descent Method Revisited. *arXiv* **2022**, arXiv:2206.07482.
28. Rondeau, X.; Thiébaut, E.; Tallon, M.; Foy, R. Phase retrieval from speckle images. *JOSA A* **2007**, *24*, 3354–3365. [CrossRef] [PubMed]
29. Johnson, P.M.; Goda, M.E.; Gamiz, V.L. Multiframe phase-diversity algorithm for active imaging. *JOSA A* **2007**, *24*, 1894–1900. [CrossRef]
30. Broyden, C.G. Quasi-Newton methods and their application to function minimisation. *Math. Comput.* **1967**, *21*, 368–381. [CrossRef]

31. Liu, D.C.; Nocedal, J. On the limited memory BFGS method for large scale optimization. *Math. Program.* **1989**, *45*, 503–528. [CrossRef]

32. Zernike, V.F. Beugungstheorie des schneidenver-fahrens und seiner verbesserten form, der phasenkontrastmethode. *Physica* **1934**, *1*, 689–704. [CrossRef]