applied optics

Co-phasing of the segmented mirror and image retrieval based on phase diversity using a modified algorithm

DAN YUE,[1,2,*] SHUYAN XU,[1] AND HAITAO NIE[1]

[1]Space Optics Department, Changchun Institute of Optics, Fine Mechanics and Physics, Chinese Academy of Sciences, Changchun, Jilin 130033, China
[2]University of Chinese Academy of Sciences, Beijing 100049, China
*Corresponding author: danzik3@126.com

Received 23 June 2015; revised 10 August 2015; accepted 12 August 2015; posted 13 August 2015 (Doc. ID 243575); published 3 September 2015

The conventional Broyden–Fletcher–Goldfarb–Shanno (BFGS) method used to solve the cost function of a phase diversity (PD) algorithm converges to a global optimum only when the cost function is convex. We present a modified BFGS method, which has fine global convergences for both convex and nonconvex functions, guarantees that the solutions will converge to the global minimum, corresponding to the actual wavefront coefficients, and apply it to minimize the PD cost function to co-phase the segmented active optics system and recover the unknown object under different noise levels. The noise amplification effect on the accuracy of the algorithm is removed by our proposed estimated strategy of the regularization parameter for the PD problem. The vast contrast results demonstrate that the modified method has a much higher accuracy than the conventional BFGS method for the nonconvex condition even under a considerably high noise level.

OCIS codes: (070.2025) Discrete optical signal processing; (100.5070) Phase retrieval; (110.1080) Active or adaptive optics; (110.3010) Image reconstruction techniques.

http://dx.doi.org/10.1364/AO.54.007917

1. INTRODUCTION

The trend in next-generation large aperture telescope design will be segmented mirror synthetic aperture optical (SAO) systems. Classic examples of such systems include the James Webb space telescope [1] and the Keck telescope [2]. The segmented primary mirror stitches a series of submirrors together to reach the optimal capabilities equivalent to a monolithic one while avoiding the considerable problems encountered in the fabrication of a monolithic primary mirror with apertures larger than 8 m. However, high-quality images equivalent to those of a monolithic mirror can only be achieved if co-phasing of the segmented mirrors occurs. Co-phasing a segmented mirror is to remove misalignments resulting from relative piston aberrations between segments and tip-tilt aberrations of each segment, which is one of the most important problems related to segmented telescopes.

Considerable effort has been directed toward the co-phasing of the segmented mirror to obtain nearly diffraction-limited performance from the total aperture. Many methods have been proposed for sensing piston and tip-tilt misalignments, such as modified Shack–Hartmann wavefront sensing (WFS) [3], the Mach–Zehnder interferometer sensor [4], the pyramid sensor [5], curvature sensing [6], the Zernike phase contrast sensor [7], and phase diversity (PD) WFS [8,9]. PD WFS stands out in the development of technologies for the co-phasing of SAO systems because traditional wavefront reconstruction using methods such as Shack–Hartmann WFS tend to break down at the mirror segment edges. Meanwhile, the additional hardware required for this technique is modest and it is highly sensitive to the continuous or discontinuous input of a distorted wavefront for a point source or extended scene.

The PD algorithm is an *a posteriori*, image-based wavefront-sensing technique that utilizes a pair of simultaneously collected images with a known phase diversity to construct the optimization cost function based on the maximum likelihood or least squares estimation theory, then jointly estimates both the system aberrations and the unknown object by minimizing the constructed cost function. Conventional nonlinear optimization methods are employed to find those misalignments that are consistent with both degraded images, such as the gradient descent method, conjugate gradient method, quasi-Newton method, neural network algorithm, and genetic algorithm [10]. A common quasi-Newton method called the Broyden–Fletcher–Goldfarb–Shanno (BFGS) method, which was

brought up by Broyden, Fletcher, Goldfarb, and Shanno in 1970, is widely recognized as one of the most effective methods to solve unconstrained optimization problems and has become a popular way to solve engineering optimization problems due to its fine stability and fast convergence. Currently, the BFGS method has an impeccable global convergence property for convex functions. However, for nonconvex functions this method will fall into local minimums and cannot yield a global optimization solution. Aimed at the situation that the cost functions are nonconvex, a modified BFGS method is presented in this paper and applied to sense the co-phase errors and restore the unknown object for segmented active optics systems under different noise levels. In order to overcome the noise amplification effect on the accuracy of the modified algorithm, we also present an efficient and feasible technique to estimate the regularization parameter for the regularized PD cost function.

This paper is organized as follows. In Section 2 we review the PD algorithm for segmented telescopes and introduce the modified BFGS method and estimate strategy of the regularization parameter. In Section 3 we describe the numerical simulations that compare the modified BFGS method to the conventional BFGS method under different noise levels and give the statistics results of the contrast experiments. In Section 4, we summarize and conclude the paper.

2. THEORY

This section first describes the basic theory of the PD algorithm and then presents the modified BFGS optimization method and estimate strategy of the regularization parameter.

A. Statement of the PD Algorithm

The PD algorithm utilizes multichannel images of the same target to jointly estimate both the system aberrations and the unknown object. One of these images is recorded in the focal plane of the optical system degraded by some unknown aberrations, such as turbulence or telescope optical aberration; the other of these images is collected in a separate channel that introduces an extra known aberration to perturb the infocus image, such as defocus, astigmatism, or coma. Figure 1 shows the optical layout of the phase diversity method used for the segmented active optics system. The quadratic defocus is introduced here since it is easier to obtain the out-of-focus image and the defocus length can be precisely measured. The use

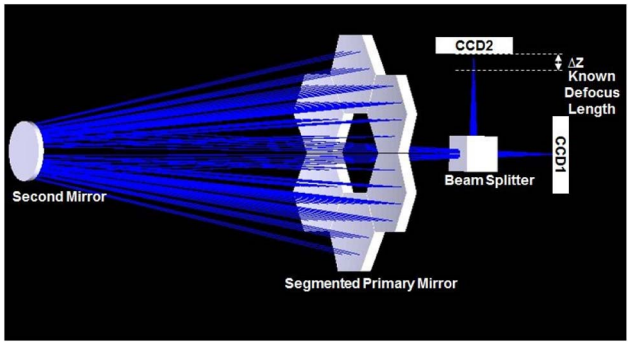


Fig. 1. Optical layout for the phase diversity method.

of multichannel images can effectively resolve the absence of the solution's uniqueness by adding constraints.

When the object or scene is illuminated with spatially incoherent quasi-monochromatic light, the imaging system can be simplified to a linear shift-invariant system, leading to the following basic imaging equation,

$$d_k(u,v) = o * s_k(u,v) + n_k(u,v) \qquad k = 1,2\ldots K, \tag{1}$$

where o is the true object, $s_k(u,v)$ and $n_k(u,v)$ are the point spread function (PSF) and additive noise term of the kth channel, respectively, and $d_k(u,v)$ is the kth detected image. The asterisk denotes two-dimensional (2D) convolution and (u,v) represents the coordinate vector in the image plane.

In order to simplify the situation, just like Keck [11] all of the subapertures are assumed to have the same shape and are perfect without high-order aberrations except co-phasing errors, namely pistons and tip-tilts. Thus, the generalized pupil function can be shown as

$$P(\varepsilon,\eta) = \sum_{n=1}^{N} p_n(\varepsilon,\eta)\exp\left[i\frac{2\pi}{\lambda}(e_n Z_1 + t_{xn} Z_2 + t_{yn} Z_3)\right], \tag{2}$$

where Z_1, Z_2 and Z_3 are piston, tip, and tilt aberrations in the Zernike polynomials, respectively, e_n, t_{xn}, and t_{yn} are the corresponding aberration coefficients of the nth submirror, respectively, (ε,η) is the coordinate vector in the pupil plane, and p_n is the binary function of the nth subaperture.

Under the near-field approximation, the PSF of the kth channel is the squared modulus of the coherent impulse response function given by Eq. (3),

$$\begin{aligned} s_k(u,v) &= |h_k(u,v)|^2, \\ &= \left|\Im\left\{\left(\sum_{n=1}^{N} p_n(\varepsilon,\eta)\exp\left[i\frac{2\pi}{\lambda}(e_n Z_1 \right.\right.\right. \\ &\quad \left.\left.\left. + t_{xn} Z_2 + t_{yn} Z_3)\right]\right) e^{i\phi_{dk}}\right\}\right|^2, \end{aligned} \tag{3}$$

where $h_k(u,v)$ denotes the coherent impulse response function and ϕ_{dk} is the known defocused aberration introduced by the kth defocused length, which is defined as Eq. (4),

$$\phi_{dk}^{\mathrm{PV}} = \begin{cases} 0 & k = 1, \\ \frac{2\pi}{\lambda}\frac{\Delta Z}{8(F^{\#})^2} & k = 2, \end{cases} \tag{4}$$

where PV stands for the peak-to-valley value and ΔZ is the known defocus length.

In order to estimate the coefficients of the Zernike aberrations and the unknown object, a cost function, presented in Eq. (5), is used to fit the data such that the total root mean square (RMS) difference between the images that are actually collected in the different channels and those assumed by the imaging mode based on the least squares approach is

$$L = \frac{1}{2}\left(\sum_{k=1}^{K} \|d_k(u,v) - o(u,v) * s_k(u,v)\|^2\right) + \frac{\gamma}{2}\|o\|^2, \tag{5}$$

where $\|\bullet\|^2$ represents the norm squared, the adding constraint $\frac{\gamma}{2}\|o\|^2$, called the regularization function, that simply fixes the norm of the object estimate, thus preventing wide oscillations that otherwise might occur due to noise amplification, and γ is a nonnegative regularization parameter.

According to the convolution theorem and the Parseval theorem, the cost function can be rewritten in the frequency domain as Eq. (6),

$$L = \frac{1}{2}\left(\sum_{k=1}^{K} \|D_k(f_u, f_v) - O(f_u, f_v)S_k(f_u, f_v)\|^2\right) + \frac{\gamma}{2}\|O\|^2, \quad (6)$$

where D_k, O, and S_k are discrete Fourier transforms of d_k, o, and s_k, respectively, and (f_u, f_v) is the coordinate vector in the frequency domain. Recall that S_k has the role of the optical transfer function (OTF) of the kth image plane.

In order to reduce the dimensions of the parameter space over which a numerical optimization is performed, the partial differential of the cost function L with respect to the object frequency spectrum O is set to zero, and then the regularized object estimate is obtained as shown below:

$$O = \frac{\sum_{k=1}^{K} D_k S_k^*}{\sum_{k=1}^{K} |S_k|^2 + \gamma}. \quad (7)$$

Substituting this expression into Eq. (6) yields the new cost function shown in Eq. (8) that does not explicitly rely on the object estimate:

$$L = \sum_{f\in\chi}\sum_{k} |D_k|^2 - \sum_{f\in\chi} \frac{\left|\sum_k D_k S_k^*\right|^2}{\sum_k |S_k|^2 + \gamma}. \quad (8)$$

Thus, the problem of reconstructing aberration coefficients of the wavefront is transferred to optimizing the cost function, namely to finding the coefficient set for which the cost function [Eq. (8)] is a global minimum.

B. Modified BFGS Method

A nonlinear optimization routine is used to minimize this cost function with respect to the unknown parameters. The BFGS method is commonly used due to its fast convergence and fine stability. The conventional BFGS method [12] has a good global convergence for convex functions; however, when the cost function is nonconvex, the solution will fall into local minimums, which leads to the failure of reconstructing the wavefront and recovering the object. This paper presents a modified BFGS method based on paper [13]; it has a good global convergence for nonconvex unconstrained optimization problems, which guarantees that the solution will converge to the global minimum, corresponding to the actual aberration coefficients. The modified BFGS method presented in this paper is shown below.

1. Make initial estimates for the aberration coefficient set $x_0 \in R^n$ and give the accuracy threshold ε. Set $B_0 = I$ and $k = 0$, where I is the identity matrix and k is the number of iterations.
2. Determine the search direction d_k by solving the linear equation [Eq. (9)],

$$B_k d_k + g_k = 0, \quad (9)$$

where B_k denotes the approximation to the Hessian matrix and g_k is the gradient vector of the cost function. Thus, $d_k = -B_k^{-1} \cdot g_k$.

3. Utilize the Armijo inexact line search, given by Eq. (10), to determine the search stepsize λ_k,

$$f(x_k + \lambda_k d_k) \le f(x_k) + \sigma\lambda_k g_k^T d_k, \quad (10)$$

where x_k is the coefficient set of the kth iteration, $f(\bullet)$ is the cost function, σ is a positive constant, and T denotes the transposition of the matrix.
4. Make $s_k = \lambda_k d_k$, then $x_{k+1} := x_k + s_k$.
5. If $\|g_{k+1}\| < \varepsilon$, terminate the algorithm.
6. Calculate y_k according to Eq. (11):

$$y_k = (g_{k+1} - g_k) + \mu[(g_{k+1} + g_k)^T s_k - 2(f(x_{k+1}) - f(x_k))]s_k/\|s_k\|^2. \quad (11)$$

The second term on the right side is the amendment term, where μ is a positive correction parameter.
7. Modify B_k according to the following correction principle:
if $y_k^T s_k/\|s_k\|^2 \ge c\|g_{k+1}\|^\alpha$,

$$B_{k+1} = B_k - B_k s_k^T B_k/(s_k^T B_k s_k) + y_k y_k^T/(y_k^T s_k), \quad (12)$$

else $B_{k+1} = B_k$; where c and α are both correction parameters.
8. Let $k := k + 1$, go to step 2 until the number of iterations k reaches the set value.

In this paper, the relevant parameters used in the modified BFGS method are : $\sigma = 0.8$; if $y_k^T s_k > 0$, $c = 10^{-6}$, else $c = 1$; if $\|g_{k+1}\| \ge 1$, $\alpha = 0.01$, else $\alpha = 2$; $\mu = 3$; the accuracy threshold $\varepsilon = 10^{-6}$, and the maximum number of iterations is set to 500.

C. Estimate Strategy of the Regularization Parameter

In order to improve the realism of the simulations, the modified algorithm is verified under different noise levels. However, the noise amplification problem must be overcome to remove its influence on the accuracy of the algorithm. One of the most effective techniques used to overcome this problem is the regularization technique.

Many strategies have been developed for the selection of the regularization parameter, such as the cone filter method [14], the constrained least squares (CLS) approach [15], the marginal estimator method [16], the joint maximum *a posteriori* metric, and Wiener filter regularization [17]. The cone filter method needs to choose the cutoff frequency that is related to the signal-to-noise ratio (SNR) of the detected images. The CLS approach needs to know the detector noise level, and the other three of the regularized metrics require *a priori* knowledge of both the object and the noise power spectra. However, both the object and the detector noise level cannot be known in the PD algorithm. Thus, how to accurately estimate these two parameters has become a key issue to overcome noise amplification effects. Some researchers assume the ratio of these two parameters is a constant, even though the power spectrum typically varies from low to high spatial frequencies by orders of magnitude; some researchers estimate the object and noise power spectra by constructing other cost functions [16], which is much overhead. We propose a more feasible and efficient way to estimate the object and noise power spectra based on the Wiener filter regularization method, which makes full use of

the information of these two parameters to overcome the noise amplification problem.

The Wiener filter regularization method defines the kth frame of a multiframe Wiener filter as

$$W_k(f_u, f_v) = \frac{S_k^*(f_u, f_v)}{c\frac{\Psi_N}{\Psi_o(f_u, f_v)} + \sum_{m=1}^{K} |S_m(f_u, f_v)|^2}, \quad (13)$$

where Ψ_N and $\Psi_o(f_u, f_v)$ are the noise and object power spectra, respectively, c is a constant that can be used to "tune" the filter to emphasize either image sharpening or noise suppression. Thus, the regularized object estimate function and cost function can be rewritten as Eqs. (14) and (15):

$$O = \frac{\sum_{k=1}^{K} D_k(f_u, f_v) S_k^*(f_u, f_v)}{c\frac{\Psi_N}{\Psi_o(f_u f_v)} + \sum_{m=1}^{K} |S_m(f_u, f_v)|^2}, \quad (14)$$

$$L = \sum_{f_u, f_v \in \chi} \sum_{k=1}^{K} |D_k(f_u, f_v)|^2 - \sum_{f_u, f_v \in \chi} \frac{\left| \sum_{j=1}^{K} D_j(f_u, f_v) S_j^*(f_u, f_v) \right|^2}{\sum_{l=1}^{K} |S_l(f_u, f_v)|^2 + c\frac{\Psi_N}{\Psi_o(f_u, f_v)}}. \quad (15)$$

We first take the ratio of the noise power spectrum and the object power spectrum as a constant set by prior experience as the regularization parameter. Then, an approximate estimate of the power spectrum of the object can be gained through minimizing the regularized cost function. Meanwhile, the noise power spectrum can also be estimated based on the estimated object power spectrum through Eq. (16),

$$\hat{\sigma}_{nk}^2 = \frac{1}{MN} \sum_{u,v} [d_k(u, v) - \hat{o}(u, v) * s_k(u, v)]^2, \quad (16)$$

where $\hat{\sigma}_{nk}^2$ is the estimated noise variance of the kth detected image, $\hat{o}$ is the estimated object, and M and N are the number of rows and columns in the array. Thus, a more accurate regularized matrix can be estimated by the ratio of the noise and object power spectra, which are ultimately substituted in Eq. (15) to reconstruct the coefficients of the co-phase errors and restore the object. This method takes full advantage of the knowledge of noise and object power spectra and the only expense is to estimate the power spectrum of the object, which can be realized by minimizing the existing Eq. (8) instead of constructing other cost functions and solving the minimum optima of them, so it is more feasible and efficient.

3. NUMERICAL SIMULATION

In this section, a number of vast numerical simulations are processed to verify the effectiveness and accuracy of the modified algorithm by measuring the co-phase errors of the segmented primary mirror and recovering the unknown object under different noise levels. Contrast experiments with the conventional BFGS method are also processed.

First, the relevant parameters of the segmented active optics simulation system are as follows: the primary mirror consists of 6 hexagon submirrors; the effective apertures of the primary mirror D and submirror d are 4 m and 1.46 m, respectively; the focal length is 40 m, thus the $F^{\#}$ of the optical system is 10.

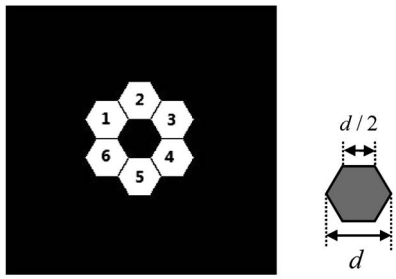


Fig. 2. Construction of the primary mirror and dimensions of the segmented subaperture.

The monochromatic wavelength is 570 nm and the defocused length is set to 600λ. The construction of the segmented primary mirror is shown in Fig. 2; the mark number is the index of subapertures.

In order to test the provided methods, 10 different images are used as observed objects in the vast number of experiments. Only five typical objects are listed in Fig. 3, for brevity. Figure 3(a) stands for point source objects, such as stars, to be observed; Figs. 3(b) and 3(c) represent objects only occupying part of the CCD target surface; specifically, the object is located in the middle of the target surface and is distinct from the background; Figs. 3(e) and 3(f) are on behalf of the panoramic images. The size of all of the tested images is 256 pixels × 256 pixels.

In order to verify the algorithm under different noise levels, we set the image peak pixel to values between 150 e^- and 3000 e^- to get five different peak pixel SNRs excluding the noise-free situation. Intensity-dependent Poisson-distributed photon noise was then added along with zero-mean, additive Gaussian CCD readout noise with a standard deviation of 15 e^-. A dark current of 0.1 e^-/s was assumed over a 1 s integration time. The peak pixel SNR is defined as

$$\mathrm{SNR} = 20 \log_{10} \frac{P(u, v)}{\sqrt{\delta_{\mathrm{read}}^2 + \delta_{\mathrm{dark}}^2}}, \quad (17)$$

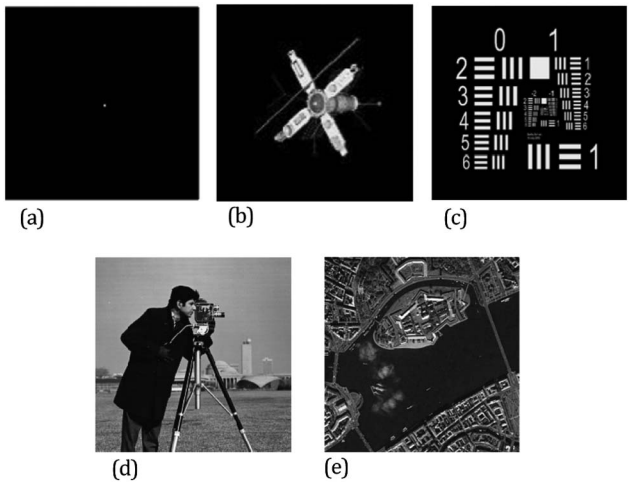


Fig. 3. Five typical objects to be observed: (a) point source, (b) marine satellite, (c) resolution test panel, (d) camera man, and (e) satellite map of an urban scene.

Table 1. Number of Photons in Peak Pixel and Corresponding SNRs

Noise added: photon noise, readout noise, and dark current noise					
Peak pixel photons (e^-)	150	300	750	1500	3,000
Peak pixel SNR	19	24	32	40	47

where $P(u, v)$ is the peak pixel value of the noise-free image, δ^2_{read} and δ^2_{dark} are the variances associated with the readout noise and the dark current noise, respectively. Table 1 shows the peak pixel photon count and the corresponding SNRs.

We first use the urban scene in Fig. 3(e) to illustrate the simulation process, then give the statistics results of all experiments.

A distorted segmented active optics system with aberration is generated by a set of random piston and tip-tilt errors, shown in Fig. 4. The corresponding degraded infocus and out-of-focus images after this optical system under a noise level of 40 peak pixel SNR are shown in Fig. 5.

According to the presented algorithm, the regularization parameter is set to 10^{-5} and substituted to Eq. (8) first. This constant is set empirically and validated by comparing the values of the cost function without and with noise with respect to one term of the aberration coefficients shown in Fig. 6. It shows that the shape of the cost function and the global minimum basically have no changes for both conditions, which proves the effectiveness of this constant during the optimization process. Then, minimizing the regularized cost function by the modified BFGS method, an approximate estimation of the object is obtained and shown in Fig. 7. The noise power spectrum is estimated by Eq. (16) at the same time. Substitution of the ratio matrix of the noise power spectrum and object power spectrum back into Eq. (15) and setting the constant $c = 1$, the coefficients of the co-phase aberrations of the segmented primary mirror are reconstructed and the object is restored by solving the minimum optima of the cost function through the modified BFGS method.

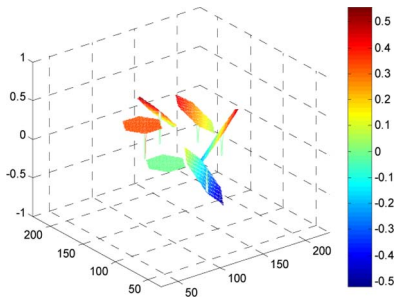
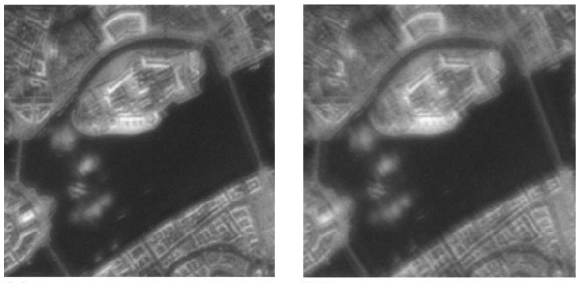
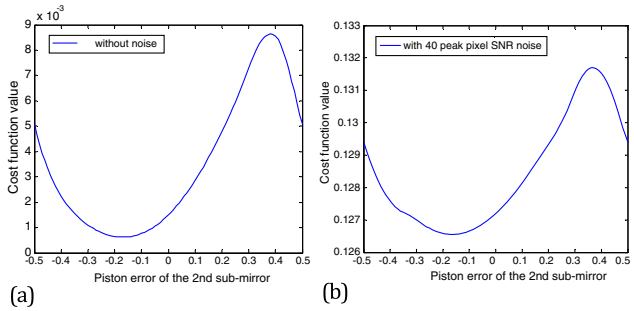


Fig. 4. Phase distribution of the original distorted wavefront.



(a) (b)

Fig. 5. Degraded (a) infocus and (b) out-of-focus images under a 40 peak pixel SNR noise level.



(a) (b)

Fig. 6. Plots of the cost function with respect to one term of aberration coefficients (a) without and (b) with noise when $\gamma = 10^{-5}$.

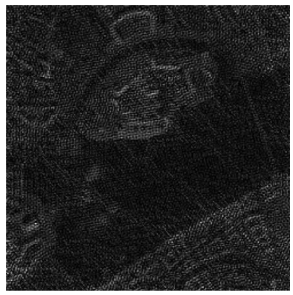


Fig. 7. Restored object using the regularization parameter 10^{-5}.

Table 2 gives the reconstructed coefficients of the co-phase errors and corresponding residual errors. Figure 8 shows the reconstructed phase distribution and residual phase distribution. Figure 9 gives the recovered object.

In order to evaluate the experimental results, the root mean square error (RMSE) on the phase and on the object are defined as Eqs. (18) and (19), respectively. A smaller RMSE value indicates a higher wavefront detection accuracy and a better image quality,

$$\text{RMSE}_\phi = \left[\frac{\sum_{k=1}^{K} (\hat{e}_k - e_k)^2 + (\hat{t}_{xk} - t_{xk})^2 + (\hat{t}_{yk} - t_{yk})^2}{3 \times K}\right]^{1/2}, \tag{18}$$

$$\text{RMSE}_o = \left\{\frac{\sum_{i=1}^{M}\sum_{j=1}^{N} [\hat{o}(i,j) - o(i,j)]^2}{\sum_{i=1}^{M}\sum_{j=1}^{N} o(i,j)^2}\right\}^{1/2}, \tag{19}$$

where $\hat{e}_k$, $\hat{t}_{xk}$, and $\hat{t}_{yk}$ are the reconstructed coefficients of the co-phase errors of the kth submirror, respectively, K is total number of submirrors, $\hat{o}$ is the recovered object, and M and N are the number of rows and columns in the array.

Table 2. Reconstructed Coefficients of the Co-phase Errors and Residual Errors

	Pupil 1	Pupil 2	Pupil 3	Pupil 4	Pupil 5	Pupil 6
Actual coefficients of co-phase errors (λ)						
Piston	0	−0.19	0.13	0.25	0.29	0.33
Tip	0	0.25	−0.3	0.14	0	0
Tilt	0	−0.23	0.28	−0.17	−0.27	0
Reconstructed coefficients of the co-phase errors (λ), 40 SNR noise						
Piston	0	−0.1902	0.1283	0.2527	0.2936	0.3285
Tip	0	0.2502	−0.2997	0.1376	1.3350e − 4	−0.0011
Tilt	0	−0.2314	0.2815	−0.1690	−0.2677	8.9277e − 4
Corresponding residual errors (λ)						
Piston	0	0.0002	0.0017	−0.0027	−0.0036	0.0015
Tip	0	−0.0002	−0.0003	0.0024	−0.0001	0.0011
Tilt	0	0.0014	−0.0015	−0.0010	−0.0023	−0.0009

From Table 2, the RMSE on the phase estimates is 1.7e − 3λ. The RMSE on the objects of Figs. 7 and 9 are 64.42% and 23.96%, respectively. It can be seen that the presented algorithm can measure the co-phasing errors with a high accuracy and the residual errors are totally within the acceptable range for the co-phasing of the segmented mirrors. A comparison of Figs. 7 and 9 indicates that the quality of the restored image obtained by using our proposed estimated regularization parameter is far better than that by using the regularization parameter as a constant.

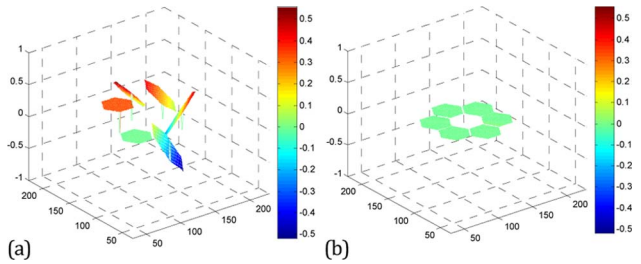


Fig. 8. (a) Reconstructed phase distribution and (b) residual phase distribution.

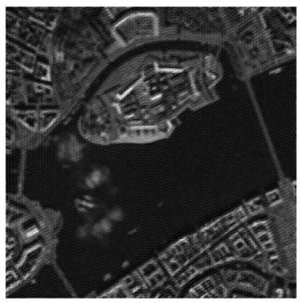


Fig. 9. Final recovered object image.

As in the example above, this paper processed simulation experiments for the 10 objects under the five different noise levels listed in Table 1 and in a noise-free situation. For each object under one certain noise level, 50 sets of random co-phase errors restricted to within $\pm 0.5\lambda$, and a corresponding distorted phase limited to 2π rad, are tested to detect the co-phase errors and restore the object by using the modified BFGS method. The statistics results of the 3000 experiments are summarized below.

Table 3 exhibits the effectiveness and accuracy of the proposed methods in measuring the co-phase errors of the segmented primary mirror. The content of the table gives the ratio of the effective reconstructions of the co-phase error coefficients, where the RMSE on the phase of $\leq 0.01\lambda$ are considered as effective reconstructions.

Table 3 shows that the effective reconstructions of the co-phase error coefficients are maintained more than 80% when the noise level reaches a 32 peak pixel SNR. However, the effective reconstructions ratio drops rapidly when the noise level reaches a 19 peak pixel SNR.

Figure 10 shows the effectiveness and accuracy of the presented methods in recovering the image for the system when using the marine satellite in Fig. 3(b) as the observed object. It

Table 3. Statistics Results of the Reconstruction Coefficients of the Co-phase Errors

		$\max\{\lvert e_1\rvert,\lvert t_{x1}\rvert,\lvert t_{y1}\rvert,\ldots\lvert e_k\rvert,\lvert t_{xk}\rvert,\lvert t_{yk}\rvert,\ldots\lvert e_6\rvert,\lvert t_{x6}\rvert,\lvert t_{y6}\rvert\}$				
Effective Reconstructions Ratio		$\leq 0.3\lambda$	$\leq 0.35\lambda$	$\leq 0.4\lambda$	$\leq 0.45\lambda$	$\leq 0.5\lambda$
Peak pixel SNR	Noise-free	100%	100%	100%	100%	100%
	47	100%	100%	100%	100%	100%
	40	100%	100%	99%	97%	96%
	32	91%	89%	86%	85%	83%
	24	79%	77%	74%	72%	69%
	19	66%	61%	57%	53%	51%

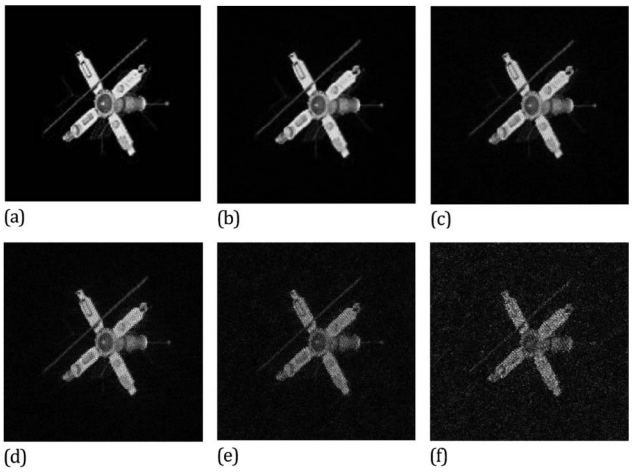


Fig. 10. Recovered images of the marine satellite object under different noise levels: (a) noise free, (b) 47 SNR, (c) 40 SNR, (d) 32 SNR, (e) 24 SNR, and (f) 19 SNR.

exhibits the recovered images of the marine satellite object under 6 different noise levels. The RMSE values on the six objects are 5.21%, 9.03%, 15.66%, 26.46%, 37.09%, and 49.78%, respectively. For the visual effect, it can be seen that as the noise level increases, the restored images have a lower contrast and more blurred edges, but the details of the restored images can still be distinguished even under a very high noise level.

In order to contrast the modified BFGS algorithm and the conventional BFGS method, we rerun the 3000 experiments using the conventional BFGS method to minimize the PD cost function under the exact same conditions. The contrast statistics results are summarized as below.

Figure 11 gives the statistics results of a set of 2500 RMSE values on the phase estimates by using a modified BFGS method and a corresponding set of 2500 RMSE values using the conventional BFGS method under five different noise levels in the form of error bars. The extra-large RMSE values on the phase estimates under different noise levels when using conventional BFGS method are caused by nonconvex conditions. It can be seen that the modified BFGS effectively promoted the nonconvex PD cost function out of local optima and converged to the global optimum even under very high noise level.

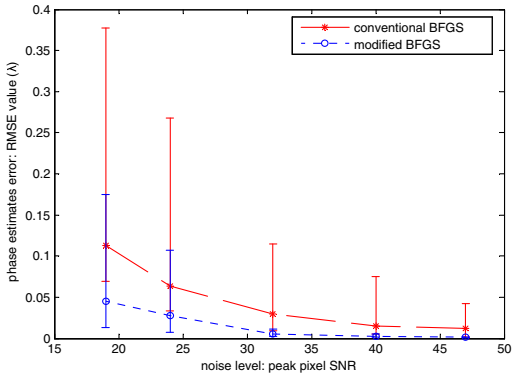


Fig. 11. Contrast error bars of the set of 2500 RMSE values on the phase estimates using the modified BFGS and the corresponding 2500 set using conventional BFGS under five noise levels.

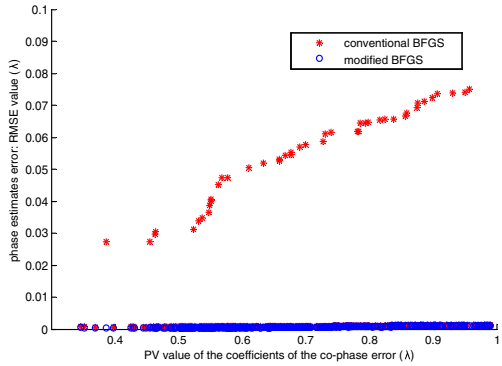


Fig. 12. Contrast scatterplots of the set of 500 RMSE values on phase estimates using the modified BFGS and the corresponding 500 set using conventional BFGS under a noise-free situation.

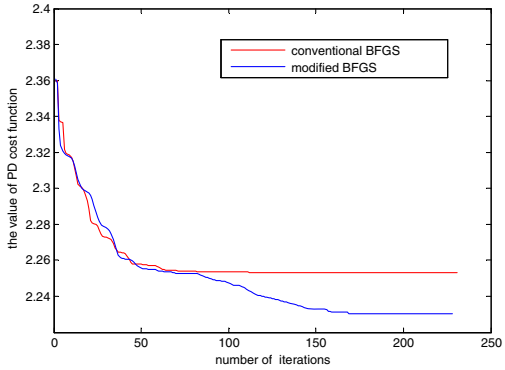


Fig. 13. Contrast plots of the convergence curves of the PD cost function using the modified BFGS and conventional BFGS for a nonconvex situation under a 32 peak pixel SNR noise level.

Figure 12 represents the scatterplots of a set of 500 RMSE values on the phase estimates by using the modified BFGS method and a corresponding set of 500 RMSE values using the conventional BFGS method under a totally noise-free situation with respect to the PV values of the co-phase error coefficients. It exhibits the accuracy of the modified algorithm aimed at nonconvex conditions more purely since it excluded the noise influence. The RMSE value of the residual co-phase errors using the modified BFGS method is around 10^{-4} orders of magnitude under the noise-free situation while that using the conventional BFGS method almost approaches 0.1λ for nonconvex conditions.

Figure 13 presents the convergence curves of the PD cost function using the modified BFGS and conventional BFGS for a nonconvex situation. It can be seen that, for a nonconvex situation, the conventional BFGS is stuck at a local optimum while the modified BFGS converges to the global optimum, corresponding to the true coefficients of the co-phase errors.

4. CONCLUSION

In this paper, we presented a modified BFGS method aimed at the nonconvex cost functions. It has fine global convergences for both convex and nonconvex functions and guarantees that

the solutions will converge to the global minimum, corresponding to the actual coefficients of the co-phase errors. We applied it to minimize the PD cost functions for co-phasing the segmented active optics system and recovering the unknown object, where different noise levels were included to improve the realism of the simulations. In order to overcome the noise amplification effect on the accuracy of the modified algorithm, we also provided an efficient estimate strategy of the regularization parameter for the PD problem. Large numbers of contrast experiments about different objects degraded by different optical aberrations under different noise levels were processed using the modified BFGS method and conventional BFGS method. The statistics results show that the modified BFGS method can effectively promote the nonconvex PD cost functions out of the local optima and converge to the global optimum even under a very high noise level, and thus has a much higher accuracy than the conventional BFGS method for nonconvex situations. For different objects under different system aberrations, when the peak pixel SNR is more than or equal to 32, the accuracy of the reconstruction of the aberration coefficients can reach 1% of the wavelength in more than 80% probability using the modified algorithm for both convex and nonconvex conditions. The future work will be concentrated on establishing an optical platform in the lab to verify the work presented in this paper.

Funding. National Natural Science Foundation of China (NSFC) (61205143).

REFERENCES

1. P. A. Lightsey, C. Atkinson, M. Clampin, and L. D. Feinberg, "James Webb Space telescope: large deployable cryogenic telescope in space," Opt. Eng. **51**, 011003 (2012).
2. G. Chanan, C. Ohara, and M. Troy, "Phasing the mirror segments of the Keck telescopes II: the narrow-band phasing algorithm," Appl. Opt. **39**, 4706–4714 (2000).
3. G. A. Chanan, "Design of the Keck Observatory alignment camera," Proc. SPIE **1036**, 59–71 (1989).
4. N. Yaitskova, K. Dohlen, P. Dierickx, and L. Montoya, "Mach–Zehnder interferometer for piston and tip–tilt sensing in segmented telescopes: theory and analytical treatment," J. Opt. Soc. Am. A **22**, 1093–1105 (2005).
5. S. Esposito, E. Pinna, A. Puglisi, A. Tozzi, and P. Stefanini, "Pyramid sensor for segmented mirror alignment," Opt. Lett. **30**, 2572–2574 (2005).
6. V. G. Orlov, S. Cuevas, F. Garfias, V. V. Voitsekhovich, and L. J. Sanchez, "Co-phasing of segmented mirror telescopes with curvature sensing," Proc. SPIE **4004**, 540–551 (2000).
7. I. Surdej, N. Yaitskova, and F. Gonte, "On-sky performance of the Zernike phase contrast sensor for the phasing of segmented telescopes," Appl. Opt. **49**, 4052–4062 (2010).
8. R. L. Kendrick, D. S. Acton, and A. L. Duncan, "Phase-diversity wave-front sensor for imaging systems," Appl. Opt. **33**, 6533–6546 (1994).
9. M. G. Löfdahl, R. L. Kendrick, A. Harwit, K. E. Mitchell, A. L. Duncan, J. H. Seldin, R. G. Paxman, and D. S. Acton, "Phase diversity experiment to measure piston misalignment on the segmented primary mirror of the Keck II telescope," Proc. SPIE **3356**, 190–201 (1998).
10. J. Cea and M. K. V. Murthy, *Optimization: Theory and Algorithms* (Springer-Verlag, 1978).
11. J. E. Nelson, T. S. Mast, and S. M. Faber, "The design of the Keck Observatory and telescope," Keck Observatory Report 90 (1985).
12. J. Nocedal and S. J. Wright, *Numerical Optimization* (Springer-Verlag, 1999), pp. 194–201.
13. D.-H. Li and M. Fukushima, "On the global convergence of the BFGS Method for nonconvex unconstrained optimization problems," SIAM J. Optim. **11**, 1054–1064 (2001).
14. D. J. Lee, "Evaluation and application of space telescope aberration sensing using phase diversity," Ph.D. dissertation, No. AFIT/DS/ENP/97-05 (Air Force Institute of Technology, 1997).
15. R. G. Paxman and J. H. Seldin, "Focal-plane alignment sensing," Final Report (Phillips Laboratory, 1993).
16. A. Blanc, L. M. Mugnier, and J. Idier, "Marginal estimation of aberrations and image restoration by use of phase diversity," J. Opt. Soc. Am. A **20**, 1035–1045 (2003).
17. M. R. Bolcar, *Phase Diversity for Segmented and Multi-Aperture Systems* (University of Rochester, 2008).