






# A Novel SPGD Algorithm for Wavefront Sensorless Adaptive Optics System

Jiaxun Li , Lianghua Wen , Hankui Liu , Guiming Wei, Xiang Cheng, Qing Li , and Bing Ran 

**Abstract**—Stochastic parallel gradient descent (SPGD) is the most frequently used optimization algorithm for correcting wavefront distortion in the wavefront sensorless adaptive optics (WFS-Less AO) system. However, the convergence speed of the SPGD algorithm becomes slow rapidly as increasing of distortion, and the probability of falling into local optimum is rising owing to the fixed gain coefficient. It cannot meet the requirement of real-time wavefront distortion correction. Therefore, a novel algorithm is proposed in this paper, called as adaptive gain stochastic parallel gradient descent (AGSPGD) based on the AMSGrad optimizer in the deep learning, to improve the convergence speed of the algorithm and to reduce the probability of falling into local optimum. The AGSPGD algorithm adopts the first-order moment and the second-order moment of the performance index, which are combined to dynamically adjust the gain. The numerical simulations are completed in this article. The results of  $D/r_0 = 2.5$  conditions demonstrate that the AGSPGD can reduce the number of iterations by 25%, and the probability of the algorithm falling into local optimum is reduced from 16% to 4%. In addition, the AGSPGD still outperforms the SPGD as  $D/r_0$  increasing.

**Index Terms**—Adaptive gain, convergence speed, deep learning, stochastic parallel gradient descent, wavefront sensorless.

## I. INTRODUCTION

ADAPTIVE optics (AO) originated in astronomy is a technology developed in 1980s. Adaptive optics system (AOS) is divided into two types: the adaptive optics system with wavefront sensor and the wavefront sensorless (WFS-Less) adaptive optics system. The WFS-Less AOS has the advantages of low cost, simple structure and convenient construction, which is widely used in laser nuclear fusion device, confocal microscope, fiber coupling, laser phase control, optical clamp, optical tracking, extended target imaging, etc. [1], [2], [3], [4], [5], [6]. When the optical system is determined, the performance of the

WFS-Less AOS mainly depends on the optimization algorithm adopted by the system.

The optimization algorithm for WFS-Less AOS is generally divided into two types: model-free and model-based, or called search based blind optimization algorithms and model-based optimization algorithms. The model-based correction algorithm converges quickly, but it has strict requirements on pupil function, and the formula derivation is very complicated, which limits its application in practical adaptive optics systems. In contrast, the realization principle of model-free algorithm is simpler and less restrictive for optical system, so it is very suitable for WFS-Less AO. Nowadays, the search-based blind optimization (model-free) algorithm mainly includes genetic algorithm (GA), simulated annealing (SA), and stochastic parallel gradient descent (SPGD). The SPGD algorithm was first applied to adaptive optics in 1997 [7], which brought up many aspects of the application [8], [9], [10], [11], [12], [13], [14], [15], [16]. The SPGD algorithm has been the focus of research in recent years owing to its simple implementation and strong comprehensive correction ability [17]. However, the convergence speed of the SPGD algorithm is slow or the result of SPGD algorithm is easy to fall into local optimum, which leads to the low correction performance, when the search space of control variables increases with the increases of the number of correction units or aberration modes [18]. In recent years, researchers have been improving the SPGD algorithm, the decoupled SPGD (DSPGD) algorithm was proposed to improve the coupling efficiency of atmospheric phase aberration compensation for tiled fiber array system [19]. A method that combines pattern recognition and the SPGD algorithm was proposed to avoid the circumstance of the local optimum [20]. Hu et al. proposed an optimization algorithm for ASPGD by combining the Adam optimizer in the field of deep learning with SPGD algorithm, which introduce the momentum to make the gain coefficient be self-adaptive. This algorithm speeds up the convergence speed and reduces the probability of times it falls into local optimum [21]. The momentum SPGD (MSPGD) algorithm was proposed to accelerate the correction process of incoherent beam combination, which not only decreases the iterative times, but also maintains the stability of the combination [22]. The MomSPGD algorithm was proposed, by deriving momentum terms from Newton's equation, to improve the convergence speed and disturbance immunity of coherent beam combinations [23].

To solve the problem that the convergence speed of the SPGD algorithm is slow or the result is easy to fall into local optimum, an optimal optimization algorithm called adaptive

Manuscript received 15 May 2023; revised 7 June 2023; accepted 11 June 2023. Date of publication 13 June 2023; date of current version 30 June 2023. This work was supported by the National Natural Science Foundation of China under Grant 61975171. (Corresponding authors: Lianghua Wen.)

Jiaxun Li and Lianghua Wen are with the School of Electronic Information Engineering, China West Normal University, Sichuan 637000, China, and also with the School of Faculty of Intelligence Manufacturing, Yibin University, Sichuan 644000, China (e-mail: 13882912635@163.com; wlh45@126.com).

Hankui Liu is with the School of Electronic Information Engineering, China West Normal University, Sichuan 637000, China (e-mail: liuhankui@cwnu.edu.cn).

Guiming Wei, Xiang Cheng, and Qing Li are with the School of Faculty of Intelligence Manufacturing, Yibin University, Sichuan 644000, China (e-mail: 18283819726@163.com; 13778959157@163.com; qiou9@163.com).

Bing Ran is with the Ordnance NCO Academy, Army Engineering University, Wuhan 430075, China (e-mail: rnbiger@126.com).

Digital Object Identifier 10.1109/JPHOT.2023.3285871

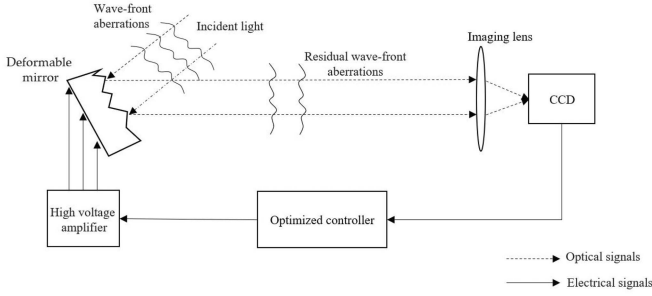


Fig. 1. Schematic of WFS-Less AO system. The optical signals are drawn as dotted lines, and the electrical signals are drawn as solid lines.

gain stochastic parallel gradient descent (AGSPGD) is proposed in this article. The AMSGrad optimizer for deep learning is integrated into the SPGD algorithm, the variation of image performance index is approximated to the gradient. The mean of the gradient is defined as the first momentum term, and the variance of the gradient is considered as the second momentum term. Then the direction and step length of the gradient descent are controlled by first momentum term and second momentum term. In addition, a single loop structure and vector parallel operation is applied to the AGSPGD algorithm, to reduce the sampling times of the far-field camera, and to speed up furtherly the system correction. It is more significant for the real-time and miniaturization, lightweight of the WFS-Less AOS.

This article is organized as follows: Section II provides the model of the WFS-Less AO system and briefly describes the working principles. Then, the existing problems of the SPGD algorithm are introduced and analyzed. A novel hybrid algorithm, AGSPGD is proposed based on the SPGD algorithm and the principle of the AMSGrad optimizer. In Section III, the simulations and the analysis are completed. In Section IV, the simulation results are compared and the comparison results are analyzed through the table. Finally, the conclusion is given in Section V, and the future work is briefly introduced.

## II. SYSTEM MODEL AND THEORETICAL ANALYSIS

In this section, the wavefront sensorless adaptive optics system model is presented to facilitate analysis.

### A. The WFS-Less AO System Model

The schematic diagram of the WFS-Less AO system is shown in Fig. 1. The WFS-Less AO system consists of the deformable mirror (DM), the imaging lens, the far-field detection CCD camera, the optimal controller, and the high-voltage amplifier. In the propagation of the beam, the wave-front phase of incident light is distorted after passing through inhomogeneous medium. The beam with wave-front distortion is reflected by the DM and is incident by the imaging lens. Then the CCD is used to obtain the far-field spot images of distorted beam. The controller is used to read the far-field spot images and the blind optimization algorithm is adopted to generate control voltage signals according to optimization indexes of far-field spot images, and the control

voltage signals are amplified by a high voltage amplifier to compel the deformable mirror. Therefore, the wavefront distortion is corrected. So, the performance of the optimal control algorithm greatly determines whether the wavefront-less sensing adaptive optical system can achieve the desired correction effect.

Generally speaking, the SPGD algorithm's gain coefficient is usually a fixed value. When the wavefront aberration changes rapidly, the result of the SPGD algorithm is easy to fall into local optimum or convergence time increases.

### B. Theoretical Analysis of AGSPGD Algorithm

To address the above problems of the SPGD algorithm, this article proposes the SPGD algorithm based on the AMSGrad optimizer called AGSPGD algorithm, which integrates the AMSGrad optimizer in deep learning and the SPGD algorithm. Then the variation of the performance index of the spot image is approximated as a gradient, the average value of the gradient is defined as the first momentum term, and the variance of the gradient is considered as the second momentum term. The first momentum term is used to control the direction of the gradient descent. When the direction of the gradient descent is positive, a positive excitation is generated, otherwise, a negative excitation is generated. The second momentum term is used to control the step size of the gradient descent method to complete the adaptive gain coefficient when the wavefront aberration changes. On the other hand, the single cycle structure and vector parallel operation are adopted by the AGSPGD algorithm, which reduces the calculation amount of the optimal controller and improves the operation speed, and furthermore, it also reduces the sampling times of the far-field camera greatly. Therefore, the convergence speed of the WFS-Less AO system is accelerated and the probability of falling into local optimum is reduced. The AGSPGD algorithm flowchart is shown in Fig. 2.

The idea of correcting wavefront aberration in the SPGD algorithm is as follows. In the  $(k-1)$  iteration, the correction voltage vectors  $u(k-1) = \{u_1, u_2, \dots, u_m\}$  are applied to the DM. Before the next iteration, a set of disturbance voltage vectors  $\Delta u(k) = \{\Delta u_1, \Delta u_2, \dots, \Delta u_m\}$  are randomly generated, according to the Bernoulli distribution. And the  $\Delta u(k)$  has a fixed amplitude of disturbance, that is,  $|\Delta u(k)| = \Delta u$ . Then, the forward perturbation voltages  $u^+ = u(k-1) + \Delta u(k)$  are applied to DM to obtain the forward performance index  $J(u^+)$ ; and the negative perturbation voltages  $u^-$  are obtained by a similar method to obtain the negative performance index  $J(u^-)$ . The formula for calculating the performance index is as follows:

$$J = \frac{\iint I(x, y)^2 dx dy}{(\iint I(x, y) dx dy)^2} \quad (1)$$

where the  $I$  is the far-field light intensity.

So, the variation of performance index is expressed as

$$\Delta J = J(u^+) - J(u^-) \quad (2)$$

Then, the correction voltage vectors applied to the DM after the  $k^{\text{th}}$  iteration is expressed as

$$u(k) = u(k-1) + \gamma \Delta J \Delta u, \quad (3)$$

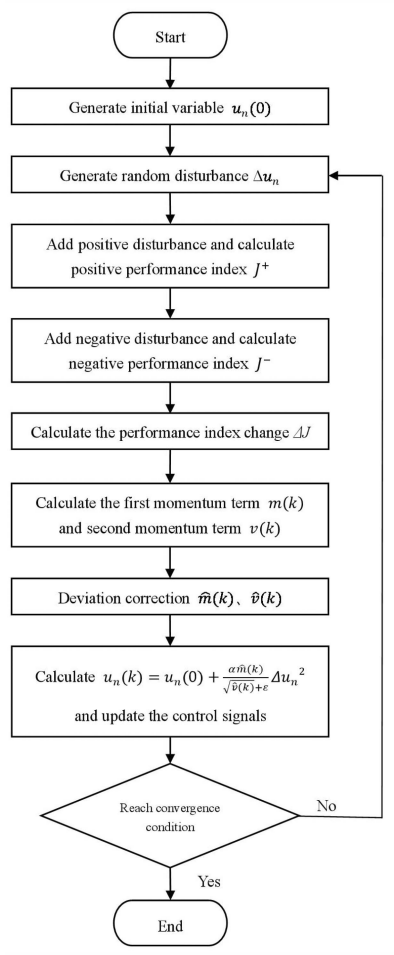


Fig. 2. Flowchart of the AGSPGD algorithm.

where the  $k$  is the number of iterations, and the  $\gamma$  is the gain coefficient, the symbol of the  $\gamma$  is determined by the optimization direction of the performance index. When the performance index is optimized to the maximum direction, the  $\gamma$  takes a positive number, otherwise, it takes a negative number. It is precisely because the  $\gamma$  is a fixed value that the SPGD algorithm is easy to fall into local optimum, or convergence time increases.

So, the AGSPGD algorithm controls the gain coefficient to adjust automatically to speed up the convergence and to reduce the probability of falling into local optimum. In the AGSPGD algorithm, the gradient is approximated as

$$g = \Delta J. \quad (4)$$

First, the momentum is introduced into the SPGD algorithm to speed up the convergence [24], [25]. In the AGSPGD algorithm, the first momentum term  $m(k)$  is constructed by adding a hyper-parameter  $\beta_1$ , and the average value of the gradient is defined as the first momentum term. Then, by following the article [26], a hyper-parameter  $\beta_2$  is added to construct the second momentum term  $v(k)$ , and the variance of the gradient is considered as the  $v(k)$ . The relevant formulas for  $m(k)$  and  $v(k)$  are as follows:

$$m(k) = \beta_1 m(k-1) + (1 - \beta_1) g(k) \Delta u(k), \quad (5)$$

$$v(k) = \beta_2 v(k-1) + (1 - \beta_2) g(k)^2 \Delta u(k)^2. \quad (6)$$

The control signals are updated by a combination of the first momentum term and the second momentum term to control the gain coefficient to adjust automatically. Therefore, the convergence of the algorithm is speeded and the probability of falling into the local optimal is reduced. However, at the beginning of the iteration, the values of the first and second momentum term are smaller than the real mean and variance, then  $\beta_1$  and  $\beta_2$  are close to 1, the error of the algorithm is very large. So, it is necessary to correct the deviation of the first momentum term and the second momentum term [26]. The relevant formulas are expressed as

$$\hat{m}(k) = m(k) / (1 - \beta_1^k), \quad (7)$$

$$\hat{v}(k) = v(k) / (1 - \beta_2^k). \quad (8)$$

Then, the maximum value of all gradients is used to update the learning rate to make the learning rate positive all the time, to further speed up the convergence. That is

$$\hat{V}(k) = \max(\hat{V}(k-1), \hat{v}(k)). \quad (9)$$

As discussed above, using the AGSPGD algorithm to update the control voltage vectors computation formula is as follows:

$$u(k) = u(k-1) + \left[ \alpha \hat{m}(k) / (\hat{V}(k) + \varepsilon) \right] \Delta u(k)^2, \quad (10)$$

where the  $k$  is the number of iterations,  $\alpha / (\hat{V}(k) + \varepsilon)$  is the adaptive gain,  $\alpha$  is the learning rate. The  $\varepsilon$  is a small constant usually set to  $10^{-8}$ , that is used to avoid the denominator of 0.

The gain coefficient  $\gamma$  in (3) is a fixed value, which is the main reason why the result of the SPGD algorithm is easy to fall into local optimum or convergence time increases. From the (10), as the gain coefficient of AGSPGD algorithm, the  $\alpha / (\hat{V}(k) + \varepsilon)$  can realize automatic adjustment. In addition,  $\Delta u(k)^2$  avoids the problem of oscillation in the late iterations of the algorithm.

The implementation of the AGSPGD algorithm is comprehensively described in Algorithm 1.

In theory, the AGSPGD algorithm changes the limitation of the SPGD algorithm on fixed gain coefficient. The momentum is used to control the direction and step size of the gradient descent, so that the convergence is speeded up and the probability of falling into local optimum is reduced. The numerical simulations will be carried out next. The results will verify the correction performance of the AGSPGD algorithm.

### III. SIMULATIONS AND ANALYSIS

In this section, the numerical simulations are performed to verify the effect of the AGSPGD algorithm on static wavefront aberrations correction. Generally, the DM is used to correct the wavefront aberrations. Therefore, the input wavefront is comprised by the Zernike modes. In the numerical simulations, select the aberrations mode with  $m = 10$ .

The wavefront phase distortion is expressed by the weighted sum of the Zernike polynomials and their corresponding coefficients. Therefore, in the following numerical simulations, the Zernike polynomials are used to represent the wavefront phase

**Algorithm 1:** Procedure of AGSPGD Algorithm.

**Input:** The learning rate  $\alpha$ , the hyper-parameters  $\beta_1$  and  $\beta_2$ , the constant  $\varepsilon$ , the amplitude of random perturbation voltages  $\Delta \mathbf{u}$ , and the maximal number of iterations  $N$ .

**Output:** Calculated control voltage vectors

$\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_m$ .

1. Initialize control voltage vectors  $\mathbf{u}_0$ , the first momentum term  $\mathbf{m}(0)$ , the second momentum term  $\mathbf{v}(0)$
2. **for**  $k = 1, 2, \dots, N$  **do**
3. Randomly generate the perturbed voltage obeying the Bernoulli distribution  $\Delta \mathbf{u}(k)$
4. Obtain the evaluation functions under perturbation voltage  $J_{\pm}(k) = J(\mathbf{u}(k-1) \pm \Delta \mathbf{u}(k))$
5. Obtain the change in the evaluation function  $\Delta J(k) = J_+(k) - J_-(k)$
6. Calculate the gradient  $\mathbf{g}(k)$  (see (4))
7. Calculate the bias-corrected first momentum term  $\hat{\mathbf{m}}(k)$  ((5) and (7))
8. Calculate the bias-corrected second momentum term  $\hat{\mathbf{v}}(k)$  (6) and (8))
9. Take the maximum of the gradient  $\hat{\mathbf{v}}(k)$  (9))
10. Update the control voltage  $\mathbf{u}(k)$  (see (10))
11. The SR value reaches the preset condition
12. **end for**

$\varphi_0(\rho, \theta)$ . The  $\varphi_0(\rho, \theta)$  can be expressed as follows [27]:

$$\varphi_0(\rho, \theta) = a_0 + a_1 Z_1(\rho, \theta) + a_2 Z_2(\rho, \theta) + \sum_{m=3}^{\infty} a_m Z_m(\rho, \theta) \quad (11)$$

where  $Z_m(\rho, \theta)$  denotes the  $m^{\text{th}}$  Zernike polynomial and  $a_k$  is the corresponding coefficient. In the Zernike polynomials, the 0<sup>th</sup> is the translation term, it can be ignored.  $Z_1$  and  $Z_2$  represent the tilt terms along X and Y directions, respectively. The tip-tilt mirror (TM) is used to correct these two terms. So, in the simulations, the 3<sup>th</sup> to 12<sup>th</sup> terms in the Zernike polynomials are modeled as the distorted wavefront, to simulate the atmospheric turbulence.

In the simulations, the Strehl Ratio (SR) is used to measure the correction performance of the two algorithms. The SR is defined as the ratio of the far-field peak intensity of an actual beam to the peak intensity of an ideal beam with the same power and uniform phase. The SR can be expressed as

$$SR = \frac{I}{I_0} \quad (12)$$

where  $I$  denote the far-field peak intensity of the actual beam, and  $I_0$  is the peak intensity of the ideal beam.

In order to analyze the feasibility of the AGSPGD algorithm. First, two algorithms are respectively used to correct the same group of randomly generated wavefront aberrations, and then the simulation results are compared. According to Roddier's

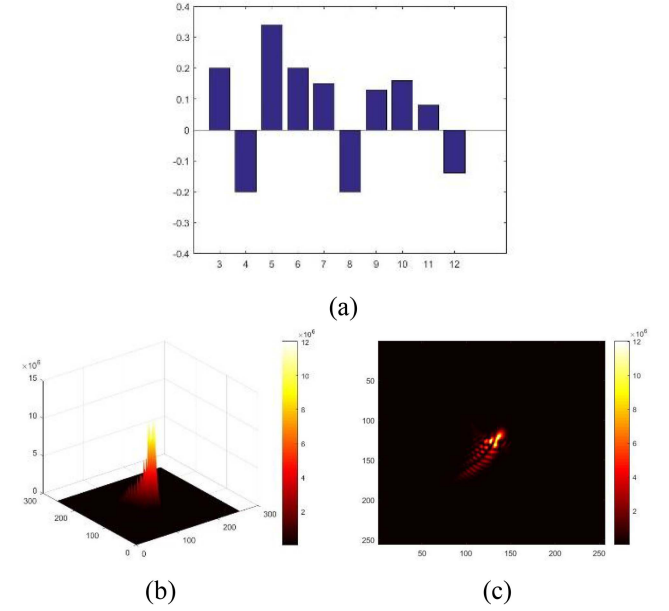


Fig. 3. Random set of simulated data: (a) Initial Zernike coefficients of the wavefront aberrations, (b) 3-D image of wavefront distortion spot, (c) Original PSF.

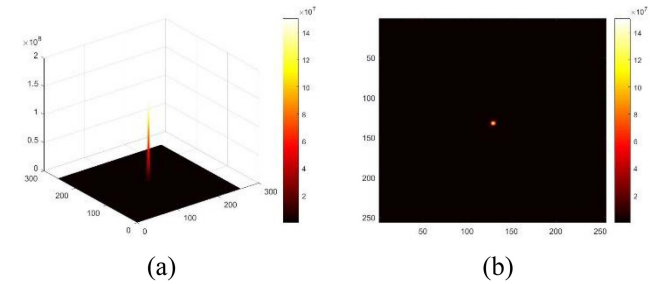


Fig. 4. Wavefront correction results of SPGD: (a) 3D image of corrected far-field spot, (b) PSF after calibration.

method [28], Zernike polynomials with 10 terms of  $a_3 - a_{12}$  are generated randomly, as shown in Fig. 3(a).

Specifically, the SPGD algorithm and the AGSPGD algorithm are used to correct the same group of random wavefront aberrations in the simulations. When the two algorithms were iterated 800 times, the correction effects of the SPGD algorithm are shown in the Fig. 4.

The correction effects of the AGSPGD algorithm are shown in the Fig. 5.

The SR iteration curves for both algorithms are shown in Fig. 6.

In the simulations, the  $SR = 0.9$  is regarded as the convergence condition of the algorithm. The simulation results show that both the SPGD algorithm and the AGSPGD algorithm can meet the convergence condition. The simulation results of the SPGD algorithm show that when the number of iterations is 800, the SPGD algorithm just reaches the convergence condition in Fig. 6(a). Then, in the AGSPGD algorithm simulation, when the number of iterations is 600, the AGSPGD algorithm reaches



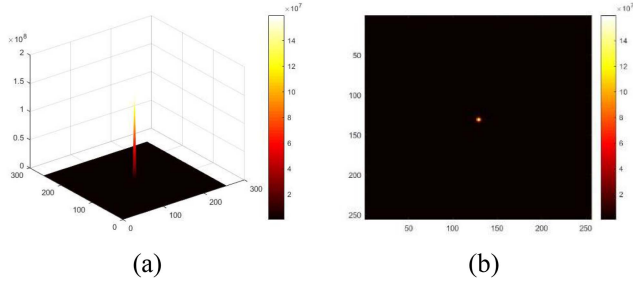


Fig. 5. Wavefront correction results of AGSPGD: (a) 3D image of corrected far-field spot, (b) PSF after calibration.

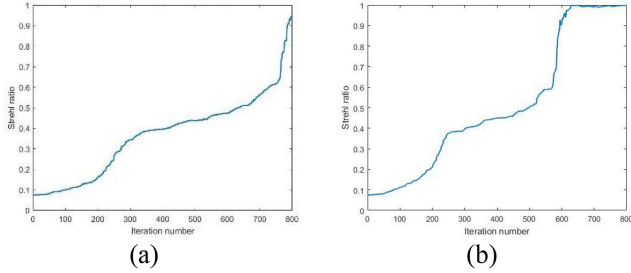


Fig. 6. The SR iteration curves: (a) SPGD algorithm, (b) AGSPGD algorithm.

the convergence condition. So, the AGSPGD algorithm reduces the number of iterations by 25% compared with the SPGD algorithm. Besides, when the number of iterations is 630 in Fig. 6(b), the SR value of the AGSPGD algorithm is close to 1, which indicates that the algorithm has fully converged. Moreover, the 3D images show that the corrected energy is more concentrated, and has increased by at least one order of magnitude. So, the above simulation results show that the AGSPGD algorithm has better correction effect than the SPGD algorithm. Therefore, the AGSPGD algorithm is worthy of further research.

In order to avoid the chance and uncertainty caused by a group of random aberrations, the following simulations are carried out. First, 100 groups of wavefront aberrations with aberration intensity  $D/r_0 = 2.5$  are randomly generated. Then, the SPGD algorithm and the AGSPGD algorithm are used to correct the wavefront aberrations. In the following simulations, the number of iterations is also set to 800, and  $SR = 0.9$  is regarded as algorithm convergence.

The simulation results of 100 groups of wavefront aberrations corrected by the SPGD algorithm are shown in Fig. 7. The red bold line represents the SR average value of 100 groups of wavefront aberrations.

Similarly, the simulation results of 100 groups of wavefront aberrations corrected by the AGSPGD algorithm are shown in Fig. 8. And the red plus bold line also represents the average value of SR.

It can be known from the above 100 groups of simulation results that: first, when the algorithm iteration times is 800, in the Fig. 7, the average value of SR is 0.79. It shows that the result of the SPGD algorithm does not meet the convergence condition. However, in the simulations of the AGSPGD algorithm,

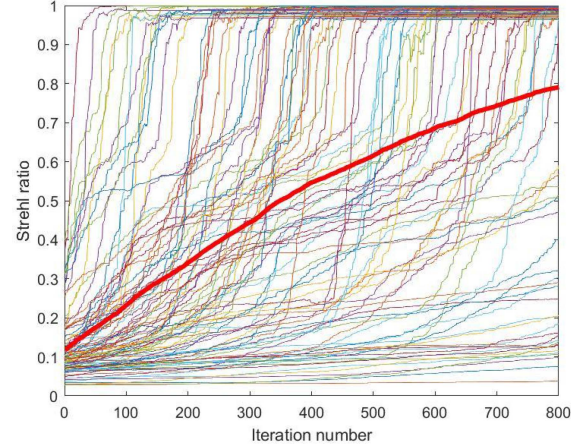


Fig. 7. Simulation results of the SPGD algorithm.

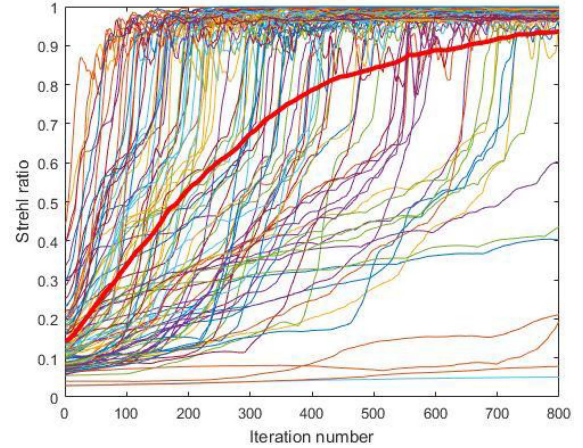


Fig. 8. Simulation results of the AGSPGD algorithm.

the average value of SR is 0.94 in Fig. 8, which reaches the convergence condition. Therefore, the simulation results of 100 groups of random wavefront aberrations show that the AGSPGD algorithm has better correction effect than the SPGD algorithm when the number of iterations is the same.

In addition, if the  $SR < 0.3$ , the result of the algorithm is considered to fall into the local optimum. According to the simulation results, when the SPGD algorithm is used for correction, 16 groups of the 100 wavefront aberrations fall into the local optimum, so the convergence probability is 84%. But only 4 groups fall into local optimum when the AGSPGD algorithm is used to correct, and the convergence probability is 96%. Obviously, the AGSPGD algorithm reduces the probability of falling into the local optimum. The simulation results of 100 groups are consistent with the above results, so the correction effect of the AGSPGD algorithm is better than the SPGD algorithm. By the way, the SPGD algorithm oscillates at the end of iteration, but the AGSPGD algorithm almost does not.

Next, according to the simulation results of the two algorithms, the group of wavefront aberrations with the worst correction effect is found and corrected separately after adjusting

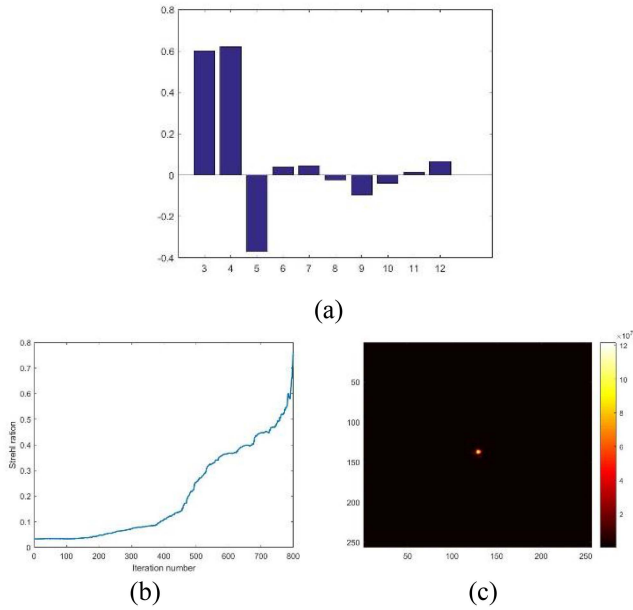


Fig. 9. Wavefront correction results of SPGD: (a) Initial Zernike coefficients of the wavefront aberrations, (b) SR iteration curve, (c) PSF after calibration.

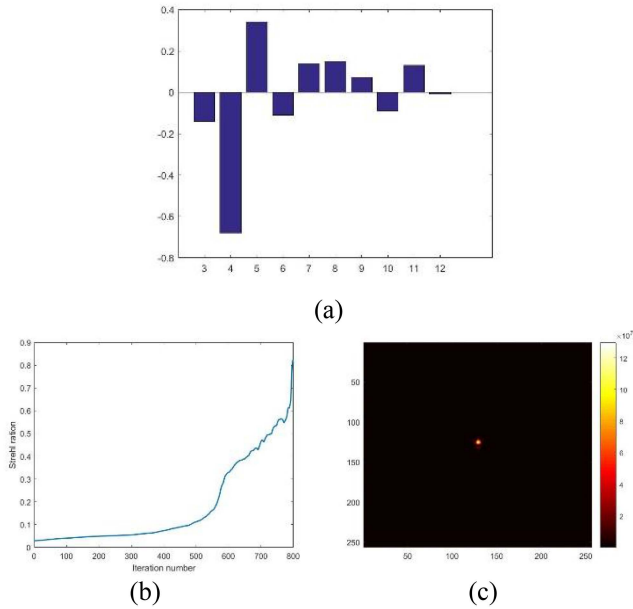


Fig. 10. Wavefront correction results of AGSPGD: (a) Initial Zernike coefficients of the wavefront aberrations, (b) SR iteration curve, (c) PSF after calibration.

the parameters  $\Delta u$  and  $\alpha$ . And the  $SR = 0.8$  is regarded as the convergence condition. The simulation results of separate correction by using the SPGD algorithm are shown in Fig. 9.

The AGSPGD algorithm is adopted to conduct the following simulations, and the simulation results are shown in Fig. 10.

As shown in Fig. 9, when the number of iterations is 800, the SR value increases from 0.03 to 0.75 by adjusting the parameters  $\Delta u$  and  $\alpha$  of the SPGD algorithm, but the convergence condition is still not reached. However, after adjusting the parameters of

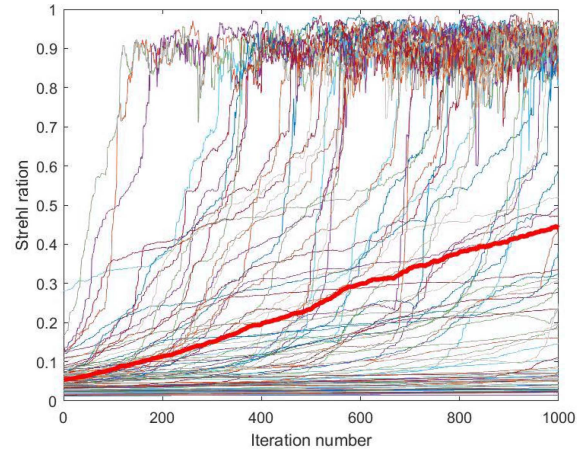


Fig. 11. Simulation results of the SPGD algorithm.

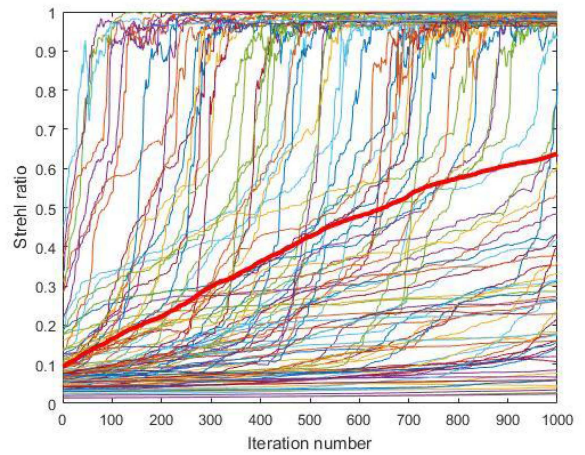


Fig. 12. Simulation results of the AGSPGD algorithm.

the AGSPGD algorithm, its SR value increases from 0.02 to 0.83 in Fig. 10. This result achieves the convergence condition. Therefore, the simulation results show that due to the high sensitivity of some parameters of the algorithms, the correction results of partial wavefront aberrations are not ideal. But there is no doubt that the correction performance of AGSPGD algorithm is better than that of SPGD algorithm.

In order to further verify the correction performance of the AGSPGD algorithm, the intensity of wavefront aberrations is increased, and the two algorithms are used to correct wavefront aberrations. Similarly, 100 groups of wavefront aberrations with intensity  $D/r_0 = 5$  are randomly generated. In the simulations, set the number of iterations to 1000.

The results of correcting 100 groups of wavefront aberrations with the SPGD algorithm are shown in Fig. 11. Consistent with the previous simulations, the red plus bold line represents the average value of SR.

The simulation results of 100 groups of wavefront aberrations corrected by the AGSPGD algorithm are shown in Fig. 12. And the red plus bold line also represents the average value of SR.

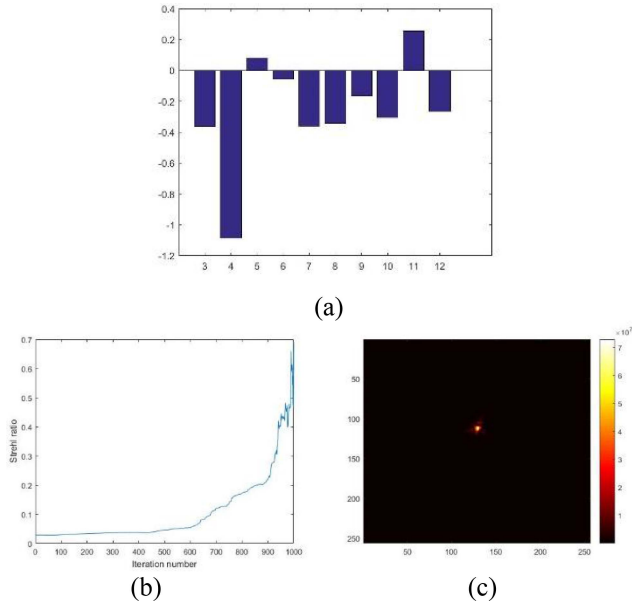


Fig. 13. Wavefront correction results of SPGD: (a) Initial Zernike coefficients of the wavefront aberrations, (b) SR iteration curve, (c) PSF after calibration.

Compared with the above simulation results, it is obvious that the enhancement of wavefront aberrations will degrade the correction performance of the AGSPGD algorithm, but the effect is significantly less than that of the SPGD algorithm. Specifically, when the intensity of wavefront aberrations is increased, the correction performance of the SPGD algorithm decreases significantly, and the average SR is only 0.45. Although the AGSPGD algorithm is also affected, its correction effect is still better than the SPGD algorithm, and the SR average value is 0.64. In addition, the simulation results in Fig. 11 show that the SPGD algorithm will produce oscillation when it reaches convergence, while the AGSPGD algorithm is relatively stable, as shown in Fig. 12. Therefore, the AGSPGD algorithm is better than the SPGD algorithm in correction performance. For the above two algorithms, the SR average value of both is less than 0.8, which may be because some parameters of the algorithm are too sensitive or the number of iterations is insufficient.

From the above simulation results, it can be seen that the performance of AGSPGD algorithm is still better than SPGD algorithm under the condition of increasing aberrations intensity. In order to verify whether the poor correction effect of the two algorithms is related to too sensitive parameters, the following simulations are carried out.

As in the above simulations, the wavefront aberrations with the worst correction effect in the simulation results of the two algorithms should be found out and corrected after adjusting the parameters  $\Delta u$  and  $\alpha$ , respectively. In the following simulations, the number of iterations is set to 1000 and the  $SR = 0.8$  is regarded as algorithm convergence.

The simulation results by using the SPGD algorithm are shown in Fig. 13.

The simulation results by using the AGSPGD algorithm are shown in Fig. 14.

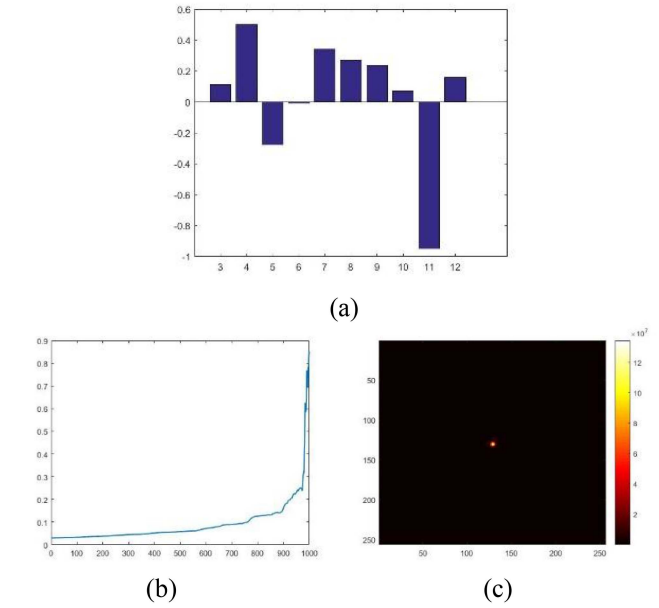


Fig. 14. Wavefront correction results of AGSPGD: (a) Initial Zernike coefficients of the wavefront aberrations, (b) SR iteration curve, (c) PSF after calibration.

TABLE I  
COMPARISON OF ITERATION TIMES OF TWO ALGORITHMS

Algorithm	Number of iterations
SPGD	800
AGSPGD	600

According to the above simulation results, the correction effect of the two algorithms is significantly improved after adjusting parameters  $\Delta u$  and  $\alpha$ . Specifically, it can be seen from Fig. 13 that the SPGD algorithm after adjusting parameters is used to correct wavefront aberrations, and the SR value increases from 0.03 to 0.69. Although the convergence condition is not reached, its correction effect has been significantly improved. When the parameters are adjusted, the simulation results show that the SR value increases from 0.028 to 0.86 by using the AGSPGD algorithm, which achieves the convergence condition, as shown in Fig. 14. These two simulation results prove the hypothesis above that the poor correction effect is related to too sensitive parameters. At the same time these results also show that the correction performance of the AGSPGD algorithm is better than the SPGD algorithm.

#### IV. COMPARATIVE ANALYSIS

To directly compare correction speed and the probability of falling into local optimum of the two algorithms, the above simulation results are compared and analyzed in this section, as follows Tables I and II.



TABLE II  
COMPARISON OF THE AVERAGE VALUE OF SR AND THE TIMES OF FALLING INTO LOCAL OPTIMUM BETWEEN THE TWO ALGORITHMS

$D/r_0$	Evaluation Indices	Algorithm	
		SPGD	AGSPGD
2.5	SR Average Value	0.79	0.94
	Local Optimum	16	4
5	SR Average Value	0.45	0.64
	Local Optimum	32	25

By comparing the data of the two algorithms in Table I, the correction speed of the AGSPGD algorithm is 25% faster than the SPGD algorithm. And based the Table II, the correction performance of the AGSPGD algorithm is better than that of the SPGD algorithm under different  $D/r_0$ . In this section, the  $SR < 0.3$  is regarded as the result of the algorithm falls into the local optimum. When  $D/r_0 = 2.5$ , the probability of the AGSPGD algorithm falling into local optimum is reduced by 12% compared with the SPGD algorithm. When  $D/r_0 = 5$ , the probability of the AGSPGD algorithm falling into local optimum is reduced by 7% compared with the SPGD algorithm.

Based on the above analysis, the AGSPGD algorithm converges faster and has a smaller probability of falling into the local optimum than the SPGD algorithm. And on the other hand, this article proves that the reason why the convergence condition is not achieved is related to the sensitivity of the parameters, rather than the unreasonable algorithm itself. Similarly, it can be guessed from the SR curve that the number of iterations will limit the convergence effect as the aberration intensity increases. Therefore, some simulation studies on improving aberration intensity and increasing iteration times will be further carried out in the future.

## V. CONCLUSION

The results demonstrate that the correction effect of the SPGD algorithm is related to the wavefront aberration intensity and the number of iterations. Moreover, the adaptability of the perturbation voltage amplitude  $\Delta u$ , the gain coefficient  $\gamma$  and other parameters of the algorithm is poor. These reasons lead to slow convergence of the result of the SPGD algorithm and even easy to fall into local optimum. Therefore, an improved SPGD algorithm called the AGSPGD algorithm is proposed in this article to accelerate the convergence of SPGD algorithm and reduce the probability of falling into the local optimum. The AGSPGD algorithm adopts adaptive gain coefficient, and its convergence speed is faster than the SPGD algorithm. Moreover, it is less affected by the change of wavefront aberration intensity. Specifically, the AGSPGD algorithm can reduce the number of iterations by 25% compared with the SPGD algorithm, and the probability of the algorithm falling into local optimum is reduced from 16% to 4%. In addition, the parameter adaptability of the AGSPGD algorithm is stronger than that of the SPGD algorithm.

In general, the AGSPGD algorithm proposed in this article has the characteristics of fast convergence speed and is not easy to fall into the local optimum.

In the future, as an optimized SPGD algorithm, the application of the AGSPGD algorithm in optical problems with higher turbulence intensity and more complex will be further studied.

## REFERENCES

- [1] J. Cao, X. Zhao, W. Liu, and H. Gu, "Performance analysis of a coherent free space optical communication system based on experiment," *Opt. Exp.*, vol. 25, no. 13, pp. 15299–15312, Jun. 2017.
- [2] M. Chen, C. Liu, and H. Xian, "Experimental demonstration of single-mode fiber coupling over relatively strong turbulence with adaptive optics," *Appl. Opt.*, vol. 54, no. 29, pp. 8722–8726, Oct. 2015.
- [3] J. Huang, H. Mei, K. Deng, L. Kang, W. Zhu, and Z. Yao, "Signal to noise ratio of free space homodyne coherent optical communication after adaptive optics compensation," *Opt. Commun.*, vol. 356, pp. 574–577, Dec. 2015.
- [4] H. Takenaka, M. Toyoshima, and Y. Takayama, "Experimental verification of fiber-coupling efficiency for satellite-to-ground atmospheric laser downlinks," *Opt. Exp.*, vol. 20, no. 14, pp. 15301–15308, 2012.
- [5] M. Li, W. Gao, and M. Cvijetic, "Slant-path coherent free space optical communications over the maritime and terrestrial atmospheres with the use of adaptive optics for beam wavefront correction," *Appl. Opt.*, vol. 56, no. 2, pp. 284–297, Jan. 2017.
- [6] L. Yang et al., "Performance analysis of 349-element adaptive optics unit for a coherent free space optical communication system," *Sci. Rep.*, vol. 9, no. 1, Sep. 2019, Art. no. 13150.
- [7] M. A. Vorontsov, G. W. Carhart, and J. C. Ricklin, "Adaptive phase-distortion correction based on parallel gradient-descent optimization," *Opt. Lett.*, vol. 22, no. 12, pp. 907–909, Jun. 1997.
- [8] M. A. Vorontsov and V. Sivokon, "Stochastic parallel-gradient-descent technique for high-resolution wave-front phase-distortion correction," *J. Opt. Soc. Amer. A*, vol. 15, no. 10, pp. 2745–2758, 1998.
- [9] M. A. Vorontsov and G. W. Carhart, "Adaptive wavefront control with asynchronous stochastic parallel gradient descent clusters," *J. Opt. Soc. Amer. A*, vol. 23, pp. 2613–2622, 2006.
- [10] E. Chen, H. Cheng, Y. An, and X. Li, "The improvement of SPGD algorithm convergence in satellite-to-ground laser communication links," *Procedia Eng.*, vol. 29, pp. 409–414, 2012.
- [11] W. Xiong, W. Xiaolin, Z. Pu, X. Xiaojun, and S. Bohong, "Numerical simulation of tilt-tip control in coherent beam combining using SPGD algorithm," *Opt. Laser Technol.*, vol. 48, pp. 343–350, 2013.
- [12] L. Lei, G. Jin, Z. Shuai, J. Zhen-hua, S. Tao, and W. Ting-feng, "Application of stochastic parallel gradient descent algorithm in laser beam shaping," *Chin. Opt.*, vol. 7, no. 2, pp. 260–266, Mar. 2014, doi: [10.3788/CO.20140702.0260](https://doi.org/10.3788/CO.20140702.0260).
- [13] K. Wu, Y. Sun, Y. Huai, S. Jia, X. Chen, and Y. Jin, "Multi-perturbation stochastic parallel gradient descent method for wavefront correction," *Opt. Exp.*, vol. 23, pp. 2933–2944, Feb. 2015.
- [14] T. Wang and X. Zhao, "Hybrid atmospheric compensation in free-space optical communication," *J. Opt. Soc. Korea*, vol. 20, no. 1, pp. 13–21, Feb. 2016.
- [15] K. Choi, Y. Kim, Y. Yun, Y. -C. Noh, and C. Jun, "Modal decomposition of measured near-field and far-field beams with partial coherence based on SPGD algorithm," in *Proc. Conf. Lasers Electro-Opt. Pacific Rim*, Sydney, NSW, Australia, Aug. 2020, pp. 1–6, doi: [10.1364/CLEOPR](https://doi.org/10.1364/CLEOPR).
- [16] K. Yang et al., "Piston error correction of sparse aperture systems using the metaheuristic stochastic parallel gradient descent algorithm," *Appl. Opt.*, vol. 59, no. 22, pp. 6505–6516, Jul. 2020.
- [17] J. Cao, X. Zhao, Z. Li, W. Liu, and Y. Song, "Stochastic parallel gradient descent laser beam control algorithm for atmospheric compensation in free space optical communication," *Optik*, vol. 125, no. 20, pp. 6142–6147, Oct. 2014.
- [18] H. Zhao, J. An, M. Yu, D. Lv, K. Kuang, and T. Zhang, "Nesterov-accelerated adaptive momentum estimation-based wavefront distortion correction algorithm," *Appl. Opt.*, vol. 60, no. 24, pp. 7177–7185, Aug. 2021.
- [19] S. L. Lachinova and M. A. Vorontsov, "Performance analysis of an adaptive phase-locked tiled fiber array in atmospheric turbulence conditions," *Proc. SPIE*, vol. 5895, pp. 124–137, 2005.



- [20] G. Yang, L. Liu, Z. Jiang, T. Wang, and J. Guo, "Improved SPGD algorithm to avoid local optimum for incoherent beam combining," *Opt. Commun.*, vol. 382, pp. 547–555, Jan. 2017.
- [21] Q. Hu, L. Zhen, Y. Mao, S. Zhu, X. Zhou, and G. Zhou, "Adaptive stochastic parallel gradient descent approach for efficient fiber coupling," *Opt. Exp.*, vol. 28, no. 9, pp. 13141–13154, Apr. 2020.
- [22] G. Yang, L. Liu, Z. Jiang, J. Guo, and T. Wang, "Incoherent beam combining based on the momentum SPGD algorithm," *Opt. Laser Technol.*, vol. 101, pp. 372–378, May 2018.
- [23] J. Song, Y. Li, D. Che, J. Guo, and T. Wang, "Coherent beam combining based on the SPGD algorithm with a momentum term," *Optik*, vol. 202, Feb. 2020, Art. no. 163650.
- [24] N. Qian, "On the momentum term in gradient descent learning algorithms," *Neural Netw.*, vol. 12, no. 1, pp. 145–151, Jan. 1999.
- [25] D. E. Rumelhart, G. E. Hinton, and R. J. Williams, "Learning representations by back-propagating errors," *Nature*, vol. 323, pp. 533–536, Oct. 1986.
- [26] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," *CoRR*, vol. abs/1412.6980, 2014.
- [27] R. J. Noll, "Zernike polynomials and atmospheric turbulence," *J. Opt. Soc. Amer.*, vol. 66, no. 3, pp. 207–211, Mar. 1976, doi: [10.1364/JOSA.66.000207](https://doi.org/10.1364/JOSA.66.000207).
- [28] N. Roddier, "Atmospheric wavefront simulation using Zernike polynomials," *Opt. Eng.*, vol. 29, no. 10, pp. 1174–1180, Oct. 1990.