applied optics

Optimization method using nodal aberration theory for coaxial imaging systems with radial basis functions based on surface slope

Shuai Zhang,[1] Liuchang Xiao,[1] Xing Zhao,[1,2,*] Lipei Song,[1] Yongji Liu,[1] Lingjie Wang,[3,4] Guangwei Shi,[3] and Weiwei Liu[1,2]

[1]Institute of Modern Optics, Nankai University, Tianjin 300350, China
[2]Tianjin Key Laboratory of Micro-scale Optical Information Science and Technology, Tianjin 300350, China
[3]Key Laboratory of Optical System Advanced Manufacturing Technology, Changchun Institute of Optics, Fine Mechanics and Physics, Chinese Academy of Sciences, Changchun 130033, China
[4]e-mail: wanglingjie@sina.com
*Corresponding author: zhaoxingtjnk@nankai.edu.cn

Received 30 December 2020; revised 10 February 2021; accepted 1 March 2021; posted 2 March 2021 (Doc. ID 418563); published 18 March 2021

The radial basis functions based on the surface slope (RBF-Slope) freeform surfaces model has demonstrated stronger fitting ability and better optical performance than the conventional RBF model. However, the large number of basis functions and optimization variables of the RBF-Slope model may result in convergence problems during optimization for optical systems consisting of freeform surfaces characterized by RBF-Slope. To overcome these drawbacks, we use Zernike polynomials to link the RBF-Slope model to aberration correction and propose a new optimization method for coaxial imaging systems using the RBF-Slope model based on nodal aberration theory (NAT). The aberrations generated by the conic parameter and Zernike terms up to $Z_{17/18}$ of the Zernike freeform surface at the non-stop surface in the coaxial imaging system are analyzed, and the gradient descent is implemented to obtain the optimal coefficients of the Zernike surface, which is then fitted by the RBF-Slope surface for further optimization. The method is applied to the optimization of a secondary mirror using the RBF-Slope model in a two-mirror telescope and proved to have better results than traditional commonly used direct optimization. This research offers an important reference for optimization using NAT and provides valuable insight into the optimization method for RBF-Slope freeform surfaces.

https://doi.org/10.1364/AO.418563

1. INTRODUCTION

In optical system design, freeform surfaces [1] that are not symmetric have more degrees of freedom than conventional rotational symmetric surfaces such as spherical and conic surfaces, which therefore reduces the aberration and simplifies the structure of the optical system. With the development of manufacturing techniques, freeform surfaces have been increasingly used in imaging systems [2–6] and nonimaging systems [7–10].

It is challenging to choose a suitable freeform surface model to connect classic aberration theory with the surface shape. Among these freeform surface models, Zernike polynomials [11], which are both continuous and orthogonal over a unit circle, have the same form as the types of wave aberrations, making them particularly applicable to the fields of surface characterization, optical design, and optical testing. Q-type polynomials developed by Forbes are orthogonalized with respect to mean square gradient over a circular aperture with the goal of facilitating measures of manufacturability [12]. XY polynomials, which are consistent with the numerical control (NC) optical expression form, have been widely selected for industrial manufacture. But when we represent an optical surface as the above three global-type freeform surfaces models, the sag value of the whole aperture will be affected by any changes of the coefficient of any term in the polynomial, which leads to deterioration of fitting ability and image quality for complicated or asymmetric surfaces. To solve this problem, the conventional radial basis function (RBF) model was proposed by Cakmakci *et al.*, which is a local descriptor of surface shape [13]. This means that the local shape is only impacted by variation of few basis function in this model. The stronger ability in fitting surfaces and aberration balancing compared to a Zernike polynomial model up to and including the 10th order in the design of single-element head-worn display (HWD) system has been demonstrated [14]. However, there are still some drawbacks in the conventional RBF model. Due to inefficient distribution and identical shape

factors in the conventional RBF model, the accuracy of characterization and the ability of aberration balancing are both limited. Therefore, the model with radial basis functions based on surface slope (RBF-Slope) was proposed by Tong *et al.* and proved stronger fitting ability and better optical performance than the conventional RBF model by improving the basis-function distribution for circular apertures and establishing a relationship between shape factor and local surface slope [15].

For the conventional RBF model and RBF-Slope model, the normal optimization process often consists of two steps: the first step is to set optimization variables such as shape factor, coefficient, and location of basis functions, and the second step is to optimize by using the default optimization function of optical design software. This process does work for the case of few basis functions. Nevertheless, in most situations, the number of basis functions of RBF-Slope surface and the variables during the optimization process are too large, which therefore leads to some problems in the optimization procedure like no convergence, low efficiency, and low convergence speed. Thus, it is of great importance to establish a new optimization method to improve the aberration-balancing ability of the RBF-Slope model. It is well known that the Zernike polynomials are directly associated with the traditional Seidel wave aberration because they have the same forms. Obviously, it can be treated as a bridge between the RBF-Slope model and wave aberration, and it can help the RBF-Slope surface be optimized with the guidance of aberration theory. As an aberration theory suitable for asymmetric optical system, the nodal aberration theory (NAT) was proposed by Shack and developed by Thompson [16–21]. By using NAT, the aberrations induced by the conic parameter and the Zernike polynomial terms on either the reflective or the refractive surfaces in the coaxial imaging systems can be quantitatively derived, which will be a useful tool to develop the optimization method for the RBF-Slope model.

In this paper, we use the Zernike polynomial whose expression is the same as the traditional Seidel aberration as a bridge between RBF-Slope and aberration correction and propose a new optimization method for coaxial imaging systems using the RBF-Slope model. In this method, the gradient descent algorithm is applied to acquire the optimal conic and Zernike term coefficients up to $Z_{17/18}$, and the initial optimization can be finished by fitting Zernike terms with the RBF-Slope model. The aberrations generated by the conic parameter and Zernike terms up to $Z_{17/18}$ at the non-stop surface in the coaxial imaging system are analyzed in Section 2. The gradient descent algorithm is explained, and the optimization method is introduced in detail in Section 3. In Section 4 we present the optimization of a second mirror using the RBF-Slope model in two-mirror telescopes. Finally, the advantages and shortcomings of this method are discussed in Section 5.

2. ABERRATIONS GENERATED BY THE CONIC PARAMETER AND ZERNIKE TERMS OF THE ZERNIKE FREEFORM SURFACE IN A COAXIAL IMAGING SYSTEM

In this section, we will quantitatively analyze the aberration generated by Zernike freeform surface terms located away from the stop. According to the previous studies by Thompson [18] and Yang [22], it can be achieved by establishing the quantitative relationship between the terms of the freeform surface at the non-stop surface and the induced aberration.

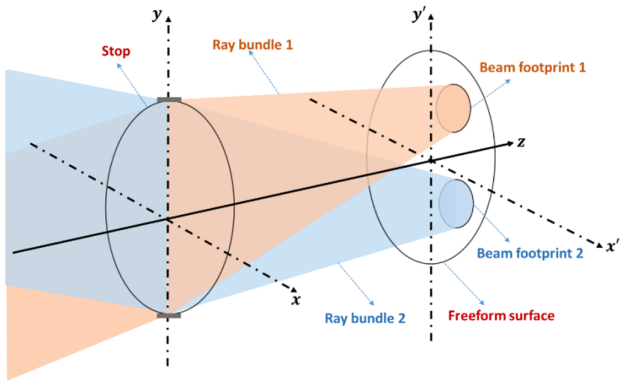

Fig. 1. Light footprints of different fields when the freeform surface is located at the non-stop surface.

A. Aberrations Generated by Zernike Freeform Surface Terms in a Coaxial Imaging System

Freeform surface terms added on a freeform surface can be seen as a zero-power thin optical plate or deformation. When the freeform surface is located at the stop, the light beam footprints of different fields are same on the surface. Thus, the induced aberration of Zernike terms to the system aberration can be seen as field constant. Base on NAT, the field coordinate H and pupil coordinate ρ in aberration terms are expressed in vector form $\vec{H} = He^{i\theta}$ and $\vec{\rho} = \rho e^{i\phi}$, where θ and φ represent the orientations of the two vectors. So the Zernike terms need to be expressed in vector form to integrate the freeform surface term into NAT. Considering Zernike terms as a kind of ϕ-polynomials that are expressed as $Z(\rho, \phi)$, they can be rewritten as $Z(\vec{\rho})$ in vector form. Since the freeform is at the stop, and the light beam from different fields shares the same area at the stop, and the aperture vector of Zernike surface $\vec{\rho}'$ is equal to the pupil vector $\vec{\rho}$. The induced aberration generated by Zernike terms of a Zernike freeform surface at the stop can be given by

$$W = \frac{n' - n}{\lambda} \vec{z}_i \cdot Z_i(\vec{\rho}') = \frac{n' - n}{\lambda} \vec{z}_i \cdot Z_i(\vec{\rho}), \tag{1}$$

where λ denotes the wavelength, $\vec{\rho}'$ is the normalized aperture vector of the Zernike surface, and $\vec{\rho}$ is the normalized pupil vector. $\vec{z}_i$ is the coefficient vector with respect to the ith Zernike terms $Z_i(\vec{\rho})$, and n and n' are the indices of refraction before and after the freeform surface. Especially in the reflective system, $n = 1$, $n' = -1$, and for odd times of reflection, the direction of the pupil coordinate rotates 180°, and the pupil vector $\vec{\rho}$ is replaced by $-\vec{\rho}$; however, the replacement just changes the sign of the generated aberrations, and the type of the aberrations will not be influenced. Therefore, we only analyze the situation in even times reflection here.

When the freeform surface is located away from the stop, the light beam footprints of different fields on the surface are different and only cover a part of the surface as shown in Fig. 1.

In this case, the aperture vector of Zernike surface $\vec{\rho}'$ is no longer equal to the pupil vector $\vec{\rho}$. Since aberrations are often

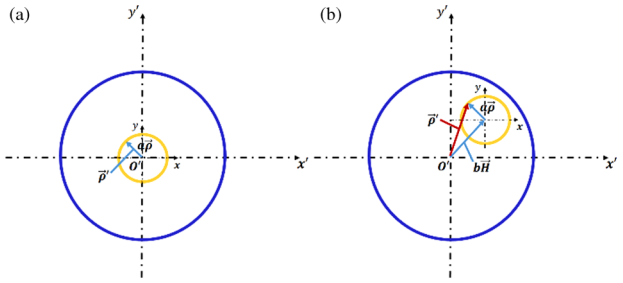

Fig. 2. Process of coordinate conversion on the freeform surface located at the non-stop surface. (a) Rays from the central field are irradiated on the freeform surface away from the stop; (b) rays from different fields are irradiated on the freeform surface away from the stop.

characterized in pupil coordinates, it is necessary to convert $\vec{\rho}'$ to $\vec{\rho}$. The process of coordinate conversion is illustrated in Fig. 2.

As shown in Fig. 2(a), when the rays from the central field are irradiated on the freeform surface away from the stop, the beam no longer completely covers the freeform surface, and the effective freeform surface aperture that interacts with the beam is determined by the actual beam trajectory irradiated on the freeform surface. Here we use the scaling factor a related to the marginal ray height to define the effective freeform surface aperture. In Fig. 2(b), when rays of different fields are irradiated on the freeform surface away from the stop, the center position of the beam footprint will shift with field of view, so the shift factor b related to the height of the chief ray is required to determine the center position of the beam footprint.

In summary, the pupil coordinate vector for different fields can be written as

$$\vec{\rho}' = a\vec{\rho} + b\vec{H} \tag{2}$$

with

$$a = \frac{y}{R}, \tag{3}$$

$$b = \frac{\bar{y}}{R}, \tag{4}$$

where y and $\bar{y}$ denote the marginal ray height and the chief ray height on the freeform surface and R is the normalized radius of the Zernike freeform surface.

Combining Eqs. (1) and (2), the aberration contribution induced by the freeform surface away from stop can be written as Eq. (5):

$$W(\vec{\rho}) = \frac{n' - n}{\lambda}\vec{z}_i \cdot Z_i(a\vec{\rho} + b\vec{H}). \tag{5}$$

From Eq. (5) the specific form of the field-dependent aberration components induced by the Zernike terms can be determined. For the sake of description, we represent $\frac{n'-n}{\lambda}\vec{z}_i$ with $\vec{C}_i$ in the following text.

Table 1. Fringe Zernike Terms up to $Z_{17/18}$ and Their Corresponding Aberrations

Term	Zernike Term	Corresponding Aberration Type
1	1	Piston (constant)
2	$\rho\cos\phi$	Distortion-tilt (x axis)
3	$\rho\sin\phi$	Distortion-tilt (y axis)
4	$2\rho^2 - 1$	Defocus
5	$\rho^2\cos(2\phi)$	Astigmatism, primary (axis at 0° or 90°)
6	$\rho^2\sin(2\phi)$	Astigmatism, primary (axis at ±45°)
7	$(3\rho^3 - 2\rho)\cos\phi$	Coma, primary (x axis)
8	$(3\rho^3 - 2\rho)\sin\phi$	Coma, primary (y axis)
9	$(6\rho^4 - 6\rho^2 + 1)$	Spherical aberration, primary
10	$\rho^3\cos(3\phi)$	Trefoil, primary (x axis)
11	$\rho^3\sin(3\phi)$	Trefoil, primary (y axis)
12	$(4\rho^4 - 3\rho^2)\cos(2\phi)$	Astigmatism, secondary (axis at 0° or 90°)
13	$(4\rho^4 - 3\rho^2)\sin(2\phi)$	Astigmatism, secondary (axis at ±45°)
14	$(10\rho^5 - 12\rho^3 + 3\rho)\cos(\phi)$	Coma, secondary (x axis)
15	$(10\rho^5 - 12\rho^3 + 3\rho)\sin(\phi)$	Coma, secondary (y axis)
16	$20\rho^6 - 30\rho^4 + 12\rho^2 - 1$	Spherical aberration, secondary
17	$\rho^4\cos(4\phi)$	Tetrafoil, primary (x axis)
18	$\rho^4\sin(4\phi)$	Tetrafoil, primary (y axis)

B. Aberration Contribution by the Conic Parameter and Zernike Terms up to $Z_{17/18}$

Based on the above theoretical basis, the induced aberration of the different Zernike terms up to $Z_{17/18}$ at the non-stop surface can be analytically derived and calculated. The first 18 fringe Zernike terms and their corresponding aberration types are given in Table 1. When the Zernike terms are added to the non-stop surface, the induced aberration types and properties can be analyzed by expanding Eq. (5) based on NAT. According to Yang's analysis of the nodal aberration properties of the coaxial imaging systems with Zernike polynomial surfaces in [22], the aberrations generated by the ith Zernike term are given in Table 2. Among these fringe Zernike terms, the terms 1–3 are the piston and tilt related to the position of surface, which will not affect the image quality. Furthermore, considering the complication and smaller contribution of fringe Zernike term 16 compared with other terms, the aberrations induced by fringe Zernike terms 1–3 and 16 are ignored in our work.

However, it is not sufficient to improve the image quality by only using Zernike terms because the conic parameter also plays an important role in correcting the aberration of the optical system. Thus, it is of great importance to quantitatively analyze the aberration generated by the conic parameter. For a freeform surface at stop surface, only spherical aberration is generated by the conic deformation. Based on Zhong's research in [23], the contribution of the conic parameter at the stop can be expressed by Eq. (6):

$$W(\vec{\rho})_{\mathrm{CONIC}} = \frac{kc^3y^4(n' - n)}{8\lambda}(\vec{\rho}\cdot\vec{\rho})^2, \tag{6}$$

Table 2. Aberrations Generated by the Conic Parameter and Zernike Terms of the Zernike Freeform Surface at the Non-stop Surface in a Coaxial Imaging System

Conic Parameter and Coefficients of Zernike Terms	Generated Aberration Types and Their Coefficients
k	$Kc^3y^4/8$ (Spherical aberration, primary)
	$Kc^3y^3\bar{y}\vec{H}/2$ (Coma, primary)
	$Kc^3y^2\bar{y}^2\vec{H}^2/4$ (Astigmatism, primary)
	$Kc^3y^2\bar{y}^2(\vec{H}\cdot\vec{H})/2$ (Defocus)
z_4	$2a^2C_4$ (Defocus)
$\vec{z}_{5/6}$	$a^2\vec{C}_{5/6}$ (Atigmatism, primary)
$\vec{z}_{7/8}$	$3a^3\vec{C}_{7/8}$ (Coma, primary)
	$6a^2b(\vec{C}_{7/8}\cdot\vec{H})$ (Defocus)
	$3a^2(\vec{C}_{7/8}\vec{H})$ (Astigmatism, primary)
z_9	$6a^4C_9$(Spherical aberration, primary)
	$24a^3b(C_9\vec{H})$ (Coma, primary)
	$12a^2b^2(C_9\vec{H}^2)$ (Astigmatism, primary)
	$24a^2b^2C_9(\vec{H}\cdot\vec{H})$ (Defocus)
$\vec{z}_{10/11}$	$a^3\vec{C}_{10/11}$ (Trefoil, primary)
	$3a^3b(\vec{C}_{10/11}\vec{H}^*)$ (Astigmatism, primary)
$\vec{z}_{12/13}$	$4a^4\vec{C}_{12/13}$ (Oblique spherical aberration)
	$12a^3b(\vec{C}_{12/13}\vec{H}^*)$ (Coma, primary)
	$4a^3b(\vec{C}_{12/13}\vec{H})$ (Trefoil, primary)
	$12a^2b^2(\lvert\vec{H}\rvert^2\vec{C}_{12/13})-3a^2\vec{C}_{12/13}$ (Astigmatism, primary)
	$12a^2b^2(\vec{C}_{12/13}\cdot\vec{H}^2)$ (Defocus)
$\vec{z}_{14/15}$	$10a^5\vec{C}_{14/15}$ (Coma, secondary)
	$20a^4b(\vec{C}_{14/15}\vec{H})$(Oblique spherical aberration)
	$30a^4b(\vec{C}_{14/15}\cdot\vec{H})$ (Spherical aberration, primary)
	$60a^3b^2(\vec{H}\cdot\vec{H})\vec{C}_{14/15}+30a^3b^2(\vec{C}^*_{14/15}\vec{H}^2)-12a^3\vec{C}_{14/15}$ (Coma, primary)
	$10a^2b^3(\vec{C}^*_{14/15}\vec{H}^3)+30a^2b^3(\vec{H}\cdot\vec{H})(\vec{C}_{14/15}\vec{H})-12a^2b(\vec{C}_{14/15}\vec{H})$ (Astigmatism, primary)
	$10a^3b^2(\vec{C}_{14/15}\vec{H})$ (Trefoil, primary)
	$60a^2b^3(\vec{H}\cdot\vec{H})(\vec{C}_{14/15}\cdot\vec{H})-24a^2b(\vec{C}_{14/15}\cdot\vec{H})$(Defocus)
$\vec{z}_{17/18}$	$a^4\vec{C}_{17/18}$ (Tetrafoil, primary)
	$4a^3b(\vec{C}_{17/18}\vec{H}^*)$ (Trefoil, primary)
	$6a^2b^2(\vec{C}_{17/18}(\vec{H}^*)^2)$ (Astigmatism, primary)

where k denotes the conic parameter and c is the curvature of the surface. Here we use K to represent $\frac{k(n'-n)}{\lambda}$ to make the expression clearer. When the conic parameter is added to the non-stop surface, the aberration generated by the conic parameter can be obtained by replacing $\vec{\rho}$ with $\vec{\rho}+\frac{\bar{y}}{y}\vec{H}$, and Eq. (6) can be written as Eq. (7):

$$
\begin{aligned}
&W(\vec{\rho})_{\mathrm{CONICnonstop}}\\
&=\frac{Kc^3y^4}{8}\left(\left(\vec{\rho}+\frac{\bar{y}}{y}\vec{H}\right)\cdot\left(\vec{\rho}+\frac{\bar{y}}{y}\vec{H}\right)\right)^2\\
&=\frac{Kc^3y^4}{8}\left(\vec{\rho}\cdot\vec{\rho}+2\frac{\bar{y}}{y}(\vec{\rho}\cdot\vec{H})+\left(\frac{\bar{y}}{y}\right)^2(\vec{H}\cdot\vec{H})\right)^2\\
&=\frac{Kc^3y^4}{8}\left((\vec{\rho}\cdot\vec{\rho})^2+4\frac{\bar{y}}{y}(\vec{H}\cdot\vec{\rho})^2+\left(\frac{\bar{y}}{y}\right)^2(\vec{H}\cdot\vec{H})^2\right.\\
&\quad\left.+4\frac{\bar{y}}{y}(\vec{H}\cdot\vec{\rho})(\vec{\rho}\cdot\vec{\rho})+2(\vec{H}\cdot\vec{H})(\vec{\rho}\cdot\vec{\rho})+4\left(\frac{\bar{y}}{y}\right)^3(\vec{H}\cdot\vec{H})(\vec{H}\cdot\vec{\rho})\right)\\
&=\frac{Kc^3y^4}{8}(\vec{\rho}\cdot\vec{\rho})^2+\frac{Kc^3y^2\bar{y}^2}{2}(\vec{H}\cdot\vec{H})(\vec{\rho}\cdot\vec{\rho})\\
&\quad+\frac{Kc^3y^2\bar{y}^2}{4}(\vec{H}^2\cdot\vec{\rho}^2)+\frac{Kc^3y^3\bar{y}}{2}(\vec{H}\cdot\vec{\rho})(\vec{\rho}\cdot\vec{\rho})\\
&\quad+\frac{Kc^3y^4}{2}\left(\frac{\bar{y}}{y}\right)^3(\vec{H}\cdot\vec{H})(\vec{H}\cdot\vec{\rho})+\frac{Kc^3y^4}{8}\left(\frac{\bar{y}}{y}\right)^2(\vec{H}\cdot\vec{H})^2
\end{aligned}
\tag{7}
$$

From Eq. (7), the aberrations generated by the conic parameter of the non-stop surfaces can be obtained in Table 2.

Using the equations in Table 2, the aberrations generated by the conic parameter and the Zernike polynomial terms on either the reflective or refractive surfaces in the coaxial imaging systems can be quantitatively calculated, which can be used as the guidance for aberration correction and optimization during the freeform optical system design. For example, the aberrations induced by fringe Zernike terms can be used to minimize the RMS wavefront error, and the image quality of the coaxial optical system might be improved by adding Zernike terms. In the following section, the quantitative relationship between the conic parameter, the Zernike terms, and the induced aberrations will be used in the initial optimization of the RBF-Slope model.

3. OPTIMIZATION METHOD OF A COAXIAL IMAGING SYSTEM USING THE RBF-SLOPE MODEL

The RBF-Slope model, which improves the basis-function distribution for circular apertures and establishes a relationship between shape factor and local surface slope, has stronger fitting ability and better optical performance than the conventional RBF model. The normal optimization method of the RBF or RBF-Slope model mainly utilizes the default optimization function of optical software to improve the imaging quality. However, in most situations the number of basis functions of the RBF-Slope model, the variables during the optimization process could be large and therefore lead to drawbacks in the optimization procedure such as no convergence, low efficiency, and low convergence speed. In order to overcome these drawbacks, we can apply the Zernike polynomials for linking the RBF-Slope model with aberration optimization in a coaxial imaging system. The optimization procedure is illustrated in Fig. 3.

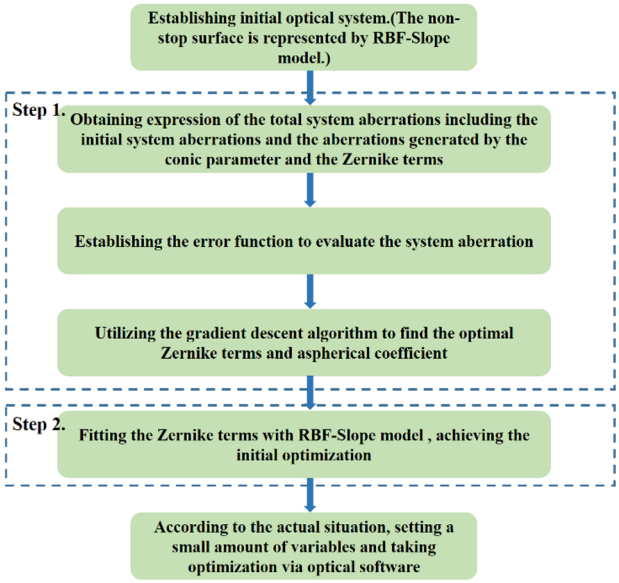

Fig. 3. Optimization procedure for the coaxial imaging system.

In a coaxial imaging system, the aberrations generated by the freeform surface located at the stop are field constant, which will be used to correct the field constant aberration. In this paper, we consider a more complicated situation in which the freeform surface will be away from the stop to generate the field-dependent aberrations. So the non-stop surface is represented by the RBF-Slop model to correct the field-dependent aberrations. After establishing the initial optical system, the initial optimization can be achieved by two steps. In step 1, we utilize the gradient descent algorithm to solve the optimal Zernike terms and conic parameter. In step 2, the initial optimization is finished by fitting the Zernike terms with the RBF-Slope model. The details of the two steps will be explained as follows.

A. Step 1 in the Initial Optimization Procedure

In the previous section, the aberrations generated by the conic parameter and Zernike terms can be quantitatively calculated by the equations in Table 2. Furthermore, we can find that when the Zernike terms and conic parameter are added to the non-stop Zernike surface, not only the aberrations corresponding to the employed Zernike term but also some other additional aberrations will be generated. Therefore, the system aberrations after adding the Zernike terms and conic parameter on the non-stop surface can be expressed as

$$W_i = W_i' + W_i'', \tag{8}$$

where W_i' denotes the matrix of the ith initial aberration coefficients corresponding to the ith Zernike term over the full field of view (FOV) for the coaxial imaging system before Zernike terms and conic parameters are added. W_i'' is the matrix of the ith additional aberration coefficients over the full FOV generated by the Zernike terms and conic parameters. W_i is the matrix of the ith total aberration coefficients over the full FOV after adding the Zernike terms and the conic parameter.

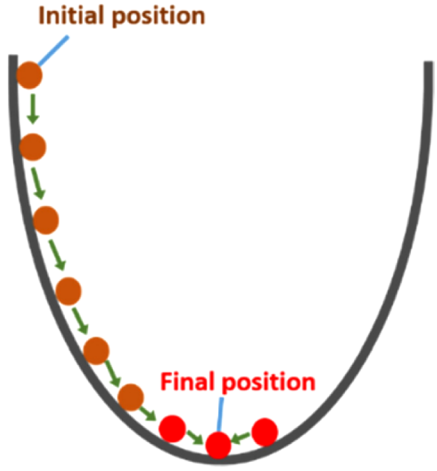

Fig. 4. Based schematic of the gradient descent algorithm.

To solve the optimal Zernike terms and conic parameter in a coaxial imaging system for the initial optimization, the aberration evaluation function needs to be established to evaluate the system aberrations. The aberration evaluation function consists of two parts. One is the sum of the square of average values of the ith aberration coefficients corresponding to the ith Zernike term up to $Z_{17/18}$ over full the FOV, which is directly related with the average values of the system aberrations. In the process of system optimization, the system is mainly limited by the low-order and third-order aberrations [24]; therefore, the other is the sum of the square of the root mean square (RMS) of the low-order and third-order aberrations corresponding to the jth Zernike term up to Z_9, which can represent the difference of the system's main aberrations in different fields. It can be written as

$$J(k,\ Z_1 \sim Z_{18}) = \sum_{i=1}^{18} \bar{W}_i^2 + \sum_{j=1}^{9} \mathrm{RMS}_j{}^2 \tag{9}$$

with

$$\bar{W}_i = \frac{\mathrm{sum}(W_i)}{q}, \tag{10}$$

$$\mathrm{RMS}_j = \sqrt{\frac{\mathrm{sum}[(W_j - \bar{W}_j)]}{q}}, \tag{11}$$

where $\bar{W}_i$ denotes average value of the ith aberration coefficients corresponding to the ith Zernike term, q is the number of fields in the coaxial imaging system, sum() is to get the sum of the matrix, RMS_j is RMS of the jth aberration coefficient corresponding to the jth Zernike term, $Z_1 \sim Z_{18}$ are the coefficients of the Zernike terms up to $Z_{17/18}$, and $J\,(k,\ Z_1 \sim Z_{18})$ is the value of the aberration evaluation function.

Then the gradient descent algorithm is used to obtain the optimal coefficients of the Zernike terms up to $Z_{17/18}$ and the conic parameter. The gradient descent algorithm is an iterative method to solve the extreme value with the direction of the gradient descent. Figure 4 illustrates the basic schematic of the gradient descent algorithm. The minimum value of the objective function can be achieved at the final position, and the unknown variables will be solved.

By using Eq. (9), the gradient function can be derived as

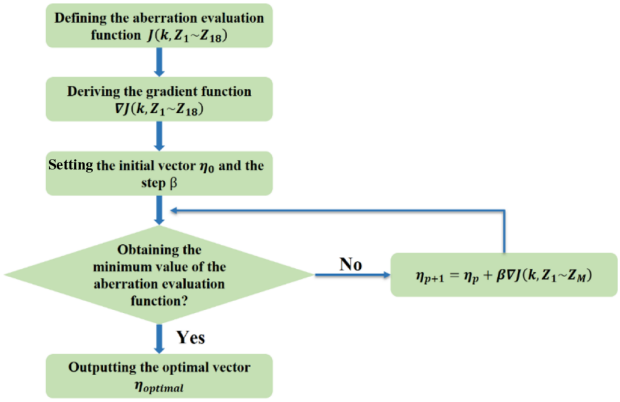

Fig. 5. Solving process for the optimal coefficients of the Zernike terms and conic parameter.

$$\nabla J(k, Z_1 \sim Z_{18}) = \left\langle \frac{\partial J}{\partial k}, \frac{\partial J}{\partial Z_1}, \frac{\partial J}{\partial Z_2}, \ldots, \frac{\partial J}{\partial Z_{17}}, \frac{\partial J}{\partial Z_{18}} \right\rangle. \tag{12}$$

Next, the iterative equation can be written as

$$\eta_{p+1} = \eta_p + \beta \nabla J(k, Z_1 \sim Z_{18}), \tag{13}$$

where η_p and η_{p+1} denote the vector $\left\langle \frac{\partial J}{\partial k}, \frac{\partial J}{\partial Z_1}, \frac{\partial J}{\partial Z_2}, \ldots, \frac{\partial J}{\partial Z_{17}}, \frac{\partial J}{\partial Z_{18}} \right\rangle$ before and after the pth iteration, respectively. β determines the step of gradient descent. First, we need to set the initial vector η_0 and the step β, which determine the descent speed of the objective function. Then, after each iteration, the value of the aberration evaluation function will decrease with the direction of the gradient descent until the minimum is obtained, meaning the optimal vector η_{optimal} is solved. Figure 5 illustrates the solving process.

B. Step 2 in the Initial Optimization Procedure

After getting the optimal coefficients of the Zernike terms and the conic parameter, the initial optimization of the RBF-Slope model can be achieved by fitting the Zernike terms with the RBF-Slope model. The RBF-Slope model can be represented by a linear combination of RBFs and a conic as

$$z(x, y) = \frac{c(x^2 + y^2)}{1 + \sqrt{1 - (1 + k)c^2(x^2 + y^2)}} + \sum_i w_i e^{-\varepsilon_i^2((x - x_{0i})^2 + (y - y_{0i})^2)}, \tag{14}$$

where c represents curvature, k is the conic constant, $z(x, y)$ denotes the sag of the freeform surface in the aperture, and (x, y) are the Cartesian coordinates. The second term represents a linear combination of RBFs, where (x_{0i}, y_{0i}) and shape factor ε_i determines the center position and the width of the basis functions. Comparing the freeform surface model with the Zernike polynomials, the first term is the same. Therefore, the RBF-Slope surface, which has the same sag distribution as the optimized Zernike surface, can be obtained by fitting the optimized Zernike terms with the second term of the RBF-slope model.

Combining with step 1 and step 2, the initial optimization procedure can be finished. Next we can set a small amount of optimization variables such as coefficient, shape factor, and location of basis functions depending on the actual situations and take optimization in the optical design software to obtain high imaging quality. To increase the convenience of optimization, the whole optimization process is integrated into a program. In this paper, we use optical design software Code V to obtain the required data in optimization, such as aberration coefficients and system parameters, and we use MATLAB to do data processing. Because the Code V API (application programming interface) uses the Microsoft Windows standard Component Object Model (COM) interface, users can execute Code V commands by MATLAB, which supports Windows COM architecture. Thus, without frequent switching between Code V and MATLAB, the integrated MATLAB program enables the initial optimization of the RBF-Slope model to be completed automatically.

With the above optimization method proposed by utilizing the Zernike terms to link the RBF-Slope model with aberration optimization, the initial optimization is simplified. It could make the system a good starting point for further optimization with optical design software. Therefore, the optimization efficiency could be improved. This is of great importance to solve problems like low convergence speed or even no convergence. It is expected that a RBF-Slope coaxial imaging system with better image quality could be obtained through using this optimization method.

4. OPTIMIZATION OF THE SECONDARY MIRROR OF A TWO-MIRROR TELESCOPE USING THE RBF SLOPE MODEL

Based on the optimization method for the RBF-Slope model proposed in Section 3, a two-mirror telescope is optimized whose second mirror is described by the RBF-Slope model. In terms of usage, this system is suitable for a wide spectrum; since the reflection system has no chromatic aberration, we use 632.8 nm monochromatic light as a design example. Specifications of the optical system are listed in Table 3, and the initial parameters for a two-mirror telescope with 1.2° × 1.2° rectangular FOV are listed in Table 4. The optical layout of this telescope is presented in Fig. 6. In order to verify the effectiveness of the proposed method, the goal of the optimization is to achieve smaller aberrations and better image quality by only optimizing the secondary mirror using the RBF-Slope model with 256 basis functions. Therefore, we focus on the optimization of basis functions of the RBF-Slope model in this paper, and the parameters of the primary mirror and the radius of the secondary mirror are fixed during the optimization procedure.

Before the optimization is implemented, a fixed conic parameter $k = 4$ and coefficient of Zernike terms $Z_i (i = 1 \sim$

Table 3. Specifications of the Two-mirror Telescope

Parameter	Value
Effective focal length	1292 mm
Field of view	1.2° × 1.2°
Working wavelength	632.8 nm

Table 4. Initial Parameters of the Two-mirror Telescope (unit: mm)

Surface	Type	Conic Constant	Radius	Thickness	Semi-aperture
PM(stop)	Conic	−0.6475	−742.7795 mm	−260.5773 mm	75 mm
SM	Sphere	None	−311.7356 mm	382.8867 mm	25 mm
FP			Infinity		

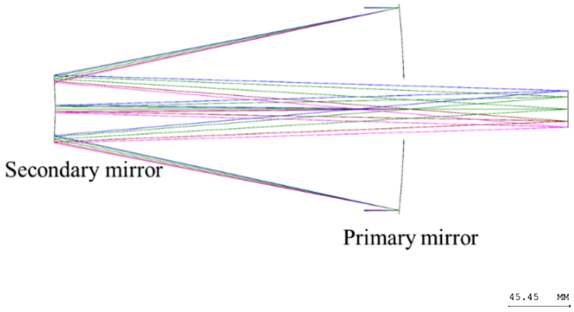

Fig. 6. Optical layout of the two-mirror telescope.

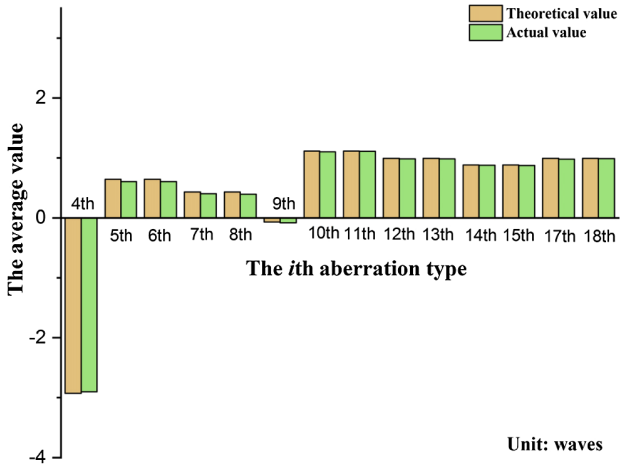

Fig. 7. Comparison of the theoretical and actual values when the Zernike terms (Zernike terms' coefficients all equal to 0.005, and conic parameter equal to 4) are added onto the secondary mirror.

18) = 0.005 are added onto the secondary mirror to validate the aberrations' contribution demonstrated in Section 2.B. The comparison of the generated ith aberration between the actual values given by CODE V and the theoretical values achieved from Table 2 are shown in Fig. 7. From the figure, it can be seen that the types and the values of the actually generated aberrations coincide with the theoretical results approximately. So the method of aberration analysis and calculation with NAT could be expected to be utilized as useful tool for coaxial system optimization.

In the optimization process, 121 field points are set within 1.2° × 1.2° field of view, the paraxial marginal ray height y and the paraxial chief ray height $\bar{y}$ on the secondary mirror can be obtained using real ray trace in Code V, and the factor a and b can be calculated by Eqs. (3) and (4) as shown in Table 5.

By using the above method of aberration analysis and calculation with NAT, the gradient descent algorithm is utilized to solve the optimal conic and Zernike coefficients. In our gradient

Table 5. Ray Height on the Secondary Mirror and the Factors *a* and *b*

y	$\bar{y}$	a	b
22.3496 mm	2.7289 mm	0.8940	0.1092

Table 6. Optimal Coefficients of Zernike Terms and the Conic Parameter

Conic Parameter and Zernike Terms	Coefficient	Zernike Fringe Term	Coefficient
k	−7.424	10	7.924e-14
1, 2, 3, 16	0	11	1.031e-14
4	−5.825e-3	12	1.987e-4
5	9.667e-5	13	−6.481e-14
6	−3.154e-14	14	7.909e-15
7	−3.262e-16	15	3.689e-14
8	2.308e-14	17	5.138e-6
9	−1.566e-3	18	−9.166e-15

descent algorithm, all the Zernike terms coefficients are set to 0.005 and the conic parameter is set to 4 initially. The step of gradient descent is 1.0×10^{-9}. Through the iterative process, the optimal coefficients of the Zernike terms and conic parameter will be obtained when the value of the objective function is minimum as shown in Table 6.

To accomplish the initial optimization by fitting the Zernike terms with the RBF-Slope model, and to perfectly transform the optimized Zernike surface into the RBF-Slope surface, it is necessary to avoid the optimization loss caused by the fitting accuracy in the fitting process as much as possible. We used 256 basis functions and 1000 sample points to fit the Zernike terms. The sample points are uniformly and regularly distributed in the aperture. Least squares was used to fit the Zernike terms with Householder transformation, which can effectively deal with the ill-condition problems of least squares [25]. Root-mean-square (RMS) error and peak-to-valley (PV) error to evaluate the fitting precision are 8.0231e-13 and 3.8854e-11, respectively. In addition, the fitting error of the RBF-Slope model is shown in Fig. 8, which demonstrates that the fitting precision is high enough for RBF-Slope to describe the shape of Zernike terms accurately. Then we can bring the fitting result to the secondary mirror and set a small amount of parameters as variables, such as coefficient, shape factor, and location of the basis functions, and we can make the following optimization automatically via optical software to achieve high imaging quality.

In order to prove the advantages of the proposed method, we compared it with the classic methods commonly used in surface optimization. Figure 9 shows the RMS wavefront error (WFE) of the system over full FOV in different cases; we can see the default optimization of the optical software is no convergence with large aberration, while the proposed optimization can

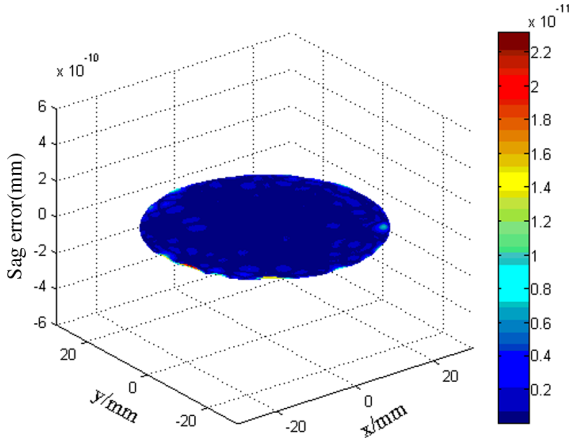

Fig. 8. Sag error as fitted with the RBF-Slope model with 256 basis functions.

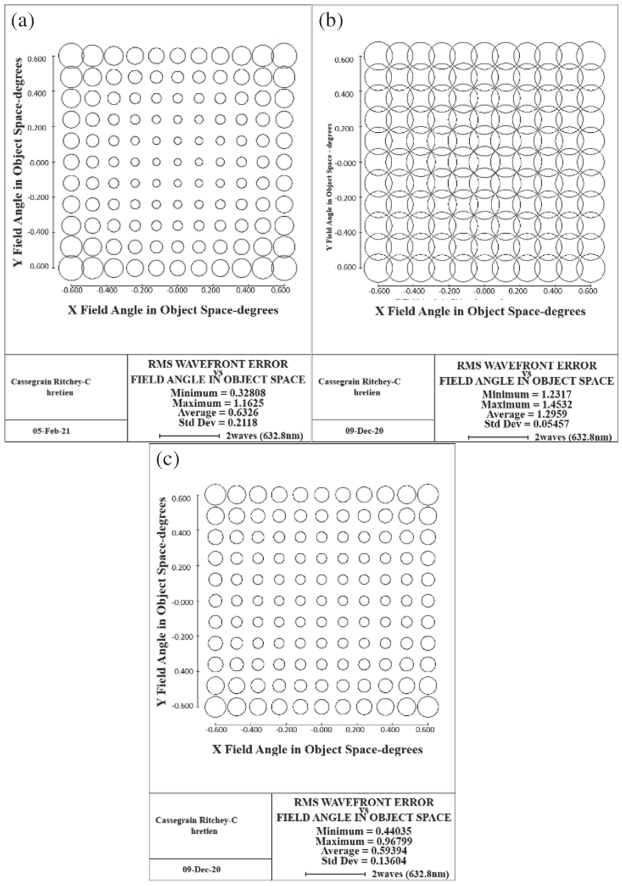

Fig. 9. RMS wavefront error (WFE) in the full FOV of the optical system (a) before optimization, (b) with optimization by the default method of optical software, and (c) with optimization by proposed method.

achieve better aberration balance and optical performance. The reason why the effect of the default optimization is worse than before optimization is that the system default optimization process does not converge due to the large number of optimization variables of RBF-Slope, which makes the optimization stagnate in the process of finding local extrema. In addition,

Table 7. Average Value of the Aberration Coefficients for the Three Situations (unit: waves)

	Before Optimization	Default Optimization	Proposed Optimization Method
RMS Wavefront Error	0.6326	1.2959	0.5939
Defocus	–0.1222	–0.3473	–0.2284
Astigmatism	0.1482	0.7235	0.2662
Coma	1.7515	1.5227	1.0986
Spherical Aberration	0.0008	4.0372	0.8322

Table 8. Standard Deviation of the Aberration Coefficients for the Three Situations (unit: waves)

	Before Optimization	Default Optimization	Proposed Optimization Method
RMS Wavefront Error	0.2118	0.0546	0.1360
Defocus	0.3183	0.6775	0.3898
Astigmatism	0.0926	0.4517	0.1574
Coma	0.6489	0.5648	0.4219
Spherical Aberration	0.0006	0.0010	0.0043

the average value and the standard deviation value of RMS WFE, defocus, astigmatism, coma, and spherical aberration for the three situations are calculated and listed in Tables 7 and 8. The average values in Table 7 show that coma is improved after optimization by the proposed method, and the uniformity of coma is tremendously improved as well as shown in Table 8. In contrast, the spherical aberration increases greatly. Actually, during the optimization procedure, it can be observed that coma is gradually compensated with the increase of spherical aberrations. Finally, the average value and uniformity of RMS wavefront error aberration are both improved with all primary aberrations balanced, which means that the overall image quality become better. Thus, the results prove that the proposed optimization method for the RBF-Slope model can effectively balance the system aberrations and solve problems in optimization such as no convergence, low efficiency, and low convergence speed.

5. CONCLUSION

In this paper, we use Zernike polynomials to link the RBF-Slope model to aberration correction and propose a new optimization method for coaxial imaging systems using the RBF-Slope model based on NAT. The method utilized Zernike terms to link the RBF-Slope model with aberration correction and optimization by NAT. Based on NAT, the aberration generated by the conic parameter and Zernike terms up to $Z_{17/18}$ can be calculated, which helps establish the quantitative relationships between the generated aberration and the system aberration. Then the gradient descent algorithm is used to obtain the optimal conic and Zernike coefficients. After that, the initial optimization can be completed by fitting the Zernike terms with the RBF-Slope model, and further optimization can be implemented automatically in optical software. Finally, the optimization result of

the secondary mirror of a two-mirror telescope using the RBF-Slope model indicated that better optical performance could be achieved, and the existing problems like slow convergence speed, no convergence, and low efficiency can be solved by the proposed optimization method. It is believed that the research in this paper offers a good reference for RBF-Slope-based coaxial imaging system optimization with NAT. In the future, we will carry out further research on the optimization method for the refraction freeform system. The optimization method for the RBF-Slope model in off-axis imaging systems using enhanced NAT and an improved optimization algorithm will be investigated, which is more useful to the applications of freeform optical systems.

Funding. National Natural Science Foundation of China (62075106, 61675100); Natural Science Foundation of Tianjin City (19JCZDJC36600); Tianjin Key R&D Program (19YFZCSY00250).

Disclosures. The authors declare that there are no conflicts of interest related to this paper.

REFERENCES

1. K. Fuerschbach, J. P. Rolland, and K. P. Thompson, "A new family of optical systems employing φ-polynomial surfaces," Opt. Express **19**, 21919–21928 (2011).
2. X. Hu and H. Hua, "High-resolution optical see-through multi-focal-plane head-mounted display using freeform optics," Opt. Express **22**, 13896–13903 (2014).
3. B. Chen and A. M. Herkommer, "Alternate optical designs for head-mounted displays with a wide field of view," Appl. Opt. **56**, 901–906 (2017).
4. Y. Nie, R. Mohedano, P. Benítez, J. Chaves, J. C. Miñano, H. Thienpont, and F. Duerr, "Multifield direct design method for ultra-short throw ratio projection optics with two tailored mirrors," Appl. Opt. **55**, 3794–3800 (2016).
5. C. Xu, D. Cheng, J. Chen, and Y. Wang, "Design of all-reflective dual-channel foveated imaging systems based on freeform optics," Appl. Opt. **55**, 2353–2362 (2016).
6. E. Muslimov, E. Hugot, W. Jahn, S. Vives, M. Ferrari, B. Chambion, D. Henry, and C. Gaschet, "Combining freeform optics and curved detectors for wide field imaging: a polynomial approach over squared aperture," Opt. Express **25**, 14598–14610 (2017).
7. Z. Feng, L. Huang, G. Jin, and M. Gong, "Designing double freeform optical surfaces for controlling both irradiance and wavefront," Opt. Express **21**, 28693–28701 (2013).
8. T. Yang, J. Zhu, X. Wu, and G. Jin, "Direct design of freeform surfaces and freeform imaging systems with a point-by-point three-dimensional construction-iteration method," Opt. Express **23**, 10233–10246 (2015).
9. S. Sorgato, R. Mohedano, J. Chaves, M. Hernández, J. Blen, D. Grabovičkić, P. Benítez, J. C. Miñano, H. Thienpont, and F. Duerr, "Compact illumination optic with three freeform surfaces for improved beam control," Opt. Express **25**, 29627–29641 (2017).
10. X. Wu, G. Jin, and J. Zhu, "Freeform illumination design model for multiple light sources simultaneously," Appl. Opt. **56**, 2405–2411 (2017).
11. R. W. Gray, C. Dunn, K. P. Thompson, and J. P. Rolland, "An analytic expression for the field dependence of Zernike polynomials in rotationally symmetric optical systems," Opt. Express **20**, 16436–16449 (2012).
12. G. W. Forbes, "Shape specification for axially symmetric optical surfaces," Opt. Express **15**, 5218–5226 (2007).
13. O. Cakmakci, I. Kaya, G. E. Fasshauer, K. P. Thompson, and J. P. Rolland, "Application of radial basis functions to represent optical freeform surfaces," Proc. SPIE **7652**, 76520A (2010).
14. O. Cakmakci, B. Moore, H. Foroosh, and J. P. Rolland, "Optimal local shape description for rotationally non-symmetric optical surface design and analysis," Opt. Express **16**, 1583–1589 (2008).
15. K. Tong, Y. Zheng, Z. Zhang, X. Zhao, B. Zhang, L. Song, L. Wang, C. Wang, and P. Wu, "Model of radial basis functions based on surface slope for optical freeform surfaces," Opt. Express **26**, 14010–14023 (2018).
16. R. V. Shack and K. P. Thompson, "Influence of alignment errors of a telescope system," Proc. SPIE **251**, 146–153 (1980).
17. K. P. Thompson, "Aberration fields in unobscured mirror systems," J. Opt. Soc. Am. **103**, 159–165 (1980).
18. K. P. Thompson, "Description of the third-order optical aberrations of near-circular pupil optical systems without symmetry," J. Opt. Soc. Am. A **22**, 1389–1401 (2005).
19. K. P. Thompson, "Multinodal fifth-order optical aberrations of optical systems without rotational symmetry: spherical aberration," J. Opt. Soc. Am. A **26**, 1090–1100 (2009).
20. K. P. Thompson, "Multinodal fifth-order optical aberrations of optical systems without rotational symmetry: the comatic aberrations," J. Opt. Soc. Am. A **27**, 1490–1504 (2010).
21. K. P. Thompson, "Multinodal fifth-order optical aberrations of optical systems without rotational symmetry: the astigmatic aberrations," J. Opt. Soc. Am. A **28**, 821–836 (2011).
22. T. Yang, J. Zhu, and G. Jin, "Nodal aberration properties of coaxial imaging systems using Zernike polynomial surfaces," J. Opt. Soc. Am. A **32**, 822–836 (2015).
23. Y. Zhong and H. Gross, "Vectorial aberrations of biconic surfaces," J. Opt. Soc. Am. A **35**, 1385–1392 (2018).
24. A. Bauer, E. M. Schiesser, and J. P. Rolland, "Starting geometry creation and design method for freeform optics," Nat. Commun. **9**, 1756 (2018).
25. X. Lin, X. G. Liu, Y. Li, and G. Wei, "A new orbit fitting algorithm of space-borne SAR based on householder transformation," in *Asian-Pacific Conference on Synthetic Aperture Radar (APSAR)* (2009), pp. 832–835.