# Method of optimizing deterministic small free-form surfaces for off-axis optical systems with extremely low aberration

IOP Publishing Journal of Optics
J. Opt. **25** (2023) 015701 (8pp) https://doi.org/10.1088/2040-8986/ac9d4b

# Method of optimizing deterministic small free-form surfaces for off-axis optical systems with extremely low aberration

**Xu Zhang[1,2], Jie Yu[1,3], Liping Wang[1,3] and Chunshui Jin[1,3,*]**

[1] Changchun Institute of Optics, Fine Mechanics and Physics, Chinese Academy of Sciences, Changchun, Jilin 130033, People's Republic of China
[2] University of Chinese Academy of Sciences, Beijing 100039, People's Republic of China
[3] State Key Laboratory of Applied Optics, Changchun, Jilin 130033, People's Republic of China

E-mail: jincs@sklao.ac.cn

Received 9 June 2022, revised 1 October 2022
Accepted for publication 25 October 2022
Published 5 December 2022

**Abstract**

A mapping model of arbitrarily shaped surfaces and the system image quality is constructed to examine the optimization of the residual aberration of an off-axis optical system to construct an off-axis optical system with very small aberrations. First, orthogonal surfaces are chosen within the arbitrary aperture. The mapping relationship between the orthogonal surfaces and system wavefront aberration is then established. Finally, the surfaces required for optimization are acquired by solving the mapping relationship through singular value decomposition and the Gauss–Newton algorithm. The residual aberration of the system is optimized using free-form surfaces with small deviations. In this paper, the residual aberration of an off-axis optical system with a non-circular aperture is optimized by adopting the above method, with the residual aberration of the system reducing from 0.549 nm (root-mean-square, RMS) to 0.443 nm (RMS) after adding a small free-form surface to a single lens, and to 0.393 nm (RMS) after adding small free-form surfaces to two lenses. The optical system with a circular aperture is optimized and the residual aberration of the system is reduced from 2 nm (RMS) to 0.47 nm (RMS).

Keywords: freeform optimization, optical design, aberration correction

(Some figures may appear in colour only in the online journal)

## 1. Introduction

The demand for high-quality imaging systems in medical diagnostics, remote sensing, and extreme ultraviolet lithography has increased with the rapid advance in optical technology. However, perfect imaging is difficult to achieve owing to the intrinsic defects of the imaging system, and aberrations are inevitable [1].

In reducing the residual wavefront aberration and improving the system's image quality, the system lenses are often changed from rotationally symmetric surfaces to free-form surfaces to provide a higher degree of freedom [2]. When choosing a free-form surface to optimize, the Zernike polynomials commonly used to analyze optical surface deviations [3] and wavefront aberrations [4] lose their advantage of orthogonality in the non-circular aperture of the off-axis system, resulting in the polynomial coupling, whereas the orthogonal Zernike surfaces in the non-circular domain are mostly used for wavefront reconstruction [5]. Additionally, traditional optimization methods for coaxial spherical systems may no longer be applicable and may need to be re-established in guiding the optimization process. Zhu Jun's team the Tsinghua University has proposed a direct method to design the freeform off-axis reflective imaging system [6] and uses a point-by-point design method for mixed-surface-type off-axis reflective imaging systems [7]. Wang Yongtian's team at the Beijing Institute of Technology has proposed a

* Author to whom any correspondence should be addressed.

 

stepwise approximation optimization strategy for free-form imaging systems, whereby the image quality is gradually improved by progressively optimizing the surface from simple to complex [8]. The team has also proposed an automatic balance optimization method for the overall image quality of the imaging system, by adding a cyclic control layer to analyze the imaging quality of the system and automatically assigning and setting the field of view weights. The method thus provides a balanced image quality throughout the field [9]. Fuerschbach *et al* analyzed the main aberration types of the imaging system according to aberration theory and set corresponding free-form surface coefficients as variables for targeted optimization, thus realizing progressive optimization [10]. The use of these common surfaces and optimization methods to optimize a small-aberration off-axis optical system and thus reduce the residual aberration of the system faces the problems of coupled surface polynomials and the lack of a link between the surfaces and the wavefront aberration of the system. These problems provide more degrees of freedom and make the system more complex, resulting in slow and time-consuming optimization and the need for constant trial and error to approach the optimal solution. The optimization process thus lacks clear directionality.

To solve the above problem, this paper proposes a free-surface optimization method for optical systems based on the mapping relationship between the non-circular orthogonal Zernike surface and the wavefront aberration of the system. Compared with other optimization methods using common free-form surface shapes, the proposed method characterizes the orthogonal Zernike surface within an arbitrary aperture according to the orthogonal decomposition of the Gram matrix, reducing the polynomial coupling within the lens aperture during the optimization process. It is explained in detail in section 2.1. Additionally, the method establishes a mapping relationship between each of the non-circular orthogonal Zernike surfaces and the wavefront aberration, and solves the coefficients of the non-circular orthogonal Zernike surface in the optimization process by adopting singular value decomposition (SVD) and Gauss–Newton algorithm. It is explained in detail in section 2.2. This method is by adding the surface which is used in the optimization to the original lenses of the system, allowing iterative optimization. Through this method, the residual wavefront aberration of a small aberration off-axis optical system is reduced. Two kinds of systems are optimized to demonstrate the proposed design method in section 3. The validity of the mapping between the non-circular orthogonal Zernike surface and the wavefront aberration and the effectiveness of the method in free-form optimization is demonstrated.

## 2. Free-form optimization method based on the relationship between a non-circular orthogonal Zernike surface and wavefront aberration mapping

### *2.1. Characterization of non-circular orthogonal Zernike surfaces*

The orthogonal decomposition of the Gram matrix to reorganize the vectors within the lens aperture and thus generate a new orthogonal basis function is adopted to characterize the non-circular orthogonal Zernike surface [5, 11, 12].

According to higher algebraic theory, a linear space is an abstraction of three-dimensional geometric space, and for any two real-variable real-valued functions $f,g$ on a linear space domain $R$, the relation,

$$(f,g) = \int_R f(r)g(r)d\bar{r} \tag{1}$$

is an inner product on a spatial domain $R$. $R$ with the inner product $(f,g)$ forms an infinite dimensional Euclidean space, and $f,g$ is a vector of Euclidean spaces. For the $n$ vectors $a_1,a_2,\ldots,a_n$of the Euclidean space, the Gram matrix is defined as $z_1,z_2,\ldots,z_n$,

$$G(a_1,a_2,\ldots,a_n) = \begin{pmatrix} (a_1,a_1) & (a_1,a_2) & \cdots & (a_1,a_n) \\ (a_2,a_1) & (a_2,a_2) & \cdots & (a_2,a_n) \\ \cdots & \cdots & \cdots & \cdots \\ (a_n,a_1) & (a_n,a_2) & \cdots & (a_n,a_n) \end{pmatrix} \tag{2}$$

a sufficient condition for $a_1,a_2,\ldots,a_n$ to be linearly independent is that the Gram matrix is positive definite [13].

Taking the Fringe Zernike polynomials as an example, for the $m$ vectors of the Euclidean space $z_1,z_2,\ldots,z_m$ constituted by the Fringe Zernike terms in the non-circular domain

$$R(\mathrm{Max}(r) \leqslant 1), \tag{3}$$

the Gram matrix is defined as,

$$G(z_1,z_2,\ldots,z_m) = \begin{pmatrix} (z_1,z_1) & (z_1,z_2) & \cdots & (z_1,z_m) \\ (z_2,z_1) & (z_2,z_2) & \cdots & (z_2,z_m) \\ \cdots & \cdots & \cdots & \cdots \\ (z_m,z_1) & (z_m,z_2) & \cdots & (z_m,z_m) \end{pmatrix} \tag{4}$$

where $z_m$ is the $m$ Fringe Zernike polynomial evaluated at points in the non-circular domain. We can analysis of linear correlation. If the $m$ vectors are linear correlation, the $n$ linearly independent vectors can be selected to form a new vector group $A = (z_1,z_2,\ldots,z_n)^T$. The vector group $A$ is orthogonalized to obtain an orthogonal set of vectors $B = (b_1,b_2,\ldots,b_n)^T$. Concerning the orthogonal decomposition of the Gram matrix, because the Gram matrix is real, symmetric, and positive definite, there must exist a unitary matrix $Q$ such that,

$$Q'GQ = \mathrm{diag}(\lambda_1,\lambda_2,\ldots,\lambda_n) \tag{5}$$

where $\lambda_1,\lambda_2,\ldots,\lambda_n$ are eigenvalues and $Q'$ is the eigenvector matrix of $G$, and $Q' = Q^{-1}Q' = Q^{-1}$. The linear transformation is written as,

$$B = Q'A \tag{6}$$

where $B = (b_1,b_2,\ldots,b_n)^T$, $A = (z_1,z_2,\ldots,z_n)^T$ and $Q'$ are the linear transformation matrices of the two vectors. It follows from the Gram matrix of vector $B$ [14] that,

$$
\begin{aligned}
G(B) &= \begin{pmatrix} (b_1,b_1) & (b_1,b_2) & \cdots & (b_1,b_n) \\ (b_2,b_1) & (b_2,b_2) & \cdots & (b_2,b_n) \\ \cdots & \cdots & \cdots & \cdots \\ (b_n,b_1) & (b_n,b_2) & \cdots & (b_n,b_n) \end{pmatrix} \\
&= Q' G(A) (\overline{Q'})' \\
&= Q' G(A) Q \\
&= \mathrm{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n)
\end{aligned} \tag{7}
$$

and

$$
|G(B)| = \prod_{i=1}^{n} (b_i, b_i). \tag{8}
$$

When the equal sign is established, the generalized Hadamard inequality theorem states that,

$$
(b_i, b_j) = 0, i \neq j, i, j = 1, 2, \ldots, n. \tag{9}
$$

$B$ is a set of orthogonal vectors in Euclidean space. At this point, the items in the Fringe Zernike polynomials are combined to form a new orthogonal basis function in the non-circular domain. In the optimization process of the optical system, the non-circular domain is the lens aperture. The non-circular orthogonal Zernike surface achieves orthogonal decomposition with $B$ as the orthogonal vector basis function. The basic functions of different lenses are different owing to the different shapes of the apertures of the different lenses.

### 2.2. Establishment of mapping relations with wavefront aberrations and solution of the optimized surface

In section 2.1, the orthogonal basis functions on each lens are obtained. The correspondence between each basis function and the wave aberration of the optical system is then established.

Each basis function takes the same coefficients and is evaluated at points in the non-circular domain to form the corresponding surface. By adding the surface to the corresponding lens through interacting software (CODEV), the effect of each basis function on the wavefront aberration of different fields is obtained, and the variation of the wavefront aberration is described by Fringe Zernike polynomials. The sensitivity matrix $S$ of the basis function on the corresponding lens concerning the wavefront aberration is thus established, and the mapping relationship between the non-circular orthogonal Zernike surface and the wavefront aberration is obtained. The sensitivity matrix is

$$
S = \begin{pmatrix}
\overbrace{\underbrace{BM_{11}}_{f_1\left\{\begin{matrix}\Delta z_1 \\ \cdots \\ \Delta z_{37}\end{matrix}\right.} \quad \underbrace{BM_{1i}}_{f_1\left\{\begin{matrix}\Delta z_1 \\ \cdots \\ \Delta z_{37}\end{matrix}\right.}}^{M_1} & \cdots & \overbrace{\underbrace{BM_{m1}}_{f_1\left\{\begin{matrix}\Delta z_1 \\ \cdots \\ \Delta z_{37}\end{matrix}\right.} \quad \underbrace{BM_{mj}}_{f_1\left\{\begin{matrix}\Delta z_1 \\ \cdots \\ \Delta z_{37}\end{matrix}\right.}}^{M_m} \\
\cdots \quad \cdots & \cdots & \cdots \quad \cdots \\
f_n\left\{\begin{matrix}\Delta z_1 \\ \cdots \\ \Delta z_{37}\end{matrix}\right. \quad f_n\left\{\begin{matrix}\Delta z_1 \\ \cdots \\ \Delta z_{37}\end{matrix}\right. & \cdots & f_n\left\{\begin{matrix}\Delta z_1 \\ \cdots \\ \Delta z_{37}\end{matrix}\right. \quad f_n\left\{\begin{matrix}\Delta z_1 \\ \cdots \\ \Delta z_{37}\end{matrix}\right.
\end{pmatrix} \tag{10}
$$

where $M_1, M_2, \ldots, M_m$ denote different lenses, $BM_{11}$, $BM_{12}, \ldots, BM_{1i}$ denote the $i$th basis function corresponding to the first lens, $BM_{m1}, BM_{m2}, \ldots, BM_{mj}$ denote the $j$th basis function corresponding to the $m$th lens, $f_1, f_2, \ldots, f_n$ denote different fields, and $\Delta z_1, \Delta z_2, \ldots, \Delta z_{37}$ denote changes in the wavefront aberration.

The difference between the wavefront aberration of the current system and the optimized target value is denoted $\Delta W$. The corresponding lens basis function coefficient matrices are denoted $CM_1, CM_2, \ldots, CM_m$. Adopting SVD [15, 16] to eliminate orthonormal basis functions that are not sensitive to the wavefront,

$$
S = UDV^T \tag{11}
$$

where $U$ and $V$ are orthogonal and $D$ is square diagonal. The pseudo inverse of $S$ is

$$
S^{-1} = VD^{-1}U^T. \tag{12}
$$

The required surfaces $SM_1, SM_2, \ldots, SM_m$ are calculated for optimization based on the Gauss–Newton algorithm. We have

$$
C = (S)^{-1} \cdot \Delta W \tag{13}
$$

where

$$\Delta W = \begin{pmatrix} f_1 \begin{Bmatrix} \Delta z_1 \\ \cdots \\ \Delta z_{37} \end{Bmatrix} \\ \cdots \\ f_n \begin{Bmatrix} \Delta z_1 \\ \cdots \\ \Delta z_{37} \end{Bmatrix} \end{pmatrix} \quad C = \begin{pmatrix} CM_1 \begin{Bmatrix} CM_{11} \\ \cdots \\ CM_{1i} \end{Bmatrix} \\ \cdots \\ CM_m \begin{Bmatrix} CM_{m1} \\ \cdots \\ CM_{mj} \end{Bmatrix} \end{pmatrix}. \tag{14}$$

Taking the $m$th lens as an example, the formula for calculating the required surface for optimization is,

$$SM_m = CM_{m1}BM_{m1} + CM_{m2}BM_{m2} + \ldots + CM_{mj}BM_{mj}. \tag{15}$$

## 3. Validation

The system is optimized using the free-form optimization process shown in figure 1 to validate the free-form optimization method based on the mapping of the orthogonal Fringe Zernike surface to wavefront aberrations and to demonstrate the effectiveness of the method in optimizing the residual aberrations of very-small-aberration off-axis optical systems.

### *3.1. Non-circular orthogonal Fringe Zernike surface for optimizing system with non-circular aperture*

The system with the parameters in table 1 shown in figure 2 is optimized. This is done in a simple-to-complex sequence. First, a small freeform deviation is added to a single lens surface. Considering the strong coupling when too many surfaces are selected for optimization, then we choose two lens surfaces to add small freeform deviations and validate that the method is also effective to reduce residual aberration. The principle of selection is to be able to correct both the residual aberrations common to each field and the individual aberrations of each field. The small deviation surface for optimizing the system is at the nanoscale. The curvature of the original surface determines the reference sphere for testing. The curvature of the original surface of M3 is 0.003 $mm^{-1}$ and the surface curvature of the original surface of M4 is 0.002 $mm^{-1}$.

*3.1.1. Adding a small free-form surface to a single lens.* The lens M3 is placed close to the lens M2 as a diaphragm in the optical system. The footprint of each field is more discrete than that of the other lenses. When M3 is chosen for adding a small free-form surface, it is useful for the correction of both specific and common residual aberrations of fields.

For calculating the surfaces to establish the mapping, basis functions should be determined. If the basis function group is orthogonally complete, the product of each term of the basis function group with the other terms tends to zero according to the definition of orthogonality. The degree of non-orthogonality determines the number of terms of basis functions. It is calculated between each basis function and each other basis function. The basis function is the set of Fringe

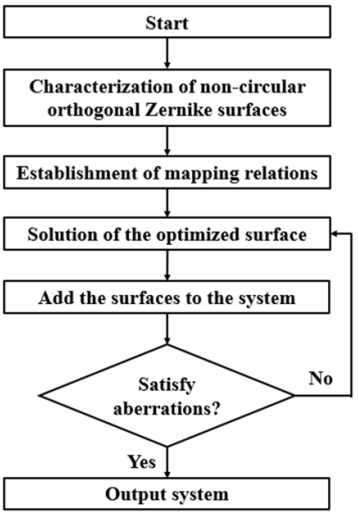


**Figure 1.** Optimization for a free-form surface.

**Table 1.** Basic system parameters.

| Item | Specification |
|---|---|
| Numerical aperture | 0.33 |
| Reduction ratio | 4 |
| Image field size | 26 mm × 2 mm |

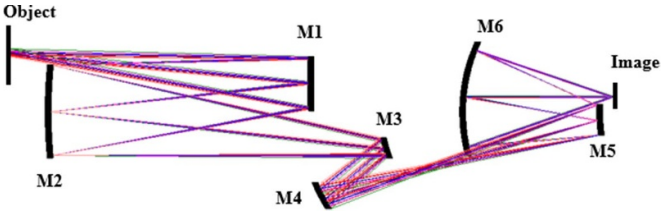


**Figure 2.** System structure.

Zernike polynomials with different numbers of terms. Taking M3 as an example, figure 3(a) shows the non-orthogonality between the basis functions comprising 192 terms of Fringe Zernike polynomials, which is close to 1. The basis functions comprising 165 terms maintain good orthogonality. We thus identify a set of basis functions that can be used to describe the non-circular orthogonal Zernike surface of M3, as shown in figure 3(c) for the 165-term basis function surface of M3.

After constructing the basis functions used to characterize the non-circular orthogonal Zernike surface of M3, the required non-circular orthogonal Zernike surface for the optimization is solved, as shown in figure 4. The surface added to optimize the system reduces the residual wavefront aberration from 0.549 nm (root-mean-square, RMS) to 0.443 nm (RMS) and residual aberration reduced by 19%.

When the circular orthogonal surface is used, free-form surface for optimization can be obtained based on the method in section 2.2. For example, 165 terms of the orthogonal Fringe Zernike are used to optimize the same system. The surface for optimization is shown in figure 5. Although the residual wavefront aberration can be 0.444 nm (RMS), the distortion

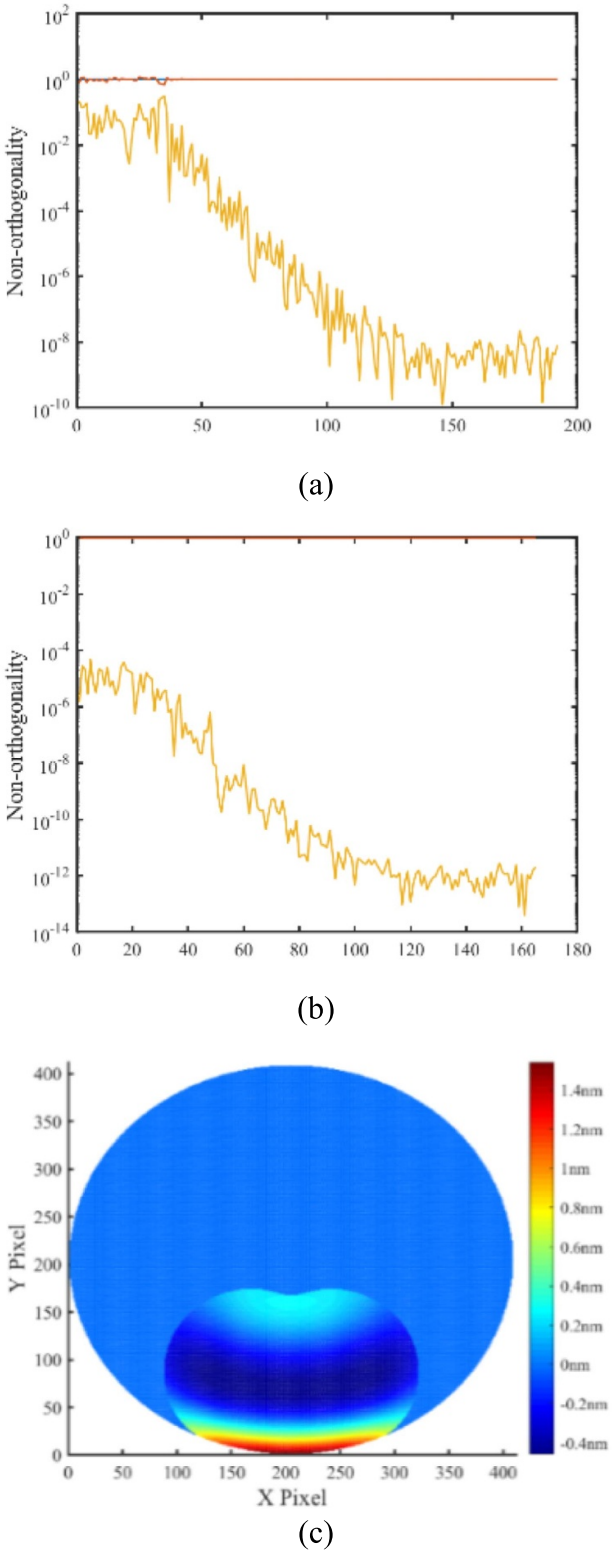


**Figure 3.** (a) Non-orthogonality between the basis functions comprising 192 terms of Fringe Zernike polynomials, (b) non-orthogonality between the basis functions comprising 165 terms of Fringe Zernike polynomials, and (c) 165-term basis function surface of the element M3.

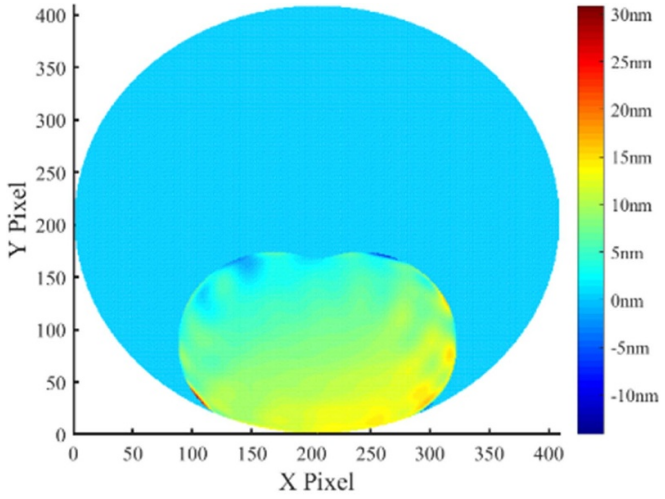


**Figure 4.** Free-form surface of M3 required for optimization.

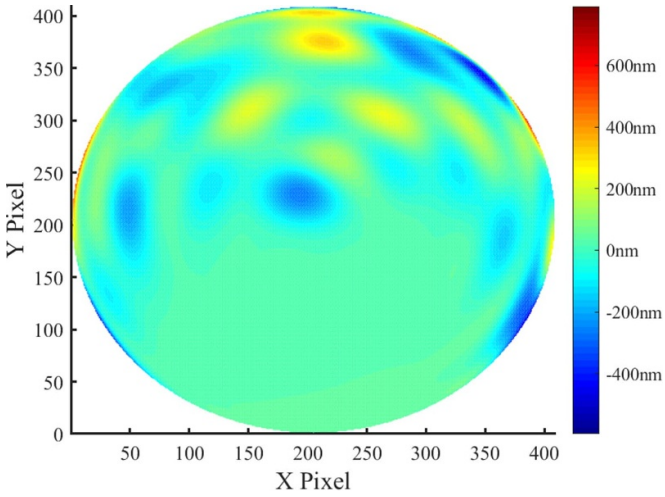


**Figure 5.** Free-form surface in circular domain of M3 required for optimization.

is 46.5 nm due to the large deviation of the surface. When the non-circular orthogonal Fringe Zernike is used, the distortion is 6.6 nm. The non-circular orthogonal Fringe Zernike can reduce the deviation, at the same time it can reduce the residual aberration. The shape of the aperture determines whether the freeform surface polynomial is circularly or non-circularly orthogonal.

*3.1.2. Adding small free-form surfaces to two lenses.* To verify the effectiveness of the method in optimizing the combination of multiple lenses, two lenses are selected to be added to small free-form surfaces. In addition to M3, the M4 is chosen to owe to its relatively independent contribution to each field and to further correct for residual wavefront aberrations in each field. The number of Fringe Zernike polynomial terms used to build the set of basis functions in the aperture of M4 is

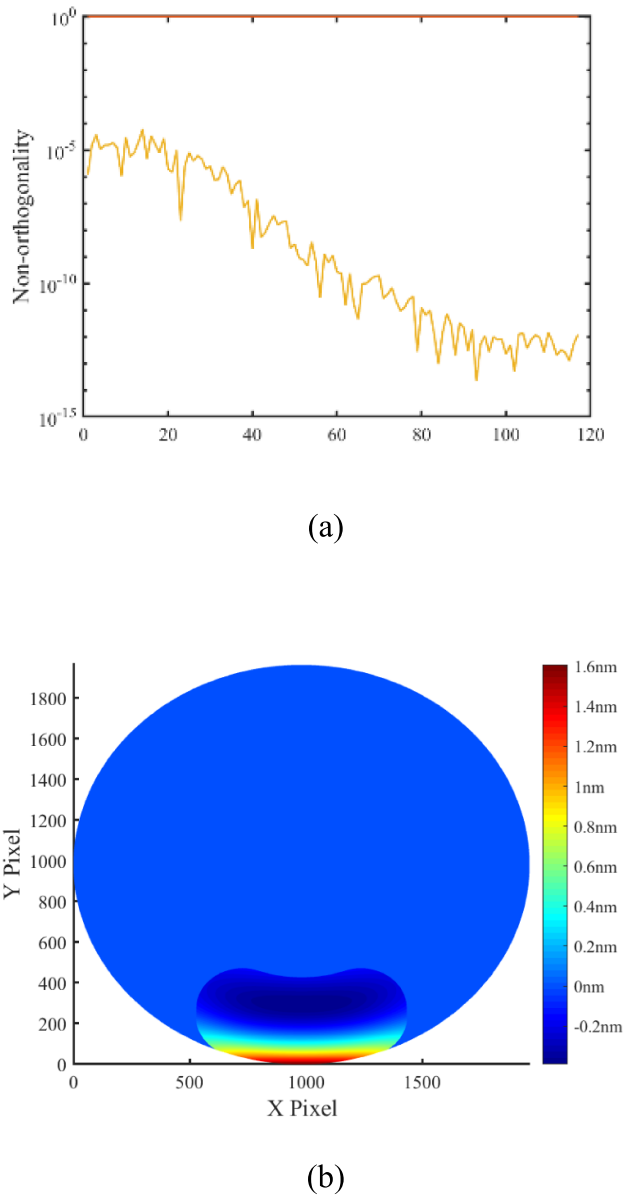


**Figure 6.** (a) Non-orthogonality between the basis functions comprising 117 terms of Fringe Zernike polynomials and (b) 117-term basis function surface of the element M4.

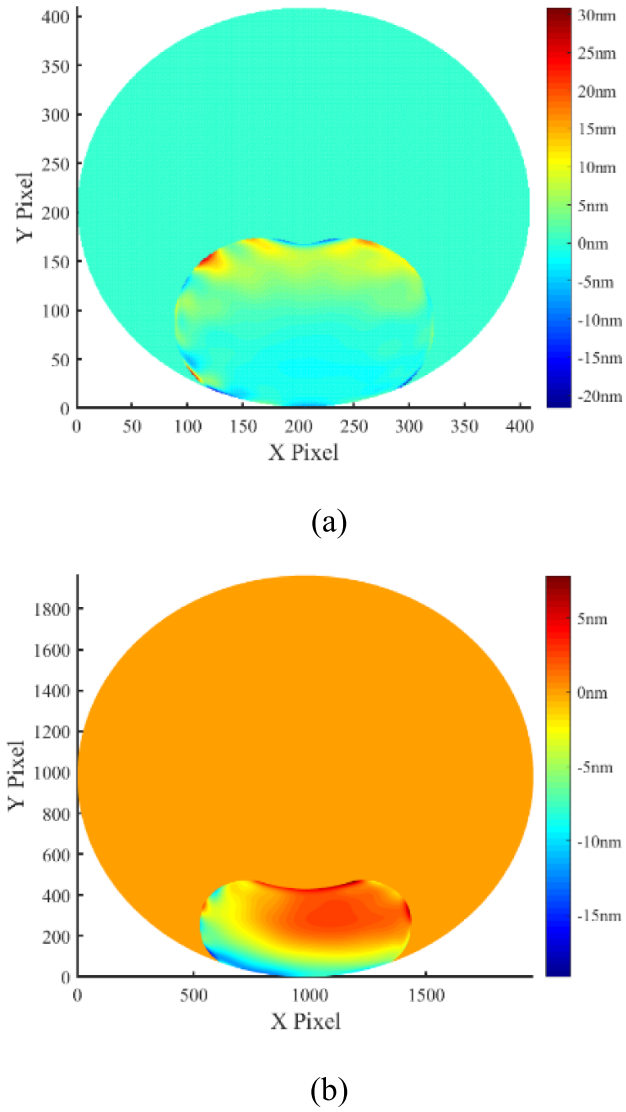


**Figure 7.** (a) Free-form surface of the element M3 required for optimization and (b) free-form surface of M4 required for optimization.

determined using the method described and the corresponding set of basis functions for M4 is built, as shown in figure 6, with 117 terms used.

Using the method described, the non-circular orthogonal Zernike surfaces of M3 and M4 for optimization are required, as shown in figure 7. The residual wavefront aberration is reduced from 0.549 nm (RMS) to 0.393 nm (RMS) by adding the surfaces to the system and residual aberration reduced by 28%.

### *3.2. Circular orthogonal Fringe Zernike surfaces for optimizing system with circular aperture*

To verify the effectiveness of the method in the circular aperture, the system with a circular aperture, as shown in figure 8, is optimized by M6. The surface curvature of the original surface of M6 is 0.001 $mm^{-1}$.

Using the method described, the non-circular orthogonal Zernike surface of M6 for optimization is required, as shown in figure 9. The residual wavefront aberration is reduced from 2 nm (RMS) to 0.47 nm (RMS) by adding the surface to the system and residual aberration reduced by 76%.

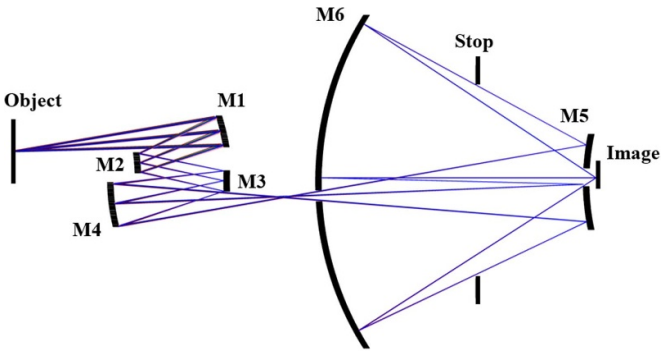


**Figure 8.** System structure.

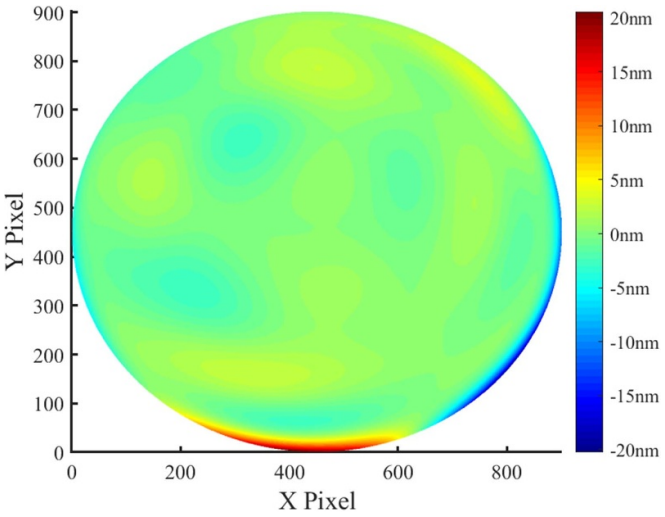


**Figure 9.** Free-form surface of the element M6 required for optimization.

## 4. Summary

This paper presented a free-form optimization method for optical systems based on the mapping relationship between the non-circular orthogonal Zernike surface and the system wavefront aberration. The presented method effectively reduces the residual wavefront aberration of very small-aberration off-axis optical systems. The method characterizes the orthogonal Zernike surface within an arbitrarily shaped aperture and applies it to optimize systems. The surface can reduce polynomial coupling and increase the freedom of the system. The mapping between the surface and wavefront aberration is obtained by establishing the sensitivity matrix of the basis functions concerning the wavefront aberration, and the optimized surface is solved by adopting SVD. This method improves the efficiency of free-form surface optimization and reduces the tedious trial-and-error process. The method was applied to the optimization of very-small-aberration off-axis optical systems with arbitrary lens aperture, and the residual aberration of the system reduced from 0.549 nm (RMS) to 0.443 nm (RMS) when using a single lens and to 0.393 nm (RMS) when using two lenses. The optical system with a circular aperture is optimized and the residual aberration of the system is reduced from 2 nm (RMS) to 0.47 nm (RMS).

For dynamic range, this method allows sub-nanometer, deep sub-nanometer wavefront aberration convergence at surface deviation in the nanoscale. The spatial resolution is up to $z_{37}$ that wavefront aberration is described by Fringe Zernike polynomials. This method is applicable in the low-frequency range. The surface for optimization cannot have the large deviation and it could destroy other indicators of the original system.

## Data availability statement

All data that support the findings of this study are included within the article (and any supplementary files).

## Acknowledgments

This research was supported by the National Science and Technology Major Project.

## References

[1] Liu Z, Lv Z and Gao L 2020 Optical system design based on lower aberration compensation method *Infrared Phys. Technol.* **105**
[2] Zhu H, Cui Q, Piao M and Zhao C2014 Design of a dual-band MWIR/LWIR circular unobscured three-mirror optical system with Zernike polynomial surfaces *Proc. SPIE* **9272** 92720W
[3] Iskander D R, Collins M J and Davis B 2001 Optimal modeling of corneal surfaces with Zernike polynomials *IEEE Trans. Biomed. Eng.* **48** 87–95
[4] Nijboer B R A 1943 The diffraction theory of optical aberrations *Physica* **10** 679–92
[5] Duan H-F, Yang Z-P and Wang S-Q 2002 Model wavefront reconstruction of Shack-Hartmann sensor on arbitrary area and wavefront expression by Zernike polynomials *Chin. J. Lasers* **29** 517–20
[6] Yang T, Zhu J, Hou W and Jin G 2014 Design method of freeform off-axis reflective imaging systems with a direct construction process *Opt. Express* **22** 9193–205
[7] Gong T, Jin G and Zhu J 2017 Point-by-point design method for mixed-surface-type off-axis reflective imaging systems with spherical, aspheric, and freeform surfaces *Opt. Express* **25** 10663–76
[8] Cheng D, Wang Y, H H and Talha M M 2009 Design of an optical see-through head-mounted display with a low f-number and large field of view using a freeform prism *Appl. Opt.* **48** 2655–68
[9] Cheng D, Wang Y and Hua H 2010 Automatic image performance balancing in lens optimization *Opt. Express* **18** 11574–88
[10] Fuerschbach K, Rolland J P and Thompson K P 2011 A new family of optical systems employing phi-polynomial surfaces *Opt. Express* **19** 21919–28

[11] Swantner W and Chow W W 1994 Gram-Schmidt orthonormalization of Zernike polynomials for general aperture shapes *Appl. Opt.* **33** 1832–7
[12] Upton R and Ellerbroek B 2004 Gram-Schmidt orthogonalization of the Zernike polynomials on apertures of arbitrary shape *Opt. Lett.* **29** 2840–2
[13] Shi X 2009 The positive semi-definite of Gram matrix and its application *J. Shaoyang Univ.* **6** 15–17
[14] Lirong Y and Sun S 1996 Gram matrix and Hadamard inequality *J. Math. Technol.* **12** 142–5
[15] Golub G and Kahan W 1965 Calculating the singular values and pseudo-inverse of a matrix *J. Soc. Ind. Appl. Math.* B **2** 205–24
[16] Akritas A G and Malaschonok G I 2004 Applications of singular-value decomposition (SVD) *Math. Comput. Simul.* **67** 15–31