

Freeform OST-HMD system with large exit pupil diameter and vision correction capability

DEWEN CHENG,¹  JIAXI DUAN,^{1,2} HAILONG CHEN,¹ HE WANG,² DANYANG LI,¹ QIWEI WANG,^{1,2} QICHAO HOU,¹ TONG YANG,¹ WEIHONG HOU,² DONGHUA WANG,² XIAOYU CHI,³ BIN JIANG,³ AND YONGTIAN WANG^{1,*}

¹Beijing Engineering Research Center of Mixed Reality and Advanced Display, School of Optics and Photonics, Beijing Institute of Technology, Beijing 100081, China

²Ned Co., Ltd., Beijing 100081, China

³Goertek Co., Ltd., Weifang 261031, China

*Corresponding author: wyt@bit.edu.cn

Received 9 August 2021; revised 31 October 2021; accepted 1 November 2021; posted 2 November 2021 (Doc. ID 440018); published 9 December 2021

Compactness and light weight, large exit pupil diameter and distance, small distortion for virtual image, and see-through light paths are pivotal factors to achieve a better, wearable experience of optical see-through head-mounted displays (OST-HMDs). In addition, light efficiency of the virtual image light path is an important factor for heat dissipation in HMD devices. This paper presents a new type of OST-HMD optical system that includes three wedge-shaped freeform prisms and two symmetric lenses. Based on a 0.71 in. microdisplay, an OST-HMD prototype with a diagonal field of view (FOV) of 45.3°, an F -number ($F/\#$) of 1.8, an exit pupil size of 12 mm \times 8 mm, and an eye relief of 18 mm is demonstrated. The maximum value of distortion of the final system is 0.6% and 0.4% for virtual image and see-through light path, respectively. The overall dimension of the optical system per eye is no larger than 30 mm (width) \times 40 mm (height) \times 14 mm (thickness), and the weight of the optical module including lenses, holder, and microdisplay is 12.8 g. The light efficiency of the virtual image light path is up to 50% higher than those of other OST-HMD optical solutions. © 2021 Chinese Laser Press

<https://doi.org/10.1364/PRJ.440018>

1. INTRODUCTION

An optical see-through head-mounted display (OST-HMD) can optically superimpose computer-generated virtual scenes onto the views of a real-world scene in mixed or augmented reality systems [1]. Rather than a video see-through approach, the OST-HMD has an actual direct view of the space in front of the user, which introduces minimal degradation to real-world scenes. The OST-HMDs are widely employed in the fields of medicine [2], manufacturing [3], scientific visualization [4], aviation [5], education, and entertainment [6].

However, it is difficult to achieve a large field of view (FOV), low $F/\#$, compact, and nonintrusive OST-HMD through a nonpupil-forming system [1]. Optical designers have tried various optical solutions for such problems, including, but not limited to, catadioptric systems, geometric waveguides, diffraction waveguides, freeform prisms, light fields, and metasurfaces. Each of these technologies has a unique disadvantage. For example, although a catadioptric lens system can avoid the effects of stray light and ghost images, achieving a compact form may deem quite challenging [7]. Similarly, the optical distortion of the projected image is difficult to correct during the design process [8–10]. The OST-HMD integrating the projection optics

and waveguide easily meets the requirements of small thickness and light weight. However, the stray light of the geometrical waveguide severely degrades the display quality when high processing technology is required [11–14]. For diffraction waveguides, it is still difficult to develop full-color, high-resolution displays that are free of artifacts such as speckles [15–20]. Improvements in design and manufacturing ability have made freeform optics broadly applied to imaging systems [21–25]. The performance of the freeform prism OST-HMD has progressed significantly. Thus, a wide FOV and low $F/\#$ OST-HMD can be achieved using a freeform prism while maintaining compactness and light weight, but large distortion and anamorphic ratio remain [26–28]. A light-field near-eye display can obtain a corrected depth perception of the virtual information in OST-HMD applications. Transparent display panels are rare and have low resolution, which cannot be achieved in a short time [29,30]. Metasurfaces have shown remarkable potential for manipulating light in imaging systems [31–36]. Metalens, produced only by nanoimprinting, can overcome the existing bottleneck imposed by the narrow FOV and bulkiness of the current systems. The limitation of processing methods is their inability to be widely promoted [37].

OST-HMD system structures are classified into four types: general catadioptric system, diffractive waveguide, geometrical waveguide, and freeform prism. Table 1 lists the module views, pros, and cons of these four solutions as well as commercially available HMD products based on these solutions, such as ODG R-9 [38], Magic Leap 1 [39], Lumus DK-52 [40], and our previous design in Ref. [26].

The development of freeform surface description and design methods has promoted a revolutionary leap for the advancement of the OST-HMD technology. Although the freeform element is thicker than waveguides, plastic material has much less specific gravity and keeps its weight. Its fundamental difference from waveguides is that light efficiency of freeform prisms is no longer restricted by diffraction efficiency or reflectance of reflective mirrors, so its light efficiency is higher than those of waveguides and general catadioptric systems. That is why the waveguide OST-HMD system needs a Pico projection system, and the overall thickness of the waveguide system is similar to or even larger than that of the freeform optics solution. Briefly, the freeform prism is an advantageous modern solution to achieve OST-HMD designs with better display performance.

In this paper, we present the design of an OST-HMD system that includes three wedge-shaped freeform prisms and two lenses. Using a 0.71-in. (1 in. = 2.54 cm) microdisplay, the system offers a diagonal FOV of 45.3°, an $F/\#$ of 1.8, an eye relief of 18 mm, an exit pupil diameter as large as

12 mm \times 8 mm, and vision correction capability. The distortion of the virtual image light path and the see-through light path is only 0.6% and 0.4%, respectively, indicating that the distortion is well controlled. The rest of the paper is organized as follows. In Section 2, we review and compare previous research using freeform prisms in HMD designs. In Sections 3–6, we describe the display principle and system specifications of our design, design of the virtual image light path, system optimization methods, and prototype testing results.

2. PREVIOUS RESEARCH

The concept and development of freeform prisms have been elaborated in previous studies [26]. Freeform prisms have shown significant advantages in the design of compact HMD systems. The wedge-shaped freeform prism provides full play to the characteristics of the total internal reflection and minimizes light loss. The correcting plate effectively corrects the see-through light path, reduces deformation of the external scene, and achieves high-quality virtual and real registration. After extensive research, the freeform prism has been applied to current OST-HMD systems and has been commercially available. Many companies, such as Olympus [41], eMagin [42], and Rockwell Collins [43], have released HMDs or optical modules based on freeform prisms.

In our previous research paper [26], we designed an OST-HMD based on a freeform prism. Along with the deep

Table 1. Comparison of Four OST-HMD Optical Solutions

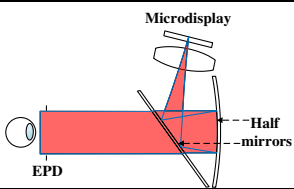
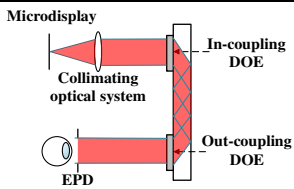
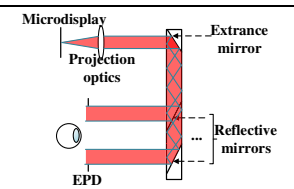
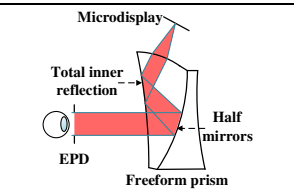
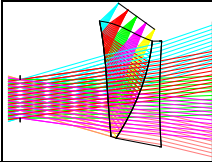
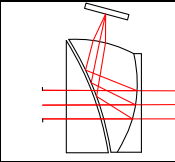
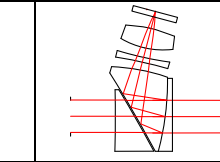
	Module view	Pros and cons	HMD Products	FOV (°) and Resolution (pixels)
General catadioptric type		High resolution. Normal light efficiency. Small see-through distortion. Less stray light.	ODG R9 [38]	50° 1920 \times 1080
Diffractive waveguide		High resolution. Low light efficiency. No see-through distortion. Being ultra-thin. Nonuniform light distribution. Serious stray light and rainbow effect.	Magic Leap 1 [39]	50° 1268 \times 720
Geometrical waveguide		High resolution. Low light efficiency. No see-through distortion. Being ultra-thin. Serious stray light.	Lumus DK-52 [40]	40° 1280 \times 720
Freeform prism		Normal resolution. High light efficiency. Normal see-through distortion. Large field of view. Uniform light distribution. Less stray light.	Our previous design [26]	53° 800 \times 600

Table 2. Specifications of Different Freeform Prisms

Module view	Previous Design	Similar Design	Present Design
			
Full diagonal FOV	53.5°	60°×36°	45.3°
Eye relief (mm)	18.25	/	18
Exit pupil diameter (mm)	8	10	12×8
Effective focal length (mm)	15	22	21.6
Diagonal image size (in.)	0.61	/	0.71
<i>F</i> /#	1.875	2.2	1.8
Distortion (of the virtual path)	12%	/	0.6%
Diopter adjustment	No	/	Yes
Thickness (mm)	12	20	14

development of the study, different types of freeform prism OST-HMDs have been commercially available [44]. However, the shortcomings of the design in Ref. [26] still intervene with the wearable experience. For example, if the distortion of the virtual image light path is up to 12%, electronic correction is required when this optical system is applied in OST-HMD applications. Although the see-through distortion is only 1.4%, the design process of the auxiliary prism is complicated and difficult, and the image quality of the real scene decreases. Diopter adjustment is also an important function for OST-HMDs. Previous designs cannot achieve diopter adjustment, causing some users to wear two pairs of glasses.

To address the problem of see-through distortion of freeform prism OST-HMD, an auxiliary lens was added between the main wedge-shape prism and the eye [45]. This auxiliary lens can also achieve the diopter adjustment function [46,47]. In Table 2, we list the optical path evolution process. In this paper, we compare our previous design, the three-piece prism design [45], and our present design across all aspects.

The current solution entails that the distortion of the virtual image light path is less than 0.6% without electronic correction, projecting a clear and undistorted image from the microdisplay to the user's eye. In addition, a larger exit pupil diameter is achieved in this new design, thereby simplifying the structure by avoiding the pupil adjustment mechanism. Due to the addition of the auxiliary lens, which makes the optical surface near the human eyes flat or spherical, it is much easier to design the see-through light path in the present design. In addition, we achieve diopter adjustment by changing the shape of the first surface of the auxiliary lens to meet the diopter needs of different users, while other optical elements remain unchanged. This process is described in Section 4.B. Although the FOV of the present design is slightly smaller than that of our previous design, it is similar to other optical solution types and is still able to meet the usage requirements of AR applications.

3. DISPLAY PRINCIPLE AND SYSTEM SPECIFICATIONS

An optical see-through HMD typically consists of an optical path for viewing a displayed virtual image and a path for viewing a real-world scene directly. The previous design included only two freeform elements: a main prism and a corrector plate. Although it had a larger FOV, lighter weight, and simpler structure, it retained unsolvable defects, as mentioned in Section 2. Therefore, it is necessary to improve the performance of previous designs. The OST-HMD optical system presented in this study is shown in Fig. 1. The overall system is set to be symmetric about the *YOZ* plane, comprising three freeform prisms and two rotationally symmetric lenses.

In the present design, two extra lenses close to the microdisplay are used to improve the image quality of the system,

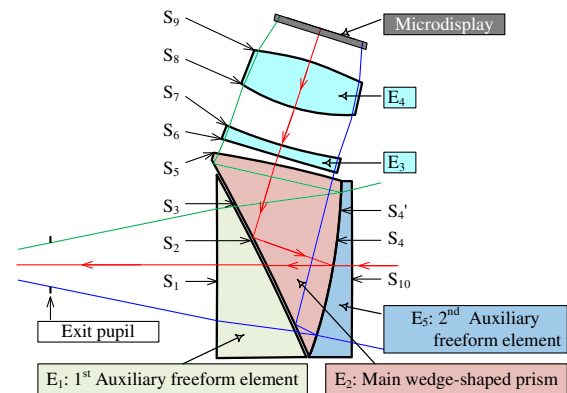


Fig. 1. Layout of the final OST-HMD system. S_1 and S_2 denote optical surfaces of the 1st auxiliary freeform element E_1 , S_3 – S_5 denote optical surfaces of the main wedge-shaped prism E_2 , S_6 – S_{10} denote the optical surfaces of two rotationally symmetric lenses E_4 and E_5 , S_4' and S_{10} denote optical surfaces of the 2nd auxiliary freeform element E_5 .

balancing the performance across the entire exit pupil. The wedge-shaped freeform main prism E_2 assumes the larger part of the optical power, magnifying the image from the microdisplay. The 1st auxiliary lens is placed between the main prism E_2 and human eye and is effective in correcting the distortion of the system, just as it is necessary to achieve the diopter adjustment function. In addition, the 1st auxiliary lens E_1 is conducive to counteracting the optical path difference of rays of different fields caused by the off-axis setup. On the right side of E_2 , the 2nd auxiliary lens E_5 is cemented with E_2 , completely correcting the see-through distortion.

As shown in Fig. 1, the wedge-shaped freeform prism E_2 and the 1st auxiliary freeform element E_1 consist of three and two optical effective surfaces, respectively. The 2nd auxiliary freeform element is similar with E_1 . The system is optimized with rays traced from the exit pupil to the microdisplay, the optically effective surfaces are denoted by S_1 and S_2 of E_1 , S_3 – S_5 of E_2 , S_6 – S_{10} of E_4 and E_5 for virtual image path. Surfaces S'_4 and S_{10} with S_1 – S_4 are for the see-through light path. Surfaces S_2 – S_5 are freeform surfaces, while others are spherical or have a flat surface.

For the elements in the optical system, the main prism and 1st auxiliary freeform element are adjacently arranged. The surface S_2 of E_1 , is proximate to the surface, S_3 , of the main prism E_2 , and there is a small air gap between them for total internal reflection condition. Surface S'_4 of the 2nd auxiliary freeform element E_5 has the same surface shape as S_4 of the main prism E_2 to ensure that the E_5 and E_2 would be perfectly cemented. Surface S_{10} of the E_5 should be parallel to the surface S_1 of E_1 , no matter with or without requirement of diopter accommodation.

The virtual image light path consists of the 1st auxiliary freeform element E_1 , the main wedge-shaped prism E_2 , two lenses E_4 and E_5 , and the microdisplay. The light emitted from the microdisplay is first transmitted through lenses E_4 and E_3 , and refracted into the main prism E_2 by the surface, S_5 , close to the lenses. After two consecutive reflections on surfaces S_3 and S_4 in the main prism, it reaches surface S_3 for the second time. Then, light is transmitted out of the surface S_3 of the main prism into the air gap between the prism E_2 and auxiliary freeform lens E_1 and enters E_1 by surface S_2 . Finally, light is projected through surface S_1 to the exit pupil, entering the human eye.

Some transmission light is lost during the first reflection on surface S_4 to reduce the power of the microdisplay. To prevent such light from entering human eyes, it is necessary to always have the light reflected by the surface, S_3 , satisfying the total internal reflection (TIR) condition. Therefore, only if there is an air gap between the surfaces S_3 and S_2 , light travels in the direction from a medium with a higher refractive index to a medium with a lower refractive index so that total internal reflection is possible.

For the see-through light path, light from the real-world scene enters surface S_{10} of the correcting plate and transmits through surface S'_4 , which is cemented with surface S_4 of the main prism, into the main prism. In addition to the light from the microdisplay, the light transmits through the body of the main prism, into the air gap, and then through the auxiliary

lens E_1 , finally exiting from the surface, S_1 , and entering the human eye.

Surface S_4 of the main prism was coated with a beam-splitting film. Consequently, the rays from the microdisplay are partially reflected by surface S_4 , similar to the rays from the real-world scene.

In the previous design [26], the surface close to the eyes is a freeform surface, causing viewing axis deviation, undesirable distortion, and other off-axis aberrations in the view of the real-world scene. Although the correcting plate corrected the deviation of the optical axis and the off-axis aberrations introduced by the freeform prism, there is still approximately 1.4% distortion of the see-through light path. The design of the correcting plate was complicated. However, in the present design, the surface close to the eyes is flat or spherical, which simplifies the design process of the see-through light path and reduces the distortion. In addition, the 1st auxiliary lens E_1 is an essential element for achieving the diopter adjustment function. For the virtual image light path, we change the shape of surface S_1 to adjust the virtual image distance. Hence, people with blurred vision can see a virtual image without wearing two pairs of glasses. For the light path, the curvature difference between surface S_1 and surface S_{10} can be matched to adapt to the user's diopter. To achieve the diopter adjustment function, a series of auxiliary lenses with different shape of surface S_1 should be produced, and customization is available to users with different diopters.

The overall specifications of the system are listed in Table 3. Based on various considerations of volume, resolution, availability, and cost, a 0.71-in. organic LED (OLED) display with a resolution of 1920×1080 pixels and an $8.2 \mu\text{m}$ pixel size is selected. In the design of visual instruments, especially binocular HMDs, a large exit pupil diameter (EPD) is essential to prevent vignetting or loss of image due to the rotation of the eyes. A large EPD is highly adaptable to the interpupillary distances among different users and avoids mechanically adjusting the interval of the binocular optics. However, a large EPD will affect the compactness, weight, and FOV of the optical system,

Table 3. Overall Parameter Requirements of the System

Parameter	Specification
Active display area	15.8 mm \times 9.0 mm
Resolution	1920 \times 1080 pixels
Effective focal length (EFL)	21.6 mm
Exit pupil diameter (EPD)	12 mm \times 8 mm
Eye relief (ERF)	18 mm
$F/\#$	1.8
Lens type	Freeform prisms + lenses
Wavelength	486.1–656.3 nm
Field of view (FOV)	45.3° (diagonal)
Virtual image distortion ^a	$\leq 0.6\%$
See-through distortion ^b	$\leq 0.4\%$
Modulation transfer function (MTF)	$\geq 40\%$ at 50 lp/mm for all fields

^aThe difference between two kinds of distortion is that virtual image distortion describes the deforming of the image displayed on the screen, resulting from the virtual image optics.

^bThe see-through distortion is about the deformation of the real scene, which cannot be corrected by the image prewarping method.

resulting in a sharp increase in the challenge of designing low $F/\#$ systems. Considering these factors, we set the exit pupil diameter to be $12\text{ mm} \times 8\text{ mm}$ and the focal length to be 21.6 mm , resulting in an $F/\#$ of 1.8. Then, a 45.3° diagonal full FOV was determined. Although the FOV is smaller than that of our previous design [26], a larger EPD provides a better wearable experience.

In a freeform optical system, the distortion can be quite large and irregular if there are no constraints. Therefore, we need to control the imaging distortion to less than 0.6% to avoid electronic correction of the initial image source. Compared with other evaluation methods of image quality, the modulation transfer function (MTF) is more suitable for evaluating the overall image sharpness. The MTF value would be better than 40% at a spatial frequency of 50 lp/mm in the full FOV.

4. DESIGN OF THE VIRTUAL IMAGE LIGHT PATH

The additional use of the 1st auxiliary freeform lens E_1 makes the design of the see-through light path very easy; this section just needs to discuss the design of the virtual image light path. Optical freeform surfaces exhibit nonrotational symmetry. In most cases, an optical system with freeform surfaces can obtain higher image quality. In contrast to traditional optical surfaces, such as spherical or aspherical surfaces, several surface description methods of freeform surfaces can be used in the design of nonrotationally symmetric optical systems. The variety of freeform surface expressions provides more flexibility during the process of designing a freeform surface optical system. We have described, in detail, the type of freeform surface representation, strategy of freeform surface design, and method of maintaining the TIR condition in previous papers [26]. Therefore, this section focuses on structural constraints in the process of designing an OST-HMD and diopter adjustment based on the 1st auxiliary freeform lens E_1 and wedge-shaped freeform prism E_2 .

A. Structure Constraints

To guarantee that the rays of all fields pass through the optical system correctly and enter the user's eyes, it is essential to control the structure of the main freeform prism, make its three surfaces form a valid prism shape, and ensure that the center and edge thicknesses meet the manufacturing requirements. The methods to optimize the physical structure of the wedge-shaped freeform prism are as follows.

By controlling the position relations of specific points and lines, we can constrain the structure of the main prism. Figure 2 illustrates the structural control method employed in this study. Two reference rays are selected to follow the constraints in the optimization process: the upper marginal ray, R_1 , of the maximum field in the positive Y direction, and the lower marginal ray, R'_2 , of the maximum field in the negative Y direction. As shown in Fig. 2, P_{a0} , P_{a1} , P_{a2} , and P_{a3} denote the intersection points of the ray R'_2 with surfaces S_3 , S_4 , and S_5 , respectively; P_{b0} , P_{b1} , and P_{b2} denote the intersection points of ray R_1 with surfaces S_3 , S_4 , and S_5 , respectively; and P_{c0} represents the intersection point of the chief ray with surface S_5 .

By constraining the points P_{a1} and P_{b0} locating down the line Y_1 , a segment $P_{a2}P_{a3}$ of lower marginal ray R'_2 , we ensure

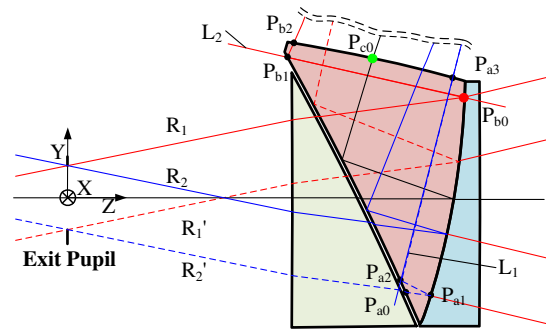


Fig. 2. Optical paths of the rays of different object fields. R_1 is the upper marginal ray of the maximum Y -direction field, and R'_2 is the lower marginal ray of the minimum Y -direction field. P_{a0} - P_{a3} are the intersection points of ray R'_2 and surfaces. P_{b0} - P_{b2} are the intersection points of ray R'_2 and surfaces. Y_1 and Y_2 are straight lines coinciding with marginal rays.

that surfaces S_3 and S_4 intersect properly so that both the upper marginal ray and the lower marginal ray can be traced through the prism without obstruction. In a similar manner, the points P_{a3} and P_{b2} must be constrained above the line Y_2 , a segment $P_{b0}P_{b1}$ of upper marginal ray R_1 . Consequently, the surface S_5 will, respectively, intersect the surfaces S_4 and S_3 correctly, avoiding the escape of the upper marginal ray and controlling the thickness of the prisms.

Commercial optical design and analysis software CODE V [48] develops a physical structure control function, referred to as JMRCC, which can be used to define the aforementioned constraints easily. This function computes the signed perpendicular distance from a point to a line in the YZ plane. It can, thus, create clearance or interference constraints when optimizing off-axis reflective/catadioptric systems. The input parameters of the JMRCC function are a single ray and a single point. The distance between the point and line is the output. Considering the mathematical and geometric rules, the output value should be positive when the point locates above the line; otherwise, the output value should be negative. As shown in Fig. 3, the JMRCC function builds a constraint about the distance between line L and point P , which can be expressed as

$$D = \text{JMRCC}(L, P). \quad (1)$$

According to the above discussion, the constraints to maintain the physical structure of the main prism can be given as follows:

$$\begin{cases} \text{JMRCC}(L_1, P_{a1}) \leq -0.1 \\ \text{JMRCC}(L_1, P_{b0}) \leq -0.1 \\ \text{JMRCC}(L_2, P_{b2}) \geq 0.1 \\ \text{JMRCC}(L_2, P_{a3}) \geq 0.1 \end{cases} \quad (2)$$

Similarly, we can construct specific constraints and use the JMRCC function to effectively ensure the shape of all prisms logically.

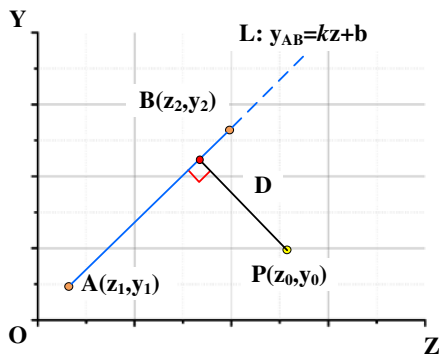


Fig. 3. Schematic diagram of JMRC basic principle, wherein L is a line defined by point A of a reference ray intersected on specified surface and its direction of propagation after the surface, and P is another point formed by another specified reference ray and surface.

B. Diopter Adjustment Based on Changing the 1st Auxiliary Freeform Lens

For most OST-HMD systems, devices did not include diopter adjustment. Users with myopia or hyperopia were required to wear two pairs of glasses simultaneously. Some devices attained diopter adaption by placing an additional corrective lens between the human eye and OST-HMD, as shown in Fig. 4(a), but such deemed tedious for users to fit it [46]. Meanwhile, the corrective lens complicates the system structure to a certain extent. Hence, the volume and weight of the device will increase accordingly. This is not in line with our original intention of lightweight and compact characteristics in our design concept.

The system presented in this paper can address this problem by customizing the 1st auxiliary freeform lens according to the users' diopeters. For users with normal diopeters, the optical surfaces, S_1 and S_{10} , can be, for example, flat surfaces to ensure that the visual image and see-through image enter the eyes in focus. For users with myopia or hyperopia, we reach diopter adjustment by changing the curvature of surface S_1 . The curvature of surface S_1 is set to negative for myopia and positive for hyperopia so that the entire optical system can meet diopter requirements for different users. Although it seems that the layout shown in Fig. 4(a) is more compact, the 1st auxiliary freeform lens E_1 presented in this paper not only benefits to the diopter adjustment, but also the image quality of the virtual light path. Details are discussed in Section 5. Also, because there is a small air gap between the E_1 and E_2 , a reasonable

mechanical structural design can realize the disassembly and replacement of E_1 .

When the user's diopter changes, the diopter adjustment function can be achieved by replacing auxiliary lens E_1 . Figure 4(b) shows an example of the freeform prism system that adjusts the diopter by changing the shape of the surface, S_1 . The positive curvature of surface S_1 corrects the user's hyperopia. Figure 4(c) shows that negative curvature of the surface S_1 corrects the user's myopia. For instance, surface S_1 is flat for users with normal vision, and for users with myopia of -4 diopter, S_1 is concave and has a radius of curvature of 131.7 mm. For users of different diopeters, this design can reduce the cost and space occupancy rate of the freeform prism group and extends the service life. For a more clear explanation, the following optimization process sets the S_1 to a flat surface.

5. OPTIMIZATION METHODS

Based on the system specifications and design methods outlined in Sections 3 and 4, the optimization process of the OST-HMD optical system is introduced in detail in this section. The optimization of the virtual image light path, correction of distortion, and balance of the system performance attained by the automatic method in the entire FOV are emphasized.

A. Optimization of the Virtual Image Light Path

The optimization process has been thoroughly analyzed in a previous paper [14], but a brief introduction is provided here. Based on the previously designed system, an auxiliary freeform lens was added to the light path. By preliminary scaling, the parameters of the initial structure gradually approximated the requirements.

Modeling this system in CODE V, the optimization process is always based on the TIR constraints, structural constraints, and the basic optical definitions such as effective focal length, outlined in Section 4. Unlike the rotationally symmetric system, the freeform prism has only single-sided symmetry. Therefore, the setup method of linear field sampling along the radial direction is no longer applicable, and the freeform prism design must be optimized for half of all FOVs sampled in a rectangular grid. An incremental optimization strategy was adopted. With the improvement in system performance during the optimization process, we gradually increased the number of field samples and the degree of freedom of the surface. The surface type of the prism is converted from spheric to aspheric, then anamorphic, and finally XY polynomial. Compared with

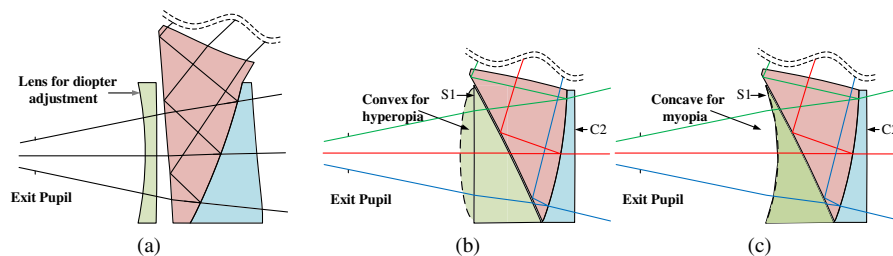


Fig. 4. Diopter adjustment diagram of freeform prism OST-HMD. (a) Traditional diopter adjustment scheme, where additional lens was added between human eye and prism. (b) Correct hyperopia through auxiliary prism with convex S_1 . (c) Correct myopia through auxiliary prism with concave S_1 .

the design in the previous paper [26], the present main prism has only two freeform surfaces, S_4 and S_5 , which reduces the difficulty of optimization.

B. Distortion Correction

In the imaging optical system, improving the imaging quality is the main task, and the ultimate goal is to obtain an image that is completely similar to the object on the focal plane. However, the actual optical system can only form a perfect image when the object point is paraxial, and a thin beam is observed. The outward expansion of the FOV and aperture destroys the concentricity of the imaging beam and produces various imaging defects; thus, the shape of the image and object cannot be completely similar.

The correction of optical distortion, in a fundamental sense, is to change the relative positional relationship between the image points on the image plane or to rearrange the image points so that the observer or camera can capture an undistorted image. The correction of distortion requires changing the conjugate relationship between the object and image points. A new conjugate relationship between the object and image can be achieved by controlling the positioning of rays reached in the image plane, resulting in a smaller distortion ratio.

Generally, for an optical system with rotational symmetry, only radial distortion should be considered. Figure 5(a) shows a schematic diagram of radial distortion, where h_{real} is the real image point in the image plane of the specified field, and h_{ideal} is the ideal image point of the same field. By controlling the distance between h_{ideal} and h_{real} , the distortion ratio can also be controlled. However, in our design, the off-axis property of the optical system causes irregular distortion. In addition to the radial direction, tangential distortion should be considered. As shown in Fig. 5(b), assuming that an ideal image without distortion is a black grid, the red dots represent the real image points of the sample fields. For a specified field, the ideal image point in the image plane was marked as P_{ideal} , and the real image point position in the image plane was marked as P_{real} . To control the distortion of our system, the constraint to make the point P_{real} close to P_{ideal} must be established.

In theory, a freeform surface has a different curvature at each point. Thus, the freeform surface can control the optical path difference, luminous flux, and even the direction and position of the rays according to the image quality requirements more easily. Therefore, the freeform surface can be used to grasp the direction of rays emitted from each FOV and control the position of each image point reached on the image plane.

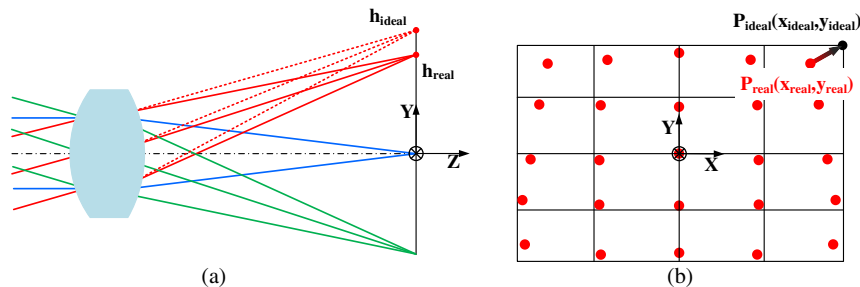


Fig. 5. Schematic diagram of distortion control. (a) For rotational symmetry system, radial distance of h_{real} and h_{ideal} affects the distortion ratio. (b) For the off-axis system, radial and tangential distances of P_{real} and P_{ideal} affect the distortion ratio.

For an OST-HMD optical system, both the distortion of the virtual image light path and see-through path must be corrected. On the one hand, the freeform system produced large distortion when projecting the image displayed in microdisplay into the human eye. Due to the off-axis property of freeform prisms, the produced distortion is also asymmetry. On the other hand, when users see the external scene through the OST-HMD, the deformation of the real-world scene caused by the see-through distortion critically affects the virtual and real registration of the system.

For the design in our previous paper [26], due to the trapezoidal shape distortion caused by the off-axis property of freeform prisms and the presence of a small amount of lens barrel distortion, the distortion grid in the virtual image light path is as high as 12% in the upper left/right corner, as shown in Fig. 6(a), which is difficult to correct. The distortion of the see-through light path was 1.4%, although a complicated design process was executed, as shown in Fig. 6(b). In addition, to control the asymmetry distortion, the effective focal lengths in the X and Y directions of the optical system are different. Consequently, there is an anamorphic ratio in both the virtual image light path and the light path, which makes the image look slightly wider than the origin, especially for the see-through light path.

However, for the current solution, the optimized distortion of the virtual image light path is now less than 0.6% before electronic correction, while the difference in the distortion rates between the X and Y directions is avoided. Regarding the see-through light path, due to the application of auxiliary lens E_1 , the distortion of the real scene is greatly decreased to 0.4% after simple optimization is implemented. Moreover, the cancelation of electronic corrections saves computing resources and reduces power consumption. The distortion diagram of the virtual image light path and the see-through light path is shown in Figs. 6(c) and 6(d), respectively, for the presented design.

Further, the MTF of the see-through light path is improved, as shown in Fig. 7. Figure 7(a) shows the MTF plot of the previous design [26], where the value is greater than 0.1 at 50 lp/mm for all sampled fields. The MTF value of the present design is 0.9 at 50 lp/mm for all fields, which implies a better see-through performance.

C. Automatic Image Performance Balancing Method

To further improve the optical performance of the system, the system performance of the entire FOV must be balanced, by continuously adjusting the weight of the sampling field and

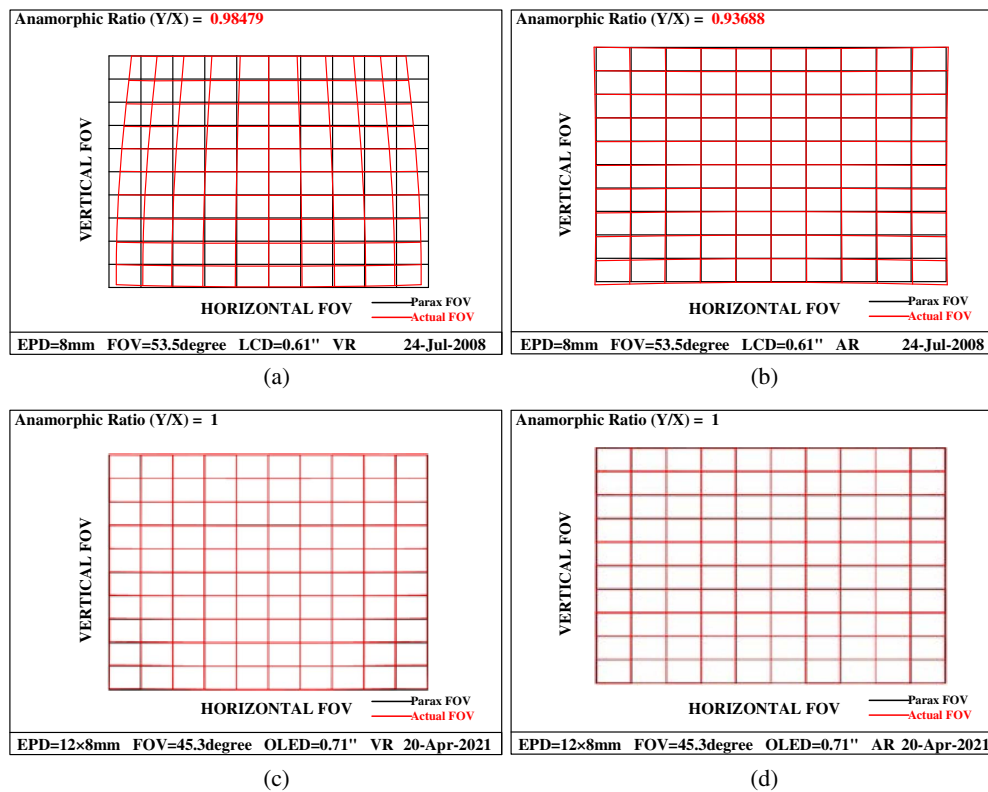


Fig. 6. Distortion grids of the previous system and present system. (a) Distortion grid of the previous virtual image light path, maximum ratio is 12% [26]. (b) Distortion grid of the previous see-through path, maximum ratio is 1.4%. [26]. (c) Distortion grid of the present virtual image light path, maximum ratio is 0.6% without anamorphosis. (d) Distortion grid of the present see-through path, the maximum ratio is 0.4% without anamorphosis.

azimuth. In most cases, the default weights provided by the software will be used for optimization, or a more appropriate set of weights will be assigned based on experience. After each optimization trial, the designer needs to adjust the weight of the sampling field and the azimuth according to the optimization results and design requirements and then use the original constraints and new constraints to optimize the system again to seek the best optimization results. This can be a long and tedious process for optical designers.

Our previous work [49] introduced an automatic image-performance balancing method. An outer loop was added for automatic weight adjustment for optimization. First, we

set the initial weights for all the sampled fields and azimuths to an equal value and evaluated the image performance before each optimization trial to calculate the weights for the next trial. The sampled fields and azimuths with performance values lower than the average were assigned higher weights and vice versa. When the exit condition is reached, the loop ends.

When designing a freeform optical system, global optimization is also a way to balance image performance. However, in this study, the design included two freeform optical surfaces; over 100 parameters were defined as variables, and 25 fields were sampled. It requires several minutes to complete a single local optimization loop and requires more than 100 h to finish

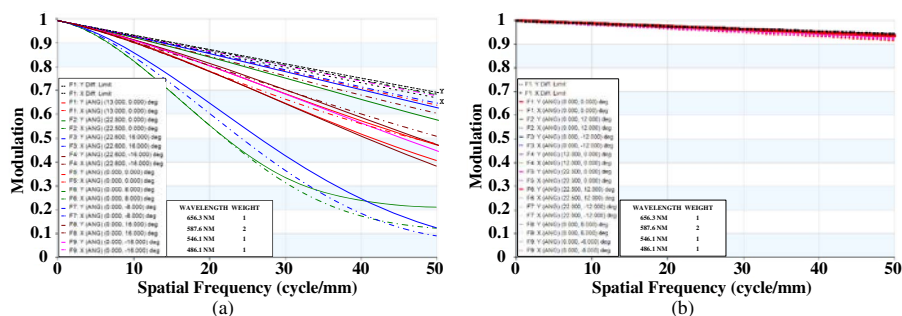


Fig. 7. See-through MTF plot of the previous and present systems. (a) MTF plot of previous design in Ref. [26], the value is higher than 0.4 for most fields at 50 lp/mm. (b) MTF plot of present design, the value is higher than 0.9 for all fields at 50 lp/mm.

a global optimization loop. However, by employing the automatic balancing method, the balancing process is completed within half an hour. Finally, through several automatic optimization trials to achieve the performance balance of the entire FOV, the overall optical performance of the system can be improved. Figure 8(a) shows the error function variations for different optimization trials of automatic balancing. The value dropped from 71 for the original design to 50 after 10 trials of optimization, and the rapid descent occurred in the first four to seven trials. After completing all optimization steps, the final system is obtained, and the layout is shown in Fig. 8(b).

The MTF plot of the optical system before and after balancing is shown in Fig. 9. Uniformity of the MTF value for full fields has been greatly improved. Besides, because of the use of a pair of positive and negative lenses, the lateral color aberration is fully corrected.

For a visual system, the relative position between human eye and exit pupil of the OST-HMD is always unstable. Eye shift or rotation caused by the difference in face shape between individuals and other factors should be considered. To evaluate the image performance when the eye shifts in the area of the exit pupil, additional MTF plots are given in Fig. 10. A major assumption is that the pupil size of the human eye is 3 mm

under normal light conditions. In the area of the exit pupil, three positions including center, 2.5 mm offset to the right from center, 2.5 mm offset upwards from center, are chosen for evaluation. The results show that the MTF values of all fields are higher than 0.4 at 50 lp/mm, indicating the MTF performance at different positions meets the design requirement.

D. Tolerance Analysis

Tolerance analysis is an essential part of optical design because it is closely related to optical fabrication, alignment, and cost. By carefully evaluating the parameter-variant/image-effect interaction and paying particular attention to the negative and positive aspects of these interactions, we can avoid overdesign and assign tight tolerances only where needed. Such analysis helps determine the acceptable limits for different fabrication and alignment errors.

Unlike the rotationally symmetric system, the tolerance analysis of the freeform system is more complex, and each surface has many optical parameters that would affect the result. In Ref. [27], we introduced the process of tolerance analysis of a system with freeform in detail. For the design presented in this paper, an overall tolerance analysis was conducted using the

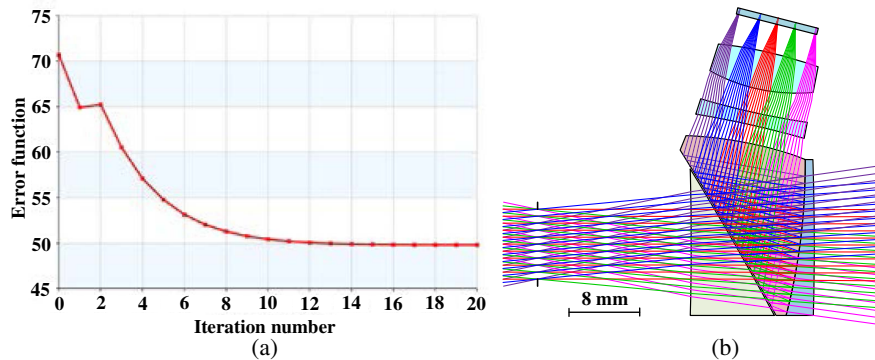


Fig. 8. (a) Error function variation curve. (b) Final optical layout of the OST-HMD.

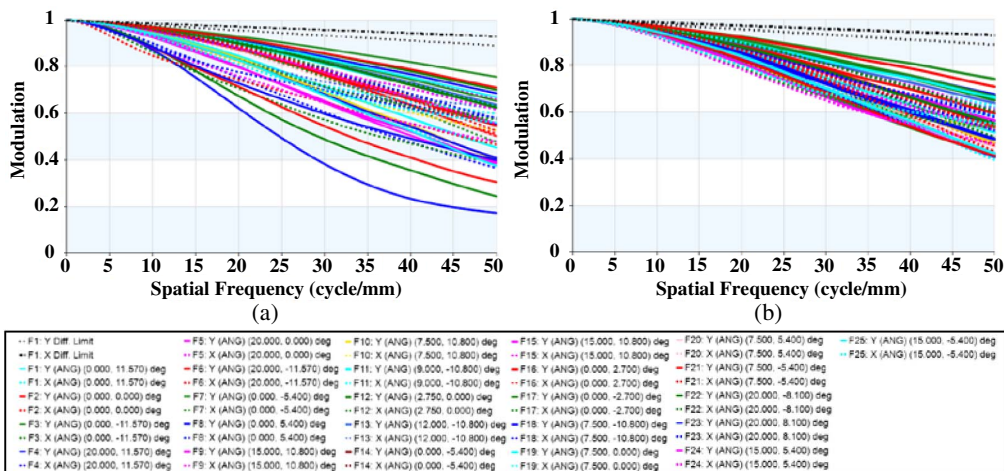


Fig. 9. MTF plot of the optical system. (a) The MTF plot before automatic balancing, MTF value is higher than 0.18 at 50 lp/mm for all fields. (b) The MTF plot after automatic balancing, MTF value is higher than 0.4 at 50 lp/mm for all fields.

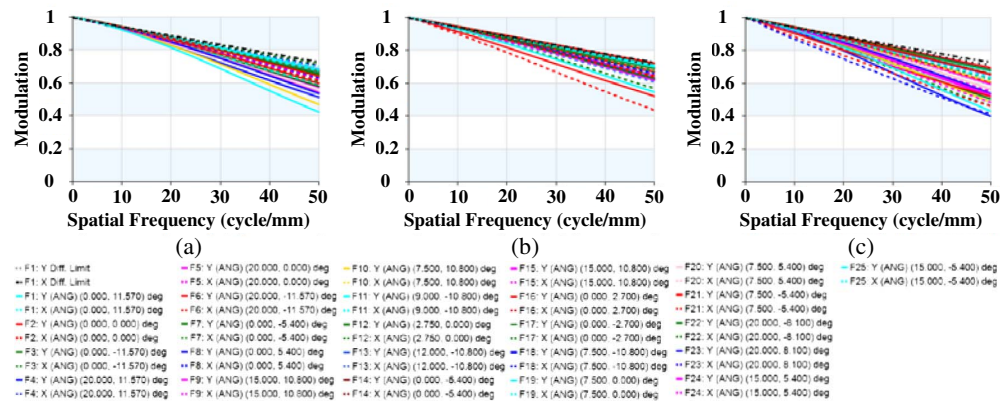


Fig. 10. MTF plots for different eye positions. (a) MTF plot when eye locates in center of eyebox. (b) MTF plot when human eye moves 2.5 mm to the right. (c) MTF plot when human eye moves up 2.5 mm.

tolerance values listed in Table 4. The position and tilt of the microdisplay are defined as compensators.

Figure 11 plots the overall analysis results of the probable changes in MTF values with different cumulative probabilities. The largest probable change in MTF with 99.9% cumulative probability is -0.23 . The MTF at 50 lp/mm is higher than 0.17. Meanwhile, the MTF with 84.1% cumulative probability is higher than 0.29.

Among the different tolerance errors, delta sag at a clear aperture (DLS) is more sensitive than the others, while the thickness of the surfaces is insensitive. The performances of F4, F6, and F23 have the greatest impact.

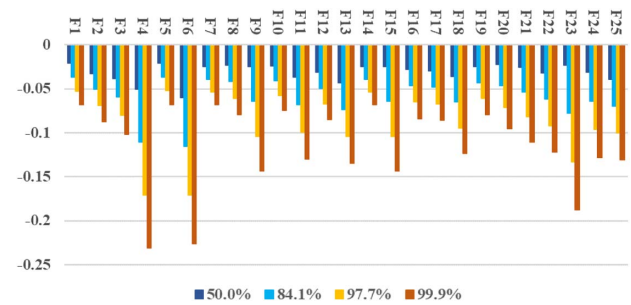


Fig. 11. Probable change of MTF value with four different cumulative probabilities for overall tolerance analysis using tolerances values listed in Table 4. F1–F25 denote the sampled fields the same as in Fig. 9.

Table 4. Tolerance Items

Tolerance Type	Location	Value	Unit
DLT–thickness delta	S_1, S_3	500	μm
DLT–thickness delta	S_2	40	μm
DLT–thickness delta	S_4 – S_9	20	μm
DLN–refractive index delta	E_1 – E_4	0.001	—
DLV–V-number delta	E_1 – E_4	0.005	—
DLX–surface X-displacement	S_3, S_5, S_9	5	μm
DLX–surface X-displacement	S_6, S_{10}	10	μm
DLX–surface X-displacement	S_1, S_2, S_4, S_7	25	μm
DLY–surface Y-displacement	S_3, S_5, S_9	5	μm
DLY–surface Y-displacement	S_6, S_{10}	10	μm
DLY–surface Y-displacement	S_1, S_2, S_4, S_7	25	μm
DLZ–surface Z-displacement	S_5, S_6, S_9	5	μm
DLZ–surface Z-displacement	S_2 – S_4, S_7	10	μm
DLZ–surface Z-displacement	S_1, S_{10}	20	μm
DLA–surface alpha tilt	S_2 – S_6	0.3	mrاد
DLA–surface alpha tilt	S_7	0.5	mrاد
DLA–surface alpha tilt	S_1, S_9, S_{10}	1	mrاد
DLB–surface beta tilt	S_2 – S_6	0.3	mrاد
DLB–surface beta tilt	S_7	0.5	mrاد
DLB–surface beta tilt	S_1, S_9, S_{10}	1	mrاد
DLG–surface gamma tilt	S_3, S_5	0.5	mrاد
DLG–surface gamma tilt	S_1, S_2, S_4, S_6 – S_{10}	5	mrاد
DLS–delta sag at clear aperture	S_2 – S_6	2	μm
DLS–delta sag at clear aperture	S_7	4	μm
DLS–delta sag at clear aperture	S_1, S_9, S_{10}	8	μm
DSR–surface roughness error	S_1 – S_{10}	5	μm

6. PROTOTYPE AND EXPERIMENTAL RESULTS

The system was fabricated using a molding approach. From Fig. 12(a), an exploded view of the optics module with all components is provided. The overall appearance is shown in Fig. 12(b). The main prism E_2 and auxiliary lens E_1 were connected by a mechanical structure while maintaining air separation. The main prism and correcting plate were glued with UV glue, and the lens holder fixed the microdisplay, lens-01 (E_3), and lens-02 (E_4). The total weight of the display module is 12.8 g, which meets the weight requirement of the OST-HMD.

To demonstrate the virtual and see-through light path image quality of the OST-HMD, simple tests are implemented. A commercial camera is used to simulate the human eyes and capture the displayed image transmitted through the OST-HMD system. The camera was placed in front of the optical module; the stop aperture and optical axis of the camera coincide with the exit pupil and viewing axis of the optical module.

An image resolution test chart shown in Fig. 13(a) was displayed at the microdisplay. Figure 13(b), a clear and nondistorted picture, was obtained, expressing the good virtual image performance of the optical module. Figure 13(c) shows the result of the fusion of a virtual cup and a real cup, demonstrating this module's potential applications in augmented reality.

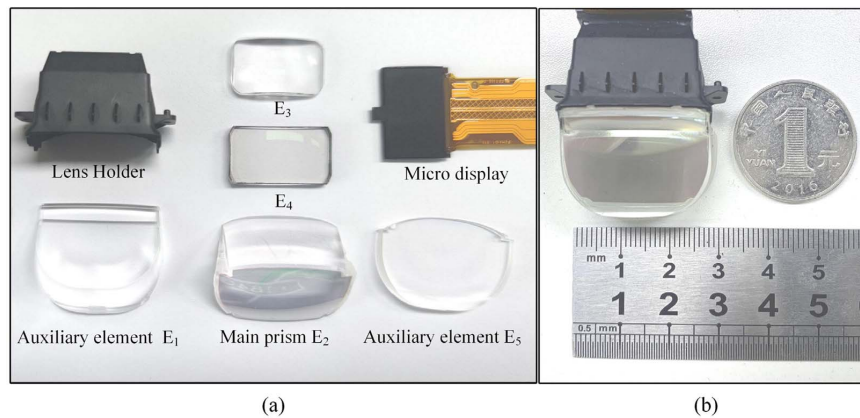


Fig. 12. Components and prototype of the optical system. (a) Exploded view showing all elements of the system. (b) Overall appearance of the prototype.

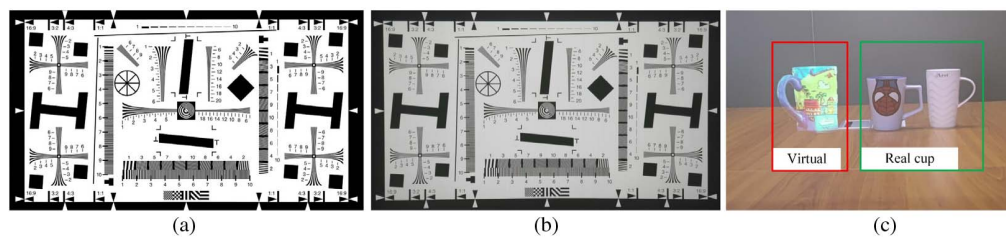


Fig. 13. Testing results of the optical prototype. (a) The input image displayed in microdisplay when testing the performance of the system. (b) Output image captured by camera at exit pupil of the system. (c) The result of fusion of virtual cup and real cup.

7. CONCLUSION

In this study, we designed an OST-HMD system using a combination of three wedge-shaped freeform prisms and two lenses. The system achieves a diagonal FOV of 45.3° and an $F/\#$ of 1.8, with an EPD of $12\text{ mm} \times 8\text{ mm}$, and an ERF of 18 mm. The overall size of the optical system is $30\text{ mm (width)} \times 40\text{ mm (height)} \times 14\text{ mm (thickness)}$, and the system has a weight of 12.8 g per eye, which meets the requirements of compact structure and light weight. Although the thickness of the optical element is up to 14 mm, the total thickness of the OST-HMD system is much less than that of the other solutions listed in Table 1 with similar FOV and EPD. The light efficiency of the virtual image light path can be up to 50%, which is much higher than that of other optical solutions, reducing the heat generation and making the system conducive to the wearing experience. Almost negligible distortion not only reduces the energy consumption due to electronic correction, but also guarantees virtual and real registration. Changing the radius of the first surface, S_1 , of the auxiliary lens, E_1 , allowed the optical system to meet the user's diopter requirements. The performance of the optical system was analyzed and was accompanied by a demonstration of an experimental prototype. The results indicate that the system is well suited to OST-HMD applications. We will develop a smaller FOV but much thinner OST-HMD system with freeform optics in a future study, and we believe the design will also greatly promote the AR industry as well.

Funding. National Key Research and Development Program of China (2017YFA0701200); National Natural Science Foundation of China (61822502); Beijing Municipal Science Technology Commission (Z201100004020011).

Acknowledgment. We thank Synopsys for providing the educational license of CODE V.

Disclosures. The authors declare no conflicts of interest.

REFERENCES

- O. Cakmakci and J. Rolland, "Head-worn displays: a review," *J. Display Technol.* **2**, 199–216 (2006).
- A. J. Lungu, W. Swinkels, L. Claesen, P. Tu, J. Egger, and X. Chen, "A review on the applications of virtual reality, augmented reality and mixed reality in surgical simulation: an extension to different kinds of surgery," *Expert Rev. Med. Devices* **18**, 47–62 (2021).
- V. Elia, M. G. Gnani, and A. Lanzilotto, "Evaluating the application of augmented reality devices in manufacturing from a process point of view: an AHP based model," *Expert Syst. Appl.* **63**, 187–197 (2016).
- H. Hua, L. D. Brown, and C. Gao, "Scape: supporting stereoscopic collaboration in augmented and projective environments," *IEEE Comput. Graph. Appl.* **24**, 66–75 (2004).
- H. Li, X. Zhang, G. Shi, H. Qu, Y. Wu, and J. Zhang, "Review and analysis of avionic helmet-mounted displays," *Opt. Eng.* **52**, 110901 (2013).
- L. Jensen and F. Konradson, "A review of the use of virtual reality head-mounted displays in education and training," *Educ. Inf. Technol.* **23**, 1515–1529 (2018).

7. J. Yang, W. Liu, W. Lv, D. Zhang, F. He, Z. Wei, and Y. Kang, "Method of achieving a wide field-of-view head-mounted display with small distortion," *Opt. Lett.* **38**, 2035–2037 (2013).
8. Z. Zheng, X. Liu, H. Li, and L. Xu, "Design and fabrication of an off-axis see-through head-mounted display with an x - y polynomial surface," *Appl. Opt.* **49**, 3661–3668 (2010).
9. L. Wei, Y. Li, J. Jing, L. Feng, and J. Zhou, "Design and fabrication of a compact off-axis see-through head-mounted display using a freeform surface," *Opt. Express* **26**, 8550–8565 (2018).
10. A. Wilson and H. Hua, "Design and demonstration of a vari-focal optical see-through head-mounted display using freeform Alvarez lenses," *Opt. Express* **27**, 15627–15637 (2019).
11. Q. Wang, D. Cheng, Q. Hou, Y. Hu, and Y. Wang, "Stray light and tolerance analysis of an ultrathin waveguide display," *Appl. Opt.* **54**, 8354–8362 (2015).
12. Q. Wang, D. Cheng, Q. Hou, L. Gu, and Y. Wang, "Design of an ultrathin, wide-angle, stray-light-free near-eye display with a dual-layer geometrical waveguide," *Opt. Express* **28**, 35376–35394 (2020).
13. D. Cheng, Y. Wang, C. Xu, W. Song, and G. Jin, "Design of an ultrathin near-eye display with geometrical waveguide and freeform optics," *Opt. Express* **22**, 20705–20719 (2014).
14. L. Gu, D. Cheng, Q. Wang, Q. Hou, and Y. Wang, "Design of a two-dimensional stray-light-free geometrical waveguide head-up display," *Appl. Opt.* **57**, 9246–9256 (2018).
15. I. Kasai, Y. Tanijiri, T. Endo, and H. Ueda, "A practical see-through head mounted display using a holographic optical element," *Opt. Rev.* **8**, 241–244 (2001).
16. C. Pan, Z. Liu, Y. Pang, X. Zheng, H. Cai, Y. Zhang, and Z. Huang, "Design of a high-performance in-coupling grating using differential evolution algorithm for waveguide display," *Opt. Express* **26**, 26646–26662 (2018).
17. Z. Liu, Y. Pang, C. Pan, and Z. Huang, "Design of a uniform-illumination binocular waveguide display with diffraction gratings and freeform optics," *Opt. Express* **25**, 30720–30731 (2017).
18. J. Xiao, J. Liu, J. Han, and Y. Wang, "Design of achromatic surface microstructure for near-eye display with diffractive waveguide," *Opt. Commun.* **452**, 411–416 (2019).
19. T. Levola, "Novel diffractive optical components for near to eye displays," *SID Symp. Dig.* **37**, 64–67 (2006).
20. J. Han, J. Liu, X. Yao, and Y. Wang, "Portable waveguide display system with a large field of view by integrating freeform elements and volume holograms," *Opt. Express* **23**, 3534–3549 (2015).
21. T. Yang, G. Jin, and J. Zhu, "Automated design of freeform imaging systems," *Light Sci. Appl.* **6**, e17081 (2017).
22. J. P. Rolland, M. A. Davies, T. J. Suleski, C. Evans, A. Bauer, J. C. Lambropoulos, and K. Falaggis, "Freeform optics for imaging," *Optica* **8**, 161–176 (2021).
23. B. Zhang, G. Jin, and J. Zhu, "Towards automatic freeform optics design: coarse and fine search of the three-mirror solution space," *Light Sci. Appl.* **10**, 65 (2021).
24. F. Duerr and H. Thienpont, "Freeform imaging systems: Fermat's principle unlocks "first time right" design," *Light Sci. Appl.* **10**, 95 (2021).
25. D. Cheng, H. Chen, T. Yang, J. Ke, Y. Li, and Y. Wang, "Optical design of a compact and high-transmittance compressive sensing imaging system enabled by freeform optics," *Chin. Opt. Lett.* **19**, 112202 (2021).
26. D. Cheng, Y. Wang, H. Hua, and M. M. Talha, "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**, 2655–2668 (2009).
27. Q. Wang, D. Cheng, Y. Wang, H. Hua, and G. Jin, "Design, tolerance, and fabrication of an optical see-through head-mounted display with free-form surface elements," *Appl. Opt.* **52**, C88–C99 (2013).
28. D. Cheng, Y. Wang, H. Hua, and J. Sasian, "Design of a wide-angle, lightweight head-mounted display using free-form optics tiling," *Opt. Lett.* **36**, 2098–2100 (2011).
29. C. Yao, D. Cheng, T. Yang, and Y. Wang, "Design of an optical see-through light-field near-eye display using a discrete lenslet array," *Opt. Express* **26**, 18292–18301 (2018).
30. C. Yao, D. Cheng, and Y. Wang, "Matrix optics representation and imaging analysis of a light-field near-eye display," *Opt. Express* **28**, 39976–39997 (2020).
31. L. Huang, J. Whitehead, S. Colburn, and A. Majumdar, "Design and analysis of extended depth of focus metalenses for achromatic computational imaging," *Photon. Res.* **8**, 1613–1623 (2020).
32. Y. Liu, Q. Yu, Z. Chen, H. Qiu, R. Chen, S. Jiang, X. He, F. Zhao, and J. Dong, "Meta-objective with sub-micrometer resolution for microendoscopes," *Photon. Res.* **9**, 106–115 (2021).
33. H. Li, X. Xiao, B. Fang, S. Gao, Z. Wang, C. Chen, Y. Zhao, S. Zhu, and T. Li, "Bandpass-filter-integrated multiwavelength achromatic metalens," *Photon. Res.* **9**, 1384–1390 (2021).
34. M. Deng, T. Ren, J. Wang, and L. Chen, "Doublet achromatic metalens for broadband optical retroreflector," *Chin. Opt. Lett.* **19**, 023601 (2021).
35. Z. Shen, S. Zhou, X. Li, S. Ge, P. Chen, W. Hu, and Y. Lu, "Liquid crystal integrated metalens with tunable chromatic aberration," *Adv. Photon.* **2**, 036002 (2020).
36. B. Xu, H. Li, S. Gao, X. Hua, C. Yang, C. Chen, F. Yan, S. Zhu, and T. Li, "Metalens-integrated compact imaging devices for wide-field microscopy," *Adv. Photon.* **2**, 066004 (2020).
37. G. Y. Lee, J. Y. Hong, S. Hwang, S. Moon, H. Kang, S. Jeon, H. Kim, J. H. Jeong, and B. Lee, "Metasurface eyepiece for augmented reality," *Nat. Commun.* **9**, 4562 (2018).
38. https://xinreality.com/wiki/R-9_Smartglasses.
39. <https://www.magicleap.com/en-us/magic-leap-1>.
40. <https://lumusvision.com/products/dk-52-2/>.
41. <https://www.olympus-global.com/en/news/2000b/nr000831fmd250e.html>.
42. <https://www.emagin.com/products/application-note>.
43. <https://www.rockwellcollins.com/>.
44. <http://www.nedplusar.com/en/index>.
45. S. J. Robbins, "Three piece prism eye-piece," U.S. patent 8,937,771 B2 (January 20, 2015).
46. D. Cheng and Q. Wang, "Free-form prism-lens group and near-eye display apparatus," U.S. patent 2019/0278087 A1 (September 12, 2019).
47. D. Cheng, H. Chen, Q. Wang, and Q. Hou, "Optical component for near-eye display," CN patent 213069354U (April 27, 2021).
48. *CODE V, Reference Manual* (Optical Research Associates, 2020).
49. D. Cheng, Y. Wang, and H. Hua, "Automatic image performance balancing in lens optimization," *Opt. Express* **18**, 11574–11588 (2010).