A retinal-projection-based near-eye display for virtual reality

Lantian Mi, Wenbo Zhang, Chao Ping Chen, Yuanchao Zhou, Yang Li, et al.

Lantian Mi, Wenbo Zhang, Chao Ping Chen, Yuanchao Zhou, Yang Li, Bing Yu, Nizamuddin Maitlo, "A retinal-projection-based near-eye display for virtual reality," Proc. SPIE 10676, Digital Optics for Immersive Displays, 106761C (21 May 2018); doi: 10.1117/12.2315672

SPIE.

Event: SPIE Photonics Europe, 2018, Strasbourg, France

A retinal-projection-based near-eye display for virtual reality

Lantian Mi, Wenbo Zhang, Chao Ping Chen*, Yuanchao Zhou, Yang Li, Bing Yu, and Nizamuddin Maitlo
Smart Display Lab, Department of Electronic Engineering, Shanghai Jiao Tong University, Shanghai, China

ABSTRACT

We propose a retinal-projection-based near-eye display for virtual reality. Our design is highlighted by an array of tiled organic light-emitting diodes and a transmissive spatial light modulator. Its design rules are to be set forth in depth, followed by the results and discussion regarding the field of view, angular pixel resolution, stereoscopic vision, modulation transfer function, distortion, contrast ratio, simulated imaging, and industrial design.

Keywords: near-eye display, head-mounted display, retinal projection, virtual reality, ultra-large field of view

1. INTRODUCTION

It has been five decades ever since the first near-eye display (NED)—dubbed as "The Sword of Damocles"—was invented [1]. Despite their success in the market of consumer electronics, today's near-eye displays for virtual reality (VR) are not totally satisfactory. In order to magnify the images of a microdisplay, lenses are widely adopted in NEDs for achieving a large field of view (FOV). However, the distortions are inevitably serious especially at the edge of viewing field. Moreover, the presence of lenses usually makes VR headsets bulky and heavy. To solve the said issues, we propose a retinal-projection-based NED for VR. Our design is highlighted by an array of tiled organic light-emitting diodes (OLEDs) and a transmissive spatial light modulator (SLM). As opposed to the conventional NEDs for VR, our NED is able to achieve a compact structure by getting rid of lenses. Its design rules are to be set forth in depth, followed by the results and discussion regarding the FOV, angular pixel resolution, stereoscopic vision, modulation transfer function (MTF), distortion, contrast ratio, simulated imaging, and industrial design.

2. DESIGN PRINCIPLE

2.1 Proposed structure

Fig. 1 is a schematic drawing of the proposed NED, which involves three major components, *i.e.* an array of OLEDs [2-4], a transmissive SLM, and an eye. OLEDs serve as the light source to illuminate the virtual image. Since the size of an individual OLED is way smaller than that of SLM, each OLED can be regarded as a point light source. The SLM is responsible for loading virtual images and modulating the intensity of light emitted from the OLEDs. OLEDs in tandem with the SLM constitute an array of miniature projectors that deliver the virtual image directly to the retina. d_s is the distance between the SLM and eye. d_o is the distance between the OLED array and eye. D is the center spacing between two adjacent OLEDs. L is the dimension of SLM.

*contact author: ccp@sjtu.edu.cn | website: sdl.sjtu.edu.cn

Digital Optics for Immersive Displays, edited by Bernard C. Kress, Wolfgang Osten, Hagen Stolle, Proc. of SPIE Vol. 10676, 106761C · © 2018 SPIE
CCC code: 0277-786X/18/$18 · doi: 10.1117/12.2315672

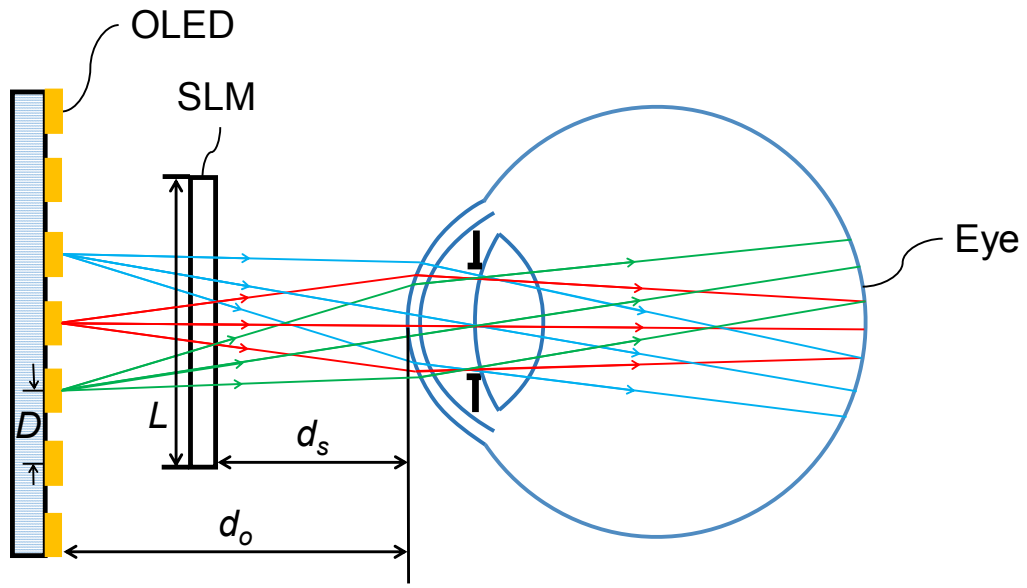

Fig. 1. Schematic drawing of the proposed NED. d_s is the distance between the SLM and eye. d_o is the distance between the array of OLEDs and eye. D is the center spacing between two adjacent OLEDs. L is the dimension of SLM.

2.2 Eye

Prior to explaining the design rules of our NED, it is essential to understand the mechanism of eye. Human eye is a zoom lens system, mainly consisting of cornea and lens as well as a light receptor—*i.e.* retina [5]. To accommodate different object distances, diopter of eye P_e could be tunable from 40.5 m^{-1} to 52.8 m^{-1} [6]. As an approximation, object distance s, image distance s' and diopter of eye P_e could be correlated via the thin-lens equation [7]

$$\frac{1}{s} + \frac{1}{s'} = P_e \tag{1}$$

where the image distance s' shall be fixed to the length of eye ball, which is about 24-25 mm for an adult. Say s'= 24.5 mm, object distance s can be calculated as a function of the diopter of eye P_e, as shown in Fig. 2. If the target value of object distance s is set as 3 m, P_e shall be 41.15 m^{-1}. For the maximum P_e= 52.8 m^{-1}, the near point—also known as the minimum object distance where sharp focusing is possible—is 8.35 cm.

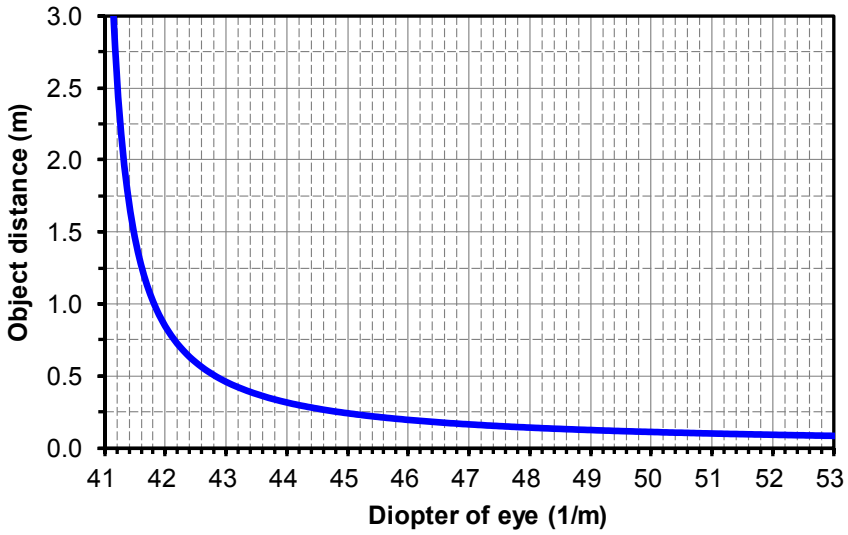

Fig. 2. Calculated object distance s as a function of the diopter of eye P_e. If the target value of object distance s is set as 3 m, P_e shall be 41.15 m^{-1}.

2.3 Design rules

The optical path diagram is depicted in Fig. 3. Let us first consider a scenario when merely a single OLED is turned on. Light rays emitting from the OLED will first encounter the SLM and illuminate its pixels, which in turn form a real object A that is composed of the actual rays (solid lines). In order for the real object A not to be directly perceived by the eye, it is required that SLM be placed out of the range of accommodation. Instead of seeing the real object A, eye will trace back along the extended virtual rays (dashed lines) to see a virtual object S formed at a distance s that is given by

$$s = \frac{s'}{P_e s' - 1} \tag{2}$$

which implies that the object distance s could be adjusted by the eye. This phenomenon would be very helpful for the image registration [8]—the alignment between the virtual and real objects—even without resorting to any zoom lens. By invoking the theorem of similar triangles, the size O of the virtual object can be easily deduced as

$$O = a_p \left(\frac{s'}{(P_e s' - 1) d_s} - 1 \right) \tag{3}$$

where a_p is the pupil size. Thus, it can be said that the size of virtual object largely hinges on the pupil size and diopter of eye. FOV of a single OLED, θ—defined as the angular extent of the virtual object—can be calculated as

$$\theta = 2 \, tan^{-1} \left(\frac{O}{2s} \right) \tag{4}$$

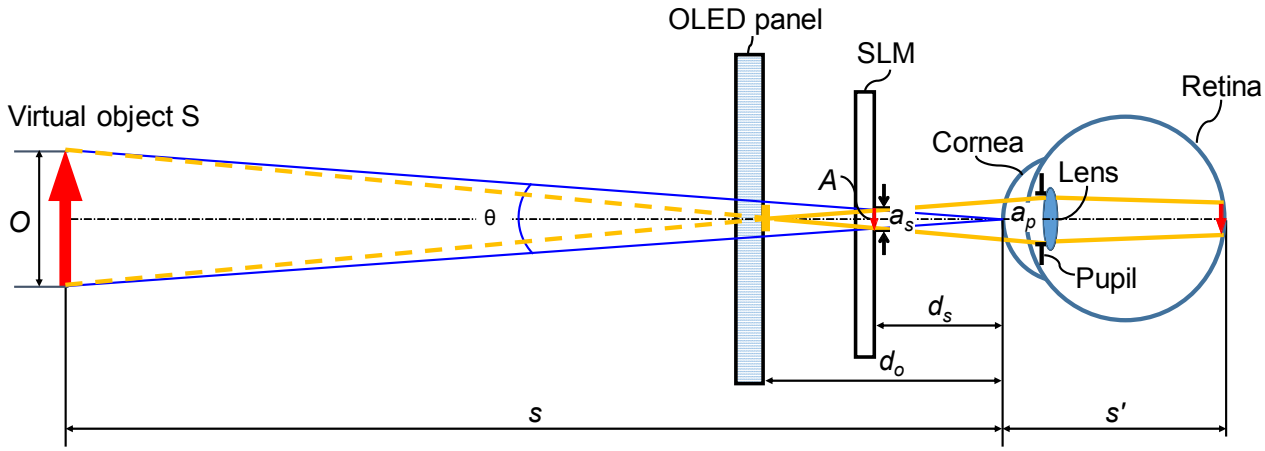

Fig. 3. Optical path diagram for imaging the virtual object, when merely a single OLED is turned on. Light rays emitting from the OLED will first encounter the SLM and illuminate its pixels, which in turn form a real object A that is composed of the actual rays (solid lines). In order for the real object A not to be directly perceived by the eye, it is required that SLM be placed out of the range of accommodation. Instead of seeing the real object A, eye will trace back along the extended virtual rays (dashed lines) to see a virtual object S formed at a distance s.

Given s= 3 m, P_e= 41.15 m^{-1}, s'= 24.5 mm, a_p= 4 mm, and d_o= 20 mm, θ is only 11.3°. For θ= 100°, d_s shall be decreased to 1.68 mm, which is apparently impractical. Hence, a tilted configuration with a multiple of OLEDs is required to realize large FOV.

Now consider a scenario when two adjacent OLEDs are simultaneously turned on, as illustrated in Fig. 4. There will be two virtual objects being imaged at the distance s. In order for these two virtual objects to be seamlessly tiled, the centers of two adjacent OLEDs should be spaced at an optimal distance D that is given by [9]

$$D = a_p \left(1 - \frac{d_s}{s} \right) \tag{5}$$

In addition, for the above condition to be strictly met, the distance d_s between the SLM and eye—should be adjusted as

$$d_s = \frac{a_p d_o}{a_p + D} \tag{6}$$

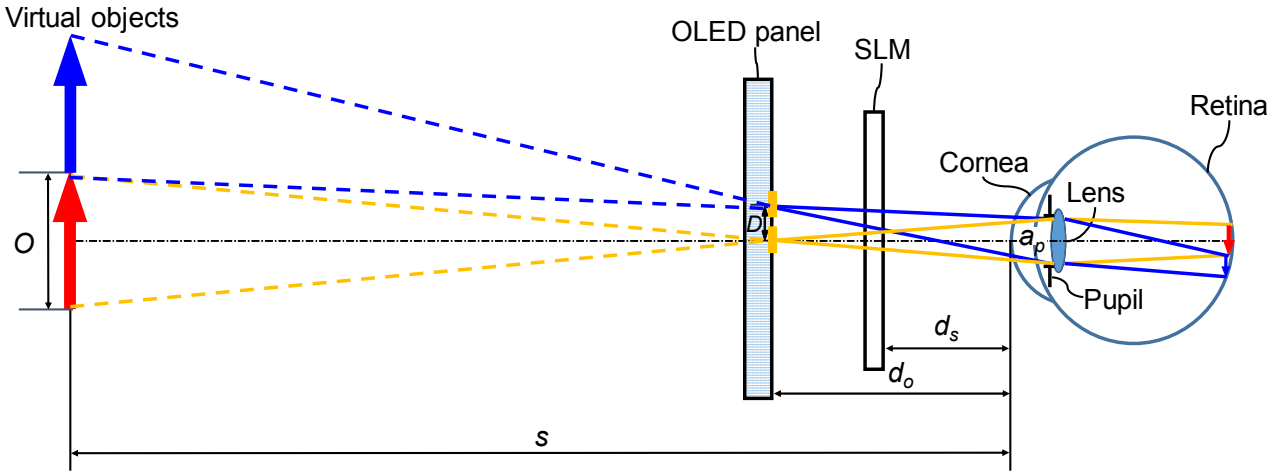

Fig. 4. Optical path diagram for imaging multiple virtual objects, when two adjacent OLEDs are simultaneously turned on. In order for these two virtual objects to be seamlessly tiled, the centers of two adjacent OLEDs should be spaced at an optimal distance *D*.

As can be inferred from Eq. (5), the optimal distance *D* will be subject to change as both pupil and diopter of eye may vary from time to time. When pupil expands or shrinks, shifting the pupil size away from the predetermined value, there will be an overlapping or gap between the neighboring virtual objects. To handle the change in pupil size, the array of tiled OLEDs shall be addressed in a way that it is able to dynamically tune the distance *D* to match with the current pupil size [9]. Since the distance d_o is usually way smaller than the object distance *s*, the change in diopter of eye, on the other hand, barely affects the distance *D* unless the object distance *s* becomes very close.

2.4 Spatial light modulator

As a viable option, a backlight-free, monochrome liquid crystal display (LCD) [10]—*e.g.* LCX036AMT (Sony)—can be used for this purpose. Its specifications are listed in Table 1, where the contrast ratio is 400 and the transmittance or transparency is 25%. For a better transparency, a higher aperture ratio is needed. If not to shrink the size of opaque thin-film transistors and electrodes, then the resolution has to be reduced, meaning that there is a tradeoff between the transparency and resolution.

Table 1. Specifications of SLM.

Object	Parameter	Value
SLM	*Resolution*	1604×1204
	L (diagonal)	1.8 inch (45.72 mm)
	L (horizontal)	36.6 mm
	L (vertical)	27.4 mm
	Pixel size	23 μm
	Transmittance	25%
	CR	400

2.5 Organic light-emitting diodes

OLEDs in tandem with a SLM constitute an array of miniature projectors that deliver the virtual image directly to the retina. Basically, each OLED can be regarded as an individual point light source. Alternatively, OLEDs can be replaced by the quantum dot light-emitting diodes [11] or other types of thin-film lighting technologies. For the tiled configuration of NED, the total number of OLEDs shall be determined by

$$\#\ of\ OLEDs = \frac{L(horizontal)\times L(vertical)}{a_s} \tag{7}$$

According to Eqs. (5) and (6), for s= 3 m, a_p= 4 mm, and d_o= 20 mm, D= 3.97 mm and d_s= 10.03 mm. Thus, a_s= 1.99 mm, for which the total number of OLEDs required for fully illuminating the SLM LC 2012 (Holoeye) shall be 19 (horizontal) by 14 (vertical). Further, in order to realize full-color NED, the above number will be tripled in one dimension by incorporating R/G/B OLEDs. As each OLED will illuminate more than one pixel of SLM, a sequential color scheme [12], in which each color of OLEDs is activated in sequence, is indispensable.

3. RESULTS AND DISCUSSION

3.1 Simulation settings

Our simulation is implemented with the optical design software Code V (Synopsys), which is based on the ray tracing [13] and capable of analyzing the imaging properties, including MTF, distortion, and imaging simulation. Our design wavelength is 550 nm. Although a configuration with tiled OLEDs is proposed, our simulation is limited to the case of a single OLED. This is because it is technically impossible to set up more than one object in Code V. As the sharpest or best vision occurs at the center of retina [14], where the fovea is located, a configuration with a single OLED being center-aligned with retina is adopted for the simulation. To avoid duplicate simulations, other cases are omitted.

The numbering of surfaces is labelled as in Fig. 5. The object represents the virtual object that is 3 m away from the eye. Surface 1 (S1) is a stop, which is used to mimic the point light source. Surfaces 2 to 6 (S2 to S6) make up the eye, of which, S2 is anterior cornea, S3 is posterior cornea, S4 is anterior lens with pupil, S5 is posterior lens, and S6 is retina. To treat the OLED, which is situated on S1, as a point light source, the aperture of S1 is decreased to 0.1 mm.

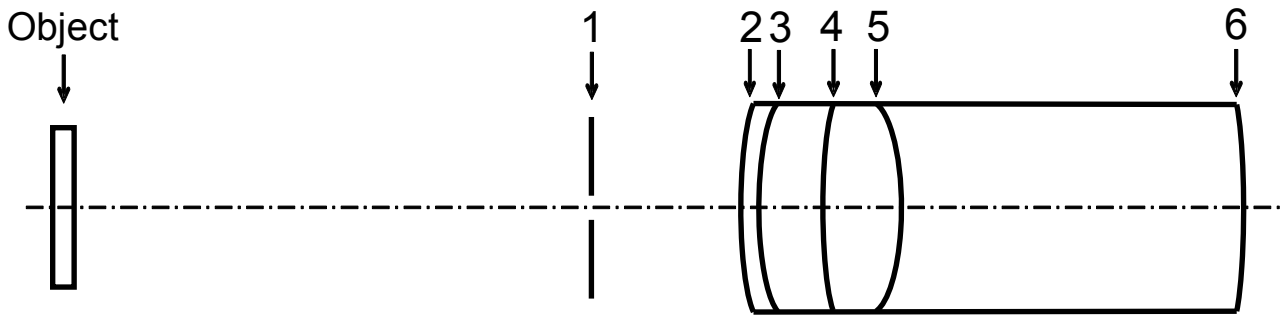

Fig. 5. The numbering of surfaces. The object represents the virtual object that is 3 m away from the eye. Surface 1 (S1) is the stop. Surfaces 2 to 6 (S2 to S6) make up the eye.

According to the said design rules, we could create an initial structure by presetting the parameters for each element. Then, an optimization, whose error function type is set as transverse ray aberration [13], is carried out by constraining the effective focal length of the eye to be 24.5 mm—*i.e.* length of eye ball. The parameters obtained after the optimization are summarized in Table 2. Besides, more detailed parameters for defining aspherical surfaces are disclosed in Table 3.

Table 2. Parameters used for the simulation.

Surface	Surface type	Radius (mm)	Thickness (mm)	Refractive index	Semi-aperture (mm)
object	sphere	infinity	2980		
1	sphere	infinity	20		0.1000
2	asphere	6.6624	0.5000	1.3760	2.0159
3	asphere	3.5491	3.1600	1.3360	2.0012
4	asphere	11.4085	3.6000	1.4085	2.0000
5	asphere	−5.5801	17.2000	1.3360	1.9968
6	sphere	−11.0000	0.0000	1.3360	1.6078

Table 3. Parameters for aspherical surfaces.

Surface	Y radius (mm)	Conic constant (K)	4th order coefficient (A)	6th order coefficient (B)	8th order coefficient (C)
2	6.6624	−0.1800	0.0002	0	0
3	3.5491	−0.6000	0.0040	0	0
4	11.4085	−0.9427	0.0005	−0.0002	−2.9563E-06
5	−5.5801	2.5161	0.0051	−0.0002	3.2283E-05

3.2 FOV & angular pixel resolution

FOV and angular pixel resolution are two key indicators for evaluating the performance of NED. FOV, defined as the angle subtended by the virtual image as viewed by the observer [15], can be expressed as the angular extent of the SLM—if SLM is fully illuminated. As a simple means of comparing the display quality to visual limit, angular pixel resolution—measured in arcminute (′)—is calculated by dividing the FOV by the number of pixels [15]. As shown in Fig. 6, FOV and angular pixel resolution could be calculated as

$$FOV = 2\,tan^{-1}\left(\frac{L}{2d_s}\right) \tag{8}$$

and

$$Angular\ pixel\ resolution = \frac{60 \cdot FOV}{\sqrt{N_h^2+N_v^2}} \tag{9}$$

where N_h and N_v are the number of pixels along the horizontal and vertical directions, respectively. For L= 1.8 inch, d_s= 10.03 mm, N_h= 1604, and N_v= 1204, FOV is 133° and angular pixel resolution is 3.9′.

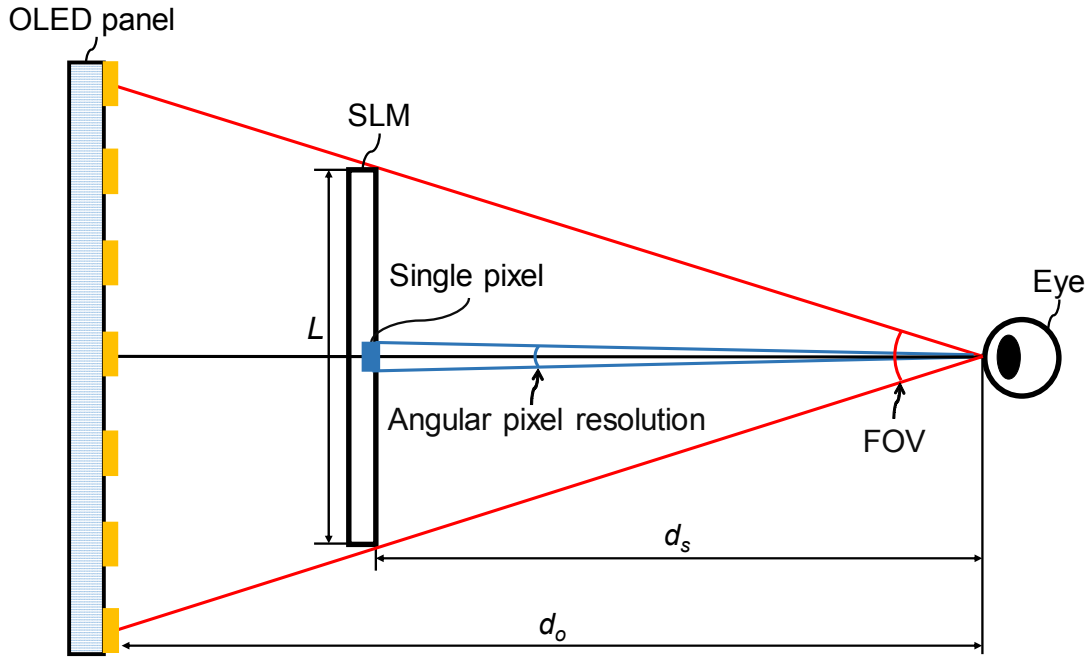

Fig. 6. Illustration of FOV and angular pixel resolution along the diagonal direction. FOV is defined as the angular extent of the SLM, whereas angular pixel resolution is defined as the angular extent of single pixel.

3.3 Stereoscopic vision

As a measure of stereoscopic vision, stereo overlap is defined as the ratio of the size of image shared by both eyes, $W_{overlap}$, to the size of image visible to one eye, W. A small stereo overlap may result in a binocular rivalry—the conflict between two images seen by the different eyes. As shown in Fig. 7, stereo overlap can be estimated by

$$Stereo\ overlap = \frac{W_{overlap}}{W} = 1 - \frac{d_p}{W} \tag{10}$$

where d_p stands for the interpupillary distance, which is by average 65 mm for an adult. Given s= 3 m, d_s= 10.03 mm, the stereo overlap is approximately 99% for our design.

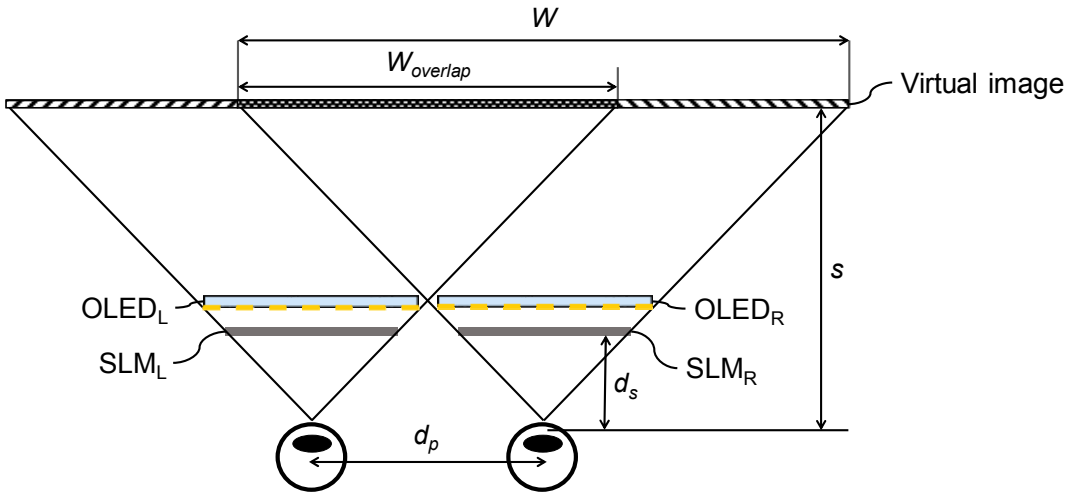

Fig. 7. Illustration of the stereo overlap for stereoscopic vision. $W_{overlap}$ stands for the size of image shared by both eyes, W the size of the virtual image observed by one eye, and d_p the interpupillary distance.

3.4 MTF & distortion

MTF and distortion are two important performance criterions for NEDs. By computing the auto-correlation of the pupil function [16], MTFs of are calculated, as shown in Fig. 8. Since the tiniest aperture in our NED—*i.e.* the pixel of SLM—is 23 μm across, which is almost two orders of magnitude larger than the wavelength, the diffraction is negligible. Moreover, owing to the rotational symmetry, the marginal angle is set as 5.65°—half of FOV of single OLED. Distortion is defined as the displacement of image height or ray location, as plotted in Fig. 9. As can be seen in Figs. (8) and (9), MTFs for all field angles are above 0.4 at 120 cycles/mm, and the distortion is below 0.73%.

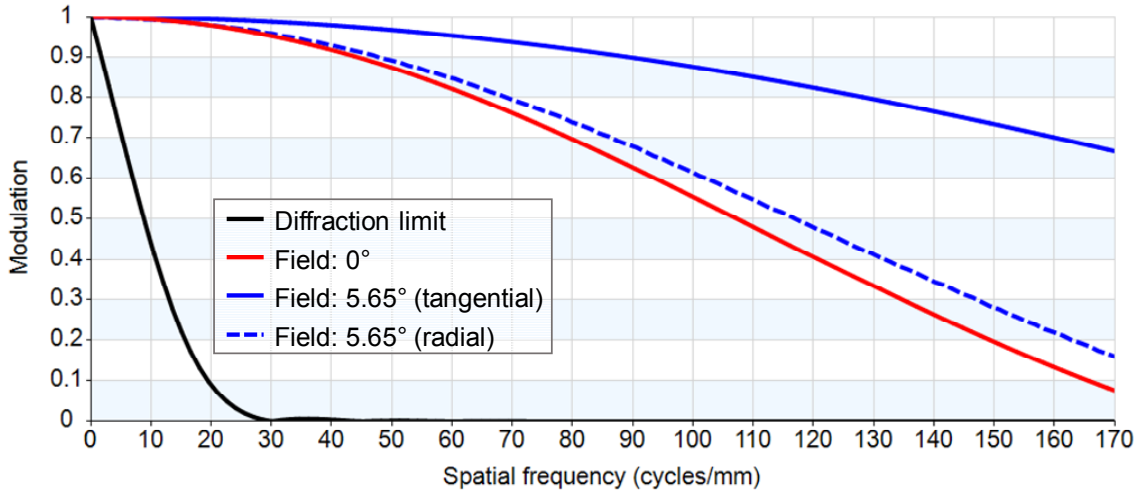

Fig. 8. Calculated MTF versus the spatial frequency. MTFs for all field angles are above 0.4 at 120 cycles/mm.

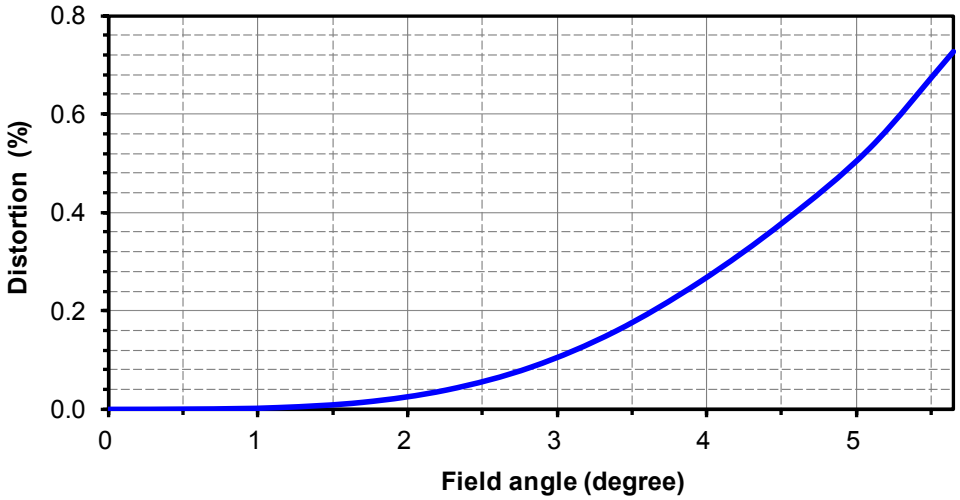

Fig. 9. Calculated distortion versus the field angle. The distortion is below 0.73%.

3.5 Contrast ratio

Contrast ratio (CR) is defined as the ratio of maximum intensity to minimum intensity, and it can be derived as [6]

$$CR = \frac{1+M\cdot MTF}{1-M\cdot MTF} \tag{11}$$

where M denotes the modulation in object, *i.e.*

$$M = \frac{CR_o-1}{CR_o+1} \tag{12}$$

where CR_o is the CR of object. In our design, CR_o is the CR of SLM. From Eqs. (11) and (12), for the spatial frequency of 21.74 cycles/mm—which corresponds to a pixel size of 23 μm—at the central angle, CR is calculated as 66.

3.6 Simulated imaging

For a qualitative analysis of imaging quality, virtual images are visualized from the imaging simulation that takes into account the effects of distortion, aberration blurring, diffraction blurring, and relative illumination, as shown in Fig. 10. By comparing the original image and simulated virtual image, it can be seen that the simulated virtual image turns out to be blurred. It has to be mentioned that the simulated virtual image is what exactly appear on the retina.

(a) (b)

Fig. 10. (a) Original image (photographer: C. P. Chen, location: Bolshoi Theater, Moscow, Russia), and (b) simulated image. By comparing the original image and simulated virtual image, it can be seen that the simulated virtual image turns out to be blurred.

3.7 Industrial design

Fig. 11 shows the proposed industrial design for our NED, which is convertible between a VR headset and a headphone. To reveal the interior, the opaque cover/hood, which is used to block the ambient light, is omitted. The temples, on which both OLEDs and SLMs are mounted, are foldable. Plus, the forearms of temples are movable so that the distances of both OLEDs and SLMs could be manually adjusted. When not in use, both OLEDs and SLMs could be packed into the headphone, thereby making it more portable. As our NED is based on the retinal projection, eye trackers to detect the gaze direction and pupil size are employed as well.

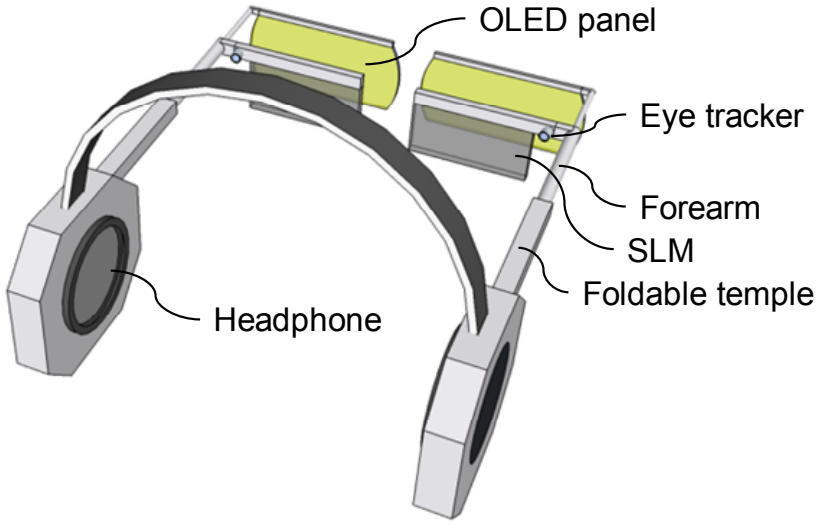

Fig. 11. Proposed industrial design for our NED. It is convertible between a VR headset and a headphone. To reveal the interior, the opaque cover/hood, which is used to block the ambient light, is omitted. The temples, on which both OLEDs and SLMs are mounted, are foldable. Plus, the forearms of temples are movable so that the distances of both OLEDs and SLMs could be manually adjusted.

4. CONCLUSIONS

A retinal-projection-based NED for VR is proposed. Its structure is highlighted by an array of tiled OLEDs and a transmissive SLM. Based on the simulation, its overall performance has been comprehensively investigated. The FOV is 133° (diagonal). The angular pixel resolution is 3.9′. The stereo overlap is 99%. The MTF is above 0.4 at 120 cycles/mm. The CR is 66. The distortion is below 0.73%. Compared with other types of NEDs [17-20], our NED exhibits several unique features and advantages. First, instead of using an eyepiece or ocular lens, our NED leans on the eye itself in imaging the virtual objects. Second, compact form factor and lightweight are expected as no lenses are present in our NED. Third, the distance and size of virtual objects hinge on the status of eye, including its diopter and pupil. Fourth, its ultra-large FOV is unparalleled.

ACKNOWLEDGMENTS

This work is supported by Science and Technology Commission of Shanghai Municipality (1701H169200), Shanghai Jiao Tong University (AF0300204, WF101103001/085), and Shanghai Rockers Inc. (15H100000157).

REFERENCES

[1] Wikipedia, “Virtual reality,” https://en.wikipedia.org/wiki/virtual_reality.
[2] S. Huang, Z. Ye, J. Lu, Y. Su, C. Chen, and G. He, “Position dependence of extraction efficiency in organic light-emitting diodes with photonic crystal structure,” Chin. Opt. Lett. **11**(6), 062302 (2013).

[3] C. Chen, H. Li, Y. Zhang, C. Moon, W. Y. Kim, and C. G. Jhun, "Thin-film encapsulation for top-emitting organic light-emitting diode with inverted structure," Chin. Opt. Lett. **12**(2), 022301 (2014).
[4] X. Wei, H.-J. Yu, J. Cao, C. Jhun, and C. P. Chen, "Efficient blue and white phosphorescent organic light-emitting diodes with a mixed host in emission layer," Mol. Cryst. Liq. Cryst. **651**(1), 108–117 (2017).
[5] Wikipedia, "Human eye," https://en.wikipedia.org/wiki/human_eye.
[6] C. P. Chen, L. Zhou, J. Ge, Y. Wu, L. Mi, Y. Wu, B. Yu, and Y. Li, "Design of retinal projection displays enabling vision correction," Opt. Express **25**(23), 28223–28235 (2017).
[7] F. L. Pedrotti, L. M. Pedrotti, and L. S. Pedrotti, *Introduction to Optics 3rd Edition* (Addison-Wesley, 2006).
[8] Y.-J. Wang, P.-J. Chen, X. Liang, and Y.-H. Lin, "Augmented reality with image registration, vision correction and sunlight readability via liquid crystal devices," Sci. Rep. **7**, 433 (2017).
[9] A. Maimone, D. Lanman, K. Rathinavel, K. Keller, D. Luebke, and H. Fuchs, "Pinlight displays: wide field of view augmented reality eyeglasses using defocused point light sources," ACM Trans. Graph. **33**(4), 89 (2014).
[10] C. P. Chen, S. P. Preman, T.-H. Yoon, and J. C. Kim, "Dual-mode operation of dual-frequency liquid crystal cell by horizontal switching," Appl. Phys. Lett. **92**, 123505 (2008).
[11] J. Cao, J.-W. Xie, X. Wei, J. Zhou, C.-P. Chen, Z.-X. Wang, and C. Jhun, "Bright hybrid white light-emitting quantum dot device with direct charge injection into quantum dot," Chin. Phys. B **25**(12), 128502 (2016).
[12] C. P. Chen, Y. Li, Y. Su, G. He, J. Lu, and L. Qian, "Transmissive interferometric display with single-layer Fabry-Pérot filter," J. Disp. Technol. **11**(9), 715–719 (2015).
[13] R. E. Fischer, B. Tadic-Galeb, and P. R. Yoder, *Optical System Design 2nd Edition* (McGraw-Hill Education, 2008).
[14] Wikipedia, "Visual acuity," https://en.wikipedia.org/wiki/Visual_acuity.
[15] J. E. Melzer and K. Moffitt, *Head Mounted Displays: Designing for the User* (McGraw-Hill Professional, 1996).
[16] H. H. Hopkins, "The numerical evaluation of the frequency response of optical systems," Proc. Phys. Soc. B **70**(10), 1002–1005 (1957).
[17] L. Zhou, C. P. Chen, Y. Wu, K. Wang, and Z. Zhang, "See-through near-eye displays for visual impairment," in *The 23rd International Display Workshops in conjunction with Asia Display* (2016), pp. 1114–1115.
[18] L. Zhou, C. P. Chen, Y. Wu, Z. Zhang, K. Wang, B. Yu, and Y. Li, "See-through near-eye displays enabling vision correction," Opt. Express **25**(3), 2130–2142 (2017).
[19] Y. Wu, C. P. Chen, L. Zhou, Y. Li, B. Yu, and H. Jin, "Design of see-through near-eye display for presbyopia," Opt. Express **25**(8), 8937–8949 (2017).
[20] Y. Wu, C. P. Chen, L. Zhou, Y. Li, B. Yu, and H. Jin, "Near-eye display for vision correction with large FOV," in *SID Display Week* (2017), pp. 767–770.