Lightweight spatial-multiplexed dual focal-plane head-mounted display using two freeform prisms

Dewen Cheng (程德文)[1], Qingfeng Wang (王庆丰)[1], Yongtian Wang (王涌天)[1*], and Guofan Jin (金国藩)[2]

[1]*School of Optoelectronics, Beijing Institute of Technology, Beijing 100081, China*

[2]*State Key Laboratory of Precision Measurement Technology and Instruments, Tsinghua University, Beijing 100084, China*

[*]*Corresponding author: wyt@bit.edu.cn*

Received August 27, 2012; accepted October 10, 2012; posted online February 6, 2013

Accommodation and convergence play critical roles in the natural viewing of three-dimensional (3D) scenes, and these must be accurately matched to avoid visual fatigue. However, conventional stereoscopic head-mounted displays lack the ability to adjust accommodation cues. This is because they only have a single, fixed image plane, but the 3D virtual objects generated by a pair of stereoscopic images are displayed at different depths, either in front or behind the focal plane. Therefore, in order to view objects clearly, the eyes are forced to converge on those objects while maintaining accommodation fixed on the image plane. By employing freeform optical surfaces, we design a lightweight and wearable spatial-multiplexed dual focal-plane head-mounted display. This display can adjust the accommodation cue in accordance with the convergence cue as well as generate the retinal blur cue. The system has great potential applications in both scientific research and commercial market.

OCIS codes: 120.2820, 220.1250, 220.3620, 330.1400, 330.7322.
doi: 10.3788/COL201311.031201.

The accommodation and convergence discrepancy (ACD) problem in stereoscopic display is a well-known scientific problem in the three-dimensional (3D) display research community[1−6]. This problem has received increasing attention in recent years[7−18]. The accommodation distance of human eyes changes in accordance with the convergence distance when the eyes gaze at natural 3D objects. Specifically, the angle between the lines of sight of the two eyes and the relaxation of the muscles are compatible. However, the harmony between accommodation and convergence is disrupted when the user observes a 3D object through a stereoscopic display. In this case, the eyes need to focus on the virtual screen of the display to view the image clearly; however, the virtual 3D objects are usually located in front or behind the virtual screen. When the accommodation and convergence distances differ, accommodation cannot be varied to match convergence. The ACD problem becomes more serious as the disparity range of the virtual 3D scene increases.

The ACD problem is determined by the essential characteristics of conventional stereoscopic display technologies according to many previous works[7−18]. Such problem is a major cause of unnatural 3D perception and visual fatigue. Fortunately, the accommodation distance (D_{v}) can be greatly changed with minor variations of the effective focal length (f') of the eyepiece or the distance l' between the eyepiece and the display device according to Eq. (1), which is given as

$$D_{\mathrm{v}} = \frac{(f' - l')}{(f' - l') \times l_{\mathrm{erf}} + f' \times l'}, \quad (1)$$

where l_{erf} is the eye relief. The items in Eq. (1) are shown in Fig. 1. Here, l' is less than f' but their difference is very small, such that D_{v} can be greatly changed with subtle variations of f' or l'.

The key issue is generating multiple continuous or discrete focal planes in a stereoscopic display; furthermore, each focal plane must correspond to a virtual screen with a specified distance. Several methods have been proposed to generate dual or multiple focal planes for stereoscopic displays. These methods can be classified into spatial-multiplexed[7−12] and time-multiplexed methods[13−21].

Spatial-multiplexed methods include the stack display and splitting method. Rolland *et al.*[7] presented a comprehensive investigation to determine the minimum number of focal planes required for a spatial-multiplexed head-mounted display (HMD). They concluded that an image source consisting of a stack of at least 14 planar displays with a total thickness of 22.5 mm is required to accommodate from 0 to 2 D. Akeley *et al.* developed a three focal-plane prototype by dividing a liquid crystal display (LCD) panel into three zones located at different physical distances with three half mirrors[8]. They studied a depth-filtering rendering technique to determine how the image intensity at each pixel should be assigned to the focal planes, thus reducing the number

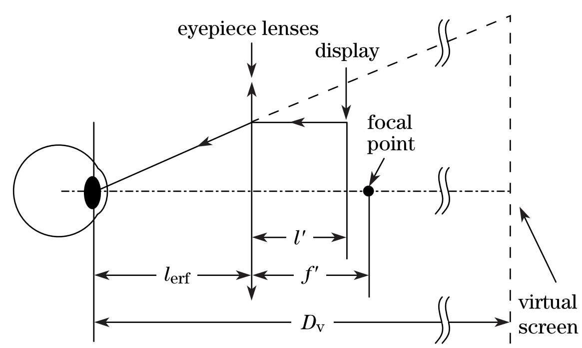

Fig. 1. Schematic diagram showing the principle of a near-eye display.

of focal planes[8,12]. MacKenzie *et al.* then carried out experiments with a similar prototype as that of Akeley *et al.*[9]. They concluded that a dioptric separation from 0.6 to 0.9 D between two adjacent focal planes is acceptable. The conclusion is identical to those obtained by the systematic and theoretical analysis of the depth-filtering multiple focal-plane display in Ref. [13]. MacKenzie *et al.* presented a detailed analysis of retinal-image formation as a function of image-plane separation, and measured accommodation responses to these depth-filtered stimuli[10]. Ravikumar *et al.* analyzed the effectiveness of different depth-filtering rules, and concluded that the linear rule is overall the best depth-weighted blending rule for multi-plane displays[11]. In spatial-multiplexed 3D displays, the focal planes physically exist simultaneously. Two-dimensional (2D) objects with different depths are rendered on the 2D focal planes at the same time, but with depth-fused filtering.

In time-multiplexed 3D displays, 2D objects with different depths are rendered sequentially on a single 2D focal plane, with a focal distance adjusted in synchronization with the depth of the object being rendered. Time-multiplexing methods include shifting the relay lens, moving the microdisplay, and using deformable mirrors, liquid lenses or birefringence lenses, among others. Shiwa *et al.* presented a mechanical way of changing the axial position of the relay lens in a binocular 3D displays[14]. Shibata *et al.* proposed to change the axial position of the microdisplay directly to generate several virtual screens[15]. Suyama *et al.* demonstrated a time-multiplexed approach using dual frequency liquid crystal varifocal lens[16] and liquid lens[17,18]. In the latter prototype, the accommodation distance can be varied from 8 D to infinity. Love *et al.* presented a four focal-plane display with a depth range of 1.8 D (from 5.09 to 6.89 D) and a refresh rate of 45 Hz using two fast switchable lenses synchronized with a cathode ray tube (CRT) display[19]. Each switchable lens consists of a ferroelectric liquid crystal (FLC) polarization filter and a birefringent lens, with two refractive indices while the incoming light has changing polarization status[16]. Mcquaide *et al.* demonstrated a dual focal-plane retinal scanning display, which uses a deformable membrane mirror (DMM) device[20]. The focus cues of each pixel are rendered by defocusing a laser beam through the DMM. Schowengerdt *et al.* implemented a volumetric display using a scanned fiber array[21].

Most of the abovementioned solutions lead to heavy and bulky systems. Moreover, some of the systems require long focal depth, which increases the complexity of the optical system and makes it difficult to reduce their size and weight. Other disadvantages are as follows: the spatial-multiplexed method uses a stack of microdisplays or a large LCD, the time-multiplexed method employs active elements, and the electronic control units are usually bulky. Another drawback is that a visible flicker may be observed in time-multiplexed prototypes if the refresh rate is not high enough, which is limited by the component with the lowest refresh rate (e.g., the display device, the graphic card, or the active optical elements, including liquid lenses, DMMs, FLCs, and so on). Thus, although the abovementioned approaches are solutions to a true 3D display, they are not suitable for wearable stereoscopic HMDs.

To circumvent the ACD problem and dramatically reduce the system volume and weight, we present a novel dual focal-plane HMD solution based on the spatial time-multiplexed method that uses freeform optics. Aside from its compact size and portability, the solution does not contain any active or mechanically moving optical elements; thus, the refresh rate is not a problem. In addition, our proposed solution can also generate the retinal blur cue because it has two real focal planes.

The proposal in this letter is closely related to the optics tiling HMD design we presented in a previous work[22]. Both of our proposed approaches have two display channels, each consisting of a freeform surface prism and a microdisplay; however, they have distinctly different functions and applications. In Ref. [22], two identical prisms are tiled at their bottom surfaces, which are non-optical surfaces. The focal distances of the two channels are equal, and the field of view (FOV) of each channel is separate. Thus, the FOV of the tiled HMD in the direction perpendicular to the tiling surface is expanded. Rather than forming a large FOV by tiling the fields, the FOVs of two display channels in this letter are completely overlapped, and the virtual screens have a depth difference of 0.6 D according to previous works[9,10,13]. In the current work, we propose four design schemes by considering the structure, design complexity, light efficiency, and stray light issues of the dual focal-plane optical system, as shown in Fig. 1. Their advantages and disadvantages are also analyzed and described in detail.

The first scheme is shown in Fig. 2(a). The freeform optical surfaces 1 and 2 (which are the cemented surfaces) are used thrice in the two display channels. This may cause stagnation in optimization due to the varied requirements in different light paths, because the optical systems of the two display channels have to be optimized

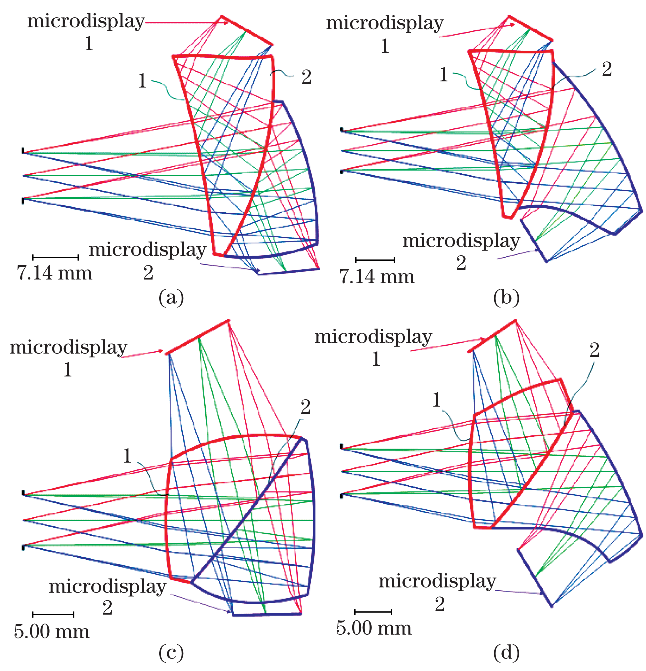

Fig. 2. 2D layouts of four configurations dual focal-plane HMDs.

simultaneously to balance their performance. In addition, the illumination of the second display channel is half that of the first display channel because the light from microdisplay 2 passes one more time through the half mirror than that from microdisplay 1. Moreover, strong stray light may be generated because the light from the two microdisplays shines directly onto each other, which is reflected back.

Similar to the first scheme, the second one shown in Fig. 2(b) consists of a wedge-shaped prism and a triangular prism. The light from the second display channel passes only once through the half mirror. When this occurs, the influence of the half mirror on system image quality is reduced, ensuring that the two display channels have the same light efficiency. However, surface 1 is still used thrice in the two light channels, and the design difficulty remains high. In the third and fourth schemes, surface 1 functions only as a refractive surface, which is shared by the two display channels. However, the third scheme has some serious stray light problems because the two microdisplays are facing each other (Fig. 2(c)), and the two display channels have different light efficiencies. The fourth scheme (Fig. 2(d)), reduces the design difficulty; it also solves the light efficiency and stray light issues. Thus, the design in this letter is based on the fourth scheme.

The dual focal-plane HMD in this letter has a FOV of 40°, an eye relief of 20 mm, and an exit pupil diameter of 6 mm. The depth of the display image is determined within 0.2 and 0.8 D according to Refs. [9,10], that is, the viewing distance ranges from 1.25 to 5 m.

Figure 3(a) shows the 2D layout of the optimized design based on the fourth scheme. As can be seen, microdisplay 1 corresponds to a virtual screen with a depth of 0.8 D, whereas microdisplay 2 generates the image on a screen with a depth of 0.2 D. Figures 3(b) and (d) show the distortion grid maps of the first and second display channel, respectively. Figures 3(c) and (e) show the modular transfer function (MTF) curves of the first and second display channels, respectively. The solid lines show the MTF in the tangential plane, while the dashed lines represent the MTF in the sagittal direction. These are evaluated on a full pupil and most of them are above 0.2 at 331 lps/mm. The design has a 10% keystone distortion at the corners for each display channel, which is canceled out by prewarpping the stereoscopic images electronically. The final 3D scene thus becomes sharp and undistorted.

This letter aims to resolve the accommodation and convergence mismatch problem in stereoscopic displays, especially for HMD systems. A novel lightweight spatial-multiplexed dual focal-plane HMD adopting freeform optics is presented and designed based on the thorough analysis and discussion of the existing solutions. Four configurations are compared, and the final configuration is selected according to the design complexity, light efficiency, and stray light issues. The optical design includes two light paths with fully overlapping FOVs, but with two focal depths. The dual focal-plane HMD system can generate a 40° FOV 3D scene with a depth range from 0.2 to 0.8 D; thus, it is capable of alleviating eye fatigue. The current solution can only generate two different focal planes, and as such, it is not a general

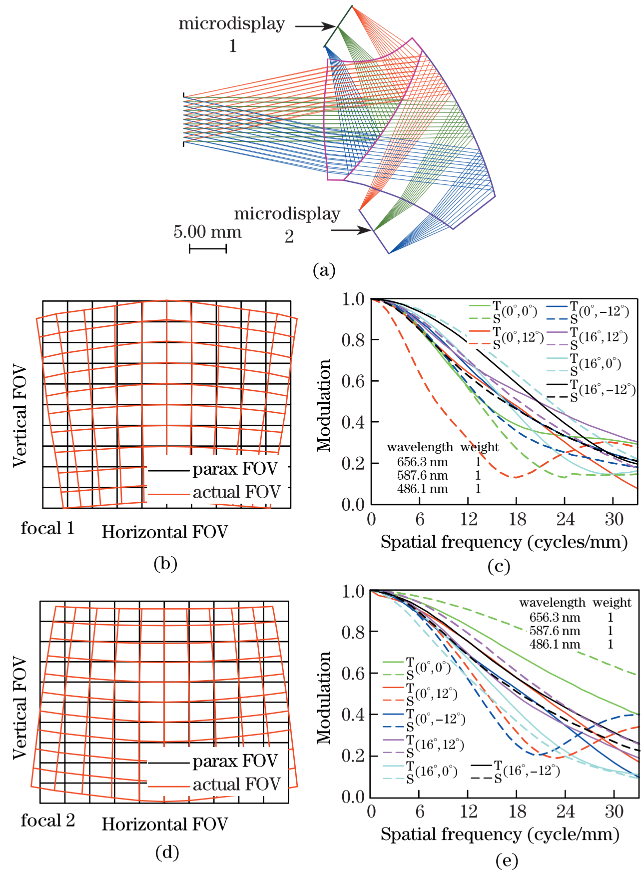

Fig. 3. Optimization results of the dual focal-plane freeform HMD system. (a) 2D layout; (b) distortion plot of the first display channel; (c) MTF curves of the first display channel; (d) distortion plot of the second display channel; (e) MTF curves of the second display channel.

solution to the accommodation and convergence mismatch problem. However, the system is small-sized, lightweight and can be easily mounted on the head of a user.

This work was supported by the National Basic Research Program of China (No. 2011CB706701) and the National Natural Science Foundation of China (Nos. 61178038 and 61205024). We would like to thank Synopsys for providing the education license of CODEV.

References

1. F. L. Kooi and A. Toet, Displays **25,** 99 (2004).
2. S. J. Watt, K. Akeley, M. O. Ernst, and M. S. Banks, J. Vis. **5,** 834 (2005).
3. D. M. Hoffman, A. R. Girshick, K. Akeley, and M. S. Banks, J. Vis. **8,** 1 (2008).
4. G. Nehmetallah, P. P. Banerjee, D. Ferree, R. Kephart, and S. Praharaj, Chin. Opt. Lett. **9,** 120004 (2011).
5. T. Yamaguchi and H. Yoshikawa, Chin. Opt. Lett. **9**, 120006 (2011).
6. M. Lambooij, W. IJsselsteijn, M. Fortuin, and I. Heynderickx, J. Imag. Sci. Technol. **53,** 30201 (2009).
7. J. P. Rolland, M. W. Krueger, and A. Goon, Appl. Opt. **39,** 3209 (2000).
8. K. Akeley, S. J. Watt, A. R. Girshick, and M. S. Banks, ACM Trans. Graph. **23,** 804 (2004).

9. K. J. MacKenzie, R. A. Dickson, and S. J. Watt, Proc. SPIE **7863,** 78631 (2011).
10. K. J. MacKenzie, D. M. Hoffman, and S. J. Watt, J. Vis. **10,** 1 (2010).
11. S. Ravikumar, K. Akeley, and M. S. Banks, Opt. Express **19,** 20940 (2011).
12. K. Akeley, "Achieving near-correct focus cues using multiple image planes", PhD. Thesis (Stanford University, 2004).
13. S. Liu and H. Hua, Opt. Express **18,** 11562 (2010).
14. S. Shiwa, K. Omura, and F. Kishino, Journal of the SID **4,** 255 (1996).
15. T. Shibata, T. Kawai, K. Ohta, M. Otsuki, N. Miyake, Y. Yoshihara, and T. Iwasaki, Journal of the SID **13,** 665 (2005).
16. S. Suyama, M. Date, and H. Takada, Jpn. J. Appl. Phys. **39,** 480 (2000).
17. S. Liu and H. Hua, Opt. Lett. **34,** 1642 (2009).
18. S. Liu, H. Hua, and D. Cheng, IEEE Trans. Visual. Comput. Graph. **16,** 381 (2010).
19. G. D. Love, D. M. Hoffman, P. J. W. Hands, J. Gao, A. K. Kirby, and M. S. Banks, Opt. Express **17,** 1716 (2009).
20. S. C. McQuaide, E. J. Seibel, J. P. Kelly, B. T. Schowengerdt, and T. A. Furness, Displays **24,** 65 (2003).
21. B. T. Schowengerdt, M. Murari, and E. J. Seibel, Journal of the SID **41,** 653 (2010).
22. D. Cheng, Y. Wang, H. Hua, and J. Sasian, Opt. Lett. **36,** 2098 (2011).