Light field head-mounted display with correct focus cue using micro structure array

Weitao Song (宋维涛), Yongtian Wang (王涌天), Dewen Cheng (程德文)*, and Yue Liu (刘 越)

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

**Corresponding author: cdwlxk@bit.edu.cn*

Received Mach 31, 2014; accepted May 7, 2014; posted online May 30, 2014

A new type of light field display is proposed using a head-mounted display (HMD) and a micro structure array (MSA, lens array or pinhole array). Each rendering point emits abundant rays from different directions into the viewer's pupil, and at one time the dense light field is generated inside the exit pupil of the HMD through the eyepiece. Therefore, the proposed method not only solves the problem of accommodation and convergence conflict in a traditional HMD, but also drastically reduces the huge data in real three-dimensional (3D) display. To demonstrate the proposed method, a prototype is developed, which is capable of giving the observer a real perception of depth.

OCIS codes: 110.0110, 120.2040, 330.7322, 330.1400.

doi: 10.3788/COL201412.060010.

Three-dimensional (3D) display technique is undergoing rapid progress in recent decades and has various applications in different fields. Many 3D displays are based on binocular parallax which is a main factor to perceive depth, including stereoscopic displays with special glasses[1] and autostereoscopic displays[2], for instance. Many psychophysical and usability studies have suggested these kinds of displays are subject to visual confusion and fatigue due to the discrepancy between accommodation and convergence[3]. Many approaches hopeful to provide 3D imagery with both correct focus and parallax cues have been presented in the past, including super multi-view displays[4,5], volumetric displays[6,7], integral imaging[8,9], multilayer displays[10], and holographic displays[11,12]. To provide the correct focus, each displayed point should emit abundant light rays which enter the viewer's pupil from different directions in order to make the shape of the crystalline lens adjust to perceive a correct depth. However, a huge amount of data is needed in such a true 3D display, which prevents the relevant technologies from being widely accepted for many demanding applications or daily usage, no matter whether these data are in the form of voxels in volumetric displays, geometrical rays in light field displays, light wave in holographic displays, or even only one directional parallax in super multi-view displays.

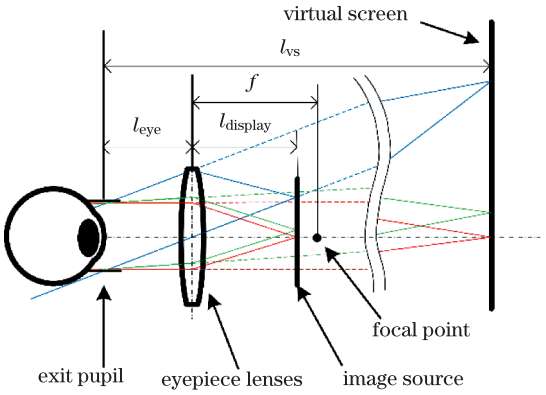

Fig. 1. Schematic diagram of the traditional HMD system.

Offering a type of solution to mobile displays for wearable computing, head-mounted displays (HMDs) have great application potential. An HMD can realize 3D display with binocular parallax, but again it suffers from the problem of the discrepancy between accommodation and convergence, and this problem becomes more evident in an optical see-through HMD for augment reality. Figure 1 shows a schematic diagram of the traditional HMD system. Only one virtual screen is generated by the micro-display device and the eyepiece, and the distance between the eye and the virtual screen (l_{vs}) can be expressed as

$$l_{vs} = \frac{(f - l_{display})\, l_{eye} + f l_{display}}{f - l_{display}}, \tag{1}$$

where f is the effective focal length of the eyepiece, $l_{display}$ is the distance between the eyepiece and the micro-display device, and l_{eye} is the eye relief. To solve this problem, many methods have been proposed to generate dual or multiple focal planes in HMD by means of spatial-multiplexed[13,14] and time-multiplexed methods[15,16]. The complexity of these systems will be increased by employing multiple micro-displays or half mirrors in the spatial-multiplexed methods, or by the active elements in the time-multiplexed methods.

In this letter, we propose a light field head-mounted display (LF-HMD) combining a traditional HMD and integral imaging, which can provide dense light field particularly at the exit pupil of the HMD. As the viewing zone for the generated rays is quite small compared to other true 3D display methods mentioned above, a relatively small amount of data is required to form the 3D image with correct depth information for the crystalline lenses of the eyes. The modules of the proposed system are fixed by its mechanical structure, and the relative positions of the eyepiece, the micro structure array (MSA) and the image source are unchanged. Therefore, we do not need to track the eye pupil but the position of the HMD when the viewer surfs around, which can ease the computing burden.

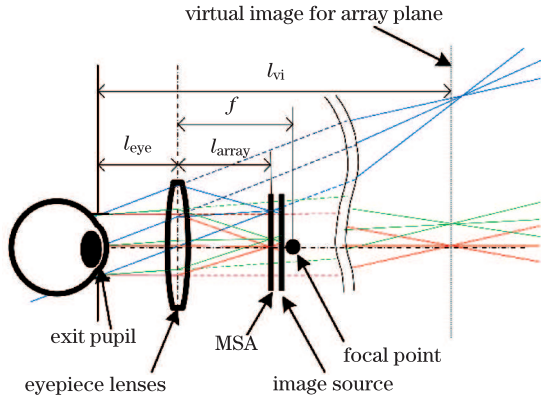

Fig. 2. Schematic diagram of the proposed LF-HMD system.

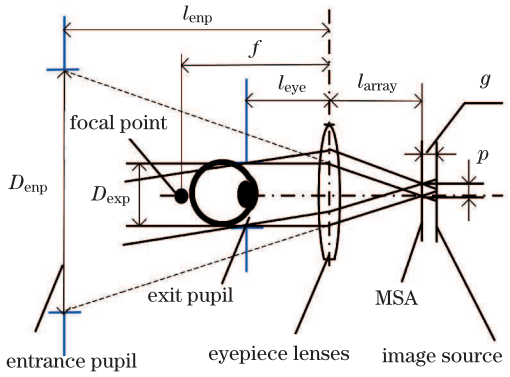

Fig. 3. Schematic diagram of exit pupil for the LF-HMD system.

Figure 2 shows the schematic diagram of the proposed LF-HMD system. A pinhole array or lens array, acted as the MSA, is inserted between the eyepiece and the image source. Compared with conventional integral display systems[17,18], in which light rays are also adjusted by convex lens or concave lens, the proposed method is used for near-eye displays by generating the dense light field particularly at the exit pupil of the HMD, thus, relatively low resolution display device can be employed to achieve a real 3D display. Moreover, more complex optics such as a freeform prism is employed in the proposed system to obtain the 3D scenery with a large field of view (FOV). The position of the virtual image for the MSA plane ($l_{\rm vi}$) can be changed by the effective focal length of the eyepiece (f) and the distance between the eyepiece and the MSA ($l_{\rm array}$), which is expressed as

$$l_{\rm vi} = \frac{(f - l_{\rm array})\, l_{\rm eye} + f l_{\rm array}}{f - l_{\rm array}}. \tag{2}$$

In order to match the zone with the generated light field and the exit pupil of eyepiece lenses, the entrance pupil of the eyepiece lenses and the entrance pupil determined by the MSA and the image source should be overlapped (as shown in Fig. 3). The entrance pupil of eyepiece lenses is the image of the exit pupil, so the distance between entrance pupil and eyepiece lenses ($l_{\rm enp}$) and the diameter of entrance pupil ($D_{\rm enp}$) can be expressed as

$$l_{\rm enp} = \frac{f l_{\rm eye}}{f - l_{\rm eye}}, \tag{3}$$

$$D_{\rm enp} = \frac{f}{f - l_{\rm eye}} D_{\rm exp}, \tag{4}$$

where $D_{\rm exp}$ is the exit pupil diameter of the eyepiece. Therefore, the pitch of the MSA p and the gap between the MSA and the image source g should satisfy the following equation:

$$D_{\rm enp} = \frac{l_{\rm enp} + l_{\rm array}}{g} p. \tag{5}$$

If the condition of Eq. 5 is not met, the exit pupil diameter of the system would be the minimum of the zone with generated rays by the MSA and the exit pupil of the eyepiece:

$$D_{\rm exp,\,system} = \min\left[D_{\rm exp}, \frac{f l_{\rm eye} + (f - l_{\rm eye})\, l_{\rm array}}{g f} p\right] \tag{6}$$

and the viewer can obtain correct light field only when he puts his eye right at the exit pupil position.

Light field generated by the proposed method is shown in Fig. 4 (both the rays in red and in green), and the size of the zone with generated light field at the exit pupil position is $D_{\rm exp,\,system}$. The angular resolution $\Delta\theta$ is determined by the pitch of the MSA and the distance between the eyepiece and the MSA, and it can be written as

$$\Delta\theta = \arctan(p / l_{\rm array}). \tag{7}$$

When the viewer puts his eye right at exit pupil, he can receive the light rays that enter his pupil. The size of the eye pupil is also plotted in Fig. 4, and the light field received by the viewer are shown in red. Only when each rendering point emits at least two rays from different directions into the viewer's pupil, can corrected perception of depth be obtained. The lateral resolution Δl at the depth of D can be expressed as

$$\Delta l = D\Delta\theta / \cos\omega, \tag{8}$$

where ω is the FOV of the displayed object.

According to the relationship in similar triangles, the depth resolution Δd at the depth of D can be obtained as

$$\Delta d = D\Delta l / E. \tag{9}$$

Substituting Δl with Eq. 8 we can see that Δd is determined by the angular resolution, the size of eye pupil and the position of displayed object

$$\Delta d = D^2 \Delta\theta / E\cos\omega. \tag{10}$$

In order to render a 3D object in our system, let the exit pupil be the x-y plane and the virtual image for array plane in the object space be the u-v plane. Also we define the x_a-y_a plane on the array plane and the x_d-y_d plane on the image source plane. The MSA consists of $K_1 \times K_2$ microelements, which are indexed with $k = [k_1,\ k_2]^T$, $k_1 \in \{0, \cdots,\ K_1 - 1\}$, $k_2 \in \{0,\ \cdots,\ K_2 - 1\}$.

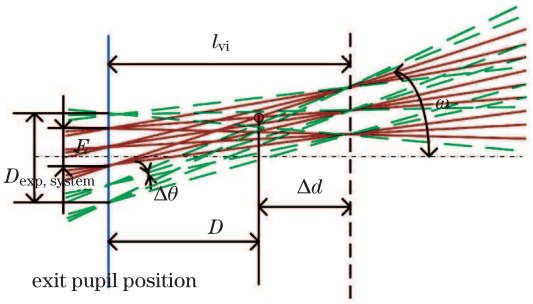

Fig. 4. Schematic diagram for resolution analysis.

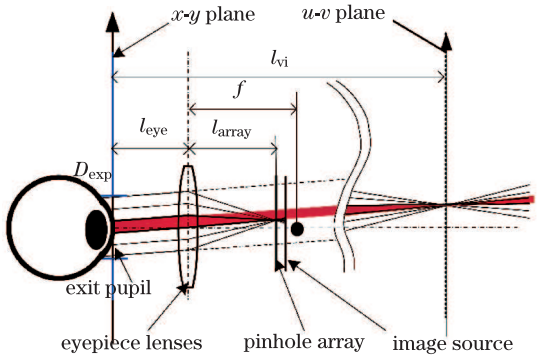

Fig. 5. Schematic diagram of the light field coordinating to one pixel in the LF-HMD system.

Their centers locate at $c_k = p_k + [x_{a0},\ y_{a0}]$ with spacing p and prime microelement coordinate $[x_{a0},\ y_{a0}]$. The element image for the microelement can be defined as a set given as

$$P = (x_d,\ y_d) \left| \left| P - c_k \cdot \frac{l_{enp} + l_{array} + g}{l_{enp} + l_{array}} \right| \right. \leqslant \frac{f g D_{exp,\ system}}{(l_{enp} + l_{array})\,(f + l_{eye})}. \quad (11)$$

The color and the intensity of pixel data on image display can be given by the integral of light field defined by the x-y plane and the u-v plane (as shown in Fig. 5). Supposing the pixel is in the element image for the k^{th} microelement, and pixel data can be given as

$$I = \frac{1}{l_{vi}^2} \int L_F\,(x, y, u, v) \cos^4\theta \mathrm{d}x\mathrm{d}y\mathrm{d}u\mathrm{d}v, \quad (12)$$

where L_F $(x,\ y,\ u,\ v)$ is the light field, which is a 4D function parameterized by the two planes as used in Ref. [19]. The integral zone in Eq. 12 can be calculated by the image on the u-sv plane for the k^{th} microelement and the projected area on the exit pupil by the pixel through the eyepiece lenses and the MSA.

To demonstrate the proposed method, a prototype has been developed using a freeform optical see-through HMD presented in Refs. [20, 21]. Figure 6 is the optical layout of the LF-HMD system. The effective focal length of the wedge-shaped freeform prism is 15 mm, and it consists of three surfaces, labeled 1, 2, and 3 respectively. A pinhole array is used as MSA, and it determines the particular direction of the rays emitted from the micro-display device. Rays with the particular direction are first refracted by surface 3 close to the pinhole array. After two consecutive reflections by surfaces 1 and 2, the rays are transmitted through the surface 1 and reach the exit pupil of the system. An auxiliary element, which consists of two freeform surfaces labeled 2 and 4, is attached to the prism to obtain a see-through view, and surface 4 is designed in order to maintain a non-distortion real-world scene.

Figure 7 shows the photographs of the freeform prism, pinhole array, micro-display device and the monocular bench prototype of an optical see-through LF-HMD. In the prototype, the image source is a 0.61" full color organic light emitting diode (OLED) with 852×600 pixels and the size of each pixel is 15×15 (μm). The pinhole array is composed of 200×150 rectangular holes in the same size as one pixel, which is manufactured by lithographic technique. The pitch is 60×60 (μm) and the distance between the pinhole array and micro-display device is 0.5 mm. The pinhole array is adjusted to be imaged at a virtual distance of 1000 mm. Image data are transferred onto the micro-display device to render a light field of inclined black-white stripes, and the distance between virtual strips and the exit pupil of the developed system ranges from 200 mm and 1000 mm (shown in Fig. 8).

Figure 8 shows the experimental setup of the prototype. Two reference cards with some content are placed in front of the FFS prism for comparison. The distances from the cards to the exit pupil of the FFS prism are 200 mm and 1000 mm, respectively.

Figure 9 gives photos of the real cards along with virtual inclined black-white stripes taken by a commercial mobile phone camera. As the pupil of the camera is only 2 mm that is even smaller than the normal size of eye's pupil, it can be used to simulate the accommodation of crystalline lens. The photos show the scenes focused on the reference targets at 200 mm and 1000 mm, respectively. When the camera focuses on the near reference

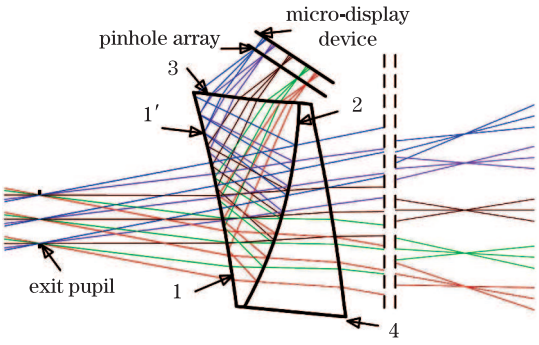

Fig. 6. Optical layout of the developed LF-HMD system.

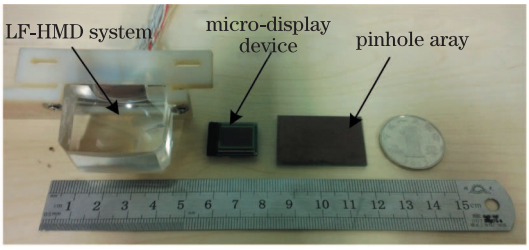

Fig. 7. Photographs of the pinhole array, micro-display device, and the monocular bench prototype of the optical see-through LF-HMD.

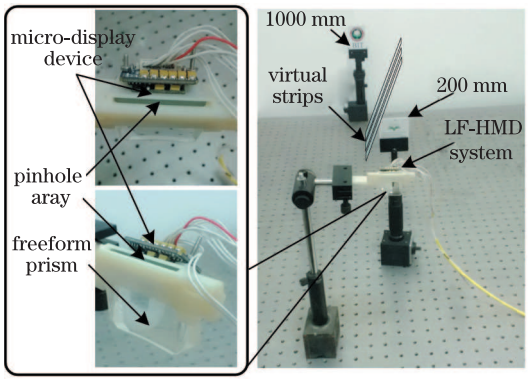

Fig. 8. Photographs of the setup for the experimental system, the monocular bench prototype from top view, and the prototype from side view.

(a)

(b)

Fig. 9. Virtual sloping black-white stripes are rendering with two real cards, one at 200 mm and the other at 1000 mm away from the exit pupil, in the constructed display system. The digital camera focuses on the reference targets at (a) 200 and (b) 1000 mm, respectively.

target, the farther card and the farther part of the virtual stripes become blur due to defocusing, vice versa. When farther reference target is in focus, the nearer card and the nearer part of the virtual stripes become blur. Thus, different from only one virtual display plane in conventional HMD, different depths can be provided by our proposed system by rendering dense light field, and only small number of image data is used in this method, Moreover, the LF-HMD can be compact and lightweight but with a large field of view and depth of perception.

In conclusion, a novel method to realize light field display using LF-HMD has been proposed and verified by a proof-of-concept experiment. In the prototype, the dense light field can be generated at the exit pupil and the viewer can obtain correct depth. We expect that this new form of light field display will be helpful in realizing a practical 3D image display system. However, some difficulties still exist in the alignment of the MSA and the micro-display device. Future works include studies on precise calibration and fast rendering method for the light field generated by the proposed system.

This work was partially supported by the National Basic Research Program of China (No. 2013CB328805), the National Science Foundation of China (NSFC, No. 61205024, 61178038), and the National Key Technology R&D Program (No. 2012BAH64F03). We would like to thank Synopsys for providing the education license of CODE V.

References

1. B. Javidi and F. Okano, *Three-dimensional Television, Video, and Display Technologies* (Springer, 2002).
2. D. Fattal, Z. Peng, T. Tran, S. Vo, M. Fiorentino, J. Brug, and R. G. Beausoleil, Nature **495,** 7441 (2013).
3. D. M. Hoffman, A. R. Girshick, K. Akeley, and M. S. Banks, J. Vis. **8**, 3 (2008).
4. Y. Kajiki, H. Yoshikawa, and T. Honda, Proc. SPIE **3012,** 154 (1997).
5. Y. Takaki and N. Nago, Opt. Express **18,** 8824 (2010).
6. E. Downing, L. Hesselink, J. Ralston, and R. Macfarlane, Science **273,** 1185 (1996).
7. W. Song, Q. Zhu, T. Huang, Y. Wang, and Y. Liu, SID Symposium Digest of Technical Papers, **44,** 318 (2013).
8. G. Lippmann, J. Phys. **7,** 821 (1908).
9. Y. Kim, K. Hong, and B. Lee, 3D Research **1,** 17 (2010).
10. G. Wetzstein, D. Lanman, W. Heidrich, and R. Raskar, ACM Trans. Graph. **30,** 95 (2011).
11. M. Stanley, M. A. Smith, A. P. Smith, P. J. Watson, S. D. Coomber, C. D. Cameron, C. W. Slinger, and A. Wood, Proc. SPIE **5249,** 297 (2004).
12. Y. Pan, Y. Wang, J. Liu, X. Li, and J. Jia, Appl. Opt. **52,** A290 (2013).
13. J. P. Rolland, M. W. Krueger, and A. Goon. Appl. Opt. **39,** 3209 (2000).
14. D. Cheng, Q. Wang, Y. Wang, and G. Jin, Chin. Opt. Lett. **11,** 031201 (2013).
15. S. Liu and H. Hua, Opt. Lett. **34,** 1642 (2009).
16. S. Suyama, M. Date, and H. Takada, Jpn. J. Appl. Phys. **39,** 480 (2000).
17. S. W. Min, M. Hahn, J. Kim, and B. Lee, Opt. Express **13,** 4358 (2005).
18. J. Hong, S.-W. Min, and B. Lee, Appl. Opt. **51,** 4201 (2012).
19. R. Ng, "Digital light field photography", PhD. Thesis (Stanford University, 2006).
20. D. Cheng, Y. Wang, H. Hua, and M. M. Talha, Appl. Opt. **48,** 2655 (2009).
21. Q. Wang, D. Cheng, Y. Wang, H. Hua, and G. Jin, Appl. Opt. **52,** C88 (2013).