Depth-expansion of Light-field Head-mounted Display using a Lens Array with Two Focal Lengths

Yasutaka Maeda[1], Daiichi Koide[1], Hisayuki Sasaki[1], Takuya Handa[1], Kensuke Hisatomi[1]

maeda.y-ki@nhk.or.jp
[1]Japan Broadcasting Corporation, 1-10-11 Kinuta, Setagaya-ku, Tokyo 157-8510, Japan
Keywords: HMD, VR, Light-field, Spatial imaging

ABSTRACT

Light-field head-mounted displays show three-dimensional (3D) images; however, the quality of the 3D images with depth is degraded by the defocus blur of rays. We propose a method to extend the depth of field using a lens array, consisting of lenses with different focal lengths in a checkerboard pattern.

1 Introduction

The issue with binocular head-mounted displays (HMDs) is the vergence-accommodation conflict (VAC) that causes visual fatigue and discomfort during prolonged use [1]. The VAC is a mismatch between the vergence and the accommodation depth; in binocular HMDs, the accommodation depth is fixed to the magnified display surface. Light-field HMDs can reconstruct focus-adjustable three-dimensional (3D) images and are expected to resolve the VAC. A light-field HMD consists of a display, a lens array (pinhole array), and an eyepiece [2,3]. The display shows elemental images consisting of slightly different views of the 3D object. An intermediate image is formed by elemental images and the lens array. By viewing the intermediate image through the eyepiece, a magnified 3D image can be seen. Intermediate and 3D images are created by the intersection of multiple rays emerging from pixels of the display. Therefore, they are formed at various depths other than the conjugate surface of the display. However, at depths away from the conjugate plane, the diffraction of rays and defocus blur of the lens array degrade the quality of the 3D image.

The depth of field (DOF) is the depth at which a 3D image can be displayed without the loss of quality. Various methods have been demonstrated to expand the DOF using tunable lenses [3,4]. The implementation of tunable lenses can increase the DOF but requires an optical relay system, which increases the size of the HMD. A method using a digitally switchable aperture array and a lens array consisting of elemental lenses with different focal lengths in the center and periphery has been proposed [6]. While this method does not affect the size of the optical system and thus is very effective, there are concerns about increased system complexity and low frame rates due to time-resolved driving.

In this study, we propose a method to improve DOF using a lens array composed of two elemental lenses with different focal lengths in a checkerboard pattern. Since the depth of the conjugate plane differs for each elemental lens, the DOF can be improved by changing elemental images according to the depth of the 3D object. This method does not increase the number of optical elements and only requires some modification of the rendering. We have verified the effectiveness of the proposed method via optical simulation.

2 Proposed method

2.1 Optical system

Figure 1 shows the schematic of the conventional light-field HMD. The distance between the display and the lens array is shorter than the focal length of the lens array. Light emitted from the pixels of the display is refracted by the lens array and eyepiece to reach the eye. After passing through the lens array and tracing the rays in the opposite direction, they are focused on a single point. The plane on which rays are focused by the lens array is called the intermediate image plane, which is the conjugate plane of the display. The relationship between the lens array and the intermediate image plane is expressed by the following equation:

$$\frac{1}{a}-\frac{1}{b}=\frac{1}{f_{\mathrm{array}}} \tag{1}$$

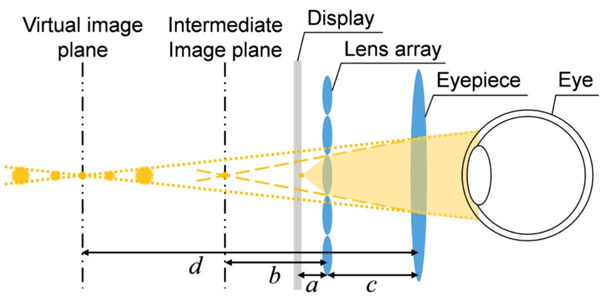


Fig. 1 Conventional light-field HMD

where a and b are the distances between a display and a lens array and a lens array and an intermediate image plane, respectively. The focal length of the lens array is denoted by f_{array}. The rays after passing through the eyepiece are also traced in the opposite direction, and they are again focused on a single point. The plane on which rays are focused by the eyepiece is called the virtual image plane, which is also the conjugate plane of the display. The relationship between the eyepiece and the virtual image plane is expressed by the following equation:

$$\frac{1}{b+c}-\frac{1}{d}=\frac{1}{f_{\text{eye}}} \tag{2}$$

where c and d are the distances between a lens array and an eyepiece and an eyepiece and a virtual image plane, respectively. The focal length of the eyepiece is denoted by f_{eye}. When the eye is focused on the virtual image plane, the rays entering the eye get focused on a point on the retina. However, if the eye is focused in front of or behind the virtual image plane, the rays get spread out on the retina. The depth of the virtual image plane should be determined by considering the depth position of the 3D image to be displayed, which is generally set at approximately 1.0 m from the eye.

Figure 2 shows the schematic of the proposed light-field HMD. The lens array consists of elemental lenses of two different focal lengths arranged in a checkerboard pattern. With the addition of elemental lenses with a focal length of f'_{array}, a new intermediate image plane and a virtual image plane are added. The distance b' between the lens array and the new intermediate image plane is considered in the following equation:

$$b'=\frac{af'_{\text{array}}}{f'_{\text{array}}-a} \tag{3}$$

The distance d' between the eyepiece and the new virtual image plane is considered in the following equation:

$$d'=\frac{(b'+c)f_{\text{eye}}}{b'+c-f_{\text{eye}}} \tag{4}$$

Once the depth of the new virtual image plane is determined, the position of the intermediate image plane can be determined from equation (4), and the required focal length can be determined from equation (3).

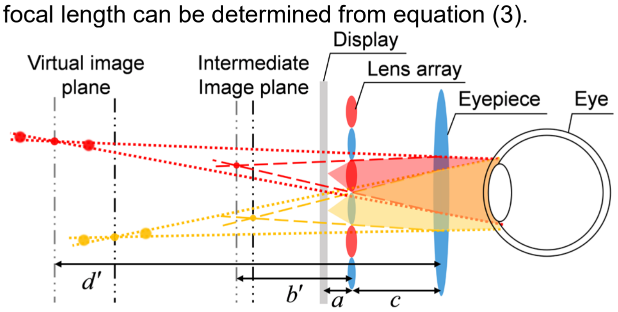


Fig. 2 Proposed light-field HMD

As shown in Fig. 2, the proposed method has two virtual image planes; even at depths where the quality of the 3D image degrades with one virtual image plane, the quality can be maintained by placing another virtual image plane. In other words, the DOF can be expanded by different elemental lenses to form a 3D image depending on the depth position.

The parameters and arrangements of the display and optical elements used in the simulation are shown in Table 1 and Table 2. Figure 3 shows the scheme of the lens array. The conventional method employed a lens array consisting of elemental lenses with a focal length of 19.0 mm, while the proposed method employed a lens array consisting of elemental lenses with focal lengths of 19.0 mm and 20.0 mm.

Table 1 Simulation conditions

Display	Pixel pitch	0.024 mm
Lens array	Pitch	1.55 mm
	Focal length	19.0 mm, 20.0 mm
	Arrangement	Orthogonal
Eyepiece	Focal length	250 mm

Table 2 Variables of the distance

Display — Lens array	16.9 mm
Lens array — Eyepiece	50.7 mm
Eyepiece — Pupil plane	60.0 mm
Eyepiece — Front virtual image plane	450 mm
Eyepiece — Rear virtual image plane	1100 mm

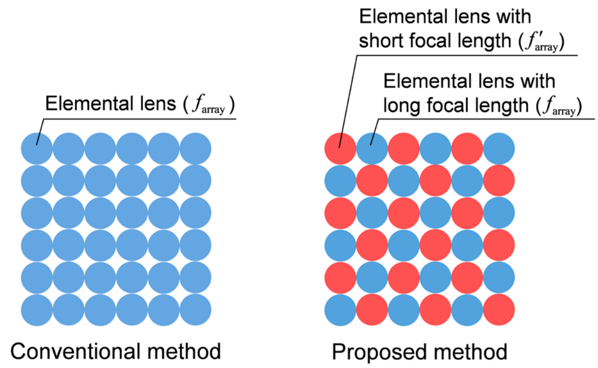


Fig. 3 Scheme of lens array

2.2 Elemental images generation

Elemental images are generated by ray tracing (Fig. 4). Pinhole arrays are used instead of lens arrays to trace the path of the chief ray through the center of each elemental lens. The display, pinhole array, eyepiece, and pupil plane are arranged. In the conventional method, elemental images were generated in all pixels using the following procedure:

i Define a ray from the display pixel center and the pinhole position.

ii Extend the defined ray to the eyepiece and refract it at the contact position.

iii Extend the refracted ray to the pupil plane and retroreflect.

iv Judge contact between the retroreflected ray and the 3D model and assign the color of the contacted area to the pixel.

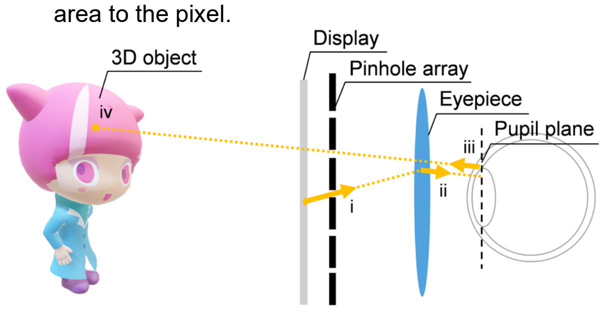


Fig. 4 Process for generating elemental images

In the proposed method, step iv needs to be modified. The lens array consists of elemental lenses with two different focal lengths, and each ray from the pixel passes through one of the element lenses. In other words, each ray from the pixel passes through either of the virtual image planes. A reference depth position between the two virtual image planes is determined, and color assignment is determined from the relationship between the 3D model and the reference depth in step iv. When the 3D model is in front of the reference depth, a color is assigned if it passes through an elemental lens having a larger focal length. Conversely, when the 3D model is deeper than the reference depth, a color is assigned if it passes through an elemental lens having a shorter focal length. In this way, elemental images of the proposed method are generated.

In the proposed method, the two virtual image planes are at depths of 450 mm and 1100 mm, and the reference plane is set at 630 mm.

3 Simulation and Results

3.1 Simulation

The retinal image is used to observe the appearance of the 3D image when viewed through the light-field HMD. The retinal image was generated via a simulation using OpticStudio 22.2 (Zemax, LLC), which is an optical design software. The display, lens array, and eyepiece were placed in the CG space of the software, as summarized in Table 1 and Table 2; moreover, an additional lens and detector were placed as a virtual crystalline lens and retina, respectively (Fig. 5). The focal length of the crystalline lens was determined by ensuring that the rays from the reconstructed 3D image are focused on the detector. The lens array, eyepiece, and crystalline lens were all ideal lenses with no optical aberrations. Elemental images were used as a surface light source, and ray tracing was performed by randomizing the starting position and direction to generate a retinal image.

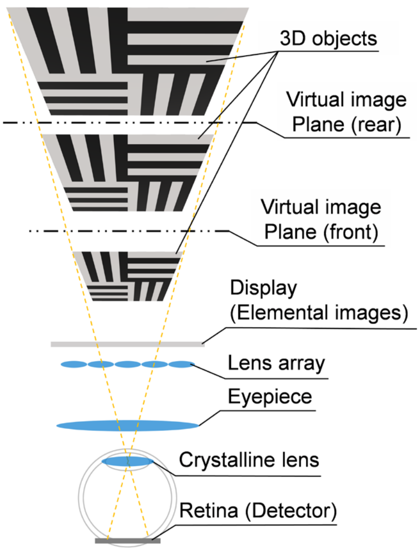


Fig. 5 Arrangement of 3D objects

2D plates with a black and white edge pattern image pasted for texture were used as 3D objects. These objects were placed at depths ranging from 400–2000 mm from the crystalline lens position (Fig. 5). The size of the 3D objects was varied such that the visual angle of the 3D images was constant at any depth. Elemental images were generated using the method described in Section 2.2, and retinal images were generated using this method. The proposed and conventional methods were compared for the sharpness of the retinal images at each depth position.

3.2 Results

Figure 6 shows elemental and retinal images generated by the conventional method using a lens array consisting of elemental lenses with a focal length of 19.0 mm. Figures 6 (b)–(d) are simulated retinal images, where the objects within the dotted lines correspond to the focus positions. In Figs. 6 (c) and (d), the 3D image at the focus position is clear. On the contrary, in Fig. 6 (b), the 3D image is blurred even though the focus is set at 400 mm. This was due to focus blur and was caused by reconstructing the 3D image at a distance from the depth of the virtual image plane.

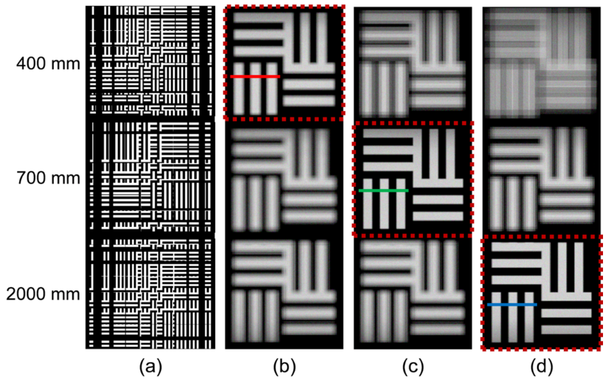


Fig. 6 (a) Generated elemental images by the conventional method and (b)–(d) simulated retinal images focused at 400 mm, 700 mm, and 2000 mm

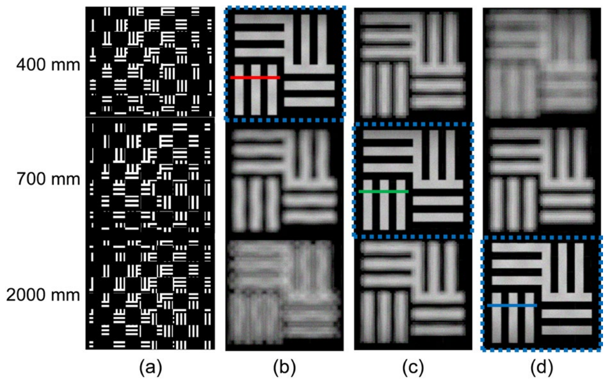
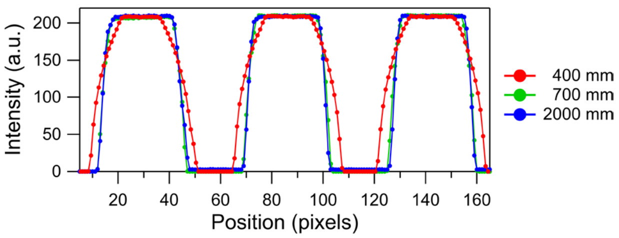


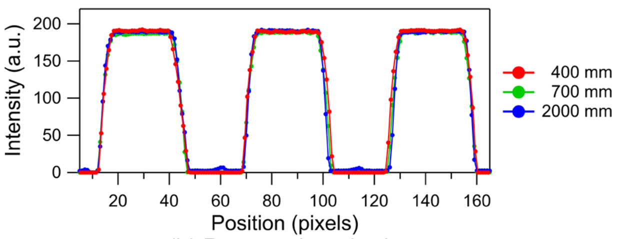
Fig. 7 (a) Generated elemental images by the proposed method and (b)–(d) simulated retinal images focused at 400 mm, 700 mm, and 2000 mm

Figure 7 (a) shows elemental images generated via the proposed method using a lens array consisting of elemental lenses with focal lengths of 19.0 mm and 20.0 mm. Because elemental lenses of different focal lengths are arranged in a checkerboard pattern, the elemental images also alternate between color-assigned and color-unassigned areas. Since the reference depth position is set at 630 mm, for a 3D object at 400 mm, the color is assigned to the pixel in front of the elemental lens with a focal length of 20.0 mm. Conversely, for 3D objects at 700 mm and 2000 mm, colors are assigned to pixels in front of elemental lenses with a focal length of 19.0 mm. Figures 7 (b)–(d) show the retinal images generated using elemental images. In addition to the 700 mm and 2000 mm depths, the quality of the 3D image is maintained at 400 mm, as shown in Fig. 7 (b). However, the brightness of the retinal image is slightly lower than the image obtained via the conventional method due to the reduced pixel utilization.

Figure 8 shows the line profiles of the red, green, and blue lines in Figs. 6 and 7. The horizontal axis is position, and the vertical axis is intensity. In the conventional method, the graph shape is rounded at 400 mm compared to the graph shape at 700 mm and 2000 mm. On the contrary, the proposed method maintained the rectangular shape even at 400 mm. This result indicates that the DOF was expanded.



(a) Conventional method



(b) Proposed method

Fig. 8 Line profile of the retinal images

4 Discussion

The degradation of 3D images at a distance from the virtual image plane was improved using a lens array with two focal lengths. As shown in Figs. 6 and 7, the proposed method maintains the quality of the 3D image at 400 mm by the optical simulation. Because the proposed method does not require additional relay optics or time-resolved driving elements as employed in previous studies, the optical system is not large, and the frame rate would be guaranteed. It inherits the feature of the simple device structure of a light-field HMD using the lens array. On the contrary, the display's pixel utilization rate decreased, resulting in lower brightness. This is considered to have little impact because the user views the HMD in a closed space and rays of light from multiple pixels reach the eyes.

5 Conclusions

We proposed a method using a lens array with two focal lengths for light-field HMDs. The DOF expansion was demonstrated via optical simulation. Since this method requires only modification of the lens array and rendering, the optics are not too large. This is an advantage in HMDs for VR where the elements are placed in front of the face. In the future, we will produce a checkerboard pattern lens array with two focal lengths and proceed to evaluate image quality using actual equipment.

References

[1] D.M. Hoffman, A.R. Girshick, K Akeley, and M.S. Banks, “Vergence-accommodation conflicts hinder visual performance and cause visual fatigue,” J. Vis., Vol. 8, No. 3:33, pp. 1–30 (2008).

[2] H. Hua and B. Javidi, “A 3D integral imaging optical see-through head-mounted display,” Opt. Express Vol. 22, No. 11, pp. 13484–13491 (2014).

[3] W. Song, Y. Wang, D. Cheng, and Y. Liu, “Light field head-mounted display with correct focus cue using micro structure array,” Chin. Opt. Let., Vol. 12, No. 6, pp. 60010–60013 (2014).

[4] X. Shen and B. Javidi, "Large depth of focus dynamic micro integral imaging for optical see-through augmented reality display using a focus-tunable lens," Appl. Opt. Vol. 57, No. 7, pp. B184–B189 (2018).

[5] H. Huang and H. Hua, “High-performance integral-imaging-based light field augmented reality display using freeform optics,” Opt. Express, Vol. 26, No. 13, pp. 17578–17590 (2018).

[6] X. Wang and H. Hua, "Depth-enhanced head-mounted light field displays based on integral imaging," Opt. Lett. Vol. 46, No. 5, pp. 985–988 (2021).