Distortion corrected tomographic near-eye displays using light field optimization

Siwoo Lee, Seungjae Lee, Dongyeon Kim, and Byoungho Lee[*]

School of Electrical and Computer Engineering, Seoul National University, Gwanak-Gu Gwanakro 1, Seoul 08826, Republic of Korea

[*]*byoungho@snu.ac.kr*

Abstract: Several multifocal displays have been proposed to provide accurate accommodation cues. However, multifocal displays have an undesirable feature, which is especially emphasized in near-eye displays configuration, that the field of views (FOVs) of the virtual planes change over depth. We demonstrate that this change in FOV causes image distortions, which reduces overall image quality, and depth perception error due to the variation of image sizes according to depths. Here, we introduce a light field optimization technique to compensate for magnification variations among the focal planes. Our approach alleviates image distortions, especially noticeable in the contents with large depth discontinuity, and reconstructs the image size to precise depths, while maintaining a specific tolerance length for the target eye relief. To verify the feasibility of the algorithm, we employ this optimization method for the tomographic near-eye display system to acquire the optimal image and backlight sequences for a volumetric scene. In general, we confirm that the structural similarity index measure of reconstructed images against ground truth increases by 20% when the eye relief is 15 mm, and the accommodation cue is appropriately stimulated at the target depth with our proposed method.

1. Introduction

An ideal three-dimensional (3D) displays have the prospects to provide a natural viewing experience analogous to the real world. Accordingly, the growth of head-mounted displays (HMDs) for virtual reality (VR) and augmented reality is advanced. However, users often feel visual fatigue after using HMDs for a long time. The commercially available HMDs provide binocular parallax to stimulate vergence cues for users to perceive depth information. Simultaneously, the users focus on a single fixated physical screen indicating the accommodation cue. The mismatch between the vergence cue and the accommodation cue is referred to as vergence accommodation conflict (VAC), which has been reported to cause visual fatigue [1,2]. A recent study investigated that the supporting accurate focus cues enhanced comfort [3].

Several methods have been proposed to mitigate VAC by providing accurate accommodation cues. The representatives of these methods are light field displays [4,5], holographic displays [6,7], varifocal displays [8,9], and multi-plane displays [10–19]. The light-field displays using microlens arrays provide focus cues and motion parallax, considering the 4D ray space of the light produced by the display [5]. However, this system suffers from a low spatial resolution due to a spatial-angular resolution trade-off. Holographic displays can reconstruct the wavefront of 3D information providing accurate focus cues. Nevertheless, the limited resolution and the pixel size of a spatial light modulator cause a trade-off between the field of view (FOV) and eye box size. The varifocal displays support one focal plane whose depth is adjusted by the user's eye-tracking data. However, this system requires an eye tracker with minimal latency since it should actively track the accommodation of the eye in real-time. Multi-plane displays provide near-correct 3D volumetric scenes by displaying images at different depth planes and are usually classified into four systems [20]. The spatially multiplexed system [10–12] physically

stacks multiple screens and display images at the appropriate depths. However, this system has a disadvantage in its large form factor. The polarization multiplexing system [13] expresses multi-depth planes using polarization states of light, providing a compact form factor without loss in refresh rate. Nevertheless, this system offers a limited number of focal planes since only two orthogonal polarization states are used. In the wavelength multiplexing system [14], although lights of different wavelengths travel to different optical paths providing multiple focal planes, this system has many challenges, such as small FOV and difficulty to express full-color display. In the time-multiplexing system [15–19], a fixed display is repositioned optically by focus-tunable optics. The images at different depths are divided temporally and provided faster than the flicker fusion rate. This system can be designed with smaller form factors than the spatially multiplexed system and can express near-correct focus cues with a wide depth range.

Recently, a tomographic near-eye display [21], which is mainly based on the principle of compressive light-field displays [22–26], has been introduced. This system is a time-multiplexing system using synchronization of digital micromirror device (DMD) and focus-tunable lens (FTL). This technique alleviates VAC by providing a nearly accurate depth cue. For this system, Lee et al. [21] suggested an algorithm to relieve artifacts at occlusion boundaries, which are observed since images at different depths are synthesized additively. However, an undesired feature of such near-eye displays is that the FOV varies relative to the depth of virtual planes if the distance from the eyepiece lens to an eye exists. This variation in FOV causes image distortions and depth perception error. Since a specific range of eye relief needs to be guaranteed due to the user's face structures or additional placement of eyeglasses, these problems should be mitigated to provide a natural viewing experience for users.

Several studies used the 4-f system to assume that the eye relief is zero, which is impractical in HMDs due to its large form factor, to alleviate this problem [21,27,28]. Rathinavel et al. [29] suggested the calibration process that samples the FOVs of focal planes and applies the scaling factor to each focal plane. However, this method requires a post-processing step to correct the FOV variations. Zhan et al. [17] suggested a light field rendering technique that naturally involves eye relief in near-eye displays. We propose a magnification compensated optimization method based on this technique considering FOV variations in multi-layer displays, especially highlighted in the tomographic near-eye display. The ray sampling space is modulated during the optimization process to compensate magnification of each plane. We obtained the optimal parameters such as the number of iterations and the illumination time. We analyzed the image quality tendency due to eye relief. The effectiveness and trade-offs caused by light field rendering, when eye relief exists, are investigated. Our approach alleviates image distortions and depth perception error due to FOV differences at different depths without additional optical elements or post-processing. In addition, we analyzed that the proposed method effectively drives the accommodation cue at the target depth and does not cause depth error. This approach can be applied to other multi-layer displays, especially emphasized in near-eye display configuration [18,21,27,28,30].

In this study, we explain the principles of tomographic near-eye systems and problems caused by magnification variation. Then the optimization method is described how FOV variations, caused by eye relief, are considered in light field sampling to compensate for the depth-dependent magnification. We access the proposed work with simulation and experiments to show that the image distortions are alleviated, and the size of images are reconstructed to precise depths. In the discussion, we verify the feasibility of our work by analyzing the accommodation response and the eye relief tolerance.

2. Background

Recently, Lee et al. [21] have introduced tomographic near-eye display. This system reproduces volumetric scenes with a time-multiplexing system using synchronization of DMD and FTL, as

illustrated in Fig. 1(a). The focal length of FTL sweeps within a specific range according to the applied voltage, allowing multiple depth expressions from a single display panel. Simultaneously, the DMD enables depth expression within the focal length range of FTL by illuminating a specific pixel at a particular depth during one cycle of FTL with high driving speed. However, since this system assumed a constant FOV for each focal plane, expected volumetric scenes are observed only when eye relief is zero in practical situations. When eye relief is not zero, the distortions are noticed in the area with depth discontinuity. Additionally, the incorrect relative size of objects at different depths can cause depth perception error [31].
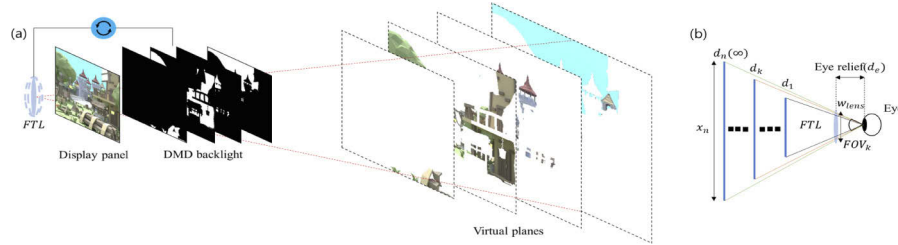

Fig. 1. (a) A system of tomographic near-eye displays. The tomographic display reproduces a volumetric scene by synchronizing a DMD and an FTL. The FTL with fast backlight modulation allows multiple depth expressions with a single RGB image. (b) FOV variations due to eye relief.

If eye relief is present, the magnification of each virtual plane among the virtual planes is different, as shown in Fig. 1(b). The FOV is determined by the limited lens aperture or the size of a magnified plane, as shown in Eq. (1).

$$\text{FOV} = 2\arctan\left[\min\left(\frac{w_{\text{lens}}}{2d_e}, \frac{x_n}{2d_n}\right)\right], \tag{1}$$

where w_{lens} is the diameter of the FTL, d_e is the eye relief, x_n is the width of the farthest virtual plane, and d_n is the distance between the farthest virtual plane and the eye. Assuming w_{lens} is not constrained, the maximum FOV is determined by the width of a farthest virtual plane. Then the maximum magnification difference, ΔM_{max}, is expressed as Eq. (2).

$$\Delta M_{\text{max}} = \frac{\tan\left(\text{FOV}_{\text{max}}\right)}{\tan\left(\text{FOV}_1\right)} - 1 = \frac{d_e}{d_1}, \tag{2}$$

where FOV_{max} is the maximum field of view, and d_1 is the distance between the nearest plane and an eye, as illustrated in Fig. 1(b). The maximum magnification difference, which is the difference between the farthest and the nearest plane, is proportional to the eye relief.

3. Method

3.1. Algorithm

Ideally, we can adjust the size of images for each depth to correct the magnification variations. However, since the state-of-the-art display can refresh at 240 Hz, though it can express four depths by sacrificing the frame rate, this is insufficient to provide continuous focus cues. Hence, we provide a single optimal RGB image shared for all expressed depths without refreshing the display. In this work, we modify the sampling space of focal planes in the optimization step to obtain the optimal 3D scenes by resolving magnification differences. In the conventional light field sampling, the rays intersecting with each pixel of the target plane are projected onto the pupil plane. On the other hand, we sample the rays from the pupil intersecting with the

magnified virtual planes. The sampling space of each of the focal planes is scaled relative to the FOV variations. We assume that the focal planes are perpendicular to the display's optical axis. Figure 2 demonstrates how rays are sampled when point grid images at 0 and 5.5 diopters (D) are aligned with each other. Note that the points appear to be far away rather than overlapped. In the kth focal plane, the magnification ratio M_k is defined as Eq. (3), and the modified sampling space, P_k', is defined as Eq. (4).

$$M_k = \frac{(d_k + d_e)\left(\frac{w_n}{d_e + d_n}\right)}{d_k \tan \theta_n} \approx 1 + \frac{d_e}{d_k}, \tag{3}$$

$$P_k'(x, y, z) = P_k(M_k x, M_k y, z), \tag{4}$$

where P_k is the original sampling space of the kth plane, d_k is the distance between the kth plane and the FTL, d_e is the eye relief, w_n is the width of the last plane, d_n is the distance between the last plane and the FTL, and θ_n is the FOV between the last plane and the pupil center. According to the pixel difference $\left(P_k'(x, y, z) - P_k(x, y, z)\right)/\Delta w_k$, the light field sampled from the multiple layers is modeled by a projection matrix, T. The projection matrix is modified to occupy a magnified finite region determined by the sample spacing to consider magnification. We describe the light field synthesis in the form of $L_i = \sum_{j=1}^{N} T_{(i,j)} (B_j \circ D)$, where N is the number of focal planes, L_i is the vectorized pupil plane images, and $T_{(i,j)}$ indicates the projection matrix between the ith perspective view and the jth depth layer. B_j is the vectorized binary backlight pattern for each focal plane, and D is the vectorized single displayed RGB image with $n \times n$ resolution.
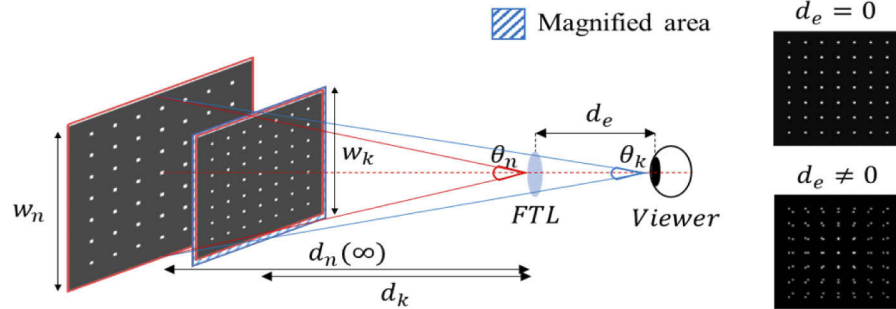

Fig. 2. Light field sampling of point grid images from the focal planes (0D and 5.5D). Due to FOV variations between virtual planes, the point grids are not aligned when viewed from optical axis.

The display image and backlight sequences are computed to reproduce the original light field in the optimization process. Since the focal stack images include the 4D light field information [32], we implemented light field synthesis with multi-view images. The light field synthesis based on multiple perspective images guarantees the tolerance in potential eye movement within the eye box [33]. Target light field is the summation of the multiple view images within the pupil and rendered with Blender. We calculate the optimal displayed image (D) and backlight patterns (B), minimizing artifacts caused by magnification variation and pupil shift. Note that the DMD only supports binary patterns, and the image on the panel is bounded between 0 and 1. These constraints make the problem NP-hard. We solve the relaxation of the NP-hard problem via alternative linear least-squares problem in Eq. (5) with the simultaneous algebraic reconstruction technique (SART) [34].

$$\min \sum_{i=1}^{p^2} \left\| c I_i - \sum_{j=1}^{N/t} T_{(i,j)} (B_j \circ D) \right\|, \tag{5}$$

where tB_j should be an integer, c is an integer constant and I_i is the vectorized target input. The constant c indicates that each pixel of backlight illuminates the corresponding display pixel by c times. In this alternative linear least-squares problem, we update the backlight sequence and RGB image independently. When the backlight sequences are set constant, the RGB image is optimized, and vice versa. The binary constraint is ignored in the optimization process to solve this problem with SART at each iteration. Instead, after each iteration step, we rounded off the backlight values so that the DMD can express binarized backlight patterns. However, we observed that the convergence of optimization is not guaranteed by the rounding process. Thus, we compare the sum of each pixel value before and after rounding. By calculating the differences in each pixel, we compensate each pixel's differences to preserve the total energy of rays at each pixel before and after rounding by comparing sum of each pixel value. The overall optimization process is summarized in Fig. 3. An in-focus RGB center perspective view and a linear blended depth map are applied for the initial condition.
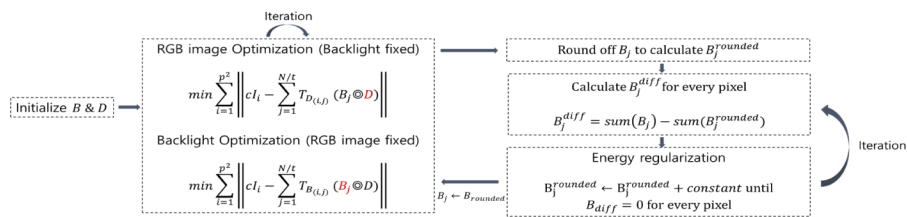

Fig. 3. Flow chart of the optimization process. The backlight sequence and RGB image are updated independently. During the rounding-off process, the energy of each pixel is regularized by updating backlight sequences. The updated backlight sequences are used in the optimization process.

Figure 4(a) demonstrates the overall similarity of retinal images against ground truth as the optimizing iteration increases. The ground truth is the focal image reconstructed with dense light fields. We calculated the display image and backlight sequences for each iteration. According to accommodation states, we rendered focal images to find the number of iterations that reproduce the optimal light field minimizing the optimization time. In the rounding off step once in every 10 iterations, the averaged structure similarity index metric (SSIM) [35] decreases drastically. In general, the average SSIM against ground truth increases as the number of iterations increases. We determined that 80 iterations are sufficient as the SSIM increment converges to less than 1% after 80 iterations. The calculation time for 80 iterations is 680 seconds.

In this system, illumination time is the number of turning on by each pixel, c, divided by the number of layers, N. Since the illumination time has a degree of freedom in the optimization process, we obtained the retinal images relative to the illumination time. In this work, the total number of optimizing layers is N/t, where t is the number of backlights reconstructed by one backlight. In other words, the N/t optimized backlights are linearly blended and are reconstructed into N binary sequences in experiments. The average SSIM of retinal images varying illumination time is measured as shown in Fig. 4(b). The highest SSIM is obtained when the illumination time is 0.3 (red circle). As the illumination time increases, we obtained the results that show reduction of overall image quality involved by blurred details of objects. This will be explained in the discussion section.

3.2. *Hardware*

The experimental setup is illustrated in Fig. 5(a). The system consists of a DLP9000X from Texas Instruments as DMD supporting a maximum frequency of 14,989 Hz for binary patterns and an LED source from Advanced Illumination. With a collimating lens, a customized prism,
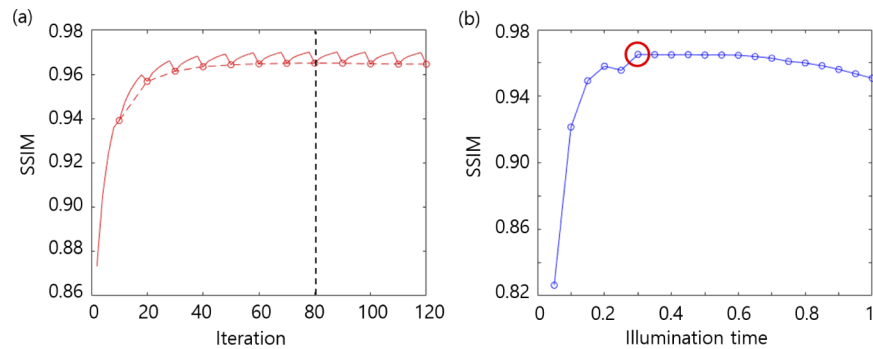
Fig. 4. (a) Average SSIM values of optimized retinal images after each iteration. We conclude that SSIM converge after 80 iterations. (b) Average SSIM of optimized retinal images according to the illumination time. We confirm that the optimal illumination time for optimization is 0.3 (red circle).

and magnification optics, the backlight patterns of the DMD are magnified and projected onto the diffuser on an LCD panel, Topfoison TF60010A model. The DMD is synchronized with the focus-tunable lens (EL10-30-TC-VIS-12D from Optotune) with a diameter of 10 mm, to display a stack of binary frames in each lens cycle. Two synchronized signals are generated using a Labview and a Data Acquisition (DAQ) board from National Instruments. The 4-f relay optics are required because it is difficult to observe where the eye relief is zero due to the CCD specification. This 4-f system is not necessary if an eye relief is accurately measured as viewed with an eye. The sinusoidal wave with a frequency of 60 Hz is applied to the FTL considering the flicker fusion rate of a human eye. The pulse train with 9600 ($2 \times 80 \times 60$) Hz is generated for the DMD signal to produce binary patterns of 80 focal planes. In the half cycle of FTL, backlight according to depth is applied, and black frames are loaded in other half cycles. The summarized synchronization process is illustrated in Fig. 5(b).
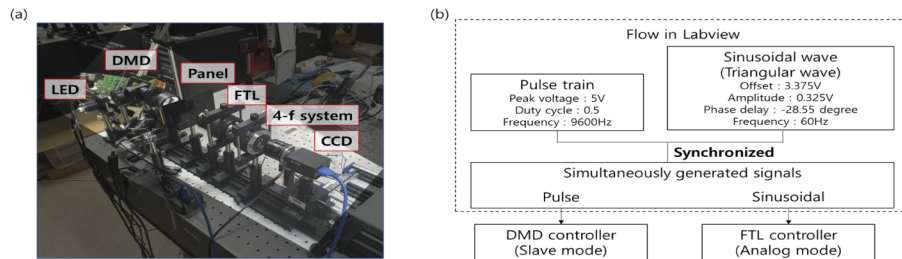

Fig. 5. (a) Experimental setup. (b) Synchronization process with Labview.

4. Results

We used a tunable lens that changes its focal length from +8.3D to +20D, proportional to the applied voltage. If the vibration speed of FTL is extremely fast in this system, the maximum and minimum values are not theoretically applied. Hence voltage values representing 0D at 5.5D were obtained empirically. After that, the lens and DMD are synchronized with a phase shift of −28.5 degrees. We compensated for the pincushion distortion of the backlight from DMD by applying pre-distortion to the backlight sequences.

For retinal reconstruction, rays are sampled from the 7×7 pupil plane within 6 mm × 6 mm through the ray tracing method. By calculating the projection matrix between the focal plane and

the retina plane, point spread functions mapped on the retina relative to accommodative states are determined. Using this projection matrix, we obtained the focal stack of images that are supposed to be seen by the viewer accommodating at focal depths.

The rays are sampled from 7×7 pupil planes within 6 mm $\times$ 6 mm pupil size for optimization. The accommodation cue may be driven more accurately but may degrade spatial resolution if we sample more pupil views [36]. Since the optimization is implemented to reproduce the target light field, it requires a high computational load. The separation between optimizing planes is 0.6D since it is sufficient to trigger accommodation at intermediate planes, minimizing the number of optimizing planes to reduce time cost [37]. The optimized backlights of 10 sequences are recomputed to 80 sequences in 300×300 resolution and expressed within 0D to 5.5D. Figure 6(a) shows the rendered retinal images with 15 mm eye relief when the accommodation is at 1.8D for Castle Scene and 2.4D for Forest Scene, respectively. The averaged SSIM values of focal planes are measured. The enlarged areas show that the optimization effectively alleviates distortions due to magnification variations. Figure 6(b) illustrates that the relative size of objects according to the depth is compensated. We implemented the experiments with Road Scenes to verify the feasibility, as illustrated in Fig. 6(c). The results show the observed images with 15 mm eye relief when the accommodation is at 0.0D. The defects in retinal images without optimization (green box) are alleviated in the optimized retinal images (red box).

As shown in Fig. 7, the artifacts observed due to the difference in image shift by eye movements and magnification variations for each depth are alleviated through optimization to guarantee the tolerance for motion parallax in the eye box. As illustrated in the red box, the proposed optimization allocates the objects to the appropriate depths by compensating relative size and mitigates the noise produced by images displayed at the inappropriate depths (red arrows), compared with occlusion blending.

Yang et al. [38] have suggested that an effective eye relief in HMDs is 22 mm to allow for all types of eyeglasses. To verify the validity of this algorithm in this condition, we demonstrated SSIM values of retinal images against ground truth as a function of eye relief from 0 mm to 42 mm. As illustrated in Fig. 8(a), SSIM decreases linearly as eye relief increases when backlights are not optimized (black) or optimized with occlusion blending (blue). On the other hand, the proposed optimization (red) leads to a slight decrease in SSIM as the eye relief increases. The reduction ratios at 24 mm eye relief are 19%, 20%, and 2% for no optimization, occlusion blending, and proposed optimization, respectively. Figure 8(b) illustrates the retinal images focused on 2.4D with eye relief 24 mm. In Fig. 8(b), the enlarged area shows where the dark gap appears due to depth discontinuity. The dark gap presents the proper position of the pillar. The results of occlusion blending, though the dark gap is softened slightly, show that the position of the pillar is not still correctly placed due to FOV variations. On the other hand, the proposed method reconstructs the objects at the appropriate locations by compensating magnifications. We conclude that the proposed method reproduces an optimal image compensating magnification variations for various target eye reliefs through these results.
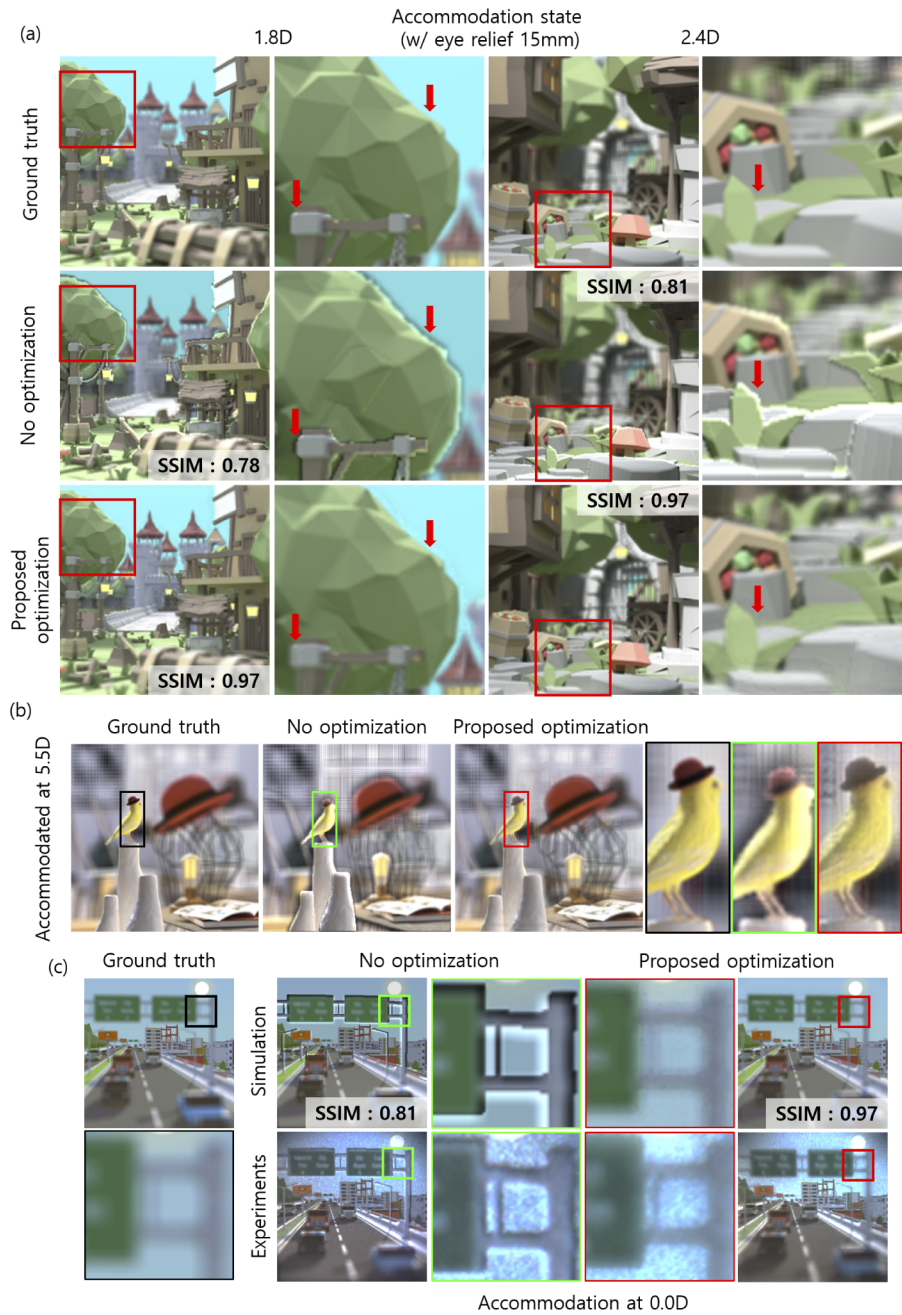
Fig. 6. (a) Retinal images of Castle Scenes at 1.8D and Forest Scenes at 2.4D before and after optimization. (b) Retinal images of Room Scenes focused on 5.5D. The relative size of objects according to the depth is compensated. (c) Simulation and experimental results with Road Scenes at 0.0D. (Source image courtesy: www.cgtrader.com)
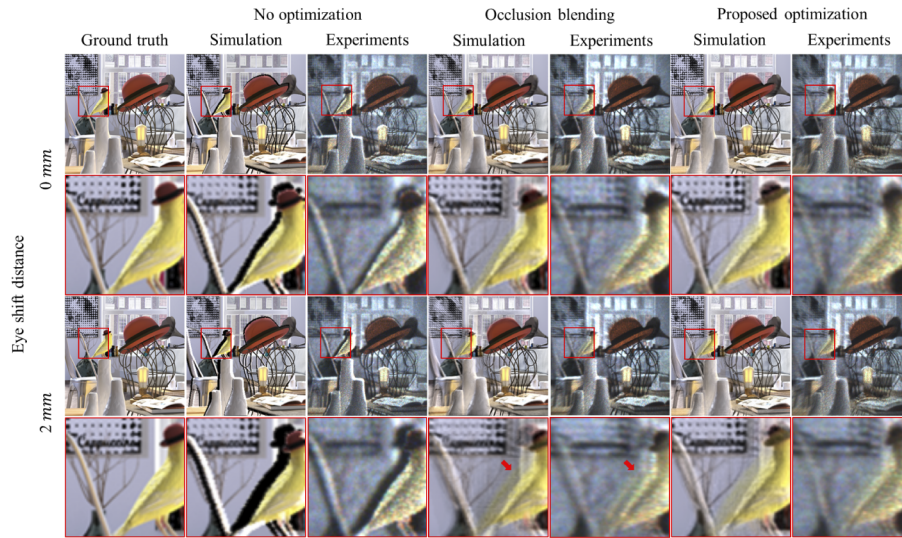
Fig. 7. We verify the feasibility of the optimization by shifting viewpoints. The artifacts due to magnification variation and viewpoint shift are mitigated.
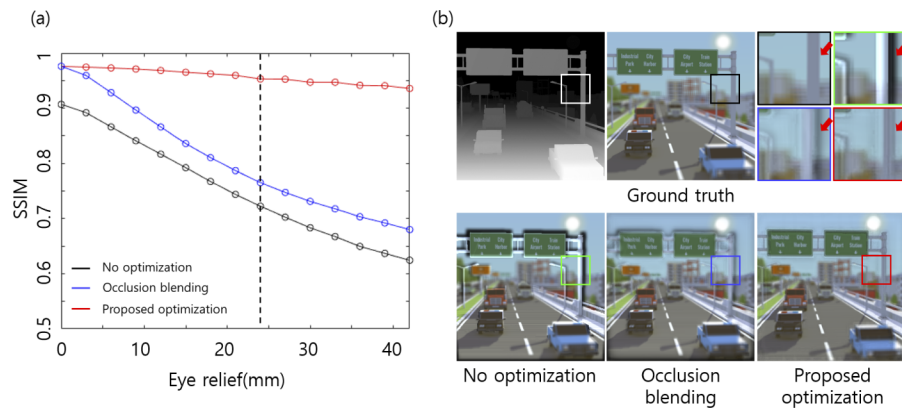

Fig. 8. (a) SSIM values of reconstructed retinal images as a function of eye relief. The black line, blue line and red line indicate SSIM values without optimization, with occlusion blending and proposed optimization, respectively. (b) Retinal images focused on 2.4D with eye relief 24 mm.

5. Discussion

5.1. *Image contrast ratio*

Light field optimization considering eye relief renders objects with a larger finite sampling space. This causes smaller magnification at a closer focal depth plane proportional to the eye relief. Since the magnification according to the depth is not considered in target multi-view images, the objects with significant magnification differences are expected to lose image contrast by optimization. To verify whether any contrast changes occur, we plot the contrast ratio curves of the reconstructed images after various optimization procedures in Fig. 9(a). A flat scene at the specific depth is provided as a volumetric target object for light field synthesis in this simulation. Figure 9(a) illustrates the contrast value relative to the depth of the flat scene in the middle-range of angular frequency, which is 5 cycles per degree (cpd), when the eye relief is 15 mm. The maximum angular frequency that can be attained in this system is limited to 10 cpd. Blue lines and red lines show the contrast ratio (dashed lines) and SSIM (solid lines) of the reconstructed retinal images with occlusion blending and the proposed optimization. The target objects are located from 0.0D to 5.5D with a separation of 0.6D. Retinal images are reconstructed corresponding to the accommodation states. To estimate the contrast ratio, the frequency component of the Fourier transformed retinal images with accommodation depth identical to target depth is normalized by that of the Fourier transformed ground truth. As demonstrated in Fig. 9(a), the proposed optimization has a slight loss of contrast in the near focal depth plane as opposed to the occlusion blending method. In the occlusion blending method, however, the SSIM value of the rendered retinal image decreases noticeably in the near focal planes. In other words, we demonstrate a trade-off in the contrast ratio in the close focal planes but benefit in the overall distortion mitigation with the optimization considering eye relief.
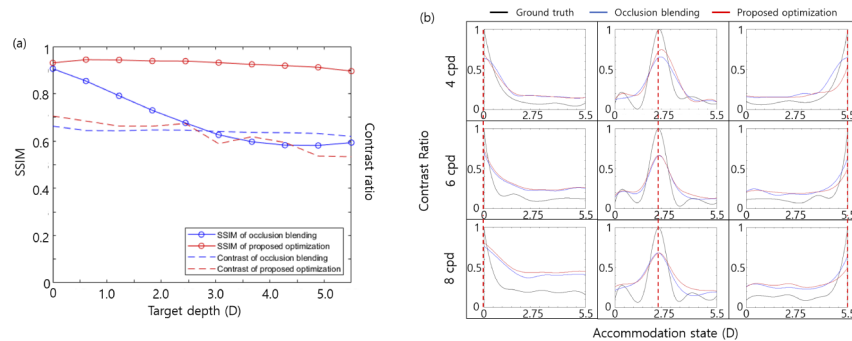

Fig. 9. (a) Contrast ratio (dashed lines) and the averaged SSIM (solid lines) of the retinal images at 5 cpd with 15 mm eye relief. The blue lines and red lines indicate values with the occlusion blending and the proposed optimization, respectively. (b) Contrast ratio of the retinal images as a function of accommodation state with target depth 0D, 2.4D and 5.5D, respectively. The red dashed line indicates the depth where the target plane is placed.

The proposed optimization compensates for eye relief while alleviating the artifacts at occlusion boundaries, but a depth error may occur as the backlight is turned on at several depth planes. Since human is supposed to accommodate in the direction of increasing image contrast to produce a sharper image [11], we consider the contrast change of retinal images as the viewer accommodates at different depths. While a sharp gradient of the contrast ratio successfully drives the accommodation, a contrast of retinal images at the target depth is expected to be maximum. To determine the depth with maximum contrast, we obtained frequency components of signals with 4 to 8 cpd from retinal images according to the accommodation states. As demonstrated in Fig. 9(b), the proposed optimization provides robustness to express correct depth information.

However, we observe the contrast loss when the target is at 5.5D. In other words, this optimization method causes a loss of contrast at a near focal depth plane, but we expect that the human can still accommodate at the target depth.

5.2. Eye relief tolerance

We measured the average SSIM of retinal images acquired by the occlusion blending and the proposed optimization to demonstrate the tolerance for eye relief, as illustrated in Fig. 10(a). A blue line and a red line show averaged SSIM values of entire retinal images optimized with occlusion blending and eye relief compensation at 15 mm, respectively. With the occlusion blending, the averaged SSIM of retinal images is maximum when eye relief is 0 and linearly decreases in proportion to differences in eye relief. On the other hand, with the proposed optimization, the averaged SSIM is maximum at the target eye relief. We determine the tolerance length as the distance at which the SSIM change is 5%. The black dashed lines indicate the eye reliefs where the SSIM is about 95% of the maximum SSIM. As shown in Fig. 10(a), this optimization method has a tolerance length of 14 mm, which is a suitable tolerance range for near-eye displays. Figure 10(b) shows that the reconstructed retinal images within the tolerance length focused on 3.1D. We determine that viewers can observe the optimal volumetric scene within a certain tolerance length around target eye relief.
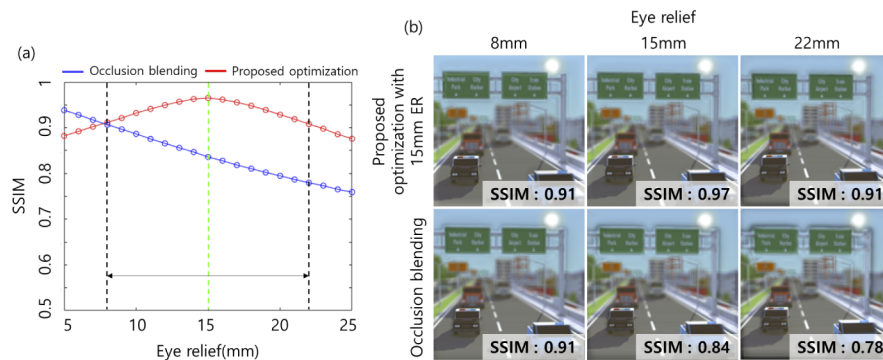

Fig. 10. (a) Average SSIM values relative to eye relief with target eye relief 15 mm (green dashed line). The blue and the red solid lines indicate SSIM of retinal images optimized with occlusion blending and target eye relief 15 mm, respectively. The black dashed lines indicate that a tolerance length of the proposed optimization is 14 mm. (b) Retinal images focused on 3.1D for 15 mm ± 7 mm (8 mm, 15 mm, 22 mm) eye relief.

5.3. Illumination time

In this system, illumination time is bounded in [0, 1]. The DMD has a particular characteristic in brightness since we can modulate illumination time to find optimal images offering higher brightness. In Section 3.1, we found that the optimal illumination time is 0.3 and the SSIM value decreases slowly as illumination time increases from the optimal value. Figure 11 shows the optimal images with illumination time 0.3 (left), 0.6 (middle), and 0.9 (right) obtained at target eye relief 15 mm. The SSIM value changes slowly with increasing illumination time. However, as Fig. 11 illustrates, we noticed that the overall contrast is reduced as the details are blurred. The details are blurred as more pixels are turned on at several depth planes, and focusing blurs affect each other. In other words, the system can provide higher brightness but with inaccurate accommodation cues. Hence, we determine that the novel optimization, considering image similarity and brightness, should be studied further in the future.
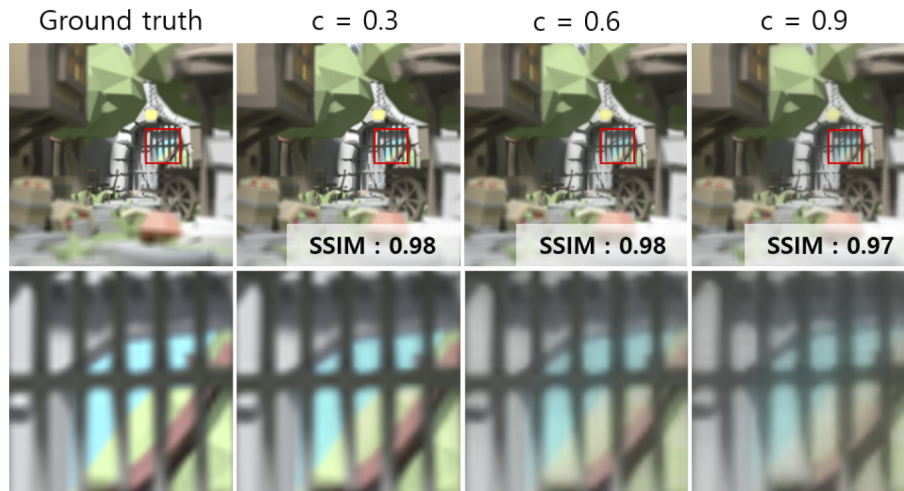
Fig. 11. Reconstructed retinal images with eye relief 15 mm and illumination time 0.3 (left), 0.6 (middle), and 0.9 (right). As illumination time increases, the overall image quality decreases though the SSIM value decreases slowly.

6. Conclusion

We have introduced a magnification compensated optimization method, which alleviates image distortions and depth perception error due to magnification variations in multifocal displays, with light field synthesis. By designing the magnified sampling space in the optimization process, we solved the undesired problem caused by FOV variations without additional optical elements or post-processing. The alleviation of image distortions and depth perception error by magnification variations has been demonstrated through simulations and experiments. Through the magnification compensated optimization, observed image distortions have been mitigated, and the relative size of objects corresponding to the depth is corrected when eye relief exists, while the optimization does not cause critical contrast ratio loss and depth error. In addition, the optimization provides the motion parallax and enlarged eye box size. Since we can optimize with the various target eye relief assuring a tolerance length, we expect this method to be utilized in several near-eye displays. Nevertheless, this optimization has a large computational load. We hope that this problem will inspire new approaches in future work.

Disclosures. The authors declare no conflicts of interest.

Data availability. Data underlying the results presented in this paper are not publicly available at this time but may be obtained from the authors upon reasonable request.

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. **8**(3), 33 (2008).
2. M. Lambooij, M. Fortuin, I. Heynderickx, and W. IJsselsteijn, "Visual discomfort and visual fatigue of stereoscopic displays: A review," J. Imaging Sci. Technol. **53**(3), 030201-1 (2009).
3. G.-A. Koulieris, B. Bui, M. S. Banks, and G. Drettakis, "Accommodation and comfort in head-mounted displays," ACM Trans. Graph. **36**(4), 1–11 (2017).

4. H. Hua and B. Javidi, "A 3D integral imaging optical see-through head-mounted display," Opt. Express **22**(11), 13484–13491 (2014).
5. D. Lanman and D. Luebke, "Near-eye light field displays," ACM Trans. Graph. **32**(6), 1–10 (2013).
6. H.-J. Yeom, H.-J. Kim, S.-B. Kim, H. Zhang, B. Li, Y.-M. Ji, S.-H. Kim, and J.-H. Park, "3D holographic head mounted display using holographic optical elements with astigmatism aberration compensation," Opt. Express **23**(25), 32025–32034 (2015).
7. A. Maimone, A. Georgiou, and J. S. Kollin, "Holographic near-eye displays for virtual and augmented reality," ACM Trans. Graph. **36**(4), 1–16 (2017).
8. N. Padmanaban, R. Konrad, T. Stramer, E. A. Cooper, and G. Wetzstein, "Optimizing virtual reality for all users through gaze-contingent and adaptive focus displays," Proc. Natl. Acad. Sci. **114**(9), 2183–2188 (2017).
9. R. Konrad, E. A. Cooper, and G. Wetzstein, "Novel optical configurations for virtual reality: Evaluating user preference and performance with focus-tunable and monovision near-eye displays," in *Proc. of the ACM conference on Human Factors in Computing Systems*, (2016), pp. 1211–1220.
10. K. Akeley, S. J. Watt, A. R. Girshick, and M. S. Banks, "A stereo display prototype with multiple focal distances," ACM Trans. Graph. **23**(3), 804–813 (2004).
11. K. J. MacKenzie, D. M. Hoffman, and S. J. Watt, "Accommodation to multiple-focal-plane displays: implications for improving stereoscopic displays and for accommodation control," Journal of Vision **10**(8), 22 (2010).
12. O. Mercier, Y. Sulai, K. Mackenzie, M. Zannoli, J. Hillis, D. Nowrouzezahrai, and D. Lanman, "Fast gaze-contingent optimal decompositions for multifocal displays," ACM Trans. Graph. **36**(6), 1–15 (2017).
13. G. Tan, T. Zhan, Y.-H. Lee, J. Xiong, and S.-T. Wu, "Polarization-multiplexed multiplane display," Opt. Lett. **43**(22), 5651–5654 (2018).
14. T. Zhan, J. Zou, M. Lu, E. Chen, and S.-T. Wu, "Wavelength-multiplexed multi-focal-plane seethrough near-eye displays," Opt. Express **27**(20), 27507–27513 (2019).
15. G. D. Love, D. M. Hoffman, P. J. Hands, J. Gao, A. K. Kirby, and M. S. Banks, "High-speed switchable lens enables the development of a volumetric stereoscopic display," Opt. Express **17**(18), 15716–15725 (2009).
16. C.-K. Lee, S. Moon, S. Lee, D. Yoo, J.-Y. Hong, and B. Lee, "Compact three-dimensional head-mounted display system with savart plate," Opt. Express **24**(17), 19531–19544 (2016).
17. T. Zhan, Y.-H. Lee, and S.-T. Wu, "High-resolution additive light field near-eye display by switchable pancharat-nam–berry phase lenses," Opt. Express **26**(4), 4863–4872 (2018).
18. J.-H. R. Chang, B. V. Kumar, and A. C. Sankaranarayanan, "Towards multifocal displays with dense focal stacks," ACM Trans. Graph. **37**(6), 1–13 (2019).
19. N. Matsuda, A. Fix, and D. Lanman, "Focal surface displays," ACM Trans. Graph. **36**(4), 1–14 (2017).
20. T. Zhan, J. Xiong, J. Zou, and S.-T. Wu, "Multifocal displays: review and prospect," PhotoniX **1**(1), 10 (2020).
21. S. Lee, Y. Jo, D. Yoo, J. Cho, D. Lee, and B. Lee, "Tomographic near-eye displays," Nat. Commun. **10**(1), 2497 (2019).
22. F.-C. Huang, D. P. Luebke, and G. Wetzstein, "The light field stereoscope: Immersive computer graphics via factored near-eye light field displays with focus cues," ACM Trans. Graph. **34**(4), 1–12 (2015).
23. D. Lanman, M. Hirsch, Y. Kim, and R. Raskar, "Content-adaptive parallax barriers: optimizing dual-layer 3D displays using low-rank light field factorization," ACM Trans. Graph. **29**(6), 1–10 (2010).
24. G. Wetzstein, D. Lanman, W. Heidrich, and R. Raskar, "Layered 3D: tomographic image synthesis for attenuation-based light field and high dynamic range displays," ACM Trans. Graph. **30**(4), 1–12 (2011).
25. D. Lanman, G. Wetzstein, M. Hirsch, W. Heidrich, and R. Raskar, "Polarization fields: dynamic light field display using multi-layer LCDs," ACM Trans. Graph. **30**(6), 1–10 (2011).
26. G. Wetzstein, D. R. Lanman, M. W. Hirsch, and R. Raskar, "Tensor displays: compressive light field synthesis using multilayer displays with directional backlighting," ACM Trans. Graph. **31**(4), 1–11 (2012).
27. Y. Jo, S. Lee, D. Yoo, S. Choi, D. Kim, and B. Lee, "Tomographic projector: large scale volumetric display with uniform viewing experiences," ACM Trans. Graph. **38**(6), 1–13 (2019).
28. D. Yoo, S. Lee, Y. Jo, J. Cho, S. Choi, and B. Lee, "Volumetric head-mounted display with locally adaptive focal blocks," IEEE Trans. Visualization and Computer Graph., doi: 10.1109/TVCG.2020.3011468 (2020).
29. K. Rathinavel, H. Wang, and H. Fuchs, "Optical calibration and distortion correction for a volumetric augmented reality display," in *Emerging Digital Micromirror Device Based Systems and Applications XII*, vol. 11294 (International Society for Optics and Photonics, 2020), p. 112940M.
30. S. Choi, S. Lee, Y. Jo, D. Yoo, D. Kim, and B. Lee, "Optimal binary representation via non-convex optimization on tomographic displays," Opt. Express **27**(17), 24362–24381 (2019).
31. J. E. Cutting and P. M. Vishton, "Perceiving layout and knowing distances: the integration, relative potency, and contextual use of different information about depth," in *Perception of Space and Motion*, (Elsevier, 1995), pp. 69–117.
32. A. Levin and F. Durand, "Linear view synthesis using a dimensionality gap light field prior," in *2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition* (2010), pp. 1831–1838.
33. D. Kim, S. Lee, S. Moon, J. Cho, Y. Jo, and B. Lee, "Hybrid multi-layer displays providing accommodation cues," Opt. Express **26**(13), 17170–17184 (2018).
34. A. H. Andersen and A. C. Kak, "Simultaneous algebraic reconstruction technique (sart): a superior implementation of the art algorithm," Ultrason. Imaging **6**(1), 81–94 (1984).

35. Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: from error visibility to structural similarity," IEEE Trans. Image Process. **13**(4), 600–612 (2004).
36. H. Huang and H. Hua, "Systematic characterization and optimization of 3D light field displays," Opt. Express **25**(16), 18508–18525 (2017).
37. S. J. Watt, K. J. MacKenzie, and L. Ryan, "Real-world stereoscopic performance in multiple-focal-plane displays: How far apart should the image planes be?" in *Stereoscopic Displays and Applications XXIII*, vol. 8288 (International Society for Optics and Photonics, 2012), p. 82881E.
38. X.-J. Yang, Z.-Q. Wang, and R.-L. Fu, "Hybrid diffractive-refractive 67-diagonal field of view optical see-through head-mounted display," Optik **116**(7), 351–355 (2005).