PHOTONICS Research

Varifocal occlusion in an optical see-through near-eye display with a single phase-only liquid crystal on silicon

Woongseob Han,[1,†] Jae-Won Lee,[2,†] Jung-Yeop Shin,[2] Myeong-Ho Choi,[1] Hak-Rin Kim,[2,3,5] and Jae-Hyeung Park[1,4,6]

[1]Department of Electrical and Computer Engineering, Inha University, 100 Inha-ro, Michuhol-gu, Incheon 22212, Republic of Korea
[2]School of Electronic and Electrical Engineering, Kyungpook National University, 80 Daehak-ro, Buk-gu, Daegu 41566, Republic of Korea
[3]School of Electronics Engineering, Kyungpook National University, 80 Daehak-ro, Buk-gu, Daegu 41566, Republic of Korea
[4]Department of Information and Communication Engineering, Inha University, 100 Inha-ro, Michuhol-gu, Incheon 22212, Republic of Korea
[5]e-mail: rineey@knu.ac.kr
[6]e-mail: jh.park@inha.ac.kr
[†]These authors contributed equally to this work.

Received 20 October 2023; revised 30 December 2023; accepted 1 January 2024; posted 3 January 2024 (Doc. ID 509948); published 1 April 2024

We propose a near-eye display optics system that supports three-dimensional mutual occlusion. By exploiting the polarization-control properties of a phase-only liquid crystal on silicon (LCoS), we achieve real see-through scene masking as well as virtual digital scene imaging using a single LCoS. Dynamic depth control of the real scene mask and virtual digital image is also achieved by using a focus tunable lens (FTL) pair of opposite curvatures. The proposed configuration using a single LCoS and opposite curvature FTL pair enables the self-alignment of the mask and image at an arbitrary depth without distorting the see-through view of the real scene. We verified the feasibility of the proposed optics using two optical benchtop setups: one with two off-the-shelf FTLs for continuous depth control, and the other with a single Pancharatnam–Berry phase-type FTL for the improved form factor.

https://doi.org/10.1364/PRJ.509948

1. INTRODUCTION

Augmented reality (AR) is regarded as the next step in information technology, connecting the real world and the virtual one [1,2]. An AR near-eye display (NED) that presents virtual 3D images on top of the real world attracts a lot of attention not only from experts in related fields but also from the public. Many technical issues of the AR NED optics, including the large form factor [3–5], the small field of view (FoV) [6–9], the limited eyebox [10–14], and the unnatural 3D imaging [15–21] have been studied in industry and academia, and there has been some progress. Studies on the mutual occlusion between virtual digital images and real physical objects, however, are still immature, revealing complicated challenges.

Occlusion (i.e., the obscuring of a rear object by the front one) is the strongest depth cue in all depth ranges [22]. Conventional AR NED configurations simply combine the light from the virtual images and the real objects, not supporting the mutual occlusion between them. While virtual images can appear to be occluded by the real objects using pre-cropped images, the real objects cannot be occluded by the virtual images because the virtual images are always translucent. Furthermore, because the system is unable to control the transmittance of the real objects, presenting semi-transparent virtual objects such as glass becomes challenging. A mutually occluded AR scene, therefore, should support masking of the real scenes with a variable masking ratio. The mutual occlusion can be distinguished by three factors: high visibility, contrast, and clear depth order perception of both real objects and virtual images.

Figure 1 shows simulated AR scenes comprising a virtual image ("merchant") and a real scene ("store"). The merchant is translucently overlaid over the store, showing a conventional AR scene in Fig. 1(a), while it occludes the store in Fig. 1(b). As illustrated in Fig. 1(a), the contrast and the visibility of the virtual image are significantly reduced by the lack of occlusion. Moreover, the depth order relationship between the real store scene and the virtual merchant image is not clear in Fig. 1(a), which confuses users. One could simply increase the luminance of the display panel to address these issues. Considering high outdoor sunlight illumination in the daytime, and the low optical efficiency of AR optics, however, the required luminance level of the display panel is much higher than currently available ones, making this approach less practical [23,24].

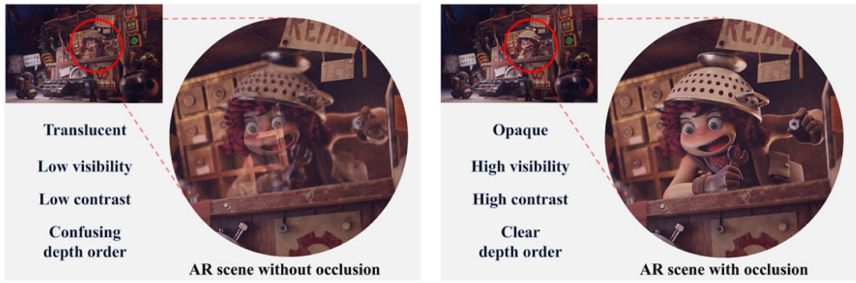

Fig. 1. Simulated AR scene (a) without occlusion and (b) with occlusion by Blender. In the pictures, the merchant and store are regarded as a virtual image and a real scene, respectively. The transparency of the merchant in (a) and (b) was set to 50% and 0%, respectively, to simulate the conventional and occlusion-supported AR scene. Source image: "The Junk Shop" by Alex Treviño. Original concept by Anaïs Maamar.

An occlusion-capable optical see-through NED (OC-OST-NED) is, instead, able to achieve superior contrast, visibility, and depth order perception of the AR scene using the currently available display panels by isolating the virtual image from the real objects, as simulated in Fig. 1(b). Therefore, the OC-OST-NEDs have great promise in terms of presenting realistic AR scenes.

The concept of the OC-OST-NED was pioneered by a few research groups [25,26]. They proposed various occlusion optics that block the light from the real objects behind the virtual images, presenting opaque virtual images to users. These occlusion optics can be classified into two categories according to the axial position of the mask blocking the real scene light. In one category, called "soft-edge masking," the mask is in the NED plane without any imaging optics for the real scene [25,27]. The real scene light is blocked directly by the mask in the NED plane. This approach has advantages in system size and weight since it can be implemented simply by adding a masking spatial light modulator (SLM) to the world-side end of the NED. However, because the depth where the real scene is blocked (= NED plane) is largely different from the depth where the virtual image is formed (> at least tens of centimeters from the NED plane), the mask appears to be significantly blurred for the users focusing on the virtual images [25]. Although commercialized products have started to adopt this approach for its simplicity [28], it is not the ultimate solution for a realistic AR experience. To the contrary, in the other category, called "hard-edge masking," the real scene is first imaged by optics and blocked by the mask. Although the real scene imaging optics increases the system form factor, they can place the optical position of the real scene mask at the virtual image depth, providing realistic masking of the real scene to the users without edge blurring problems [26]. As a prerequisite factor for an ideal OST-NED, various types of optics that support hard-edge occlusion masking have been proposed [29–37].

The hard-edge OC-OST-NED was first demonstrated by Kiyokawa *et al.* [26] using a ring-shaped prototype. In their work, a transmissive liquid crystal display (LCD) was utilized as a pixelated real scene mask. Exploiting a $4f$ system, the optical position of the mask plane was successfully matched to the virtual image plane. However, the optical system with $4f$ optics made the system bulky, hindering its practical usage. To reduce the overall system size, various research has been done that reduces the number of components [29–31] or replaces them with specially designed ones [32–34]. For example, a double-pass folding structure has been widely adopted to reduce the number of lenses and display panels [31,35,36]. Attempts to replace the conventional lenses with microlens arrays (MLAs) have also been proven to be effective in reducing the size of the system by shortening the gap in the $4f$ optics [34]. However, this research mainly focused on the bulkiness of the system. A limitation shared by these approaches is that the optical position of the mask is fixed at a certain distance, and is not supporting varifocal property.

The OC-OST-NED supporting a varifocal hard-edge mask is becoming an emerging topic with the advance of monocular 3D display technology [38–42]. Hamasaki and Itoh [43] proposed varifocal occlusion optics using a physically shifted mask. Although they verified the 3D hard-edge occlusion mask, physically sliding the mask on an electric linear stage makes the overall system prone to issues related to mechanical movement. Rathinavel *et al.* [44] successfully demonstrated varifocal occluding optics using two focus-tunable lenses (FTLs) and display panels without physical movement. Their work, however, was not implemented in the NED configuration, requiring more complex optics for its application to a practical OC-OST-NED. Exploiting the phase-only SLM as a dynamic free-form lens to generate an occlusion mask of continuous depth was proposed by Hiroi *et al.* [45]. Recently, Chae *et al.* [46] suggested a varifocal occlusion optics system using MLAs and FTLs, achieving small form factors. However, their works only showed the varifocal masking without experimental demonstration of the display of the corresponding virtual images.

In this paper, we propose a new optics system for the OC-OST-NED with a varifocal hard-edge mask. The proposed method supports the masking of a real scene and the imaging of the digital virtual image at variable depths using an FTL pair of opposite curvature, a single liquid crystal on silicon (LCoS) and a folding optic. The two FTLs have the opposite focal power with the same magnitude, and they enable varifocal masking and imaging while leaving the depth of the unmasked real scene not affected. The LCoS is used as both a pixelated mask for the real scene and a display for the virtual images by controlling the polarization state of the incident light. The double-pass folding optic contributes to the small form factor by minimizing the number of components.

The proposed optic places both the masking plane of the real scene and the imaging plane of the virtual images at the focal length of the FTL. Considering wide focus tunable range of currently available FTLs, the proposed varifocal optic achieves a satisfactory depth range for OC-OST-NED. To utilize a single phase-only LCoS as a mask and a display, a polarization-control method using a half-wave plate (HWP) and a linear polarizer (LP) is adopted. Since the physical locations of the mask and the image are identical, the resultant optical positions are automatically matched, spatially in a pixel-by-pixel dynamic manner, supporting the hard-edge masking naturally. The number of components required for the implementation is reduced by the polarization-based folding optics. The proposed design is, to our best knowledge, the first proposal of a compact OC-OST-NED configuration that supports varifocal imaging and hard-edge masking to the real world in front of the user without any mechanical movement.

We demonstrate the proposed method using two optical benchtop setups. In the first setup, we place two off-the-shelf liquid-based FTLs outside the folding optics to check the feasibility of our method. The continuous focal power variation of the liquid-based FTLs enables the virtual image display and real scene masking in continuous depth planes. In the second setup, a single Pancharatnam–Berry phase (PBP) type FTL that exhibits opposite focal power to different polarizations is used inside the folding optics. The polarization dependency of the PBP lens replaces two liquid-based FTLs outside the folding optics with a single PBP-type FTL inside the folding optics, improving the form factor of the system. The double-pass configuration of the folding optics also compensates the wavelength dependent focus error of the PBP lens automatically. In the following sections, we explain the principle of the proposed method and present the experimental results of the two benchtop setups.

2. PROPOSED VARIFOCAL OCCLUSION-SUPPORT SCHEME

A. Principle of the Proposed Varifocal Occlusion Optics

Figure 2 depicts a simplified scheme of the proposed varifocal occlusion optics. The proposed optical system is a $4f$ relay with two FTLs, two convex lenses, and a mask. Two FTLs are placed at the front and rear focal planes of the relay and they have variable but opposite focal power, i.e., $+f_v$ and $-f_v$, respectively. The two convex lenses have a fixed focal length f_o, and the mask is located at the Fourier plane of the $4f$ relay. Suppose that point sources are located at the focal length $+f_v$ in front of the first FTL, as shown in Fig. 2(a). The light from the point sources is collimated by the first FTL of $+f_v$, and focused on the mask plane by the convex lens. The physical mask then blocks the light from the selected point sources. After masking, the convex lens and the second FTL of the focal length $-f_v$ form the virtual image of the point sources at f_v distance behind the second FTL. Note that the final virtual image distance from the second FTL (i.e., f_v in this case) is the same as the distance of the original point sources from the first FTL. This is true not only for the point sources at f_v distance but also for the point sources at arbitrary distance z ($\neq f_v$).

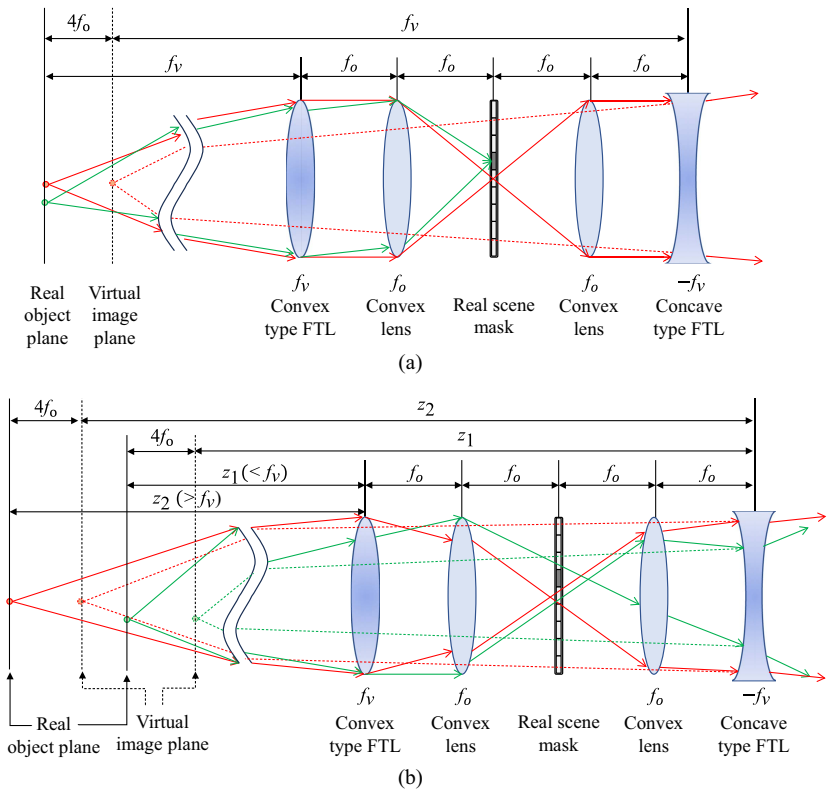

Fig. 2. Simplified optics of the proposed varifocal occlusion when the light source is located at (a) focal length and (b) other distance from the convex type FTL.

Figure 2(b) shows a situation when the point sources are located at an arbitrary distance z ($\neq f_v$) from the first FTL. The ray transfer matrix of the proposed optics is given by

$$L_{-f_v} P_{f_o} L_{f_0} P_{2f_0} L_{f_o} P_{f_o} L_{f_v} = -I, \tag{1}$$

where

$$P_d = \begin{bmatrix} 1 & d \\ 0 & 1 \end{bmatrix}, \qquad L_f = \begin{bmatrix} 1 & 0 \\ -\frac{1}{f} & 1 \end{bmatrix}, \qquad I = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}. \tag{2}$$

Here, P_d, L_f, and I are the propagation, lens, and identity matrix, respectively. The resultant negative identity matrix in Eq. (1) proves that the final virtual image distance from the second FTL is always the same as the original object distance from the first FTL, regardless of the variable focal length $\pm f_v$ of the FTLs if the optical power of two FTLs has the same magnitude but opposite signs. This indicates that the real scene maintains its original depth after passing through the proposed optics, while the objects at the f_v distance are selectively blocked by the mask. By changing the focal length f_v of the FTLs, the depth of the masking plane can be controlled without introducing the depth distortion of the real scene. Note that the negative sign of the identity matrix in Eq. (1) implies that the real scene is inverted after the optics. This inversion is corrected by using additional inverting optics in the actual implementations.

The proposed optic requires two FTLs with the opposite focal power, i.e., $+f_v$ and $-f_v$, respectively. This could be implemented either using two FTLs working in synchronization or using a single FTL exhibiting opposite focal power depending on the polarizations. In the following sections, we explain these two implementations.

B. Varifocal Occlusion Optics Using Two Liquid-Based FTLs

Figure 3 shows the schematic diagram of the two-FTLs based implementation of the proposed OC-OST-NED optics. As described in Fig. 3, a convex-type FTL of $+f_v$ and a concave-type FTL of $-f_v$ are deployed at the world-side and eye-side ends of the system, respectively. A Schmidt–Pechan prism is also employed to compensate for the real scene inversion explained in the previous section. Light from the real scene is s-polarized by an LP, pre-inverted by the prism, and reflected toward the left side of the system by a polarizing beam splitter (PBS). The optics at left, indicated by a dotted rectangle in Fig. 3 are the part where the mutual occlusion and virtual scene imaging are performed. At first, a convex lens located at a focal length distance from the convex-type FTL converges the incident light from the real scene toward an LCoS that corresponds to the real scene mask in Fig. 2. The polarization angle of this light is adjusted by an HWP behind the convex lens such that the LCoS works as a per-pixel polarization modulator. The light reflected from the LCoS with the pixel-wise modulated polarization passes through the PBS again, being the intensity-modulated light. In the case of the virtual image, light emitted from the light source on the top of the left side optics in Fig. 3 is linearly polarized by an LP and reflected by the LCoS with spatially different polarizations. This polarization distribution is also transformed to the intensity distribution by passing through the PBS. Note that the real scene masking and virtual image display are performed by the single LCoS with a proper choice of the LP and the HWP angles. More details on this polarizer angle selection are explained in Section 2.D.

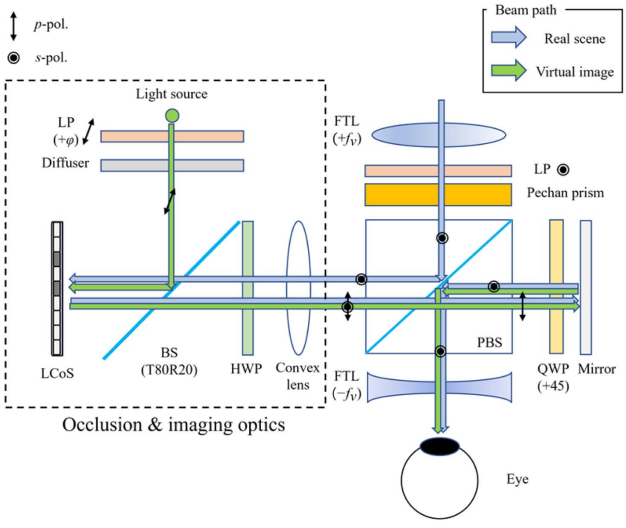

Fig. 3. Schematic diagram of the two-FTLs optics.

The output of the optics on the left (i.e., the part in the dotted rectangle in Fig. 3) is the polarization-modulated light of the real scene and virtual images. It is first transformed to the intensity-modulated light by passing through the PBS. It then propagates toward the eye after being reflected by a mirror with a quarter-wave plate (QWP). A concave type FTL of the focal length $-f_v$ at the eye-side finally restores the depth of the real scene, forming the mask and virtual images at f_v distance.

Figure 4 shows the detailed light trajectory with the polarization states for the masking and nonmasking LCoS pixels in the case of a binary virtual image. When the light is incident on the nonmasking pixel, as shown in Fig. 4(a), the real scene light passes through the PBS and enters the eye pupil of the user. The virtual image light from the light source, however, is reflected by the PBS toward the world side, not reaching the eye, as depicted in Fig. 4(a). On the contrary, for the masking pixel, the real scene light is reflected from the PBS toward the world side while the virtual image light reaches the eye, as shown in Fig. 4(b). Consequently, the proposed system exhibits varifocal masking and display of the real scene and binary virtual images simultaneously. In case of the grayscale virtual images, a time-multiplexing scheme can be adopted, as will be explained in Section 2.D.

C. Varifocal Occlusion Optics Using a Single Polarization-Dependent PBP-Type FTL

Figure 5 depicts our second OC-OST-NED optics configuration, implementable with a single FTL. In this configuration, a single polarization-dependent PBP-type FTL is placed inside the folding optics, instead of the two polarization-independent FTLs outside the folding optics in the first configuration. The PBP-type FTL exhibits positive $+f_v$ and negative $-f_v$ focal power depending on the incident polarization, replacing two FTLs in the first configuration.

We implement the PBP-type FTL with the desired property by using a stack of multiple PBP lenses and switchable HWPs

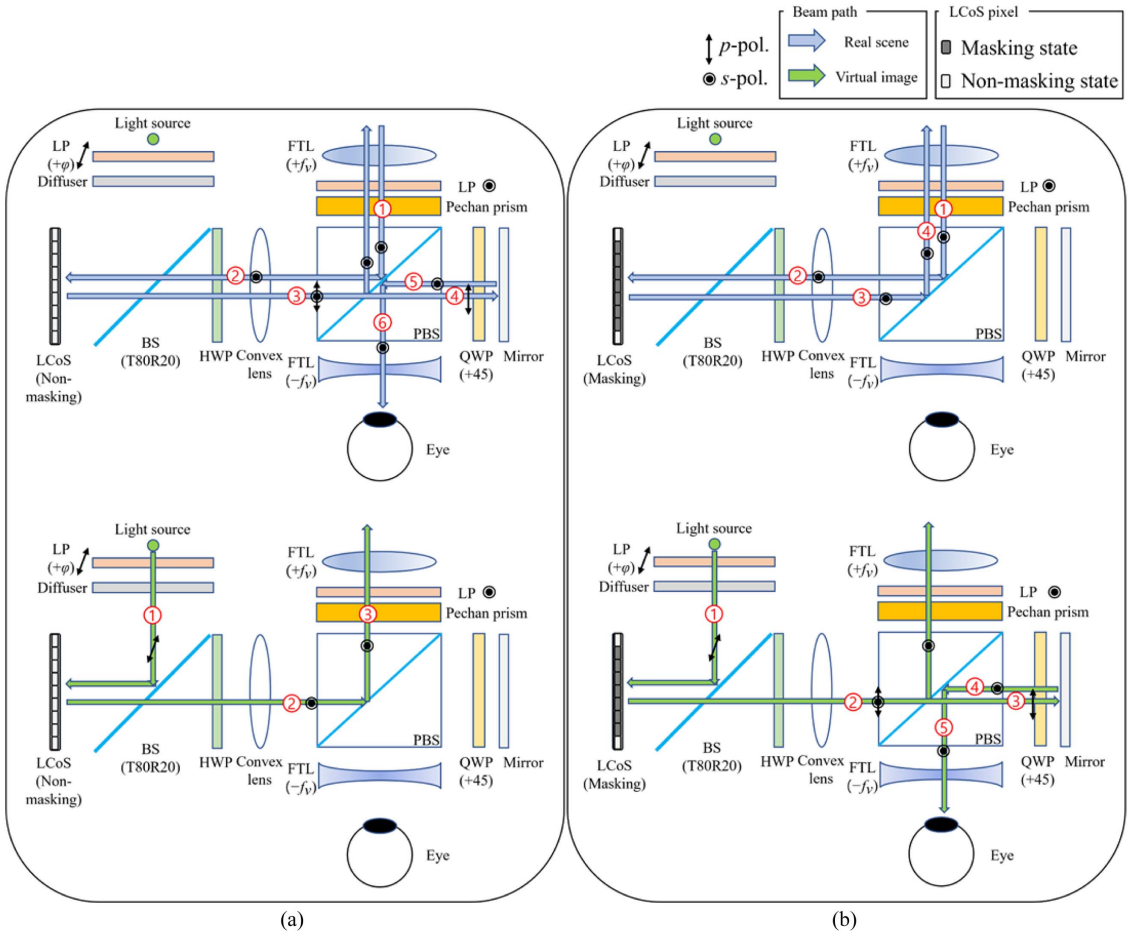

Fig. 4. Schematic diagram of the two liquid-based FTL optics with light path of real objects and binary virtual image reflected on (a) a nonmasking pixel, and (b) a masking pixel of the LCoS. Blue and green lines indicate the light trajectory of real objects and virtual images, respectively.

(SHWPs) [47], with two additional QWPs at the input and output ends of the device. The individual PBP lens shows the focal power of the same magnitude but opposite sign depending on the incident light polarization. By switching the SHWPs, different focal power combinations of the PBP lenses in the stack are selected, realizing the varifocal property. The maximum number of the available focal lengths is given by the square of the number of the PBP lenses in the stack. In the proposed method, however, only half is used to keep the same polarization for the input and output of the FTL. The detailed structure and the polarization changes inside the PBP-type FTL are explained in Appendix A.

D. Polarization-Control Scheme for Mutual Occlusion Optics

In the proposed scheme, a single LCoS dynamically serves for both real scene masking and virtual image display with pixel-by-pixel operation. Figure 6 shows the polarization states of the real scene and the virtual image light in the proposed configurations in Figs. 3 and 5. Note that as the optic is unfolded, a coordinate axis (i.e., the x axis) is reversed after the LCoS reflection in Fig. 6. From Fig. 6 and the derivation explained in Appendix B, the Jones vectors of the output light can be obtained by

$$J_{\text{real.output}} = \begin{bmatrix} J_{rs} \\ J_{rp} \end{bmatrix} = \exp\left(j\frac{2\pi}{\lambda}dn_0\right)\begin{bmatrix} \cos^2 2\theta + e^{j\delta(x,y)}\sin^2 2\theta \\ -\cos 2\theta \ \sin 2\theta(1 - e^{j\delta(x,y)}) \end{bmatrix}, \tag{3}$$

$$J_{\text{virt.output}} = \begin{bmatrix} J_{vs} \\ J_{vp} \end{bmatrix} = \exp\left(j\frac{2\pi}{\lambda}dn_0\right)\begin{bmatrix} \cos 2\theta \cos\varphi + e^{j\delta(x,y)}\sin 2\theta \ \sin\varphi \\ -\sin 2\theta \cos\varphi + e^{j\delta(x,y)}\cos 2\theta \sin\varphi \end{bmatrix}, \tag{4}$$

where,

$$\delta(x,y) = \frac{2\pi}{\lambda}d[n_e(x,y) - n_o]. \tag{5}$$

In Eqs. (3)–(5), $J_{\text{real.output}}$ and $J_{\text{virt.output}}$ are the Jones vectors of the real object and virtual image light just before returning to the PBS, as highlighted by a red rectangle in Fig. 6, respectively. δ is the phase retardation of the LCoS and n_o, n_e are the ordinary and extraordinary refractive indices of the LC. λ, θ, and φ

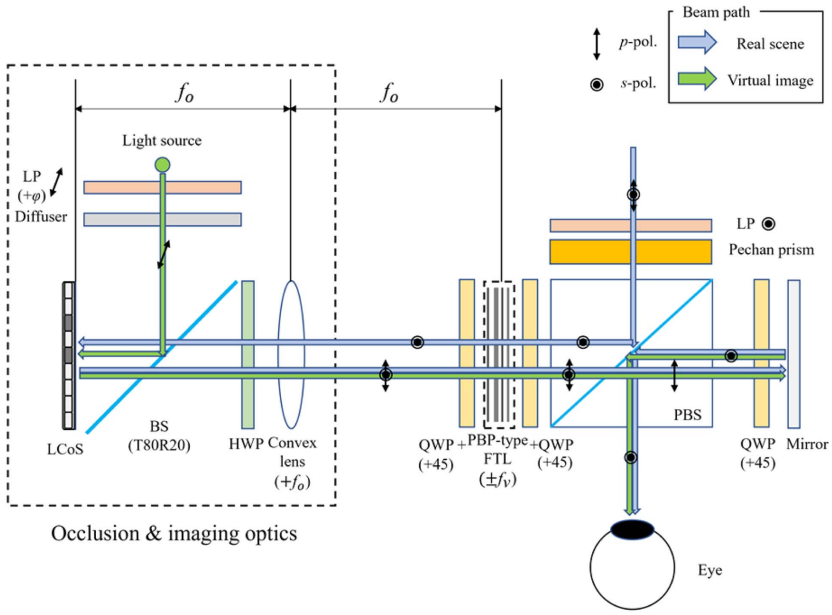

Fig. 5. Schematic diagram of a single FTL optics.

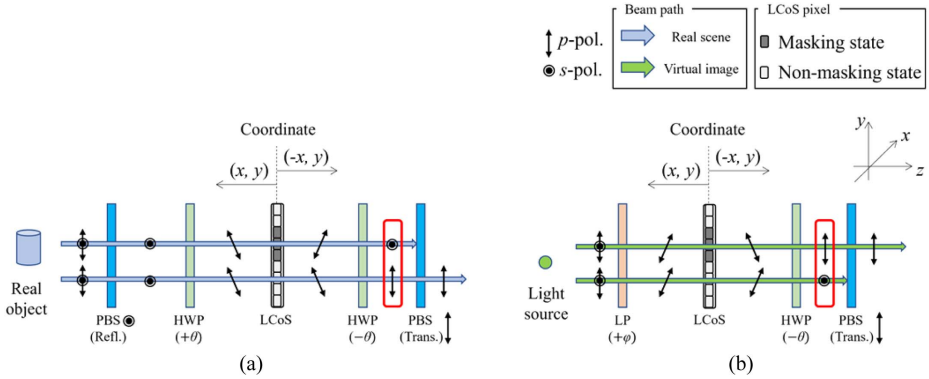

Fig. 6. Polarization optics of the proposed method with light path of (a) the real object and (b) the virtual image. Red bold rectangle indicates output polarization of each case.

are the wavelength of the incident light, the fast axis angle of the HWP, and the transmission axis angle of the LP, respectively. The first and the second row of the $J_{\text{real.output}}$, $J_{\text{virt.output}}$ indicate the s-polarized component and p-polarized component, respectively. Note that the p-polarized component of the $J_{\text{real.output}}$, $J_{\text{virt.output}}$ passes through the PBS and eventually reaches the user's eye, as illustrated in Figs. 4 and 5. Thus, the second row components of the $J_{\text{real.output}}$, $J_{\text{virt.output}}$ in Eqs. (3) and (4) (i.e., J_{rp} and J_{vp}), represent the light observed by the user's eye.

Suppose a simple case where the light is monochromatic, and the virtual image is binary. An LCoS pixel works as a binary switch in this case. In the masking state, the real scene is blocked while the virtual image is observed, i.e., $J_{rp} = 0$ and $J_{vp} \neq 0$. In the nonmasking state, the virtual image is blocked, and the real scene enters the eye, i.e., $J_{rp} \neq 0$ and $J_{vp} = 0$. From Eqs. (3) and (4), it can be easily achieved by setting the HWP and LP angles to be $\varphi = -2\theta$, and controlling the LCoS phase retardation δ to be $\delta = 0$ or 2π for the masking and $\delta = \pi$ for the nonmasking, respectively. In this case, J_{rp} and J_{vp} are given by $J_{rp} = 0$, $J_{vp} = -\sin 4\theta$ in the masking state and $J_{rp} = -\sin 4\theta$ and $J_{vp} = 0$ in the nonmasking state. Therefore, in the binary virtual image case, the desired operation with a 100% occlusion ratio is achieved, except for the case of $\theta = n\pi/4$ (n, integer).

Presenting grayscale virtual images, however, reduces the occlusion ratio of the mask. To display the grayscale images, J_{vp} should be in the range of [0,1], which is achievable by controlling the LCoS phase retardation δ in the range of $[\pi, 2\pi]$. In this condition, the J_{rp} cannot be maintained to be 0, and some of the real scene light is leaked to be visible to the user. Figure 7 shows the visibility of the real scene and virtual image with respect to the variation of δ when $\varphi = -2\theta$. The varying real scene visibility according to the LCoS phase retardation δ in

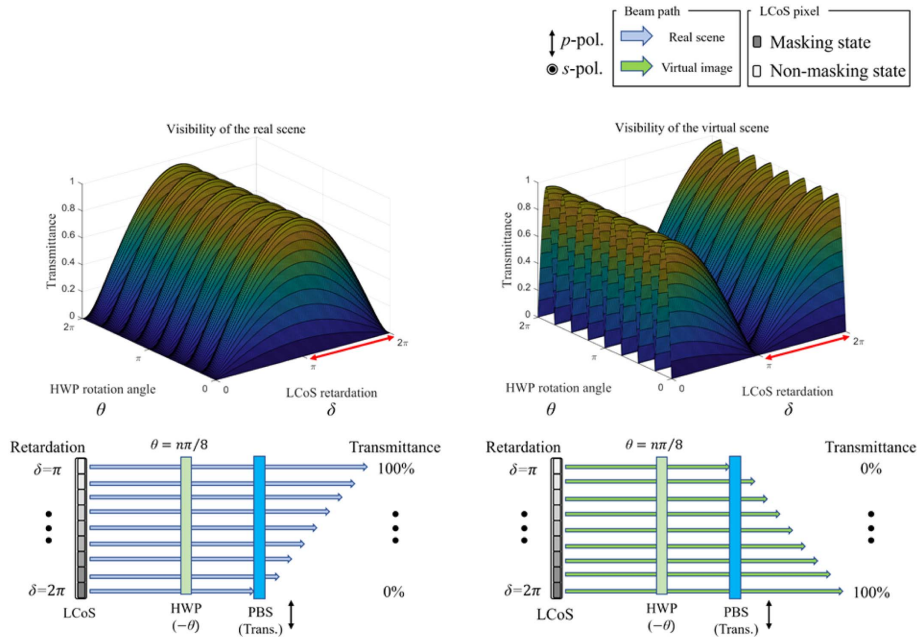

Fig. 7. Visibility of the real scene and virtual scene according to the LCoS retardation and HWP rotation angle when $\varphi = -2\theta$. Bold red double-sided arrows indicate the retardation range used in grayscale imaging.

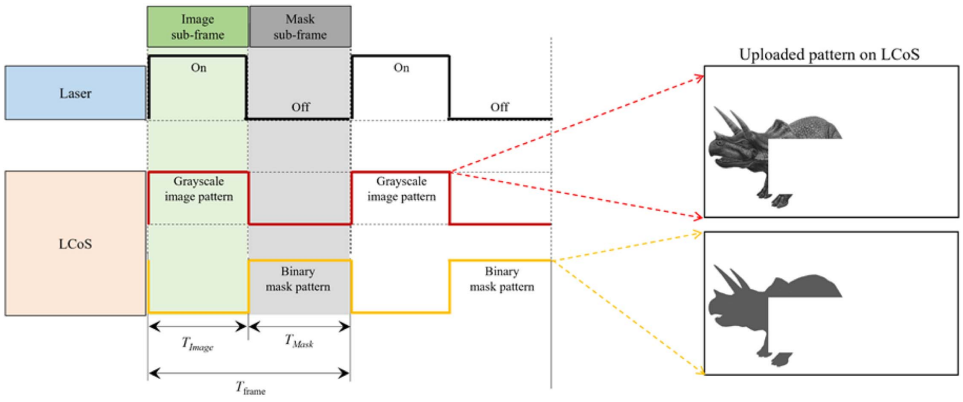

Fig. 8. Operation scheme of the time-multiplexed grayscale imaging.

Fig. 7 implies that the occlusion ratio is reduced and dependent on the grayscale value of the virtual image.

To overcome this issue, we apply a time-multiplexing scheme to achieve grayscale imaging with real scene occlusion, as shown in Fig. 8. We divide a single frame into two subframes (i.e., a mask and an image subframe), and the light source for the virtual image is turned on only during the image subframe. In the mask subframe, the LCoS presents the binary mask pattern, completely blocking the real scene in the desired area. In the image subframe, the LCoS is modulated to present the grayscale virtual image. This time-multiplexing scheme increases the overall occlusion ratio and sustains the image contrast, as will be demonstrated in Section 4.B. In the actual operation, the wavelength dependency of the LCoS phase modulation δ also affects the occlusion performance. Further discussion on the wavelength dependency and the full-color operation is detailed in Appendix B. Also note that the time-multiplexing scheme widens the degree of freedom such that the LP angle φ is no longer limited to -2θ. In the following experiments, we maintain the condition of $\varphi = -2\theta$ for the consistency between the binary imaging and the grayscale imaging. An exploration of the other possible choices of the LP angle φ is detailed in Appendix C.

3. HARDWARE IMPLEMENTATION

A. Verification Setup of Optical System with Two Liquid-Based FTLs

Figure 9 shows a photo of the experimental setup of two FTL optics. Two commercialized liquid-based FTLs (Optotune, EL-16-40-TC) were placed at the start and the end of the system. A Pechan prism with an LP was deployed right behind the

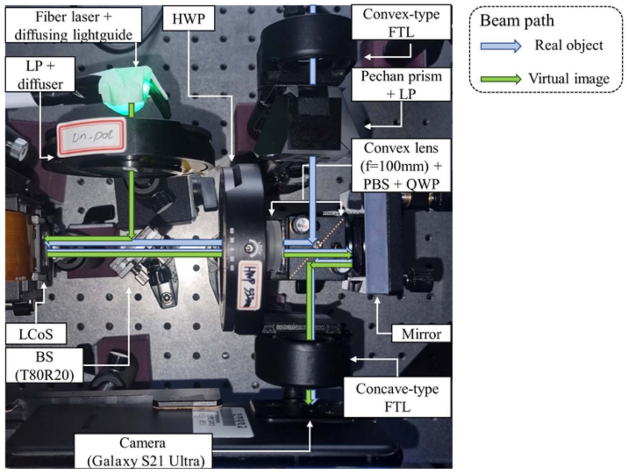

Fig. 9. Picture of the experiment setup of two-FTLs optics.

convex-type FTL to pre-invert the real scene. A convex lens with a focal length of $f_o = 100$ mm, a PBS, and an achromatic QWP (Edmund Optics, #46-558) were arranged on a custom mount that was fabricated using a 3D printer to minimize the air gap between the components. A zeroth-order HWP was placed on the left side of the PBS to rotate the angle of the linear polarization of the input light. A BS with 80% transmittance and 20% reflectance was used to obtain the brighter real scene than a conventional 50/50 BS. A phase-only LCoS (HOLOEYE Photonics, GAEA-2) was used to display the virtual image and to mask the real scene by modulating the polarization state of the incident light. A full-color laser (FISBA, READYBeam™) was employed as a virtual image light source of our setup. A diffusing lightguide and an LP were additionally used to achieve uniform illumination on the entire panel area of the LCoS, and to set the laser polarization, respectively. Finally, a camera (Samsung, Galaxy S21 Ultra) behind the concave-type FTL was placed at the eye position to capture the results.

B. Validation Setup of Optical System with a Single PBP-Type FTL

Figure 10 shows the implementation of our second optics with a single PBP-type FTL. The left part of the setup is the same as that of the first setup shown in Fig. 9. The main changes are in the central part, highlighted by the red rectangle in Fig. 10. In this part, we replaced two commercialized liquid-based FTLs with a single homemade PBP-type FTL consisting of a PBP lens and SHWP stacks. The fabricated PBP-type FTL has two PBP lenses of ±100 cm and ±70 cm focal lengths, two SHWPs made by twisted-nematic (TN) cells, and four QWP films (Edmund Optics, WP140HE, #25-369), providing two sets of positive and negative focal lengths to the system. Here, by using the film-based QWPs, the thickness form factor of the stacked module was considerably improved. The SHWP was operated by applying a voltage to the TN cells using a two-channel function generator (GW Instek, AFG-2225). See Appendix A for the fabrication process of the FTL. The FTL was placed on a custom 3D-printed mount with an achromatic QWP. We also made a PBS mount to put two more achromatic QWPs on both sides of the PBS.

4. EXPERIMENT RESULTS

In this section, we show the results from two experiment setups. Sections 4.A–4.C show the results from two liquid-based FTLs optics to verify the varifocal imaging with a depth-matched mask, grayscale presentation using time-multiplexing, and full-color imaging. The result from the second experiment setup is shown in Section 4.D to demonstrate the feasibility of single PBP-type FTL optics.

A. Varifocal Occlusion

We first examined the masking performance of the proposed optics by using the binary states of the LCoS, as explained in Section 2.D. In this experiment, a real scene target is placed at 50 cm in front of the world-side FTL, and the focal power of two FTLs was set to be ±2D (diopter), presenting the virtual image and the mask on a real target plane. The result in Fig. 11 demonstrates the capability of the simultaneous imaging and masking operations using the single LCoS.

Second, we tested the depth-varying image and mask by changing the focal power of two FTLs. The focal length of two FTLs was set to ±33, ±50, ±100 cm, and optical infinity, such that the image and mask are formed at the corresponding distances, i.e., 33, 50, 100 cm, and optical infinity, respectively. The real scene target was placed at a distance of 50 cm (= 2D), as in the previous experiment in Fig. 11.

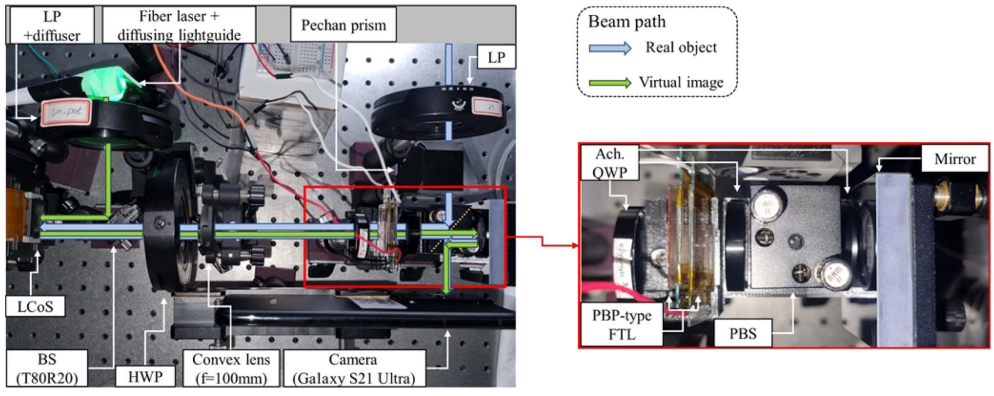

Fig. 10. Experiment setup of the single FTL optics.

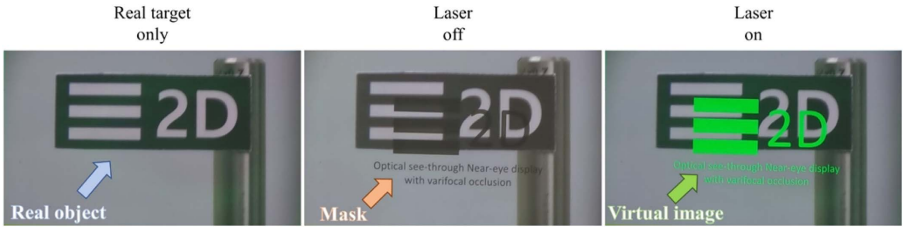

Fig. 11. Captured photos of the occlusion test with/without the illumination from the fiber laser.

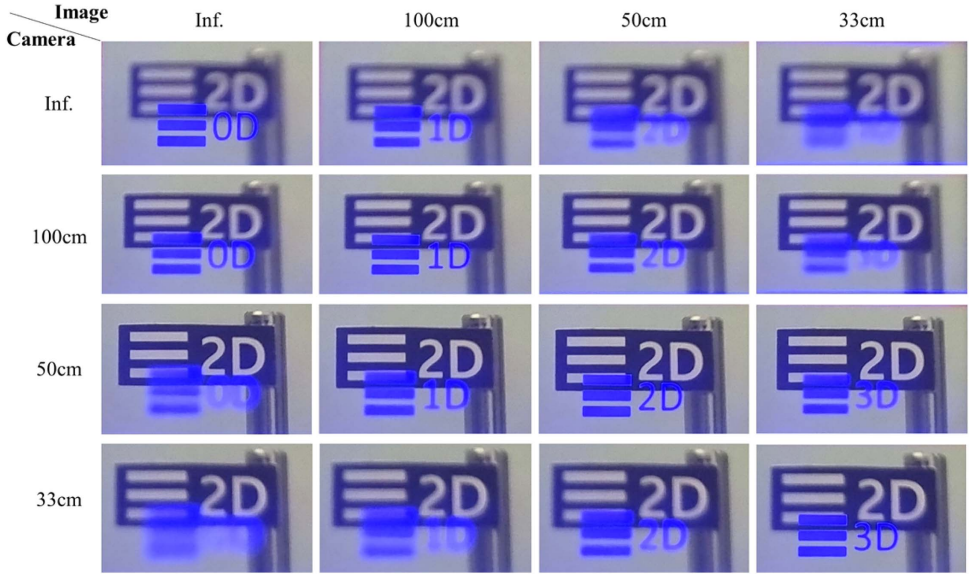

Fig. 12. Captured photos of the real scene at 50 cm (= 2D) while the virtual image is at 33, 50, 100 cm, and optical infinity.

Figure 12 shows the experiment results of the varifocal imaging and occlusion test. In Fig. 12, it can be observed that the virtual image is focused and blurred according to the camera focus at each image distance, which indicates the successful presentation of the varifocal images. Note that when the camera focus is at the real object distance (i.e., a distance of 50 cm, as in the third row in Fig. 12), the real object maintains its focus regardless of the focal power variation of the two FTLs. This demonstrates that the proposed optics changes the virtual image and the mask distance without distorting the real scene depth, as we explained in Section 2.A. Also note that the depth of the virtual scene image and the real scene mask does not need to be selected from a pre-determined discrete set but can be controlled continuously within a wide range, owing to the continuous focal power variation of the liquid-type FTLs used in our implementation.

B. Grayscale Imaging Using Temporal Multiplexing

Figure 13 shows the experiment result of the grayscale imaging with the varifocal hard-edge mask scheme. A virtual dinosaur

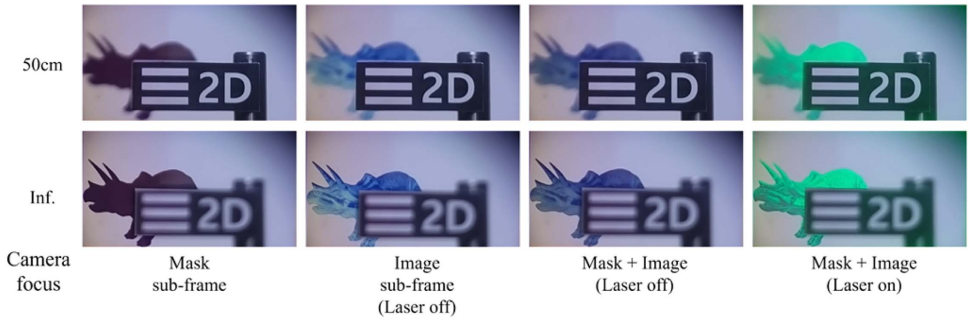

Fig. 13. Time-multiplexed experimental result for grayscale imaging. In the time multiplexing, the mask and image subframes are alternately displayed at 10 Hz. The pictures were taken with a camera setting of ISO 50 and a 1 s shutter speed.

image was displayed at the optical infinity. The real scene consists of a target object at 50 cm and a white background at the optical infinity. The virtual dinosaur image at the optical infinity was pre-cropped to account for its occlusion by the real target object at 50 cm. To represent the dinosaur in grayscale, we use the time-multiplexing scheme explained in Section 2.D. The first and second columns in Fig. 13 show the observed images in the mask subframe and image subframe without a virtual image light source, respectively. In the mask subframe, the pixel values of the LCoS were set to achieve the maximum occlusion. The result in the first column of Fig. 13 demonstrates the proper occlusion of the real white background as expected. In the image subframe, the LCoS displays the virtual dinosaur image. As the pixel values of the LCoS pattern in this subframe are determined to present the grayscale virtual image, the occlusion of the real scene becomes nonuniform, being slightly degraded. However, by temporal multiplexing of the mask and the image subframe, the overall occlusion still maintains a satisfactory level, as shown in the third column of Fig. 13. Finally, by turning on the virtual image light source in the image subframe, the grayscale virtual dinosaur was displayed with high quality, as in the last column in Fig. 13.

The experiment results of the grayscale imaging with more complex real scenes at different depths are presented in Fig. 14. Two real targets, a yellow fish and a background ocean with an orange fish, were placed at 25 cm and 50 cm in front of the system, respectively. A virtual image (i.e., a green fish) was displayed at a distance of 33 cm. As before, the virtual green fish image at 33 cm was pre-cropped to mimic its occlusion by the real yellow fish at 25 cm. The result in Fig. 14 reveals that the virtual green fish image occludes the real orange fish and the ocean behind it, enabling clear depth perception between the virtual image and the real targets.

Figure 15 highlights the effect of the occlusion for an AR scene that consists of a real background ("beach") and a virtual image ("dinosaur"). The result in Fig. 15 clearly shows that the occlusion-supported image of the proposed system effectively portrays a realistic AR scene, while the conventional simply merged image without the occlusion shows low visibility and contrast. Note that the luminance of the virtual image was set to be the same in both cases, and the enhanced visibility and contrast in Fig. 15 are solely because of the occlusion enabled by the proposed method.

C. Full-Color Grayscale Imaging

Figure 16 shows the full-color imaging result. The proposed method achieves the masking and display by giving pixel-wise phase retardation to the incident light. Since the amount of the phase retardation for a given LCoS pixel voltage is dependent on the wavelength of the incident light, the display pattern of the LCoS was optimized for R, G, and B color separately for full-color imaging.

The result is shown in Fig. 16. The real target ("satellite") and background ("space") were placed at a distance of 30 cm

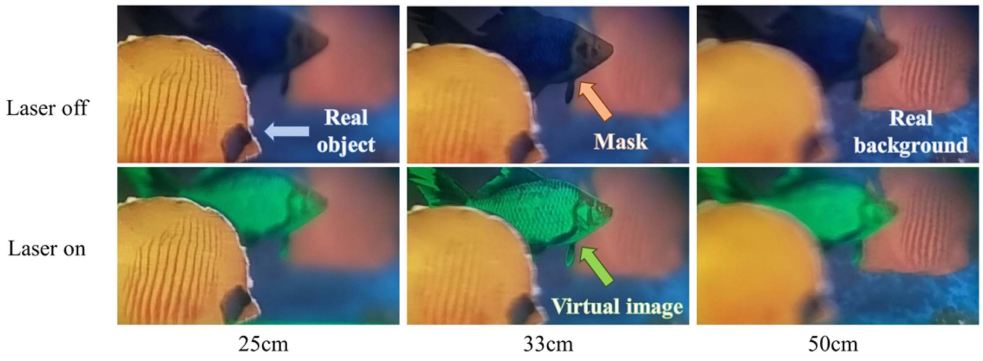

Fig. 14. Experiment result of the grayscale imaging with the time-multiplexing scheme. The real target of the yellow fish and the background is located at 25 cm and 50 cm while the virtual images are at 33 cm.

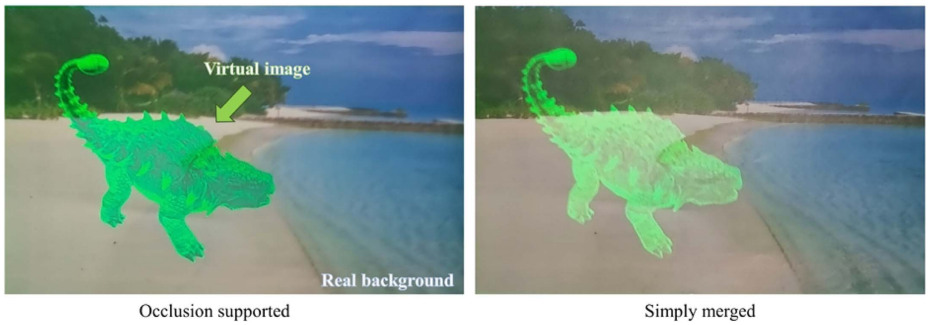

Fig. 15. Comparison of the AR scene with occlusion and without occlusion. Both pictures were taken with the same camera setting of ISO 50 and a 1 s shutter speed.

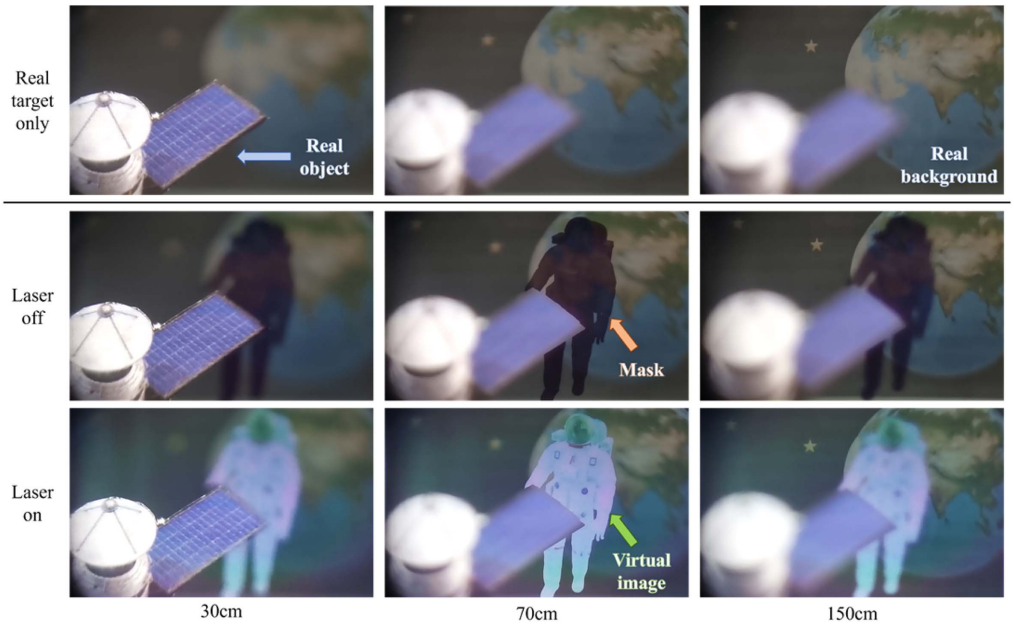

Fig. 16. Experiment results of the full-color imaging. The real target of the satellite and background is placed at 30 cm and 150 cm while the virtual image of the spaceman is at 70 cm. Camera setting: ISO 50 and 2 s shutter speed.

and 1.5 m, respectively, as presented in the first row of Fig. 16. The virtual image ("spaceman") with the corresponding mask was displayed at 70 cm, between the "satellite" and "space." The virtual "spaceman" image at the middle-depth plane was pre-cropped again, considering the "satellite" real object in the near-depth plane. As evident from Fig. 16, the full-color virtual image display with adequate occlusion of the real scene can be obtained by the proposed method. Note that for full-color image presentation, the time-multiplexing scheme in Section 2.D was employed to display grayscale images of the RGB color channels sequentially. The camera captured the full sequence of the RGB grayscale image subframes and mask subframes in a single shot with a setting of ISO 50 and a 2 s shutter speed. Also note that the RGB image subframes were obtained only by adjusting the voltage levels applied to LCoS pixels, leaving the HWP and the LP angles unchanged. Therefore, the real-time full-color imaging with a varifocal hard-edge mask would be enabled by utilizing the phase-only LCoS with a fast refresh rate.

D. Experiment Results Using the Optical System with a Single PBP-Type FTL

The demonstrations in the previous section can also be realized by using only a single PBP-type FTL module instead of two liquid-type FTLs. We placed a real target at 50 cm and captured the result using an optical benchtop setup in Fig. 10. As mentioned in Section 3.B, our homemade PBP-type FTL has two PBP lenses with focal lengths of ± 1 m and ± 70 cm, which would result in two focal lengths of 0.41 m and 2.38 m if all elements are cemented without a gap. In the implementation of the device, the air gap between the PBP lenses shifts the focal lengths slightly, achieving 0.45 m and 2.3 m in experiment.

Figure 17 shows the experimental results. The virtual image with a hard-edge mask was displayed at varying depths successfully using only the single PBP-type FTL, yielding results equivalent to the previous two-FTLs setup. Note that the depth and size of the real target also remained still regardless of the focal power of the PBP-type FTL, proving the distortion-free see-through view of the proposed configuration. Also note that the color aberration of the real scene caused by the wavelength-dependent focal power of a usual PBP lens-based optics is not present in the proposed optics because the color aberration in the forward pass through the PBP lens in the FTL is automatically compensated in the backward pass of the same PBP lens in the proposed optics.

Figure 18 shows the experiment results of displaying the grayscale image. As before, the time-multiplexing technique was applied to represent the grayscale of the image while preserving reasonable occlusion. The captured photos in Fig. 18 show the AR scene with a grayscale virtual image (bird) that occludes the real background (beach). These results confirm the proper occlusion of the real background and high-quality display of the virtual image similar to the results from two-FTL-based implementation, proving the validity of the proposed method with a single FTL.

5. DISCUSSION

Varifocal occlusion-support AR image presentation of the proposed single LCoS based OC-OST-NED system was successfully demonstrated using both a liquid-type FTL pair configuration and a single PBP-type FTL configuration. We believe that the current implementation can be further improved.

In the presented implementation, the refresh rate of the grayscale imaging was 10 Hz, which is lower than 60 Hz of standard displays. The low-speed imaging stems from the limited operation speed of the LCoS in the Python-based driving

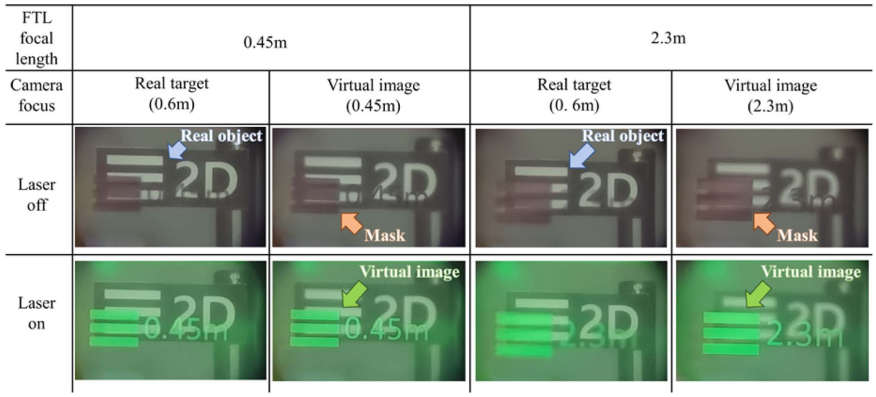

Fig. 17. Feasibility test of varifocal occlusion using a single PBP-type FTL.

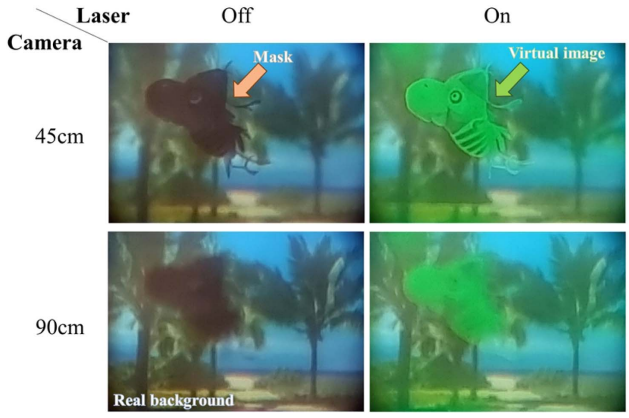

Fig. 18. Grayscale image presentation using a single PBP-type FTL setup (Visualization 1).

mode that was used to synchronize the LCoS and the laser light source. We believe that grayscale imaging with a higher refresh rate can be achieved by adopting faster LCoS or implementing a more sophisticated synchronization method.

The FoV of our implementation is approximately 2.45° in the diagonal direction. Three factors limit the FoV of our setup: the small ray angle tolerance of the Pechan prism, the small size of the LCoS panel, and the long focal length of the convex lens. The Pechan prism can be removed by adding a convex lens between the PBS and mirror, forming another $4f$ optics to invert the real scene. The focal length of the convex lens on the left part of the optics can be shortened by substituting the 80/20 BS with a light source to a waveguide-based backlight [41]. By adopting the waveguide-type backlight and a shorter-focal-length convex lens, the distance between the LCoS and the convex lens can be significantly reduced, achieving a wider FoV. For instance, suppose that a 25 mm aperture optic without the Pechan prism and a convex lens of 30 mm are implemented with a 1 mm thickness waveguide-type backlight. We estimate that the maximum diagonal FoV of 33.01° can be achieved with the current 0.7" LCoS panel by shortening the gap between the convex lens and the LCoS from 100 mm to 30 mm. To expand the FoV further, enlarging the LCoS panel size is required. We expect that the size of the LCoS will be increased considering the rapid advance of the manufacturing technology [23], potentially allowing further expansion of the FoV in the proposed method.

The dimensions of each system are approximately 15 cm × 3 cm × 15 cm (width × height × depth) in the two-FTL setup and 23 cm × 3 cm × 8 cm in the single FTL setup, provided that the circuit board size of the LCoS is disregarded. A more compact system can be achieved by shortening the long focal length of the convex lens. Assuming that the waveguide-based backlight is adopted, and the Pechan prism is replaced by another $4f$ system as explained above, the convex lens of 30 mm focal length can be adopted. With these assumptions, the size of both systems can be optimized to 10 cm × 3 cm × 10 cm, and 13 cm × 3 cm × 3 cm, respectively. Note that the optimized dimension was calculated using the 25 mm aperture optic. If the lens diameter and the PBS size are reduced (e.g., 10 mm), the system form factor can be reduced further with the sacrifice of the FoV. Replacing the PBS combiner with a more compact combiner can be also considered to reduce the system size and enhance the integration with commercialized devices. Given that the commercialized AR NEDs mainly use compact waveguide combiners, our method could be integrated into contemporary AR NEDs by substituting the PBS with the waveguide combiner equipped with achromatic couplers. However, the implementation with a minimized system form factor is out of the scope of this paper and we will leave it for future work.

In the current implementation, the available depth planes of the virtual images in the single PBP-type FTL setup are limited to two, insufficient for exhibiting an immersive 3D experience. Assuming additional stacks of the PBP lens and TN cell are incorporated into the current PBP-type FTL, the number of the depth planes would increase, enabling the expression of a 3D scene with higher depth resolution. In this paper, we focus on confirming the principle and feasibility of the implementation using the PBP-type FTL. We will leave the depth range extension of the single PBP-type FTL setup as a topic for future work.

The full-color image representation was only demonstrated using the two-FTL setup. In the single FTL implementation, the fabricated PBP lens is optimized to green color, evoking

chromatic aberration, and hindering the full-color imaging. Although the chromatic aberration of the PBP-type FTL in a real scene path is self-compensated by the double-pass folding structure as explained in Appendix A, the chromatic aberration of the virtual images that experience only a single pass is not compensated. Given that the recent work by Luo *et al.* [48] experimentally demonstrated an achromatic PBP lens, however, we believe an achromatic PBP-type FTL will be developed soon and can be incorporated into our work.

In the case of the two-FTLs implementation, the depth of the real scene is slightly shifted in the axial direction by the amount of the gap between two FTLs. However, this discrepancy is small because the folding optics adopted in our configuration lessens a physical gap between the convex type FTL and the concave type FTL. We expect the amount of axial shift could be smaller than 5 cm by detaching the optical components from the mount and reducing the air gap between them. Note that the single PBP-type FTL implementation shows no axial shift from the varifocal optics since the physical gap between two FTLs is zero.

Finally, compared to the varifocal occlusion results with the two-FTLs based OC-OST-NED, image degradation with blurring is observed (Figs. 17 and 18) in the single PBP-type FTL scheme, which needs to be improved. Unlike the two-FTLs optics operated by the polarization-insensitive focus-tuning liquid lens units (Figs. 3 and 9), the PBP-based single FTL scheme can provide a much-improved form factor by using the polarization-dependent bifocal switching and polarization-changing abilities of the PBP lens. But precise control on each polarization-changing step between the polarization-dependent units is essential for better pixel-wise control of the varifocal self-aligned mutual occlusion effects, as explained in detail in Appendix A. In the implementation of the compact bi-stack module for the single FTL optics, four layers of the film-based achromatic QWPs were additionally used, compared to the two-FTLs optics, which can be ascribed to the image degradation caused by the generation of multiple surface-reflective stray rays and distortions in the polarization state. The bulk optics of QWP used in both experiment schemes exhibits superior anti-reflection characteristics of less than 1% in surface reflection. In contrast, the employed film-type QWPs have surface reflection properties at approximately 10% while generating haze. Furthermore, the film-type achromatic QWP exhibits relatively high chromatic dispersion characteristics, leading to polarization control distortions deviating from the ideal QWP behavior. However, we expect that a significant improvement in image qualities in both real and AR scenes is achievable by adopting superior achromatic film-type QWPs with anti-reflection coatings over the whole visible range.

6. CONCLUSION

In this paper, we proposed an OC-OST-NED that supports virtual scene imaging and real scene masking at variable depths by using a single phase-only LCoS and varifocal optics. The opposite curvature FTL pair of the proposed varifocal occlusion optics enables the optical relay of the real scene without depth aberration and the display of the virtual image with the corresponding mask in the wide depth range required for the NED applications. The single LCoS of the proposed optics works both for the virtual image display and real scene masking, enabling self-alignment of the image and mask. The double-pass folding configuration reduces the form factor of the proposed system. The feasibility of the proposed system was verified using two optical implementations: one with two commercialized liquid-type FTLs outside the folding optics and the other one with a single homemade PBP-type FTL inside the folding optics. Both implementations demonstrated that grayscale virtual images can be displayed in a wide depth range with proper occlusion of the real scene.

APPENDIX A: FABRICATION, STRUCTURE, AND POLARIZATION VARIATION OF PBP-TYPE FTLs

The fabrication process of the PBP lens used in this study is shown in Fig. 19(a). First, the photo-alignment layer solution is coated on a cleaned glass substrate through spin coating, and the solvent of the photo-alignment layer solution is evaporated through heat treatment at about 120°C inside a glove box in a N_2 environment to form a photo-alignment layer. Then, using two beams of circularly polarized lights orthogonal to each other, a holographic polarization interference pattern is irradiated onto the photo-alignment layer, where the linearly polarized PBP profiles are recorded with the lens-modulated objective beam. A UV-curable reactive mesogen (RM) solution

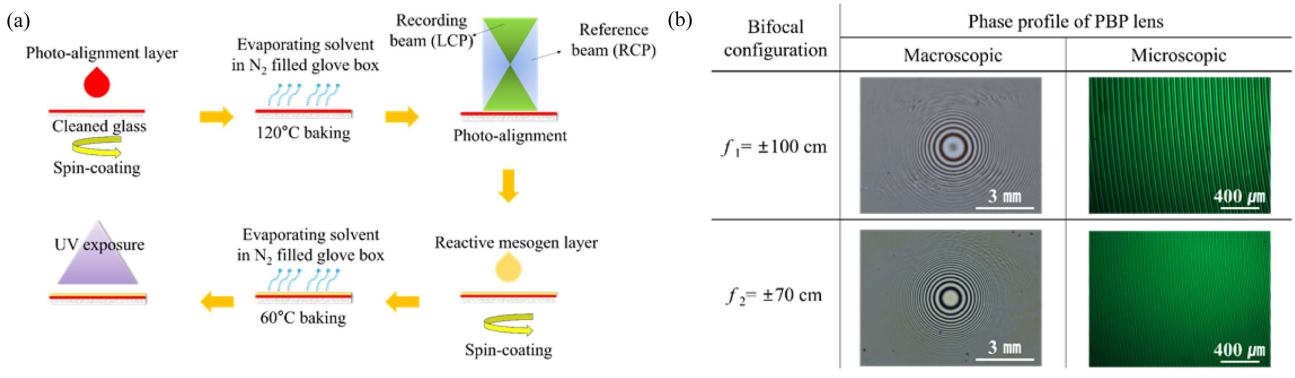

Fig. 19. (a) Fabrication process of the PBP lens and (b) phase profiles of the PBP lens with the polarization-dependent focal lengths of ±100 cm and ±70 cm, taken through the macroscopic CCD image sensor and polarizing optical microscope, where the PBP lenses are placed between the crossed polarizers.

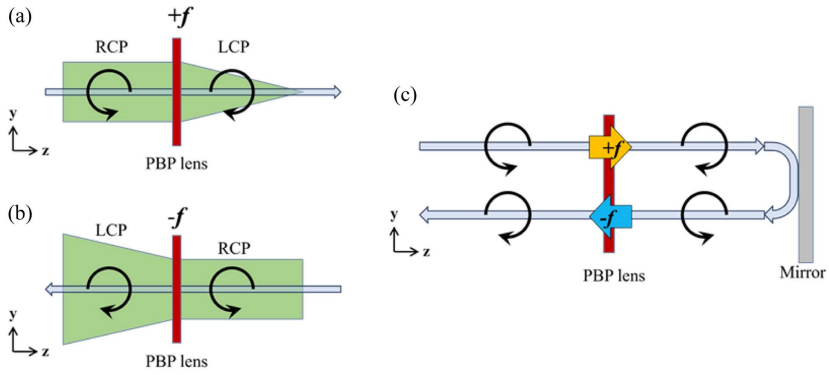

Fig. 20. (a) and (b) The polarization-dependent focusing ($+f$) and defocusing properties of the PBP lens and their polarization conversion states according to the light incident directions. (c) Simplified analysis of the polarization states and magnification cancellation effects after the round-trip light propagation through the single PBP lens by the mirror reflection.

is spin-coated on the PBP-recorded photo-alignment layer and the RM molecules are aligned to match the PBP profile information recorded on the photo-alignment layer. After the RM coating, a heat treatment is performed at about 60°C to remove the solvent remaining within the RM layer. Finally, a UV is irradiated to polymerize the RM layer. We fabricated the PBP lens with the two focal length conditions of ±70 cm and ±100 cm. Figure 19(b) shows the radially symmetric RM textures, which are taken through the macroscopic CCD image sensor and polarizing optical microscope between the crossed polarizers. The PBP lens with a shorter focal length shows a larger PBP modulation change in the radial phase profile.

The PBP lens operates in different lens modes depending on both the circular polarization state and the traveling direction of the light incident on the PBP lens. For example, as shown in Fig. 20(a), when the right-circularly polarized (RCP) light traveling in the $+z$ direction passes through the PBP lens, the PBP lens operates as a convex lens with a focal length of $+f$ and the output polarization is changed into the left-circularly polarized (LCP) state because the PBP lens is made with the HWP retarder condition. When the LCP light is incident in the $+z$ direction, the PBP lens operates as the concave lens with a focal length of $-f$. However, when the RCP light is incident in the $-z$ direction at the same PBP lens, the PBP lens makes phase modulation as the concave lens, as shown in Fig. 20(b), because the incident beam experiences conjugated PBP profiles by the RM orientations with the mirror inversion. Figure 20(c) shows the polarization change sequences and PBP-modulated focusing and subsequent defocusing effects in the folded optics by the mirror. This allows real scenes to be delivered to a user without magnification distortion, irrespective of the focal length condition of the PBP lens tuned for the AR depth-imaging. By utilizing the polarization-dependent PBP modulation and polarization conversion effects, the PBP lens can provide the self-aligned occlusion functionality in a single optic module by employing the LCoS as the pixelated reflecting plane in the proposed folded optics.

Figure 21 describes the detailed optical components that make up the PBP-type FTL module shown in Figs. 5 and 10. As illustrated in Fig. 21, the polarization states of light change as they pass through each component for the self-aligned, depth-switchable occlusion function in the AR imaging. The PBP-type FTL module consists of the PBP lens 1 with a focal length of ±100 cm, PBP lens 2 with a focal length of ±70 cm, two SHWPs, and four achromatic QWP passive films. According to the SHWP operation conditions, the polarization states along the beam pass for the unmasked real scene and the virtual image with occlusion are shown in Fig. 21 in detail. Before the final PBS step, the real scene area to be masked is modulated into the s polarization state that cannot be delivered to a user's eyes. Both the unmasked real scenes and the virtual images are modulated into the p-polarization state, however, allowing pass-through to the eyes after the PBS image combiner.

APPENDIX B: ANALYSIS OF POLARIZATION-BASED OCCLUSION USING THE JONES MATRIX

Figure 22 shows the unfolded optics of the proposed polarization-based occlusion. The Jones matrices of the components and the Jones vectors of the light in Fig. 22 are given by

$$J_{\mathrm{LP}}^{(\theta)} = \begin{bmatrix} \cos\theta \\ \sin\theta \end{bmatrix}, \quad T_{\mathrm{LP}}^{(\theta)} = \begin{bmatrix} \cos^2\theta & \cos\theta\sin\theta \\ \cos\theta\sin\theta & \sin^2\theta \end{bmatrix},$$
$$T_{\mathrm{HWP}}^{(\theta)} = \begin{bmatrix} \cos 2\theta & \sin 2\theta \\ \sin 2\theta & -\cos 2\theta \end{bmatrix}, \quad T_{\mathrm{Mirror}} = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix},$$
$$T_{\mathrm{LC}} = \begin{bmatrix} \exp\left(j\frac{2\pi}{\lambda}n_0 d\right) & 0 \\ 0 & \exp\left(j\frac{2\pi}{\lambda}n_e(x,y)d\right) \end{bmatrix}$$
$$= \exp\left(j\frac{2\pi}{\lambda}n_0 d\right)\begin{bmatrix} 1 & 0 \\ 0 & e^{j\delta'(x,y)} \end{bmatrix},$$
$$T_{\mathrm{LCoS}}^{\delta} = T_{\mathrm{LC}}T_{\mathrm{Mirror}}T_{\mathrm{LC}} = \begin{bmatrix} 1 & 0 \\ 0 & e^{j\delta'(x,y)} \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$$
$$\times \begin{bmatrix} 1 & 0 \\ 0 & e^{j\delta'(x,y)} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & -e^{j\delta(x,y)} \end{bmatrix}, \tag{B1}$$

where $J_{\mathrm{LP}}^{(\theta)}$ is the Jones vector of the linearly polarized light with a rotation angle of θ to the x axis and $T_{\mathrm{LP}}^{(\theta)}$, $T_{\mathrm{HWP}}^{(\theta)}$, and T_{Mirror}

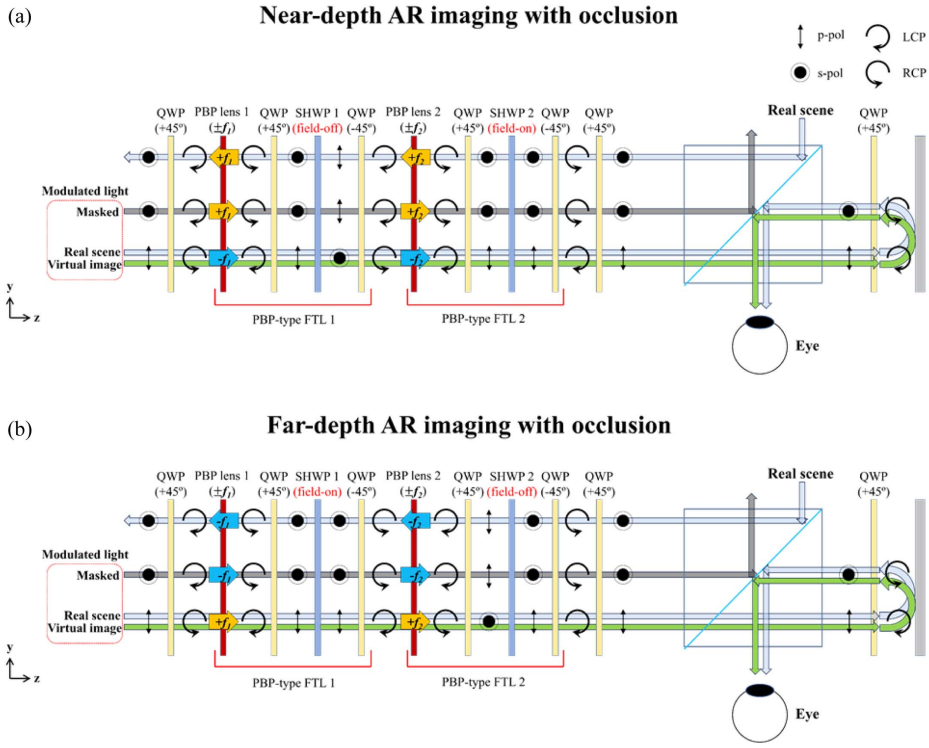

Fig. 21. Polarization change schemes of real scenes and virtual AR images for varifocal self-aligned mutual occlusion effects constructed by the PBP-type FTL module, where two PBP-type FTLs are cascaded and their incident polarization states are controlled by the LC-based SHWP1 and SHWP2 cells for (a) the near-depth and (b) far-depth AR imaging switching with pixelized occlusions.

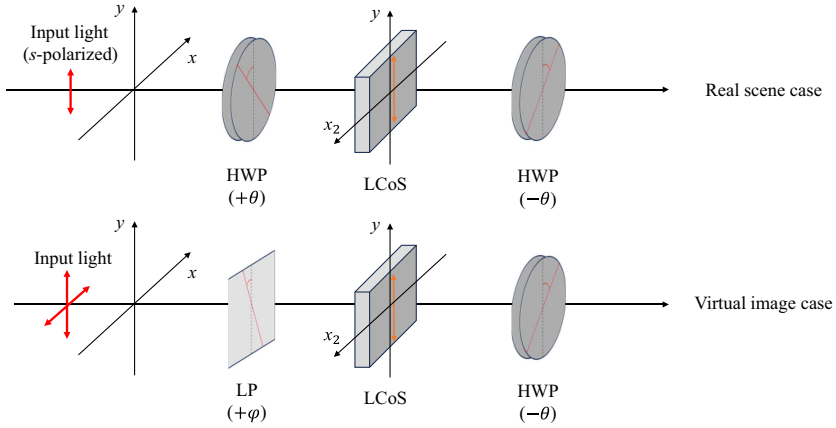

Fig. 22. Unfolded optics of polarization-based occlusion. Bold orange arrow in the LCoS indicates the direction of the LCs aligned in parallel and an x coordinate is reversed after reflection in the reflective-type LCoS.

are the Jones matrices of the LP, achromatic HWP with rotation angle θ, and the mirror, respectively. T_{LC} is the Jones matrix of the LC layer operated in a parallel aligned nematic mode, where $\delta'(x,y) = 2\pi[n_e(x,y) - n_o]d/\lambda$ is the phase retardation in each pixel. d is thickness of LC layer in the reflective-type LCoS, λ is the wavelength of the incident light, and n_o, n_e are the ordinary and extraordinary refractive indices of the LC. T_{LCoS} is the Jones matrix of the reflective type LCoS, consisting of the sequence of an LC layer, a mirror, and an LC layer. Using these matrices, the output Jones vector of the real scene, $J_{\mathrm{real.output}}$, and virtual scene, $J_{\mathrm{virt.output}}$, can be described as

$$J_{\mathrm{real.output}} = \begin{bmatrix} J_{rs} \\ J_{rp} \end{bmatrix} = T_{\mathrm{HWP}}^{(-\theta)} T_{\mathrm{LCoS}}^{(\delta)} T_{\mathrm{HWP}}^{(\theta)} J_{\mathrm{LP}}^{(0)}$$
$$= \exp\left(j\frac{2\pi}{\lambda}n_0 d\right)\begin{bmatrix} \cos^2 2\theta + e^{j\delta(x,y)}\sin^2 2\theta \\ -\cos 2\theta \sin 2\theta(1 - e^{j\delta(x,y)}) \end{bmatrix}, \quad \textbf{(B2)}$$

$$J_{\text{virt.output}} = \begin{bmatrix} J_{vs} \\ J_{vp} \end{bmatrix} = T_{\text{HWP}}^{(-\theta)} T_{\text{LCoS}}^{(\delta)} T_{\text{HWP}}^{(\theta)} J_{\text{unpolarized}}$$
$$= \exp\left(j\frac{2\pi}{\lambda} n_0 d\right) \begin{bmatrix} \cos 2\theta \cos\varphi + e^{j\delta(x,y)} \sin 2\theta \sin\varphi \\ -\sin 2\theta \cos\varphi + e^{j\delta(x,y)} \cos 2\theta \sin\varphi \end{bmatrix}. \quad \textbf{(B3)}$$

In the proposed optics, the PBS located after the polarization-control optics reflects the s-polarized components of the lights (i.e., J_{rs} and J_{vs} in the above equation) toward the world side. Only the p-polarized components (i.e., J_{rp} and J_{vp}) are observed by the user's eye. As explained in Section 2.D, the requirements that must be satisfied for the simultaneous real scene masking and virtual image displaying are

$$J_{rp} = -\cos 2\theta \sin 2\theta(1 - e^{j\delta(x,y)}) = 0, \quad \textbf{(B4)}$$

$$J_{vp} = -\sin 2\theta \cos\varphi + e^{j\delta(x,y)} \cos 2\theta \sin\varphi \neq 0 \quad \textbf{(B5)}$$

in the masking state, and

$$J_{rp} = -\cos 2\theta \sin 2\theta(1 - e^{j\delta(x,y)}) \neq 0, \quad \textbf{(B6)}$$

$$J_{vp} = -\sin 2\theta \cos\varphi + e^{j\delta(x,y)} \cos 2\theta \sin\varphi = 0 \quad \textbf{(B7)}$$

in the nonmasking state. These requirements in Eqs. (B4)–(B7) lead to the LCoS phase retardation conditions $\delta(x,y) = 2\pi$ and $\theta \neq n\pi/4$ (n: integer) in the masking state, and $\delta(x,y) = \pi$ and $\varphi = -2\theta$ in the nonmasking state, respectively. However, the phase retardation δ is dependent on the wavelength, resulting in a defective occlusion of the real scene in the full-color range. Therefore, choosing a proper value of θ that minimizes the effect of the wavelength dependency of the phase retardation δ on the final occlusion performance is required.

Suppose the simplest case at first, where both a real scene and a virtual image light are monochromatic. When the phase retardation is optimized to the input monochromatic wavelength (i.e., masking/nonmasking LCoS pixel delays the phase of the incident light to π and 2π, respectively), the visibility of the real scene and virtual image shows exact opposite behaviors, as shown in Fig. 23.

Figure 23 shows that by setting the rotation angle θ of the HWP to $n\pi/8$ (n, odd integer), the maximum visibility of the real and virtual scene can be obtained with perfect occlusion in this simple monochromatic case. In the actual condition, however, real scene light has a continuous spectrum, which requires optimization of the θ, considering the overall visibility variation in the visible wavelength range.

Figure 24 shows the visibility of the real scene over the visible wavelength range when the phase retardation of the LCoS is optimized to each RGB color. The transmittance in Fig. 24 is the ratio of the light passing through the PBS after the polarization-based occlusion optics. As shown in the graphs, the light of the real objects in the full-color range reveals nonuniform transmittance in the nonmasking pixels, and increased transmittance (= leakage) in the masking pixels as the wavelength deviates further from the optimized wavelength.

Figure 25 shows the visibility of the real scene averaged over full-color range with respect to the variation of θ. Red and blue lines indicate the transmittance of the real scene reflected from masking pixel and nonmasking pixel of the LCoS, respectively. The vertical bars in Fig. 25 represent the variance of the transmittance over the full-color spectrum. As depicted in Fig. 25, the average transmittance of both a masking pixel and a nonmasking pixel increases when the θ is close to $\pi/8$, requiring compromise between the see-through real scene transmittance and the occlusion performance. In our experiments, we chose θ between $\pi/16$ and $\pi/8$. Note that all experimental results in Section 4 show highly enhanced virtual image visibility and contrast with clear see-through real scene, even in the full-color case despite this compromise.

APPENDIX C: MODIFIED OPERATION SCHEME BASED ON THE POLARIZATION CHANGE OF THE VIRTUAL IMAGE LIGHT SOURCE

Up to now, the LP angle φ for the virtual image light source was set to be twice that of the HWP angle θ with a negative sign, i.e., $\varphi = -2\theta$. In this appendix, we discuss another possible choice of the LP angle φ and the resultant modification of the time-multiplexed operation scheme. The equations in

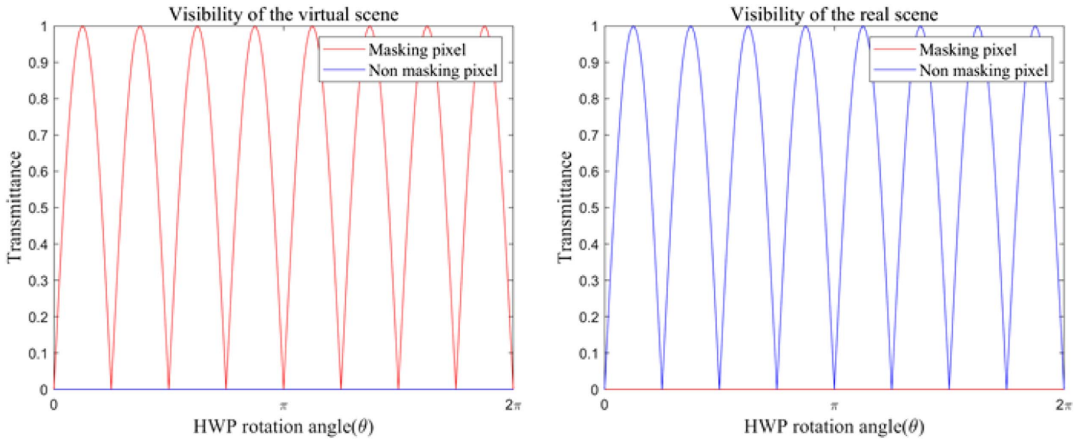

Fig. 23. Visibility of the real scene and the virtual scene when the incident light is monochromatic.

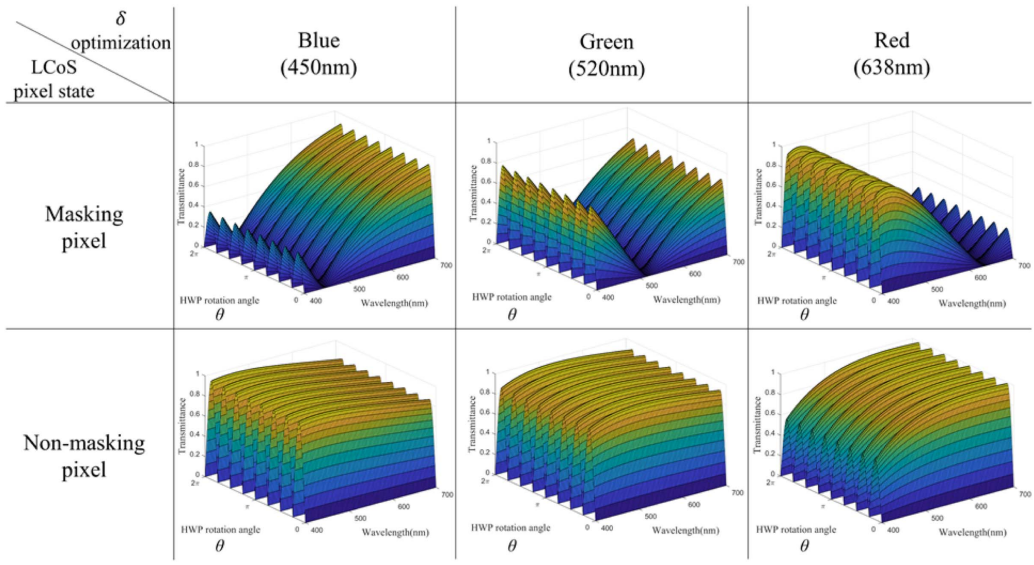

Fig. 24. Visibility of the real scene in the full-color spectrum when the phase retardation of the LCoS is optimized to each RGB color.

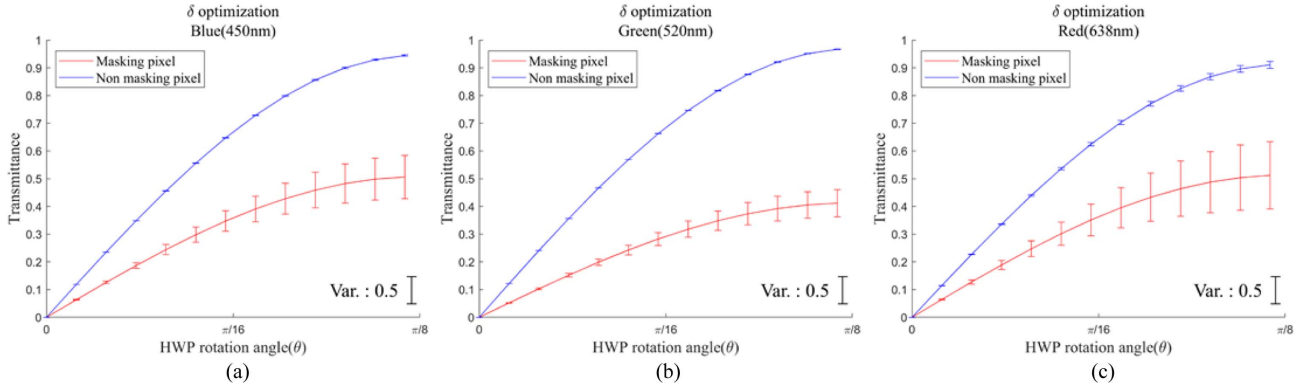

Fig. 25. Average visibility of the real scene over the full-color range when the phase retardation of the LCoS is optimized to wavelengths of (a) 450 nm, (b) 520 nm, and (c) 638 nm. The vertical bars indicate the variance of the real scene visibility over the full-color range at each HWP rotation angle.

Section 2.D and Appendix B were formulated to obscure the real scene within the virtual image area in a single frame without time multiplexing. However, given that the time-multiplexing scheme, which sequentially blocks the real scene and presents the virtual image, is applied, the equations above possess an extra degree of freedom. In the time-multiplexed operation, the mask subframe primarily contributes to the occlusion of the real scene. Consequently, the pixels of a virtual grayscale image in the image subframe are exempt from the masking function, which also allows for the presentation of some portion of the real scene. This means that the condition in Eq. (B4) that the virtual image area should block the real scene is no longer applied; i.e., the LP angle φ can be varied, not limited to -2θ.

From Eqs. (B5)–(B7) without Eq. (B4), we can find another possible choice of $\varphi = +2\theta$. Figure 26 shows the visibility of the real scene $|J_{rp}|$ and the virtual image $|J_{vp}|$ with respect to the HWP angle θ and the LCoS phase retardation δ in the case of $\varphi = +2\theta$. By comparing Figs. 26 and 7, it can be found that the lower gray level (darker) virtual image part occludes the real scene stronger in this new condition $\varphi = +2\theta$, as shown in Fig. 26, while the higher gray level (brighter) virtual image part occludes the real scene stronger in the original condition $\varphi = -2\theta$, as shown in Fig. 7.

Note that this new characteristic can be beneficial when the brightness of the real scene is much higher than the virtual light source. Due to the time multiplexing in the grayscale virtual image case, some of the real scene light is not completely occluded in the image subframe but is leaked into the observed virtual image, degrading the contrast. In the original condition $\varphi = -2\theta$, this leakage is stronger in the dark virtual image area, affecting the virtual image contrast more, while it is stronger in the bright virtual image area in the new condition $\varphi = -2\theta$, affecting the contrast less. An experimental demonstration of

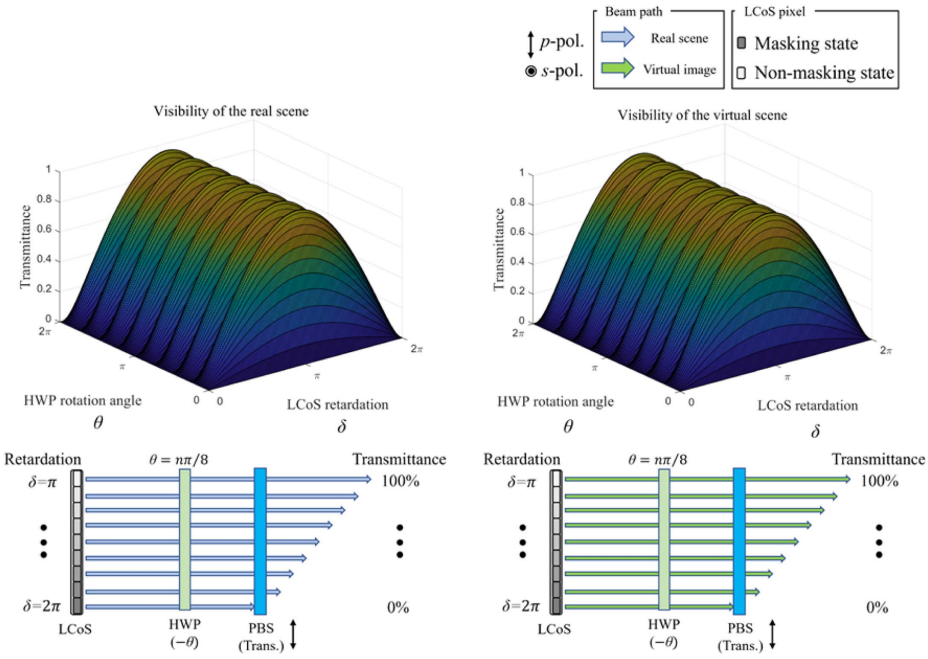

Fig. 26. Visibility of the real scene and virtual scene according to the LCoS retardation and HWP rotation angle when $\varphi = 2\theta$.

the enhanced contrast under highly bright real scene condition using the new choice $\varphi = +2\theta$ is given in Appendix D.

APPENDIX D: MEASUREMENT OF THE OCCLUSION AND CONTRAST

We measured the occlusion ratio of the occlusion mask as well as the contrast ratio of the virtual image. In the measuring process, we used a two liquid-based FTLs implementation with a CMOS camera (BFS-PGE-200S6-C, FLIR) for accurate measurement of the occlusion performance.

Figure 27 shows the experimental setup for the occlusion ratio measurement and its result. We placed a white background in front of the system and displayed a binary stripe pattern as a virtual image. Without the illumination of the virtual light source, we captured the result and designated the region of interest (ROI), as shown in Fig. 27(b). The intensity distribution of the center horizontal line in the ROI in Fig. 27(c) shows

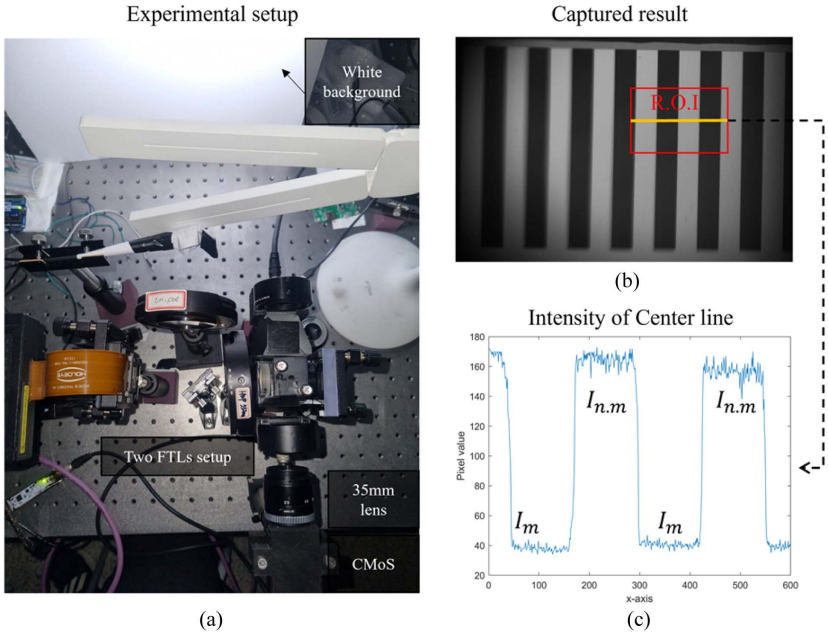

Fig. 27. (a) Experimental setup for measuring the occlusion ratio. (b) Captured result and ROI used to calculate the occlusion ratio of the masking pixel. (c) Intensity distribution of the center line in the ROI indicated as a yellow line.

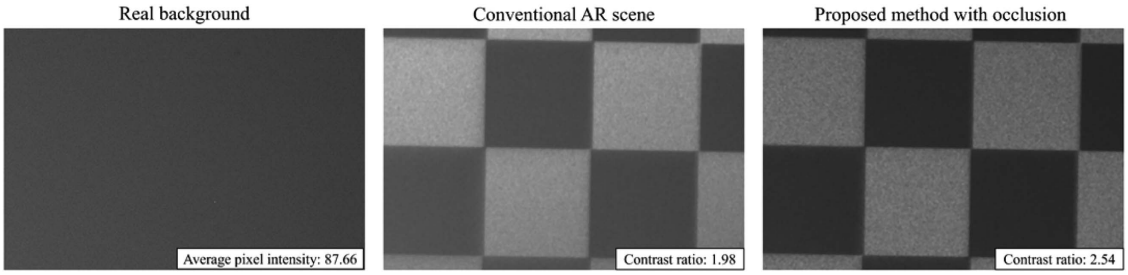

Fig. 28. Experimental result of the contrast measurement under moderate real background luminance when $\varphi = -2\theta$.

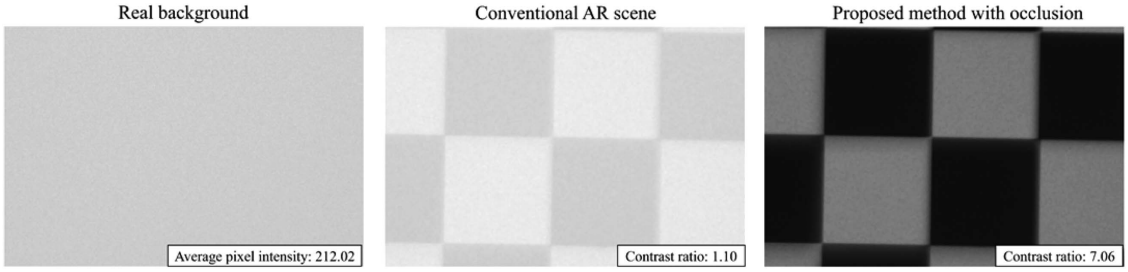

Fig. 29. Experimental result of the contrast measurement under strong real background luminance when $\varphi = 2\theta$.

that the masking pixel effectively reduced the luminance of the real scene. The measured average CMOS camera pixel value in the nonmasking and masking area [i.e., $I_{n,m}$ and I_m in Fig. 27(c)] was 158.55 and 40.55, respectively. The resultant occlusion ratio is, therefore, estimated to be (158.55–40.55)/158.55 = 74.42%.

To calculate the contrast of the virtual image in the AR scene, we displayed a checkerboard pattern as a virtual image and captured the result on a white background real object using the same experimental setup shown in Fig. 27(a). Note that the luminance level of the real scene significantly affects the contrast of the virtual image in the AR scene. Hence, we tested two cases: a moderate real background luminance case similar to the condition in Section 4, and a strong real background luminance case much brighter than the virtual image light source.

Figure 28 shows the captured results in the moderate real background luminance case. To maintain the real scene luminance consistency between the conventional and the proposed method, we captured the virtual image and real background separately and merged them together computationally in the conventional AR scene. The average CMOS camera pixel value of the real background was measured to be 87.66 in this case. As shown in Fig. 28, the proposed method gives a higher contrast ratio (=2.54) than the conventional one (=1.98), proving the validity of the occlusion scheme. The quantitative contrast ratio was calculated from the CMOS camera pixel values averaged in the black and white areas, respectively.

Figure 29 shows the experiment result in the strong real background luminance case. The average CMOS camera pixel value of the real background was measured to be 212.02, which is much higher than the previous case. In this case, for the proposed method, we used the $\varphi = +2\theta$ LP angle configuration explained in Appendix C to validate its feasibility. The resultant contrast ratio was calculated to be 7.06 in the proposed method with $\varphi = +2\theta$ and 1.10 in the conventional AR without occlusion, respectively, showing a clear enhancement of the contrast.

Funding. National Research Foundation of Korea (2022R1A2C2013455); Defense Acquisition Program Administration and Ministry of Trade, Industry and Energy of Korean government (UM22203RD2); Korea Creative Content Agency (R2022020033); Institute of Information & Communications Technology Planning & Evaluation (2020-0-00929).

Acknowledgment. This research was partly financially supported by the National Research Foundation of Korea (NRF) grant funded by the Korea Government (MSIT) (Research on occlusion-capable holographic augmented reality 3D near-eye display, 30%), Institute of Civil Military Technology Cooperation funded by the Defense Acquisition Program Administration and Ministry of Trade, Industry and Energy of the Korean government (Technical development of simultaneous acquisition of 2D/3D information based on fast-switching microlens array applicable to unmanned aerial vehicle, 50%), Culture, Sports and Tourism R&D Program through the Korea Creative Content Agency grant funded by the Ministry of Culture, Sports and Tourism in 2022 (Project Name: Development of immersive contents anti-counterfeiting and holography technology with micro patterning, 10%), and Institute of Information & Communications Technology

Planning & Evaluation (IITP) grant funded by the Korea government (MSIT) (Development of digital hologram window reconstruction, 10%).

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. C. Chang, K. Bang, G. Wetzstein, *et al.*, "Toward the next-generation VR/AR optics: a review of holographic near-eye displays from a human-centric perspective," Optica **7**, 1563–1578 (2020).
2. J. Xiong, E. L. Hsiang, Z. He, *et al.*, "Augmented reality and virtual reality displays: emerging technologies and future perspectives," Light Sci. Appl. **10**, 216 (2021).
3. S. Moon, S.-W. Nam, Y. Jeong, *et al.*, "Compact augmented reality combiner using Pancharatnam-Berry phase lens," IEEE Photon. Technol. Lett. **32**, 235–238 (2020).
4. Y. H. Lee, G. Tan, K. Yin, *et al.*, "Compact see-through near-eye display with depth adaption," J. Soc. Inf. Disp. **26**, 64–70 (2018).
5. D. Cheng, Y. Wang, C. Xu, *et al.*, "Design of an ultra-thin near-eye display with geometrical waveguide and freeform optics," Opt. Express **22**, 20705–20719 (2014).
6. C. Yoo, K. Bang, M. Chae, *et al.*, "Extended-viewing-angle waveguide near-eye display with a polarization-dependent steering combiner," Opt. Lett. **45**, 2870–2873 (2020).
7. A. Maimone, D. Lanman, K. Rathinavel, *et al.*, "Pinlight displays: wide field of view augmented reality eyeglasses using defocused point light sources," ACM Trans. Graph. **33**, 89 (2014).
8. X. Duan, J. Liu, X. Shi, *et al.*, "Full-color see-through near-eye holographic display with 80° field of view and an expanded eye-box," Opt. Express **28**, 31316–31329 (2020).
9. J. Xiong, G. Tan, T. Zhan, *et al.*, "Breaking the field-of-view limit in augmented reality with a scanning waveguide display," OSA Contin. **3**, 2730–2740 (2020).
10. X. Shi, J. Liu, Z. Zhang, *et al.*, "Extending eyebox with tunable viewpoints for see-through near-eye display," Opt. Express **29**, 11613–11626 (2021).
11. J. Jeong, J. Lee, C. Yoo, *et al.*, "Holographically customized optical combiner for eye-box extended near-eye display," Opt. Express **27**, 38006–38018 (2019).
12. T. Lin, T. Zhan, J. Zou, *et al.*, "Maxwellian near-eye display with an expanded eyebox," Opt. Express **28**, 38616–38625 (2020).
13. Y. Jo, C. Yoo, K. Bang, *et al.*, "Eye-box extended retinal projection type near-eye display with multiple independent viewpoints," Appl. Opt. **60**, A268–A276 (2021).
14. C. Jang, K. Bang, G. Li, *et al.*, "Holographic near-eye display with expanded eye-box," ACM Trans. Graph. **37**, 195 (2019).
15. H. Do, Y. M. Kim, and S.-W. Min, "Focus-free head-mounted display based on Maxwellian view using retroreflector film," Appl. Opt. **58**, 2882–2889 (2019).
16. T. Ueno and Y. Takaki, "Super multi-view near-eye display to solve vergence-accommodation conflict," Opt. Express **26**, 30703–30715 (2018).
17. C. Martinez, V. Krotov, B. Meynard, *et al.*, "See-through holographic retinal projection display concept," Optica **5**, 1200–1209 (2018).
18. A. Maimone, A. Georgiou, and J. S. Kollin, "Holographic near-eye displays for virtual and augmented reality," ACM Trans. Graph. **36**, 85 (2017).
19. C. Jang, K. Bang, S. Moon, *et al.*, "Retinal 3D: augmented reality near-eye display via pupil-tracked light field projection on retina," ACM Trans. Graph. **36**, 190 (2017).
20. C. Chang, W. Cui, J. Park, *et al.*, "Computational holographic Maxwellian near-eye display with an expanded eyebox," Sci. Rep. **9**, 18749 (2019).
21. D. Dunn, C. Tippets, and K. Torell, "Wide field of view varifocal near-eye display using see-through deformable membrane mirrors," IEEE Trans. Vis. Comput. Graph. **23**, 1322–1331 (2017).
22. 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.
23. E.-L. Hsiang, Z. Yang, Q. Yang, *et al.*, "AR/VR light engines: perspectives and challenges," Adv. Opt. Photon. **14**, 783–861 (2022).
24. B. C. Kress, *Optical Architectures for Augmented-, Virtual-, and Mixed-Reality Headsets* (SPIE, 2020).
25. E. Tatham, "Technical opinion: getting the best of both real and virtual worlds," Commun. ACM **42**, 96–98 (1999).
26. K. Kiyokawa, M. Billinghurst, B. Campbell, *et al.*, "An occlusion capable optical see-through head mount display for supporting co-located collaboration," in *2nd IEEE and ACM International Symposium on Mixed and Augmented Reality* (2003), pp. 1–9.
27. Y. Itoh, T. Hamasaki, and M. Sugimoto, "Occlusion leak compensation for optical see-through displays using a single-layer transmissive spatial light modulator," IEEE Trans. Vis. Comput. Graph. **23**, 2463–2473 (2017).
28. V. Mathur, J. N. Haddock, T. Diehl, *et al.*, "Ambient light management systems and methods for wearable devices," U.S. Patent, 11,561,401 (24 January 2023).
29. A. Wilson and H. Hua, "Design and prototype of an augmented reality display with per-pixel mutual occlusion capability," Opt. Express **25**, 30539–30549 (2017).
30. B. Krajancich, N. Padmanaban, and G. Wetzstein, "Factored occlusion: single spatial light modulator occlusion-capable optical see-through augmented reality display," IEEE Trans. Visual. Comput. Graph. **26**, 1871–1879 (2020).
31. Y.-G. Ju, M.-H. Choi, P. Liu, *et al.*, "Occlusion-capable optical-see-through near-eye display using a single digital micromirror device," Opt. Lett. **45**, 3361–3364 (2020).
32. O. Cakmakci, Y. Ha, and J. P. Rolland, "A compact optical see-through head-worn display with occlusion support," in *Proceedings of the Third IEEE and ACM International Symposium on Mixed and Augmented Reality (ISMAR)* (2004), pp.16–25.
33. C. Gao, Y. Lin, and H. Hua, "Occlusion capable optical see-through head-mounted display using freeform optics," in *IEEE International Symposium on Mixed and Augmented Reality (ISMAR)* (2012), pp. 281–282.
34. M. Chae, K. Bang, Y. Jo, *et al.*, "Occlusion-capable see-through display without the screen-door effect using a photochromic mask," Opt. Lett. **46**, 4554–4557 (2021).
35. A. Wilson and H. Hua, "Design of a pupil-matched occlusion-capable optical see-through wearable display," IEEE Trans. Vis. Comput. Graph. **28**, 4113–4126 (2021).
36. Y. Zhang, X. Hu, K. Kiyokawa, *et al.*, "Add-on occlusion: turning off-the-shelf optical see-through head-mounted displays occlusion-capable," IEEE Trans. Vis. Comput. Graph. **29**, 2700–2709 (2023).
37. Y. Zhang, X. Hu, K. Kiyokawa, *et al.*, "Realizing mutual occlusion in a wide field-of-view for optical see-through augmented reality displays based on a paired-ellipsoidal-mirror structure," Opt. Express **29**, 42751–42761 (2021).
38. L. Shi, B. Li, C. Kim, *et al.*, "Towards real-time photorealistic 3D holography with deep neural networks," Nature **591**, 234–239 (2021).
39. S. Lee, Y. Jo, D. Yoo, *et al.*, "Tomographic near-eye displays," Nat. Commun. **10**, 2497 (2019).
40. S. Choi, M. Gopakumar, Y. Peng, *et al.*, "Time-multiplexed neural holography: a flexible framework for holographic near-eye displays with fast heavily-quantized spatial light modulators," ACM Trans. Graph. **41**, 119 (2022).
41. J. Kim, M. Gopakumar, S. Choi, *et al.*, "Holographic glasses for virtual reality," in *ACM SIGGRAPH 2022 Conference Proceedings* (2022), paper 33.
42. D. Lanman and D. Luebke, "Near-eye light field displays," ACM Trans. Graph. **32**, 220 (2013).

43. T. Hamasaki and Y. Itoh, "Varifocal occlusion for optical see-through head-mounted displays using a slide occlusion mask," IEEE Trans. Vis. Comput. Graph. **25**, 1961–1969 (2019).
44. K. Rathinavel, G. Wetzstein, and H. Fuchs, "Varifocal occlusion-capable optical see-through augmented reality display based on focus-tunable optics," IEEE Trans. Vis. Comput. Graph. **25**, 3125–3134 (2019).
45. Y. Hiroi, T. Kaminokado, S. Ono, *et al.*, "Focal surface occlusion," Opt. Express **29**, 36581–36597 (2021).
46. M. Chae, J. Shin, Y. Jo, *et al.*, "Implementation of varifocal occlusion using lens arrays and focus-tunable lenses," Proc. SPIE **12443**, 124430G (2023).
47. L. Lu, S. C. McEldowney, and P. Saarikko, "Focus adjusting Pancharatnam Berry phase liquid crystal lenses in a head-mounted display," U.S. Patent, 10,379,419 (13 August 2019).
48. Z. Luo, Y. Li, J. Semmen, *et al.*, "Achromatic diffractive liquid-crystal optics for virtual reality displays," Light Sci. Appl. **12**, 230 (2023).