Optics Letters

Occlusion-capable optical-see-through near-eye display using a single digital micromirror device

YEON-GYEONG JU,¹ MYEONG-HO CHOI,¹ PENGYU LIU,² BRANDON HELLMAN,² TED L. LEE,² YUZURU TAKASHIMA,² AND JAE-HYEUNG PARK^{1,*}

¹Department of Electrical and Computer Engineering, Inha University, 100 Inha-ro, Michuhol-gu, Incheon 22212, South Korea

²University of Arizona, 1630 E. University Boulevard, Tucson, Arizona 85721, USA

*Corresponding author: jh.park@inha.ac.kr

Received 19 March 2020; revised 24 April 2020; accepted 1 May 2020; posted 12 May 2020 (Doc. ID 393194); published 17 June 2020

Occlusion of a real scene by displayed virtual images mitigates incorrect depth cues and enhances image visibility in augmented reality applications. In this Letter, we propose a novel optical scheme for the occlusion-capable optical-see-through near-eye display. The proposed scheme uses only a single spatial light modulator, as the real-scene mask and virtual image display simultaneously. A polarization-based double-pass configuration is also combined, enabling a compact implementation. The proposed scheme is verified by optical experiments which demonstrate a 60 Hz red-green-blue video display with a 4-bit depth for each color channel and per-pixel dynamic occlusion of a 90.6% maximum occlusion ratio. © 2020 Optical Society of America

<https://doi.org/10.1364/OL.393194>

Provided under the terms of the [OSA Open Access Publishing Agreement](#)

Near-eye displays (NEDs) are wearable devices presenting virtual imagery to users. In augmented reality (AR) applications, the NEDs are usually required to have optical-see-through (OST) configurations so that light from the real scene is transmitted toward the user's eye—not obscured as is the case in virtual reality applications. Virtual images displayed by a spatial light modulator (SLM) are superimposed onto the transmitted real scene by an optical combiner, giving a mixed scene to a user. Most OST NEDs are limited to an additive mode: virtual images are only added to the real scene without a masking capability. Virtual images without conjugate masks are always translucent and cannot occlude the real scene behind them. Occlusion is a strong perceptual cue describing the depth order of objects, so insufficient occlusion degrades scene reality. The contrast and visibility of displayed virtual images are also reduced significantly without occlusion, especially in bright ambient light conditions.

Several works have been reported for occlusion-capable OST NEDs [1–10]. Typical architectures form a conjugate image of the real scene onto an SLM for masking, and a second SLM forms the virtual images to be displayed. The two SLMs are then added by an optical combiner before projection into the observer's eye. After the preliminary work of Kiyokawa *et al.* [1], various implementations, including the use of a freeform prism [2], a compact package using stock lenses [3] or a double-pass

configuration [4], the use of an *x*-prism [5], and mechanical sliding of the masking SLM for varifocal occlusion [6] have been reported. Direct placement of the masking SLM in front of the eye has also been proposed [7] with the use of a microlens array [8] or computational stacked display configuration [9] for three-dimensional (3D) occlusion. Finally, the use of a digital micromirror device (DMD) as both the masking SLM and the optical combiner has also been reported [10].

Although various configurations have been proposed, all of the previous works are limited by requiring two SLMs: one masks the real scene, and the other displays the virtual image. The use of two separate SLMs not only increases system volume and complexity, but also requires precise pixel-by-pixel alignment for exact overlay of the displayed virtual image onto the real-scene mask. Furthermore, the additional image relay required by the two-SLM architecture reduces system brightness and contrast by reduced throughput and increased stray light.

In this Letter, we propose a novel OST NED scheme which enables both occlusion and display using a single SLM. The proposed scheme uses a DMD as the SLM in a polarization-based double-pass optical architecture. Employing its binary on-off states, the individual pixel mirror of the DMD in this configuration directs light either from the real scene or from a light-emitting diode (LED) toward the eye, enabling simultaneous occlusion and display with per-pixel accuracy [11,12].

Note that a closely related concurrent work by Krajancich, *et al.* has been reported [13]. The binary DMD mirror angle states are similarly employed to selectively direct light either from the real scene or from the LED toward the eye. The optical configuration, however, simply projects the real scene and the LED illumination onto the DMD from largely tilted opposite directions, and the DMD is observed from the surface-normal direction. The large tilt angle of the real-scene projection with respect to the DMD normal creates a large focus mismatch across the field of view (FOV) which would require correction optics. The off-axis reflection geometry of the real-scene light requires separated imaging optics for incident and reflected real-scene light, increasing system volume. While the basic single-SLM occlusion-display idea was verified by directly observing the DMD without any projection optics, there was

Fig. 1. Schematic diagram of the proposed method.

no development of the optical system toward implementing the concept into an OST NED application. Additionally, real-time operation was not demonstrated due to device frame rate and computational complexity limitations. In contrast, we present an OST NED system using a compact polarization-based double-pass configuration and a single DMD which physically forms the optical image of the mixed scene (occluded-real and virtual) far from the eye. Our unique DMD arrangement reduces the tilt angle of the real-scene projection by a half, relieving the focus mismatch problem significantly. The on-axis reflection geometry enables sharing of a single imaging optic, both for the incident and for the reflected real-scene light. Real-time operation with a red-green-blue (RGB) video display is also first demonstrated with dynamic occlusion.

Figure 1 illustrates a schematic diagram of the proposed system. Light from the real scene is polarized by a linear polarizer, reflected by a polarizing beam splitter (PBS), and focused by an imaging lens onto the DMD. The DMD is also illuminated by an LED from a different angle. Each pixel mirror of the DMD reflects the light either from the real scene or from the LED back along the optical axis. This reflected light is directed toward the eye through the PBS and an additional mirror by polarization state changes as illustrated in Fig. 1. A Schmidt-Pechan prism placed before the eye finally reverses the display horizontally and vertically, presenting upright images to the eye.

Figure 2 shows the light path around the DMD in the proposed system. The red and blue arrows depict the paths of light from the real scene and from the LED, respectively. The DMD is a binary device having pixelated mirrors. According to the applied “on” and “off” signal, each pixel mirror is controlled to have either a positive ($= +\theta_t$) or negative ($= -\theta_t$) tilt angle with respect to the DMD normal. In the proposed system, the DMD itself is also tilted by angle θ_t with respect to the optical axis, as shown in Fig. 2(a). The LED illuminates the DMD at an angle $3\theta_t$ with respect to the DMD normal, or equivalently $4\theta_t$, with respect to the optical axis.

In operation, the “on” state mirrors reflect light from the real scene directly back along the optical axis. This light is then captured by the imaging lens again for projection to the eye. The “off” state mirrors reflect light from the real scene toward the LED—at a $4\theta_t$ angle from the optical axis to prevent recapture by the imaging lens. Then the observer can only see real-scene regions corresponding to “on” state pixels, while “off” state regions are blocked. If the LED is on, “off” state pixel mirrors reflect LED light toward the optical axis, filling the blocked real-scene region, as shown in Fig. 2(b). Therefore, each pixel mirror can reflect either the real-scene light or the LED light

Fig. 2. Real-scene occlusion and virtual image display by the DMD. (a) Light path of the real scene and the LED; (b) example of the observed scene according to the state of the pixel mirrors.

Fig. 3. Operation scheme for RGB ($N_c = 3$) images with a 4-bit ($N_b = 4$) depth.

toward the optical axis selectively, enabling real-scene occlusion and virtual image display.

A time-multiplexing scheme is implemented to independently control the occlusion ratio of the real scene and the gray scale of the virtual images. As illustrated in the top of Fig. 3, the scheme divides each frame time T_{frame} into three segments: one for the virtual image display, T_{display} , another for its compensation, $T_{\text{compensation}}$, and the other for the real-scene occlusion, $T_{\text{occlusion}}$.

In the display segment, the LED is turned on, and the DMD displays the binary patterns representing bit planes of the virtual scene. For an image of N_b bit depth for each of the N_c color channels, N_b sub-frames of equal length $T_{\text{sub-frame}}$ are used for each color to display corresponding bit planes, consuming $N_c \times N_b$ sub-frames in total, i.e., $T_{\text{display}} = N_c N_b T_{\text{sub-frame}}$. With a constant I_o , the LED intensity is modulated to $2^n I_o$ for the n^{th} bit plane ($n = 0, 1, \dots, N_b - 1$) of each color channel, as shown in Fig. 3.

In the compensation segment, the LED is turned off, and the DMD displays $N_c \times N_b$ binary patterns in $T_{\text{compensation}} = T_{\text{display}} = N_c N_b T_{\text{sub-frame}}$. Within the occluding mask area, the binary patterns are complementary to the corresponding bit plane images in the display segment, making every pixel of the virtual scene have the same 50% duty-cycle during the two segments, i.e., $T_{\text{display}} + T_{\text{compensation}} = 2N_c N_b T_{\text{sub-frame}}$,

Fig. 4. Experimental benchtop setup.

regardless of its pixel value. This removes the dependency of the final occlusion ratio on the displayed virtual images, enabling uniform masking of the real scene.

In the occlusion segment, the LED remains turned off, and the DMD displays a binary mask pattern of occluding “off” pixels and non-occluding “on” pixels. The occlusion ratio can also be controlled by adjusting the “off” pixel duty-cycle (i.e., all “off” pixels can be turned back “on”). Denoting the pixel “off” time during the occlusion segment as $T_{\text{off,occ}}$, the total pixel off time across the whole three segments is given by $T_{\text{off}} = T_{\text{off,occ}} + (T_{\text{display}} + T_{\text{compensation}}) \times 0.5 = T_{\text{off,occ}} + N_c N_b T_{\text{sub-frame}}$. The occlusion ratio is $T_{\text{off}}/T_{\text{frame}} = (T_{\text{off,occ}} + N_c N_b T_{\text{sub-frame}})/T_{\text{frame}}$. Considering $0 \leq T_{\text{off,occ}} \leq T_{\text{occlusion}} (= T_{\text{frame}} - 2N_c N_b T_{\text{sub-frame}})$, the controllable range of the occlusion ratio is given by $N_c N_b T_{\text{sub-frame}}/T_{\text{frame}} \leq T_{\text{off}}/T_{\text{frame}} \leq 1 - N_c N_b T_{\text{sub-frame}}/T_{\text{frame}}$.

We verified the proposed method experimentally. Figure 4 shows a photo of the experimental setup. The DLPLCR6500EVM from Texas Instruments with a 9,523 Hz refresh rate ($= T_{\text{sub-frame}} = 1/9523 = 105 \mu\text{s}$) was employed as the DMD and driver. Although the DMD has a 0.65 inch diagonal size with 1920×1080 resolution and $7.56 \mu\text{m}$ pixel pitch, we used only a central 2.3 mm diameter area which corresponds to the FOV of our system. The on/off tilt angle of the pixel mirrors of the DMD is $\theta_i = 12$ deg. As illustrated in Figs. 1 and 2, the DMD was aligned with a tilt such that the “on” pixel mirrors reflect the real-scene light directly back to along the optical axis of the double-pass configuration. We used a 477 nm wavelength blue LED (EP470S04, Thorlabs) for single-color operation and a RGB LED whose wavelengths are 630, 530, and 470 nm (R05RGB-C-A001, SENWIN) for full color operation. A homemade diffusive light guide of rolled and bent white paper was used with the LED to achieve uniform illumination on the active DMD area of 2.3 mm diameter, as shown in the magnified view in Fig. 4. A Canon EF 50 mm F/1.8 II lens was used as the imaging lens between the PBS and the DMD. A 500 mm focal length concave lens was additionally used before the camera to project the mixed scene (occluded-real and virtual) back to the approximate original distance of the real scene.

An Arduino Uno microcontroller synchronizes the DMD and the LED, triggering the DMD actuation between frames and driving the LED. For the LED intensity modulation shown in Fig. 3, we implemented a pulse-width modulation where the LED on time for the n^{th} bit plane is given by $2^n/2^{N_b-1} T_{\text{sub-frame}}$, where $n = 0, 1, \dots, N_b - 1$. In order to achieve a sufficient occlusion ratio for RGB-color images, we limited the bit depth N_b to 4-bit in the experiment. With $T_{\text{sub-frame}} = 105 \mu\text{s}$, the

Fig. 5. Objects and images used in the experiment. (a) Real object, (b) “pirate parrot” image extracted from the Microsoft Office 3D model sample library, and (c) image data loaded to the DMD.

Fig. 6. Experimental result. Images were captured with the LED off and sub-frame length increased ($T_{\text{sub-frame}} = 420 \mu\text{s}$) for clear demonstration of the non-uniform occlusion artifact (a) before applying the compensation segment and (b) after applying the compensation segment.

maximum occlusion ratio we obtained in the experiment is 96.9% for a single-color ($N_c = 1$), 4-bit ($N_b = 4$), 60 Hz display and 90.6% for a RGB ($N_c = 3$), 4-bit ($N_b = 4$), 60 Hz display where the numbers are slightly smaller than the theoretical maximum values $1 - N_c N_b T_{\text{sub-frame}}/T_{\text{frame}}$ due to additional $T_{\text{sub-frame}}$ used for the DMD triggering.

Figure 5 shows the real scene and the virtual image used in the experiment. The object shown in Fig. 5(a) was located about 900 mm in front of the setup. The pirate parrot image shown in Fig. 5(b) was displayed, occluding the real object in Fig. 5(a). The mask, bit plane, and compensation images loaded to the DMD in the real-scene occlusion, the virtual image display, and the compensation segments, respectively, are shown in Fig. 5(c).

Figure 6 is the experimental result showing the effect of the compensation segment. When the images in Fig. 6 were captured, the LED was constantly off, and the sub-frame length was deliberately increased to $T_{\text{sub-frame}} = 420 \mu\text{s}$ for clear demonstration. If the compensation segment is not used, the overall occlusion ratio $T_{\text{off}}/T_{\text{frame}}$ of each pixel is dependent on the bit plane patterns displayed in the display segment, making the non-uniform occlusion artifacts, as shown in Fig. 6(a). With the compensation segment, this dependency is removed, achieving the uniform occlusion effect, as demonstrated in Fig. 6(b).

Figure 7 shows the experimental results for a single-color ($N_c = 1$), 4-bit ($N_b = 4$) and RGB-color ($N_c = 3$), 4-bit ($N_b = 4$) display with $T_{\text{sub-frame}} = 105 \mu\text{s}$. We captured the images while operating the system for minimum, medium, and maximum occlusion ratios: $T_{\text{off}}/T_{\text{frame}} = 2.5\%$, 49.7%, and 96.9% for the single-color case shown in Fig. 7(a), and $T_{\text{off}}/T_{\text{frame}} = 7.6\%$, 49.1%, and 90.6% for the RGB-color case

Fig. 7. Experimental results for a 4-bit ($N_b = 4$), 60 Hz display with different occlusion ratios. (a) Single color ($N_c = 1$) and (b) RGB color ($N_c = 3$). The lower row images were captured with LED operation. For movies, see [Visualization 1](#) for single color and [Visualization 2](#) for RGB.

shown in Fig. 7(b). In both Figs. 7(a) and 7(b), the upper row shows images captured when the LED is constantly off, showing only the occlusion effect, and the lower row is the images captured with the LED operation showing the virtual image addition. Figure 7 clearly demonstrates the successful occlusion of the real scene by the displayed virtual images. From the comparison between different occlusion ratios, the enhanced visibility of the displayed virtual images by the occlusion is also confirmed. The captured movies (see [Visualization 1](#) and [Visualization 2](#)) demonstrate real-time operation of the 60 Hz moving picture display with exact pixel-matched occlusion successfully.

Simultaneous occlusion and display by a single SLM in an OST NED architecture are successfully verified, though the current experimental implementation has limitations. The system has a limited FOV, measured to be 2.9° which is too restrictive for NED applications. A dedicated opto-mechanic design with freeform optics, rather than the use of the standard bulk optics with individual mounts in our benchtop, would increase the FOV significantly. Our experimental verification is limited to 4-bit RGB images with a 90.6% maximum occlusion ratio due to the 9,523 Hz refresh rate of the DMD used in the experiment. This limitation, however, can be overcome easily by using a DMD with a higher refresh rate. For example, a DMD with

a 23 KHz refresh rate which is commercially available would enable 8-bit RGB images with a 93.7% maximum occlusion ratio. The DMD plane is tilted by $\theta_i = 12^\circ$ which causes a focus mismatch across the FOV. We expect that a focus plane tilt correction could be incorporated into the imaging optics by accommodating a Scheimpflug imaging scheme or by using a prism in front of the DMD. Finally, the polarization optics in our double-pass configuration cuts off 50% of the incident light and exhibits the extinction ratio degradation at larger FOV. This optical loss lowers brightness and contrast of the OST real scene, and its improvement is a topic of further research.

We have proposed an occlusion-capable OST NED. The proposed NED uses a single DMD and the polarization-based double-pass configuration. The single DMD works as both the real-scene mask and the virtual image display through time-multiplexing, enabling pixel-matched occlusion. The combination of the polarization-dependent double-pass configuration and the use of the single DMD contribute to the realization of the compact occlusion-supporting AR OST NED. The proposed system is verified experimentally, demonstrating a 60 Hz real-time 4-bit RGB video display with dynamic real-scene occlusion with a 90.6% maximum occlusion ratio.

Funding. Institute for Information and Communications Technology Promotion (2017-0-00417, GK20D0100, IITP-2018-2015-0-00448); National Research Foundation of Korea (NRF-2017R1A2B2011084).

Disclosures. The authors declare no conflicts of interest.

REFERENCES

1. K. Kiyokawa, M. Billingham, B. Campbell, and E. Woods, in *Proceedings of the Second IEEE and ACM International Symposium on Mixed and Augmented Reality* (IEEE, 2003), p. 133.
2. C. Gao, Y. Lin, and H. Hua, *Proc. SPIE* **8735**, 87350F (2013).
3. A. Wilson and H. Hua, *Opt. Express* **25**, 30539 (2017).
4. A. Wilson, in *Frontiers in Optics* (Optical Society of America, 2019), paper FTh3A.2.
5. O. Cakmakci, Y. Ha, and J. P. Rolland, in *Proceedings of the Third IEEE and ACM International Symposium on Mixed and Augmented Reality* (IEEE, 2004), p. 16.
6. T. Hamasaki and Y. Itoh, *IEEE Trans. Vis. Comput. Graph.* **25**, 1961 (2019).
7. Y. Itoh, T. Hamasaki, and M. Sugimoto, *IEEE Trans. Vis. Comput. Graph.* **23**, 2463 (2017).
8. Y. Yamaguchi and Y. Takaki, *Appl. Opt.* **55**, A144 (2016).
9. A. Maimone and H. Fuchs, in *Proceedings of 2013 IEEE International Symposium on Mixed and Augmented Reality* (IEEE, 2013), p. 29.
10. K. Kim, D. Heo, and J. Hahn, in *Proceedings of the 7th International Conference on Photonics, Optics and Laser Technology (INSTICC)* (2019), pp. 299.
11. Y. G. Ju, M. H. Choi, P. Liu, B. Hellman, T. L. Lee, Y. Takashima, and J. H. Park, in *Proceedings of the 9th Korea-Japan Workshop on Digital Holography and Information Photonics (OSK, OSJ, and KOPTI)* (2019), pp. 75.
12. Y. G. Ju, M. H. Choi, P. Liu, B. Hellman, T. L. Lee, Y. Takashima, and J. H. Park, in *Proceedings of Digital Holography and 3-D Imaging* (Optical Society of America, 2020).
13. B. Krajancich, N. Padmanaban, and G. Wetzstein, *IEEE Trans. Vis. Comput. Graph.* **26**, 1871 (2020).