





Realizing mutual occlusion in a wide field-of-view for optical see-through augmented reality displays based on a paired-ellipsoidal-mirror structure

YAN ZHANG,^{1,*}  XIAODAN HU,¹ KIYOSHI KIYOKAWA,¹ NAOYA ISOYAMA,¹ HIDEAKI UCHIYAMA,¹ AND HONG HUA² 

¹Graduate School of Science and Technology, Nara Institute of Science and Technology, 8916-5 Takayama, Ikoma, Nara 6300192, Japan

²James Wyant College of Optical Sciences, University of Arizona, 1630 E University Blvd, Tucson, AZ 85721, USA

*yan.zhang.zq6@is.naist.jp

Abstract: Mutual occlusion is an essential feature for augmented reality (AR) displays for allowing the virtual content to be clearly perceived under an excessively illuminated environment. Although a few works have been done to facilitate the performance of occlusion-capable optical see-through augmented reality (OC-OST-AR) displays, the realization of mutual occlusion in a wide field-of-view (FOV) is still challenging. Divergent from typical hard-edge occlusion and soft edge-occlusion designs, we propose the paired-ellipsoidal-mirror (PEM) structure. The proposed system is allowed to support either hard-edge occlusion or enhanced soft-edge occlusion in a wide FOV by optionally fixing a spatial light modulator (SLM) before the entrance pupil or at an inner focal plane. The numerical aperture (NA) of the system is efficiently increased by the combination of paired ellipsoidal mirror imaging and aperture stop restriction. With proof-of-concept prototypes built, virtual display in a FOV of $H160^\circ \times V74^\circ$ and mutual occlusion in a FOV of $H122^\circ \times V74^\circ$ are demonstrated with a basic design, respectively. Furthermore, a mixed FOV of $H95.3^\circ \times V52.9^\circ$ is demonstrated by an optimized design with vertical parallax reduction and virtual display improvement.

© 2021 Optica Publishing Group under the terms of the [Optica Open Access Publishing Agreement](#)

1. Introduction

Augmented reality (AR) is the technology that combines digital information with the physical world. Optical see-through augmented reality (OST-AR) displays are considered as an efficient platform to support AR experience for users. Despite the fact that the imaging systems of OST-AR displays have been optimized by various works [1,2], lack of modulation on the real scene is still a crucial issue that prevents AR contents from being provided vividly. Due to the usage of optical combiners and the consideration of balancing power consumption, state-of-the-art OST-AR displays are designed to project virtual images with sufficient illuminance in general indoor scenarios (e.g., the maximum illuminance $\sim 200\text{lux}$ of HoloLens 2) while the environment illuminance easily increases to another order of magnitude when an OST-AR display works outdoor (e.g., the illuminance of 10000lux for full daylight) [3]. AR contents displayed upon a bright background suffer from low visibility by the highly transparent virtual images and have poor color fidelity since a perceived image is a mix of the background light and the projected light. Besides, lack of occlusion between real objects and virtual objects makes the user to lose crucial perceptual cues. The real scene needs to be modulated as the virtual content for solving this issue. Consequently, mutual occlusion is introduced into typical OST-AR displays to address those challenges. Occlusion-capable optical see-through augmented reality displays (OC-OST-AR displays) implement spatial light modulators (SLMs) to block incident light from the real world.

With a variety of OC-OST-AR displays proposed, a compact structure with decent occlusion performance has been realized [4]. However, real scenes of OC-OST-AR displays can not be preprocessed. Thus the occlusion-capable FOV is more restricted by the optical aberration of imaging systems than the virtual display FOV. Therefore, FOVs of existing OC-OST-AR displays are difficult to reach the range of common wide-view OST-AR displays [5,6].

Common methods of mutual occlusion in OC-OST-AR displays are known as hard-edge occlusion and soft-edge occlusion. In the hard-edge occlusion approach, a spatial light modulator (SLM) is positioned at an inner focal plane of the imaging system. Mutual occlusion is conducted with a focused image of the real scene so that pixelated precision is achieved. Recent progress of hard-edge occlusion improves the graphics performance of OC-OST-AR displays by implementing reflective SLMs. Liquid crystal on silicon (LCoS) devices [7] and digital mirror devices (DMDs) [8,9], rather than transmissive SLMs are used to render the occlusion pattern. A movable lens design [10] and the usage of tunable lenses [11] combine a vari-focal display and an occlusion capability in OC-OST-AR displays. Nevertheless, limited numerical aperture (NA) of lenses used in hard-edge occlusion systems prevents the expansion of FOV. One of the solutions to overcome the restriction from the nature of lenses is conducting mutual occlusion with only SLMs, which is the soft-edge occlusion approach. Although occlusion patterns rendered in this way appear severely blurred due to the out-of-focus of the SLM plane, soft-edge occlusion systems provide wider FOVs than hard-edge occlusion systems because the real scene is no longer processed by lenses. And this lens-less design also brings a smaller form-factor to OC-OST-AR displays. Targeting to overcome the poor occlusion performance in soft-edge occlusion systems, a few methods based on ingenious image processing, including light-field occlusion with a multi-layer SLMs structure [12] and video-compensation occlusion [13], are proposed. Better occlusion performance has been achieved, but heavier computation and additional latency become issues. And the occlusion-capable FOV is still restricted by the distance between the SLMs and the pupil.

Motivated by the work by Yang et al. [14], a paired-ellipsoidal-mirror (PEM) structure that supports wide-view mutual occlusion in AR displays is proposed in this paper. Compared to the double-parabolic mirror system proposed by Zhang et al., the PEM structure achieves a wider FOV with fewer optical elements [15]. A more cost-effective OC-OST-AR display is expected to be developed with the PEM structure. With beam propagation through the imaging system analyzed, mutual occlusion can be realized by either locating an SLM at an inner focal plane or before the entrance pupil of the system. The former configuration conducts mutual occlusion in the hard-edge occlusion way, and the latter solution conducts mutual occlusion in the enhanced soft-edge way. The paired ellipsoidal mirrors function the objective lens and the eyepiece to increase the NA of the imaging system significantly. Hence, the PEM system focuses the real scene into a much smaller space than typical OC-OST-AR displays, allowing the dimension-limited SLMs to conduct hard-edge occlusion in a wide FOV. An aperture stop at the shared focus of the paired ellipsoidal mirrors shrinks propagation beams and projects a virtual pinhole onto the user's pupil. Thus, the typical soft-edge occlusion approach is enhanced to higher precision and wider FOV by an optical way, which does not require additional computation. A folded structure is designed to reduce the vertical parallax, and an algorithm for virtual display is developed to compensate for optical aberration and to attenuate the defocus. Proof-of-concept prototypes are built with a basic design and an optimized structure, respectively. The maximum virtual display FOV of $H160^\circ \times V74^\circ$ and an occlusion-capable FOV of $H122^\circ \times V74^\circ$ are demonstrated with the basic prototype. A mixed FOV of $H95.3^\circ \times V52.9^\circ$ is demonstrated with the optimized prototype.

2. Optical design

The schematic diagram of the proposed system is shown as Fig. 1. Chief rays emitted from the real scene are drawn with red lines, and the virtual image projected by the LCD is drawn with blue lines. Optical elements, including two paired ellipsoidal mirrors and a lens pair consisting

of a concave lens and a convex lens, are drawn with solid lines for the usable parts and with dash lines for the redundant parts. The redundant parts are not related to beam propagation but exist in commercial products, which increases the form-factor of the proposed system. An optical combiner between the concave lens and the convex lens is used to merge the real scene and the virtual image. The optical path in the vertical direction is folded by two mirrors so that the vertical parallax is reduced. The LCD is tilted to partially compensate for the defocus of projected virtual images. A mirror located before the LCD is used to further compress the optical path of the virtual display. A pinhole mask is positioned at the back focus of the convex lens that works as the aperture stop of the system. Hence, the entrance pupil and the exit pupil are defined as images of the pinhole projected by the front elements group and the back ellipsoidal mirror, respectively. Mutual occlusion is conducted by fixing an SLM either in front of the entrance pupil ($SLM_{(op1)}$) or above the lens pair ($SLM_{(op2)}$), which leads the system to work in the enhanced soft-edge occlusion way or the hard-edge occlusion way.

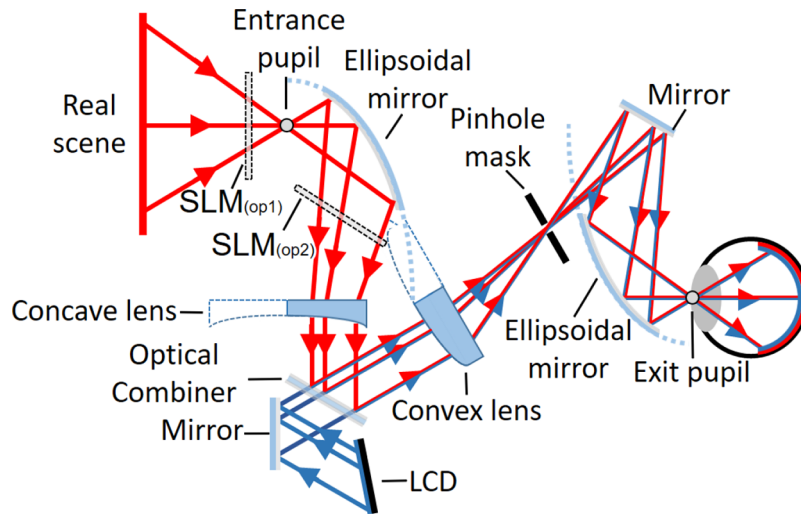


Fig. 1. The schematic diagram of the proposed system. SLM_{op1} and SLM_{op2} are optional positions for an SLM to render mutual occlusion. The system works in an enhanced soft-edge occlusion way when the SLM is located at SLM_{op1} , or in the hard-edge occlusion way when the SLM is located at SLM_{op2} .

The unfolded structure of the proposed system is shown as Fig. 2. Rays emitted from the front real scene and the LCD are similarly drawn as red and blue lines. The paired ellipsoidal mirrors have the same major axis length of a and minor axis length of b . Foci of the paired ellipsoidal mirrors are labeled as F_1 , F_2 , F_3 , and F_4 , respectively. P_{o1} and P_{o2} are the two feasible occlusion panel planes, P_d is the inclined LCD plane, and P_m is the pinhole mask plane. The concave lens and the convex lens are signed as L_v and L_x , and their focal lengths are f_v and f_x , respectively.

When the occlusion panel is positioned at P_{o1} , the proposed system conducts mutual occlusion similar to typical soft-edge OC-OST-AR displays. Nevertheless, the whole imaging system functions a virtual pinhole (with extra parallax) on the user's pupil that shrinks propagation beams to improve the occlusion precision while keeping a wide FOV, though it reduces the image brightness at the same time. The occlusion pattern can be computed from the projection between the target real scene and physical pixels on the occlusion panel, which benefits the proposed system with minimum computation.

In the case that the occlusion panel is positioned at the inclined plane P_{o2} , the proposed system works in the hard-edge occlusion way. Light emitted from a real scene pixel is transformed by

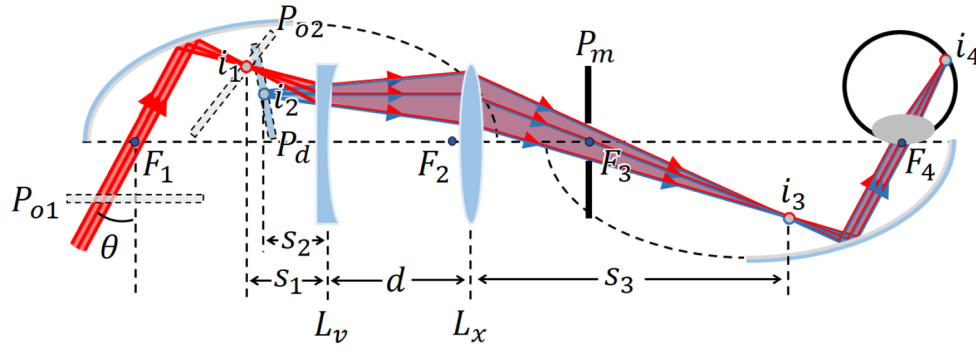


Fig. 2. Unfolded structure of the proposed system.

optical elements as intermediate images points i_1 , i_3 , and i_4 . The entire beam can be blocked by modulating the corresponding pixel on the occlusion panel at P_{o2} . Occlusion patterns in this configuration can be derived from the imaging equation of the ellipsoidal mirror. For attenuating computation, the ellipsoidal mirror is treated as a series of continuous tiny curved mirrors that spreads along an ellipsoidal surface. The focal length of each curved mirror is given as:

$$f_e = \frac{R}{2} = \frac{(2ar - r^2)^{\frac{3}{2}}}{2ab} \quad (1)$$

Where R is the radius of curvature of a certain curved mirror, and it can be further derived from r , which is the distance from the curved mirror to the focus. Let the incident angle θ clock-wise from the perpendicular through F_1 to the chief ray of a beam be plus, r can be calculated with the polar equation of an ellipse, which is:

$$r = \frac{a(1 - e^2)}{1 - e \sin \theta} \quad (2)$$

Where $e = c/a$ is the eccentricity of the ellipse. The projection between a real scene pixel and the focused image point i_1 is generated with Eq. (1) and Eq. (2), then an occlusion pattern for an arbitrary see-through vision area can be computed.

A relay lens pair between the two ellipsoidal mirrors are installed to guarantee a parallel incoming beam through F_1 be recovered correctly at F_4 . Let focal lengths of the two lenses be as $f_v = -f_x = -f$, and the back foci of L_v and L_x are fixed at F_2 and F_3 , respectively. The spacing d between the two lenses is given as [16]:

$$d = \frac{2f^2}{s_1 + f} \quad (3)$$

Where s_1 is the distance from i_1 to L_v . With a constant s_1 , the system generates a clear see-through view and renders a sharp occlusion pattern simultaneously. However, Eq. (1) and Eq. (2) indicate that i_1 is different from each incident angle θ altogether, which leads to a volumetric instead of planar distribution of i_1 around P_{o2} . Hence, a planar occlusion panel inevitably faces the defocus at most see-through FOVs. What is more, the projection from i_1 to the conjugate image point i_3 causes an inverse proportion between s_1 and s_3 . As a result, the generated real scene pixel i_4 on the retina becomes out-of-focus soon when a deviation from the focused FOV occurs. Therefore, the spatial location of P_{o2} should refer to the detailed distribution of i_1 versus different FOVs for attenuating the defocus of the occlusion pattern.

Assuming that an ellipsoidal mirror with a major axis length of 61.9mm and a minor axis length of 36.4mm, the distribution of i_1 versus a see-through FOV of $H60^\circ \times V60^\circ$ is shown as

the upper colorful surface in Fig. 3. The z coordinate represents distances from image points to the L_v plane. Along with the FOV varies from $H0^\circ \times V30^\circ$ to $H0^\circ \times V-30^\circ$, s_1 increases from 9.2mm to 44.5mm, contours of s_1 versus each FOV are drawn below for better understanding. Therefore, the occlusion panel fixed at an inclined P_{o2} plane, which is shown as Fig. 2, reduces the defocus of each occlusion pixel so that it expands the hard-edge occlusion-capable FOV. In order to test the occlusion performance with this configuration, a transparent USFA 1951 target is used to simulate an occlusion panel. The experiment results show a sharp occlusion pattern in a designated FOV with the target correctly placed (see Supplement 1). Unfortunately, the dimensions of available SLM panels are too large to be placed at P_{o2} for the ellipsoidal mirror used in our prototype. Therefore, the SLM is chosen to be placed at P_{o1} , and the prototype works in the enhanced soft-edge occlusion way.

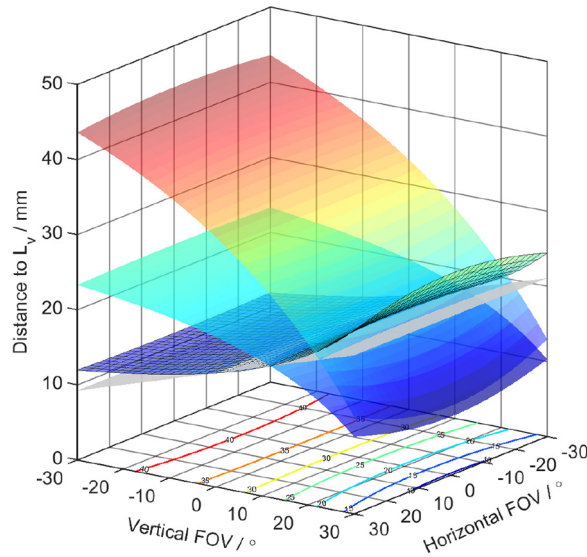


Fig. 3. Ideal pixel distributions within a FOV of $H60^\circ \times V60^\circ$ calculated for an occlusion panel at P_{o2} (the upper colorful surface) and a display panel at P_d (the meshed surface). The z coordinate represents the distance to the concave lens L_v . When a planar display panel is used, the optimized layout of P_d is shown as the grey plane with a rotated angle of 11.2° to minimize the defocus.

After being imaged by the concave lens, real scene pixels spread within a narrower range shown as the lower colorful surface in Fig. 3. Hence, the virtual image projected by a planar LCD at P_d suffers less defocus than the occlusion pattern at P_o . Moreover, the image point i_2 , which is projected by the pixel i_4 at the retina through the second ellipsoidal mirror and the convex lens, is traced with the same FOV. The distribution of i_2 is shown as the meshed surface in Fig. 3, which is a space between the vertex of 26.8mm at the FOV of $H0^\circ \times V30^\circ$ and the nadir of 11.7mm at the FOV of $H0^\circ \times V-30^\circ$. A flexible display panel can be used to display a virtual image free of defocus, while it also makes the imaging system costly. In the case that planar display panels are used for projecting virtual images, an optimized grey plane with a tilt angle of 11.2° in Fig. 3 is calculated. With the FOV of $H0^\circ \times V0^\circ$ kept to be in-focus, RMS (root mean square) of defocus between pixels aligned with the optimized P_d plane and the ideal distribution of i_2 is minimized.

3. Prototype configuration and analysis

A previous prototype based on the unfolded structure has demonstrated a virtual display FOV of $H160^\circ \times V74^\circ$ and an occlusion-capable FOV of $H122^\circ \times V74^\circ$ (see Supplement 1). The bench-top

prototype shown in Fig. 4(a) is built based on the optimized design. The paired ellipsoidal mirrors have the same major axis length of 61.9mm and minor axis length of 36.4mm as the above simulation. A SONY LCX017 panel, which has the resolution of 1024×768, the dimension of 36.8mm×27.6mm, and the transmittance of 10% (measured with an additional linear polarizing film) is fixed before the entrance pupil for rendering occlusion patterns. Hence, the prototype works in the enhanced soft-edge occlusion way. A tilted SHARP LS029B3SX02 panel whose resolution and dimension are 1440×1440 and 51.84mm×51.84mm is directly placed before a half mirror working as the optical combiner. As shown in Fig. 1, a reflective mirror can be additionally installed to compress the light path of virtual display further. The lens pair is composed of a plano-concave lens with a focal length of −50mm and a diameter of 40mm, and an aspheric lens with a focal length of 50mm and a diameter of 50mm. The half mirror and a following flat mirror are tilted by an angle of 45°. A compressed vertical parallax of 129.6mm and an enlarged horizontal parallax of 130.8mm are given by measuring the distances between the entrance pupil and the exit pupil in the vertical and horizontal directions, respectively. Further reduction of the vertical parallax requires customized optical elements that allow the half mirror to reflect rays from the plano-concave lens upward, which is the ideal case drawn with solid lines in Fig. 1. The iris diaphragm is set as 3mm. Based on ZEMAX simulation, the see-through efficiency is 0.5% compared to naked eyes with a pupil diameter of 2mm.

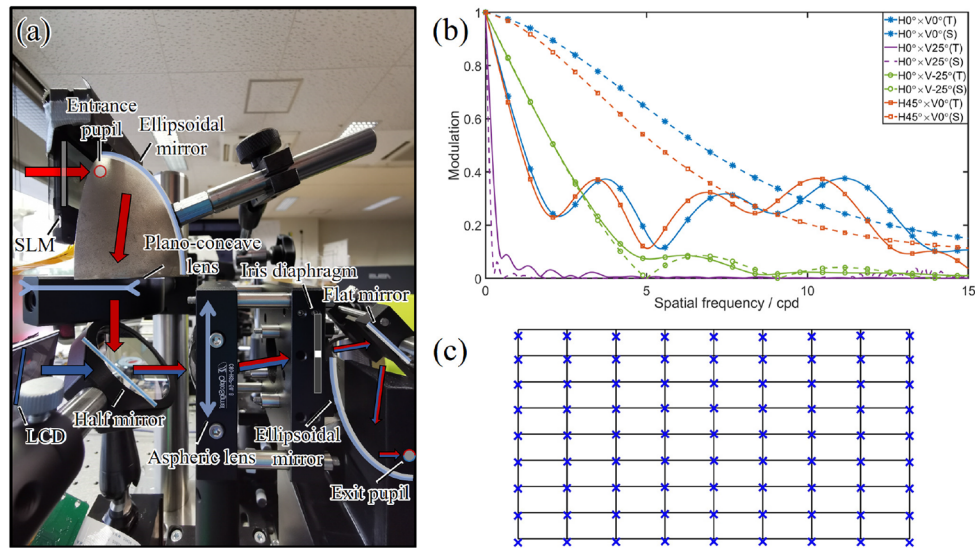


Fig. 4. (a) The bench-top prototype of the proposed method. The SLM is placed before the entrance pupil. A half mirror and a flat mirror are both tilted by 45°. Minimizing the vertical parallax requires the half mirror and the flat mirror to be rotated further, which is obstructed by redundant parts of other optical elements. (b) The MTF curves at different see-through FOVs and (c) the gird distortion within the see-through FOV of H95.3°×V52.9° are both calculated by ZEMAX. The pinhole aperture is set as 3mm, and the diffraction by the SLM is ignored in the calculation.

The modulation transfer function (MTF) curves for both the tangential (T) and the sagittal (S) planes at different see-through FOVs are simulated by ZEAMX, as shown in Fig. 4(b). The diffraction by the SLM is ignored in the simulation. The FOV of H0°×V0° is adjusted to achieve the best image quality, which reaches MTF30 at a spatial frequency of 11.5 cycle per degree (cpd). With the FOV expanding to H45°×V0°, the MTF curve slightly deteriorates. MTF30 here is achieved at the spatial frequency of 7.6cpd, which indicates that the proposed system

keeps stable image quality for each horizontal vision plane. In comparison, the MTF curve is easily influenced by the FOV shifting vertically. MTF30 for the FOV of $H0^\circ \times V-25^\circ$ occurs at the spatial frequency of 3.1cpd and even at 0.2cpd for the FOV of $H0^\circ \times V25^\circ$. Consequently, the proposed system only provides a sharp see-through view along with a band-like area around the designated vision, and the upper vision shows worse image quality than the lower vision. In practice, the transmissive LCD installed in the prototype further deteriorates the image quality due to a low fill factor (less than 50% typically). A transmitted beam through an activated LCD pixel is expected to have the intensity of the 1st order diffraction as high as 13.5% to the ideal zero diffraction [17]. Hence, the see-through view observed through a transmissive LCD is worse than the simulation.

The grid distortion within a FOV of $H95.3^\circ \times V52.9^\circ$ is shown as Fig. 4(c). Overall, the proposed system keeps minor distortion. A maximum distortion of -7.3% is observed at the central top vision. In order to show the image quality and distortion of the prototype more clearly, we used a smartphone camera of HUAWEI P30 Pro to record perceived images (see Supplement 1). Images shown in this paper are captured by another web camera for a wider FOV.

As discussed above, the LCD for projecting virtual images is tilted by 11.2° to attenuate the mismatch between the actual planar display panel and the ideal curved display surface. The experiment results with and without LCD tilt are shown in Fig. 5. A red teapot is used as the target image. Displayed images by the LCD are rendered in real-time by our algorithm (based on OpenGL API) with aberration compensation. A simple projection on the vertical pixels of the displayed image is conducted to render the 11.2° tilted image (blue labeled) from the initial format (orange labeled). The grey boundary corresponds to the designated FOV for virtual display ($H105^\circ \times V105^\circ$ in the experiment). With the see-through scene blocked, recorded images are taken with a FOV of $H95.3^\circ \times V52.9^\circ$, and the central view is focused. The two virtual images have the same distortion as the see-through view shown in Fig. 7. In order to highlight the difference between perceived virtual images with and without LCD tilt, we additionally put a

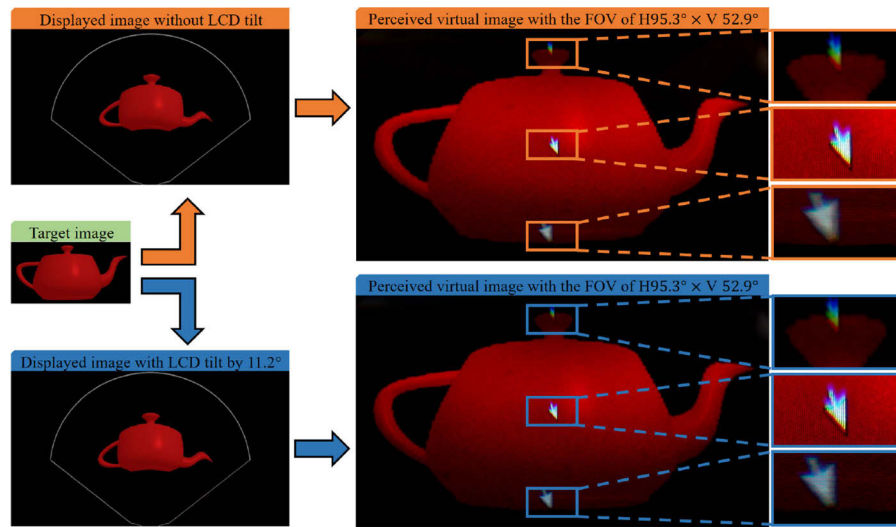


Fig. 5. Virtual images displayed with and without LCD tilt. A red teapot is given as the target image. Projected images on the LCD are separately shown as the non-tilt format (orange labeled) and 11.2° tilt format (blue labeled). Perceived images with a FOV of $H95.3^\circ \times V52.9^\circ$ are shown at the right, whose distortion is the same as the see-through view in Fig. 7. Zoom-in images of the upper, middle and lower areas are placed beside. A cursor is additionally put to highlight differences.

cursor at the top, middle, and bottom parts of each projected image to take zoom-in images, which are shown at the rightmost column. In comparison, a noticeable improvement of image resolution is observed in the lower vision when the LCD is tilted by 11.2° , while the middle and upper visions show a minor difference.

The entrance pupil and the exit pupil of the proposed system are given as projections of the pinhole by the former elements group and the latter ellipsoidal mirror, respectively. We calculated footprints of beams from various FOVs at the entrance pupil and the exit pupil with different pinhole apertures by ZEMAX. Footprint diagrams are shown in Fig. 6. In general, footprints at the entrance pupil and the exit pupil show the same dimensions. Besides the extension caused by directly increasing pinhole aperture, footprints also enlarge with the FOV shifting from $H0^\circ \times V-25^\circ$ to $H0^\circ \times V25^\circ$. In addition, circular footprints at $H0^\circ \times V0^\circ$ are projected into ellipses at $H45^\circ \times V0^\circ$. In the case that an SLM is placed at P_{o1} , blocking a pixel from the real scene requires all corresponding rays through the imaging system to be cut off. Thus occlusion-capable pixel sizes are directly determined by footprints of imaging beams at the P_{o1} plane, which is the same as footprints at the entrance pupil for parallel beams. Considering the pinhole aperture of 3mm, the occlusion-capable pixel sizes are 0.56mm @ $H0^\circ \times V0^\circ$, 0.40mm @ $H0^\circ \times V-25^\circ$, 0.96mm @ $H0^\circ \times V25^\circ$, and 0.68mm @ $H45^\circ \times V0^\circ$ (the average of the horizontal and vertical dimensions). We used a black teapot as the occlusion pattern and located it at different FOVs. A transparent film is hung before the prototype to slightly uniform the real scene brightness

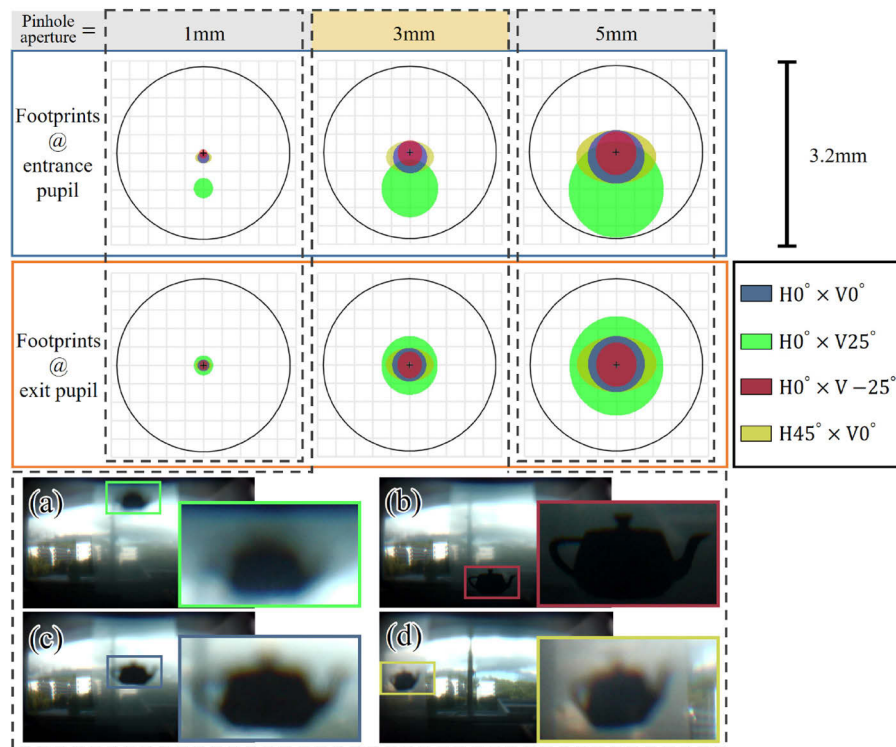


Fig. 6. Footprint diagrams at the entrance pupil and the exit pupil with different pinhole apertures are shown above. Beams emitted from four different FOVs are drawn with different colors. All circles in the figures indicate a range of 3mm. Our prototype works with the pinhole aperture of 3mm, a FOV of $H95.3^\circ \times V52.9^\circ$ is recorded. The occlusion performance by locating the same occlusion pattern at the (a) top, (b) bottom, (c) middle, and (d) left vision areas are shown below.

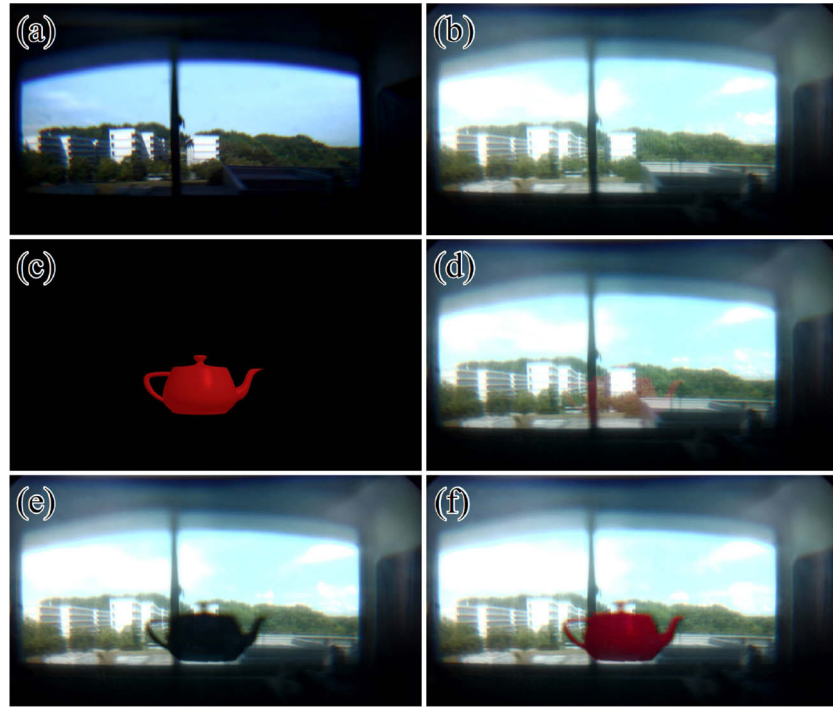


Fig. 7. Mutual occlusion is demonstrated with a FOV of $H95.3^{\circ} \times V52.9^{\circ}$. (a) The see-through scene taken without the transmissive SLM installed, and (b) the see-through scene taken with the SLM mounted before the entrance pupil. The camera is set with different exposure mode when takes the two photos. (c) The input virtual content, and (d) a transparent image is recorded by the camera. The occlusion pattern is shown as (e), and the recorded virtual image displayed with occlusion turns to be opaque in (f).

that distinguishes occlusion patterns from the background. Experiment results conducted by the prototype with the pinhole aperture of 3mm are shown in the figures below. A FOV of $H95.3^{\circ} \times V52.9^{\circ}$ is recorded in the figures. From Fig. 6(a) to Fig. 6(c), the occlusion pattern is moved to the top, bottom, and middle visions, respectively. With the vision moving downward, the occlusion performance is improved from a soft-edge occlusion-like level to a hard-edge occlusion-like level. Additionally, the occlusion pattern is moved to the leftmost of the FOV in Fig. 6(d). Despite that the occlusion pattern is similarly distorted as the real scene, the precision keeps stable while the contrast slightly decreases.

The footprints at the exit pupil in Fig. 6 depict the eye-box of the proposed system. In the case that the entire FOV is observed, the user's pupil overlaps the exit pupil, which gives the eye-box of 0.19mm, 0.56mm, and 0.96mm (taking the footprint of the central view as an average) of the system with pinhole apertures of 1mm, 3mm, and 5mm, respectively. Due to the narrow imaging beams, the focus state of the user's eyes impacts the image resolution minor. Users with eye problems such as low vision are allowed to have a clear see-through view without eyeglasses [16].

Fig. 7 shows mutual occlusion conducted by the prototype. The pinhole aperture is chosen as 3mm. A see-through scene without the transmissive SLM installed is shown as Fig. 7(a). After the SLM is installed, the image quality is deteriorated due to the severe diffraction caused by artifacts in the SLM panel, which is shown as Fig. 7(b). Notice that the camera was switched to a different exposure mode, so that makes Fig. 7(b) look brighter than Fig. 7(a). Figure 7(c) shows the input image of the rendering flow. The displayed image by the prototype without

mutual occlusion is shown as Fig. 7(d). The red teapot looks highly transparent on the bright background, which makes it mostly invisible (see [Visualization 1](#)). Furthermore, missed detail information, such as the lighting effect, reduces the realism of the teapot. Figure 7(e) shows the occlusion pattern displayed by the transmissive SLM. The pattern has gradient resolution along with the vertical direction. The virtual image displayed with mutual occlusion is shown in Fig. 7(f). With the background being blocked, both the framework and the lighting effect of the teapot are clearly perceived (see [Visualization 2](#)). Besides, the variation of pinhole aperture impacts system performance. A smaller pinhole brings better occlusion performance and image quality while it also darkens the image brightness conversely (see [Supplement 1](#)).

4. Conclusion

A PEM structure for realizing wide-view mutual occlusion in AR displays is proposed. The system is optional to conduct hard-edge occlusion or enhanced soft-edge occlusion by switching the position of the SLM. Bench-top prototypes with enhanced soft-edge occlusion are built, occlusion rendered in a wide FOV of $H122^\circ \times V74^\circ$ is demonstrated. A few issues need to be addressed in future research. The ellipsoidal mirrors increase NA of imaging systems but cause a larger form-factor at the same time. The relay lens pair achieves a sharp band-like vision and an overall small-distorted image but fails to focus the whole wide FOV. Although both planar SLM and LCD can be tilted to meet the curved focal surface of ellipsoidal mirrors, an optimized rendering algorithm and flexible displays are expected to improve system performance further.

Funding. Japan Society for the Promotion of Science (18H04116).

Disclosures. The authors declare that there are no conflicts of interest related to this article.

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.

Supplemental document. See [Supplement 1](#) for supporting content.

References

1. G.-Y. Lee, J.-Y. Hong, S. Hwang, S. Moon, H. Kang, S. Jeon, H. Kim, J.-H. Jeong, and B. Lee, "Metasurface eyepiece for augmented reality," *Nat. Commun.* **9**(1), 4562 (2018).
2. A. Wilson and H. Hua, "Design and demonstration of a vari-focal optical see-through head-mounted display using freeform alvarez lenses," *Opt. Express* **27**(11), 15627–15637 (2019).
3. A. Erickson, K. Kim, G. Bruder, and G. F. Welch, "Exploring the limitations of environment lighting on optical see-through head-mounted displays," in *Symposium on Spatial User Interaction* (2020), pp. 1–8.
4. A. Wilson and H. Hua, "Design of a pupil-matched occlusion-capable optical see-through wearable display," *IEEE Trans. Vis. Comp. Graph.* (2021), early access.
5. K. Kiyokawa, M. Billinghurst, B. Campbell, and E. Woods, "An occlusion capable optical see-through head mount display for supporting co-located collaboration," in *The Second IEEE and ACM International Symposium on Mixed and Augmented Reality, 2003. Proceedings*, (IEEE, 2003), pp. 133–141.
6. D. Dunn, C. Tippetts, K. Torell, P. Kellnhofer, K. Akşit, P. Didyk, K. Myszkowski, D. Luebke, and H. Fuchs, "Wide field of view varifocal near-eye display using see-through deformable membrane mirrors," *IEEE Trans. Visual. Comput. Graphics* **23**(4), 1322–1331 (2017).
7. A. Wilson and H. Hua, "Design and prototype of an augmented reality display with per-pixel mutual occlusion capability," *Opt. Express* **25**(24), 30539–30549 (2017).
8. 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. Graphics* **26**(5), 1871–1879 (2020).
9. Y.-G. Ju, M.-H. Choi, P. Liu, B. Hellman, T. L. Lee, Y. Takashima, and J.-H. Park, "Occlusion-capable optical-see-through near-eye display using a single digital micromirror device," *Opt. Lett.* **45**(13), 3361–3364 (2020).
10. T. Hamasaki and Y. Itoh, "Varifocal occlusion for optical see-through head-mounted displays using a slide occlusion mask," *IEEE Trans. Visual. Comput. Graphics* **25**(5), 1961–1969 (2019).
11. K. Rathinavel, G. Wetzstein, and H. Fuchs, "Varifocal occlusion-capable optical see-through augmented reality display based on focus-tunable optics," *IEEE Trans. Visual. Comput. Graphics* **25**(11), 3125–3134 (2019).
12. A. Maimone and H. Fuchs, "Computational augmented reality eyeglasses," in *2013 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, (IEEE, 2013), pp. 29–38.
13. 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. Visual. Comput. Graphics* **23**(11), 2463–2473 (2017).

14. J. Yang, W. Liu, W. Lv, D. Zhang, F. He, Z. Wei, and Y. Kang, "Method of achieving a wide field-of-view head-mounted display with small distortion," *Opt. Lett.* **38**(12), 2035–2037 (2013).
15. Y. Zhang, X. Hu, K. Kiyokawa, N. Isoyama, N. Sakata, and H. Hua, "Optical see-through augmented reality displays with wide field of view and hard-edge occlusion by using paired conical reflectors," *Opt. Lett.* **46**(17), 4208–4211 (2021).
16. Y. Zhang, N. Isoyama, N. Sakata, K. Kiyokawa, and H. Hua, "Super wide-view optical see-through head mounted displays with per-pixel occlusion capability," in *2020 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, (IEEE, 2020), pp. 301–311.
17. X. Wang, Y. Qin, H. Hua, Y.-H. Lee, and S.-T. Wu, "Digitally switchable multi-focal lens using freeform optics," *Opt. Express* **26**(8), 11007–11017 (2018).