




# Add-on Occlusion: Turning Off-the-Shelf Optical See-through Head-mounted Displays Occlusion-capable

Yan Zhang , Xiaodan Hu , Kiyoshi Kiyokawa , and Xubo Yang

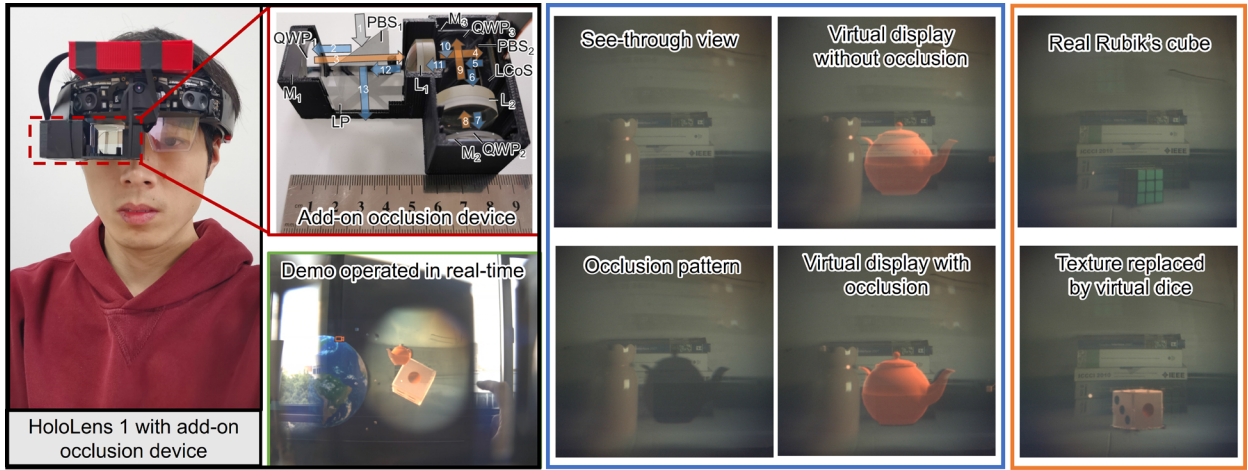


Fig. 1: **Left:** HoloLens 1 is assembled with the add-on occlusion device (right eye's view). The visor of HoloLens 1 is taken off, and the device is attached to the front surface of the waveguide combiner. The optical architecture is shown beside. The optical path is highly compressed by reusing optical elements by modulating light polarization between the p-direction (orange arrows) and the s-direction (blue arrows) repeatedly. A liquid crystal on silicon (LCoS) device is used to display the occlusion pattern, the driver board is put in the red box above. All the components are mounted in a 3D-printed housing, leading to a volume of  $8.5 \times 5.5 \times 3 \text{ cm}^3$ . See Fig. 3 for details of the optical structure. The occlusion pattern is rendered by post-processing virtual images streamed from HoloLens 1. A demo of object manipulation is shown in the image bounded by green lines. See the movie clip for the real-time operation with the demo. **Middle:** Mutual occlusion is demonstrated by the assembled system. The Spatial Awareness System of the Mixed Reality Toolkit (MRTK) is used to build the real-world mesh, and the occlusion pattern is generated with the virtual object in real-time. **Right:** A Rubik's cube is placed in the real scene, then transformed to be a dice by replacing the surface texture.

**Abstract**—The occlusion-capable optical see-through head-mounted display (OC-OSTHMD) is actively developed in recent years since it allows mutual occlusion between virtual objects and the physical world to be correctly presented in augmented reality (AR). However, implementing occlusion with the special type of OSTHMDs prevents the appealing feature from the wide application. In this paper, a novel approach for realizing mutual occlusion for common OSTHMDs is proposed. A wearable device with per-pixel occlusion capability is designed. OSTHMD devices are upgraded to be occlusion-capable by attaching the device before optical combiners. A prototype with HoloLens 1 is built. The virtual display with mutual occlusion is demonstrated in real-time. A color correction algorithm is proposed to mitigate the color aberration caused by the occlusion device. Potential applications, including the texture replacement of real objects and the more realistic semi-transparent objects display, are demonstrated. The proposed system is expected to realize a universal implementation of mutual occlusion in AR.

**Index Terms**—Augmented reality, near-to-eye displays, occlusion displays, head-mounted displays, diminished reality, color blending

## 1 INTRODUCTION

Augmented reality (AR) is a technology that allows users to interact with digital contents registered in the physical world in real-time. The

optical see-through head-mounted display (OSTHMD) is a promising platform to support daily AR applications since it provides users hands-free wearable experience [28]. Optical combiners in OSTHMDs transmit the light emitted from the physical world while reflecting or diffracting the light projected by the imaging systems to the users' pupils. The additive display feature of OSTHMDs benefits users by obtaining latency-free physical world visions with rich graphics information (frame rate, color gamut, and contrast ratio etc.) of the real scene kept, yet results in the severe color blending of the displayed virtual image pixels and overlapped real scene pixels [7, 8, 13, 27]. Virtual contents displayed by OSTHMDs, hereby, suffer weaker perceptual cues, lower contrast ratio, and a narrower color gamut with the illuminance of the real background increase [6, 30–32].

Occlusion-capable optical see-through head-mounted displays (OC-OSTHMDs) are proposed to address this issue. Different from common OSTHMDs that pass the real image to users' pupil directly, OC-OSTHMDs build optical architectures to manipulate the Alpha channel

of the real image by integrated spatial light modulators (SLMs). Real scene pixels overlapped by the displayed virtual image are erased by setting the Alpha channel of the corresponding area to zero value, so that it prevents the contrast ratio and color gamut deterioration of the virtual image from color blending with the real background. Furthermore, the registered virtual object in the AR scene looks more realistic since mutual occlusion with real objects is supported according to the depth information.

Although state-of-the-art designs achieve more appealing performance with a varifocal occlusion feature and the wide field of view (FOV), the practical application of OC-OSTHMDs is hindered by the challenge of reducing the form-factor. The basic imaging system of common OST-HMDs projects digital images shown by a display panel to users' eyes. As a comparison, an optical architecture of OC-OSTHMDs firstly generates an inverted image of the input real scene at the SLM plane to implement the per-pixel occlusion, then erects the processed real scene with a relay lens group [15]. A following lens group is added to function the projection of both the erected real image and the displayed digital image to users' eyes as in common OSTHMDs. As a consequence, an OC-OSTHMD system becomes more costly on the budget and the system volume than common OSTHMDs. What is more, monochromatic optical components, such as holographic optical elements (HOEs) [1, 39] and waveguide combiners [4, 24, 25], downsize the imaging systems of OSTHMDs efficiently, yet suffer severe color aberration in OC-OSTHMDs that work with the visible spectrum. Meanwhile, commercial OSTHMDs have demonstrated to support passable visibility under indoor illumination [6] and the enhanced color appearance and spatial perception of virtual contents via the software approach [5, 10, 19, 22, 38]. OC-OSTHMDs with the built-in occlusion feature may be less critical but highly costly for some indoor AR applications. Integration of the occlusion feature and OSTHMDs in an efficient way is still challenging.

Targeting to facilitate a universal application of the occlusion feature among the OSTHMD platform, we propose add-on occlusion in this paper. Distinguished from the current approach that includes the occlusion function and the virtual display in an OC-OSTHMD device, the proposed method displays virtual images by an off-the-shelf OSTHMD, and per-pixel occlusion on the real scene is achieved by a separated device. The external occlusion device is built in a 3D-printed housing, achieving a maximum FOV of  $H28.7^\circ \times V16.4^\circ$ . A wearable prototype is assembled by attaching the occlusion module on the HoloLens 1 device, which is shown in the left image of Fig. 1. The occlusion pattern is generated from the streamed virtual image of HoloLens 1, then calibrated to match the displayed virtual contents in the AR scene. As a result, the prototype makes HoloLens 1, which is initially an OSTHMD, equivalently realize virtual display with per-pixel occlusion as an OC-OSTHMD.

The primary contributions of the present paper are highlighted as follows:

- 1) We firstly demonstrate the feasibility of realizing per-pixel occlusion on common OSTHMDs via a novel approach of attaching an external device to an off-the-shelf product, showing the potentiality of universal applications of the occlusion function.
- 2) We propose a compact design for the external occlusion device. The fabricated occlusion module is attached on a HoloLens 1 device, proving a wearable form-factor for the daily AR usage.
- 3) We realize the real-time occlusion with a virtual display on the HoloLens 1 device by post-processing, making software development kits (SDKs) designed for common AR applications compatible with the add-on occlusion system.
- 4) We propose a color correction method for the fabricated occlusion module to compensate the color aberration of the processed real scene.
- 5) We discuss potential applications of per-pixel occlusion on the HoloLens 1 platform. The texture replacement of real objects and a

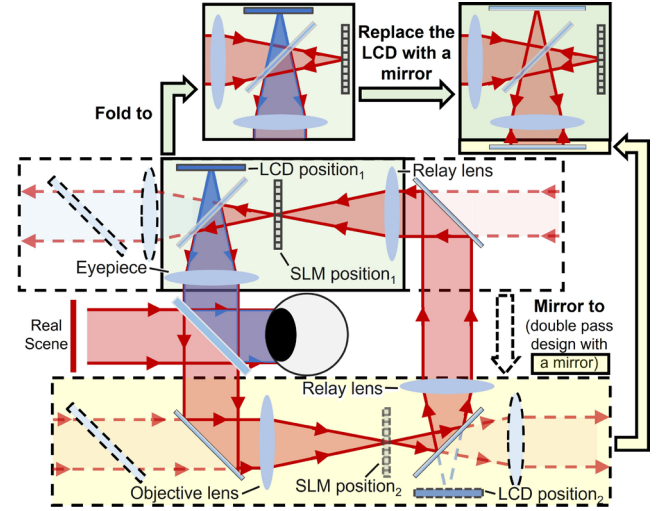


Fig. 2: The ring-shaped structure of OC-OSTHMDs proposed by Kiyokawa et al. [15]. A 4-f structure composed of an objective lens, a relay lens pair, an eyepiece, and an SLM is essential for realizing per-pixel occlusion. The ray bundle from the real scene is processed by the SLM, then merges with the light from the LCD. Semi-transparent optical elements drawn in dash lines are the optical conjugation of physical optical elements. The whole architecture can be separated into two mirrored parts in the dash rectangles, allowing either part that is physically built to form another part optically by using the double-pass design [14, 34]. The lens pair in the green rectangle can be further folded into half volumes. The downsized structure to realize per-pixel occlusion in the proposed add-on occlusion device is achieved by using the double-pass design and the structure folding simultaneously. See Fig. 3 for the detailed optical architecture.

more vivid graphics of semi-transparent virtual objects are demonstrated with the prototype.

## 2 RELATED WORK

### 2.1 OC-OSTHMDs with per-pixel occlusion

The OC-OSTHMD is firstly proposed by Kiyokawa et al. to tackle the challenge of presenting mutual occlusion between the registered virtual object and real objects correctly in the AR scene [16, 17]. The typical optical architecture to realize per-pixel occlusion is introduced with their ELMO-4 system, as shown in Fig. 2 [15]. The red ray bundle emitted from the real scene is focused by an objective lens to be pixels on an inverted image, a relay lens pair is required to erect the image, then an eyepiece is added to recover the transmitted real scene pixels to be the initial ray bundle that is suitable for human perception. An SLM is placed either to overlap the inverted image generated by the objective lens (SLM position<sub>2</sub>) or the erected image generated by the relay lens group (SLM position<sub>1</sub>). Thus, the transmitted optical power of each real image pixel is controlled by modulating the grayscale of the pattern shown on the SLM. Considering typical SLMs with an 8-bit depth grayscale, an OC-OSTHMD is able to set the Alpha channel of the real scene within the 0 – 255 range.

A few works are done based on the structure of the ELMO-4 system. Wilson and Hua achieve better image quality by replacing the transmissive LCD panel in the ELMO-4 system with a liquid crystal on silicon (LCoS) device [33]; Krajancich et al. downsize the system by utilizing a digital micro-mirror device (DMD) [18], the optimized design with smaller form-factor and eliminated horizontal viewpoint offset is done by Ju et al. [14]; Hamasaki and Itoh introduce the varifocal occlusion feature into OC-OSTHMD by mechanically sliding the lenses in the optical system [9], and Rathinavel et al. realize the same function by implementing tunable lenses [29]; Zhang et al. demonstrated an enlarged FOV of the OC-OSTHMD by alternatively using concave mirrors as the objective lens and the eyepiece [35–37]. Very soon, Wilson

OC-OSTHMDs	Components (for single eye occlusion, SLMs and display panels not included)	System layout	Potential volume	Parallax	Highlighted feature
Kiyokawa et al. (2003) [15]	4 lenses, 1 prism, 1 PBS, 3 flat mirrors	ring-shaped design	8f	parallax-free	none
Wilson and Hua (2017) [33]	3 lenses, 1 prism, 1 PBS, 2 flat mirrors	light path folding	4f	longitudinal parallax ↓	none
Hamasaki and Itoh (2019) [9]	3 lenses, 1 beam splitter, 1 actuator	naive layout	4f	longitudinal parallax lateral & longitudinal	varifocal occlusion
Rathinavel et al. (2019) [29]	2 tunable lenses, 4 lenses, 1 PBS	naive layout	4f +	longitudinal parallax	varifocal occlusion
Ju et al. (2020) [14]	1 lens, 1 prism, 1 PBS, 1 LP, 1 flat mirrors, 2 QWPs	double-pass design	3f	longitudinal parallax ↓	none
Zhang et al. (2021) [35]	2 parabolic mirrors, 4 lenses, 1 prism, 1 PBS, 1 pinhole, 2 flat mirrors	ring-shaped design	8f ↑	longitudinal parallax ↓	wide FOV
Zhang et al. (2021) [36,37]	2 ellipsoidal mirrors, 2 lenses, 1 half mirror, 1 pinhole	light path folding	4f ↑ +	lateral & longitudinal parallax ↓	wide FOV
Wilson and Hua (2021) [34]	2 lenses, 1 half mirror, 1 PBS, 2 LPs, 1 QWP	double-pass design	3f	parallax-free	none
Our system	2 lenses, 2 PBSs, 3 QWPs, 1 LP, 3 flat mirrors	folded ring-shaped & double-pass design	2f +	parallax-free	add-on design

Table 1: Comparison of form-factor and parallax of OC-OSTHMDs with per-pixel occlusion capability. SLMs and display panels are not included. Potential volumes are counted as multiples of the focal length of lenses (e.g., “8f” indicates eight times the focal length). “+” indicates extra spacings (shorter than 1f) of optical elements necessary for building OC-OSTHMDs, and “↑” indicates any optical element in OC-OSTHMDs that has extraordinary size. The parallax of OC-OSTHMDs is marked with “↓” when it is reduced by utilizing optimized system layouts.

and Hua introduce a more compact layout of the OC-OSTHMD by using a double-pass design of lenses, a significant enhancement of the contrast ratio and color gamut of the displayed virtual image by enabling per-pixel occlusion is also demonstrated with their wearable prototype [34].

In summary, the feasibility and the importance of including the occlusion function in OSTHMDs have been well demonstrated. Nevertheless, mainstream OSTHMDs less benefit from the technology since the requirement of the dedicated optical architecture binds the occlusion function with OC-OSTHMDs tightly. The proposed design is, hereby, distinguished from all the previous works for the novel approach of building an external device for off-the-shelf OSTHMDs to obtain the per-pixel occlusion function.

## 2.2 Parallax-free design for OC-OSTHMDs

OC-OSTHMDs process the real scene by inputting it into an imaging system with the occlusion function and generating a recovered image for users perceiving. Where the image inputs and outputs the system are the entrance pupil (virtual viewpoint) and the exit pupil (real viewpoint), respectively. A user putting his/her physical eye on the real viewpoint perceives the physical world as if the physical eye was placed on the virtual viewpoint. A mismatch between the physical eye and the virtual eye causes parallax, leading to confusing spatial perception of the physical world of users. Therefore, a parallax-free design is a key for an OC-OSTHMD to provide users with correct spatial perception. Among all the OC-OSTHMDs proposed, the first one by Kiyokawa et al. [15] and the latest one by Wilson and Hua [34] achieve the parallax-free feature by using the ring-shaped structure and the double-pass structure, respectively.

A couple of OC-OSTHMDs realize the viewpoint overlapping by avoiding the usage of the lens-based optical structure. Maimone and Fuchs propose an OC-OSTHMD by using a stack of transmissive LCD panels to realize occlusion via a light field way [20]. The virtual viewpoint and the real viewpoint are inherently overlapped in this design since no lens structure exists. However, ray bundles emitted from real scene pixels fail to be blocked precisely by the SLMs for the absence of lens focusing. Hence, the design implements occlusion in lower precision than typical OC-OSTHMDs with per-pixel occlusion. Itoh et

al. propose a video compensation method to increase the precision of the lens-free OC-OSTHMD, while considerable latency is introduced for computing compensation images [12].

The add-on occlusion device presented in this paper also keeps the advantage of parallax-free of OC-OSTHMDs. As shown in Fig. 2, the optical system is derived from the ring-shaped architecture proposed by Kiyokawa et al. [15]. The original system is separated into two mirrored parts in dashed rectangles. An equivalent system can be achieved by physically building either of the two parts, then optically duplicating another part by using the double-pass design [14,34]. What is more, the lens pair in the green rectangle can be compressed into a half volume by reflecting the light back from the SLM plane. As result, the final design supports per-pixel occlusion while taking 1/4 of the original volume. The detailed optical design is shown in Fig. 3, where polarizing optical elements are used for the double-pass design and the folded lens layout.

Table 1 summarizes potential volume and parallax of existing OC-OSTHMDs, a few systems that are not ready for near-to-eye displays are not included [3,18]. The potential volumes are counted as the multiples of the focal length of lenses since lens systems are included in all OC-OSTHMDs to realize per-pixel occlusion. We further assume that various devices use the same lenses with a common numerical aperture (NA) of 0.5. Hence, the rest of the optical elements in OC-OSTHMDs, such as flat mirrors and polarizing beam splitters (PBSs), are counted as a volume of 1f for each one. For example, the ring-shaped OC-OSTHMD is marked as a 8f volume since it needs the 4f lens structure for implementing occlusion while having another 4f space to locate mirrors for making the device parallax-free. In comparison, the proposed system realizes per-pixel occlusion within the smallest volume while keeping parallax-free.

## 2.3 Projector-based occlusion displays

Projector-based augmented reality systems utilize external projectors to modulate environment illumination of the AR scene to realize mutual occlusion. Bimber et al. develop a stationary system that locates the AR scene in a covered space. Environment illumination is only given by a video projector, so that occlusion to real objects is equivalently realized by cutting off the light towards the object surface [2]. Maimone et al. propose a similar way, which enforces the illumination of non-occluded



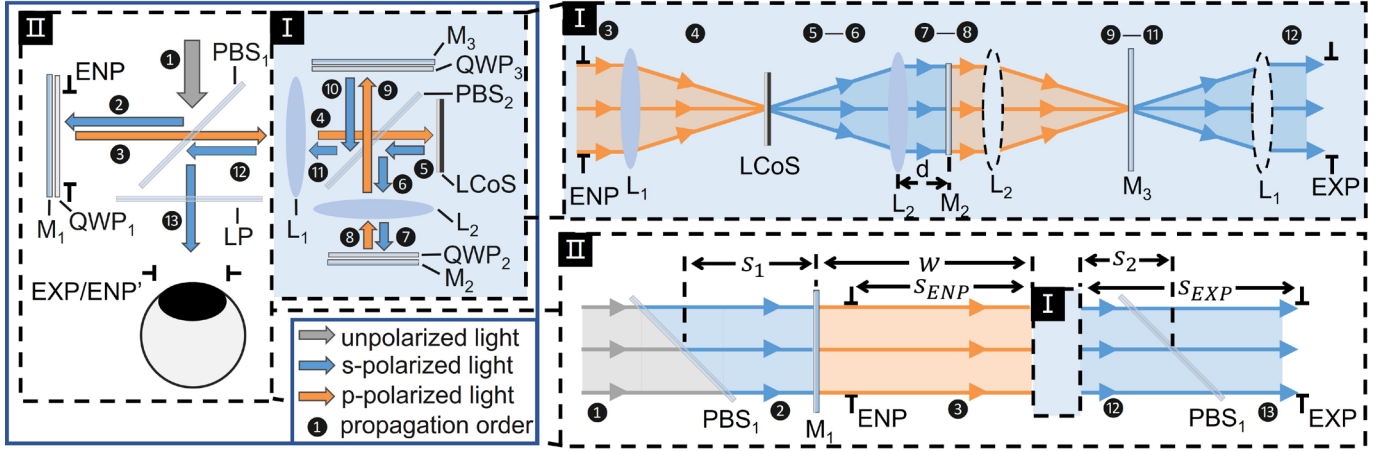


Fig. 3: Optical design of the proposed system. The unpolarized light, s-polarized light, and p-polarized light are drawn in grey, blue, and orange, respectively. **Left:** The entire system with the real scene light propagating with a sequence by (1) to (13). Part I and part II are the structures for realizing per-pixel occlusion with an LCoS and achieving the parallax-free feature, respectively. **Top-right:** The equivalent system of part I. Propagation light repeatedly passes through convex lenses  $L_1$  and  $L_2$ , forming a typical optical architecture of OC-OSTHMDs. The entrance pupil (ENP) and the exit pupil (EXP) are formed symmetrically by the mirror  $M_2$ , which is the aperture stop in the structure. **Bottom-right:** The equivalent system of part II. The proposed system achieves the parallax-free feature by placing the mirror  $M_1$  at ENP.

areas in the real scene to achieve the mutual occlusion of a virtual portrait for a telepresence system [21]. Inami et al. utilize a system consisting of a head-mounted projector (HMP) device and a haptic object covered by retro-reflective material. The retro-reflective screen reflects the projected synthetic image and blocks the backward real scene so that it realizes mutual occlusion in the AR scene [11].

In comparison, the proposed system realizes mutual occlusion by similarly using external devices as in projector-based systems but is much more universal since it only modulates the light approaching users' eyes with a wearable size.

### 3 OPTICAL DESIGN

#### 3.1 System layouts

The optical design of the proposed system is shown in Fig. 3. The entire device bounded in the left blue rectangle is composed of two parts, which are the part I for implementing the per-pixel occlusion to the real scene and the part II for eliminating the viewpoint offset, respectively. The incident light from the real scene (drawn as the grey arrow) is firstly linearly polarized by a PBS, then passes through the system with a sequence indicated from (1) to (13). Combinations of a mirror ( $M_{1-3}$ ) and a quarter wave plate ( $QWP_{1-2}$ ) switch the light polarization between s-direction (drawn as blue arrows) and p-direction (drawn as orange arrows). Thus, propagation light transmitted from a PBS is changed to be reflected at the same PBS surface after being guided backward by the mirror attached with a QWP, and vice versa. Consequently, the cubic space taken by  $PBS_2$  in part I is used repeatedly to achieve an optical path with four times length, leading to a highly compressed system layout.

The equivalent system of part I is shown in the beside dotted box with a blue background. The blue and orange ray bundles indicate the s-polarized light and p-polarized light, respectively. An LCoS is placed at the back focal plane of a convex lens  $L_1$ . Per-pixel occlusion on the input real scene is realized by shuttering pixels on the LCoS. Although only a convex lens  $L_2$  is physically placed after  $L_1$ . Both  $L_1$  and  $L_2$  are repeatedly used based on a double-pass structure, so that part I works equivalently as the typical optical architecture of OC-OSTHMDs to output an erected real scene with mutual occlusion. The physical lens pair of  $L_1$  and  $L_2$  are fixed by sharing the same focus, generating an afocal system with a unit magnification of the input real image by the equivalent 4-f structure. The  $M_2$  works as the aperture stop of the imaging system. Thus, the entrance pupil (ENP) and the exit pupil (EXP) are located symmetrically with positions equally determined by the distance of  $M_2$  to  $L_2$ .

Although part I is all the need for realizing the per-pixel occlusion, part II is still necessary for eliminating the viewpoint offset for providing an appealing AR experience. Light in part II passes through ENP and EXP of part I sequentially. A linear polarizer (LP) is placed before EXP to block the stray light reflected from the LCoS. An equivalent structure of Part II is shown in the bottom-right dotted box. The unpolarized light is similarly drawn as a grey ray bundle. Letting the distance from  $PBS_1$  to  $M_1$  and  $L_1$  be  $S_1$  and  $S_2$ , respectively, the overlap of the conjugated entrance pupil ENP' and EXP requires:

$$\begin{aligned} 2S_1 + S_2 - S_{ENP} &= S_{EXP} - S_2 \\ S_1 + S_2 &= \frac{S_{ENP} + S_{EXP}}{2} \end{aligned} \quad (1)$$

where  $S_{ENP}$  and  $S_{EXP}$  are the distances of ENP and EXP to  $L_1$ . Since ENP and EXP are symmetrically projected by the aperture stop  $M_2$  via the same lens pair in part I, which can be derived as:

$$S_{ENP} = S_{EXP} = f_1 + \frac{f_2 - d}{f_2} \cdot \frac{f_1^2}{f_2} \quad (2)$$

where  $f_1$  and  $f_2$  are the focal lengths of  $L_1$  and  $L_2$ , respectively, and  $d$  is the distance between  $L_2$  and  $M_2$ . Besides, the distance between  $M_1$  and  $L_1$  is expressed as:

$$W = S_1 + S_2 \quad (3)$$

As a consequence, the constraint for a parallax-free design can be derived by substituting Eq. 2 and Eq. 3 into Eq. 1:

$$W = S_{ENP} = S_{EXP} \quad (4)$$

indicating that the parallax-free merit is guaranteed in the proposed design as long as  $M_1$  is placed at EXP.

#### 3.2 Optical performance analysis

The modulation transfer function (MTF) curves both for the tangential (T) and the sagittal (S) planes within a horizontal FOV of  $20^\circ$  are calculated by ZEMAX (one of the most used optical simulation software), as is shown in Fig. 4(a). The central FOV reaches MTF30 (where the average modulation for the T and S planes drops to 0.3) at the spatial frequency of 26.5 cycles per degree (cpd). A FOV of  $H5^\circ$  roughly keeps the performance. With the FOV enlarged to  $H10^\circ$ , the MTF curves deteriorate significantly, which reaches MTF30 @ 6.5cpd. MTF30 for the FOV approaching  $H20^\circ$  can further decrease to the

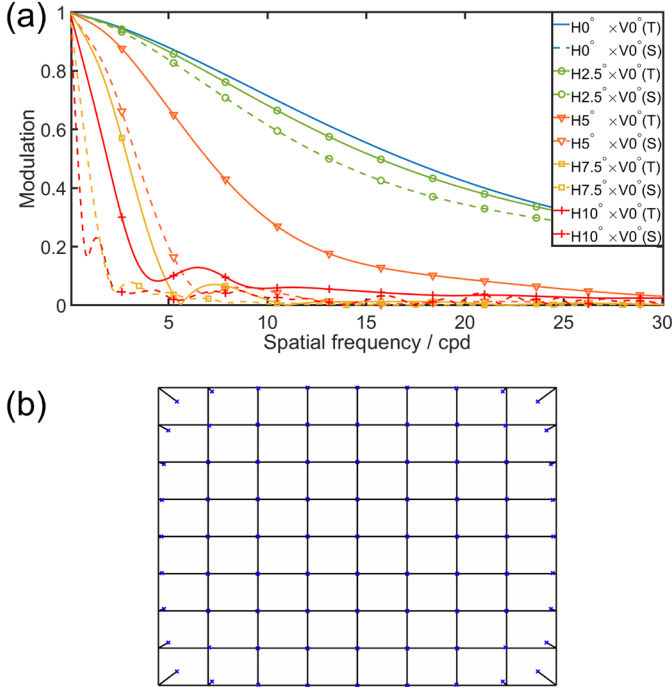


Fig. 4: (a) The calculated MTF curves at the tangential (T) and sagittal (S) planes of the proposed add-on occlusion device within horizontal FOVs rising to  $20^\circ$ . (b) Grid distortion of the add-on occlusion device within a FOV of  $H20^\circ \times V15^\circ$ .

spatial frequency of 2cpd, indicating a blurred real scene would be perceived when the FOV exceeds  $H10^\circ$ . Additionally, the calculated grid distortion within a FOV of  $H20^\circ \times V15^\circ$  is shown in Fig. 4(b). The central view of the proposed add-on occlusion device expects to provide a distortion-free image, and the maximum distortion is calculated as  $-9.3\%$  at the corners of the FOV.

### 3.3 Versatility of the device

The add-on occlusion device is fixed before the optical combiner when being integrated with an OSTHMD, realizing per-pixel occlusion on the real scene while allowing the OSTHMD to display virtual images as usual. As a consequence, the proposed method is compatible with any OSTHMDs in theory. However, the versatility of the add-on occlusion device depends on the adjustable eye relief range, which is the distance from LP to EXP in Fig. 3, and the specific optical architectures of OSTHMDs. A feasible integration of an add-on occlusion device and an OSTHMD requires exit pupils of both systems overlapped. Considering  $L_1$  and  $L_2$  with focal lengths of 30mm and 35mm, respectively, a maximum  $S_{EXP}$  of 56mm can be achieved by setting  $d = 0$ mm, giving the adjustable eye relief range up to 46mm (assuming a lens aperture of 10mm for  $L_1$ ). Thus, the add-on occlusion device is able to be used for OSTHMDs with common eye reliefs (10 - 20mm) and optical combiners in moderate thickness, such as HoloLens 1 & 2, Magic Leap 1, and Lumus Maximus with waveguide combiners, Moverio with the free-from optical combiner, and Nreal Light built on birdbath optics, while it does not work on OSTHMDs that utilize huge curved mirrors as optical combiners (e.g., NorthStar). What is more, the virtual display can be integrated into the add-on occlusion device by replacing  $M_3$  with a half mirror and fixing a display panel at the same plane, allowing an OC-OSTHMD with an eyeglass-like volume to be built.

## 4 PROTOTYPING AND MUTUAL OCCLUSION DEMONSTRATION

### 4.1 Prototype of the add-on occlusion device

The prototype of the add-on occlusion device is shown in Fig. 1. A 3D-printed housing with a volume of  $8.5 \times 5.5 \times 3 \text{ cm}^3$  is fabricated to install optical elements and the LCoS. Aspherized achromatic lenses

with the same aperture of 25mm and focal lengths of 30mm and 35mm are utilized as  $L_1$  and  $L_2$  (#49662 and #49663, Edmund Optics). Wire grid polarizing films (#20235, Edmund Optics) are used both for the PBS and LP since they have spectrally flat performance at the angle of incidences up to  $45^\circ$ . Zero-order QWPs are used for the prototype to reduce the chromatic aberration of the transmitted real scene. The LCoS has Full HD resolution, a contrast ratio of 2000:1, and an active area of  $0.69''$ . Different from most of OC-OSTHMDs that only used the PBS to realize occlusion by the LCoS, PBSs in the proposed system are multiply used to achieve a compressed optical path. The transmittance and reflectance of the PBS are, thereby, crucial for the device to keep a bright see-through view. Given with an average reflectance of 90% for s-polarized light and an average transmittance of 89% for p-polarized light of PBSs (@ the incident angle of  $45^\circ$ ) and a reflectance of 75% of LCoS @ 532nm, the prototype achieves a see-through efficiency of 19.5%.

### 4.2 Occlusion pattern rendering with HoloLens 1

The correct mutual occlusion between the virtual objects and the physical world requires real-time depth sensing and self-localization. The latter is typically achieved by simultaneous localization and mapping (SLAM) technology. Fortunately, HoloLens 1 provides an easily accessed way to implement SLAM system by using the Mixed Reality Toolkit (MRTK). We build a demo of the proposed system by using Unity (Version 2019.4) with MRTK v2. Paired Unity projects are operated on HoloLens 1 and a local PC simultaneously. The virtual scene is rendered by the HoloLens 1 project with the Spatial Awareness system enabled, allowing the occlusion of displayed virtual objects by the captured mesh cloud of the surrounding environment to be implemented in real-time. Frames rendered by the HoloLens 1 project are not only used for displaying but also sent to the local PC connected with the add-on occlusion device over the local area network for generating occlusion patterns. The peer-to-peer real-time streaming is realized by importing MixedReality-WebRTC SDK into both the Unity projects on HoloLens 1 and the local PC. In detail, virtual images rendered in HoloLens 1 are coded into YUV frames, then sent to a server (operated by the node-dss repository) that also runs on the local PC. The stored frames in the node-dss server are grabbed by the MixedReality-WebRTC component of the Unity project for rendering occlusion patterns, then processed by the customized shader as 2D textures. An occlusion pattern is generated by binarizing the texture pixels in the fragment shader stage:

$$[R \quad G \quad B] = \begin{cases} [1 \quad 1 \quad 1] & \text{grayscale} < 0.09 \\ [0 \quad 0 \quad 0] & \text{otherwise} \end{cases} \quad (5)$$

where the threshold of 0.09 is an imperial value since HoloLens 1 does not render the background of the virtual scene as full black. Meanwhile, the LCoS for displaying the occlusion pattern is placed at the back focal plane of the objective lens  $L_1$  as shown in Fig. 3, requiring the rendered occlusion pattern to be rotated by  $180^\circ$ . In addition, the occlusion pattern is enlarged by 1.05 times to overlap with the focused real image by  $L_1$ . Fig. 5 shows the streamed virtual scene of a teapot and the occlusion pattern generated by the local PC. The occlusion pattern is only flipped left and right because the LCoS is physically inverted in the prototype.

### 4.3 System calibration

In order to assemble the add-on occlusion device on HoloLens 1, we take off the visor and attach the prototype on the front surface of the waveguide combiner, so that the user's pupil is allowed to put in both the eye-boxes of the prototype and HoloLens 1. The occlusion pattern displayed by the add-on occlusion device is expected to share the same world coordinate as the virtual image space by HoloLens 1, so that it overlaps with the virtual object in the user's vision. However, the mismatch shown in the bottom-left image of Fig. 5 easily occurs (the white dot is the pointer of HoloLens 1, and it also appears in all the following figures), since the apparent size of a virtual object in the user's vision is different from that in the image space and precise assembly of the prototype with HoloLens 1 is extremely difficult. Therefore,

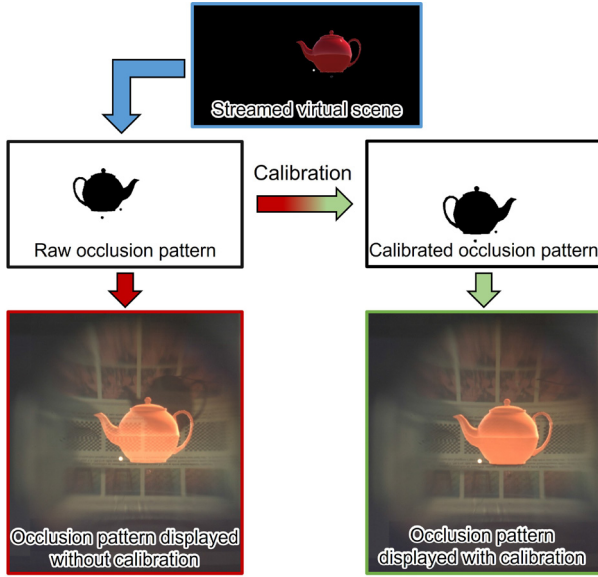


Fig. 5: Calibration of the proposed system. A raw occlusion pattern is generated from the streamed virtual scene from HoloLens 1, suffering the mismatch to the displayed teapot. Then, the calibrated occlusion pattern is generated to achieve an exact overlap with the teapot.

we further calibrate the system after a physical alignment by using the customized shader for rendering the occlusion pattern. During the calibration, the virtual object is placed in the central vision to minimize the distortion on the occlusion pattern caused by the optical system of the add-on occlusion device. Material properties are defined in the shader to leverage the projection from the raw occlusion pattern pixels to the calibrated occlusion pattern pixels output by the fragment shader. Hence, the orientation and dimension of the occlusion pattern can be precisely adjusted in real-time by manually changing the defined properties' values in the inspector window of the Unity project, making the occlusion pattern visually overlap with the virtual image, as shown in the green rectangle in Fig. 5.

#### 4.4 Mutual occlusion demonstration

Since the invisible part of a virtual object blocked by real objects is omitted during the virtual image rendering, the invisible part of the virtual object presents natural mutual occlusion with the calibrated occlusion pattern enabled. The demonstration of mutual occlusion by the prototype with HoloLens 1 is shown in the middle-left of Fig. 1. The red teapot is placed between a pink vase and a stack of books. A semi-transparent teapot is perceived when only the virtual display by HoloLens 1 is activated. Then, an occlusion pattern is shown to block the books in the background. As a result, the teapot shows higher visibility and better lighting effect with the occlusion pattern enabled.

### 5 COLOR CORRECTION

#### 5.1 Color aberration in the proposed system

Benefiting from the massive usage of the polarizing optical elements, the proposed add-on occlusion device compresses the optical path efficiently, so that it builds in a more compact volume than previous OC-OSTHMDs. However, most polarizing optical elements work well with designed wavelengths but show much worse performance when the incident light has a broader spectrum. As shown in Fig. 3, the linear polarized light is generated by the LP of the prototype, and the polarized light passes through the two PBSs in the system by multiple switching the polarization direction with QWPs. Nevertheless, QWPs used in the prototype have a designed wavelength of 532nm. The propagation light with the visible spectrum suffers varied retardance linearly proportional to its wavelength shifting from 532nm, which means only the light with the designed wavelength (green light) completes the polarization

switches well, while light with the shifted wavelengths (red light and blue light) becomes partially polarized light after transmitted by QWPs. As a result, green light from the real scene transmits the proposed system efficiently, yet red and blue light are mostly blocked by the PBSs, resulting in the see-through view with color aberration shown in Fig. 1.

An intuitive way to solve the color aberration problem is the implementation of achromatic optical elements in the proposed add-on occlusion device. However, the higher cost of the achromatic optical elements may negatively impact the universal applications of the device (e.g., an achromatic QWP has a price twice higher than common products).

Alternatively, the software approach seems promising to mitigate the color aberration issue in the proposed system. Langlotz et al. have demonstrated a color correction method that reduces the color aberration of virtual images caused by color blending with the real background, where the optimal virtual image displayed by an OSTHMD is computed in terms of the captured background image and the target virtual image [19]. Considering the real scene presented by the see-through view of the add-on occlusion device as a pixelated image, a software-based color correction method can be similarly implemented. Nevertheless, unlike virtual images that are entirely projected by OSTHMDs, real scenes are presented by the physical world, making the direct modulation of pixel values infeasible.

Therefore, the color correction method introduced in the following section modulates the real scene pixel values in an indirect way. A computed virtual image for implementing color correction is overlaid on the original real scene by the HoloLens 1, blending into an optimal real scene with mitigated color aberration for the user perceiving. A similar approach of overlaying a computed virtual image onto the real scene is utilized to modulate the brightness of the real scene by Hiroi et al. [10]. In comparison, our method focuses on modulating the color instead of the brightness of the real scene.

#### 5.2 HoloLens-based color correction

The procedure of color correction for the proposed system is shown in Fig. 6. An X-rite color checker placed under common indoor illumination is used as the experiment target. Fig. 6 (a) shows the photo of the color checker directly taken by an industry camera with sRGB color space, which is the target image for the color correction, and Fig. 6 (c) shows the scene image with color aberration captured through the add-on occlusion device. With knowing (R,G,B) triplets of each pixel on the target and the scene image, color aberration can be completed by letting the latter to approach the former as closely as possible. However, triplets in the sRGB color space follows a non-linear projection from the light intensity, so that a linear processing with the capture triplet values for color correction is hindered. Fortunately, the CIE1931 xyY color space is built on the linear mapping from the light intensity to the perceived color, which requires minimum computation to implement the color correction. In order to get the color coordinate of a pixel in the xyY color space, the CIE1931 tristimulus XYZ needs to be calculated first, which can be done by using:

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = M_T \times \begin{bmatrix} R \\ G \\ B \end{bmatrix} \quad (6)$$

where  $M_T$  is a  $3 \times 3$  transform matrix [26]. Normalized tristimulus can be calculated from triplets  $[R_r \ G_r \ B_r]^T$  and  $[R_s \ G_s \ B_s]^T$  of the target image and the scene image. Then, the xy coordinates of the pixel color for both images in the CIE1931 xy chromaticity diagram are calculated with  $x = \frac{X}{X+Y+Z}$  and  $y = \frac{Y}{X+Y+Z}$ . As shown in Fig 6 (d), the black triangle and the red triangle are drawn by connecting the xy coordinates of the primary colors captured in the target image and the scene image. Perfect color correction, hereby, expects to move each pixel in the red triangle to the corresponding coordinates in the black triangle.

Despite normalized tristimulus being achieved with Eq. 6, the absolute tristimulus still needs to know for manipulating the xy coordinates



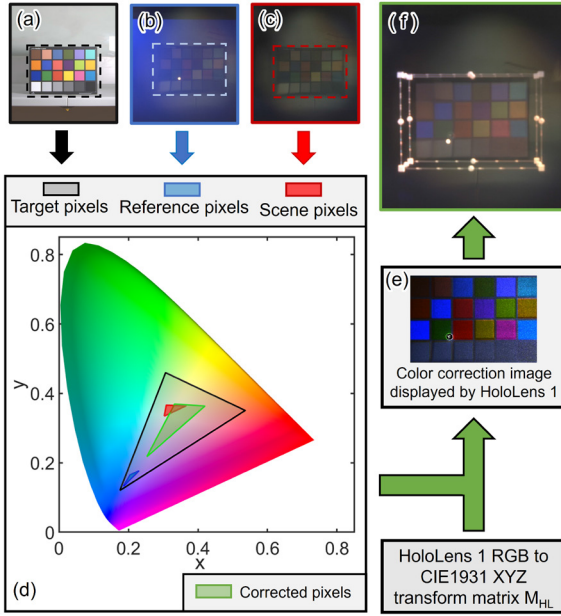


Fig. 6: Color correction of the proposed add-on occlusion system. The correction is realized by overlapping a computed virtual image on the real scene with HoloLens 1. An X-rite color checker is used as the test target in the experiment. (a) The target image taken without the add-on occlusion device and (c) the scene image taken through the device. (b) HoloLens 1 adds a background color on the scene image as a reference to compute the tristimulus of each scene pixel. The color correction is implemented by mapping pixel coordinates of the scene pixels to the target pixels in the CIE1931 xyY color space (d). A computed color compensation image is shown in (e), and more vivid colors are perceived by overlapping it on the real color checker as shown in (f), leading to the enlarged color gamut of the see-through view from the red triangle to the green triangle in (d).

of the scene pixels. Hence, a reference image is taken by the same camera, as shown in Fig. 6 (b). HoloLens 1 is used to display a background image with a triplet of  $[0 \ 0 \ 0.5]^T$ , the mixed color and scene pixels and the displayed pixels are captured. We also measured the transform matrix  $M_{HL}$  of HoloLens 1 by locating a spectrometer at the center of the eyebox, recording tristimulus when color triplets of  $[1 \ 0 \ 0]^T$ ,  $[0 \ 1 \ 0]^T$ ,  $[0 \ 0 \ 1]^T$ , and  $[1 \ 1 \ 1]^T$  are input. Thus, absolute tristimulus  $[X_{HL} \ Y_{HL} \ Z_{HL}]^T$  of pixels displayed by HoloLens 1 can be calculated by substituting  $M_{HL}$  into Eq. 6, and the absolute tristimulus of an arbitrary scene pixel  $[X_s \ Y_s \ Z_s]^T$  is obtained via an inverse mapping from its xy coordinate  $(x_{ref}, y_{ref})$  in the blue triangle of the reference image to  $(x_s, y_s)$  in the red triangle in the chromaticity diagram, which is expressed by:

$$\begin{aligned} X_s &= \frac{X_{HL} - (X_{HL} + Y_{HL} + Z_{HL}) \cdot x_{ref}}{\frac{x_{ref}}{x_s} - 1} \\ Y_s &= X_s \cdot \frac{y_s}{x_s} \\ Z_s &= X_s \cdot \frac{1 - x_s - y_s}{x_s} \end{aligned} \quad (7)$$

Since the color correction is done by adding the computed compensation pixel onto the scene pixel with HoloLens 1. The xy coordinate  $(x_c, y_c)$  of the pixel displayed by HoloLens 1 is solved by the mapping from the coordinate  $(x_s, y_s)$  to the corresponding position  $(x_t, y_t)$  given

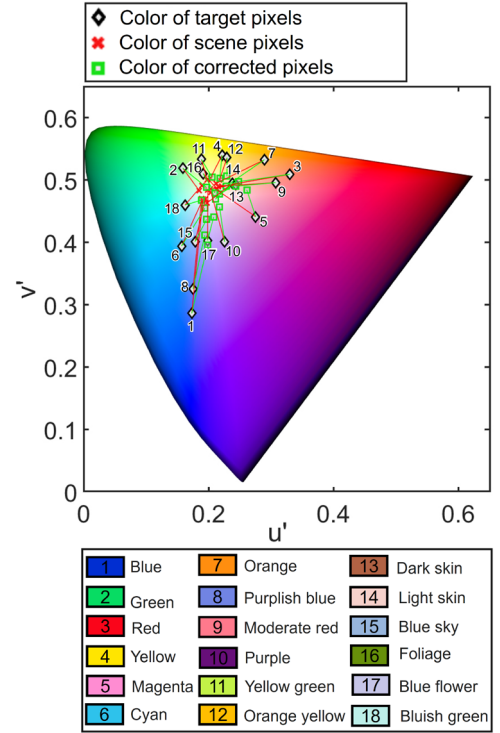


Fig. 7: Coordinates in the CIE1976  $u'v'$  chromaticity diagram of pixels picked from each color patch of the X-rite color checker from the target image (black diamonds), original scene image (red crosses), and the corrected image (green rectangles). The picked colors are indicated by numbers beside the target pixels. Symbols that refer to the same color in different images are connected by solid lines. An average color difference mitigation from  $\delta E = 56.3$  to  $\delta E = 43$  is achieved by implementing the color correction method to the scene image.

by the target image:

$$\begin{aligned} y_c &= \frac{k \cdot y_s y_t}{(k-1)y_s - y_t} \\ x_c &= \frac{y_c(x_t - x_s)}{k \cdot y_s} + x_t \end{aligned} \quad (8)$$

where  $k = Y_s/Y_c$  is the ratio of the luminance of the scene pixel to the compensation pixel. Given the preset luminance of HoloLens 1, the tristimulus  $[X_c \ Y_c \ Z_c]^T$  of the compensation pixel can be easily derived. Then, the input triplet  $[R_c \ G_c \ B_c]^T$  of HoloLens 1 for implementing color correction is computed by substituting  $[X_c \ Y_c \ Z_c]^T$  into Eq. 6. Fig. 6 (e) shows the computed image for correcting the color aberration of the color checker in Fig. 6 (b). The factor  $k$  is set as 1 in the computation. The experiment result is shown in Fig. 6 (f). We implement Fig. 6 (e) as the 2D texture of a virtual quad object and overlap it on the real color checker by using the Object Manipulator system provided by MRTK. The color checker with color correction looks more vivid and proves the bigger color gamut shown as the green triangle in Fig. 6 (d).

Fig. 7 shows the coordinates of pixels picked from each color patch of the X-rite color checker of different images in the CIE1976  $u'v'$  chromaticity diagram. The picked color patches 1 – 18 are indicated below and marked as black diamonds in the diagram by calculating from pixels in the target image. The same color patches are picked in the original scene image and the corrected image, color coordinates are labeled as red crosses and green rectangles, respectively. We further calculated color differences of the scene pixels and the corrected pixels to the target pixels in the  $La^*b^*$  color space for each color set, showing a considerable decrease from  $\delta E = 56.3$  to  $\delta E = 43$  averagely for all the picked color patches. In addition, the colors close to the blue

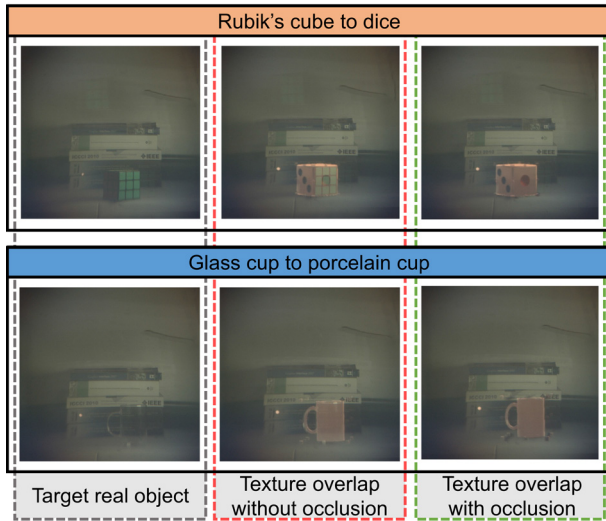


Fig. 8: Texture replacement of real objects by using the add-on occlusion system. **Left:** A Rubik's cube (upper row) and a glass cup (lower row) are placed in the real scene as target objects. **Middle:** Textures displayed by HoloLens are overlapped onto the target objects. However, the absence of occlusion on the target objects causes the mixture between the original textures and the overlapped textures. **Right:** Occlusion is implemented to set the Alpha channel of corresponding areas in the real scene to be zero. Thus, the original textures of target objects are erased. Hence, a dice (upper row) and a porcelain cup (lower row) are displayed realistically in the test scene.

spectrum benefit from the color correction mostly, the average color difference of color patches No.1, No.6, No.8, No.10, No.15, and No.17 is reduced from  $\delta E = 53.6$  to  $\delta E = 34.7$ .

In summary, the color aberration of the proposed add-on occlusion device is mitigated by using the color correction method. Although only the implementation with a static scene is demonstrated, a dynamic application is expected to develop based on the proposed method since it requires minimum computation for computing color correction images. Additionally, few instruments, which are only a camera to capture reference images from the eyebox, are required to implement the color correction in real-time, making the color correction method cost-efficient.

## 6 APPLICATIONS BASED ON ADD-ON OCCLUSION DEVICE

### 6.1 Texture replacement of real objects

The per-pixel occlusion brought by the proposed add-on occlusion device benefits OSTHMDs from not only the more decent virtual display with mutual occlusion but also a series of applications that were not feasible with the OSTHMD platform before. Considering the real scene as an RGBA image, OSTHMDs with the add-on occlusion device achieve the capability to fully modulate a user's vision, where the Red, Green, and Blue channels are manipulated by adding virtual pixels from OSTHMDs and the Alpha channel is manipulated by setting the transparency of the occlusion pattern. Thus, the AR experience provided by current OSTHMDs is pushed to the typical virtual reality (VR) and diminished reality (DR) scope one step further while keeping the unique advantage of a latency-free real-world vision [23].

Figure 8 shows the demonstration of texture replacement by using the prototype add-on occlusion system with HoloLens 1. A Rubik's cube (upper row) and a glass cup (lower row) are used as the target objects, respectively, as shown in the left row. Virtual objects with the same mesh as the target objects are generated, then mapped with the surface textures of dice and a porcelain cup. Simple overlaps of the virtual objects on the target objects are shown in the middle column. The texture projected onto the Rubik's cube suffers from the color blending with the original surface texture, and the texture projected

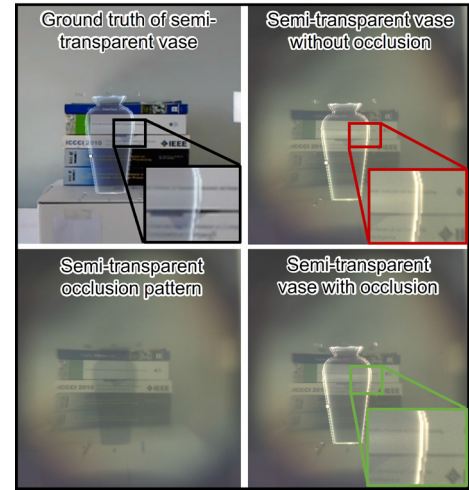


Fig. 9: A semi-transparent vase ( $\text{Alpha} = 0.13$ ) is displayed by using a semi-transparent occlusion pattern. **Top-Left:** The ground-truth image recorded by the Mixed Reality Capture (MRC) system of HoloLens 1, the zoom-in image shows a slightly darker background behind the vase due to its non-zero Alpha channel. **Top-right:** The vase is displayed by the HoloLens 1 without occlusion, and the zoom-in image shows the same brightness between the background covered and uncovered by the vase. **Bottom-left:** Occlusion pattern with an  $\text{Alpha} = 0.2$  is used to darken the background. **Bottom-right:** The vase is shown with the semi-transparent occlusion pattern. The zoom-in image shows the background covered by the vase similarly becomes darker as in the ground truth.

onto the glass cup is inversely overlapped by the specular lighting of the original surface texture while the background keeps omitting. As a result, the texture replacement without occlusion on the real scene provides a less realistic user experience. In comparison, the right column shows the texture replacement is implemented with occlusion. Surface textures of the target objects are erased by the add-on occlusion device. As a consequence, displayed textures by HoloLens 1 fully cover the target objects. The original Rubik's cube and glass cup in the real scene are replaced by the dice and the porcelain in the captured vision.

### 6.2 More realistic display for semi-transparent virtual objects

A semi-transparent object in the real world partially transmits light according to the transparency, generating a dimmed see-through background image. A semi-transparent virtual object expects to show a similar light-dimming effect for a more realistic display. However, OSTHMDs are difficult to display a semi-transparent object with the designed Alpha channel due to the lack of light-cutting capability. The occlusion function, hereby, benefits OSTHMDs with more realistic semi-transparent virtual objects.

Fig. 9 shows a semi-transparent vase with a uniform Alpha channel of 0.13 displayed by the proposed system. A ground truth image is captured by using the Mixed Reality Capture (MRC) system of HoloLens 1. A zoom-in image shows the background behind the vase appears darker than the surrounding uncovered area. In comparison, the vase displayed by the HoloLens 1 without occlusion minor impacts the background brightness as shown in the top-right image. In order to display the vase more realistically, a semi-transparent occlusion pattern with a uniform Alpha channel of 0.2 is added to the area covered by the vase. The result is shown in the bottom-right image. The displayed semi-transparent vase darkens the backward area similar to the ground truth image.



## 7 DISCUSSION AND FUTURE WORKS

### 7.1 FOV limitation

The add-on occlusion device requires different eye reliefs for compatibility with various OSTHMDs, which is done by having different  $d$  in Eq. 2. However, the variation of the distance  $d$  changes the device FOV at the same time. For example, the prototype with HoloLens 1 is chosen to have an eye relief of 17mm, decreasing the FOV from the maximum to  $H19.8^\circ \times V11.1^\circ$ . To the best of our knowledge, the largest FOV of an OC-OSTHMD built with traditional lenses is  $H34^\circ \times V25.5^\circ$  [34], which is still smaller than some OSTHMDs (e.g.,  $H43^\circ \times V29^\circ$  of HoloLens 2). Further expanding the FOV requires conical mirrors with extraordinary dimension [35–37]. Making entire FOVs of various OSTHMDs occlusion-capable with the add-on occlusion device is still challenging.

### 7.2 Device transparency

According to the extinction ratio  $> 1415:1$  of the  $PBS_1$  and the LP used in the prototype, the add-on occlusion has theoretic transparency of less than 0.1%, expecting to display occlusion patterns with a high contrast ratio. However, the measured transparency of the prototype is around 7%, which is similarly caused by the partially polarized nature after the visible light passes the QWPs with the single designed wavelength as the color aberration in the prototype. Therefore, optimizing the proposed add-on occlusion device with all achromatic optical elements benefits the system's performance in many aspects, though the cost will increase considerably. The upgraded prototype with the achromatic design will be built in the future.

### 7.3 Binocular occlusion

Currently, the prototype adds monocular occlusion capability to HoloLens 1. The vision brought by the eyes without occlusion appears brighter since it does not suffer from the optical power loss by the propagation through the add-on occlusion device. In order to keep a consistent binocular view, the add-on occlusion device needs to be mirrored to the left eye. The parallax of the binocular view should be also computed when renders occlusion patterns. For the next step, we will build a binocular system with the mirrored add-on occlusion device on HoloLens 1, and explore how the AR display benefits from the binocular occlusion.

### 7.4 System latency

The prototype generates the occlusion pattern by processing the streamed virtual image from HoloLens 1. The post-processing method makes a large number of SDKs for HoloLens 1 be directly used in the add-on occlusion system, yet faces the challenge of high latency. An occlusion pattern is generated around 0.4s after the virtual image is displayed by HoloLens 1 in our experiment. A decent AR experience with occlusion requires minimum latency. Hence, the rendering pipeline of HoloLens 1 devices may need to be modified to render the virtual image and the corresponding occlusion pattern in the same frame. Meanwhile, keeping compatible with common SDKs is also crucial for practical real-time occlusion.

### 7.5 Integration with computer vision algorithms

Sec. 6.1 proves the feasibility of the proposed system to implement texture replacement on real objects. However, the experiment is completed with the static target. Any movement of the target object causes the mismatch between the displayed texture and the target surface immediately, a dynamic application is currently infeasible with the proposed system. Computer vision algorithms may implement into the application to dynamically bind the displayed texture with the target, or even automatically replace designated real objects in the user's vision.

### 7.6 Built-in solution

The compact design of the proposed system also allows a built-in solution for OSTHMDs. In this case, the decreased environment illumination and resolution of the see-through view may negatively impact the user experience in some application scenarios (e.g., outdoor night).

Fortunately, the occlusion function of the proposed device can be easily turned off by rotating the LP by  $90^\circ$  to have the same polarization direction as  $PBS_1$ . Thus, the light from the front passes through part II directly without suffering the limited transmittance and optical aberration brought by part II. Furthermore, an algorithm may be introduced into the add-on occlusion system that lets the system switch between occlusion on/off status smartly, making a better user experience.

## 8 CONCLUSION

The occlusion function has been proven to benefit OSTHMDs in many aspects. However, a universal solution for introducing the feature into the OSTHMD platform still needs to develop. Starting from the well-demonstrated approach of realizing occlusion with the special type of OSTHMDs, OC-OSTHMD, we propose a novel method by using an add-on occlusion device to implement the occlusion function on an off-the-shelf OSTHMD device. A prototype is built, then assembled with HoloLens 1 device. A color correction method is utilized to reduce the color aberration caused by the massive usage of polarizing optical elements. The real-time virtual display with occlusion is demonstrated by the integrated system. Occlusion patterns are rendered via a post-processing approach. Thus, all SDKs designed for HoloLens 1 are compatible with the proposed device, showing the potential of universal applications. Furthermore, HoloLens 1 with the add-on occlusion device is demonstrated to achieve novel features in static experiments. Based on the easy-accessed solution of realizing occlusion proposed, we expect the technique of occlusion for OSTHMD devices attracts more attention from the community and is more widely used to benefit AR applications.

## ACKNOWLEDGMENTS

This work was partially supported by the National Key Research and Development Program of China (2018YFB1004902) and JSPS KAKENHI Grant (JP22H00539).

## REFERENCES

- [1] K. Akşit, W. Lopes, J. Kim, P. Shirley, and D. Luebke. Near-eye varifocal augmented reality display using see-through screens. *ACM Transactions on Graphics (TOG)*, 36(6):1–13, 2017. 2
- [2] O. Bimber and B. Fröhlich. Occlusion shadows: Using projected light to generate realistic occlusion effects for view-dependent optical see-through displays. In *Proceedings. International Symposium on Mixed and Augmented Reality*, pp. 186–319. IEEE, 2002. 3
- [3] O. Cakmakci, Y. Ha, and J. P. Rolland. A compact optical see-through head-worn display with occlusion support. In *Third IEEE and ACM International Symposium on Mixed and Augmented Reality*, pp. 16–25. IEEE, 2004. 3
- [4] W. Cui, C. Chang, and L. Gao. Development of an ultra-compact optical combiner for augmented reality using geometric phase lenses. *Optics Letters*, 45(10):2808–2811, 2020. 2
- [5] C. Diaz, M. Walker, D. A. Szafir, and D. Szafir. Designing for depth perceptions in augmented reality. In *2017 IEEE international symposium on mixed and augmented reality (ISMAR)*, pp. 111–122. IEEE, 2017. 2
- [6] 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*, pp. 1–8, 2020. 1, 2
- [7] J. L. Gabbard, M. Smith, C. Merenda, G. Burnett, and D. R. Large. A perceptual color-matching method for examining color blending in augmented reality head-up display graphics. *IEEE Transactions on Visualization and Computer Graphics*, 2020. 1
- [8] J. L. Gabbard, J. E. Swan, and A. Zarger. Color blending in outdoor optical see-through ar: The effect of real-world backgrounds on user interface color. In *2013 IEEE Virtual Reality (VR)*, pp. 157–158. IEEE, 2013. 1
- [9] T. Hamasaki and Y. Itoh. Varifocal occlusion for optical see-through head-mounted displays using a slide occlusion mask. *IEEE transactions on visualization and computer graphics*, 25(5):1961–1969, 2019. 2, 3
- [10] Y. Hiroi, Y. Itoh, T. Hamasaki, and M. Sugimoto. Adaptivisor: Assisting eye adaptation via occlusive optical see-through head-mounted displays. In *Proceedings of the 8th Augmented Human International Conference, AH '17*. Association for Computing Machinery, New York, NY, USA, 2017. doi: 10.1145/3041164.3041178. 2, 6

- [11] M. Inami, N. Kawakami, D. Sekiguchi, Y. Yanagida, T. Maeda, and S. Tachi. Visuo-haptic display using head-mounted projector. In *Proceedings IEEE Virtual Reality 2000 (Cat. No. 00CB37048)*, pp. 233–240. IEEE, 2000. 4
- [12] Y. Itoh, T. Hamasaki, and M. Sugimoto. Occlusion leak compensation for optical see-through displays using a single-layer transmissive spatial light modulator. *IEEE transactions on visualization and computer graphics*, 23(11):2463–2473, 2017. 3
- [13] Y. Itoh, T. Langlotz, D. Iwai, K. Kiyokawa, and T. Amano. Light attenuation display: Subtractive see-through near-eye display via spatial color filtering. *IEEE transactions on visualization and computer graphics*, 25(5):1951–1960, 2019. 1
- [14] 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. *Optics letters*, 45(13):3361–3364, 2020. 2, 3
- [15] 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.*, pp. 133–141. IEEE, 2003. 2, 3
- [16] K. Kiyokawa, Y. Kurata, and H. Ohno. An optical see-through display for mutual occlusion of real and virtual environments. In *Proceedings IEEE and ACM International Symposium on Augmented Reality (ISAR 2000)*, pp. 60–67. IEEE, 2000. 2
- [17] K. Kiyokawa, Y. Kurata, and H. Ohno. An optical see-through display for mutual occlusion with a real-time stereovision system. *Computers & Graphics*, 25(5):765–779, 2001. 2
- [18] B. Krajancich, N. Padmanaban, and G. Wetzstein. Factored occlusion: Single spatial light modulator occlusion-capable optical see-through augmented reality display. *IEEE transactions on visualization and computer graphics*, 26(5):1871–1879, 2020. 2, 3
- [19] T. Langlotz, M. Cook, and H. Regenbrecht. Real-time radiometric compensation for optical see-through head-mounted displays. *IEEE Transactions on Visualization and Computer Graphics*, 22(11):2385–2394, 2016. doi: 10.1109/TVCG.2016.2593781 2, 6
- [20] A. Maimone and H. Fuchs. Computational augmented reality eyeglasses. In *2013 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, pp. 29–38. IEEE, 2013. 3
- [21] A. Maimone, X. Yang, N. Dierk, A. State, M. Dou, and H. Fuchs. General-purpose telepresence with head-worn optical see-through displays and projector-based lighting. In *2013 IEEE Virtual Reality (VR)*, pp. 23–26. IEEE, 2013. 4
- [22] S. Mori, S. Ikeda, A. Plopski, and C. Sandor. Brightview: Increasing perceived brightness of optical see-through head-mounted displays through unnoticeable incident light reduction. In *2018 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, pp. 251–258. IEEE, 2018. 2
- [23] S. Mori, S. Ikeda, and H. Saito. A survey of diminished reality: Techniques for visually concealing, eliminating, and seeing through real objects. *IPSI Transactions on Computer Vision and Applications*, 9(1):1–14, 2017. 8
- [24] S. Odinokov, M. Shishova, V. Markin, D. Lushnikov, A. Zherdev, A. Solomashenko, D. Kuzmin, N. Nikonov, and S. Ivanov. Augmented reality display based on photo-thermo-refractive glass planar waveguide. *Optics Express*, 28(12):17581–17594, 2020. 2
- [25] C. Pan, Z. Liu, Y. Pang, X. Zheng, H. Cai, Y. Zhang, and Z. Huang. Design of a high-performance in-coupling grating using differential evolution algorithm for waveguide display. *Optics Express*, 26(20):26646–26662, 2018. 2
- [26] D. Pascale. Rgb coordinates of the macbeth colorchecker. *The BabelColor Company*, 6, 2006. 6
- [27] T. C. Peck, J. J. Good, A. Erickson, I. Bynum, and G. Bruder. Effects of transparency on perceived humanness: Implications for rendering skin tones using optical see-through displays. *IEEE Transactions on Visualization and Computer Graphics*, 28(5):2179–2189, 2022. 1
- [28] J. Polvi, T. Taketomi, A. Moteki, T. Yoshitake, T. Fukuoka, G. Yamamoto, C. Sandor, and H. Kato. Handheld guides in inspection tasks: Augmented reality versus picture. *IEEE transactions on visualization and computer graphics*, 24(7):2118–2128, 2017. 1
- [29] K. Rathinavel, G. Wetzstein, and H. Fuchs. Varifocal occlusion-capable optical see-through augmented reality display based on focus-tunable optics. *IEEE transactions on visualization and computer graphics*, 25(11):3125–3134, 2019. 2, 3
- [30] G. Singh, S. R. Ellis, and J. E. Swan. The effect of focal distance, age, and brightness on near-field augmented reality depth matching. *IEEE transactions on visualization and computer graphics*, 26(2):1385–1398, 2018. 1
- [31] J. E. Swan, A. Jones, E. Kolstad, M. A. Livingston, and H. S. Smallman. Egocentric depth judgments in optical, see-through augmented reality. *IEEE transactions on visualization and computer graphics*, 13(3):429–442, 2007. 1
- [32] J. E. Swan, G. Singh, and S. R. Ellis. Matching and reaching depth judgments with real and augmented reality targets. *IEEE transactions on visualization and computer graphics*, 21(11):1289–1298, 2015. 1
- [33] A. Wilson and H. Hua. Design and prototype of an augmented reality display with per-pixel mutual occlusion capability. *Optics express*, 25(24):30539–30549, 2017. 2, 3
- [34] A. Wilson and H. Hua. Design of a pupil-matched occlusion-capable optical see-through wearable display. *IEEE Transactions on Visualization and Computer Graphics*, 2021. 2, 3, 9
- [35] 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. *Optics letters*, 46(17):4208–4211, 2021. 2, 3, 9
- [36] Y. Zhang, X. Hu, K. Kiyokawa, N. Isoyama, H. Uchiyama, and H. Hua. Realizing mutual occlusion in a wide field-of-view for optical see-through augmented reality displays based on a paired-ellipsoidal-mirror structure. *Optics Express*, 29(26):42751–42761, 2021. 2, 3, 9
- [37] 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)*, pp. 301–311. IEEE, 2020. 2, 3, 9
- [38] Y. Zhang, R. Wang, E. Y. Peng, W. Hua, and H. Bao. Color contrast enhanced rendering for optical see-through head-mounted displays. *IEEE Transactions on Visualization and Computer Graphics*, 2021. 2
- [39] P. Zhou, Y. Li, S. Liu, and Y. Su. Compact design for optical-see-through holographic displays employing holographic optical elements. *Optics Express*, 26(18):22866–22876, 2018. 2