Doctoral Dissertation

Design and Prototyping of Wide Field of View Occlusion-capable Optical See-through Augmented Reality Displays by Using Paired Conical Reflectors

Zhang Yan

December 11, 2021

Graduate School of Science and Technology
Nara Institute of Science and Technology

A Doctoral Dissertation
submitted to Graduate School of Science and Technology,
Nara Institute of Science and Technology
in partial fulfillment of the requirements for the degree of
Doctor of ENGINEERING

Zhang Yan

Thesis Committee:

Kiyoshi Kiyokawa

(Professor, Division of Information Science)

Yasuhiro Mukaigawa

(Professor, Division of Information Science)

Hideaki Uchiyama

(Associate Professor, Division of Information Science)

Naoya Isoyama

(Assistant Professor, Division of Information Science)

Design and Prototyping of Wide Field of View Occlusion-capable Optical See-through Augmented Reality Displays by Using Paired Conical Reflectors*

Zhang Yan

Abstract

Mutual occlusion is an essential technology for augmented reality (AR) displays. Virtual objects displayed with mutual occlusion can provide correct perceptual cues and keep adequate visibility under an excessively illuminated environment. Although a few works have been developed, realizing mutual occlusion in a wide field of view (FOV) is still challenging to state-of-the-art occlusion-capable optical see-through augmented reality (OC-OST-AR) displays. In this dissertation, we addressed this issue by building OC-OST-AR displays based on a paired-conical-reflector structure. With different types of spatial light modulators (SLMs) being utilized, two sub-structures, including transmissive SLM-based paired-ellipsoidal-mirror (T-PEM) structure and reflective SLM-based paired-paraboloidal-mirror (R-PPM) structure, are proposed. The paired conical reflectors work with a tiny aperture stop to increase the numerical aperture (NA) of AR imaging systems significantly. Then, typical hard-edge or enhanced soft-edge occlusion can be conducted in a wide FOV by fixing an SLM before the entrance pupil or at an inner focal plane. Proof-of-concept prototypes are built for both T-PEM and R-PPM structures. Enhanced soft-edge occlusion is demonstrated in a monocular FOV of H122°×V74° with the T-PEM system, and hard-edge occlusion is demonstrated in a monocular FOV of H83.5°×V53.1° with the R-PPM system.

*Doctoral Dissertation, Graduate School of Science and Technology,
Nara Institute of Science and Technology, December 11, 2021.

Keywords:

Augmented reality, Virtual reality, Optical see-through head-mounted display, Near-eye display

Contents

List of Figures

1 Introduction

Augmented reality (AR) is the technology that merges digital content and the physical world. With the development in past decades, AR technologies involved with a variety of application scenarios, including navigation [59], entertainment [26], education [6], inspection task [89, 23], et al., have been proposed based on hand-held displays. And the widespread use of smartphones and tablets in recent years facilitates the boost of the civil AR market significantly. Meanwhile, optical see-through augmented reality (OST-AR) displays, which enable the AR content and the physical world to be perceived encumbrance-free at the same time, attract attention from a large group of researchers. User studies have shown the unique advantages of OST-AR displays and reported a general preference on OST-AR displays to hand-held AR displays in some application scenarios [63]. OST-AR displays are thereby considered as a promising platform for AR applications and are actively developed nowadays.

OST-AR displays are firstly utilized for military purposes. The early research of OST-AR displays can be back far as the 1950s, which is the assistant system for jet pilots [14]. Furthermore, the commercial products of OST-AR displays emerged since the 1990s, researchers and engineers began to explore solutions of OST-AR displays suitable for daily usage [29]. To date, although there is no way to achieve appealing performance in all aspects, various technologies were proposed to solve specific issues in OST-AR displays. Waveguide combiners are widely used in existing commercial OST-AR displays to downsize the system volume [61]; holographic display [90, 52], light field display [53, 24, 25, 13], retinal projection display [11, 85, 87], vari-focal display [81, 8, 1], and tomographic display [33, 66] are proposed to solve the well-known vergence-accommodation conflict (VAC); free-form optical components[79] and pinlight display [54] show the potential to build OST-AR displays with a wide field of view (FOV). Works that integrate

several technologies show a more dominant performance in some aspects [49, 45, 28].

However, the development of ideal OST-AR displays is hindered by two crucial issues. The insufficient image brightness and poor perceptual cues of the projected AR content deteriorate the user experience dramatically, and OST-AR displays targeting to address the two issues are still less developed. The former issue comes from the limited illuminance of display elements (liquid crystal display (LCD), organic light-emitting diode (OLED), or laser diode)), the poor transmittance of the imaging system, and the excessively ambient lighting. Although commercial OST-AR displays generally attached a tinted visor before the optical combiner to the attenuate background light, the projected AR content is still too dark to be perceived clearly. It is reported that the latest OST-AR display, HoloLens 2, only provides the maximum illuminance of 203lux for the AR content, while the background lighting propagates through the front tinted visor could range from <1lux in a dark room to >3000lux in the outdoor environment. In consideration of the visual contrast sensitivity of human beings, HoloLens 2 is merely to render the AR content with sufficient visibility in common indoor scenarios where the maximum illuminance of the transmitted environment light through the tinted visor is around 310lux [19, 32, 84].

The latter issue comes from the lack of occlusion between virtual objects and real objects. In the physical world, orders of different objects in depth can be recognized by observing mutual overlaps among them, which provides strong perceptual cues to observers. Regarding the AR scene, the occlusion of a virtual object by a real object is easy to be conducted as long as depth information of the physical world is included in the same rendering pipeline of the virtual scene. Those depth data can be acquired from a static scene or by real-time range sensors such as light detection and ranging (LiDAR) scanners. Then, the occluded virtual object can be rendered correctly by automatically dismissing the parts behind real objects. In contrast, the overlapping of a virtual object on a real object is much more complex in OST-AR displays. Directly projecting the foreground virtual object on the background real object causes a semi-transparent AR content, and the projected virtual is mostly like to be invisible due to the noticeable illuminance gap to natural-lighted real objects. Psychophysics research has shown that users

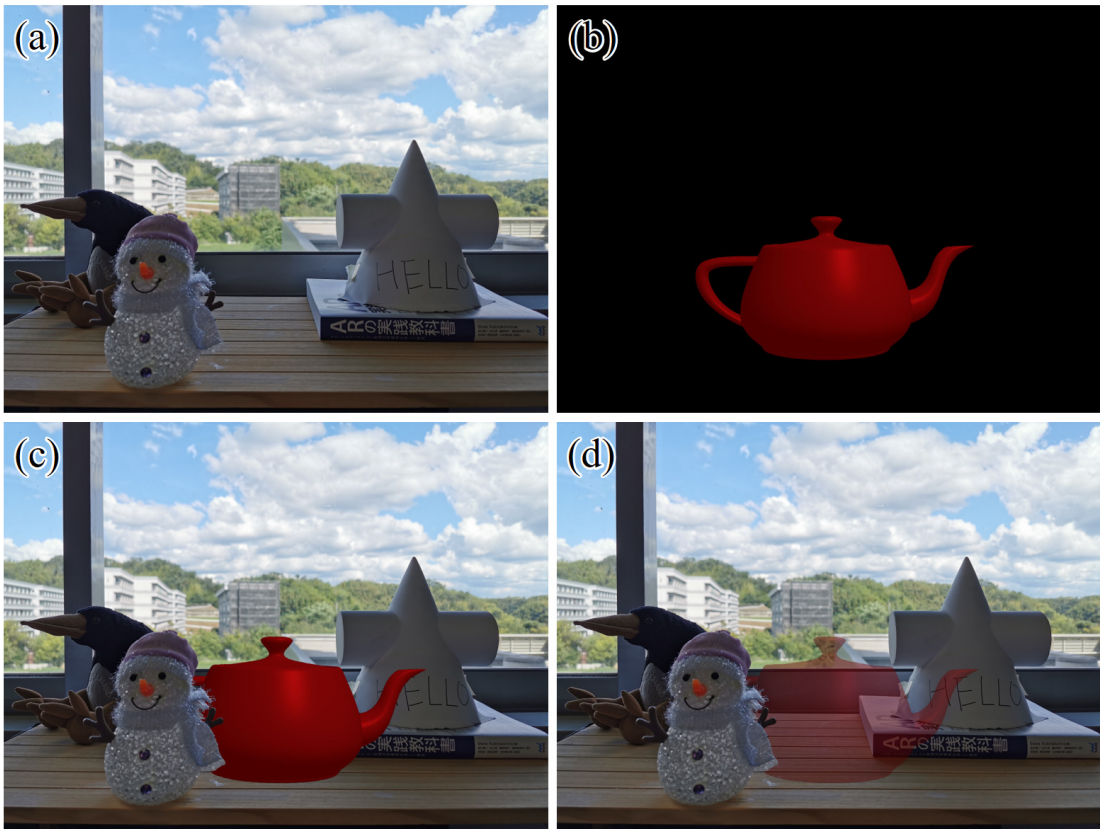

Figure 1.1: A virtual image **(b)** is displayed with a real scene **(a)** by OST-AR displays. The target case **(c)** shows that the virtual image is displayed with correct occlusion, and the real case **(d)** shows the virtual image is displayed without mutual occlusion.

are confused by the incorrect occlusion in AR applications [70, 62, 15].

These issues can be easily addressed in a video see-through augmented reality (VST-AR) display, which electronically combines AR content with a real image captured by a video camera mounted on the user's head [31]. Since the user's natural vision is entirely blocked while the combined image is projected into the pupil, VST-AR displays can modulate the real scene on a pixel-by-pixel basis. Namely, occlusion between the virtual object and the real object can be implemented without difficulty. Besides, the common image processing (e.g., intensity and tint corrections, and controlling the blending ratio) can also be conducted since the "real scene" provided by a VST-AR display is virtually a se-

ries of recorded video frames, which may provide a better AR experience in some application scenarios. Nevertheless, a VST-AR display reduces the rich information content of the physical world because the display provides low spatial and temporal resolution, limited depth-of-field, fixed-focus, and so forth. Moreover, a user may temporarily lose his or her vision under a system failure of VST-AR displays. In comparison, OST-AR displays allow the user to observe the physical world in an optically direct way. Hence, the visual information from the physical world is maximum kept, and a see-through view is likely to remain even when an OST-AR display is powered off.

Fig. 1.1 depicts how the lack of occlusion negatively impacts the visualization of AR scenes. With the real environment shown in Fig. 1.1 (a), a red teapot in Fig. 1.1 (b) is implemented as the AR content. Assume that the depth map locates the teapot between the snowman and the sculpture, the target visual effect perceived by users is shown as Fig. 1.1 (c). The teapot is partially occluded by the snowman while covering the back sculpture, which provides a correct perceptual cue to users. Meanwhile, both specular and diffuse lighting effects on the teapot are supposed to be observed clearly, which enhances the realism of the AR scene. However, existing commercial OST-AR displays only support a one-way occlusion from a real object to a virtual object by dismissing corresponding parts of the virtual object. The light emitted from the occluded parts of the real object reaches users' eyes then generates a mixed image with the projected virtual image. As a result, incorrect perceptual cues are provided to users by those OST-AR displays. What is more, the bright background reduces the contrast ratio of the displayed AR content. Thus, some visual information, such as the lighting effects, is most likely to be absent, resulting in the deterioration of user experience.

Therefore, mutual occlusion is proposed to address the two issues in OST-AR displays. The challenge to provide a complete mutual occlusion effect is to support decent occlusion from virtual objects to the real scene. Hence, this dissertation omits the realization of the occlusion from real objects to virtual objects since it has been well addressed in current OST-AR displays. Compared to the occlusion on virtual objects, implementing occlusion on the real scene becomes much more complex for OST-AR imaging systems. Unlike VST-AR displays, the real image in OST-AR displays directly comes from the physical

world. In this case, rendering mutual occlusion for the real image is equivalent to modulating the light emitted from the physical world. Occlusion-capable optical see-through augmented reality (OC-OST-AR) displays are, therefore, built with ingenious optical architectures to realize the modulation in real-time. In addition to projecting the AR content into users' eyes, the incoming beams from the real scene are also captured, then modulated by spatial light modulators (SLMs) in the imaging systems. Occlusion between the displayed AR content and the physical world can be rendered as long as correct depth maps are implemented. Similarly, an arbitrary area of the background is also allowed to be darkened by setting the alpha channel for the incoming real image, which guarantees the high visibility of AR content regardless of various environment illuminance. In most cases, the processed real scene and the rendered virtual image are emerged by an optical combiner positioned before users' eyes. Then, a mixed image enters the pupil to generate the desired AR scene. An early system for conducting mutual occlusion in OST-AR displays was developed by Tatham [73] where a transparent LCD panel is used as the active mask. Opaque virtual objects are superimposed over the real scene, and virtual cast shadows are generated on a real object by using pixels on the active mask with intermediate transparency. A similar idea has been employed in a stationary AR display [58]. To date, a number of OC-OST-AR displays have been studied for tackling the problem of mutual occlusion. Approaches are categorized as hard-edge occlusion and soft-edge occlusion based on different optical architectures. In terms of different mechanisms used for conducting occlusion, the occlusion pattern is rendered on a pixel-by-pixel basis in the hard-edge occlusion systems so that edges look "hard," namely, clear and sharp. While the occlusion pattern in the soft-edge occlusion systems has a low resolution, thereby shows "soft," in other words, blurred edges.

Nevertheless, existing OC-OST-AR displays are hard to support a wide FOV because of their special optical architectures for modulating the real scene [88, 27]. Whereas a wide FOV has been proven as a crucial feature for AR displays [41]. An ideal OC-OST-AR display is expected to support high-quality mutual occlusion in a wide FOV, which is a very challenging task based on nowadays research. In this dissertation, we address this challenge by building OC-OST-AR displays in a paired-conical-reflector (PCR) architecture. Contributions are highlighted in

follows:

1. A wide FOV is firstly achieved in OC-OST-AR displays by introducing paired conical reflectors into the imaging systems. The FOV restriction by the limited numerical aperture (NA) in typical hard-edge occlusion systems is overcome, and the occlusion precision and FOVs in typical soft-edge occlusion systems are increased simultaneously.

2. A transmissive SLM-based paired-ellipsoidal-mirror (T-PEM) sub-structure is proposed. The design is optionally to operate in the hard-edge occlusion way or the enhanced soft-edge occlusion way according to different positions of the SLM. Bench-top prototypes following the enhanced soft-edge occlusion way is built, a wide occlusion-capable FOV of H122°×V74° is demonstrated.

3. A reflective SLM-based paired-paraboloidal-mirror (R-PPM) sub-structure is proposed. The system works in the hard-edge occlusion way. The optimized image quality, see-through efficiency, and occlusion performance are achieved by using an LCoS as the SLM. A bench-top prototype is built and a wide occlusion-capable FOV of H83.5°×V53.1° is demonstrated.

4. A rendering framework for operating both the T-PEM and R-PPM system is proposed. The projection from an arbitrary target pixel to the pixels on the display panel and SLMs are driven, then implemented as a texture in the fragment shader stage. The optical aberration for virtual display and occlusion rendering are compensated by using two-step mapping in the fragment shader.

2 Related works

2.1 Hard-edge occlusion

In order to modulate beams from the real scene, SLMs are essential for building OC-OST-AR displays. In terms of the location of SLMs in the imaging system, an OC-OST-AR display can be categorized as hard-edge occlusion systems or soft-edge occlusion systems. Either of the two approaches has unique features.

In the case that an SLM is positioned at an inner focal plane of the imaging system, which is illustrated in Fig. 2.1 (a), an OC-OST-AR display works in the hard-edge occlusion way. A general optical layout for hard-edge occlusion systems composes an objective lens, a relay lens group, an eyepiece, and an optical combiner. The objective lens converts incident beams into image points at the focal plane, and an SLM mounted here can block arbitrary areas in the real scene. The relay lens group works for erecting the inverted real image by the objective lens, and the optical combiner before the eyeball is used to emerge the transmitted real scene and the projected virtual image. Because each digital pixel of the occlusion pattern provided by the SLM corresponds to individual object points, pixel-wise occlusion can be conducted by opening and shutting the pixels on the SLM panel at positions where real objects should appear and disappear, respectively. Therefore, a sharp occlusion pattern can be perceived by the eyes, as illustrated in Fig. 2.1 (b). Furthermore, by superimposing the virtual image emitted from the projector onto the occlusion pattern, the AR content is naturally displayed, as illustrated in Fig. 2.1 (c).

The concept of hard-edge occlusion was first proposed by Kiyokawa et al. in 2000 [37, 38]. The optical structure is composed of a pair of convex lenses, an erecting prism, and a transmissive LCD positioned at the back focal plane of the objective lens. The whole structure is designed with a unit magnification, so that

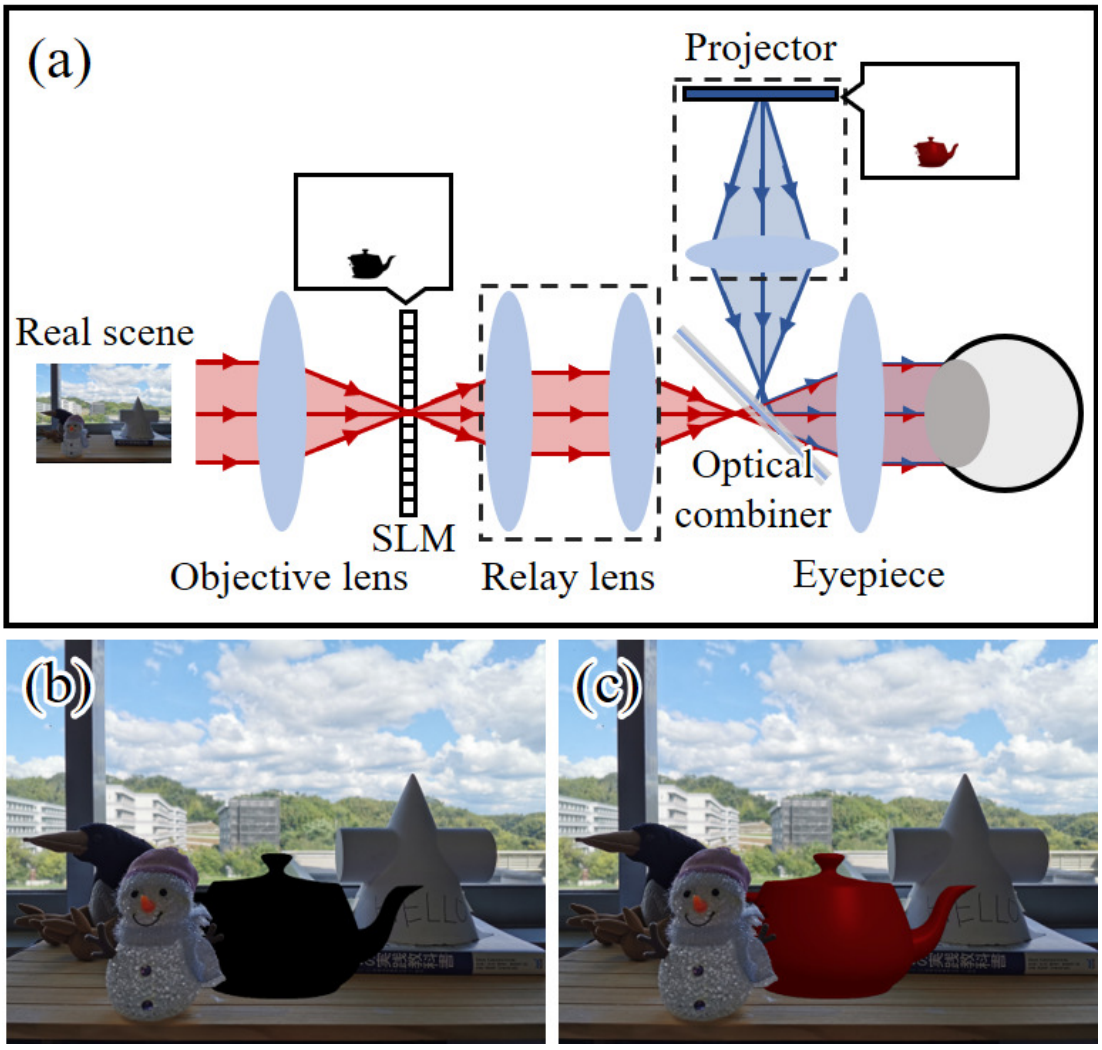

Figure 2.1: **(a)** A general architecture of hard-edge occlusion OC-OST-AR displays, **(b)** a sharp occlusion pattern perceived by eyes, and **(c)** a virtual image is superimposed on the pattern.

allows the system to render a sharp occlusion pattern at a designated depth while keeping the original see-through scene. A more compact design was proposed by Cakmakci et al. later [7]. By using a reflective SLM, a polarizer, and an X-prism, the system volume is efficiently reduced. The polarizer sets incident light as s-polarization, which allows the first polarized coating plane inside the X-prism reflects the light towards the reflective SLM. The SLM locates at the back focal

plane of the objective lens to conduct hard-edge occlusion. Reflected light by the SLM is converted to be p-polarization, then the propagation direction can be transformed into horizontal by the second polarized coating plane. Compared with the initial design by Kiyokawa et al., the usage of X-prism folds the optical path, so that downsizes the imaging system. Nevertheless, the design fails to erect the perceived image due to the absence of a relay lens group.

Restricted by the performance of the early SLM panels, few designs of hard-edge occlusion were proposed in the next decade. OC-OST-AR displays with hard-edge occlusion attract researchers' sights again after reflective SLMs with more powerful performance are developed. Gao et al. suggested utilizing a liquid crystal on silicon (LCoS) device rather than transmissive SLM as the occlusion panel in 2012 [20]. An obvious advantage of using LCoS for modulating the real scene is the high reflectance, which can be higher than 70%, while the value is usually less than 10% in transmissive SLMs. In addition, artifacts inside a transmissive SLM cause severe diffraction that leads to a deteriorated see-through vision, while LCoS-based occlusion suffers less from the diffraction, thereby, provides better image quality. However, due to expensive free-from optical elements used in Gao's design, a prototype is not built. The following work is done by Wilson and Hua [82]. By optimizing the structure with commercial lenses and prisms, a compact prototype is built. A decent image quality (1.24 arcmin) for AR display is demonstrated within a FOV of diagonal 30°. An alternative device to support high-quality see-through view for hard-edge occlusion systems is the digital micro-mirror device (DMD). Different from LCoS-based structures, the ultra-high refresh rate of DMDs allows the panel to be not only used for rendering occlusion but also projecting the virtual image via time-multiplexing control. Thus, DMD-based OC-OST-ARs can be built with fewer elements by throwing additional display panels away. Krajancich et al. exhibit a prototype that firstly demonstrates the feasibility of using DMD as both the occlusion panel and the display panel [39]. Very soon, an optimized structure is proposed by Ju et al. [30]. The polarization-based double-pass structure makes the system achieve a smaller form-factor and be free of vertical parallax. In addition, the time-multiplexing control of the DMD is further developed that prove the capability to display full-color videos.

Apart from the progress of hard-edge occlusion OC-OST-AR displays by utilizing emerging hardware, some new features are also introduced with the latest works. One of them is the vari-focal hard-edge occlusion capability that can be implemented with common vari-focal AR displays to further enhance the depth perception of users. The first system is proposed by Hamasaki and Itoh in 2019 [22]. Instead of attaching the SLM at the fixed focal plane of the objective lens, they mount a transmissive LCD on an electric slider that allows the occlusion pattern to be projected at various depths by mechanically moving the panel along the optical axis. Their proof-of-concept prototype has demonstrated the capability of rendering mutual occlusion within a depth range from 25cm to infinity. Soon after this, Rathinavel et al. also realized vari-focal occlusion by implementing tunable-focus optics [65]. Different from Hamasaki's work, an LCoS in this design is fixed at a physically static position, while the optical depth of the occlusion pattern is controlled by modulating a focus-tunable lens integrated into the eyepiece. In order to keep a unit magnification of the real scene, a secondary focus-tunable lens is similarly assembled into the objective lens and correspondingly driven in real-time. Both of the two designs support vari-focal AR displays via either computational [22] or optical [65] approaches.

The attempt to keep a wide see-through view was done by Wilson and Hua in 2021 [83]. In their design, an s-polarizer is fixed in front of a polarizing beam splitter. The following optical structure is vertically arranged so that a polarized see-through light is reflected toward the objective lens. In addition to essential components for realizing hard-edge occlusion, a quarter-wave plate is integrated into the system, which allows non-occluded see-through light modulated by the LCoS to pass then be reflected by the polarizing beam splitter. Meanwhile, the illuminance of the peripheral scene is controlled by a crossed polarizer, which makes it seamlessly merge with the modulated central view. Though mutual occlusion is still restricted in a narrow central FOV, this design shows a promising way to develop practical occlusion-capable OST-AR displays, since the further expansion of the occlusion-capable FOV may not only be challenging but also costly.

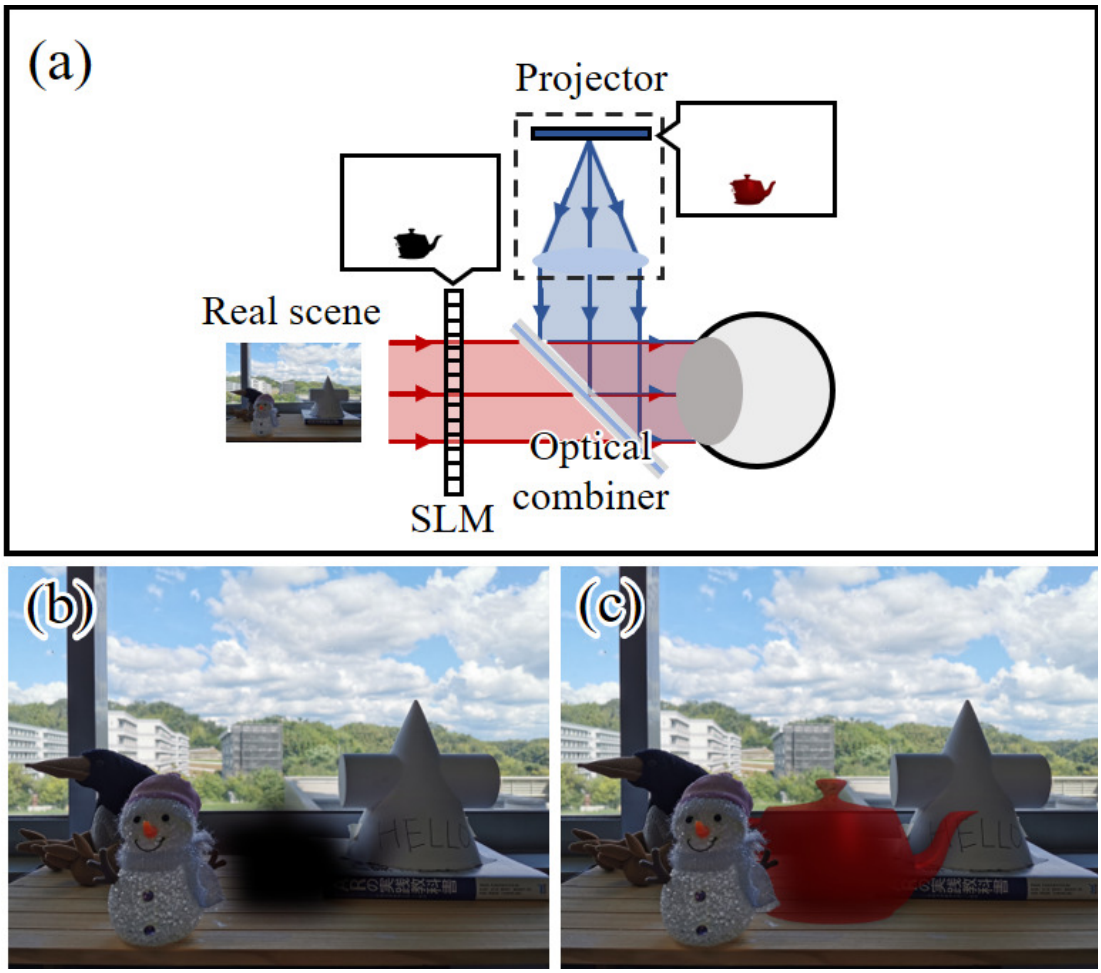

Figure 2.2: **(a)** A general architecture of soft-edge occlusion OC-OST-AR displays, **(b)** a blurred occlusion pattern perceived by eyes, and **(c)** a virtual image is superimposed on the pattern.

2.2 Soft-edge occlusion

Although the hard-edge occlusion approach makes OST-AR displays precisely render the occlusion pattern on a pixel-by-pixel basis, the massive application of optical elements leads to a relatively bulky device. In order to provide mutual occlusion capability with a smaller form-factor, soft-edge occlusion is proposed. Divergent from hard-edge occlusion systems that locate the SLM at an inner focal plane of the imaging system, SLMs in soft-edge occlusion systems are directly

fixed before the users' eyes. Fig. 2.2 (a) depicts a general structure of OC-OST-AR displays built in the soft-edge occlusion way. The raw beam from the real scene propagates through the front SLM, and an optical combiner is similarly used to combine the processed real scene and the projected virtual image. While soft-edge occlusion systems have a smaller form-factor, out-of-focus of the SLM plane leads to a blurred occlusion pattern be perceived, as illustrated in Fig. 2.2 (b). A virtual image displayed on the blurred occlusion pattern suffers from insufficient occlusion and unnatural mutual occlusion to real objects, which is illustrated as Fig. 2.2 (c).

An attempt to optimize the occlusion performance in soft-edge occlusion OC-OST-AR displays was proposed by Maimone et al. in 2013 [51]. Their system is composed of a shutter layer, a backlight, and two or more thin transmissive SLM layers packaged in an eye-glasses-like volume. An occlusion pixel in the display is not directly given by a physical pixel on either SLM but integrated by a series of pixels of the SLM stack via a light-field manner with time-multiplication, thus it supports a sharper occlusion pattern than normal soft-edge occlusion displays. Although the image quality of this display is significantly deteriorated by the massive implementation of SLMs, the unique light-field mechanism allows the occlusion pattern to be rendered with multiple focal depths simultaneously. Itoh et al. presented their video-overlap method in 2017 [27]. Rather than optimizing occlusion via more ingenious usage of SLMs, their system enlarges the occlusion pattern for guaranteeing sufficient occlusion area under the out-of-focus state. Although occlusion leak is introduced by the oversized pattern at the same time, it is addressed by overlapping a synthesized image with both the virtual content and corresponding video see-through scene. This method achieves pixel-wise occlusion, while the latency is close to typical VST-AR displays due to the image processing.

2.3 Mutual occlusion in a wide FOV

Both hard-edge occlusion systems and soft-edge occlusion systems have demonstrated the feasibility to support mutual occlusion in AR. With efforts in past decades, OC-OST-AR displays with better occlusion performance and smaller

form-factor have been developed. However, a crucial issue, which is the narrow FOV, is less explored in previous researches. Table. 2.1 shows FOVs of existing designs for OC-OST-AR displays introduced in the two sections above and provides a comparison to some wide-view OST-AR displays. In a nutshell, Those OC-OST-AR displays built with the hard-edge occlusion approach provide FOVs less than 40°, and soft-edge occlusion designs show wider FOVs close to the range of general wide-view OST-AR displays.

It is well-known that the monocular FOV of human beings spans in an extremely large range, which is more than 210° horizontally and 120° vertically [5]. Although the binocular FOV is reduced to be 120° horizontally [40], it is still much wider than FOVs achieved by existing OC-OST-AR displays. The limited FOV results in the difficultly of displaying large-scale AR contents, which may require massive head movement that causes users fatigue. In addition, the influence of FOV in AR search task has been evaluated, shows a wide FOV promotes users' performance significantly [34, 41, 67, 74]. Therefore, it is vital to consider the realization of mutual occlusion in a wide FOV for keeping the visibility of AR contents in the periphery human vision.

However, both the hard-edge occlusion approach and soft-edge occlusion approach face challenges when researchers try to expand the FOV. Regarding hard-edge occlusion systems, optical elements, such as lenses [22] and prisms [20], are used as objective lenses and eyepiece lenses. Beams from the real scene are focused by the objective lens, thus available for SLMs to render pixelated occlusion. However, lenses and prisms have limited NA, incident light beyond this range suffers from severe optical aberration [55]. It is easy to address the aberration in virtual reality (VR) displays by coding the projected image with compensation algorithms, while the real scene of OC-OST-AR displays is projected by the physical world so that it can not be pre-processed. Therefore, the scale of the real scene that can be well-imaged by optical elements is restricted, which causes the narrow FOV in existing hard-edge occlusion OC-OST-AR displays. Even though emerging technologies (e.g., metalen) have been demonstrated to expand the FOV of imaging systems further, it is still far away from building a wide-view hard-edge occlusion OC-OST-AR displays [43, 42].

In regard to soft-edge occlusion systems, the form-factor is generally smaller

OST-AR	Occlusion mechanism	FOV (diagonal)
Kiyokawa et al. [38]	hard-edge occlusion	30°
Cakmakci et al. [7]	hard-edge occlusion	40°
Cao et al. [20]	hard-edge occlusion	40°
Wilson and Hua. [82]	hard-edge occlusion	30°
Krajancich et al. [39]	hard-edge occlusion	8.7°
Ju et al. [30]	hard-edge occlusion	2.7°
Rathinavel et al. [65]	hard-edge occlusion	15°
Hamashaki and Itoh. [22]	hard-edge occlusion	<40°
Wilson and Hua [83]	hard-edge occlusion	40°
Maimone et al. [51]	soft-edge occlusion	65°
Itoh et al. [27]	soft-edge occlusion	70°-80°
Kiyokawa [36]	none	146°
Maimone et al. [54]	none	110°
Dunn et al. [17]	none	100°
Aksit et al. [1]	none	>60°

Table 2.1: FOVs and occlusion mechanisms of existing OC-OST-AR displays and a comparison to some wide-view OST-AR displays.

than OC-OST-AR displays built in the hard-edge occlusion way since no lenses or prisms are used (except the projector part). However, beams emitted from different pixels of the real scene are not focused. Hence, severe overlapping occurs among different imaging beams. Namely, precise occlusion to individual real scene pixels is theoretically impossible since light for imaging nearby pixels

would be partially blocked simultaneously. Existing OC-OST-AR displays with soft-edge occlusion attenuate the problem by introducing more complex occlusion processing algorithms, while it increases the computation cost that causes additional latency [27, 51]. What is more, FOVs of soft-edge occlusion systems are restricted by the size of SLMs. Further expanding the FOV requires either a large dimension or an eye-contact layout of the SLM. The former solution is inefficient because the FOV in soft-edge occlusion systems is the tangent function of the SLM size, and the latter aggravates the blur of the perceived occlusion pattern due to more severe defocus and imaging beam overlap.

Wide-view OST-AR displays have been actively developed in recent years, and AR is integrated into various application scenarios. Rendering AR content with sufficient visibility and correct perceptual cues in a wide FOV is vital for practical applications of AR. Mutual occlusion technology is the key to solve these issues but realizing high-quality occlusion in a wide FOV for OC-OST-AR displays is still less explored. Restricted by mechanisms of existing approaches, no solution has been reported so far. In this dissertation, we build the AR imaging system based on the PCR structure that significantly increases the NA of displays. By using the paired conical reflectors as the objective lens and the eyepiece, two optical architectures, which are the T-PEM structure and the R-PPM structure, are proposed. Bench-top prototypes are built, the capability of mutual occlusion in a super wide FOV is proved with both sub-structures. Specifically, the prototype built on the T-PEM structure works in the enhanced soft-edge occlusion way, a FOV of H160°×V74° for the virtual display and a FOV of H122°×V74° for mutual occlusion are demonstrated. The prototype built on the T-PEM structure works in the hard-edge occlusion way, a FOV of 83.5°×V53.1° for both mutual occlusion and the virtual display is demonstrated. The proposed work is expected to provide novel thinking for the community on realizing mutual occlusion in a wide FOV.

3 Transmissive SLM-based paired-ellipsoidal-mirror (T-PEM) system

3.1 Optical architecture

As discussed in Sec. 2.3, there are two approaches to conduct mutual occlusion for OC-OST-AR displays. In order to achieve a wide FOV, the hard-edge occlusion approach is firstly selected with the reasons following:

1. Although soft-edge occlusion systems support a wider FOV than hard-edge occlusion systems, existing researches seem to approach a limit, which is still smaller than the natural view of humans, as shown in Table. 2.1.

2. The occlusion performance demonstrated with hard-edge occlusion systems shows dominant advantages, which is difficult for soft-edge occlusion systems to catch up by merely optimizing the occlusion rendering algorithm.

Consequently, an optical structure is still essential for realizing high-quality mutual occlusion in a wide FOV. In order to modulate a sizeable real scene, a target imaging system is expected to meet the following requirements:

1. The use of lenses as either the objective lens and the eyepiece should be avoided due to the limited NA.

2. Beams emitted from a vast range should be focused with limited aberration regardless of various incident angles. Thus an SLM can be fixed to render pixelated occlusion.

3. The real scene is projected with a unit magnification and erected correspondingly by the optical system.

With this consideration, we firstly reviewed optical architectures of near-eye displays with wide FOVs. Apart from those works with chromatic laser sources, which are not suitable for the real scene formed by achromatic light, imaging systems built on catadioptric optics in VR displays seem to meet the above requirements partially.

A catadioptric optical system is a type of optical structure where lenses and curved mirrors are combined. Utilization of the non-spherical curved mirrors (e.g., ellipsoidal mirrors [80], paraboloidal mirrors [68, 44], and hyperboloidal mirrors [91, 12]) help imaging systems get rid of the limitation of lens performance, thus making it possible to expand the FOVs further. To date, a large number of catadioptric displays for VR implementation have been developed. Nagahara et al. proposed a wide FOV closed-type HMD with a hyperboloidal mirror and an ellipsoidal mirror yielding a FOV of H180°×V60° [60]. Kiyokawa proposed a wide FOV OST-AR display with a pair of hyperboloidal half-silvered mirrors that achieves a maximum horizontal FOV of 146.2° [35]. The use of non-spherical mirrors in these catadioptric VR displays increases the NA of imaging systems efficiently, thus allows a super-wide FOV to be achieved. Even though the projected virtual image needs to be pre-processed to compensate the optical aberration of the imaging systems, it shows the potential to build OC-OST-AR displays with similar wide FOV since the first requirement of a high NA is satisfied.

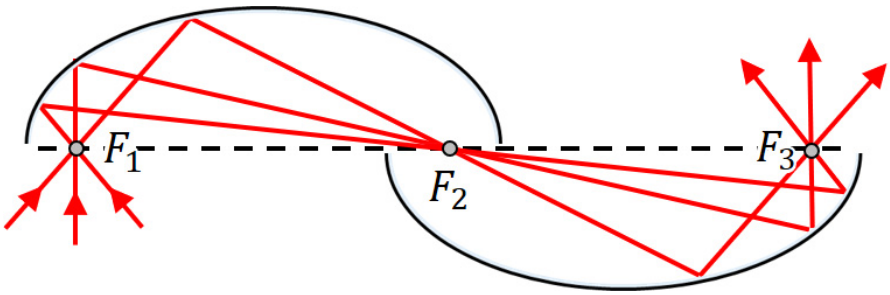

Figure 3.1: Rays propagate through each focus of two joint ellipsoidal mirrors.

Then, we focused on overcoming the optical aberration in common catadioptric

displays. A VR display with wide FOV and small distortion by Yang et al. attracted our attention [86]. Two ellipsoidal mirrors with the same conic constant are used in the display, a FOV of H160°×V120° with a binocular overlap of 80° is demonstrated by the simulation. Although stacked lenses are used in the design for attenuating extra optical aberration, the structure with two joint ellipsoidal mirrors shows the potential to achieve an ultra-high NA and project an unit magnified scene distortion-free, which meets two of the requirements above for building OC-OST-AR displays with wide FOV. The ray propagation through the basic structure is shown in Fig. 3.1. Two ellipsoidal mirrors are joint together with the same focus F_2, rays entering through an individual focus F_1 of an ellipsoidal mirror are transmitted to an individual focus F_3 of another ellipsoidal mirror without optical aberration.

Preliminary experiments are conducted with the same layout Fig. 3.1, experiment results are shown in Fig. 3.2. The two ellipsoidal mirrors are halves from a whole one, and a pinhole mask is positioned at the focus of the upper ellipsoidal mirror to block stray light. A test target displayed by a laptop is located before the system, the recorded image by fixing a camera at the focus of the lower ellipsoidal mirror is shown as Fig. 3.2 (b). The perceived image is nearly distortion-free yet with extremely low brightness, which is not sufficient to provide a visible image in practice.

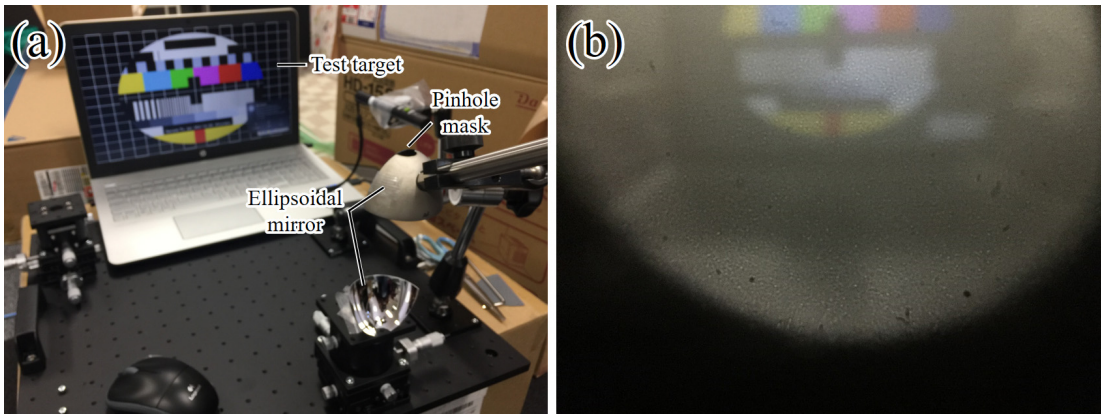

Figure 3.2: Preliminary experiment with two joint ellipsoidal mirrors.

Consequently, an optimized structure is essential for projecting a real scene

with better image quality. Then, the image property of an ellipsoidal mirror is analyzed, which provides insight for both the optimization design and occlusion rendering.

An ellipsoidal mirror is able to image a wide-view scene, but only the input image given as a point at either focus can be displayed distortion-free at the other focus, otherwise optical aberration will impact on the displayed images. The imaging property of ellipsoidal mirrors has been analyzed by some previous works. Chang et al. proposed that the dominant aberration of a single ellipsoidal mirror is primary astigmatism, and it can be eliminated by using a double ellipsoidal mirror structure [9]. Liu et al. discussed the aberration of an ellipsoidal mirror in the high-aperture case, then suggested that the sensitivity to the aberration of an ellipsoidal mirror is similar to a single lens [47, 48]. Massey et al. derived the ray transfer matrix of an ellipsoidal mirror, which is similar to an ideal curved mirror [56].

The previous works above indicate that the imaging property of the ellipsoidal mirror can be considered as a single lens or a curved mirror despite the aberration, and dominant aberration can be eliminated by jointing two ellipsoidal mirrors together. In the following, the analysis and optimization of the basic two-ellipsoidal-mirror structure in Fig. 3.1 are done based on the assumption that any tiny parts of an ellipsoidal mirror can be dealt as a reflective mirror.

The optimized optical architecture is shown in Fig. 3.3. Rays from an object point o and an LCD at P_d are drawn as red and blue lens, respectively. Two paired ellipsoidal mirrors are used as the objective lens and the eyepiece, which increases the NA of the imaging system significantly. Different from the simple structure that only consists of the two ellipsoidal mirrors, an extra lens pair composed by L_v and L_x is added. In terms of the imaging property shown in Fig. 3.1, chief rays pass through F_1 with large incident angles are shrunk into a limited area after reflected by an ellipsoidal mirror. The converged angle of the reflected rays cone depends on the parameters of the ellipsoidal mirror and is much smaller than original incident angles at F_1. Hence, the addition of the lens pair after the ellipsoidal mirror introduces minimal distortion. For example, considering an ellipsoidal mirror with a major axis of 61.9mm and a minor axis of 35.4mm, which provides a FOV of H180° × V145° theoretically, the maximum incident angle for

rays reaching the lens pair is only 35°. The foci of two ellipsoidal mirrors are labeled as F_1, F_2, F_3, and F_4, and a pinhole mask P_m is positioned at the F_3 to function as an aperture stop that cuts stay light off.

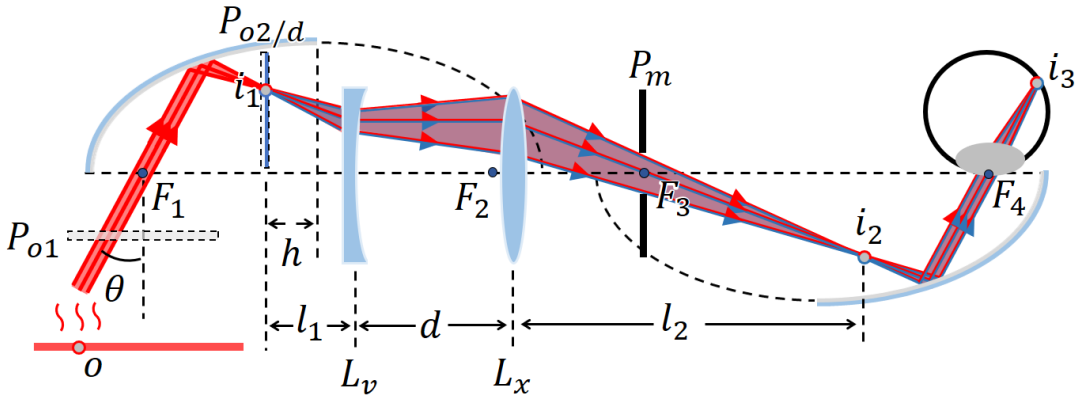

Figure 3.3: The basic optical architecture for the T-PEM system.

In the case that the beam from object point o is reflected by arbitrary part on the ellipsoidal mirror with a major axis length of a and a minor axis length of b, focal length f_e of the corresponding curved mirror is given as:

$$f_e = \frac{R}{2} = \frac{(2ar - r^2)^{\frac{3}{2}}}{2ab} \tag{3.1}$$

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 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:

$$r = \frac{a(1 - e^2)}{1 - e\sin\theta} \tag{3.2}$$

Where $e = c/a = \sqrt{a^2 - b^2}/a$ is the eccentricity of the ellipse. And the position of the image point can be calculated with the Gaussian lens equation:

$$q = \frac{p.f}{p - f} \tag{3.3}$$

where q and p are the image distance and the object distance, respectively. The projection between a real scene pixel and the focused image pointi_1 is generated with Eq.3.1, Eq.3.2, and Eq.3.3.

Then, parameters of the lens pair are determined based on the process of a ray being transformed by the lens pair component. We implement the ray transfer matrix to derive the propagation process of an incident ray. In this method, a lens $\mathbf{M}$ with a focus length f, a free space $\mathbf{S}$ and a distance l are given as:

$$\mathbf{M} = \begin{bmatrix} 1 & 0 \\ -\frac{1}{f} & 1 \end{bmatrix}, \mathbf{S} = \begin{bmatrix} 1 & l \\ 0 & 1 \end{bmatrix} \tag{3.4}$$

And a ray is expressed by the distance x and a deviation angle θ with regard to the optical axis:

$$\mathbf{I} = \begin{bmatrix} x \\ \theta \end{bmatrix} \tag{3.5}$$

Here, the light propagation process with the lens pair is given as:

$$\mathbf{I_2} = \mathbf{S_{l_2}} \mathbf{M_x} \mathbf{S_d} \mathbf{M_v} \mathbf{S_{l_1}} \mathbf{I_1} \tag{3.6}$$

Where $\mathbf{I_1}$ and $\mathbf{I_2}$ are the matrices of the incident ray and the exit ray, $\mathbf{M_v}$ and $\mathbf{M_x}$ are the matrices of the two lenses, and $\mathbf{S_{I_1}}$, $\mathbf{S_d}$, and $\mathbf{S_{I_2}}$ represent the free spaces during the light propagation path in order.

Then the ray transfer matrix of the whole optical component can be derived as:

$$\begin{bmatrix} x_2 \\ \theta_2 \end{bmatrix} = \begin{bmatrix} A & B \\ C & D \end{bmatrix} \begin{bmatrix} x_1 \\ \theta_1 \end{bmatrix} \tag{3.7}$$

where the elements A, B, C, and D are given by:

$$\begin{aligned} A &= 1 + d\frac{l_2 - f_x}{f_v f_x} - \frac{l_2(f_v + f_x)}{f_v f_x} \\ B &= d + l_1 + l_2 - d\frac{l_2 f_v + l_1 f_x - l_1 l_2}{f_v f_x} - \frac{l_1 l_2(f_v + f_x)}{f_v f_x} \\ C &= \frac{d}{f_v f_x} - \frac{f_v + f_x}{f_v f_x} \\ D &= 1 + d\frac{l_1 - f_v}{f_v f_x} - \frac{l_1(f_v + f_x)}{f_v f_x} \end{aligned} \tag{3.8}$$

Where f_v and f_x are focal lengths of L_v and L_x. Eq. 3.8 can be simplified by setting $f_v = -f_x = -f$, which is given by the condition that the lens pair is constituted by a concave lens and a convex lens with an opposite focal length. In this case, the exit ray of the lens pair is given as follows:

$$\begin{bmatrix} x_2 \\ \theta_2 \end{bmatrix} = \begin{bmatrix} 1 - d\frac{l_2 - f}{f^2} & d + l_1 + l_2 + d\frac{(l_1 - l_2)f - l_1 l_2}{f^2} \\ -\frac{d}{f^2} & 1 - d\frac{l_1 + f}{f^2} \end{bmatrix} \begin{bmatrix} x_1 \\ \theta_1 \end{bmatrix} \tag{3.9}$$

In order to allow an incident ray to transmit the optical component without deviation, a solution that satisfies $x_2 = -x_1$ and $\theta_2 = \theta_1$ needs to be found, which is given as:

$$\begin{aligned} \theta_1 &= \frac{x_1}{l_1 + f} \\ l_2 &= 2f + l_1 \end{aligned} \tag{3.10}$$

Eq. 3.10 indicates that only the ray propagating toward the back focus of the concave lens at q_1 keeps the original propagation direction, and it will reach the conjugate position at the plane behind the convex lens with a distance of $(2f + l_1)$. In this case, the spacing of the two lenses does not impact on the exit ray. Therefore, the layout of the two lenses for ensuring the chief rays propagate without deviation is constrained as overlapping the back focus of the concave lens f_v on the focus of the first ellipsoidal mirror F_2. Furthermore, the position of the second ellipsoidal mirror needs to be placed so that its focus F_3 is positioned at the front focus of the convex lens f_x. Then, the shared focus of the basic structure in Fig. 3.1 is virtually transferred to the point F_3 in the optimized design. The pinhole mask is placed here to balance image brightness and resolution.

Another parameter that needs to be addressed in the optimization structure is the spacing d between the two lenses. To satisfy the requirement 2 of the optimization structure design, exit rays at i_2 and incident rays at i_1 should satisfy $x_2 = -x_1$, while the deviation angle θ_2 is allowed to be slightly different from θ_1. Then the spacing d can be derived with the restriction:

$$d = \frac{2f^2}{l_1 + f} = \frac{2f^2}{h + c} \tag{3.11}$$

Where h is the distance from i_1 to the minor axis of the ellipsoidal mirror. Hence, a parameter l_1 (or h) can be calculated by substituting f, d into Eq. 3.11. Thus, the system is set to focus the real object at an specific FOV and distance according to Eq. 3.3. Then, an SLM can be either put at P_{o1} or P_{o2} to render mutual occlusion.

In the case that an SLM is positioned at P_{o1}, the system mostly works in the soft-edge occlusion, while the pinhole mask P_m working as an aperture stop shrinks propagation beams. The entrance pupil and the exit pupil of the system are virtually the projection of the pinhole by the former optical element group and the latter ellipsoidal mirror. Although some parallax exists, the entire system functions as a virtual pinhole on the pupil that enhances the occlusion precision by compressing each incident beam. Different from the configuration with a physical pinhole mask positioned before the eyes, the FOV perceived through the projected virtual pinhole can be maximum kept since the mask depth and the distance to the pupil can be exactly zero. As a consequence, the system works in the enhanced soft-edge occlusion way by locating an SLM at P_{o1}.

When an SLM is positioned at P_{o2}, the beam from an object point is focused by the ellipsoidal mirror as the image point i_1 at the P_{o2} plane. The SLM here is allowed to block the entire beam by shuttering off the corresponding digital pixel. Namely, the real scene is blocked by the SLM based on a pixel-by-pixel basis. Hence, the system works in the common hard-edge occlusion way. Meanwhile, the paired ellipsoidal mirrors increase NA significantly. Thus the system achieves a wide see-through view while rendering precise occlusion.

In order to display virtual images in the same wide FOV at the same time, an LCD panel is also positioned at P_d for both configurations. Light projected by a

pixel on the LCD is similarly imaged by the following optical elements as the real scene rays, which is shown in Fig. 3.3. In practice, an optical combiner should be additionally used to emerge the transmitted real scene and the projected virtual image at the plane $P_{o2/d}$, and a computational algorithm is also essential for eliminating the optical aberration on virtual display due to the single-ellipsoidal-mirror imaging.

3.2 System analysis and prototyping

3.2.1 Basic layout

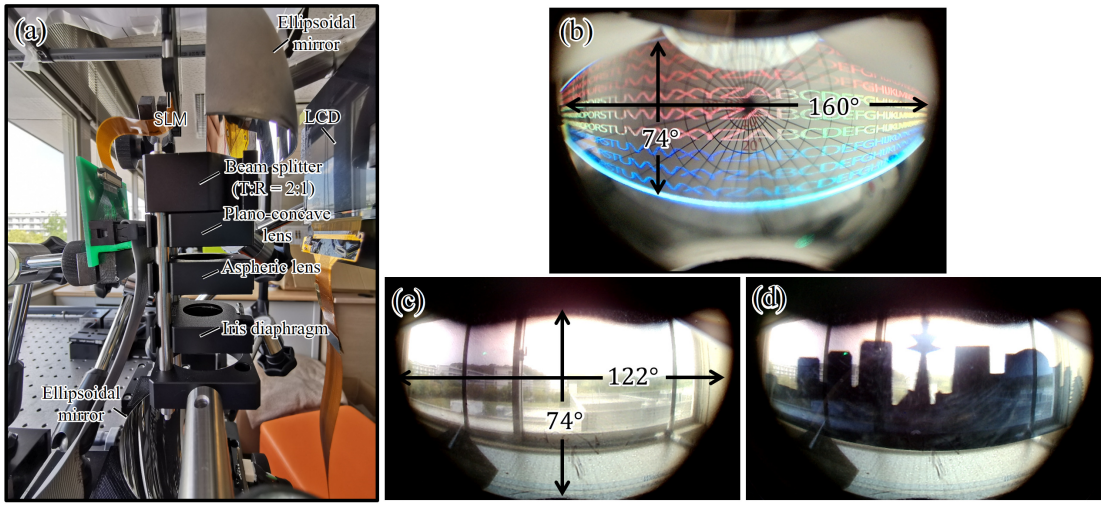

Figure 3.4: A prototype based on the basic layout **(a)**. A virtual image composed by letters **(b)**, a see-through view **(c)**, and an occlusion pattern **(d)** recorded by a fish-eye camera.

The first prototype that is directly built on the basic layout shown in Fig. 3.3. An beam splitter with a ratio 2:1 of transmittance to reflectance is additionally used as the optical combiner at the plane $P_{o2/d}$. As shown in Fig. 3.4 (a), two ellipsoidal mirrors manufactured by Phoenix Electroformed Products, Inc., whose major and minor axes are 61.9mm and 35.4mm, respectively, are assembled vertically. According to the discussion above, the lens pair should be composed of a concave lens and a convex lens, whose focal lengths are opposite to each other.

Here, the prototype utilizes a VIS 0° coated plano-concave lens and a plastic aspheric lens by Edmund Optics, whose focal lengths are −25mm and 25mm, respectively. Unfortunately, apertures of the two lenses we use are not enough to allow all incident rays to pass through so that the FOV of the prototype is smaller than the theoretical maximum. The pinhole is given by an iris diaphragm from SIGMAKOKI, whose aperture is adjustable from 1mm to 15mm. The SLM used in the prototype is 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). An SHARP LS029B3SX02 LCD panel with 705PPI and a dimension of 51.84mm×51.84mm is fixed behind the first ellipsoidal mirror. A virtual image is projected from the LCD then reflected into the system by a 20mm dielectric beam splitter cube manufactured by SIGMAKOKI, which guarantees unpolarized light with a reflectance-to-transmittance ratio of 1 : 2 giving transmittance of around 70% for light wavelength at 550nm.

Output images recorded by a 1080p fish-eye camera with a focal length of 3.7mm, f/2.1, minimum illumination of 0.1lux, and exposure time of 0.03s are shown in Fig. 3.4. Fig. 3.4 (a) shows a virtual image consists of letters is displayed in a FOV of H160°×V74° with compensated distortion. The SLM is temporarily taken off, and a chart with a series of concentric circles is placed 20cm in front of the entrance pupil of the system so that circles appear in the see-through view by an interval of 10°. Notice that apparent large distortion is primarily due to the fish-eye camera, and some above areas fail to be displayed because of the limited size of the beam splitter cube. In addition, the transmittance and reflectance of dielectric beam splitter vary with wavelength, polarization, and incident angles, which results in chromatic aberration that appears on the recorded virtual image. Due to the limited dimension of the SLM, the occlusion-capable FOV is shrunk to a range of H122°×V74°. And recorded images of a real scene without and with an occlusion pattern are shown as Fig. 3.4 (b) and Fig. 3.4 (c), respectively. The real scene is efficiently blocked by the occlusion pattern shown on the SLM.

As mentioned above, the prototype built with a constant lens spacing d is set to focus the reflected image point i_1 at different planes with heights h. An object point focused into deviated planes by the ellipsoidal mirror or a digital pixel projected by an LCD mismatching the designated plane are displayed with worse

image quality, which can be observed in Fig. 3.4 (b). The efficient calibration of the system requires a quantification of the projection from the incident angle θ and to the focusing plane height h with different object distances h being chosen. The calculation is done with incident angles θ from $-90°$ to $55°$ and distances of 0mm, 1mm, 10mm, 100mm, and 1000mm (equivalent to infinity), and parameters of ellipsoidal mirrors used in the prototype is selected. Results are shown in Fig. 3.5, where y coordinates and x coordinates refer to the heights h and incident angles θ, respectively. When the object distance p is reduced to 0mm (means the object locates at F_1), the focusing plane is fixed at F_2, which is the well-known property of an ellipsoidal mirror. With the object distance p decreases from 1000mm to 0mm, curves tend to move down and be more flattened. Specific dynamic ranges of focusing plane shift with object distance change are indicated as colorful bars at incidents angles of $50°$, $30°$, $0°$, $-30°$, and $-60°$. The range shrinks with the object lifting from the bottom of the FOV to the top. A significant rise of the focusing plane is observed within the object distance increasing from 0mm to 100mm, and it becomes roughly stable when the object goes even far.

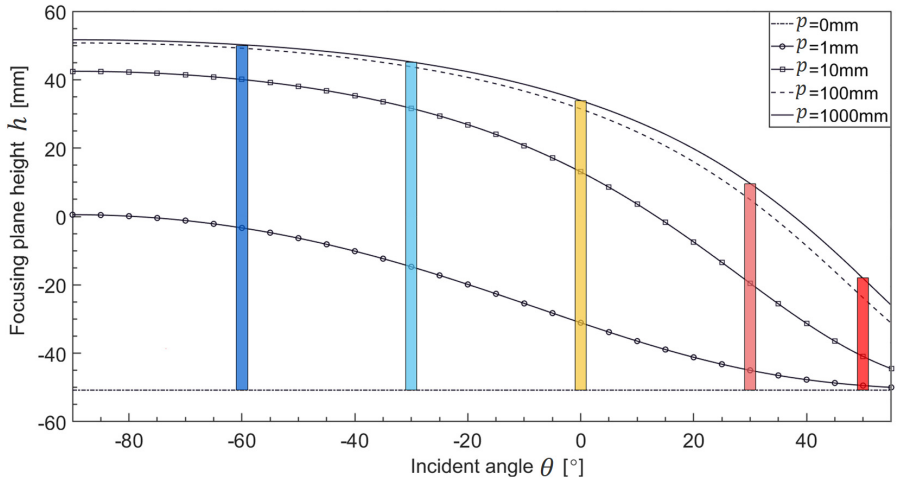

Figure 3.5: Object points o with different distances and incident angles θ are reflected by an ellipsoidal mirror into image points i_1 at heights of h.

In practice, the incident ray emitted from the upper of the FOV is focused on a relatively lower plane by the first ellipsoidal mirror as shown in Fig. 3.5, then

the distance h is going to be shorter. So a higher value of d is needed for giving a sharp vision in the upper FOV according to Eq. 3.11. In contrast, a smaller value of d should be set to observe the lower FOV clearly. Besides, the h also decreases with the object point o moving from far to near as shown in Fig. 3.5, a longer spacing is more suitable for observing a near image. Meanwhile, the plane $P_{o2/d}$ shifts correspondingly with a configuration with different d selected. The spacing of the lens pair needs to be set appropriately for different application scenarios.

Apart from the lens spacing, the image quality is also influenced by the pinhole aperture that works as an aperture stop in the system. Hence, we tested the prototype with different lens spacings and pinhole aperture, see-through views are recorded as shown in Fig. 3.6. The SLM is taken off in the experiment to have a better vision. A camera with an aperture of 2mm and a FOV of H89.8°×V50.6° is used to record the output images, and it is set to focus at distances from very near (65mm) to infinity sequentially. The perceived image through an imaging system is also influenced by the focal distance of the recording camera in general. However, our proposed system works with a flexible eye-box that is able to be shrunk less than 1mm with pinhole aperture decreasing, so the effect with the change in the camera focus in the proposed system may differ from that in normal imaging systems. Therefore, the camera focus is also taken into account in the experiment.

Captured images of the see-through view are shown in Fig. 3.6. Three targets are used to demonstrate the change of the perceived image; a transparent 1951 USAF test chart located at a distance of 65mm, an ISO-12223 test chart located at a distance of 1000mm, and a building right in front of the window at a distance of 100m approximately, which is considered as infinitely away from the system. The pinhole is enlarged from minimum 1mm to maximum 5mm, and the lens spacing is set to 15.5mm, 13.5mm, and 12mm with each aperture value. The camera focus is fixed to 1000mm when the lens spacing is 12mm or 15.5mm, whereas three levels of the camera focus are tested when the lens spacing is 13.5mm. In general, the observed image is more blurred with the pinhole aperture increasing. The highest resolution is obtained with the aperture of 1mm regardless of the lens spacing and the camera focus. In this case, the system is similar to a pinhole imaging system so that the three targets at different distances are observed clearly at the same

Figure 3.6: See-through views tested with various pinhole apertures, lens spacings, and the camera focuses.

time. The sharpest area of the whole FOV is determined by the lens spacing. The lower, middle, and upper parts of FOV are observed with the highest resolution by setting the lens spacing as 12mm, 13.5mm, and 15.5mm, respectively. Moreover, this effect becomes more apparent with the bigger pinhole. The focus of the whole system also varies with the lens spacing. The nearest 1951 USAF chart is sharper with the lens spacing of 15.5mm than the lens spacing of 12mm when the pinhole aperture is bigger than 1mm. What is more, the system shows similar resolution for all three targets at different distances when the lens spacing and the pinhole aperture are determined. Fig. 3.5 also indicates similar optical power for objects with distances from 100mm to 1000mm. Therefore, the performance of the proposed system is majorly FOV-dependent. An optimized structure is introduced in the following by assuming all real scenes locates infinite.

3.2.2 Optimized design and prototype

Although the basic structure in Fig. 3.3 is demonstrated to support a super-wide FOV and the corresponding mutual occlusion capability, a system built with this structure suffers from a minimum vertical parallax of 203.2mm due to the vertical layout of the optical elements as shown in Fig. 3.4 (a), which negatively impacts the AR experience. In addition, the virtual image projected by an LCD at P_d shows similar blurry as the see-through scene because the FOV-dependent image quality of the proposed structure. In this section, an optimized structure is designed for reducing the viewpoint shift and enhancing the virtual display. A bench-top prototype is built and the system performance is analyzed. Finally, mutual occlusion is demonstrated in the T-PEM system.

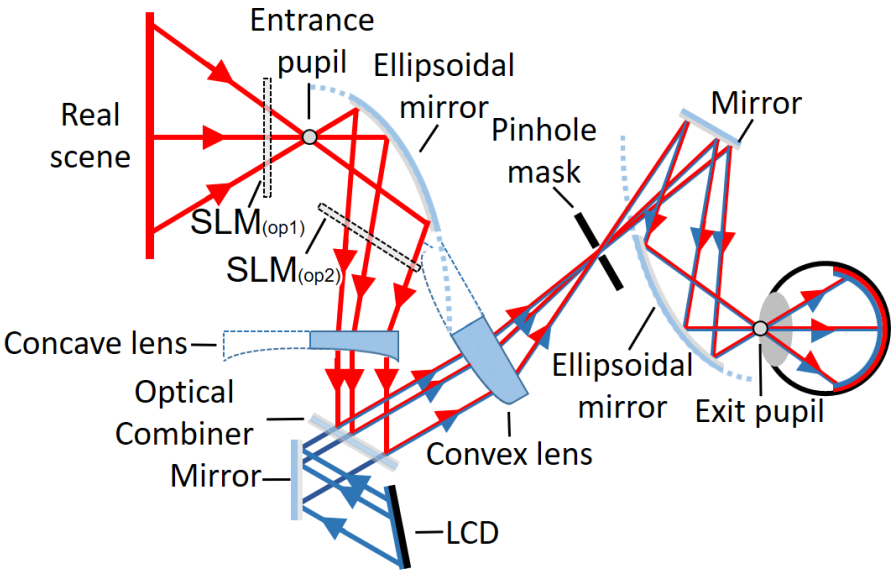

Figure 3.7: The schematic diagram of the T-PEM system.

The schematic diagram of the optimized structure is shown as Fig. 3.7. 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 consist of a concave lens and a convex lens, are drawn with solid lines for the usable parts and 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 system. An optical combiner between the concave lens and the convex lens is used to merge the real scene and the virtual image. The LCD is placed here rather than above the concave lens because of smaller defocus. Rays from the real scene and the LCD firstly are guided to up-right by the optical combiner, then downward by the mirror above the second ellipsoidal mirror. The optical path with the vertical direction is partially transformed into the horizontal direction so that the vertical parallax is reduced. A mirror locating before the LCD is used to further compress the optical path of virtual display. Mutual occlusion is conducted by fixing an SLM either in front of the entrance pupil ($\mathrm{SLM}_{(\mathrm{op1})}$) or above the lens pair ($\mathrm{SLM}_{(\mathrm{op2})}$), which leads the system to work in the enhanced soft-edge occlusion way or the hard-edge occlusion way. Moreover, both the $\mathrm{SLM}_{(\mathrm{op2})}$ plane and the LCD are tilted to partially compensate the defocus from designated inner focal planes. Hence, the hard-edge occlusion performance and the virtual display can be improved simultaneously.

In order to derive the tilt angle of the LCD panel, we calculate the distribution of transmitted image points through the ellipsoidal mirror and the concave lens versus a see-through FOV of H60°×V60°, which is shown in Fig. 3.8. Incident beams are treated as parallel since a minor effect on system performance is brought by the object distance. The z coordinate represents distances from image points to the L_v plane. The upper colorful surface depicts the distribution of reflected image points by the ellipsoidal mirror (l_1 in Fig. 3.3). Along with the FOV varies from H0°×V30° to H0°×V-30°, the distance increases from 9.2mm to 44.5mm, contours are drawn below for better understanding. Therefore, the occlusion panel fixed at an inclined $\mathrm{SLM}_{(\mathrm{op2})}$ plane, which is shown as Fig. 3.7, reduces the defocus of each occlusion pixel so that it expands the hard-edge occlusion-capable FOV.

After being imaged by the concave lens, real scene pixels spread within a narrower range shown as the lower colorful surface in Fig. 3.8. Hence, the virtual image by projected the LCD at the tilted plane P_d suffers less defocus than by the basic layout shown in Fig. 3.3. Moreover, transmitted image points from the pixel i_3 at the retina back through the second ellipsoidal mirror and the convex lens is traced with the same FOV. The distribution is shown as the meshed sur-

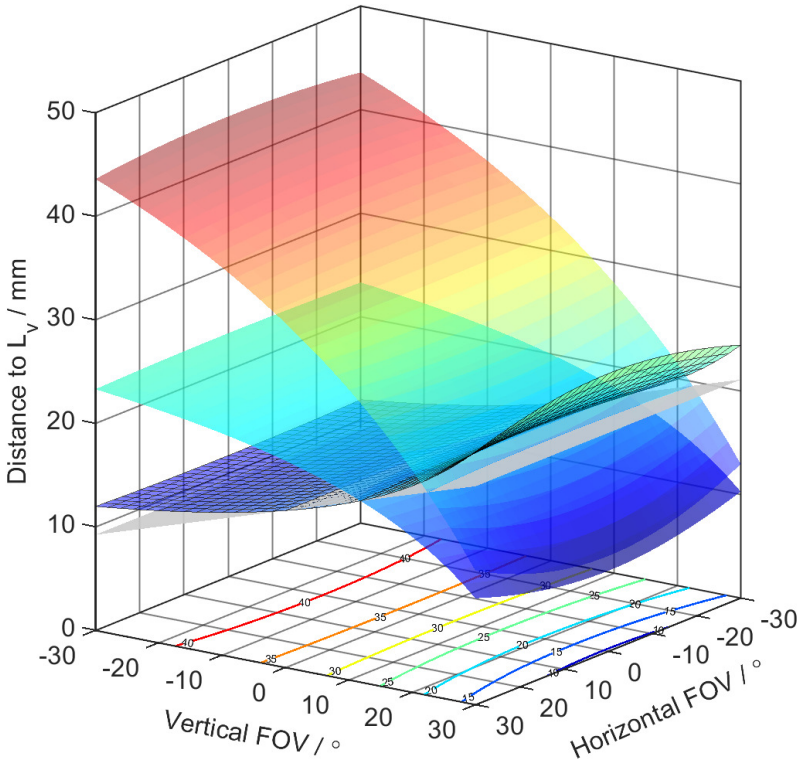

Figure 3.8: Distributions of transmitted image points by tracing beams through the ellipsoidal as the objective lens (upper colorful surface), then the concave lens (lower colorful surface), and inversely from the retina pixel to the LCD plane (meshed surface).

face in Fig. 3.8, which spreads in space from the vertex of 26.8mm at the FOV of H0°×V30° to the nadir of 11.7mm at the FOV of H0°×V-30°. 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. 3.8 is calculated. With the FOV of H0°×V0° kept to be in-focus, RMS (root mean square) of defocus between pixels aligned with the optimized P_d plane and the ideal distribution is minimized.

Fig. 3.9 (a) shows the bench-top prototype that is built based on the optimized structure. The paired ellipsoidal mirrors, the transmissive SLM, and the LCD are all the same as the those used in the previous prototype. As shown in Fig. 3.7,

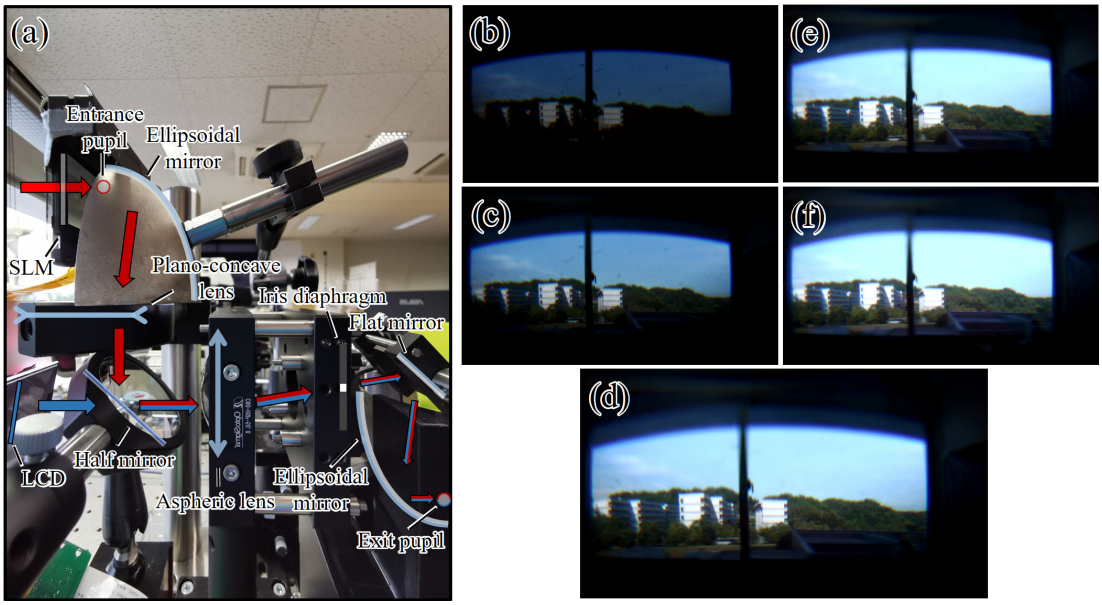

Figure 3.9: **(a)** The bench-top prototype of the proposed method. **(b)** - **(d)** See-through views are recorded without the front SLM by enlarging the pinhole aperture from 1mm to 5mm..

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°, which leads to a compressed vertical parallax of 129.6mm and an enlarged horizontal parallax of 130.8mm. 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. 3.7.

The preliminary experiment has indicated a noticeable resolution change on the perceived image by setting different pinhole apertures. The image taken with a pinhole aperture of 3mm shows a moderate level of resolution, as shown in Fig. 3.6. Moreover, the pinhole aperture also impacts the image brightness since it functions the aperture stop in the system. Therefore, we test the influence on image brightness caused by pinhole aperture change from 1mm to 5mm. Experiment results are shown at the right of Fig. 3.9. The camera is set with fixed

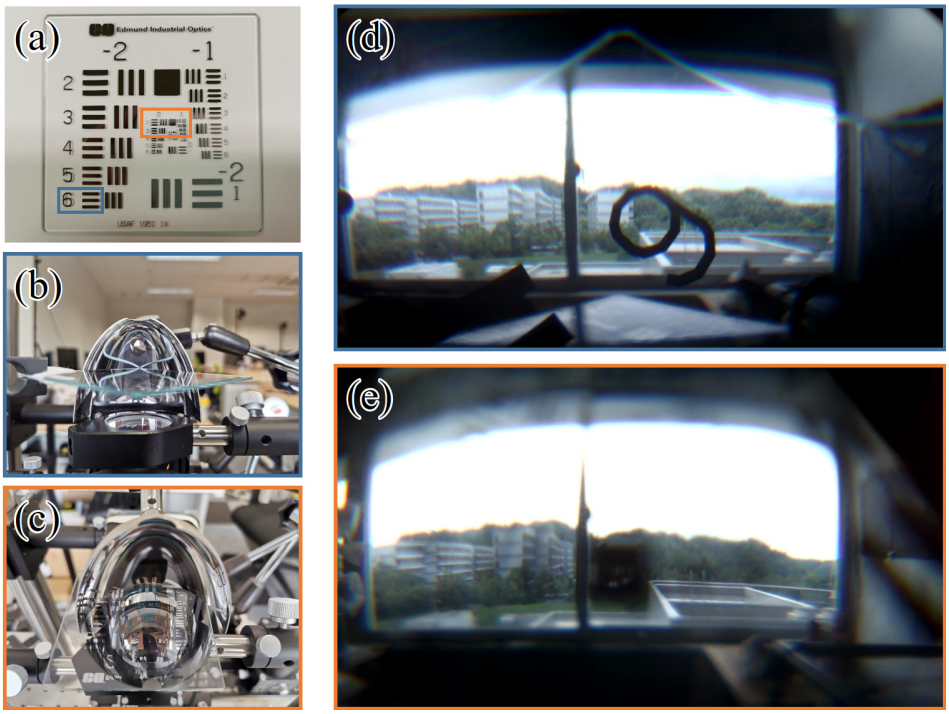

Figure 3.10: Hard-edge occlusion test with the optimized prototype.

exposure parameters, and the transmissive SLM is taken off in the experiment for better visions. with the pinhole aperture being enlarged, the image looks highly dark in Fig. 3.9 (b) and turns bright Fig. 3.9 (f). And the improvement of image brightness is majorly achieved by increasing the pinhole aperture from 1mm to 3mm, then roughly keeps stable when the pinhole is further enlarged. The simulation by ZEMAX also shows a similar trend of see-through efficiencies versus pinhole apertures for the prototype without the SLM, which is from 0.6% @ 1mm to 12.3% @ 5mm compared to naked eyes with a pupil diameter of 2mm, and a moderate pinhole aperture of 3mm achieves a relative high see-through efficiency of 5%. With taking the loss of optical power caused by a transmittance 10% of the SLM, the global see-through efficiency of the prototype is around 0.5% when setting the pinhole aperture as 3mm. Consequently, we set the pinhole aperture as 3mm for the prototype, and the following experiments are all taken with this configuration in the case that no announcement is additionally made.

Although hard-edge occlusion is feasible in the T-PEM system by fixing the

occlusion panel at $SLM_{(op2)}$, the large dimension of the transmissive SLM used for our prototype prevents the installation inside the ellipsoidal mirror. Therefore, we used a transparent USFA 1951 target as the occlusion panel alternatively to test the hard-edge occlusion performance. The experiment results are shown in Fig. 3.10. The recorded FOV by the camera is H95.3°×V52.9°. Patterns bounded with a blue rectangle and an orange rectangle in Fig. 3.10 (a) have the resolution of 0.445 line pairs per mm (lp/mm) and 1.12lp/mm (for the largest line pair), respectively. The blue-bounded pattern is located at the focal plane for image pixels from the lower part of the recorded FOV, which is shown as Fig. 3.10 (b). A sharp pattern is perceived in Fig. 3.10 (d). In the case that the orange-bounded pattern deviates the correct focal plane, which is shown as Fig. 3.10 (c) that the target is placed beneath the ellipsoidal mirror, the occlusion pattern becomes severely blurred as shown in Fig. 3.10 (e). In conclusion, hard-edge occlusion for a designated FOV with the prototype can be realized by alternatively fixing a customized SLM panel at the corresponding focal plane. Based on the analysis of the imaging property of the ellipsoidal mirror, a tilted SLM panel is expected to expand the hard-edge occlusion-capable. In conclusion, hard-edge occlusion can be conducted by the T-PEM system, while the dimensions of available SLM panels are too large to be placed at SLM_{po2} in our prototype. Therefore, the SLM is chosen to be placed at SLM_{po1}, thus the prototype works in the enhanced soft-edge occlusion way.

The modulation transfer function (MTF) curves for both the tangential (T) and the sagittal (S) planes at different see-through FOVs are calculated by ZEMAX, as shown in Fig. 3.11 (a), the diffraction by the SLM is ignored. 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°×V-25° occurs at the spatial frequency of 3.1cpd and even at 0.2cpd for the FOV of H0°×V25°. Consequently, the proposed system only provides a sharp see-through view along with a band-like area around the designated vision, and

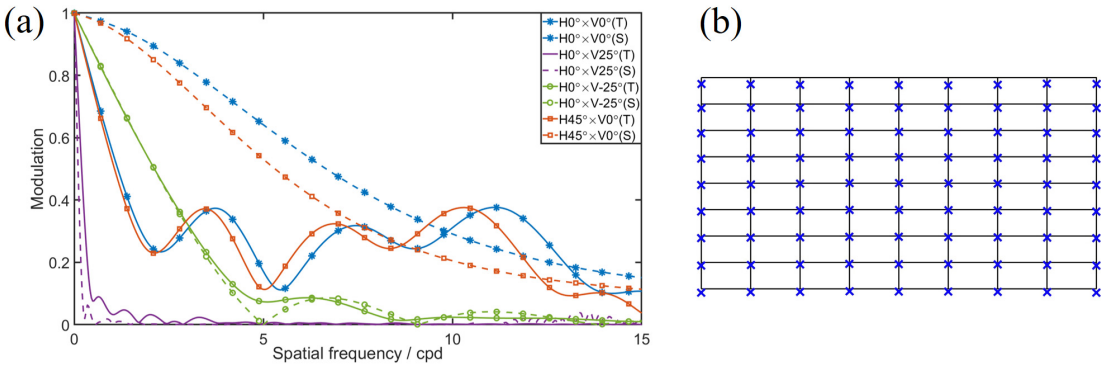

Figure 3.11: **(a)** The MTF curves at different see-through FOVs and **(b)** the gird distortion within the see-through FOV of H95.3°×V52.9°.

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%). 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 zero diffraction [78]. Hence, the see-through view observed through a transmissive LCD is worse than the simulation. The grid distortion within a FOV of H95.3°×V52.9° is shown as Fig. 3.11 (b). Overall, the proposed system keeps minor distortion. A maximum distortion of −7.3% is observed at the central top vision.

In practice, misalignment between different optical elements may occur in the optical system building, thus the system performance will slightly deteriorate from the theoretical values indicated with Fig. 3.11. Regarding the T-PEM system, the system performance is minor influenced by the deviation of optical elements along with the optical axis. In the case that on-axis misalignment occurs, the focused FOV of the T-PEM system shifts from the designated area due to the volumetric distribution of the focused image points by the ellipsoidal mirror, while the image resolution and distortion roughly keeps stable. The off-axis misalignment impacts the system performance a lot, and the two ellipsoidal mirrors in the T-PEM system are more likely to have deviations in practice because they are usually mounted in customized frameworks, thus the fixation and calibration are more difficult than the lenses and mirrors mounted in standard optical mounts.

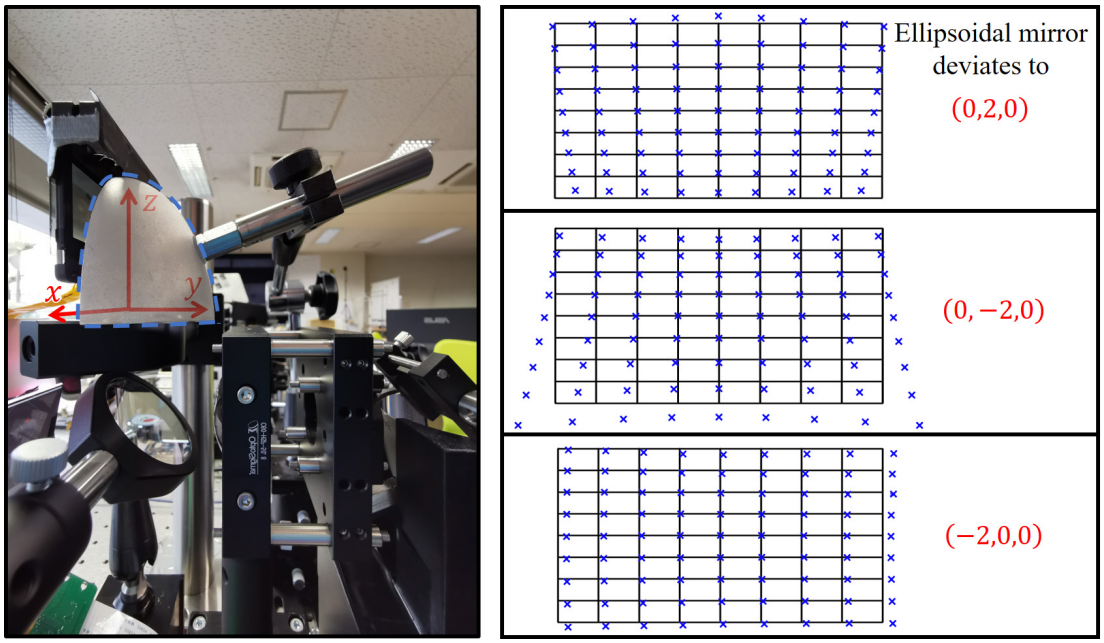

Figure 3.12: Enlarged distortion by the misalignment of the ellipsoidal mirror based on the ZEMAX simulation. The ellipsoidal mirror is set to have deviations of 2mm (top-right) and −2mm (middle-right) in the y direction, and −2mm in the x direction (bottom-right).

According to the simulation by ZEMAX, image resolution is less influenced by the off-axis misalignment, while an obvious distortion increase in the see-through view is brought by the horizontal deviation of the two ellipsoidal mirrors in the side elliptic segmentation plane. Fig. 3.12 shows different grid distortion of the T-PEM system by moving the bounded ellipsoidal mirror away from the ideal location in the xy plane. Compared with the minimum distortion shown in Fig. 3.11, a maximum distortion of −11.9% occurs at the bottom-left and the bottom-right corners of the simulated FOV of H95.3°×V52.9° when the ellipsoidal mirror is moved by 2mm in y direction. Due to the non-symmetric elliptic segmentation of ellipsoidal mirrors, the enlarged distortion becomes more notifiable with the maximum value 23.5% at the bottom-left and the bottom-right corners of the tested FOV with the −2mm deviation in y direction. The deviation in the x direction shows the symmetric influence on the image quality. The maximum distortion in the tested FOV is increased to −7.8% with the deviation of −2mm.

In addition, tilts of the ellipsoidal mirror by the z axis also bring some extra distortion, the see-through view is distorted more severely with the forward or backward tilt happens.

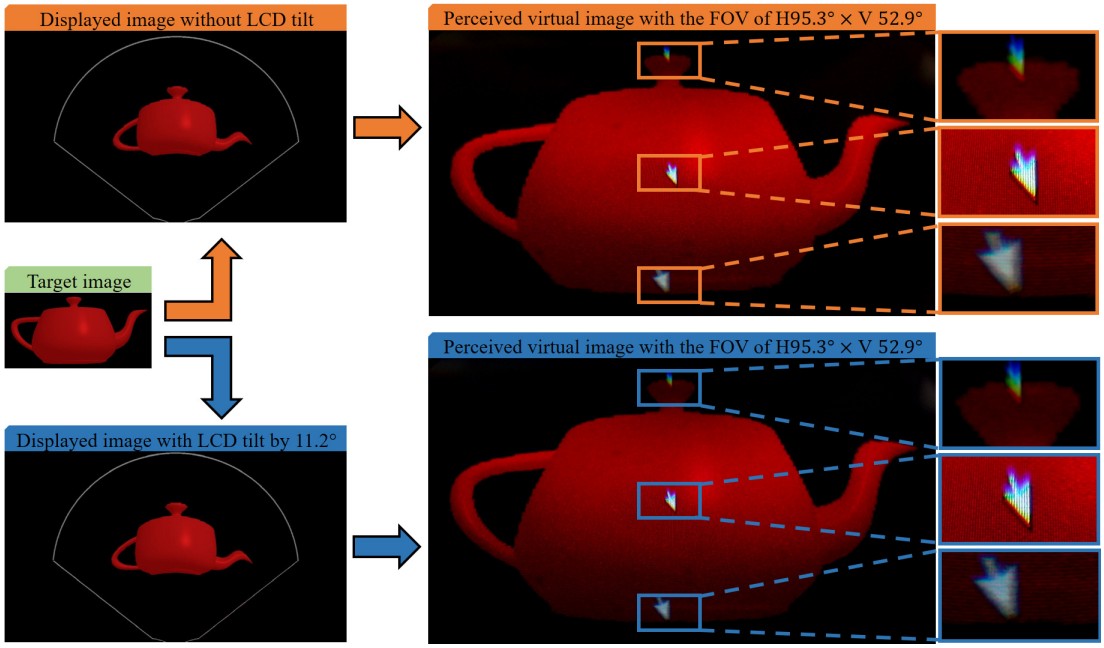

Figure 3.13: Virtual images displayed with and without LCD tilt. A red teapot is given as the target image.

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. 3.13. A red teapot is used as the target image. Displayed images by the LCD are real-time rendered 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°×V105° in the experiment). With the see-through scene blocked, recorded images are taken with a FOV of H95.3°×V52.9°, and the central view is focused. The two virtual images have the same distortion as the see-through view. In order to highlight the difference between perceived virtual images with and without LCD tilt, we additionally put a cursor at the up, mid-

dle, and down parts of each projected image to take zoom-in images, which are shown at the most right 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.

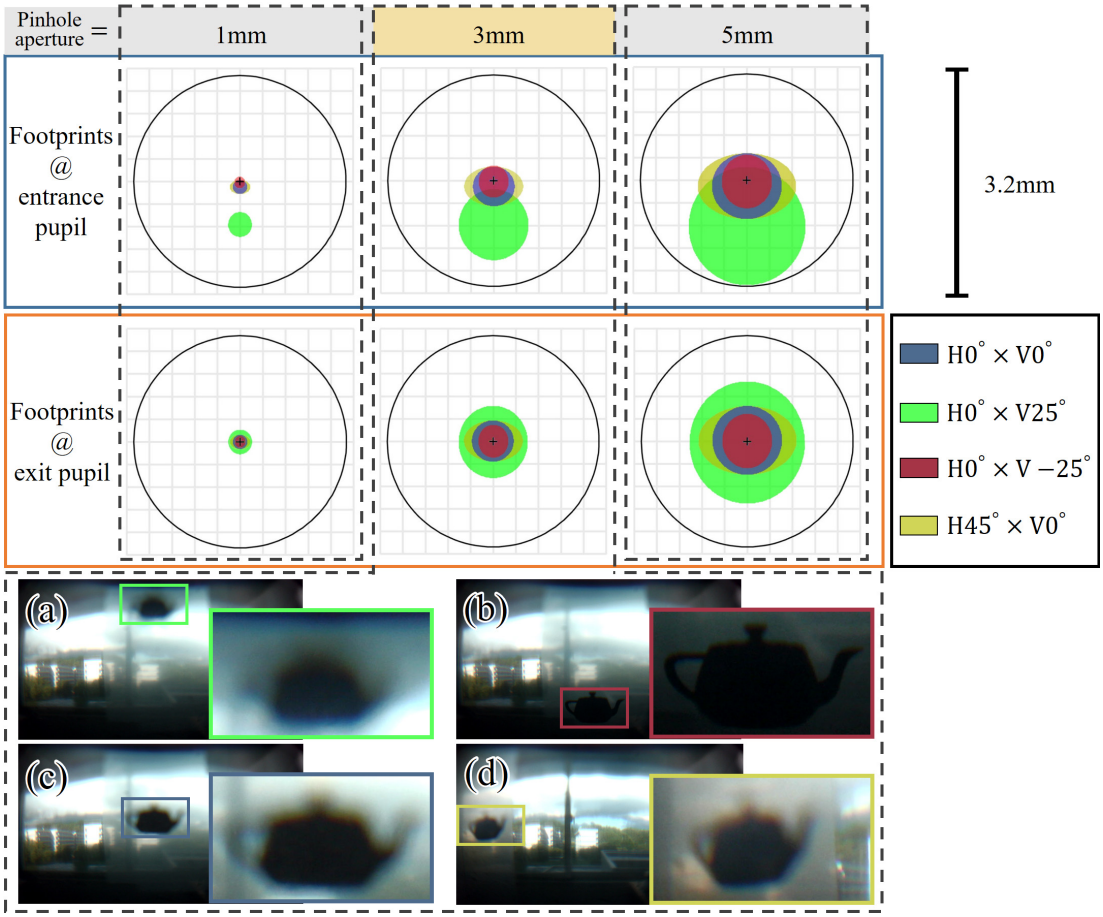

Figure 3.14: Footprints diagrams at the entrance pupil and the exit pupil with different pinhole apertures are shown above. And the occlusion performance by locating the same occlusion pattern different vision areas are shown below.

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 ZE-

MAX. Footprints diagrams are shown in Fig. 3.14. 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°×V-25° to H0°×V25°. In addition, circular footprints at H0°×V0° are projected into ellipses at H45°×V0°. In the case that an SLM is placed at SLM_{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 SLM_{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°×V0°, 0.40mm @ H0°×V-25°, 0.96mm @ H0°×V25°, and 0.68mm @ H45°×V0° (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 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°×V52.9° is recorded in the figures. From Fig. 3.14 (a) to Fig. 3.14 (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 most left of the FOV in Fig. 3.14 (d). Despite 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. 3.14 depict the eye-box of the proposed system. In the case that all FOVs are observed, the user's pupil overlaps the exit pupil, which gives the eye-box of 0.19mm, 0.56mm, and 0.96mm (take 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 pupil impacts the image resolution minor as shown in Fig. 3.6. Users with eyes problems like short-sight are allowed to have a clear see-through view without eyeglasses.

Fig. 3.15 shows mutual occlusion is conducted by the prototype. The pinhole

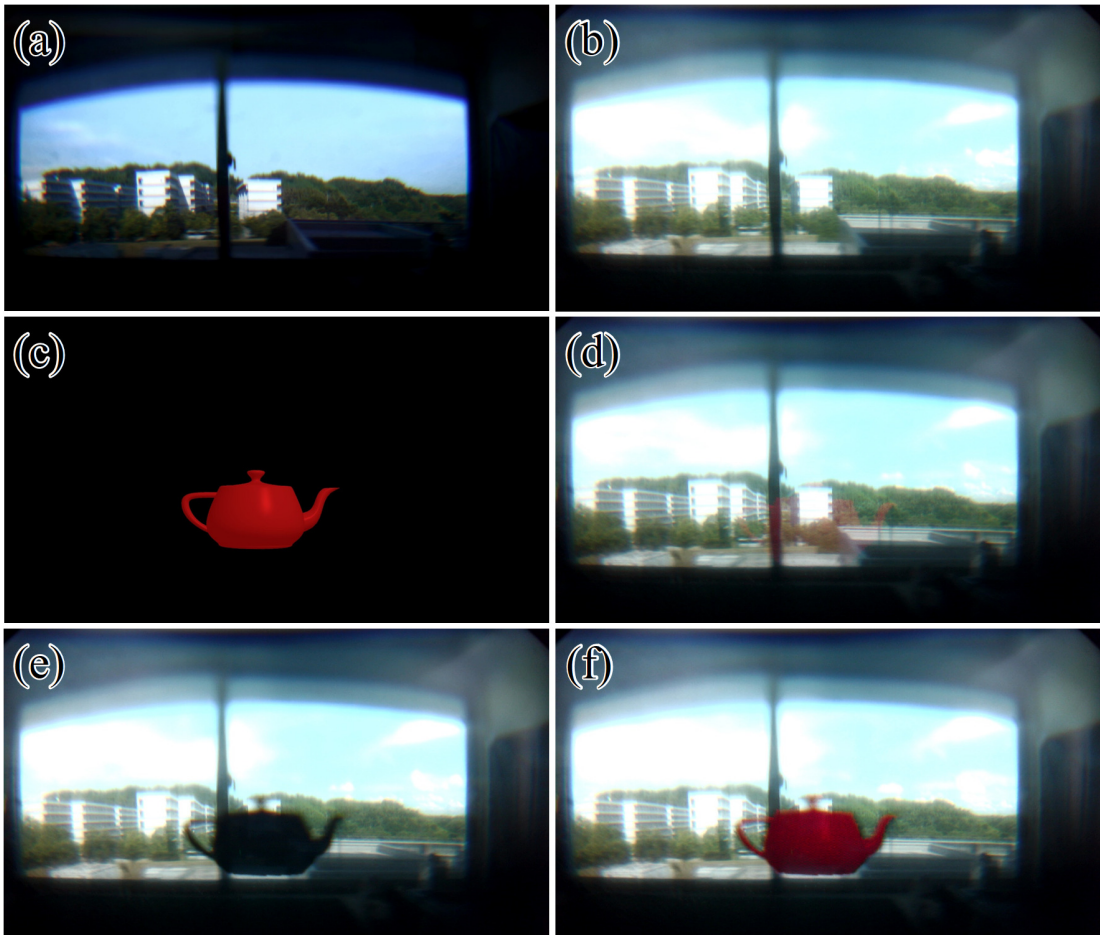

Figure 3.15: Mutual occlusion is demonstrated with a FOV of H95.3°×V52.9°.

aperture is chosen as 3mm. A see-through scene without the occlusion panel that mounted before is shown as Fig. 3.15 (a). After the transmissive 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. 3.15 (b). Notice that the camera was switched to a different exposure mode, so that makes Fig. 3.15 (b) looks brighter than Fig. 3.15 (a). Fig. 3.15 (c) shows the input image of the rendering pipeline. The displayed image by the prototype without mutual occlusion is shown as Fig. 3.15 (d). The red teapot looks highly transparent on the bright background, which makes it mostly invisible. Furthermore, missed detail information, such as the lighting effect, reduces the realism of the teapot. Fig. 3.15 (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. 3.15 (f). With the background being blocked, both the framework and the lighting effect of the teapot are clearly perceived. Besides, the variation of pinhole aperture impacts system performance as mentioned above.

We also chose pinhole apertures of 1mm and 5mm to test the corresponding display and occlusion performance. Experiment results are shown in Fig. 3.16. Figures taken with the pinhole aperture of 1mm and 5mm are placed above and below, respectively. In comparison with the occlusion performance shown in Fig. 3.15, the configuration with 1mm pinhole apertures conducts occlusion with the performance close to hard-edge occlusion approaches, and the configuration with 5mm pinhole apertures rendered the occlusion pattern with lower resolution and contrast. However, the image quality of the see-through real scene and the projected virtual image also varies with the pinhole aperture change, which has been analyzed before.

3.3 System limitations

In summary, the T-PEM system for realizing wide-view mutual occlusion in OST-AR displays is introduced in this chapter. The NA of the imaging system is efficiently increased by the combination of paired ellipsoidal mirror structure and aperture stop restriction. In terms of beam propagation being analyzed, hard-edge occlusion or enhanced soft-edge occlusion is optional to conduct. And prototypes built in the enhanced soft-edge occlusion way are analyzed and tested. As a result, virtual display in a FOV of H160°×V74° and mutual occlusion in a FOV of H122°×V74° are demonstrated with a basic design, respectively. Furthermore, a mixed FOV of H95.3°×V52.9° is demonstrated by an optimized design with vertical parallax reduction and virtual display improvement.

However, some limitations still need to be addressed for approaching a practical solution. Although the optimized design has reduced the vertical parallax significantly, it is hard to build a vertical parallax-free OC-OST-AR display with the current T-PEM system. What is more, no matter the T-PEM system is chosen to

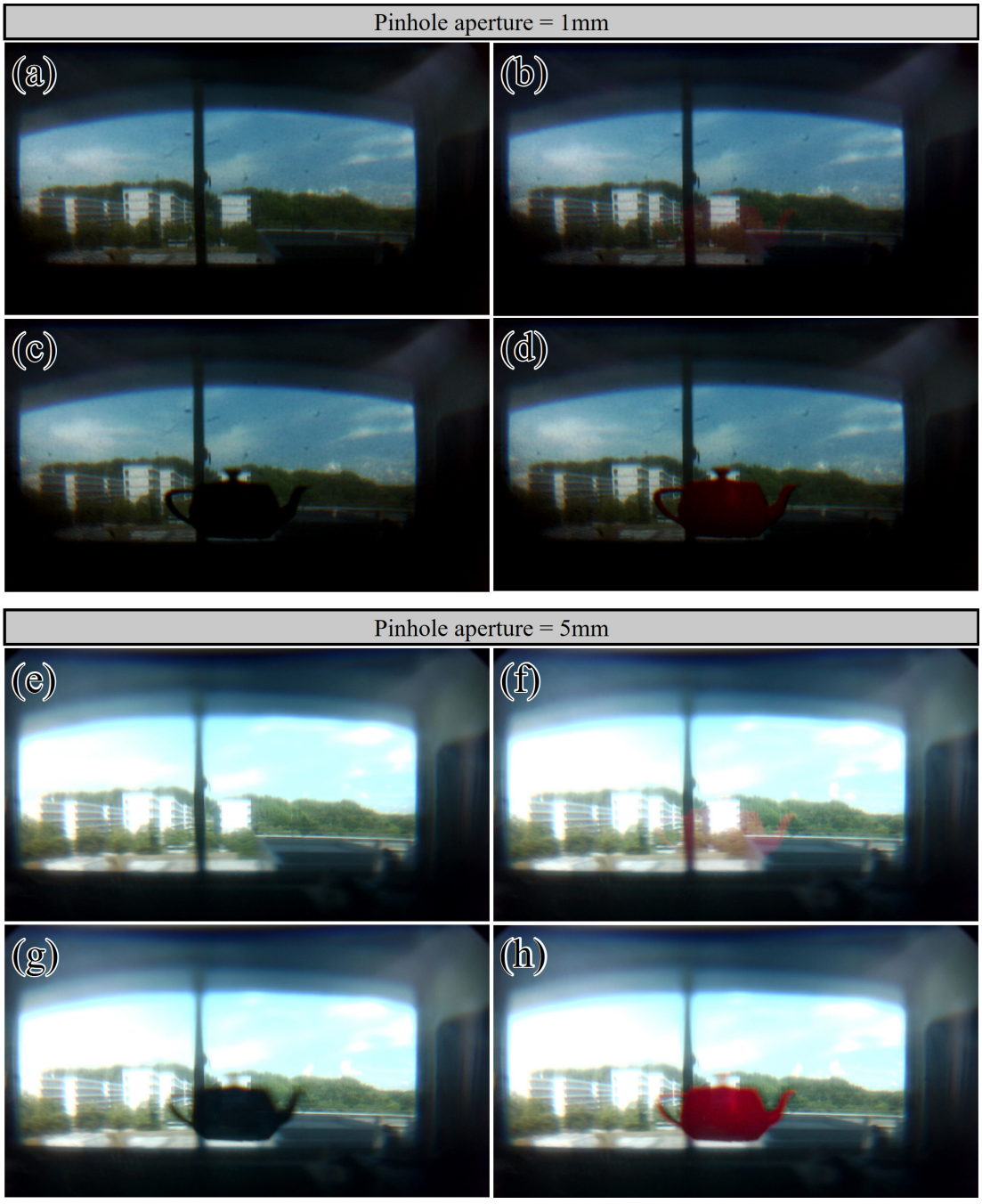

Figure 3.16: Mutual occlusion is conducted with pinhole apertures of 1mm and 5mm.

work in the hard-edge occlusion way or the enhanced soft-edge occlusion way, the diffraction and low transmittance of the transmissive SLM for rendering occlusion patterns is a considerable challenge for improving system performance. Even though some works have reported more powerful liquid crystal modules that can reach a transmittance around 70%, pixel-wise modulation has not been demonstrated [71, 72]. The current progress of liquid crystal is expected to support global dimming in OST-AR displays for enhancing the visibility under excessive environment light, but the mutual occlusion between the real object and the AR content is still an issue since the pixelated modulation is necessary [57, 3, 57]. In the future, practical solutions for OC-OST-AR displays can be developed from the proposed T-PEM system with further development on liquid crystal technology. While a more feasible solution currently may be the use of reflective SLMs, since many OC-OST-AR displays have reported better image quality and occlusion performance by using the LCoS [82, 83]. However, the converged ray bundles reflected by the ellipsoidal mirror mismatches the mechanism of LCoS. A new structure needs to be proposed for addressing the two issues.

Therefore, we turned to exploring the combination of LCoS technologies and the PCR structure. The new design is expected to support not only mutual occlusion in a wide FOV but also a parallax-free view while providing better performance than the T-PEM system. As a result, the reflective SLM-based paired-paraboloidal-mirror system is proposed. The details are shown in the next chapter.

4 Reflective SLM-based paired-paraboloidal-mirror (R-PPM) structure

4.1 Optical architecture

The use of LCoS as reflective SLMs has shown substantial advantages, including better occlusion precision, higher contrast ratio, and less artifacts diffraction, on transmissive SLMs [82, 83, 65, 38, 37]. OC-OST-AR displays built on the PCR structure is expected to achieve better system performance in the case of an LCoS rather than a transmissive SLM is implemented as the occlusion panel. Nevertheless, it is hard to continue building an OC-OST-AR display by simply combining the joint ellipsoidal mirrors and an LCoS. Rays entering through a focus of an ellipsoidal mirror are reflected towards another focus as shown in Fig. 3.1. While an LCoS usually works with rays coming from the right front, amplitude or phase of the incident light are modulated after reflected from the LCoS surface [21, 77]. However, the reflected ray bundle is converged by the ellipsoidal mirror. Hence, replacing the transmissive SLM with an LCoS in the T-PEM system is not feasible. Integrating the LCoS into the PCR structure requires a new optical architecture.

Therefore, we turned to build a system on another type of conical reflector, which is the paraboloidal mirror. Paraboloidal mirrors have been widely used in many imaging systems. A large number of paraboloidal mirror-based spectroscopy [69, 10] and microscopy [46, 16, 44] are developed in decades. Paraboloidal mirrors are demonstrated to break the theoretical FOV limitation of typical lens-based systems in those systems. Motivated by previous researches,

a R-PPM structure for building wide-view OC-OST-AR displays is designed. Compared to the T-PEM structure, both the image quality and the occlusion performance are improved in the R-PPM system. Moreover, the R-PPM system is designed with a loop layout. As a result, a vertical parallax-free OC-OST-AR display can be developed from the proposed structure.

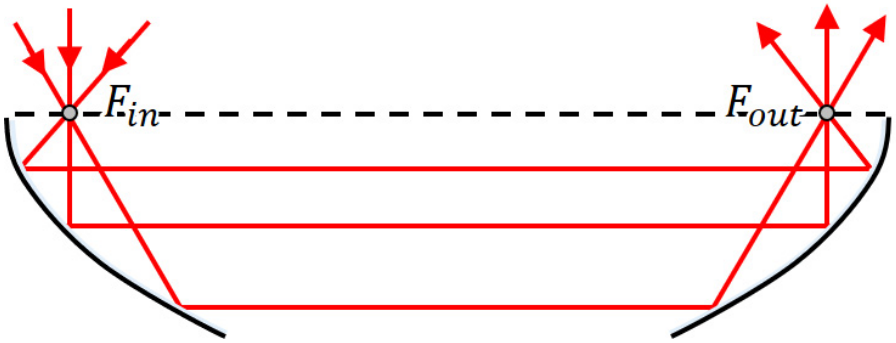

Figure 4.1: Rays propagate through each focus of two opposite paraboloidal mirrors.

A basic layout with two paired paraboloidal mirrors is shown in Fig. 4.1. The two paraboloidal mirrors are placed oppositely with sharing the same optical axis, and foci of the two paraboloidal mirrors are labeled as F_{in} and F_{out}, respectively. It is well-known that paraboloidal mirrors reflect rays passing through the focus into parallel. Thus the paired paraboloidal mirror layout shown in Fig. 4.1 transfers incident rays from F_{in} into another focus F_{out} with twice reflection by the paraboloidal mirrors, which is similar to chief rays by the joint ellipsoidal mirror structure. Whereas the exit angle of a transmitted ray is different from the incident angle, a more complex optical architecture is required to keep a real scene be erected displayed with the paired paraboloidal mirrors. What is more, a paraboloidal mirror transmits incident beams in a different manner from an ellipsoidal mirror, while hard-edge occlusion needs to be conducted in an inner focal plane of an imaging system. Thus, analysis of the imaging property of a paraboloidal mirror is essential for building the R-PPM system.

We firstly addressed the challenge of see-through view erecting by introducing a loop layout of optical elements and implementing a roof prism. The drawback of the vertical parallax in the T-PEM system is solved at the same time. A

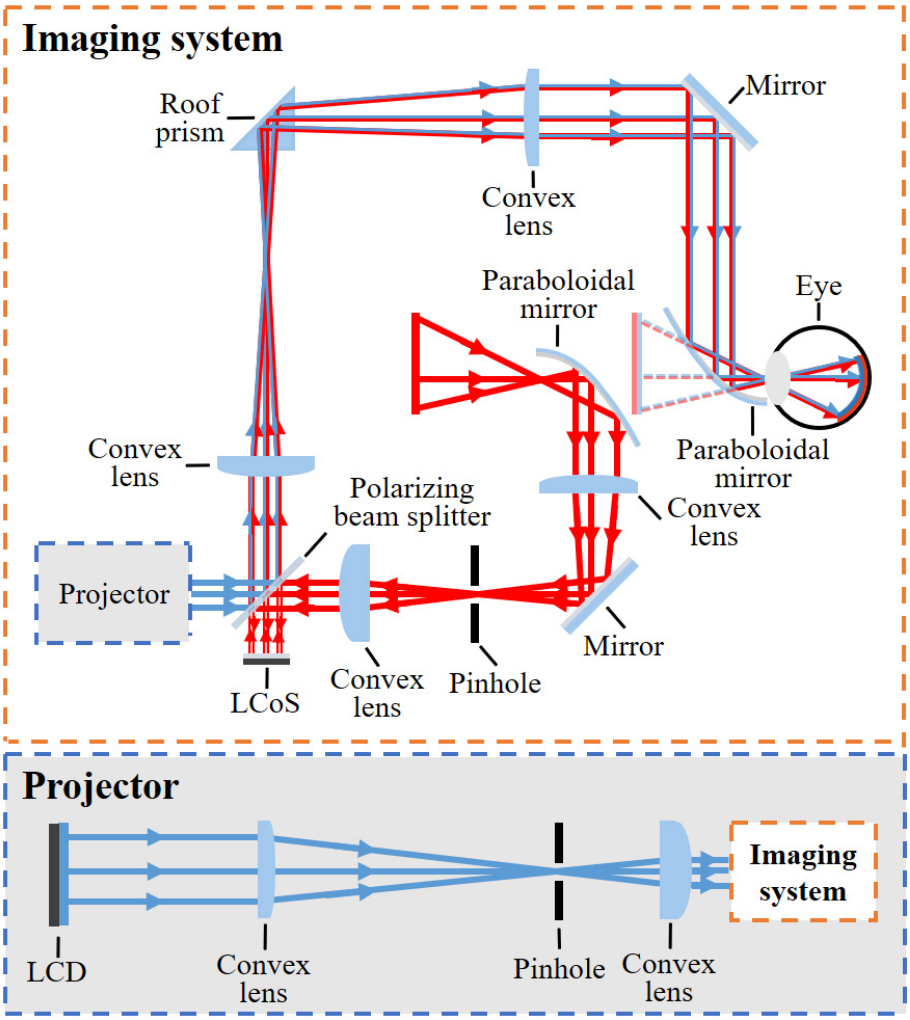

Figure 4.2: Schematic diagram of the R-PPM system. Rays from the real scene and the LCD are marked as red and blue lines.

schematic diagram of the proposed R-PPM system is shown in Fig. 4.2. The entire imaging system bounded with the orange rectangle is drawn above. Red chief rays from various directions enter the system through the focus of the first paraboloidal mirror. In terms of the well-known property of the paraboloidal mirror, these chief rays at large incident angles are transformed into parallel rays before reaching the LCoS. Off-axis aberrations, such as coma, are thereby significantly reduced even with a large FOV. A 4-lens structure that consists of

two paired convex lenses is designed to adjust the optical path of the system for eliminating the vertical parallax. Furthermore, the magnification of the input image is controlled by setting the focal lengths of these lenses since existing LCoSs are usually dimension limited. A roof prism on the top-left corner flips the propagation image left-and-right that ensures the eye, which is located at the focus of the second paraboloidal mirror, observes the input image correctly. Similar to the T-PEM system, an adjustable pinhole mask is used as an aperture stop to balance the brightness and resolution of the display image here.

The projector for projecting virtual images is separately drawn in the blue rectangle below. Virtual images are merged with the real scene by a polarizing beam splitter upon the LCoS. Similar to the imaging process of the real scene, a lens pair is implemented to shrink the original digital image given by the LCD in the projector, so that the requirement of the pixel density of the LCD for displaying high-resolution virtual images is attenuated. And the trade-off between the brightness and resolution of the virtual image is also controlled by a pinhole between the lens pair. Overall, the proposed system enables that sharp virtual images and hard-edge occlusion are rendered simultaneously with a wide see-through view.

Although the overall layout of the R-PPM system is designed, the detailed arrangement of each optical element still needs to be determined since the LCoS should be located at the plane that the propagated real scene is focused. Therefore, the analysis of the imaging property of a paraboloidal mirror is essential for deriving the beam propagation through the whole system. In terms of previous research, a paraboloidal mirror functions as a high NA lens when light from the object incident through the focus [2, 75, 76]. Thus, the reflection of a beam through the focus by the paraboloidal mirror can also be processed by the Gaussian lens equation shown in Eq. 3.3. However, the projection of an object point by the paraboloidal mirror is calculated more simply in this dissertation. In the case of dominant aberrations by the paraboloidal mirror (astigmatism and coma) being ignored, the focal length of a beam reflected by the paraboloidal mirror is equal to the distance between the focus and the intersected point of the chief ray at the mirror surface when the propagation beam is slim (since we use a pinhole to block most of the input beam in the system) [64]. This property of the

paraboloidal mirror leads to a more direct way to solve the projection between the target AR content and the projected images by the LCD and the LCoS. Thus alleviate the computation of the image and occlusion rendering in the R-PPM system.

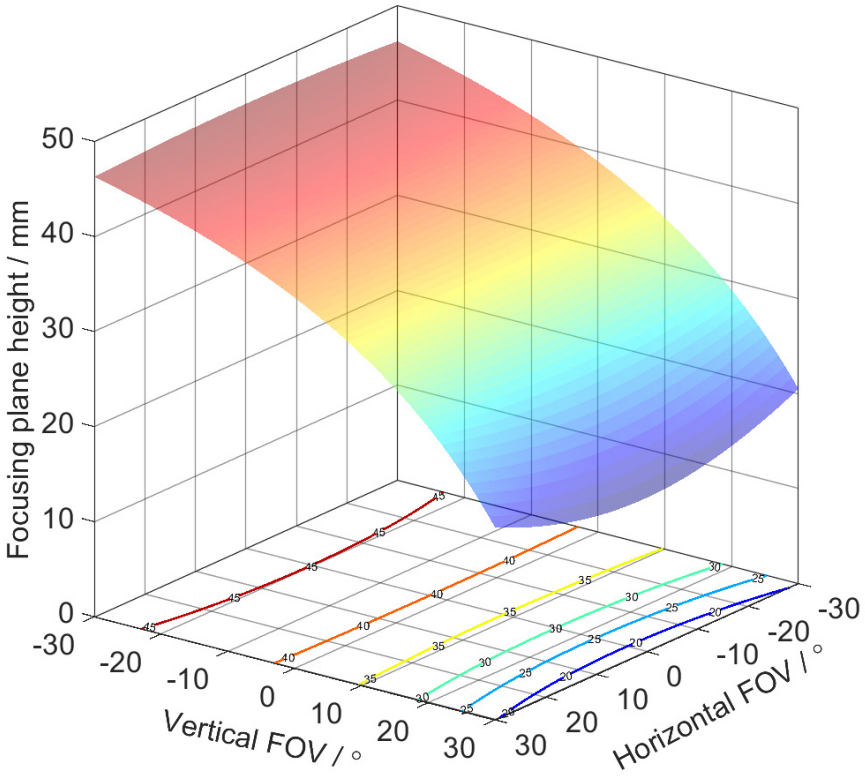

Figure 4.3: Distribution of reflected image points by a paraboloidal mirror. Incident beams are from a FOV of H60°×V60° and enter the paraboloidal mirror through the focus.

Then, we calculated focusing plane heights of object pints within a FOV of H60°×V60° reflected by a paraboloidal mirror. The focal length of the paraboloidal mirrors is set as 6mm, and the result is shown in Fig. 4.3. In order to show the comparison to ellipsoidal mirrors, the focusing plane height in Fig. 4.3 is measured from a plane below the latus rectum by a distance of 50.8mm, which is also the focal length of the ellipsoidal mirror used in the T-PEM system. And

contours are also drawn beneath for better visualization. In general, the distribution of the reflected image points by a paraboloidal mirror is similar to the case by an ellipsoidal mirror, while reflected image points from the same FOV spread in a narrower space whose height is from 46.8mm @ H0°×V-30° to 16.4mm @ H0°×V30°. Thus, patterns shown by a planar occlusion panel suffer less defocus in the case that an SLM follows the paraboloidal mirror directly, which is similar to the layout in the T-PEM system. Therefore, the R-PPM system is expected to render hard-edge occlusion more precisely. Furthermore, the LCoS in the R-PPM system is positioned after a lens pair. As a result, the transmitted real scene pixels spread in an even narrower space, allowing hard-edge occlusion to be rendered in a wide FOV.

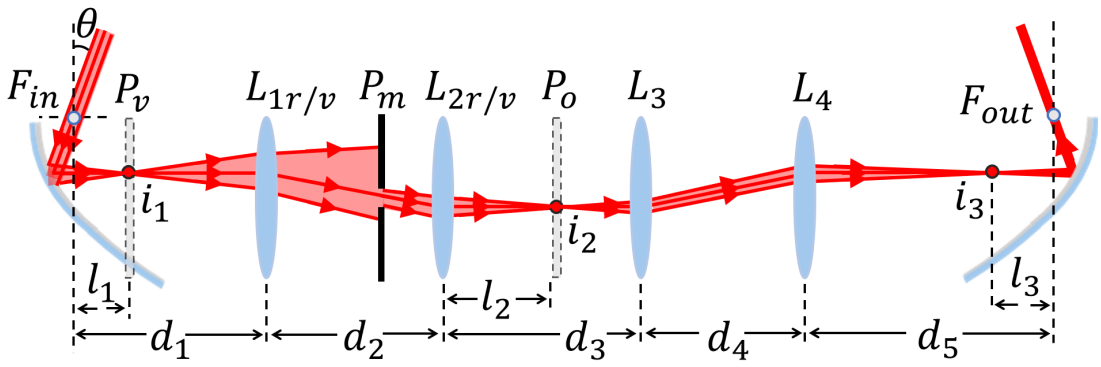

Figure 4.4: The unfolded optical layout of R-PPM structure.

The exact layout of the proposed R-PPM system is shown with the unfolded structure in Fig. 4.4. Besides the two paired paraboloidal mirrors, optical layouts for the real scene and virtual display are both built with 4 convex lenses from L_1 to L_4. In order to make a sharp occlusion mask for the see-through view while rendering virtual content with sufficient resolution, L_1 and L_2 are chosen with different parameters, where L_{1r}, L_{2r} are used for the see-through view and L_{1v}, L_{2v} are used for virtual display. The beam from an arbitrary direction of the real scene enters the system through the focus of the first paraboloidal mirror F_{in} with the angle θ then exits to the user pupil through the focus of the second paraboloidal mirror F_{out}, and it is focused at points i_1, i_2, and i_3 through the optical path. P_v and P_o are the LCD for virtual display and the LCoS for

occlusion rendering, respectively. d_1 and d_5 are distances from F_{in} and F_{out} to L_1 and L_2, respectively. And d_2, d_3, and d_4 are the spacings between L_1, L_2, L_3, and L_4, respectively. Resulting from the fact that chief rays through F_{in} are transformed into parallel rays, and the mechanism of the LCoS, the first lens pair of L_1 and L_2 are designed as a double telecentric lens by placing the pinhole plane P_m at the shared focal plane, which gives the lens spacing is:

$$d_2 = f_1 + f_2 \tag{4.1}$$

where f_1 and f_2 are the focal lengths of L_1 and L_2. In addition, it is easy to derive the subsequent design since the whole optical layout must keep the global magnification M to be 1 when the two paraboloidal mirrors have same optical attributes. Then we have:

$$d_4 = f_3 + f_4 \tag{4.2}$$

$$M = M_{12} \times M_{34} = \frac{f_2}{f_1} \times \frac{f_4}{f_3} = 1 \tag{4.3}$$

where f_3 and f_4 are the focal lengths of L_3 and L_4, M_{12} and M_{34} are the magnification by L_1, L_2 and L_3, L_4, respectively. Let angle θ measured clockwise from the perpendicular to the chief ray as negative, the distance l_1 from F_{in} to P_v is thereby only determined by θ:

$$l_1 = 2f_p\Big(\tan^2\theta + \tan\theta\sqrt{1+\tan^2\theta}\Big) + \frac{2f_p}{1-\sin\theta} \tag{4.4}$$

where f_p is the focal length of paraboloidal mirrors, which should be a constant for certain configurations. Then it is easy to derive l_2 by the Gaussian lens formula:

$$l_2 = -\frac{pf_2^2 - f_1f_2(f_1+f_2)}{f_1^2} \tag{4.5}$$

where $p = d_1 - l_1$ is the object distance referring to L_1. Eq. 4.5 gives the location of the occlusion panel, and layout of the paraboloidal mirror for the exit pupil should follow:

$$l_3 = d_5 - \Big[(\frac{f_4}{f_3})^2(f_2+f_3-d_3) + (\frac{f_2f_4}{f_1f_3})^2(f_1-p) + f_4)\Big] \tag{4.6}$$

We simplify the equation by substituting Eq. 4.3 into Eq. 4.6:

$$l_3 = d_5 - [M_{34}^2(f_2 + f_3 - d_3) + f_1 + f_4 - p] \tag{4.7}$$

The system expects to let the exit beam at F_{out} be recovered into parallel, which guarantees the image to be well-focused by the user's eye. Therefore, we have:

$$l_1 = l_3 \tag{4.8}$$

As a result, the optical layout should follow the constraint that is given as:

$$2l_1 = d_5 + d_1 - f_1 - f_4 - M_{34}^2(f_2 + f_3 - d_3) \tag{4.9}$$

Eq. 4.1, Eq. 4.2, Eq. 4.3, Eq. 4.4, and Eq. 4.9 determine the configuration of the proposed system altogether. Since l_1 varies with different incident angles, it is obvious that the system performance also differs from each FOV.

4.2 System prototyping and analysis

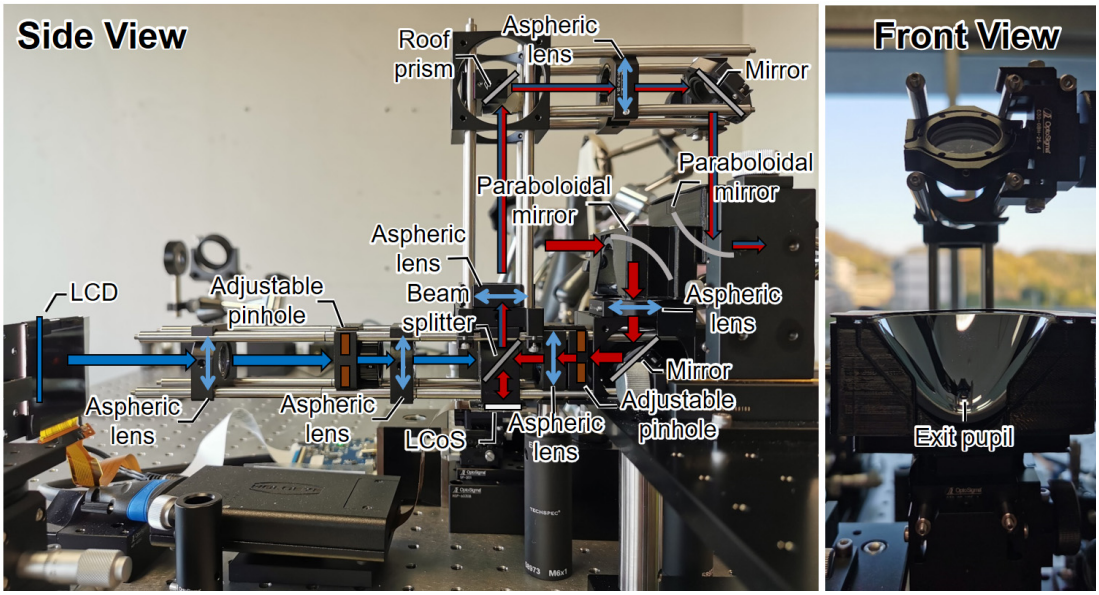

Figure 4.5: Side view (left) and front view (right) of the prototype.

The side view (left) and the front view (right) of a monocular bench-top prototype are shown in Fig. 4.5. Rays from the real scene and the LCD are labeled as red arrows and blue arrows, respectively. The paired paraboloidal mirrors have the same focal length of 6mm and height of 25mm. All lenses used in the prototype have a diameter of 25mm and a clear aperture of 23mm. The LCoS working as the SLM has the dimension of 15.36mm×7.64mm, the resolution of 1920×1080, and the reflectance of 73% (@633nm). It is noticeable that though the form-factor is increased by implementing the two paraboloidal mirrors, around 3/4 of them are redundant due to the limited dimension of other components. The paired lenses L_{1r} and L_{2r} have the focal length of 50mm and 25mm, which gives the magnification M_{12} as 1/2, thus allows a wider FOV with the dimension-limited LCoS. Accordingly, the focal length of the following paired lenses L_3 and L_4 is set as 50mm and 100mm to keep a unit magnification for the see-through view. Due to the attributes of the LCD (the active area is 51.84mm×51.84mm, and the resolution is 1440×1440) used for projecting virtual images, the lens pair in the projector has the focal length 75mm for L_{1v} and 25mm for L_{2v} that gives the magnification of 1/3. The virtual image is rendered based on OpenGL with the algorithm for compensating the optical aberration by single paraboloidal mirror imaging and the global magnification of 2/3 for virtual display. Vision areas of the real scene are selectively focused by shifting the LCoS plane up-and-down, and the virtual image is focused by moving the convex lens L_{2v} horizontally.

Fig. 4.6 shows polychromatic modulation transfer function (MTF) curves given by ZEMAX in both tangential (T) and sagittal (S) planes. The simulated system is set to focus at the lower central view and to work with the pinhole aperture of 1mm. In general, MTF curves show the best performance around the central FOV and become worse with the FOV expanding. MTF30 (Where the average modulation of T plane and S plane drops to 30%) for the FOV of H0°×V0° occurs at the spatial frequency of 30.4cpd (cycles per degree). Along with the vertical direction, MTF30 drops rapidly in the upper vision because of the severe defocus, as shown in Fig.3. It varies from 9.8cpd @ H0°×V5° to 1.6cpd @ H0°×V15°. Conversely, the image quality tends to be much more stable in the lower vision. The MTF curve always keeps above 30% until 35cpd. It even has MTF30 at 6.1cpd for the FOV of H0°×V−30°. Regarding horizontal direc-

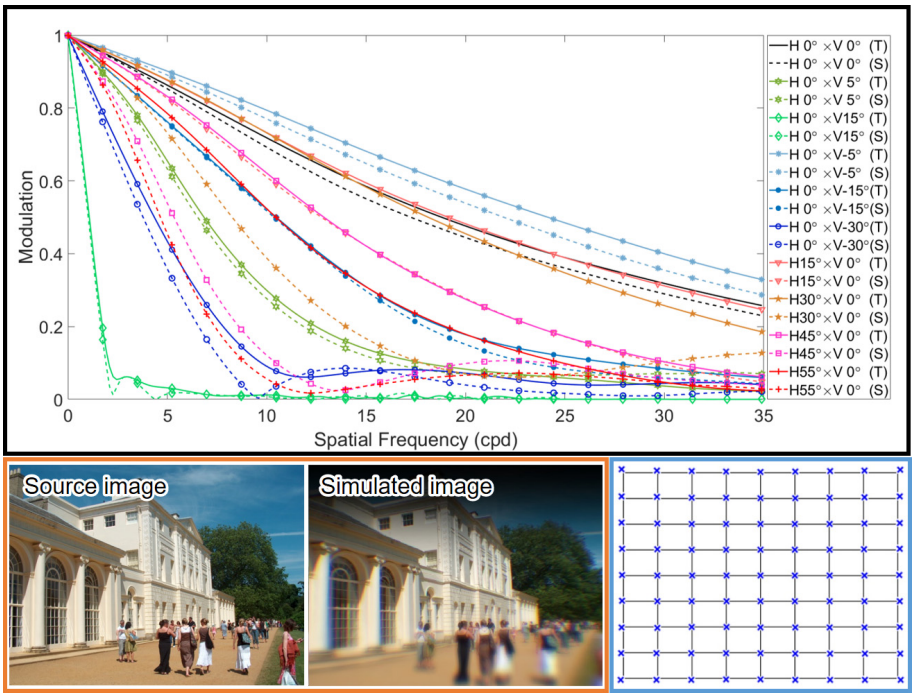

Figure 4.6: (Up) MTF curves for both tangential and sagittal planes. (Bottom-left) A source image is given with an FOV of H85.3°×V64.0°. (Bottom-middle) The simulated image and (bottom-right) the grid distortion calculated by ZEMAX.

tion, MTF30 @ H15°×V0° and H30°×V0° are given with the spatial frequency of 23.9cpd and 18.3cpd, respectively. When the vision further expands to the theoretical maximum of our prototype, MTF30 is found at 9.8cpd for the FOV of H55°×V0°. In general, the proposed system can provide decent image quality for AR experience with most of lower vision, which is also indicated by the smoother variation of defocus on P_o plane in Fig. 4.6. A source image given with a FOV of H85.3°×V64.0° is shown at the bottom-left of Fig. 4.6, the simulated image and the grid distortion by ZEMAX are adjacently placed. A maximum distortion of 2.3% occurs at the top-left and top-right corners of the FOV.

Similar to the T-PEM system, the on-axis deviation of the optical elements impacts the image quality minor. All the lenses, mirrors, and polarizing beam splitter in the R-PPM system are fixed tightly in the cage system, thus off-axis deviations of those components are also constrained. Image resolution roughly

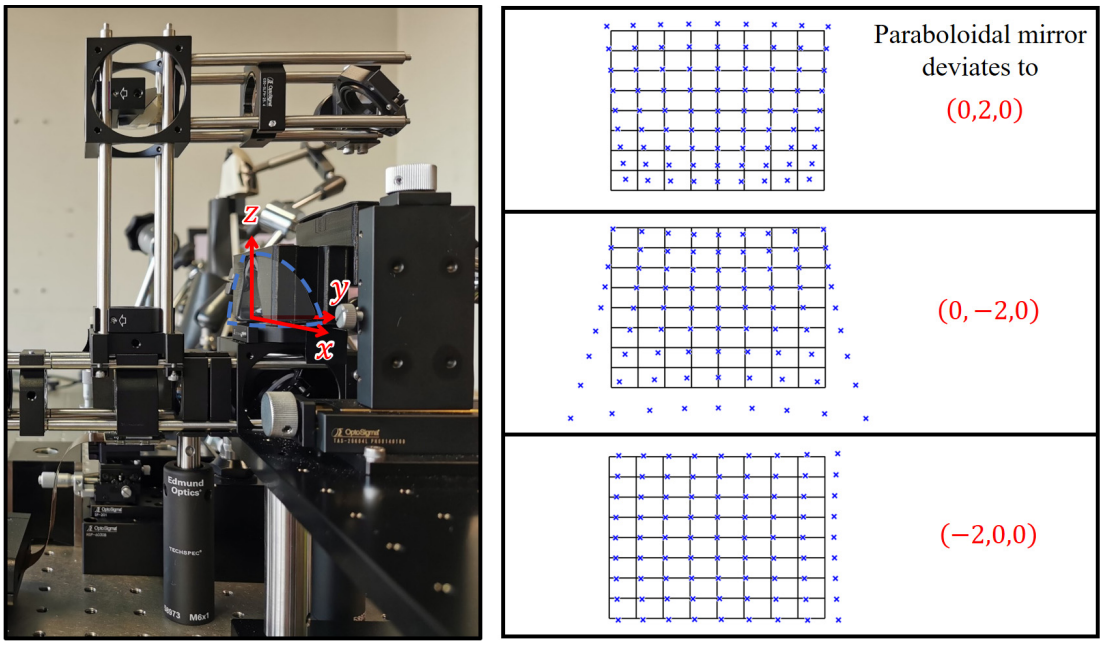

Figure 4.7: Enlarged distortion by the misalignment of the paraboloidal mirror based on the ZEMAX simulation. The paraboloidal mirror is set to have deviations of 2mm (top-right) and −2mm (middle-right) in the y direction, and −2mm in the x direction (bottom-right).

keeps stable with tiny deviations of optical elements in the R-PPM system, and noticeable deterioration of the image distortion is caused by the off-axis deviation of the paraboloidal mirror. Fig. 4.7 shows the grid distortion of the R-PPM system in the same FOV of H85.3°×V64.0° simulated by ZEMAX. The bounded paraboloidal mirror is moved in the xy plane. Compared with the grid distortion indicated by the ideal case in Fig. 4.6, deviations of 2mm and −2mm in the y direction result in obvious distortion enlargement, where the maximum distortion is increased to be −11.9% and 38.3%, respectively, at the bottom-left and the bottom-right corners of the tested FOV. And a −2mm deviation in the x direction enlarges the maximum distortion to 10.2% as shown in the bottom-right image in Fig. 4.7.

The theoretical FOV of the proposed system is calculated as Fig. 4.8 (a). The grey area shows the ideal FOV achieved when all the light through F_{in} is transmitted by the system. Due to the insufficient clear aperture of lenses (23mm), the

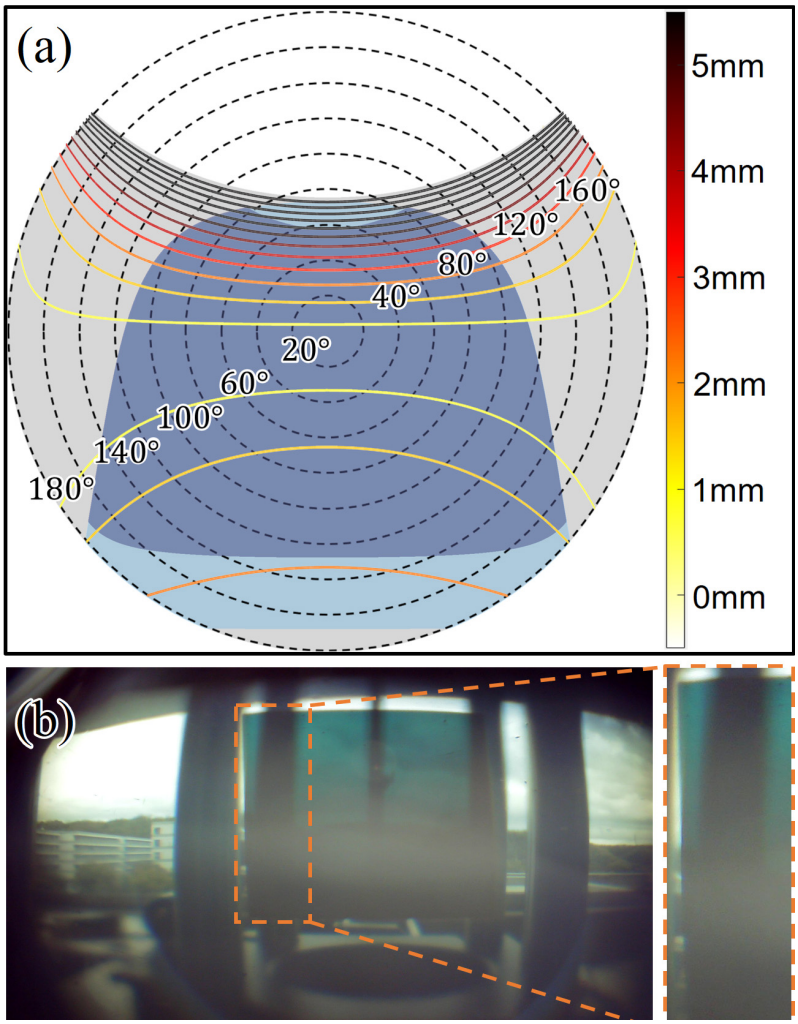

Figure 4.8: **(a)** Theoretical FOVs supported by the paraboloidal mirror (grey), lens aperture (light blue), and the LCoS (dark blue), respectively. Contours with the beside legend indicate the defocus of real scene pixels of each FOV at the plane P_o. **(b)** The contrast of the occlusion pattern decreases gradually from up to down of the FOV.

FOV is shrunk into a light-blue area around H114°×V120°. With further considering the dimension of the LCoS, the FOV shrinks to the dark-blue area with an average width of 114° and a non-symmetrical height from 32° to −63°. Contours with the beside legend indicate the defocus of real scene pixels on the LCoS plane regarding each FOV. Considering the FOV that is recorded with the prototype,

the occlusion pattern suffers a maximum defocus of 5mm to real scene pixels in the upper half-FOV and a maximum defocus of 1.5mm to real scene pixels in the lower half-FOV. Consequently, the occlusion performance, as well as the image quality, are still FOV-dependent in the R-PPM system, and the system shows better performance in the lower half-FOV than the upper half-FOV. However, the defocus majorly reduces the contrast ratio rather than the resolution of the occlusion pattern since the LCoS plane is always in focus by the latter optical elements. A large rectangle occlusion pattern rendered by the prototype with a pinhole aperture of 1mm is shown in Fig. 4.8 (b). The zoom-in image at the right shows an overall sharp occlusion performance. Meanwhile, the lower part of the occluded area looks entirely black while the upper part looks slightly transparent, while results from the gradient defocus of the occlusion pattern.

Although different pinhole apertures also lead to the variation in the occlusion performance in the R-PPM system, the sharpness of the occlusion pattern is less influenced by the pinhole aperture than the T-PEM system. Sub-figures in the right column of Fig. 4.9 show the same occlusion pattern rendered by the prototype with pinhole apertures of 1mm, 3mm, and 5mm. Even though the see-through background becomes more and more blurred, the occlusion pattern keeps sharp since masks projected by the LCoS transmit only the latter half of the prototype. Hence, the pinhole installed in the former half of the prototype only impacts the image quality of the generated real scene. However, a beam from the FOV of H0°×V0° is focused into a larger spot on the LCoS when the pinhole aperture expands. Sub-figures in the left column of Fig. 4.9 shows the corresponding simulated footprints on the LCoS plane, and a scalar of 500um is attached beside. The footprints are given with RMS radii of 5.9um @ pinhole aperture of 1mm, 11.7um @ pinhole aperture of 3mm, 56um @ pinhole aperture of 5mm. As a result, the occlusion pattern rendered with a larger pinhole aperture tends to have a lower contrast ratio, as shown in the recorded images beside.

Similar to the T-PEM system, the entrance pupil and the exit pupil of the R-PPM system are given as the projection of the pinhole by the former and the latter optical element groups, respectively. Different from the T-PEM system, the exit pupil is considered as the eye-box of the system, while the entrance pupil does not matter on the occlusion precision since the R-PPM system works in the

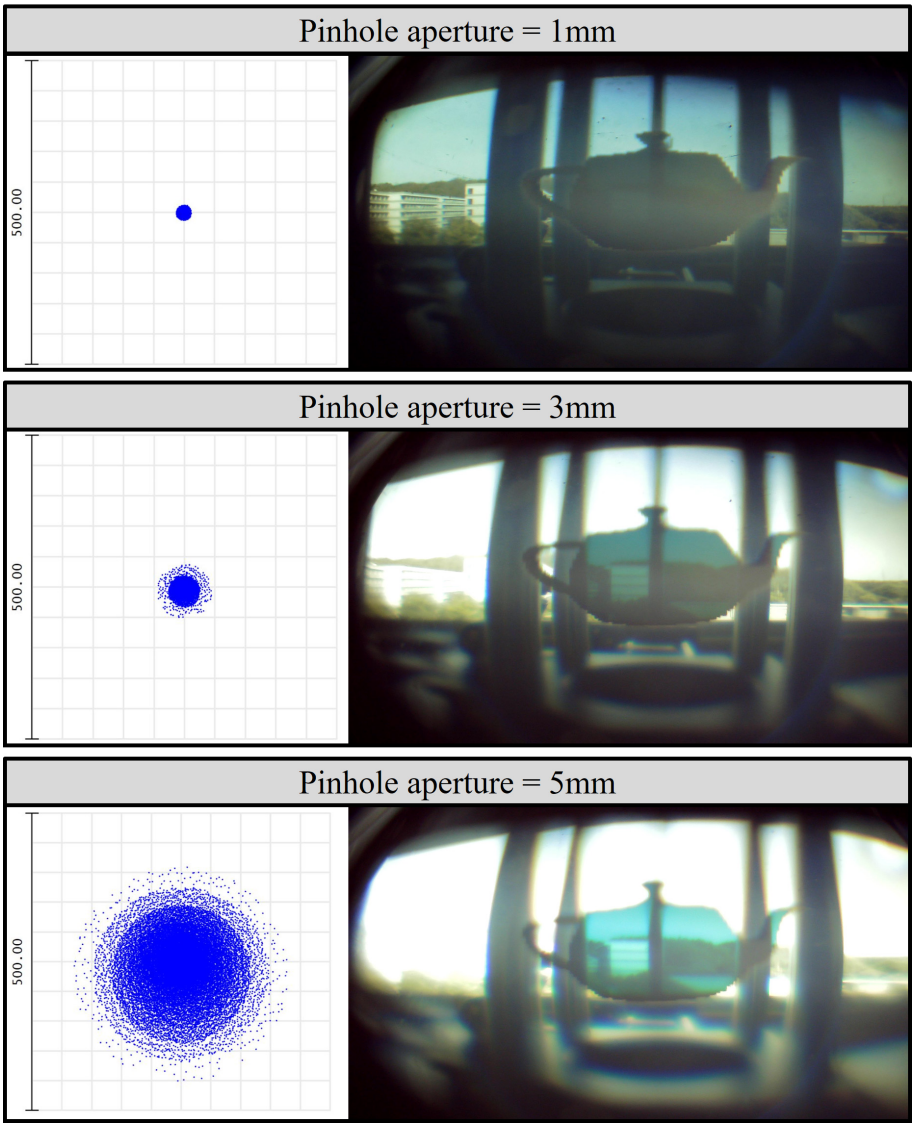

Figure 4.9: (Left) Footprints of a beam from the central FOV on the LCoS plane. (Right) Occlusion patterns rendered by prototypes with different pinhole apertures. The scalar beside is given as 500um.

hard-edge occlusion way. Hence, we only show footprints at exit pupils calculated by ZEMAX, and pinhole apertures are set as 1mm, 3mm, and 5mm. The simulation results are shown in Fig. 4.10. Light from H0°×V0°, H0°×V±25°, and H42.5°×V0°, which are the edges of the prototype FOV, is drawn with different colors. A scalar of 3mm is given beside, and circles in each sub-figure have the same diameter of 3mm. Since the light from different FOVs overlaps areas with various sizes, we take the footprint given by the FOV of H0°×V0° as the efficient eye-box. In general, the eye-box enlarges with the pinhole aperture increase, which from the minimum size of 0.4mm given by 1mm pinhole aperture, then the moderate size of 1.2mm given by 3mm pinhole aperture, to the maximum size of 1.8mm given by 5mm pinhole aperture. In addition, the see-through efficiencies also varies with different pinhole aperture settings. With the polarizing beam splitter used as the optical combiner and a transmittance of 73% of the LCoS, the see-through efficiencies are 0.6%, 5.1%, and 14.2% at different pinhole apertures of 1mm, 3mm, and 5mm (compare to naked eyes with a pupil of 2mm).

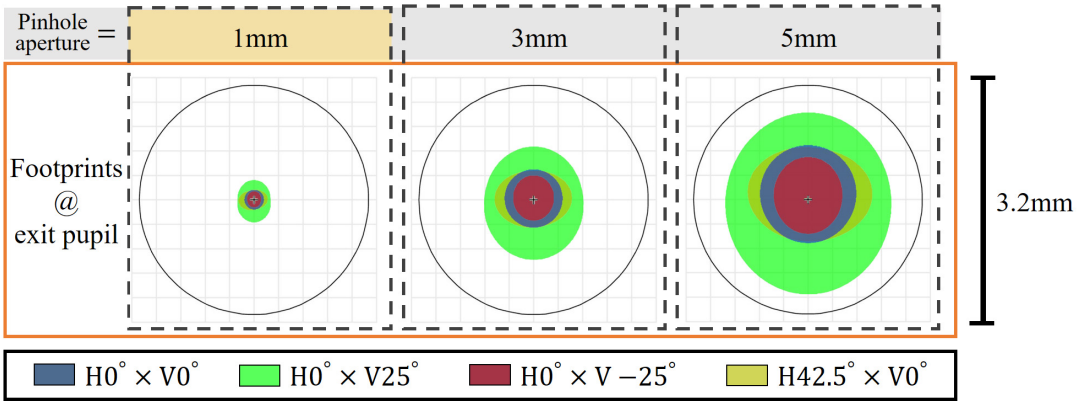

Figure 4.10: Exit pupils of the R-PPM prototype calculated with the pinhole aperture setting as 1mm, 3mm, and 5mm.

Divergent from the T-PEM system, the virtual image in the R-PPM system is projected by a projector with an individual pinhole mask. Therefore, the resolution and brightness of the generated virtual image can be adjusted separately. Fig. 4.11 shows a rectangle with words inside is displayed by the R-PPM prototype. The input image is given as Fig. 4.11 (a), a corresponding occlusion mask is also rendered to keep the virtual image visible. From Fig. 4.11 (b) to Fig. 4.11

Figure 4.11: **(a)** A rectangle with words inside is used for AR content. **(b) - (f)** Virtual images are displayed by the prototype with pinhole apertures from 1mm to 5mm.

(f), the pinhole aperture increases from 1mm to 5mm by a step of 1mm. As a result, the recorded virtual image becomes brighter but more blurred at the same time. In addition, obvious bloom also appears when the pinhole aperture exceeds 3mm. Consequently, we choose 3mm as the common pinhole aperture of the prototype for virtual display.

Fig. 4.12 shows hard-edge occlusion conducted by our prototype. A FOV of H83.5°×V53.1° is recorded by a webcam with a minimum illumination of 0.01lux. The redundant parts of paraboloidal mirrors and the additional mounts prevent a sufficient calibration, which leads to extra distortion in recorded images. The pinhole aperture for the real scene and the virtual display are set as 1mm and

Figure 4.12: Hard-edge occlusion is demonstrated by the prototype of the R-PPM system with the recorded FOV of H83.5°×V53.1°.

3mm, respectively. The real scene directly recorded by the camera at the entrance pupil is shown in Fig. 4.12 (A), and the see-through view recorded through the prototype is shown in Fig. 4.12 (B). The vertical parallax has been mostly eliminated. A red teapot with lighting effects is used as the target image as shown in Fig. 4.12 (C). The bright spot by specular lighting and the dark surface by diffuse lighting is bounded with blue and yellow dot-lines, respectively. When the virtual content is displayed without occlusion, which is shown in Fig. 4.12 (D), the image goes to be highly transparent, and the teapot also looks less realistic because effects from both diffuse lighting and specular lighting are missed. To solve the problem, a sharp occlusion pattern is given by the LCoS as shown in Fig. 4.12 (E). Again, the virtual image is projected with the same illumination by the LCD while keeping the background blocked by the occlusion pattern, and the recorded scene is shown in Fig. 4.12 (F). The teapot becomes not only more visible with the opaque surface but also more realistic with the bright spot and the dark bottom being perceived.

4.3 System limitations

In summary, the R-PPM system supports a wide-view hard-edge occlusion. By using the paired paraboloidal mirrors, NA of optical architectures in typical hard-edge occlusion OC-OST-AR displays is increased efficiently. Thus a large range of the real scene can be focused and pixelated occlusion can be conducted by locating an LCoS at an inner focal plane of the imaging system. Compared to the T-PEM system, the image quality of the transmitted real scene is improved and the see-through efficiency also increases significantly. The bench-top prototype has demonstrated a virtual display with the hard-edge occlusion in a FOV of H83.5°×V53.1°. However, some limitations still need to be addressed in future

The biggest challenge relates to the R-PPM system is the large form-factor. At this point, two paraboloidal mirrors, six lenses, a roof prism, a beam splitter, and two mirrors are used in the prototype, so that leads to a heavy system. Although the implementation of some optical elements, such as Fresnel lenses or metalenses, can alleviate the system weight, worse image quality or higher cost can be another issue. Meanwhile, the long optical path designed for erecting the

see-through view propagating through the paired paraboloidal mirrors increases the system volume. Some near-to-eye imaging systems achieve compact structures by using polarization optics to fold the optical path [4]. A similar design can be used on the current R-PPM structure. Nevertheless, real scene erect and sufficient see-through efficiency are considerable challenges in such a system.

Another problem is the variable contrast ratio of the occlusion pattern in the R-PPM system. The system is established in the hard-edge occlusion way so that it overcomes the unstable occlusion precision in the T-PEM system. However, the deteriorated contrast of the occlusion pattern through the vertical FOV and with the pinhole expansion brings a new challenge. The decrease of contrast caused by the defocus of masks projected by the LCoS may be addressed by tilting the reflective SLM, which is similar to the operation for enhancing the image resolution of the virtual display in the R-PPM system, but the performance of an LCoS strongly depends on the incident angle, the modulation properties are less influenced only when incident angle deviates within 10° [50]. In this case, DMD may be another choice to conduct the pixelated occlusion with defocus attenuation, and the optical layout of the R-PPM system should be modified correspondingly for an unencumbered see-through view.

5 Image rendering and occlusion processing

In the previous two chapters, optical architectures of T-PEM and R-PPM systems are proposed, and the light propagation through the two imaging systems is analyzed. Then, the projected image by the LCD for virtual display and the occlusion pattern by the LCoS for blocking the real scene can be both computed. With the following sections, processes for image rendering and occlusion generation for our prototypes are introduced. A rendering flow is designed for allowing both T-PEM and R-PPM systems to work in real-time. In order to compensate for the optical aberration of rendered contents, samplers containing the projection of pixel coordinates between target images and displayed images are used. The transformation from the original scene to the displayed scene is done with two-step mapping in the fragment shader. Occlusion patterns are generated based on whether the program works for the T-PEM prototype with the enhanced soft-edge occlusion or the R-PPM prototype with the hard-edge occlusion. Finally, the displayed virtual image and the occlusion pattern are output in the same render loop for synchronization.

5.1 Rendering flow

The rendering flow of both the T-PEM and R-PPM systems is shown in Fig. 5.1. OpenGL API is implemented to build the rendering framework. In addition to the same stages shown with grey arrows at the beginning, the following procedures for the T-PEM system and the R-PPM system are depicted with orange and blue arrows, respectively. The input data is composed of the mesh data for the virtual objects and the system attributes (e.g., display FOV, system category,

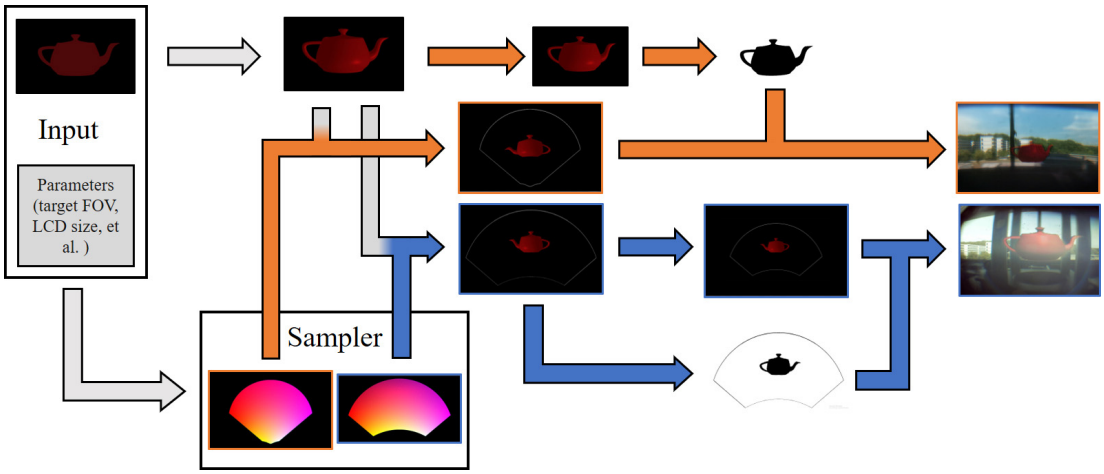

Figure 5.1: Rendering flow of the proposed T-PEM and R-PPM systems.

mirror parameters, panel dimension, et al.). The mesh data is firstly used to draw the 3D display scene. Specular and diffuse Lighting are added with the well-known Phong lighting model. Shadow is also rendered for the virtual object so that the realism is further enhanced. Notice that the casting shadow by the virtual object to the real scene is not rendered since we do not model the real scene during the rendering process. Meanwhile, system parameters are input as constant and utilized to calculate the image transformation from the target AR content to the digital pixels on the LCD (in the T-PEM system) or the LCoS (in the R-PPM system). Based on the operating system selected, a corresponding 2D sampler consisting of projections of coordinates from the displayed pixels to the target pixels is generated. The bottom rectangle in Fig. 5.1 shows two samplers calculated for T-PEM (left) and R-PPM (right) systems. The samplers are treated as RGB images, and only the red and green channels are utilized to store the projected pixel coordinates x and y. Besides, the two samplers are flipped left-right because displayed images on the LCDs are optically flipped by the imaging systems. Then, the generated sampler and the rendered virtual scene are input into a fragment shader as textures. A displayed pixel on the LCD is firstly mapped to a normalized coordinate in the designated FOV, then accesses the RGB data stored in the texture of the rendered virtual scene. Based on the two-step mapping, displayed images in prototypes can be computed in real-time

since mapping between displayed pixels and the target AR scene only depends on the system parameters and keeps stable during operation.

Algorithm 1 Fragment shader for rendering displayed images

Input *textureSampler*, *textureScene*, *viewportSize*[2], *tiltAngle*[2]
Output *color*[4]

1: $displayCoord = gl\_FragCoord.xy$ / $viewportSize * \cos(tiltAngle)$
2: $targetCoord$ = texture2D($textureSampler$, $displayCoord$)
3: **if** $targetCoord.st$ == vec2(0, 0) **then**
4: $color$ = vec4(0, 0, 0, 0)
5: **else**
6: $color$ = texture2D($textureScene, targetCoord.st$)

The two-step mapping for rendering the displayed virtual image is shown in Algorithm 1. In addition to the sampler of the projection from displayed pixels to the target scene and the texture of the original virtual scene, the tilt angle for minimizing the defocus of displayed virtual images in the T-PEM system is also input. Fig. 5.2 shows the projection from the displayed pixel on the tilted LCD to the target scene pixel. The LCD is rotated around the pixel referring to the FOV in H0°×V0° of the target scene, then *tiltAngle* in Algorithm 1 is given as the intersected angle between the original LCD plane and the tilted LCD plane. The blue pixel on the tilted LCD is firstly mapped to a red pixel on the original LCD by a cosine function. Notice that *tiltAngle* in the fragment shader is given in radian, an angle given in degree should be times $\pi/180$. Thus, the displayed pixel can be further mapped to the target scene pixel since the projection from a pixel on the original LCD to the target scene is included in the sampler. In addition, pixels in the sampler that correspond to the AR scene out of the designated FOV are filled in black, as shown in Fig. 5.1. Hence, the output variable *color* is given with vec4(0,0,0,0) when the mapped target scene pixel in Algorithm 1 has a coordinate of (0,0). Otherwise, the target scene pixel will be further mapped to a texture that contains a virtual scene rendered preliminarily. What is more, OpenGL expects the origin of a texture located at the left-bottom corner while the input textures in Algorithm. 1 have origins at the left-top corner, the two-step mapping automatically erects the output image since each fragment

are flipped upside-down twice during the processing. As a result, transformed images are generated as shown in Fig. 5.1, and grey boundaries in the images are the projection of the designated FOVs.

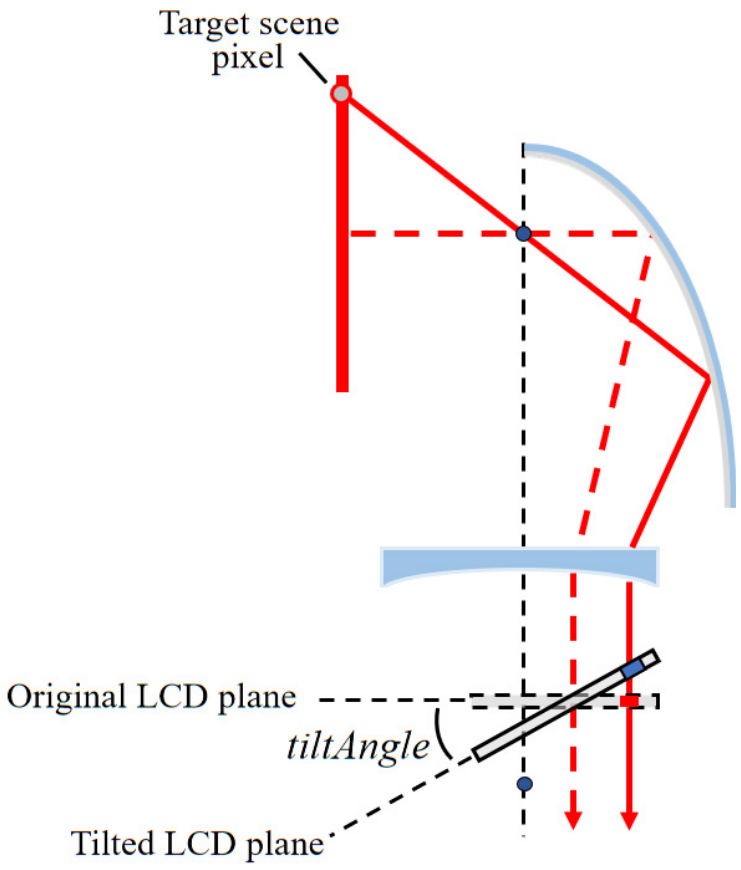

Figure 5.2: Projection between the displayed pixel on the tilted LCD to the target scene pixel in the T-PEM system.

As long as the displayed image is rendered, occlusion patterns on SLMs are also computed based on either T-PEM or R-PPM systems are operated. In the case of the T-PEM system, the occlusion pattern is generated from the original virtual image since the transmissive SLM is positioned before the entrance pupil of the imaging system, which requires the occlusion pattern to be shrunk according to the SLM position. In the case of the R-PPM system, the occlusion pattern is computed from the displayed virtual image due to the same optical position

of the projected virtual image plane and the LCoS plane. Moreover, displayed virtual images in the R-PPM system and occlusion patterns in the T-PEM system are resized since SLMs have different dimensions from LCDs in both prototypes. Finally, the virtual image and occlusion pattern are displayed sequentially in the same render loop, and the synthetic AR scene can be observed through the prototype.

5.2 Sampler for compensating optical aberration

In order to display distortion-free AR contents in the proposed systems, virtual images must be rendered with an algorithm to compensate for extra optical aberrations. Implementing a fragment shader that computes transformations through the imaging system for each generated pixel is a straightforward way to address this issue. However, the heavy computation may cause additional delays in the real-time rendering. Moreover, the optical aberration for the proposed imaging system is fixed after all components are mounted, complex real-time computation in the shader can be avoided as long as the projection from each target pixel to the displayed pixel is derived. Therefore, the compensation for optical aberration in the proposed systems is conducted by two-step mapping in the fragment shader. A sampler is preliminarily calculated before the render loop starts, then input into the fragment shader as 2D texture as shown in Algorithm. 1. Hence, the pixel coordinate transformation for each input fragment can be computed efficiently.

In this section, we show the sampler generation algorithm for rendering a distortion-free virtual image and a corresponding occlusion pattern in the R-PPM system. The system layout is derived in chapter 4, and the projection of a pixel through the imaging system can be calculated by solving the chief ray propagation. Let a pixel in the target AR scene transmits through each optical component onto the LCoS, which is shown as Fig. 5.3. Assuming the virtual image is displayed with a resolution of $RES_w \times RES_h$ and fills a rectangular FOV of $FOV_h \times FOV_v$, the spatial orientation of an arbitrary pixel P_{in} at ith row and jth column refers to a Cartesian coordinate system whose origin at the focus of

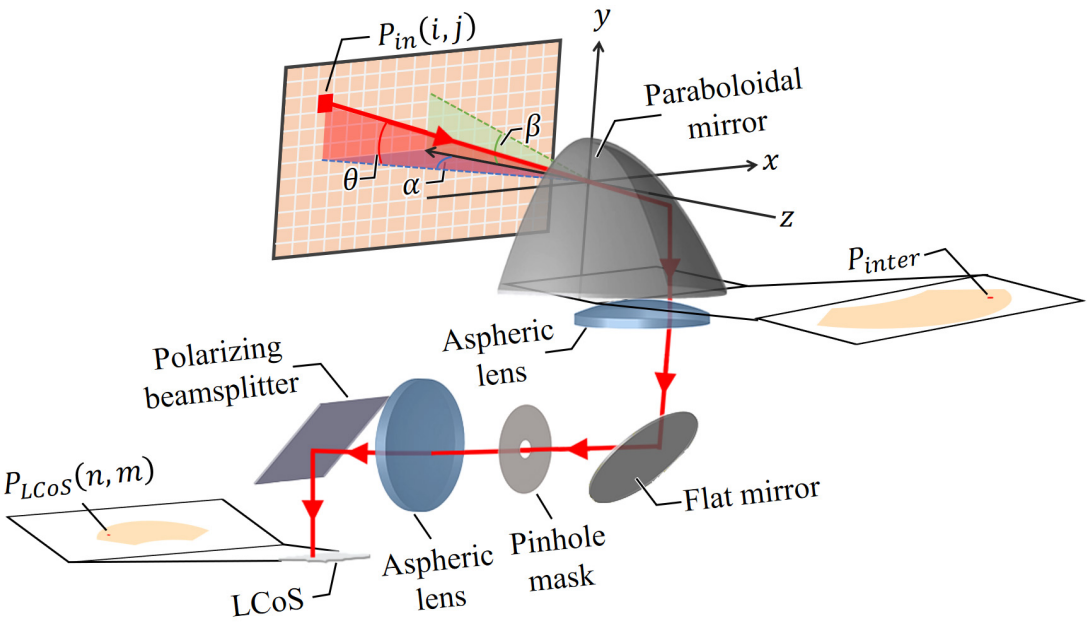

Figure 5.3: A pixel is projected from the target AR scene onto the LCoS in the R-PPM system.

the paraboloidal mirror can be given as:

$$\alpha = -\frac{FOV_h}{2} + i \times \frac{FOV_h}{(RES_w - 1)} \tag{5.1}$$

$$\beta = \frac{FOV_v}{2} - j \times \frac{FOV_v}{(RES_h - 1)} \tag{5.2}$$

Then, the incident angle of P_{in} can be derived:

$$\theta = \arctan\left(\cos\alpha \cdot \tan\beta\right) \tag{5.3}$$

Hence, the chief ray emitted from the input pixel is reflected by the paraboloidal mirror with a focal length of f into an intermediate image. Coordinates of the projected pixel P_{inter} are given as:

$$x_{inter} = 2f \cdot (-\tan\theta - \sqrt{1 + \tan^2\theta}) \cdot \sin\alpha \tag{5.4}$$

$$z_{inter} = 2f \cdot (-\tan\theta - \sqrt{1 + \tan^2\theta}) \cdot \cos\alpha \tag{5.5}$$

Since the two following lens pair functions a telecentric lens, the projection of the intermediate image by the lens pair can be solved with:

$$x_{out} = -M \cdot x_{inter} \tag{5.6}$$

$$z_{out} = -M \cdot (z_{inter} + L_{deflect}) \tag{5.7}$$

Where M is the magnification of the lens pair, the minus before M indicates that the intermediate image is flipped both up-down and left-right by the lens pair. And $L_{deflect}$ is the deflection of the optical axis of the lens pair from the optical axis of the paraboloidal mirror. According to the configuration of our prototype, $L_{deflect}$ is set as 12.7mm that lets the optical axis of the lens pair overlap the chief ray from the central pixel. Finally, the mapped pixel on the LCoS can be computed as follows:

$$\begin{aligned} P_{LCoS}(n,m) = \Bigg(& \left\lceil RES_h \cdot (\frac{1}{2} - \frac{z_{out}}{LCoS_h}) \right\rfloor, \\ & \left\lceil RES_w \cdot (\frac{1}{2} + \frac{x_{out}}{LCoS_w}) \right\rfloor \Bigg) \end{aligned} \tag{5.8}$$

Where $LCoS_w$ and $LCoS_h$ are the physical dimensions of the LCoS panel. Hereby, the mapping from a displayed pixel $P_{LCoS}(i,j)$ on the LCoS to the corresponding input pixel $P_{in}(i,j)$ in the target AR content is solved. Considering that the sampler is treated as a texture in the fragment shader by OpenGL, mapped coordinates should be normalized further.

Moreover, a displayed virtual image by the LCD in the R-PPM system propagates through a lens pair as shown in Fig. 4.2. The transmitted virtual image is flipped both horizontally and vertically by the two convex lenses, then upside-down again by the following beamsplitter. Hence, the sampler is expected to have a similar pixel coordinate transformation since the virtual image in the R-PPM system is directly rendered from the calculated sampler. Consequently, the output sampler is generated by modifying Eq. 5.8 slightly:

$$\begin{aligned} P_{out}(i,j) = \Bigg(& \left\lceil RES_h \cdot (\frac{1}{2} - \frac{z_{out}}{LCoS_h}) \right\rfloor, \\ & \left\lceil RES_w \cdot (\frac{1}{2} - \frac{x_{out}}{LCoS_w}) \right\rfloor \Bigg) \end{aligned} \tag{5.9}$$

And a corresponding occlusion pattern is generated by binary the rendered virtual image. Notice that the occlusion pattern needs to be flipped left-right to recover the pixel positions given by P_{LCoS}. What is more, the magnification of the projected virtual image by the following two convex lenses is compensated in the latter resize step in terms of the physical dimensions of the LCD and the LCoS in the R-PPM system.

With the projection given with Eq. 5.9, the mapping between a target pixel and the displayed pixel is generated. Nevertheless, a sampler directly generated in this way leaves a large number of pixels unfilled, which is shown in Fig. 5.4. Since the lower FOV is mapped to a larger area on the LCoS, blanks tend to occur more frequently in the upper part of the sampler. As a result, a virtual image is rendered with similar seams when it moves to those insufficiently filled areas. An inverse mapping from pixels on the LCoS to target pixels in the expected AR content is expected to generate a full-filled sampler, while it also brings massive redundant data since pixels that exceed the designated FOV can not be displayed by the prototype. Therefore, we chose to optimize the sampler by post-processing. The sampler is traversed right after generated, and an unfilled pixel is assigned with the value from nearby pixels. A sampler can be traversed several times until the optimized sampler is sufficiently filled. Our prototype operates the optimization twice, and the optimized sampler is shown in the lower rectangle in Fig. 5.4. The zoom-in image of the central upper area shows a higher fill factor. Although the optimized sampler still leaves some blanks, the virtual image shown below is adequately filled. Furthermore, since the sampler is generated preliminarily, the extra optimization procedure does not require additional computation during the render loop. Thereby, the R-PPM system can still render the AR content in real-time.

Similarly, the sampler in the T-PEM system can be computed by mapping the target AR scene spanning a designated FOV onto the LCD panel. Due to the larger pixel pitch of the LCD in the T-PPM system, the sampler is easily fully filled, as shown in Fig. 5.5. Hence, the sampler is output to the fragment shader without post-processing. And a rendered virtual image is shown at the right in Fig. 5.5.

Notice that we only developed the rendering flow for virtual display with oc-

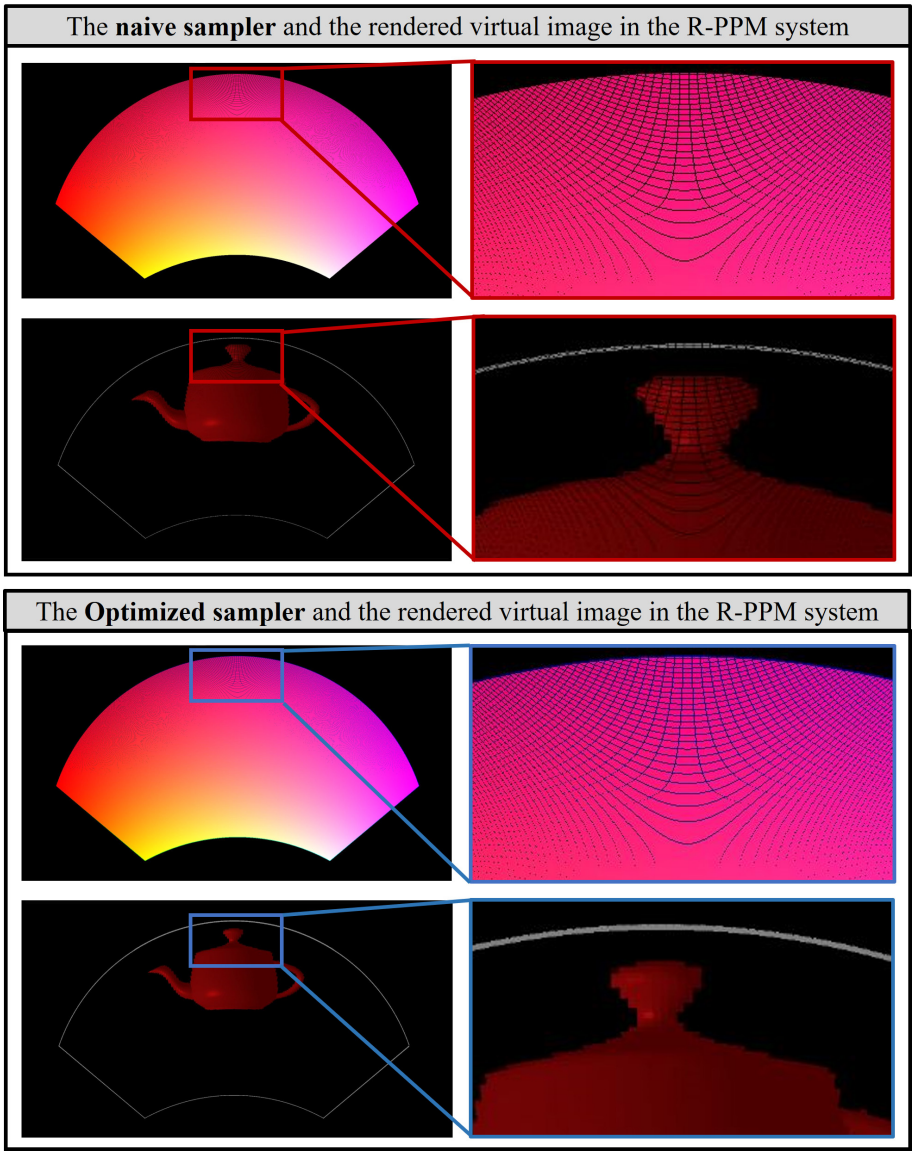

Figure 5.4: A virtual image rendered without (up) and with (below) sampler optimization in the R-PPM system.

clusion on the real scene. Thus the visibility of the AR content in a bright environment can be improved significantly. However, mutual occlusion between the virtual objects and the real objects has not been tested in our prototypes. In order to enable mutual occlusion in the prototypes, a depth sensor needs to be integrated into the systems. Thus the mesh data of the ambient environment

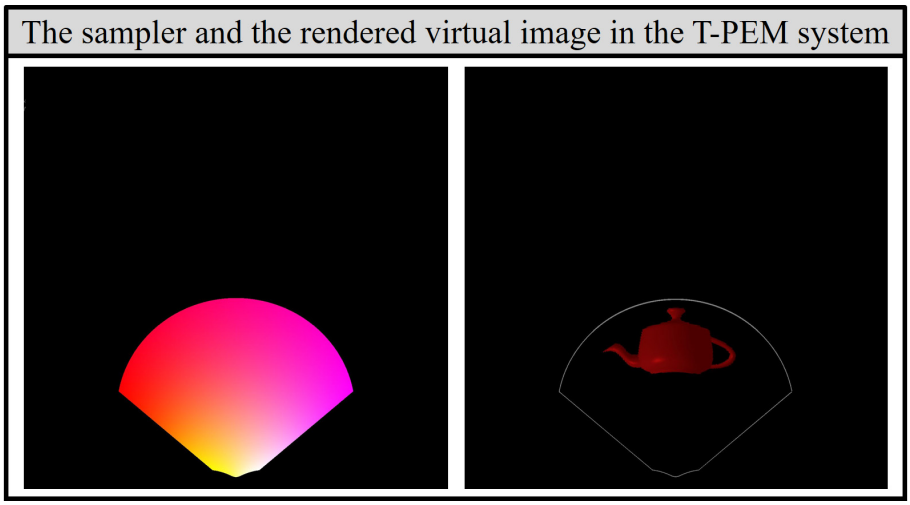

Figure 5.5: The sampler (left) and the rendered virtual image (right) in the T-PEM system.

can be captured, and the real scene can be built in the graphics space. Then, the original virtual scene input to the render flow is automatically occluded by real objects, and the following procedures conduct the occlusion from the virtual object to the real scene as the existing prototype shows. Finally, AR with high visibility and mutual occlusion can be realized based on the prototypes.

6 Conclusion

In this dissertation, two structures are proposed for OC-OST-AR displays. Both of them are expected to render mutual occlusion in wider FOVs than existing systems. Bench-top prototypes are built, the T-PEM system renders occlusion by a transmissive SLM in the enhanced soft-edge occlusion way and the R-PPM system renders occlusion by a LCoS in the hard-edge occlusion way. An algorithm with two-step mapping is proposed, real-time virtual display and occlusion are realized. The T-PEM system is demonstrated to render virtual images in a FOV of H160°×V74° and conduct occlusion in a FOV of H122°×V74°. And the R-PPM system is demonstrated to achieve a FOV of 83.5°×V53.1° for both occlusion and virtual display.

T-PEM system			
Pinhole aperture	1mm	3mm	5mm
Eye-box	0.19mm	0.56mm	0.96mm
See-through efficiency	0.1%	0.5%	1.2%
R-PPM system			
Pinhole aperture	1mm	3mm	5mm
Eye-box	0.4mm	1.2mm	1.8mm
See-through efficiency	0.6%	5.1%	14.2%

Table 6.1: Eye-boxes and see-through efficiencies of the T-PEM and R-PPM systems with different pinhole apertures.

Whereas wide-view occlusion has been demonstrated in both prototypes, the T-PEM and R-PPM systems show different system performances. Table. 6.1 lists the eye-boxes and the see-through efficiencies of the two systems with pinhole apertures of 1mm, 3mm, and 5mm. The R-PPM system shows noticeable advantages on both two attributes. Benefit from the usage of LCoS as the occlusion panel, a see-through efficiency of 14.2% can be obtained by the R-PPM system with a pinhole aperture of 5mm, which is close to the see-through efficiency of existing commercial OST-HMDs (e.g., $\sim 25\%$ for Hololens 1 and Hololens 2). In this regard, a wearable OC-OST-AR display is expected to be developed from the R-PPM system. Thus AR application scenarios can be expanded to not only indoor but also outdoor environments. In comparison, the T-PEM system provides the maximum see-through efficiency of 1.2% in the configuration of a 5mm pinhole aperture due to the low transmittance of transmissive SLMs. Hence, it may be difficult for the T-PEM system to operate indoor AR applications. Application scenarios in the outdoor environment, such as sight-view and construction, are potential directions for building stationary OC-OST-AR displays from the T-PEM system. Regarding the eye-box, though the R-PPM system has a larger dimension than the T-PEM system, it is far from the value, which is 1mm in general, to support a comfortable user experience for head-mounted displays. Hence, more efforts need to be made to expand the eye-box in current prototypes in the future.

Based on different mechanisms, the T-PEM and R-PPM systems present discriminate occlusion performances. The comparison of occlusion rendered by prototypes with different pinhole apertures and at different FOVs in the T-PEM and R-PPM systems is shown in Fig. 6.1. In general, occlusion rendered by the R-PPM system in the hard-edge occlusion way tends to be more stable with the variations of pinhole aperture and FOV. With the pinhole aperture changing from 1mm to 5mm, the T-PEM system renders the occlusion pattern more blurred, resulting from the enlargement of the occlusion-capable pixel with the pinhole aperture increase. Differently, the occlusion pattern rendered by the R-PPM system looks sharp regardless of different pinhole apertures since the pattern projected by the LCoS transmits the imaging system without passing through the pinhole. Although focused light spots on the LCoS are blocked insufficiently

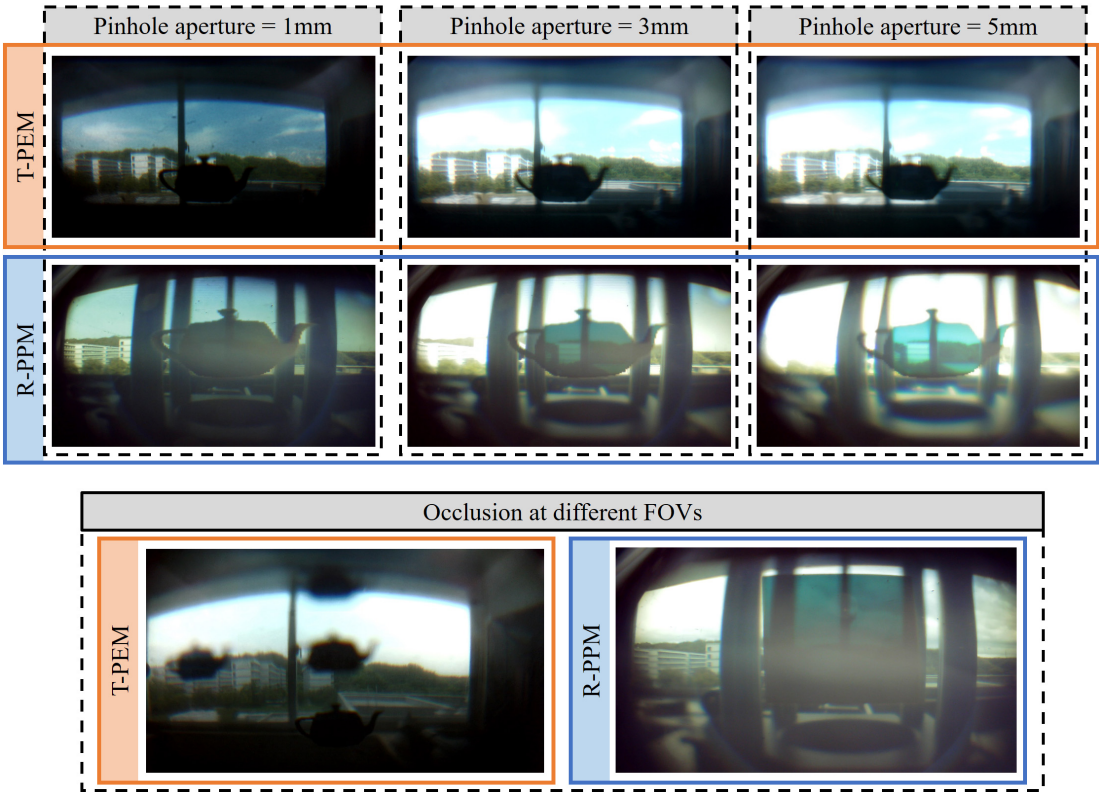

Figure 6.1: Occlusion is rendered in the T-PEM system and the R-PPM system. Different pinhole apertures (up) and FOVs (below) are tested.

when the R-PPM system works with a large pinhole aperture, which causes the decreased contrast ratio of occlusion patterns in Fig. 6.1, the system performance is less influenced by the deterioration. The R-PPM system is expected to be configured with a small pinhole aperture under excessive illumination and adjusted to a larger pinhole aperture when the ambient light is darkened. Thus, the occlusion pattern keeps high contrast in the bright environment where mutual occlusion is necessary for making AR contents visible, and changes to be low contrast in the dark environment where mutual occlusion is mainly used for enhancing users' perception but impacts the visibility of AR contents less.

Besides, the occlusion precision varies along the vertical FOV in the T-PEM system. As discussed in Section 3.2.2, beams emitted from the lower vision are shrunk into more slim than those from the upper vision by the T-PEM system. Consequently, the lower half FOV provides a sharper occlusion pattern in the

T-PEM system, as shown in Fig. 6.1, and the occlusion precision keeps roughly stable in the horizontal direction. As a comparison, the occlusion performance in the R-PPM system is less FOV dependent, which is shown in Fig. 6.1. The occlusion precision in the R-PPM system keeps stably high with the vertical shift of the FOV, while the contrast ratio gradually decreases when the occlusion pattern is moved upward.

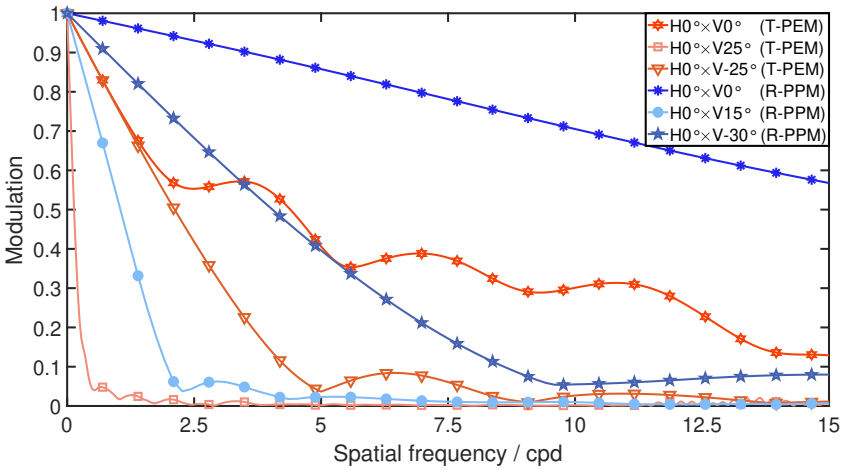

Figure 6.2: Average MTF curves for the T-PEM and R-PPM systems.

Regarding the image quality of the see-through view, the R-PPM system also shows a dominant advantage over the T-PEM. Fig. 6.2 depicts average MTF curves (by ZEMAX) in a similar vertical FOV range of prototypes built with the two systems. In general, the two systems both achieve the highest image quality in the central FOV. Due to the non-symmetric surface in the vertical direction of conical mirrors, the image quality becomes worse with the FOV shifting upward or downward, while MTF curves keep roughly stable in the horizontal plane in both systems as discussed in Section. 3.2.2 and Section. 4.2. Considering a spatial resolution of 40ppd is required for near-eye displays, the R-PPM system is expected to support a decent AR experience with MTF30 achieved at spatial frequencies of 30.4cpd @ H0°×V0° and 6.1cpd @ H0°×V-30°. Consequently, the existing prototype for the R-PPM system is sufficient to transmit the real scene in the lower FOV. Although the T-PEM system shows worse image quality, fewer optical components are utilized in the structure, which results in a much lower

cost than the R-PPM system.

Using large dimension lenses would improve the image quality in the R-PPM system, while the form-factor is increased correspondingly. Based on the consideration that the visual acuity of human beings also deteriorates with the vision expanding, an optimized design with better image quality is expected to be developed from the R-PPM structure. Thus, a wide-view OC-OST-AR display allowing outdoor application and supporting strong perceptual cues could be released. The major obstacle for developing a wearable version from the existing R-PPM system is the large form-factor. The usage of the paired conical mirrors helps the imaging system to outgrow the NA limitation so that realizes occlusion in a wide FOV, but the extra volume introduced by the conical mirrors inevitably increases the form-factor. The system volume of the existing prototype reaches the range of mixed reality (MR) displays rather than common OST-AR displays. A more compact design should be proposed to alleviate the influence by using conical reflectors.

What is more, hyperboloidal mirrors may also be used to build a wide-view OC-OST-AR display by following the PCR structure. Hyperboloidal mirrors are widely used in panoramic cameras [11, 91], and a few OST-AR displays have proved wide FOVs by integrating hyperboloidal reflectors into the imaging systems [36, 18, 60]. Therefore, a paired-hyperboloidal-mirror system could be similarly developed as the T-PEM and the R-PPM systems. However, different imaging properties of the hyperboloidal mirror from the ellipsoidal mirror and the paraboloidal mirror bring some challenges in designing such an imaging system. Two potential optical architectures for the paired-hyperboloidal-mirror system are shown in Fig. 6.3, where chief rays of the incident beams are expected to transmit systems free of aberration. Using the outer surface of a hyperboloidal reflector as a reflective mirror is common in many hyperboloidal mirror-based systems. The ray bundle reflected by the hyperboloidal mirror is convergent so that a similar architecture as the T-PEM system can be built to process the imaging light, as shown in Fig. 6.3 (a). The two paired hyperboloidal mirrors share the same focus F_2, then chief rays toward the first focus F_1 are recovered at another focus F_3. However, those rays do not pass through F_1 physically in this optical architecture, thus recovered rays do not form a physical eyebox at F_3. As near-

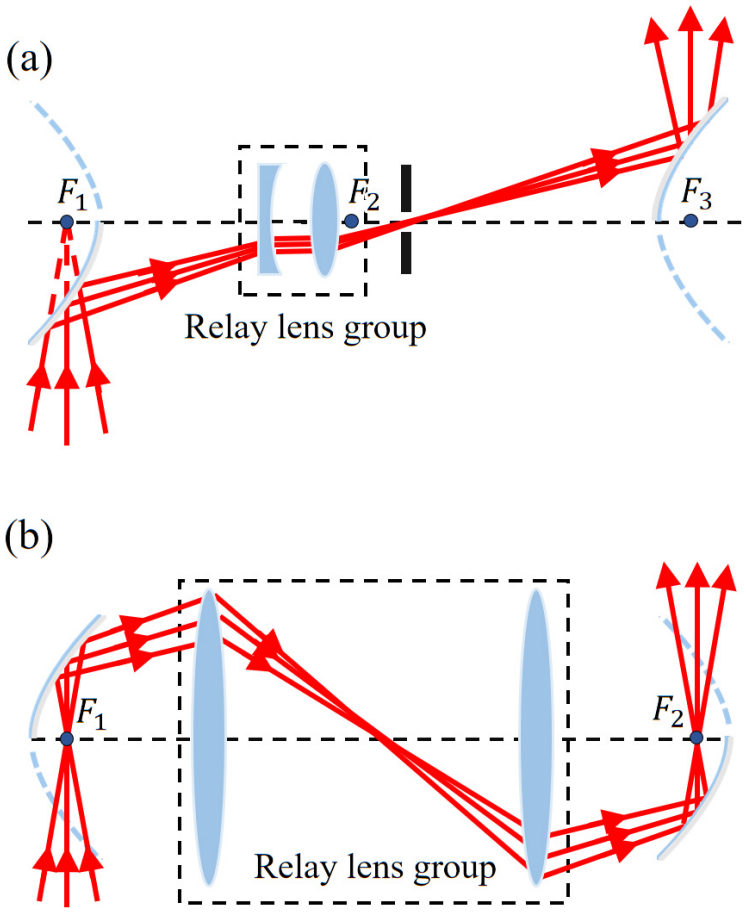

Figure 6.3: Potential optical architectures for a paired-hyperboloidal-mirror system. **(a)** The outer surface and **(b)** the inner surface of hyperboloidal mirrors are used to reflect incident rays, respectively.

to-eye displays, such a physical eyebox is essential for users to locate their eyes to receive imaging beams. Hence, a modified design should be proposed to provide a physical eye-box for practical applications. As an alternative, the inner surface of a hyperboloidal reflector can be also used to reflect incident rays, as shown in Fig. 6.3 (b). Chief rays enter the imaging system through the focus F_1 in this architecture, thus a physical eyebox will be generated at the focus F_2. However, the reflected chief rays are divergent in this case, forcing the following optical elements to be given with an extremely large dimension. Therefore, building a wide-view OC-OST-AR display with a paired-hyperboloidal-mirror structure

requires more ingenious designs.

Nowadays, the community focuses on AR technologies for indoor usage due to the poor image quality of OST-AR displays in a bright environment. The application of AR is thereby significantly restricted since a large number of human activities are conducted outdoor. The OC-OST-AR display is the only technology that has been demonstrated to address this issue. Nevertheless, existing OC-OST-AR displays suffer from the nature of imaging lenses, hence do not support a wide FOV. The PCR structure proposed in this dissertation is an attempt to overcome this limitation and, to the best of our knowledge, is the only approach that realizes occlusion in a wide FOV so far. Though a better image quality and a smaller form-factor should be achieved with future research, an OC-OST-AR display that supports most daily application scenarios could eventually be developed from the PCR structure. Then, AR technologies can be implemented to augment the reality of human society.

Acknowledgements

In the past three years, I enjoyed the journey to pursue a doctoral degree at Nara Institute of Science and Technology. At the point of boarding the airplane to Japan, I wished to have a meaningful time and do some research that can enjoy myself in the doctoral program. And at this point this time, I would say that wish has already become the wonderful life in the time at NAIST. At the end of the doctoral dissertation, I would like to thank all people who gave me favors in this period.

It is my parents that always stand by me. Their selfless supports relieved and encouraged me in those harsh times during the pursuit of the doctoral degree. Without the full understanding of my parents, it would be impossible for me to start this wonderful journey three years ago, and I would also suffer many unpredictable difficulties in the life of studying abroad. It is my fortune to be the son of them. Thank you so much, father and mother.

It is my honor to be a student of Cybernetics and Reality Engineering (CARE) Laboratory. My supervisor Prof. Kiyoshi Kiyokawa is the one that guides me to open the gate of the academy. Although he told me that he would like to treat me as a collaborator rather than a student in our first meeting, Kiyokawa sensei did teach me a lot in the past three years. I learned from him not only the knowledge and skills to be a qualified researcher but also the way to be an excellent person in life. With the influence of Kiyokawa sensei, I began to keep serious and rigorous for academic activities, learned to be strict with myself and be lenient to others, and tried to overcome laziness. The experience working with him will be a treasure for the rest of my life and encourage me to keep outstanding.

I would also thank my co-supervisors at NAIST sincerely. Prof. Yasuhiro Mukaigawa kindly helped me when I applied for the scholarship and made con-

structive comments on my research project with his professional knowledge. In a large number of lab meetings, Assoc. Prof. Hideaki Uchiyama and Asst. Prof. Naoya Isoyama never skimped to share their experience and insight on research with me, which benefited me in developing my own research taste. Also, thanks to Assoc. Prof. Nobuchika Sakata, who has transferred to Ryukoku University. He helped me not only in the main part of research but also in many trivial but essential works about conducting experiments when I established the first prototype for my research project. And Prof. Hong Hua at University of Arizona, who collaborated with me in my research project. I would deeply thank her for giving me the instruction on the optical design of AR/VR displays, sharing a lot of experience on imaging system building and analysis, and proposing many precious comments for paper writing. All of them have made non-trivial attributions on the final accomplishment of the doctoral dissertation.

And thanks to all members of CARE Lab. When I first arrived in Japan, it was them to help me be familiar with the university. Their enthusiasm helped me have an easy lifestyle transition and smoothly launch the first step of my research. Our excellent secretaries Mina Nakamura and Yoshie Sato, helped me deal with numerous procedures in campus life and research activities. Their patient work is the basis for me to concentrate on research and continue making more achievements. Some of them have left NAIST and started their new career. Here, I would like to show all of them my sincere thanks and blessings.

I am glad to have so many friends at NAIST. Life for a doctoral student is monotonous most of the time, and the COVID brought extra worries to my daily life. Nevertheless, I could relax in those times spent with my friends. We enjoyed the beautiful landscapes of Japan together and experienced a warm and exciting campus life. I sincerely wish them a happy life after graduating from NAIST, whenever and wherever they are. And also thanks to my girlfriend. Although we are in different countries and a reunion even becomes extremely difficult after the unpredictable COVID occurred. We always tried to understand each other and keep optimistic. She encouraged me when I fell into negative thoughts and derived loneliness from me in those stressful days. We have confronted one challenge after another together during the past three years.

At last, I also want to thank all faculty members of NAIST. Although I only

have contracts with a few of them, their hard work is essential to have a wonderful campus life. Furthermore, I was always treated with kindness and patience by NAIST staff. As a foreign student, they are engaged in providing me with a comfortable, enjoyable, and distinguishable study abroad experience. And they always respond to me in time whenever I need help. At the end of the 3-year study, I would say that coming here and studying for a doctoral degree is one of the best choices I have ever made.

Again, my deepest thank to all of you!

Bibliography

[1] Kaan Akşit et al. “Near-Eye Varifocal Augmented Reality Display Using See-through Screens”. In: *ACM Trans. Graph.* 36.6 (Nov. 2017). ISSN: 0730-0301.

[2] Pedro Arguijo and Marija Strojnik Scholl. “Exact ray-trace beam for an off-axis paraboloid surface”. In: *Applied optics* 42.16 (2003), pp. 3284–3289.

[3] Giovanni Avveduto, Franco Tecchia, and Henry Fuchs. “Real-world occlusion in optical see-through AR displays”. In: *Proceedings of the 23rd ACM Symposium on Virtual Reality Software and Technology*. 2017, pp. 1–10.

[4] Kiseung Bang et al. “Lenslet VR: Thin, Flat and Wide-FOV Virtual Reality Display Using Fresnel Lens and Lenslet Array”. In: *IEEE Transactions on Visualization and Computer Graphics* 27.5 (2021), pp. 2545–2554.

[5] Woodrow Barfield et al. “Comparison of human sensory capabilities with technical specifications of virtual environment equipment”. In: *Presence: Teleoperators & Virtual Environments* 4.4 (1995), pp. 329–356.

[6] Mark Billinghurst and Andreas Duenser. “Augmented reality in the classroom”. In: *Computer* 45.7 (2012), pp. 56–63.

[7] Ozan Cakmakci, Yonggang Ha, and Jannick 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*. IEEE. 2004, pp. 16–25.

[8] Praneeth Chakravarthula et al. “Focusar: Auto-focus augmented reality eyeglasses for both real world and virtual imagery”. In: *IEEE transactions on visualization and computer graphics* 24.11 (2018), pp. 2906–2916.

[9] Seunghyuk Chang. *Geometrical theory of aberrations for classical offset reflector antennas and telescopes*. University of Southern California, 2003.

[10] N Chelwani et al. “Off-axis parabolic mirror optics for polarized Raman spectroscopy at low temperature”. In: *Applied Physics Letters* 110.19 (2017), p. 193504.

[11] Chao Ping Chen et al. “Design of retinal projection displays enabling vision correction”. In: *Optics Express* 25.23 (2017), pp. 28223–28235.

[12] Xu Chen et al. “A Structured-Light-Based Panoramic Depth Camera”. In: *2018 IEEE International Conference on Real-time Computing and Robotics (RCAR)*. IEEE. 2018, pp. 102–107.

[13] Ping-Yen Chou et al. “Hybrid light field head-mounted display using time-multiplexed liquid crystal lens array for resolution enhancement”. In: *Optics express* 27.2 (2019), pp. 1164–1177.

[14] WS Colburn and BJ Chang. *Holographic Combiners for Head-Up Displays*. Tech. rep. ENVIRONMENTAL RESEARCH INST OF MICHIGAN ANN ARBOR RADAR DIV, 1977.

[15] Catherine Diaz et al. “Designing for depth perceptions in augmented reality”. In: *2017 IEEE international symposium on mixed and augmented reality (ISMAR)*. IEEE. 2017, pp. 111–122.

[16] A Drechsler et al. “Confocal microscopy with a high numerical aperture parabolic mirror”. In: *Optics express* 9.12 (2001), pp. 637–644.

[17] D. Dunn et al. “Wide Field Of View Varifocal Near-Eye Display Using See-Through Deformable Membrane Mirrors”. In: *IEEE Transactions on Visualization and Computer Graphics* 23.4 (2017), pp. 1322–1331.

[18] David Dunn et al. “Wide field of view varifocal near-eye display using see-through deformable membrane mirrors”. In: *IEEE transactions on visualization and computer graphics* 23.4 (2017), pp. 1322–1331.

[19] Austin Erickson et al. “Exploring the limitations of environment lighting on optical see-through head-mounted displays”. In: *Symposium on Spatial User Interaction*. 2020, pp. 1–8.

[20] Chunyu Gao, Yuxiang Lin, and Hong Hua. "Occlusion capable optical see-through head-mounted display using freeform optics". In: *2012 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*. IEEE. 2012, pp. 281–282.

[21] Gunther Haas. "40-2: Invited Paper: Microdisplays for Augmented and Virtual Reality". In: *SID Symposium Digest of Technical Papers*. Vol. 49. 1. Wiley Online Library. 2018, pp. 506–509.

[22] Takumi Hamasaki and Yuta Itoh. "Varifocal occlusion for optical see-through head-mounted displays using a slide occlusion mask". In: *IEEE transactions on visualization and computer graphics* 25.5 (2019), pp. 1961–1969.

[23] Steven J Henderson and Steven Feiner. "Evaluating the benefits of augmented reality for task localization in maintenance of an armored personnel carrier turret". In: *2009 8th IEEE International Symposium on Mixed and Augmented Reality*. IEEE. 2009, pp. 135–144.

[24] Hekun Huang and Hong Hua. "Generalized methods and strategies for modeling and optimizing the optics of 3D head-mounted light field displays". In: *Optics express* 27.18 (2019), pp. 25154–25171.

[25] Hekun Huang and Hong Hua. "High-performance integral-imaging-based light field augmented reality display using freeform optics". In: *Optics express* 26.13 (2018), pp. 17578–17590.

[26] Duy-Nguyen Ta Huynh et al. "Art of defense: a collaborative handheld augmented reality board game". In: *Proceedings of the 2009 ACM SIGGRAPH symposium on video games*. 2009, pp. 135–142.

[27] Yuta Itoh, Takumi Hamasaki, and Maki Sugimoto. "Occlusion leak compensation for optical see-through displays using a single-layer transmissive spatial light modulator". In: *IEEE transactions on visualization and computer graphics* 23.11 (2017), pp. 2463–2473.

[28] Changwon Jang et al. "Retinal 3D: augmented reality near-eye display via pupil-tracked light field projection on retina". In: *ACM Transactions on Graphics (TOG)* 36.6 (2017), pp. 1–13.

[29] Jason Jerald. *The VR book: Human-centered design for virtual reality*. Morgan & Claypool, 2015.

[30] Yeon-Gyeong Ju et al. "Occlusion-capable optical-see-through near-eye display using a single digital micromirror device". In: *Optics letters* 45.13 (2020), pp. 3361–3364.

[31] Masayuki Kanbara et al. "A stereoscopic video see-through augmented reality system based on real-time vision-based registration". In: *Proceedings IEEE Virtual Reality 2000 (Cat. No. 00CB37048)*. IEEE. 2000, pp. 255–262.

[32] DH Kelly. "Visual contrast sensitivity". In: *Optica Acta: International Journal of Optics* 24.2 (1977), pp. 107–129.

[33] Minkwan Kim et al. "Compact tomographic near-eye display using a MEMS scanning mirror". In: *Optics Letters* 46.17 (2021), pp. 4176–4179.

[34] Naohiro Kishishita et al. "Analysing the effects of a wide field of view augmented reality display on search performance in divided attention tasks". In: *2014 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*. IEEE. 2014, pp. 177–186.

[35] K. Kiyokawa. "A Wide Field-of-view Head Mounted Projective Display using Hyperbolic Half-silvered Mirrors". In: *2007 6th IEEE and ACM International Symposium on Mixed and Augmented Reality*. Nov. 2007, pp. 207–210.

[36] Kiyoshi Kiyokawa. "A wide field-of-view head mounted projective display using hyperbolic half-silvered mirrors". In: *2007 6th IEEE and ACM International Symposium on Mixed and Augmented Reality*. IEEE. 2007, pp. 207–210.

[37] Kiyoshi Kiyokawa, Yoshinori Kurata, and Hiroyuki 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)*. IEEE. 2000, pp. 60–67.

[38] Kiyoshi Kiyokawa et al. “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.

[39] Brooke Krajancich, Nitish Padmanaban, and Gordon Wetzstein. “Factored occlusion: Single spatial light modulator occlusion-capable optical see-through augmented reality display”. In: *IEEE transactions on visualization and computer graphics* 26.5 (2020), pp. 1871–1879.

[40] Bernard C Kress. “Optical architectures for augmented-, virtual-, and mixed-reality headsets”. In: (2020).

[41] Ernst Kruijff et al. “The influence of label design on search performance and noticeability in wide field of view augmented reality displays”. In: *IEEE transactions on visualization and computer graphics* 25.9 (2018), pp. 2821–2837.

[42] Shoufeng Lan et al. “Metasurfaces for near-eye augmented reality”. In: *ACS Photonics* 6.4 (2019), pp. 864–870.

[43] Gun-Yeal Lee et al. “Metasurface eyepiece for augmented reality”. In: *Nature communications* 9.1 (2018), pp. 1–10.

[44] Hwihyeong Lee, Byong Hyuk Chon, and Hee Kyung Ahn. “Reflective Fourier ptychographic microscopy using a parabolic mirror”. In: *Optics Express* 27.23 (2019), pp. 34382–34391.

[45] Jin Su Lee, Yoo Kwang Kim, and Yong Hyub Won. “Time multiplexing technique of holographic view and Maxwellian view using a liquid lens in the optical see-through head mounted display”. In: *Optics Express* 26.2 (2018), pp. 2149–2159.

[46] M Andreas Lieb and Alfred J Meixner. “A high numerical aperture parabolic mirror as imaging device for confocal microscopy”. In: *Optics express* 8.7 (2001), pp. 458–474.

[47] Jian Liu et al. “Focusing of an elliptical mirror based system with aberrations”. In: *Journal of Optics* 15.10 (Sept. 2013), p. 105709.

[48] Jian Liu et al. “Focusing properties of elliptical mirror with an aperture angle greater than ”. In: *Optical Engineering* 53.6 (2013), pp. 1–7.

[49] Zeyang Liu et al. “Design of a uniform-illumination binocular waveguide display with diffraction gratings and freeform optics”. In: *Optics express* 25.24 (2017), pp. 30720–30731.

[50] A Lizana et al. “Influence of the incident angle in the performance of liquid crystal on silicon displays”. In: *Optics Express* 17.10 (2009), pp. 8491–8505.

[51] Andrew Maimone and Henry Fuchs. “Computational augmented reality eyeglasses”. In: *2013 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*. IEEE. 2013, pp. 29–38.

[52] Andrew Maimone, Andreas Georgiou, and Joel S Kollin. “Holographic near-eye displays for virtual and augmented reality”. In: *ACM Transactions on Graphics (Tog)* 36.4 (2017), pp. 1–16.

[53] Andrew Maimone et al. “Focus 3D: Compressive accommodation display.” In: *ACM Trans. Graph.* 32.5 (2013), pp. 153–1.

[54] Andrew Maimone et al. “Pinlight displays: wide field of view augmented reality eyeglasses using defocused point light sources”. In: *ACM SIGGRAPH 2014 Emerging Technologies*. 2014, pp. 1–1.

[55] Daniel Malacara-Hernández, Zacarıas Malacara-Hernández, and Zacarias Malacara. *Handbook of optical design*. CRC Press, 2003.

[56] G. A. Massey and A. E. Siegman. “Reflection and Refraction of Gaussian Light Beams at Tilted Ellipsoidal Surfaces”. In: *Appl. Opt.* 8.5 (May 1969), pp. 975–978.

[57] Shohei Mori et al. “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)*. IEEE. 2018, pp. 251–258.

[58] Jurriaan D Mulder. “Realistic occlusion effects in mirror-based co-located augmented reality systems”. In: *IEEE Proceedings. VR 2005. Virtual Reality, 2005.* IEEE. 2005, pp. 203–208.

[59] Alessandro Mulloni, Hartmut Seichter, and Dieter Schmalstieg. “Handheld augmented reality indoor navigation with activity-based instructions”. In: *Proceedings of the 13th international conference on human computer interaction with mobile devices and services*. 2011, pp. 211–220.

[60] H. Nagahara, Y. Yagi, and M. Yachida. “Wide field of view catadioptrical head-mounted display”. In: *Proceedings 2003 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2003) (Cat. No.03CH37453)*. Vol. 4. Oct. 2003, 3738–3743 vol.3.

[61] Cheng Pan et al. “Design of a high-performance in-coupling grating using differential evolution algorithm for waveguide display”. In: *Optics express* 26.20 (2018), pp. 26646–26662.

[62] Jiamin Ping et al. “Effects of shading model and opacity on depth perception in optical see-through augmented reality”. In: *Journal of the Society for Information Display* 28.11 (2020), pp. 892–904.

[63] Jarkko Polvi et al. “Handheld guides in inspection tasks: Augmented reality versus picture”. In: *IEEE transactions on visualization and computer graphics* 24.7 (2017), pp. 2118–2128.

[64] Charles Lane Poor. “The aberration of parabolic mirrors”. In: *The Astronomical Journal* 18 (1897), pp. 89–91.

[65] Kishore Rathinavel, Gordon Wetzstein, and Henry Fuchs. “Varifocal occlusion-capable optical see-through augmented reality display based on focus-tunable optics”. In: *IEEE transactions on visualization and computer graphics* 25.11 (2019), pp. 3125–3134.

[66] Kishore Rathinavel et al. “An extended depth-at-field volumetric near-eye augmented reality display”. In: *IEEE transactions on visualization and computer graphics* 24.11 (2018), pp. 2857–2866.

[67] Donghao Ren et al. “Evaluating wide-field-of-view augmented reality with mixed reality simulation”. In: *2016 IEEE Virtual Reality (VR)*. IEEE. 2016, pp. 93–102.

[68] Yusuke Sando et al. “Super-wide viewing-zone holographic 3D display using a convex parabolic mirror”. In: *Scientific reports* 8.1 (2018), pp. 1–8.

[69] C Stanciu, M Sackrow, and AJ Meixner. “High NA particle-and tip-enhanced nanoscale Raman spectroscopy with a parabolic-mirror microscope”. In: *Journal of microscopy* 229.2 (2008), pp. 247–253.

[70] J Edward Swan, Gurjot Singh, and Stephen R Ellis. “Matching and reaching depth judgments with real and augmented reality targets”. In: *IEEE transactions on visualization and computer graphics* 21.11 (2015), pp. 1289–1298.

[71] Javed Rouf Talukder, Yun-Han Lee, and Shin-Tson Wu. “Photo-responsive dye-doped liquid crystals for smart windows”. In: *Optics express* 27.4 (2019), pp. 4480–4487.

[72] Javed Rouf Talukder, Hung-Yuan Lin, and Shin-Tson Wu. “Photo-and electrical-responsive liquid crystal smart dimmer for augmented reality displays”. In: *Optics express* 27.13 (2019), pp. 18169–18179.

[73] Eric W Tatham. “Technical opinion: getting the best of both real and virtual worlds”. In: *Communications of the ACM* 42.9 (1999), pp. 96–98.

[74] Christina Trepkowski et al. “The effect of narrow field of view and information density on visual search performance in augmented reality”. In: *2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*. IEEE. 2019, pp. 575–584.

[75] Peter Varga and Peter Török. “Focusing of electromagnetic waves by paraboloid mirrors. I. Theory”. In: *JOSA A* 17.11 (2000), pp. 2081–2089.

[76] Peter Varga and Peter Török. “Focusing of electromagnetic waves by paraboloid mirrors. II. Numerical results”. In: *JOSA A* 17.11 (2000), pp. 2090–2095.

[77] David Vettese. “Liquid crystal on silicon”. In: *Nature Photonics* 4.11 (2010), pp. 752–754.

[78] Xuan Wang et al. “Digitally switchable multi-focal lens using freeform optics”. In: *Optics express* 26.8 (2018), pp. 11007–11017.

[79] Lidong Wei et al. “Design and fabrication of a compact off-axis see-through head-mounted display using a freeform surface”. In: *Optics express* 26.7 (2018), pp. 8550–8565.

[80] Lidong Wei et al. “Design and fabrication of a compact off-axis see-through head-mounted display using a freeform surface”. In: *Opt. Express* 26.7 (Apr. 2018), pp. 8550–8565.

[81] Austin Wilson and Hong Hua. “Design and demonstration of a vari-focal optical see-through head-mounted display using freeform Alvarez lenses”. In: *Optics express* 27.11 (2019), pp. 15627–15637.

[82] Austin Wilson and Hong Hua. “Design and prototype of an augmented reality display with per-pixel mutual occlusion capability”. In: *Optics express* 25.24 (2017), pp. 30539–30549.

[83] Austin Wilson and Hong Hua. “Design of a Pupil-Matched Occlusion-Capable Optical See-Through Wearable Display”. In: *IEEE Transactions on Visualization and Computer Graphics* (2021).

[84] Christine E Wright and Nevilee Drasdo. “The influence of age on the spatial and temporal contrast sensitivity function”. In: *Documenta Ophthalmologica* 59.4 (1985), pp. 385–395.

[85] Yuhang Wu et al. “Design of retinal-projection-based near-eye display with contact lens”. In: *Optics express* 26.9 (2018), pp. 11553–11567.

[86] Jianming Yang et al. “Method of achieving a wide field-of-view head-mounted display with small distortion”. In: *Optics letters* 38.12 (2013), pp. 2035–2037.

[87] Chanhyung Yoo et al. “Retinal projection type lightguide-based near-eye display with switchable viewpoints”. In: *Optics express* 28.3 (2020), pp. 3116–3135.

[88] Yan Zhang et al. “Optical see-through augmented reality displays with wide field of view and hard-edge occlusion by using paired conical reflectors”. In: *Optics Letters* 46.17 (2021), pp. 4208–4211.

[89] Xianjun Sam Zheng et al. “Eye-wearable technology for machine maintenance: Effects of display position and hands-free operation”. In: *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*. 2015, pp. 2125–2134.

[90] Pengcheng Zhou et al. “Compact design for optical-see-through holographic displays employing holographic optical elements”. In: *Optics express* 26.18 (2018), pp. 22866–22876.

[91] Zoran Zivkovic and Olaf Booij. “How did we built our hyperbolic mirror omnidirectional camera-practical issues and basic geometry”. In: *University of Amsterdam, Tech. Rep. IAS-UVA-05-04* (2005).

Publication List

Peer reviewed journal paper

[1] Yan Zhang, Xiaodan Hu, Kiyoshi Kiyokawa, Naoya Isoyama, Nobuchika Sakata, and Hong Hua, "Optical see-through augmented reality displays with wide field of view and hard-edge occlusion by using paired conical reflectors," Opt. Lett. 46, 4208-4211 (2021).

[2] Yan Zhang, Xiaodan Hu, Kiyoshi Kiyokawa, Naoya Isoyama, Hideaki Uchiyama, and Hong Hua, "Realizing mutual occlusion in a wide field-of-view for optical see-through augmented reality displays based on a paired-ellipsoidal-mirror structure," Opt. Express 29, 42751-42761 (2021).

Peer reviewed journal paper (Letters, technical report etc.)

[3] Yan Zhang and Kiyoshi Kiyokawa: “Occlusion Displays,” in Handbook of Visual Display Technology, 3rd Edition, Springer. (to appear)

Peer reviewed international conference

[4] 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," 2020 IEEE International Symposium on Mixed and Augmented Reality (ISMAR), 2020, pp. 301-311.

[5] Yan Zhang, Kiyoshi Kiyokawa, Naoya Isoyama, and Nobuchika Sakata, “Design and Prototyping of Wide Field of View Optical See-through Head-Mounted Displays with Per-pixel Occlusion Capability,” 2019 IEEE International Symposium on Mixed and Augmented Reality (ISMAR) Adjunct, Oct., 2019.

[6] X. Hu, Y. Zhang, N. Isoyama, N. Sakata and K. Kiyokawa, "Design and Prototyping of Computational Sunglasses for Autism Spectrum Disorders," 2021 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW), 2021, pp. 581-582.

Patent (filed)

[7] Yan Zhang, Kiyoshi Kiyokawa, and Nobuchika Sakata. "シースルー型ディスプレイ装置." PCT/JP2020/38512.

[8] Yan Zhang, Kiyoshi Kiyokawa, and Nobuchika Sakata. "シースルー型ディ

スプレ イ装置." 特願2019-188256.