

# Perception-Driven Soft-Edge Occlusion for Optical See-Through Head-Mounted Displays

Xiaodan Hu, Yan Zhang, Alexander Plopski, Yuta Itoh, Monica Perusquía-Hernández, Naoya Isoyama, Hideaki Uchiyama, and Kiyoshi Kiyokawa

**Abstract**—Systems with occlusion capabilities, such as those used in vision augmentation, image processing, and optical see-through head-mounted display (OST-HMD), have gained popularity. Achieving precise (hard-edge) occlusion in these systems is challenging, often requiring complex optical designs and bulky volumes. On the other hand, utilizing a single transparent liquid crystal display (LCD) is a simple approach to create occlusion masks. However, the generated mask will appear defocused (soft-edge) resulting in insufficient blocking or occlusion leakage. In our work, we delve into the perception of soft-edge occlusion by the human visual system and present a preference-based optimal expansion method that minimizes perceived occlusion leakage. In a user study involving 20 participants, we made a noteworthy observation that the human eye perceives a sharper edge blur of the occlusion mask when individuals see through it and gaze at a far distance, in contrast to the camera system's observation. Moreover, our study revealed significant individual differences in the perception of soft-edge masks in human vision when focusing. These differences may lead to varying degrees of demand for mask size among individuals. Our evaluation demonstrates that our method successfully accounts for individual differences and achieves optimal masking effects at arbitrary distances and pupil sizes.

**Index Terms**—Occlusion, blur perception, user-preferred model

## 1 INTRODUCTION

Spatial light modulators (SLMs) with selective occlusion capability have been utilized in vision augmentation [1], [2], image processing [3], [4], and optical see-through head-mounted displays (OST-HMDs) [5], [6], [7]. As a type of SLM, a transparent liquid crystal display (LCD) is often used for light modulation due to its simple structure, light weight, and easy control [8], [9], [10]. By adjusting the transmittance of each liquid crystal (LC) to flexibly filter incident light, transparent LCDs change how the scene appears to humans by reducing scene contrast [1], [2], [11], highlighting emphasized areas [4], or enhancing virtual imagery [1], [12].

Some OST-HMDs employ transparent LCDs to achieve occlusion. However, most of them require complex optical systems to generate sharp occlusion masks (hard-edge occlusion), as shown in Fig. 1a [5]. Otherwise, the occlusion mask displayed by the single transparent LCD becomes blurred (soft-edge occlusion) due to the out-of-focus LCD plane in the users' eyes, as shown in Fig. 1b [13]. This effect is caused by the LCD plane and the focus plane being far apart from each other, resulting in significant blurring. This blurring process can be conceptualized as a convolution of a kernel [14], leading to a reduction in the intensity of the

mask's edge resulting in insufficient occlusion.

One solution to avoid insufficient occlusion is to expand the occlusion mask in a morphological way [3], [12], where the morphological filter size should be approximately equal to the aperture or pupil radius. However, applying the aperture radius as the expansion radius would cause occlusion leakage; that is, the occlusion mask tends to completely block the incident beam to the eyes and leak into the background [6], [12], [15].

Although an efficient way to overcome insufficient occlusion and leaked occlusion issues still needs exploration, devices with a single transparent LCD have found applications in commercial products, e.g., Magic Leap 2<sup>1</sup> [16], and have the potential to mitigate eye problems of users [2], [11].

An optimal method of rendering the occlusion mask that minimizes insufficient and leaked occlusion issues is important for improving the user experience in single transparent LCD-based applications. The impact of occlusion leaking and compensation algorithms have been studied on camera systems [11], [12], [17]. However, blur perception of the human visual system is much more complex than that of a naïve camera system. Therefore, humans may perceive the occlusion mask differently from camera systems [18], [19]. So far, there is no quantitative evaluation of the blur perception for users using an occlusion-capable device with a single transparent LCD.

In this paper, we present the first user study evaluating soft-edge occlusion perception. Additionally, we introduce a novel user preference-based expansion strategy capable of determining the optimal expansion radius for users across various distances and pupil sizes. This is achieved through a single calibration of the focusing plane's calibration pattern

- Xiaodan Hu, Alexander Plopski are with the Graz University of Technology, Graz 8010, Austria. E-mail: {xiaodan.hu, alexander.plopski}@tugraz.at
- Monica Perusquía-Hernández, Hideaki Uchiyama, and Kiyoshi Kiyokawa are with Nara Institute of Science and Technology (NAIST), Ikoma, Nara 6300192, Japan. E-mail: {m.perusquia, hideaki.uchiyama, kiyo}@is.naist.jp
- Yan Zhang is with Shanghai Jiao Tong University, Shanghai 200240, China. E-mail: yan-zh@sjtu.edu.cn
- Yuta Itoh is with the University of Tokyo, Bunkyo, Tokyo 1138654, Japan. E-mail: yuta.itoh@iii.u-tokyo.ac.jp
- Naoya Isoyama is with Otsuma Women's University, Chiyoda, Tokyo 1028357, Japan. E-mail: isoyama@otsuma.ac.jp

1. <https://www.magicleap.com/magic-leap-2>

at a known distance. Moreover, we undertake a comparison of four occlusion mask expansion strategies: no expansion, aperture size-based, point-spread-function-based, and user preference-based. The study encompasses a within-subjects analysis involving 20 participants, assessing these strategies across different distances (close: 46 cm, far: 1.67 m) and locations (central, periphery) in the user's view.

Our results show that, first, human vision is less sensitive to the blurry border caused by a defocused occlusion mask than a camera system, meaning an expansion that may cause occlusion leakage observed on a camera system with the same aperture as a user's pupil size will not necessarily lead to significant leakage in human vision. Second, due to varying abilities in discriminating blur, the demand and tolerance for mask size and its variation vary greatly among individuals. Lastly, the optimization algorithm based on user preferences can predict optimal masking at any distance and pupil size.

Our main contributions include the following:

- A novel user preference-based expansion strategy that yields complete masking in human vision for the soft-edge occlusion using a single-layered transparent LCD.
- A user study that reveals how four expansion strategies for soft-edge occlusion masks are perceived by users.
- A discussion on the factors influencing users' perception of defocused occlusion masks in comparison to camera systems and other users, as well as potential application scenarios for soft-edge occlusion based on our findings and expansion strategy.

## 2 RELATED WORK

In this section, we briefly introduce some occlusion-capable (OC) devices implemented using a single-layer LCD. We then describe methods used to optimize soft-edge occlusion on these devices, as well as user studies done with these devices. Finally, we briefly review how human visual perception differs from camera systems.

### 2.1 OC Devices Using a Single-layer Transparent LCD

As an inexpensive and practical component, transparent LCDs have been a popular tool for manufacturing OC devices.

**Image processing.** Nayar and Branzoi adopted a concept of placing a transparent LCD in front of a camera, using occlusion to filter the incident light that would produce overexposure to achieve high dynamic range imaging [3]. Wetzstein *et al.* carried this idea forward and used transparent LCD-based occlusion to facilitate more image processing including contrast manipulation, color adjustment, object highlighting, etc. [4].

**Augmented reality.** Tatham proposed an augmented reality (AR) display using a single-layer transparent LCD for mutual occlusion between real and virtual objects [13]. Kiyokawa *et al.* developed the first hard-edge OC OST-HMD [5]. Selective blocking of ambient light with occlusion can effectively reduce the effect of ghost-like virtual content

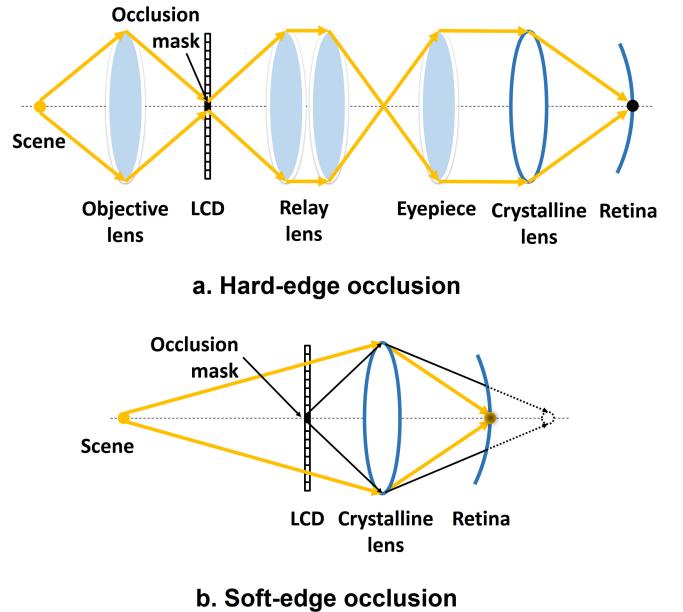


Fig. 1. The schematics of the hard-edge occlusion and the soft-edge occlusion. In the structure with hard-edge occlusion, refraction through multiple lenses can cause a point in the scene to be blocked by a point of equal size on the LCD, which is imaged on the retina as the occlusion pattern on the LCD. However, in the soft-edge occlusion structure, if the same mask is displayed, it will be imaged on the retina in an extremely blurred form due to the defocus of the human eye, which eventually leads to incomplete occlusion.

and make it more vivid. Hiroi *et al.* expanded on this concept, not only modulating the light entering the eyes with occlusion on a single transparent LCD, but also overlaying virtual images from an OST-HMD to enhance the underexposed regions [1]. Their team also used the same structure of overlaying both occlusion and virtual images to compensate for the artifacts generated by occlusion leak [12].

Maimone *et al.* introduced the occlusion feature in their pinlight displays [20]. Although the transparent LCD in their HMD modulates the image on the retina, they also adopted an occlusion mask sub-frame to improve the see-through ability. Recently, Magic Leap 2 applied an occlusion feature to the product, making it the most commercially successful OST-HMD with an occlusion technology to date [16].

**Human-centered computing.** Hu *et al.* assisted photophobics by filtering out environment glare with a transparent LCD, while avoiding excessive shading to ensure a comfortable visual experience [2], [11]. Zhang *et al.* utilized two transparent LCD panels as the lens to construct glasses for simulating visual impairments, specifically central vision loss (occlusion mask in the center) and peripheral vision loss (occlusion mask at the periphery) [10]. They also utilized polymer dispersed LC film to develop smart glasses that effectively block peripheral vision, leading to a significant reduction in the sensation of motion sickness within virtual environments [21]. Moreover, they achieved gaze guidance by utilizing the rendering of a hollow occlusion mask on the LCD, enabling the transfer of eye gaze from one person to another [22].

## 2.2 Optimization for Soft-edge Occlusion

The human eye has only limited depth of field and thus cannot focus on objects placed at different distances simultaneously. This results in a blurred occlusion mask that cannot provide sufficient blocking. Therefore, many researchers have been interested in pixel-perfect hard-edge OC OST-HMDs ever since the first of its kind was developed [5]. However, such optical designs require several components, making the entire HMD bulky as shown in Fig. 1a. This remains an unsolved challenge. Additionally, the limited numerical aperture of lenses imposes restrictions on the field of view (FOV) and can lead to distortion. In contrast, soft-edge occlusion offers several advantages, including a lightweight design, a larger FOV, and minimal distortion [9]. Therefore, the soft-edge occlusion strategy is better suited for comfortable wearable devices intended for everyday use. However, a drawback of soft-edge occlusion is the significant blurring of the occlusion mask due to the eyes' limited depth of field.

Some studies have focused on making optical designs smaller while retaining hard-edge occlusion, such as folding optical paths [7], [23], while others investigated ways to optimize soft-edge occlusion [6], [9], [11].

In theory, if the user looks at infinity, a point on the transparent LCD will be defocused as a result (shown in Fig. 1b). This defocusing process is equivalent to convolving the point spread function (PSF) [14], and the size of the convolution kernel can be approximated as the size of the aperture (see Section 3.4). Hence, Nayar and Branzoi [3] opted to create a deconvolved occlusion mask so that it would theoretically become sharp after further convolution caused by defocusing. However, since the actual transmittance values were far beyond the dynamic range of the LCD, they compromised to perform an expanded operation on the mask in a morphological way, while the expansion radius is that of the aperture. Itoh *et al.* addressed occlusion leakage on OST-HMDs by capturing the user's view of the scene with a camera and overlaying the captured content onto the areas where the occlusion leaked beyond the intended area [12]. To achieve effective occlusion without leakage, Hu *et al.* recently conducted experiments using a camera system to simulate the effect of the defocused occlusion mask and compared it with the range to be occluded. By applying weighted optimization, they were able to determine the optimal expanded radius of the occlusion mask based on the camera's observation [11].

## 2.3 User Studies in Soft-edge Occlusion-capable Devices

Since cameras may reproduce some aspects of human vision, most studies on soft-edge occlusion in the past paid more attention to displaying the visual effect produced by the camera system and analyzing it objectively using histograms, contrast, peak signal-to-noise ratio, etc [1], [3], [4], [6], [12], [24]. Although not many, several studies have conducted user studies.

Wetzstein *et al.* achieved optical image processing using the transparent LCD and validated the effectiveness through a user study [4]. They proposed two prototypes: a scope-like prototype forming hard-edge occlusion and a single-layered

window-style prototype forming soft-edge occlusion. As they applied the hard-edge occlusion prototype for their experiment, the visual perception of the soft-edge occlusion is unknown.

Hiroi *et al.* conducted a preliminary user study for soft-edge occlusion on a single-layered transparent LCD with limited participants [1]. They raised the issue that defocused blurry masks lead to unclear recognition by the user. Itoh *et al.* also conducted an informal test trial with two users [12]. Nevertheless, the blurred mask itself was not directly observed, hence this was not evaluated by the human eye because they overlaid a compensation image onto the occlusion mask.

Zhang *et al.* designed a comprehensive user study to evaluate the performance of the vision loss simulation [10], while the participants had a limited focal distance (25 cm) that perceived a sharp mask. Moreover, their study did not require any real object to be blocked. In this manner, the participants were not concerned about whether and how the mask was blurred.

To the best of our knowledge, no researcher has designed detailed experiments to verify how human vision perceives soft-edge occlusion. In this work, we design and conduct a psychophysical experiment to evaluate the soft-edge occlusion mask in human vision.

## 2.4 Human Visual Perception of Digital Image

It is known that cameras cannot replace human vision, and individual differences can also significantly impact how people perceive digital images [25]. Many psychophysics experiments have been conducted to verify these differences and have shown some critical evidence.

**Individual differences.** The individual differences in contrast sensitivity is a popular topic for exploring normal visual processes and clinical visual disorders. Psychophysical studies have confirmed that the individual differences in contrast sensitivity and acuity exist due to inattention, adaption, and myopia [26]. Hence, it is reasonable to expect individual differences in the perception of soft-edge occlusion. Besides, Grzeczkowski *et al.* summarized that there is no strong correlation between the seemingly very similar visual performances, such as luminance contrast, color, and spatial frequency [27]. This evidence prompts us to create a model that relied entirely on the user's calibration instead of finding correlations inside several visual performances.

**Optical blur model.** Watson *et al.* [19] revealed that optical blur is usually different from Gaussian blur, which is commonly used in many defocus simulation tasks [28], [29]. With visual aberration, the PSF of the human eye can be quite complicated. Thus, in this paper, we do not use, for example, a Gaussian function to simulate an optical blur model; instead, we fit the model based on the user's perception and their preference.

**Blur discrimination.** Several studies investigated out-of-focus blur, similar to that observed in soft-edge occlusions. Sebastian *et al.* [18] physically placed two images at different distances, and then asked participants to discriminate which image was more blurred. They found that images with fewer details yield lower defocus discrimination sensitivity (i.e., it is more difficult to determine which image

is more blurred). This might be related to contrast and luminance ratio to some extent. Moreover, higher standard defocus levels may lead to a higher discrimination sensitivity. However, there is still a large individual difference in the ability to discriminate defocus blur. Another interesting finding is in the discrimination of motion blur. Burr *et al.* [30] summarized that, although moving objects do indeed produce blur as captured by cameras, the reason why humans cannot detect blur is not because of special mechanisms in human vision that remove blur, but because the visual system is unable to make the necessary discrimination of whether a moving object is blurred or not. Similar to motion blur, because individuals' attention is typically focused on distant objects rather than the occlusion mask in front of their eyes, the blur captured by the camera may not be perceived as accurately by the human eye.

All of the above evidence suggests that it may be insufficient to use a camera to capture blurred soft-edge occlusion to determine the effectiveness of a near-eye occlusion-capable device. It is necessary to design a comprehensive experiment to explore how human vision perceives soft-edge occlusion.

### 3 SYSTEM OVERVIEW AND METHODS

In this section, a concise overview of the experimental prototype designs is presented in Section 3.1, followed by a comprehensive description of four distinct expansion strategies in Sections 3.2 to 3.5.

#### 3.1 Experimental Prototype Designs

Hu *et al.* proposed a benchtop prototype [2], [11] that shares a similar architecture with our experimental prototype. Our experimental system is composed of a chinrest, and an eye tracker to obtain the pupil diameter, as depicted in Fig. 4c. In the experimental prototype, a scene camera is fixed on the chinrest to detect the ambient brightness, and a transparent LCD is attached to the eye box to form a monocular structure.

#### 3.2 Naïve Mask

The camera captures the environment and transfers the image data to a modulation function, which generates a naïve mask [11]. The naïve mask is calculated solely based on scene luminance through geometric calibration without additional expansion. It can be expressed by the following equation:

$$M_n = m(I_{SC}), \quad (1)$$

where  $I_{SC}$  represents the pixel intensity obtained by the scene camera after calibration, and  $m(\cdot)$  denotes the modulation function proposed in [11] (Since the modulation is not the focus of this paper, the specific modulation function employed to control the variables in our experiments will be presented in Section 4.2.3). The visualization of  $M_n$  is shown as the dash-dotted black line in Fig. 3a. Subsequently, the naïve mask can be expanded by applying an expansion radius calculated using the aperture-based, PSF-based, or our proposed expansion method.

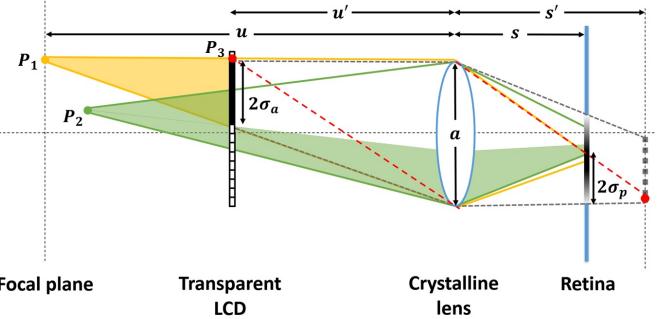


Fig. 2. The schematics of complete blocking by an aperture-based expanded mask and the reason for occlusion leakage. When the human eye focuses on the focal plane of  $P_1$ , the mask (black line and the point  $P_3$ ) will be imaged behind the retina (gray dotted line), resulting in a blurry disk on the retina (gradient black line). At the same time, the mask also obscures light from  $P_2$ . This can cause occlusion leakage, which refers to the blocking of nearby content simultaneously.

#### 3.3 Aperture-based Expansion

As shown in Fig. 2, to block the beam emanating from a point at infinity, an occlusion mask of radius  $\sigma_a$  needs to be rendered on the LCD [12] (see supplementary material for more details). It can be easily calculated that if  $u \gg u'$ , where  $u$  and  $u'$  are the distances from the crystalline lens to the focal plane and the LCD plane, respectively,  $\sigma_a$  can be approximately equal to the pupil radius  $a$  [1], [12]. We refer to this processing of a naïve mask, dilated morphologically according to the aperture or pupil size, as an *aperture-based* expanded mask. The expansion radius  $\sigma_a$  can be calculated as:

$$\sigma_a = \frac{a}{2} \cdot (1 - \frac{u'}{u}), \quad (2)$$

The expansion is achieved by convolving a circular kernel with radius  $\sigma_a$  onto the naïve mask, which is represented as:

$$M_a = M_n \otimes F(\sigma_a), \quad (3)$$

where  $M_a$  refers to the aperture-based expanded mask (shown as the purple solid line in Fig. 3b), and  $F(\cdot)$  is a circular kernel represented as:

$$F(\sigma_a)(x, y) = \begin{cases} 1, & x^2 + y^2 \leq \sigma_a \\ 0, & \text{otherwise.} \end{cases} \quad (4)$$

When an occlusion mask is formed in this dilated manner and the human eye focuses on the focal plane of  $P_1$ , the mask will be imaged behind the retina, resulting in a blurry disk on the retina. Since it blocks light from other directions, such as light from  $P_2$  in Fig. 2, it can cause occlusion leakage, which refers to the blocking of nearby content simultaneously.

#### 3.4 PSF-based Expansion

Consider a simulated disk-like occlusion mask after being out of focus. This mask is shown by the dash-dotted black line as a naïve mask and the blue line as its defocused result in Fig. 3a. The process of defocusing can be represented as convolving the naïve mask with the PSF of a human eye, given by the equation

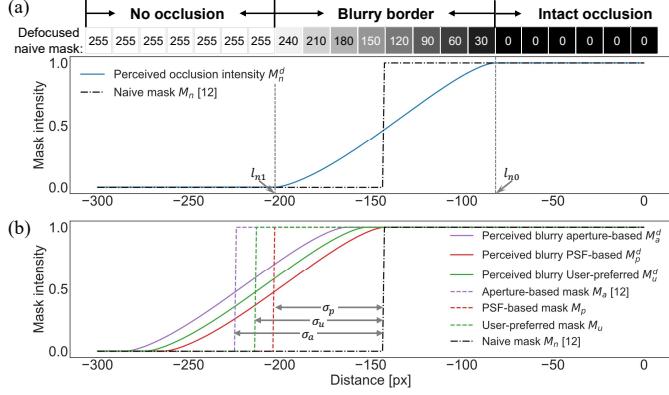


Fig. 3. (a) The intensity profile of the naïve mask as displayed on the LCD and a defocusing simulation of its perceived occlusion intensity using a circular averaging (pillbox) filter. The color blocks on the figure represent pixel values of the defocused naïve mask, from which the defocused mask can be divided into three parts: the white no occlusion region, the gradient blurry border region, and the intact occlusion region that maintains the ability to occlude. The diameter of the convolution kernel spans the distance from the threshold between the intact occlusion and blurry border  $l_{n0}$  to the edge  $l_{n1}$  where the mask disappears, coinciding with the blurry border of the naïve mask. (b) The intensity profile of the naïve mask, three expanded occlusion masks, and the corresponding defocusing simulations.  $\sigma_p$ ,  $\sigma_u$ , and  $\sigma_a$  denote the expansion radii of the PSF-based expanded mask, user-preferred expanded mask, and aperture-based expanded mask, respectively. Note that the intact occlusion region of the defocused PSF-based expanded mask coincides with the naïve mask.

$$M_n^d = M_n \otimes H, \quad (5)$$

where  $M_n^d$  is the defocused naïve mask, shown as the blue line in Fig. 3a, and  $H$  is the PSF acting as the convolution kernel. Notice that the exact form of  $H$  is not necessary and blurring only occurs at the border. The blurry border, defined as the distance between the no occlusion region and the intact occlusion region (from  $l_{n1}$  to  $l_{n0}$  in Fig. 3a), is equal to the diameter of the convolution kernel  $H$ . Similarly, the defocused aperture-based expanded mask can be represented as  $M_a^d = M_a \otimes H$ , shown as the purple solid line in Fig. 3b.

Therefore, it is also possible to determine the radius of the PSF as an expansion radius, represented as  $\sigma_p$  in Fig. 3b. This indicates the use of only the intact occlusion mask to block the incoming light. The PSF-based expanded mask  $M_p$  (shown as the red solid line in Fig. 3b) and the defocused version  $M_p^d$  (shown as the red dashed line in Fig. 3b) can be created by referring to Equations 4 and 5:

$$M_p = M_n \otimes F(\sigma_p), \quad M_p^d = M_p \otimes H. \quad (6)$$

In Fig. 3b, the intact occlusion region of  $M_p^d$  is overlaid with  $M_n$ , indicating the use of the intact occlusion intensity of  $M_p^d$  to block the light source (reverse intensity of the naïve mask).

In the imaging process of a point  $P_3$  on the LCD, as shown in Fig. 2, the point will spread into a disk, which is also the PSF with radius  $\sigma_p$  on the retina. Using geometric calculations, we can determine the PSF-based expansion radius  $\sigma_p$  as follows:

$$\sigma_p = \frac{a}{2} \cdot \left(1 - \frac{s}{s'}\right), \quad (7)$$

where  $s$  represents the transverse diameter of the eyeball, and  $s'$  represents the distance from the crystalline lens to the imaging plane of the LCD panel (shown by the gray dotted line in Fig. 2). Combining Equation 7 and the thin lens equations

$$\frac{1}{u} + \frac{1}{s} = \frac{1}{f}, \quad \frac{1}{u'} + \frac{1}{s'} = \frac{1}{f},$$

we can derive the following equation for the PSF-based expansion radius  $\sigma_p$ :

$$\sigma_p = \frac{as}{2} \cdot \left( \frac{1}{u'} - \frac{1}{u} \right). \quad (8)$$

### 3.5 User-preferred Expansion

The user-preferred expansion radius is based on each individual's current pupil size. That is, the optimal occlusion mask for each user is set to block the target regions without perceived leaking. While implementing the system, we noticed an interesting phenomenon: when we observed distant objects through the occlusion mask, the blurring at the edges of the mask appeared to be sharper. In other words, the range of blurry edges that can be perceived is smaller, which makes it possible for users to determine an expansion radius that can perform complete occlusion according to their perception. We will provide further explanations for this observation in the Section 6.1.

We provide a detailed procedure for determining this expansion radius in Section 4.3. If we assume that a defocused expanded occlusion mask according to user preference is illustrated by the green dashed line in Fig. 3b, then the user-preferred mask is an expansion of the naïve mask with an expansion radius of  $\sigma_u$ .

When dilating the naïve mask by a PSF-based expansion radius  $\sigma_p$ , the intact occlusion region can be effectively used as an occlusion mask, as explained in Section 3.4. However, different individuals may have varying perceptions when recognizing the blurry border, which can introduce biases to the ideal conditions. As a result, the relationship between the pupil size and the ideal expansion radius  $\sigma_p$  needs to be adjusted, leading to the following rewritten equation from Equation 8 as:

$$\sigma_u = \frac{a}{2} \cdot (s + \Delta s) \cdot \left( \frac{1}{u'} - \frac{1}{u} \right) + b. \quad (9)$$

Here,  $\Delta s$  accounts for imaging bias due to individual differences, and  $b$  represents a global bias. In the experiment, we asked the participants to determine their preferred expansion radius through a calibration process multiple times under different illuminance conditions. The biases  $\Delta s$  and  $b$  can be obtained by performing a linear regression on the collected data for each participant.

Once the expansion radius is determined, similar to Equation 6, the user-preferred expanded mask  $M_u$  (shown as the green solid line in Fig. 3b) and its defocused version  $M_u^d$  (shown as the green dashed line in Fig. 3b) can be represented as:

$$M_u = M_n \otimes F(\sigma_u), \quad M_u^d = M_u \otimes H. \quad (10)$$

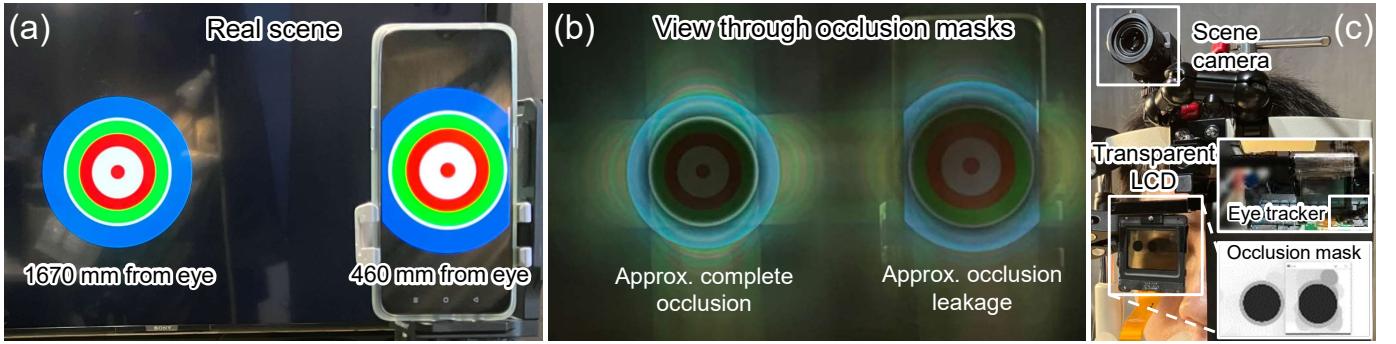


Fig. 4. Experimental scenes. (a) Two quantitative patterns are displayed: one on a television positioned 1670 mm from the user's eye and another on a smartphone positioned 460 mm from the eye. The focal plane is set on the television. (b) A view of the pattern through two occlusion masks, captured by an iPhone SE3 (3.99 mm focal length, f/1.8) with the focal plane positioned on the television (pattern on the left). Both occlusion masks are of the same size and are able to completely occlude the left pattern. Since the right pattern is closer, the required expansion radius for complete occlusion is smaller compared to the left pattern. As a result, the right view shows a slight occlusion leakage effect. (c) Experimental setup: a scene camera and an eye tracker are used to compute the optimal occlusion mask displayed on a single transparent LCD.

## 4 USER STUDY

In this section, we provide an introduction to the basic information of the participants and elaborate on the recruitment process in detail in Section 4.1. We further present a comprehensive overview of the experimental setup in Section 4.2, which encompasses various aspects such as the configuration of the experimental environment (Section 4.2.1), hardware setup (Section 4.2.2), the utilization of quantitative patterns for obtaining user perceptions (Section 4.2.3), the evaluation approach of the four occlusion masks at different distances and FOVs (Section 4.2.4), and the necessary preparations prior to the experiment (Section 4.2.5). Lastly, in Section 4.3, we provide a detailed description of the experimental procedure, which consists of calibration and evaluation tasks.

### 4.1 Participants

We recruited 20 participants (8 female) from students and faculty of Nara Institute of Science and Technology via email advertisement. The mean age of the participants was 26.9 with a standard deviation of 4.06, and a range of 22–37. The user study was approved by a university ethics committee. Of the participants, eleven had normal vision and nine were nearsighted. Among the myopic participants, seven wore contact lenses, and the remaining two wore eyeglasses to correct their vision to normal. Participants with color blindness were excluded from the study. Additionally, we later learned that two of the participants in our study were diagnosed with attention deficit hyperactivity disorder (ADHD), and their results exhibited unconventional patterns.

### 4.2 Experimental Configuration

#### 4.2.1 Experimental Environment

The experiment was conducted in a darkroom with dimensions of  $2.5 \times 1.8 \times 2.0 \text{ m}^3$ . The illuminance was controlled by two Philips Hue smart bulbs<sup>2</sup>, which could be remotely modulated using a smartphone application from outside the darkroom. However, we discovered a hardware limitation where the transparent LCD would only transmit 10% of the

light without displaying the mask, hindering the effective stimulation of the pupil and making it difficult to obtain data on smaller pupils. To address this issue, we included a light source from a floor lamp on the user's side to stimulate the pupil through the gap between the eye and the LCD.

We obtained the illuminance measurements in Table 1 by placing the illuminance meter at the eye box and the sensor surface facing the LCD. However, due to the presence of the plastic protective bar around the sensor, the light rays were partially obstructed and resulted in inaccurate measurements. To mitigate this issue, we positioned the sensor slightly further away from the LCD than the actual position of the human eye. Therefore, while the illuminance values in Table 1 may not precisely represent the exact illuminance on the human eye, they still provide a reliable indication of the luminance environment during the experiment.

To accommodate space limitations, a table with a chinrest and a height-adjustable stool was set up at one end of the darkroom. At the other end, a 55-inch Sony KJ-55X85J television ( $3840 \times 2160$  pixels, with a typical brightness of  $563 \text{ cd/m}^2$ ) was positioned to allow for maximum object distance. The distance  $u_f$  from the average eyebox to the screen panel was approximately 1.67 meters. Additionally, a 6.4-inch OPPP Reno A smartphone ( $2340 \times 1080$  pixels, with a typical brightness of  $430 \text{ cd/m}^2$ ) was placed vertically at a distance  $u_c$  of 46 cm from the average eyebox, supported by a height-adjustable tripod, to verify the effectiveness of the proposed method at different distances (shown in Fig. 4a). The table and the chinrest were both height-adjustable, and a strap with Velcro was wrapped around the post of the chinrest to fixate the head.

TABLE 1  
Six illuminance conditions, with "B" representing the smart bulb and "L" representing the floor lamp.

	IC 1	IC 2	IC 3	IC 4	IC 5	IC 6
Lighting	0	1 B	2 B	2 B, L	1 B, L	L
Illum. [lux]	0.03	140	230	520	360	270

2. <https://www.philips-hue.com/en-us>

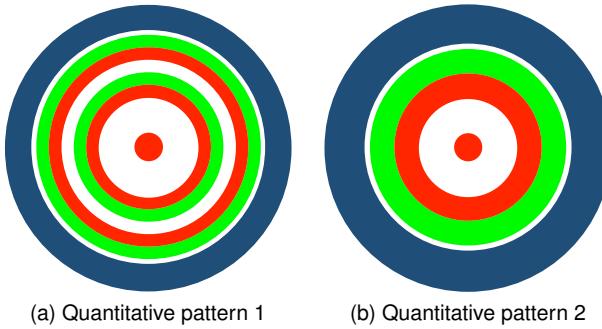


Fig. 5. Quantitative pattern for calibration and evaluation. We define a complete occlusion as a state in which users are able to see only the outermost white ring with clarity, while the inner rings are obscured by the occlusion mask and appear darker. Pattern 1 is utilized for finer calibration. However, due to the limited size of the phone screen, we can only display pattern 2 for evaluation tasks.

#### 4.2.2 Hardware

The components of the experimental prototype are all fixed around the chinrest, as illustrated in Fig. 4c. A monochrome camera, produced by Hikvision and equipped with a SONY IMX 287 sensor ( $720 \times 540$  pixels, 12-bit depth) and an 8.5mm f/1.3-f/16 lens, is secured to the crossbar above the chinrest using an articulating arm. The camera is connected through a USB 3.0 link and transmits a real-time video stream in a multi-threaded manner via a software development kit provided by the manufacturer. A SONY LCX017266 LCD panel ( $1024 \times 768$  pixels,  $36.9 \times 27.6$ mm active area, 705 dpi, 60 Hz, monochrome) with a linear polarizer affixed is clamped to the front of the chinrest. The transmittance of the LCD panel is 21% when powered on, and it decreases to 10% when an additional linear polarizer is attached due to the original LCD panel having only one polarizer. The lowest transmittance of the LCD is approximately 0.02% when rendering pixels with a pixel intensity of 0. To measure pupil size, an eye tracker from Pupil Labs (Pupil Core<sup>3</sup>) is attached to the side post of the chinrest using an articulating arm. The eye camera of the eye tracker captures images from the bottom up, while the angle of the tracker can be adjusted manually to accommodate different eye positions of the participants.

#### 4.2.3 Quantitative Patterns

As the scene observed by the human eye is inaccessible, we have designed two quantitative patterns resembling an archery target (Fig. 5). The pattern enables users to adjust the size of the occlusion mask according to their preferences, in order to achieve the desired expansion radius. Additionally, users can indicate the size of the occlusion mask they are able to see, which is used to evaluate whether the size of the mask aligns with their perception.

Pattern 1 (illustrated in Fig. 5a) consists of five thick rings (outermost green to innermost red) and a thin white ring displayed on the television, allowing users to calibrate the size of the occlusion mask. We define a complete occlusion as a state in which users can clearly see only the outermost white ring, while the inner rings are obscured by the occlusion mask and become darker (an approximate view is

shown as Fig. 4 on the left). However, due to the limited size of a smartphone screen, displaying Pattern 1 may make it difficult for users to distinguish the rings. Therefore, we have reduced the number of rings to three (as shown in Fig. 5b) in Pattern 2 and have rendered it on the smartphone screen for evaluation. The widths of the rings in Pattern 1 and Pattern 2 are both  $0.75^\circ$ , while the outermost white rings are measured at  $0.1^\circ$  from an average eyebox.

To ensure consistent pixel intensities of the occlusion masks under varying lighting conditions, we maximized the screen brightness of the television and the smartphone, as well as increased the exposure time of the scene camera. These adjustments enabled the modulation function to generate a circular naïve mask based solely on the portion within the outermost white ring. The pixel intensities of the resulting naïve mask remained constant at 0 across different illuminance conditions, representing the state with the lowest transmittance. Notably, even with the lowest transmittance, the human eye can still perceive specific colors that are obstructed by the mask due to the high brightness of the screen.

As longer wavelengths are known to have stronger penetrating power [31], we conducted tests to evaluate the recognition of several long-wavelength colors when observed through the LCD with the lowest transmittance. Finally, we selected red and green colors for the rings as they were found to be easily distinguishable by the human eye. White was used as a contrasting color to the black occlusion mask, and a red dot was placed in the center of the patterns to assist users in maintaining their gaze on it, and ensuring a consistent focusing distance.

#### 4.2.4 Verification across Different Distances and Visual Fields

We verified the naïve mask and three kinds of expanded occlusion masks at a close distance (CD)  $u_c$  and a far distance (FD)  $u_f$ , central visual field (CVF), and peripheral visual field (PVF), respectively, as shown in Fig. 6. After the user fine-tunes the mask positions, the four positions of the patterns ( $L_1$  to  $L_4$ ), the two occlusion mask positions ( $M_1$  and  $M_2$ ), and the position of the human eye should be nearly at the same height. Additionally,  $L_1$  and  $L_3$  in the middle should be in the same viewing direction as the center occlusion mask, whereas  $L_2$  and  $L_4$  in the periphery should be in the same viewing direction as the right occlusion mask.

Using the geometric relationship, the distance between  $L_3$  and  $L_4$  ( $d_c$ ) and the distance between  $M_1$  and  $M_2$  ( $d_l$ ) can be easily calculated once the distance between  $L_1$  and  $L_2$  ( $d_f$ ) is established. Specifically, we can obtain these distances by the following equations:

$$d_c = d_f \cdot \frac{u_c}{u_f}, \quad d_l = d_f \cdot \frac{u'}{u_f}. \quad (11)$$

From Equations 2, 8 and 9, it can be derived that different depths lead to different expansion radius. In fact, we can obtain the depth information of the scene through a depth camera, thus assigning a corresponding convolution kernel  $F$  to each pixel in the naïve mask generated by the scene. However, this approach is computationally intensive and can lead to significant latency. To evaluate the effectiveness

3. <https://pupil-labs.com/products/core/>

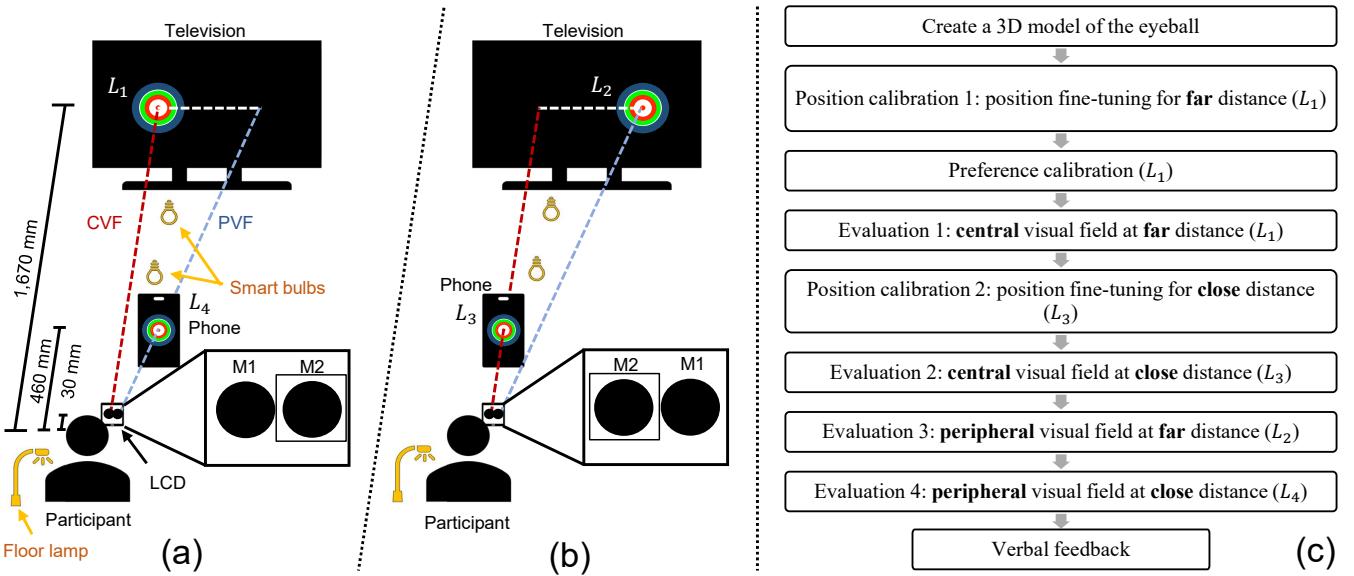


Fig. 6. Experimental configuration (a,b) and procedure (c). Participants were shown a target to focus on within their central visual field (CVF). This target was either shown on a television (a) or on a mobile phone (b). During calibration, participants had to judge how well the mask M1 occludes the target shown on the television in the calibration condition  $L_1$ . During the experiment, participants had to judge how well mask M2 occludes the targets shown on the television ( $L_1, L_2$ ) and on the mobile phone ( $L_3, L_4$ ). Lighting conditions were controlled through a floor lamp and smart bulbs in the environment.

of the four occlusion masks at different distances and fields of view in real time, we created two windows to render the occlusion masks with different expansion radii. As the screen width of the smartphone is much smaller than that of the television, we placed the window rendering the occlusion mask of the smartphone on top of the television window, as shown in Fig. 6a for Mask 1 (television window) and Mask 2 (smartphone window). Both windows are movable by the instruction of a keyboard.

#### 4.2.5 Calibration between the Scene Camera and the Human Eye

To calibrate the relationship between the scene camera and the human eye, we utilized the same Hikvision camera to simulate the human eye and placed it at the position of the eye box. The mapping between the two cameras was easily achieved using a 2D homography matrix. We displayed a chessboard pattern on both the television and the smartphone screens respectively and calculated the homography matrices using the library functions in OpenCV-Python in two parts. At this point, we had ensured that the parallel planes of the television, smartphone, and the LCD panel were mapped. However, since the position of the participants' eyes varied, they needed to manually fine-tune the position of the mask afterward. The fine-tuning process is described in detail in Section 4.3.1.

### 4.3 Experimental Procedure

The experiment was mainly divided into calibration and evaluation tasks. During these tasks, participants were instructed to focus their gaze on the red dot of quantitative patterns in the central visual field, i.e.,  $L_1$  or  $L_3$  as shown in Fig. 6. Participants were also instructed not to focus on the occlusion mask, keeping it out of focus. Following the completion of the entire experiment, we collected verbal

feedback from the participants. Since the experiment requires constant gaze, which increases eye strain, participants were informed that they could stop the experiment at any time. Also, participants wearing contact lenses were advised to take eye drops before the experiment to prevent dry eyes due to excessive eye use during the experiment. Once the participants had adjusted the height of the stool and chinrest to their desired positions, we rotated the angle of the eye tracker and verified that the eye camera was able to successfully detect their pupils. We then secured their head in place with a strap and utilized the Pupil Capture application provided by Pupil Labs to create a 3D model of their eyeballs. The experiment began and was recorded after ensuring that the pupil was tracked consistently.

#### 4.3.1 Calibration Tasks

The calibration tasks were divided into two parts: position fine-tuning and size determination of the occlusion mask. The position fine-tuning needed to be performed twice for the far (Position calibration 1 in Fig. 6c) and close distance (Position calibration 2 in Fig. 6c), while the size determination (Preference calibration in Fig. 6c) is only needed to be performed once with the Pattern 1 on the television screen to create the individual model by Equation 9.

**Position Calibration** During the position fine-tuning (Position calibration 1), participants fixated on the red dot in Pattern 1 in the central view displaced on the television and used a wireless keyboard to move the window of Mask 1 until it was concentric with Pattern 1. We put raised rubber stickers on the keyboard's control keys ('w', 'a', 's', and 'd') to make it easier for participants to press keys out of sight. During Position calibration 2, participants repeated the same process but focused on the Pattern 2 displayed on the phone screen. After each calibration, we recorded the position of the windows.

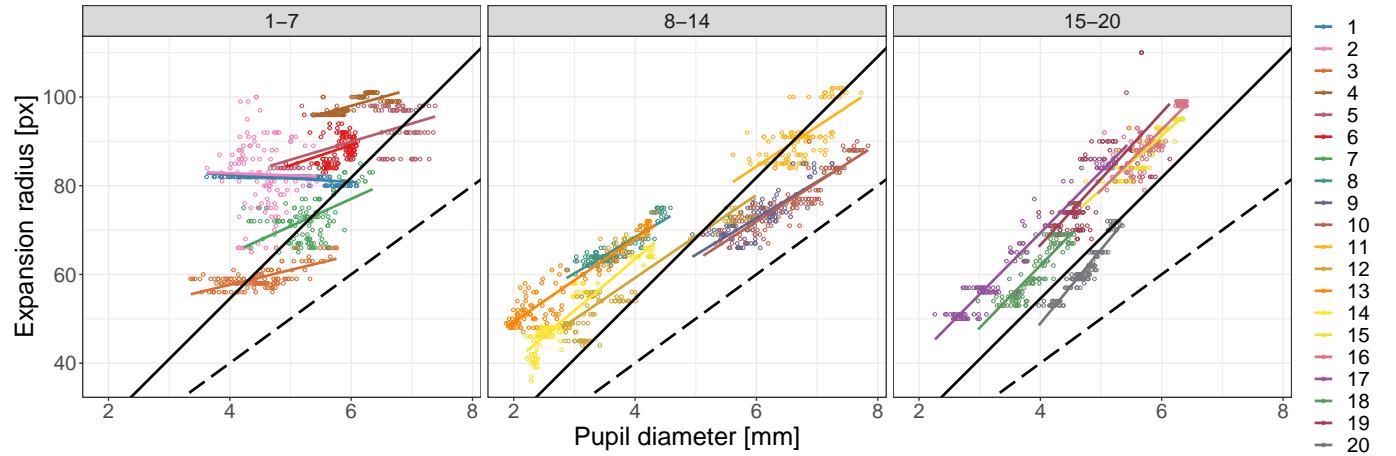


Fig. 7. Visualization and linear fit of calibration results from 20 participants. The participants' IDs are ordered in the ascending order of the slope of their linear models, and the data are divided into three sub-figures, each containing 7, 7, and 6 sets of data, respectively. The solid and dashed black lines indicate the linear models of aperture-based and PSF-based expanded occlusion masks, respectively. (Note that we did not visualize the naïve mask in the figures because according to Equation 1, the naïve mask is independent of the pupil diameter, i.e., it is a horizontal line with y-axis equal to 0.)

**Preference Calibration** After Position calibration 1, we lowered the curtain of the darkroom and let participants decide their preferred mask size based on their current pupil size. The determination was performed under six illuminance conditions (IC) listed in Table 1.

Participants were given the freedom to adjust the occlusion mask to their desired size. They could resize the mask using two keys ('J' and 'I') on the wireless keyboard, and each key press would either dilate or erode the mask by one pixel. Since the pupil size was constantly changing, we instructed participants to press a control key ('p') on the keyboard as soon as they saw a complete occlusion mask in their visual field. At this moment, the pupil size and expansion radius was saved. Furthermore, since the pupil size may also change only slightly over time, or the current mask may remain a complete mask for the participant's perception for a long time. In this case, we encouraged participants to save the current state of the pupil and occlusion mask every 1-2 s as long as they keep seeing a complete mask. However, once they perceived that the mask was insufficient or leaking out, they would resize it to the appropriate size and save it again.

We collected a total of 20 sets of data, which included the participant's current pupil diameter and corresponding expansion radius for each illumination condition. To ensure accurate measurements, we instructed the participants to close their eyes for 30 seconds whenever switching between illumination levels, allowing them to rest and stabilize their visual state.

#### 4.3.2 Evaluation Tasks

After completing the preference calibration, we conducted a linear regression on the 120 sets of data and derived the bias values  $\Delta s$  and  $b$  in Equation 9. Using these values, we calculated the corresponding  $\sigma_u$  for each participant. During the evaluation task, we displayed four types of occlusion masks (naïve, aperture-based, PSF-based, and user-preferred masks) on the transparent LCD and asked

participants one question: **How many rings are visible/not blocked in your field of view?** Participants responded on a scale from  $-1$  to  $3$ . Since our primary concern is whether the mask exhibits any leakage or not, we assigned a single scale of  $-1$  to represent occlusion leak. A value of  $0$  indicates complete occlusion, while values  $1$  to  $3$  represent the number of visible internal rings.

To prevent participants from accidentally changing the already calibrated position, we instructed them to use the keyboard keys to communicate their answers. The following instructions were given:

- If participants can see  $n$  internal rings, they should press the 'J' button  $n$  times.
- If participants notice a leaked occlusion, meaning that they cannot see the outermost white ring in their field of view, they should press the 'I' button once.
- If participants can observe a complete occlusion, where they can only see the outermost white ring but no internal rings, they should press the 'p' button to communicate their response to us.

The evaluation tasks were performed across two distances, two visual fields (center and periphery), and three illuminance conditions ( $2 \times 2 \times 3 \times 4$  (masks) = 48 trials). The three illuminance conditions included IC 1, IC 4, and dynamic illuminance controlled by the application of smart bulbs, which allowed for changeable pupil sizes. The order of mask presentation was counterbalanced among participants to mitigate potential order-related biases.

## 5 RESULTS AND ANALYSIS

### 5.1 Linear Models Obtained from Calibration Data

To visually represent individual differences in participants' perception of the proposed occlusion mask, we have included the data collected during calibration in Fig. 7 and fitted it with a linear regression model. We arranged the 20 participants' IDs in descending order of slope and plotted

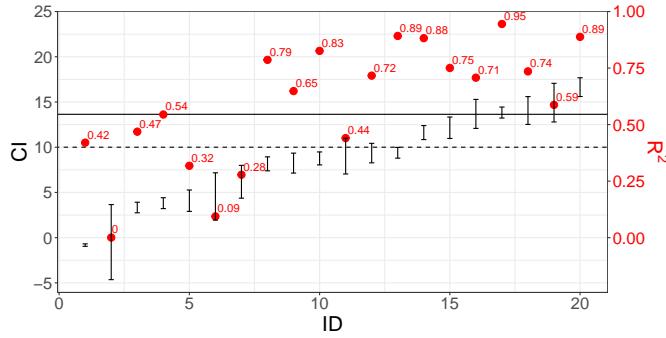


Fig. 8. Coefficient of determination ( $R^2$ ) and confidence interval (CI) for the slope coefficient of 20 linear models used to estimate the performance of the linear regression. The error bars represent the confidence intervals for the slopes, and the red dots indicate the  $R^2$  value, with the specific value indicated above each dot. The solid and dashed lines signify the slopes of the aperture-based (13.639) and PSF-based (10.002) models, respectively.

them in three sub-figures. Our findings show that the linear models with similar slopes tend to overlap and are predominantly situated close to the aperture-based model.

In the case of the two participants with ADHD, their linear models differed significantly from those of participants with typical development. This was evident by the distribution of sample points for their mask size being restricted to a narrow range; [80, 82] for participant #1 and [96, 101] for participant #4, while their pupil size continued to fluctuate.

We evaluated the efficacy of the linear fit by assessing the coefficient of determination ( $R^2$ ) and the confidence interval of the slope for each linear model. The results are presented in Fig. 8. The mean  $R^2$  value is 0.61, and the median is 0.72, indicating that the linear models are moderately effective at fitting users' perceptions. Participants #2 and #6 were wearing contact lenses that caused their eyes to blink incessantly due to dryness during prolonged gaze, causing their pupils to change significantly and interfering with the perception of the defocused mask. If we excluded these two users and the two participants with ADHD, the mean  $R^2$  value would be 0.69, and the median would be 0.74.

Furthermore, Fig. 8 shows a trend where higher confidence intervals of the slope correspond to higher  $R^2$  values, implying that for most participants, the slope consistent with their visual perception is comparable to the models of the aperture-based and PSF-based expanded occlusion masks. However, participants with IDs greater than eight, i.e., those with relatively high coefficients of determination, had slopes ranging between [8.24, 16.64] with a standard deviation of 3.74, indicating considerable individual differences.

We also plotted the slope and intercept data for the 20 users as a scatter plot, as shown in Fig. 9. We observed that the scatter in the plot exhibited a decreasing trend that appeared to be well-suited to a linear model. This seems to be an intriguing discovery, but we are unable to give its possible reasons.

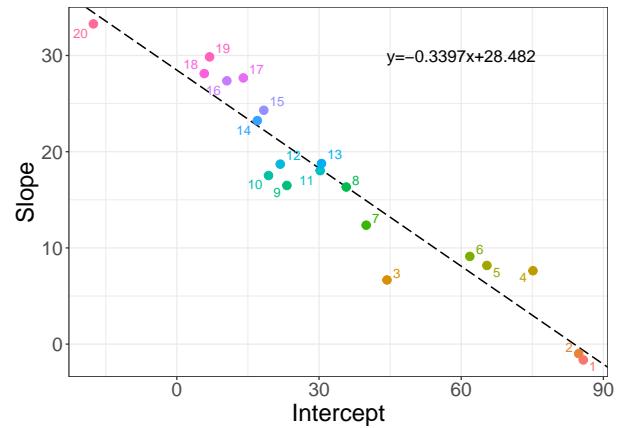


Fig. 9. A scatter plot showing the intercepts and slopes of the linear regression models for the 20 participants. The IDs of each participant are represented on the corresponding points.

## 5.2 Results of the Evaluation for Four Types of Occlusion Masks

To compare the subjective rating for four types of occlusion masks, we conducted the Kruskal-Wallis (KW) test under four conditions (CVF, PVF, FD, and CD) and post-hoc pairwise comparisons using the Dunn tests with Benjamini-Hochberg  $p$ -value correction. Furthermore, the data of participants with ADHD was analyzed separately first. Finally, the data of all participants were analyzed jointly. The results revealed significant differences between the four masks in all six analyses and is summarized in Fig. 10 and Fig. 11.

For the four conditions, the KW test indicated significant differences between the masks (CVF:  $H(3) = 296.64, p < 2.2 \times 10^{-16}$ ; PVF:  $H(3) = 290.96, p < 2.2 \times 10^{-16}$ ; FD:  $H(3) = 312.52, p < 2.2 \times 10^{-16}$ ; CD:  $H(3) = 275.3, p < 2.2 \times 10^{-16}$ ), with the Dunn tests revealing that scores of the proposed mask were significantly different from those of the aperture-based mask (CVF:  $p = 8.56 \times 10^{-4}$ ; PVF:  $p = 3.57 \times 10^{-4}$ ; FD:  $p = 9.33 \times 10^{-4}$ ; CD:  $p = 2.72 \times 10^{-4}$ ), the PSF-based mask (CVF:  $p < 2.2 \times 10^{-16}$ ; PVF:  $p < 2.2 \times 10^{-16}$ ; FD:  $p < 2.2 \times 10^{-16}$ ; CD:  $p < 2.2 \times 10^{-16}$ ) and the naïve mask (CVF:  $p < 2.2 \times 10^{-16}$ ; PVF:  $p < 2.2 \times 10^{-16}$ ; FD:  $p < 2.2 \times 10^{-16}$ ; CD:  $p < 2.2 \times 10^{-16}$ ). Similarly, for the two participants with ADHD, the Dunn test also showed that the proposed mask significantly differed from aperture-based ( $p = 7.14 \times 10^{-7}$ ), PSF-based ( $p = 7.63 \times 10^{-10}$ ) and naïve masks ( $p = 1.10 \times 10^{-13}$ ). Finally, for all ratings, the Dunn test still showed a significant difference between the proposed mask and the other three masks (aperture-based:  $p = 8.67 \times 10^{-7}$ , PSF-based:  $p < 2.2 \times 10^{-16}$ , naïve:  $p < 2.2 \times 10^{-16}$ ). We also presented histograms of the user scores for the four masks across the four conditions, as well as for participants with ADHD and all data, which is shown in Fig. 11.

Additionally, participants verbally reported that they were able to observe a relatively sharp mask. This enabled them to easily identify the border of the mask and perform the preference calibration. When we showed the participants the defocused masks captured by the camera using a similar aperture size, they all reported observing masks without such wide blurred borders.

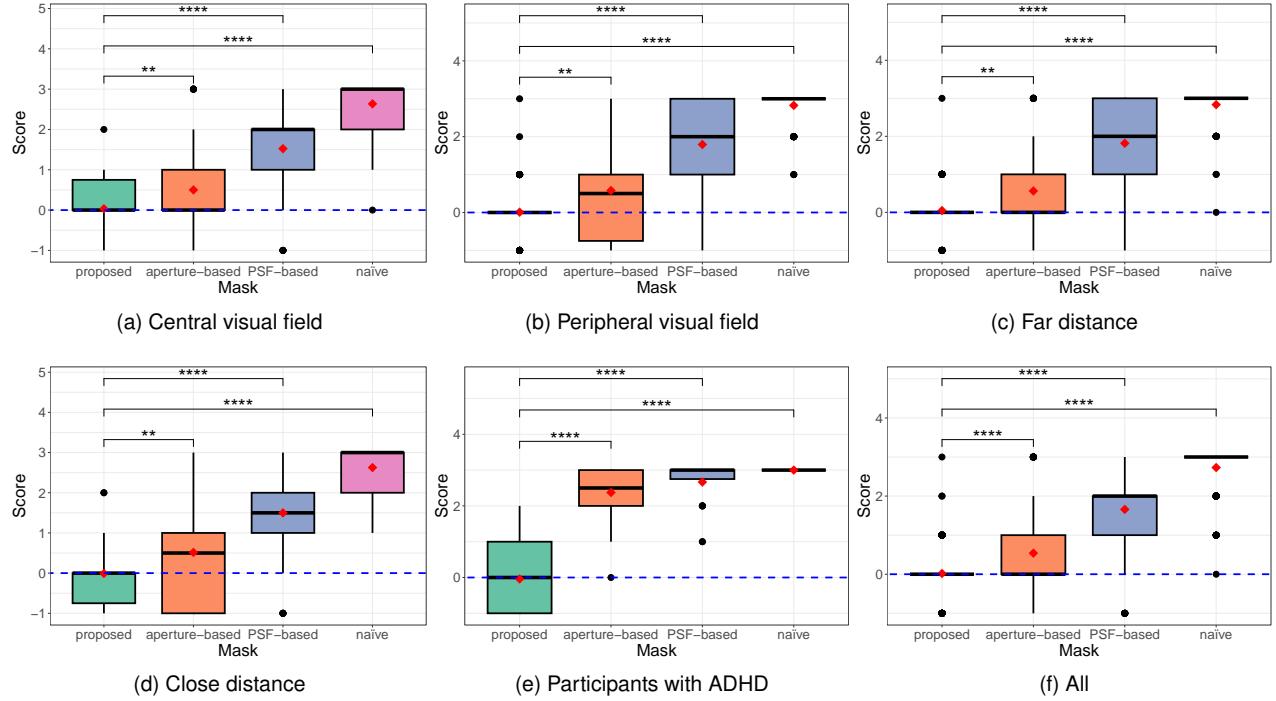


Fig. 10. Box plots showing users' subjective evaluations of the four occlusion masks. A score of 0 indicates complete occlusion, a positive score indicates the number of visible internal rings, and a score of  $-1$  indicates occlusion leakage. The red dots within the boxes represent the mean scores for each occlusion mask. Remarks: [significance: [\*\*\*\*]  $p < 10^{-4}$ , [\*\*\*]  $p < 0.001$ , and [\*\*]  $p < 0.01$ ].

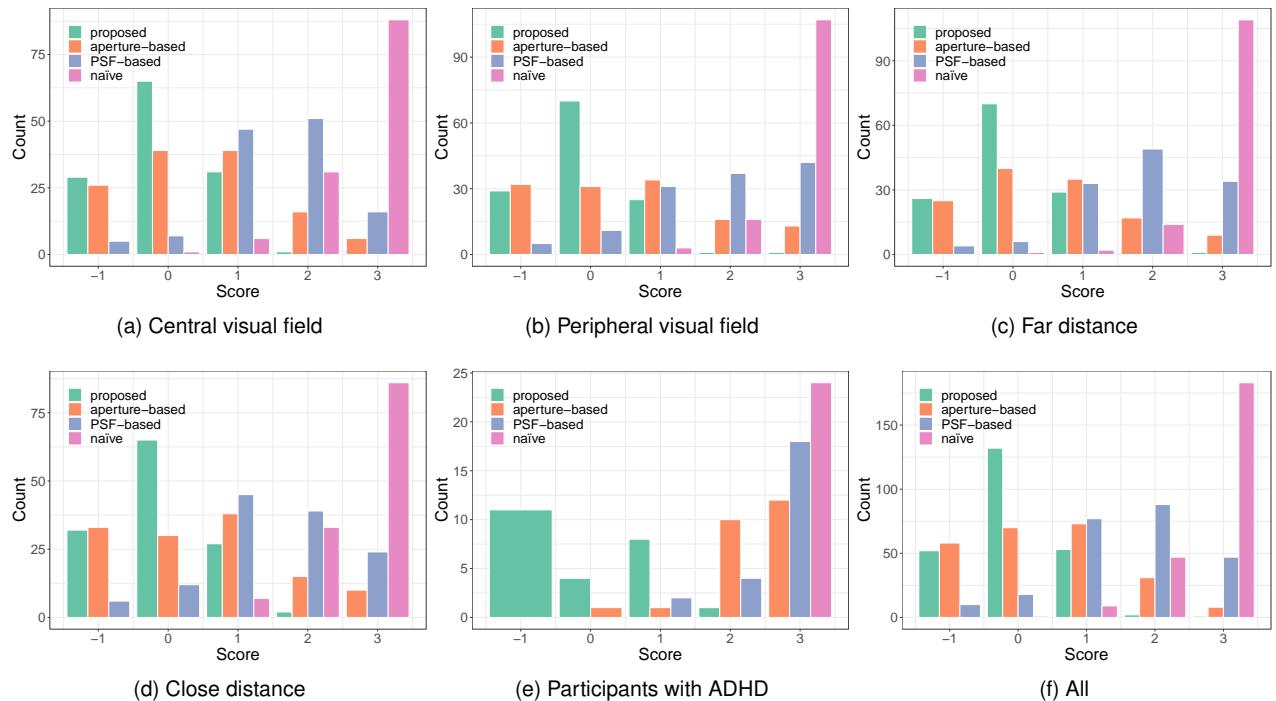


Fig. 11. Histograms showing users' subjective evaluations of the four occlusion masks. A score of 0 indicates complete occlusion, a positive score indicates the number of visible internal rings, and a score of  $-1$  indicates occlusion leakage. Scores of 1 to 3 represent the number of visible internal rings.

## 6 DISCUSSIONS

In this section, we delve into the idiosyncrasies of human visual perception in Section 6.1, including the potential reasons for users' enhanced ability to perceive sharper edge blur and the possible causes of individual differences observed in this experiment. Furthermore, we discuss the application of soft-edge occlusion in AR in Section 6.2 based on the results we have uncovered and the expansion strategy proposed. We recommend reading this section along with the Supplementary Material, where we explain in detail the geometric optics of our system and how human visual perception differs from geometric optics.

### 6.1 Human Visual Perception of Defocused Mask

#### 6.1.1 Sharper Edge Blur

Human vision's ability to recognize sharp edges is closely linked to the steepness of the image gradient [32]. While implementing the system, we observed that the edge blur of the mask appeared sharper when gazing at distant objects. This observation was further confirmed by participants' verbal feedback in Section 5.2, where they mentioned discerning a defocused occlusion mask with edges that appeared relatively sharper compared to the image captured by the camera using a similar aperture size. During the calibration and evaluation tasks, participants demonstrated the capability to discern regions with steep gradients, leading them to perceive those regions as sharp edges. The presence of these edges played a crucial role in enabling participants to perform the tasks smoothly, and the results achieved were statistically significant.

The results obtained from aperture-based masks calculated based on geometrical optics further validate this phenomenon. Itoh *et al.* previously observed occlusion leak, which occurs for enlarged masks calculated using this approach [12]. However, in our experiment, the median of participants' rating was above 0, and some participants even rated it as two or three rings visible (Fig. 11). This suggests that some participants were able to perceive an insufficient occlusion through the mask calculated theoretically.

It is established that human vision exhibits lower sensitivity in the peripheral field of view compared to the central field of view [33]. The central part of vision, known as the foveal vision, encompasses a viewing angle of around  $5^\circ$  [34]. This implies that for a viewer positioned at a distance of 30 mm from the LCD, a mask displayed on the LCD with an approximate radius of 1.3 mm would align with the highest visual acuity zone within this viewing angle. Consequently, the perception of blurred edges could potentially be challenging due to the limitations of peripheral vision in discerning fine details.

Additionally, the human eye has a higher threshold for discriminating blur when perceiving artificial images such as the Maltese cross [18]. In our experiment, the occlusion mask is also an artificial image and shares similarities with the Maltese cross, with clear separation between large light and dark color blocks. As a result, it can be more challenging for the human eye to distinguish between light and dark changes at the blurred borders of the occlusion mask. This discrepancy may lead to a perceived dark-to-light transition at the mask borders, which appears relatively sharp to the

human eye, while the camera system may capture wider blurred borders.

It is important to note that the human eye does not directly perceive sharp edges, but rather, it considers the least blur it observes as sharp [35]. Human vision incorporates a mechanism for adapting to blur through neural compensation for defocus conditions [36], [37]. This adaptive mechanism is particularly observed in myopic individuals and allows them to perceive a sharper view when they are not wearing glasses for an extended period. This adaptation to blur can also occur within short durations, even within three minutes [38]. As a result, we hypothesized that during our experiment, participants might have experienced a sharper observation due to this adaptation to blur.

From a geometric optics perspective, on one hand, the effect of partial occlusion [39] can lead to the visibility of objects. When an object is partially occluded by another object and the occluded object is in focus, it tends to remain visible to the observer. The presence of the occluder reduces the effective pupil size and affects the PSF of the optical system. This also explains the fact that in our experiments, participants' recognition of the pattern remained clear when observing the pattern that was partially occluded by the occlusion masks.

On the other hand, Charman's findings [40] highlight an important aspect of human vision that goes beyond geometric optics. While geometric optics may suggest that objects do not necessarily improve their contrast on the retina as they get closer, the actual contrast of the retinal image can be influenced by various factors, such as diffraction and aberration. These optical phenomena can lead to an increase in contrast, making details easier to perceive.

Recently, Olekiw *et al.* [41] demonstrated that V4 neurons in the primate visual cortex exhibit tuning to both the curvature of object boundaries and the degree of blur at these boundaries. Their findings suggest that the primate visual system, including humans, possesses specialized neural mechanisms to process blurred edges. This indicates that the perceptual experience of blur might not align perfectly with theoretical optical calculations.

In conclusion, the observed phenomenon of 'sharper edge blur' in our study can be explained through a range of indirect speculations. Factors such as the human vision's sensitivity to steep gradients, neural mechanisms for processing blurred edges, and optical effects like diffraction, aberration, and partial occlusion have been considered. However, it is important to note that further investigation to establish their direct correlation between these explanations and the observed phenomenon is necessary. The intricate interplay of physiological, neural, and optical aspects in shaping the perception of edge blur underscores the need for comprehensive future research to clarify the mechanisms underlying this intriguing phenomenon.

#### 6.1.2 Individual Differences

In Fig. 7, it is evident that there were significant differences in the linear models of the participants, and we suspect these differences were attributed to their contrast sensitivity. Contrast sensitivity refers to the ability of the human visual system to distinguish between variations in light intensity. Participants with lower contrast sensitivity were less likely

to perceive changes in the border of the mask due to changes in size, resulting in less adjustment in the calibration and thus lower slopes. We also noticed that most participants with lower slopes were in their 30s, while contrast sensitivity is known to be strongest at around the age of 20 years old [42]. At the same time, it should be noted that individuals with ADHD have an atypical vision that also affects their contrast sensitivity [43]. Based on verbal feedback from participants with ADHD (#1 and #4), they have reported difficulties in distinguishing details in low-light conditions, which could be attributed to their lower contrast sensitivity. Therefore, it is not surprising that the linear models of participants with ADHD had a very low slope.

The difference caused by contrast sensitivity could also be reflected in the results showing that the proposed method performs better in the periphery. This is probably because the human eye has lower contrast sensitivity in the peripheral vision compared to the central vision [44]. As a result, the human eye is more forgiving of peripheral masking and it is easier for the participants to perceive a complete mask.

However, it should be noted that our available experimental data did not directly model the participants' contrast sensitivity function. Therefore, the hypotheses regarding contrast sensitivity presented in this study are based on inferences drawn from previous research. These inferences suggest that contrast sensitivity tends to be lower in the peripheral visual field, among older individuals and in people with ADHD. To provide more conclusive evidence, further exploration and investigation are required.

We also noticed that even when the linear model was calibrated based entirely on the user's preferences, there were still errors introduced in observing the complete occlusion at the same pupil size, especially for the same position of the pattern, as when it was calibrated. We suspect that these errors are not only due to inaccuracies in the fit of the linear model, but also because during calibration the participants were focusing on the pattern the entire time, whereas during the evaluation the user only briefly glances at the pattern. A recent study has shown that perceptual processes are different when the human eye processes information briefly compared to when it is fixated on an object, and that these two perceptual processes are not integrated [45]. This error was more pronounced in the results of the participants with ADHD.

## 6.2 Utility of Soft-edge Occlusion

Occlusion capability is pivotal in OST-HMDs, significantly enhancing visual performance by rendering semi-transparent virtual objects with a more vivid and solid appearance. This feature effectively reduces visual interference from background color blending and enriches the depth perception, alongside providing benefits such as light attenuation [11], gaze guidance [22], and visual noise reduction [46]. Additionally, the ability to overlay new virtual content onto occluded areas facilitates texture replacement, diminished reality, and the depiction of translucent objects [23]. However, Magic Leap 2 is currently the only commercial product that incorporates this feature, primarily due to the complexity of the optical structure required for occlusion, which adds significant bulk to OST-HMDs.

Despite this, innovations by Zhang *et al.* [23] have led to the development of an add-on module for hard-edge occlusion compatible with HoloLens 1, though its usability is curtailed by a narrow FOV and considerable size and weight.

In contrast, soft-edge occlusion implemented with transmissive SLMs avoids these disadvantages. It offers a lightweight, easily controllable solution with a wider FOV. One drawback of using soft-edge occlusion in OST-HMDs is that it does not block incident light as precisely as hard-edge occlusion, aligning with the virtual contents. Our experiments have demonstrated that the soft-edge occlusion mask is perceived as a relatively sharp border in human vision. Through appropriate expansion, a satisfactory masking effect, akin to the capabilities of hard-edge occlusion, can be achieved, highlighting the potential of transmissive SLMs in making occlusion-capable OST-HMDs more accessible and practical for a wider range of applications.

Nevertheless, since the common device to achieve soft-edge occlusion is the transmissive LCD [1], [6], [10], [12], significant limitations of the transmissive LCD include its low transmittance, diffraction effects, and spatial resolution constraints. The low light transmittance is attributed to two main factors: the inherent light intensity loss during modulation by the LC material, and further reduction due to the incorporation of front and rear polarizers. The observed diffraction effects are a consequence of dividing the LCs into discrete cells to establish pixels. The unavoidable spaces between these LC cells lead to diffraction. Notably, when these LC cells are inherently large, they yield a significantly lower spatial resolution, which in turn leads to the manifestation of aliasing artifacts.

To overcome the inherent limitations of transmissive LCDs, numerous researchers have shifted towards reflective SLMs, such as digital micromirror device (DMD) [47] and liquid crystal on silicon (LCoS) [7], [48], [49], for implementing occlusion capabilities, typically utilizing hard-edge occlusion. This shift markedly diminishes diffraction effects and mitigates light loss. However, the adoption of reflective SLMs for occlusion purposes often necessitates a lens system to accurately direct the light. This addition, in turn, contributes to increased volume and weight of the device, and compromises the FOV, introducing distortion.

Meanwhile, Magic Leap 2 innovates by transitioning from traditional square pixel cells to curved ones, aiming to reduce diffraction spikes [50]. However, while this design change mitigates diffraction to some extent, it does not fully resolve the inherent limitations associated with the liquid crystal material, including those related to light transmittance and the diffraction effect itself.

Recently, Chae *et al.* introduced a novel occlusion-capable see-through display by employing photochromic materials [51]. Building on this, Ooi *et al.* developed a compact HMD that incorporates soft-edge occlusion using photochromic materials as well [24]. Utilizing these materials, both projects effectively addressed the issues of low transmittance, diffraction effects, and spatial resolution constraints. While it is worth noting that these approaches require UV light irradiation to activate the occlusion mask display and involve a longer delay in mask appearance and disappearance, it presents a promising solution for overcoming existing challenges.

Our experimental findings provide a strong basis for the feasibility of soft-edge occlusion masks, despite the limitations observed with the transmissive LCDs used. We remain confident that future advancements, including the potential replacement with photochromic materials, will lead to notable improvements in display performance.

## 7 LIMITATIONS

### 7.1 Effect of Saccade on Static Mask

We chose not to capture the real-time pupil position using eye tracking and display the dynamic mask to prevent any errors and delays caused by the eye tracker that may affect the participant's calibration and evaluation. Moreover, dynamic occlusion can also affect the participant's accurate judgment of mask border in calibration. However, we observed that participants produced severe saccades during the experiment while the static mask was being displayed. The saccade is a rapid, conjugate eye movement that shifts the center of gaze from one part of the visual field to another, and it is uncontrollable [52]. We calculated the dispersion of all participants for given duration windows (80–220 ms) based on the dispersion-based algorithms [53] provided by Pupil Labs, with a mean of 0.43° and a standard deviation of 0.62°. These frequent saccades cause a shift in the position of the mask, which can lead to an expansion or reduction in the effective size of the mask.

### 7.2 Errors in Eye Tracking

In our experiments, we had to position the eye tracker outside the LCD display area to avoid obstructing the participants' visual field. However, due to the limited FOV of the camera, we encountered issues with the accuracy of the 3D model of the eyeball, resulting in incorrect pupil size estimation. The aperture-based and PSF-based methods are highly dependent on precise pupil size estimation, which underscores the advantages of our proposed method that only requires the relative relationship between pupil size and expansion radius. Despite fixing the participants' head position to minimize errors, there is still a possibility of anomalous 3D models of the eyeball due to prolonged eye closure during the experiment, which can result in errors for all three methods. Although we rely on re-experimentation to avoid errors in our experiments, the accuracy of eye tracking remains crucial for daily use. Future research may consider addressing these limitations to improve the accuracy and reliability of each method.

Additionally, we found that even under identical lighting conditions, the pupil size of the same participant remained uncontrollable. This is because pupil size is influenced by a variety of factors beyond ambient illumination, including the luminance level of the focused surface [54], eye fatigue [55], and the cognitive effort required for control tasks [56]. During the experiment, we observed that participants' pupils exhibited noticeable fluctuations in size even under constant illumination. To account for this, we conducted the calibration and evaluation under varying illumination to cover a wider range of pupil sizes. Nonetheless, it is possible that fluctuations in pupil size and the corresponding adjustments to the mask size had some effect

on the reported results. In the future, faster mask adaptation, such as tracking the eye with an event camera [57], [58], and shorter tasks where the pupil size remains constant could provide further insights.

## 8 CONCLUSION

We proposed a method to achieve optimal masking effects independently of the distance to the object to be covered and the user's pupil size. This method relies on soft-edge occlusion developed on a single transparent LCD. We evaluated the performance of these occlusion masks and calibrated them according to perceived occlusion. The results suggest that human blur perception is mild when compared to images captured with a camera. Finally, this study demonstrated the potential of soft-edge occlusion masks to achieve optimal masking effects.

## ACKNOWLEDGMENTS

This work was supported by JSPS KAKENHI (Grant Number 22H00539), JST SPRING (Grant Number JPMJSP2140), Japan, and Snap Inc. We would like to thank Michael Anderson for assisting with grammar checking and language polishing of this manuscript.

## REFERENCES

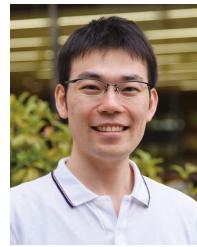
- [1] Y. Hiroi, Y. Itoh, T. Hamasaki, and M. Sugimoto, "Adaptivisor: Assisting eye adaptation via occlusive optical see-through head-mounted displays," in *Proceedings of the 8th Augmented Human International Conference*, 2017, pp. 1–9.
- [2] X. Hu, Y. Zhang, N. Isoyama, N. Sakata, and K. Kiyokawa, "Design and prototyping of computational sunglasses for autism spectrum disorders," in *2021 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW)*. IEEE, 2021, pp. 581–582.
- [3] S. K. Nayar and V. Branzoi, "Adaptive dynamic range imaging: Optical control of pixel exposures over space and time," in *Proceedings of the IEEE International Conference on Computer Vision*, vol. 2, 2003, pp. 1168–1175.
- [4] G. Wetzstein, W. Heidrich, and D. Luebke, "Optical image processing using light modulation displays," *Computer Graphics Forum*, vol. 29, no. 6, pp. 1934–1944, 2010.
- [5] K. Kiyokawa, Y. Kurata, and H. Ohno, "An optical see-through display for mutual occlusion of real and virtual environments," in *Proceedings IEEE and ACM International Symposium on Augmented Reality (ISAR 2000)*. IEEE, 2000, pp. 60–67.
- [6] A. Maimone and H. Fuchs, "Computational augmented reality eyeglasses," in *2013 IEEE International Symposium on Mixed and Augmented Reality, ISMAR 2013*, 2013, pp. 29–38.
- [7] A. Wilson and H. Hua, "Design of a pupil-matched occlusion-capable optical see-through wearable display," *IEEE Transactions on Visualization and Computer Graphics*, vol. 28, no. 12, pp. 4113–4126, 2021.
- [8] R. R. Hainich and O. Bimber, *Displays: fundamentals & applications*. CRC press, 2016.
- [9] Y. Zhang, X. Hu, K. Kiyokawa, N. Isoyama, H. Uchiyama, and H. Hua, "Realizing mutual occlusion in a wide field-of-view for optical see-through augmented reality displays based on a paired-ellipsoidal-mirror structure," *Optics Express*, vol. 29, no. 26, pp. 42 751–42 761, 2021.
- [10] Q. Zhang, G. Barbareschi, Y. Huang, J. Li, Y. S. Pai, J. Ward, and K. Kunze, "Seeing our blind spots: Smart glasses-based simulation to increase design students' awareness of visual impairment," in *Proceedings of the 35th Annual ACM Symposium on User Interface Software and Technology*, 2022, pp. 1–14.
- [11] X. Hu, Y. Zhang, N. Isoyama, H. Uchiyama, N. Sakata, and K. Kiyokawa, "Smart dimming sunglasses for photophobia using spatial light modulator," *arXiv preprint arXiv:2304.07013*, 2023.

- [12] Y. Itoh, T. Hamasaki, and M. Sugimoto, "Occlusion Leak Compensation for Optical See-Through Displays Using a Single-Layer Transmissive Spatial Light Modulator," *IEEE Transactions on Visualization and Computer Graphics*, vol. 23, no. 11, pp. 2463–2473, 2017.
- [13] E. W. Tatham, "Optical occlusion and shadows in a 'see-through' augmented reality display," in *1999 IEEE International Conference on Information Visualization (Cat. No. PR00210)*. IEEE, 1999, pp. 128–131.
- [14] J. W. Goodman, *Introduction to Fourier optics*. Roberts and Company publishers, 2005.
- [15] B. C. Kress, "Optical architectures for augmented-, virtual-, and mixed-reality headsets," (*No Title*), 2020.
- [16] K. R. Curtis, "Unveiling magic leap 2's advanced ar platform and revolutionary optics," in *SPIE AR, VR, MR Industry Talks 2022*, vol. 11932. SPIE, 2022, p. 119320P.
- [17] H. Mannami, R. Sagawa, Y. Mukaigawa, and T. YasushiYagi, "Wide dynamic range by filtering with transmissive liquid crystal," *Transactions of the Society of Instrument and Control Engineers*, vol. 41, pp. 1026–1035, 2005.
- [18] S. Sebastian, J. Burge, and W. S. Geisler, "Defocus blur discrimination in natural images with natural optics," *Journal of Vision*, vol. 15, no. 5, pp. 16–16, 2015.
- [19] A. B. Watson and A. J. Ahumada, "Blur clarified: A review and synthesis of blur discrimination," *Journal of vision*, vol. 11, no. 5, pp. 10–10, 2011.
- [20] A. Maimone, D. Lanman, K. Rathinavel, K. Keller, D. Luebke, and H. Fuchs, "Pinlight displays: wide field of view augmented reality eyeglasses using defocused point light sources," in *ACM SIGGRAPH 2014 emerging technologies*, 2014, pp. 1–1.
- [21] Q. Zhang, H. Yamamura, H. Baldauf, D. Zheng, K. Chen, J. Yamaka, and K. Kunze, "Tunnel vision-dynamic peripheral vision blocking glasses for reducing motion sickness symptoms," in *Proceedings of the 2021 ACM International Symposium on Wearable Computers*, 2021, pp. 48–52.
- [22] Q. Zhang, Y. Huang, G. Chernyshov, J. Li, Y. S. Pai, and K. Kunze, "Gazesync: Eye movement transfer using an optical eye tracker and monochrome liquid crystal displays," in *27th International Conference on Intelligent User Interfaces*, 2022, pp. 54–57.
- [23] Y. Zhang, X. Hu, K. Kiyokawa, and X. Yang, "Add-on occlusion: Turning off-the-shelf optical see-through head-mounted displays occlusion-capable," *IEEE Transactions on Visualization and Computer Graphics*, 2023.
- [24] C.-W. Ooi, Y. Hiroi, and Y. Itoh, "A compact photochromic occlusion capable see-through display with holographic lenses," in *2023 IEEE Conference Virtual Reality and 3D User Interfaces (VR)*. IEEE, 2023, pp. 237–242.
- [25] E. Kruijff, J. E. Swan, and S. Feiner, "Perceptual issues in augmented reality revisited," in *2010 IEEE International Symposium on Mixed and Augmented Reality*. IEEE, 2010, pp. 3–12.
- [26] D. H. Baker, "What is the primary cause of individual differences in contrast sensitivity?" *PLoS one*, vol. 8, no. 7, p. e69536, 2013.
- [27] L. Grzeczkowski, A. M. Clarke, G. Francis, F. W. Mast, and M. H. Herzog, "About individual differences in vision," *Vision research*, vol. 141, pp. 282–292, 2017.
- [28] K. Oshima, K. R. Moser, D. C. Rompapas, J. E. Swan, S. Ikeda, G. Yamamoto, T. Taketomi, C. Sandor, and H. Kato, "Sharpview: Improved clarity of defocused content on optical see-through head-mounted displays," in *2016 IEEE Symposium on 3D User Interfaces (3DUI)*. IEEE, 2016, pp. 173–181.
- [29] Y. Itoh and G. Klinker, "Vision enhancement: defocus correction via optical see-through head-mounted displays," in *Proceedings of the 6th Augmented Human International Conference*, 2015, pp. 1–8.
- [30] D. C. Burr and M. Morgan, "Motion deblurring in human vision," *Proceedings of the royal society of London. Series B: Biological Sciences*, vol. 264, no. 1380, pp. 431–436, 1997.
- [31] H. Moseley, "Ultraviolet and visible radiation transmission properties of some types of protective eyewear," *Physics in Medicine & Biology*, vol. 30, no. 2, p. 177, 1985.
- [32] H. Von Helmholtz, *Treatise on Physiological Optics, volume III*. Courier Corporation, 2013, vol. 3.
- [33] E. E. Stewart, M. Valsecchi, and A. C. Schütz, "A review of interactions between peripheral and foveal vision," *Journal of vision*, vol. 20, no. 12, pp. 2–2, 2020.
- [34] M. Millodot, *Dictionary of optometry and visual science*. Elsevier Health Sciences, 2014.
- [35] R. Watt, "An outline of the primal sketch in human vision," *Pattern recognition letters*, vol. 5, no. 2, pp. 139–150, 1987.
- [36] M. Mon-Williams, J. R. Tresilian, N. C. Strang, P. Kochhar, and J. P. Wann, "Improving vision: neural compensation for optical defocus," *Proceedings of the Royal Society of London. Series B: Biological Sciences*, vol. 265, no. 1390, pp. 71–77, 1998.
- [37] M. P. Cufflin and E. A. Mallen, "Blur adaptation: clinical and refractive considerations," *Clinical and Experimental Optometry*, vol. 103, no. 1, pp. 104–111, 2020.
- [38] M. A. Webster, M. A. Georgeson, and S. M. Webster, "Neural adjustments to image blur," *Nature neuroscience*, vol. 5, no. 9, pp. 839–840, 2002.
- [39] M. Zannoli, G. D. Love, R. Narain, and M. S. Banks, "Blur and the perception of depth at occlusions," *Journal of Vision*, vol. 16, no. 6, pp. 1–25, 2016.
- [40] W. Charman, "Near vision, lags of accommodation and myopia," *Ophthalmic and Physiological Optics*, vol. 19, no. 2, pp. 126–133, 1999.
- [41] T. D. Oleskiw, A. Nowack, and A. Pasupathy, "Joint coding of shape and blur in area v4," *Nature communications*, vol. 9, no. 1, p. 466, 2018.
- [42] C. Owsley, R. Sekuler, and D. Siemsen, "Contrast sensitivity throughout adulthood," *Vision research*, vol. 23, no. 7, pp. 689–699, 1983.
- [43] A. Bellato, J. Perna, P. S. Ganapathy, M. Solmi, A. Zampieri, S. Cortese, and S. V. Faraone, "Association between adhd and vision problems. a systematic review and meta-analysis," *Molecular Psychiatry*, vol. 28, no. 1, pp. 410–422, 2023.
- [44] B. L. LUNDH, G. LENNERSTRAND, and G. DEREFELDT, "Central and peripheral normal contrast sensitivity for static and dynamic sinusoidal gratings," *Acta Ophthalmologica*, vol. 61, no. 2, pp. 171–182, 1983.
- [45] Naoshige Uchida, Adam Kepecs, and Zachary F. Mainen, "Seeing at a glance, smelling in a whiff: rapid forms of perceptual decision making," *Nature Reviews Neuroscience*, vol. 7, no. June, pp. 485–491, 2006.
- [46] J. Hong, T. Langlotz, J. Sutton, and H. Regenbrecht, "Visual noise cancellation: Exploring visual discomfort and opportunities for vision augmentations," *ACM Transactions on Computer-Human Interaction*, vol. 31, no. 2, pp. 1–26, 2024.
- [47] Y.-G. Ju, M.-H. Choi, P. Liu, B. Hellman, T. L. Lee, Y. Takashima, and J.-H. Park, "Occlusion-capable optical-see-through near-eye display using a single digital micromirror device," *Optics letters*, vol. 45, no. 13, pp. 3361–3364, 2020.
- [48] A. Wilson and H. Hua, "Design and prototype of an augmented reality display with per-pixel mutual occlusion capability," *Optics express*, vol. 25, no. 24, pp. 30 539–30 549, 2017.
- [49] Y. Zhang, X. Hu, K. Kiyokawa, N. Isoyama, N. Sakata, and H. Hua, "Optical see-through augmented reality displays with wide field of view and hard-edge occlusion by using paired conical reflectors," *Optics letters*, vol. 46, no. 17, pp. 4208–4211, 2021.
- [50] A. I. Russell, J. I. Trisnadi, V. Mathur, M. R. Johnson, and C. Carlisle, "Geometries for mitigating artifacts in see-through pixel arrays," U.S. Patent 00,038,72 A1, Jan. 7, 2021.
- [51] M. Chae, K. Bang, Y. Jo, C. Yoo, and B. Lee, "Occlusion-capable see-through display without the screen-door effect using a photochromic mask," *Optics Letters*, vol. 46, no. 18, pp. 4554–4557, 2021.
- [52] B. Cassin, M. L. Rubin, and S. Solomon, *Dictionary of eye terminology*. Triad Publishing Company Gainsville, 1984, vol. 10.
- [53] D. D. Salvucci and J. H. Goldberg, "Identifying fixations and saccades in eye-tracking protocols," in *Proceedings of the 2000 symposium on Eye tracking research & applications*, 2000, pp. 71–78.
- [54] P. Bindu, M. Pereverzeva, and S. O. Murray, "Pupil size reflects the focus of feature-based attention," *Journal of neurophysiology*, vol. 112, no. 12, pp. 3046–3052, 2014.
- [55] R. Zargari Marandi, P. Madeleine, Ø. Omland, N. Vuillerme, and A. Samani, "Eye movement characteristics reflected fatigue development in both young and elderly individuals," *Scientific reports*, vol. 8, no. 1, p. 13148, 2018.
- [56] P. van der Wel and H. Van Steenbergen, "Pupil dilation as an index of effort in cognitive control tasks: A review," *Psychonomic bulletin & review*, vol. 25, pp. 2005–2015, 2018.
- [57] A. N. Angelopoulos, J. N. Martel, A. P. Kohli, J. Conradt, and G. Wetzstein, "Event based, near-eye gaze tracking beyond 10,000 hz," *IEEE Transactions on Visualization and Computer Graphics (Proc. VR)*, 2021.

- [58] T. Zhang, Y. Shen, G. Zhao, L. Wang, X. Chen, L. Bai, and Y. Zhou, "Swift-eye: Towards anti-blink pupil tracking for precise and robust high-frequency near-eye movement analysis with event cameras," *IEEE Transactions on Visualization and Computer Graphics*, 2024.
- [59] S. Mori, S. Ikeda, A. Plopski, and C. Sandor, "Brightview: Increasing perceived brightness of optical see-through head-mounted displays through unnoticeable incident light reduction," in *2018 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*. IEEE, 2018, pp. 251–258.
- [60] M. S. Arefin, N. Phillips, A. Plopski, J. L. Gabbard, and J. E. S. II, "The effect of context switching, focal switching distance, binocular and monocular viewing, and transient focal blur on human performance in optical see-through augmented reality," *IEEE Transactions on Visualization and Computer Graphics*, vol. PP, pp. 1–1, 2022.
- [61] B. Wang and K. J. Ciuffreda, "Foveal blur discrimination of the human eye," *Ophthalmic and Physiological Optics*, vol. 25, no. 1, pp. 45–51, 2005.



**Xiaodan Hu** (Member, IEEE) is currently a postdoctoral researcher at TU Graz. She received her MSc (2021) and Ph.D. (2024) from Nara Institute of Science and Technology, Japan. Her research interests include occlusion-capable OST-HMDs, vision augmentation, human visual perception, and near-eye displays.



**Yuta Itoh** (Member, IEEE) is a Project Associate Professor at the University of Tokyo, Japan. His research interest is in vision augmentation which aims to support and enhance human vision via augmented reality technology including see-through near-eye displays. Before joining the university, he was an Assistant Professor at the Tokyo Institute of Technology (2017-2021). He also worked as a Project Assistant Professor at Keio University (2016-2017). He received his Dr. rer. nat. from the Technical University of Munich in 2016. He also spent two years as a researcher at Multimedia Lab. in Toshiba Corp. (2011-2013).



**Monica Perusquia-Hernández** received her BSc in electronic systems engineering (2009) from the Instituto Tecnológico y de Estudios Superiores de Monterrey, Mexico; her MSc in human-technology interaction (2012) and the professional doctorate in engineering in user-system interaction (2014) from the Eindhoven University of Technology, the Netherlands. In 2018, she obtained her Ph.D. in Human Informatics from the University of Tsukuba, Japan. She is an Assistant Professor at the Nara Institute of Science and Technology. Her research interests include affective computing, biosignal processing, augmented human technology, and artificial intelligence.



**Naoya Isoyama** received his Ph.D. in engineering from Kobe University in 2015. After working at Aoyama Gakuin University, Kobe University, and Nara Institute of Science and Technology, he is a lecturer at Otsuma Women's University. His research interests include wearable computing, ubiquitous computing, and entertainment computing.



**Hideaki Uchiyama** received his Ph.D. in engineering from Keio University in 2010. After working at INRIA, Toshiba, and Kyushu University, he is an Associate Professor at Nara Institute of Science and Technology. His research interests include computer vision, inertial sensing, and their applications.



**Kiyoshi Kiyokawa** (Member, IEEE) received his Ph.D. degree in information systems from Nara Institute of Science and Technology (NAIST) in 1998. After working at the National Institute of Information and Communications Technology (NICT) and Osaka University, he is currently a Professor at NAIST, leading the Cybernetics and Reality Engineering Laboratory. His research interests include virtual reality, augmented reality, human augmentation, 3D user interfaces, CSCW, and context awareness. He is also a Fellow of the Virtual Reality Society of Japan. He is an associate editor-in-chief of IEEE TVCG and an inductee of the IEEE VGTC Virtual Reality Academy (Inaugural Class).



**Alexander Plopski** received his Ph.D. in information science and technology from Osaka University, Japan, in 2016. After working at the Nara Institute of Science and Technology and the University of Otago, he joined TU Graz in 2022 as an Assistant Professor. His research interests include computer vision, eye tracking, and augmented, mixed, and virtual reality.