Pinhole Occlusion: Enhancing Soft-Edge Occlusion Using a Dynamic Pinhole Array

Xiaodan Hu*
NAIST

Yan Zhang†
Shanghai Jiao Tong University

Monica Perusquía-Hernández‡
NAIST

Yutaro Hirao§
NAIST

Hideaki Uchiyama¶
NAIST

Kiyoshi Kiyokawa‖
NAIST

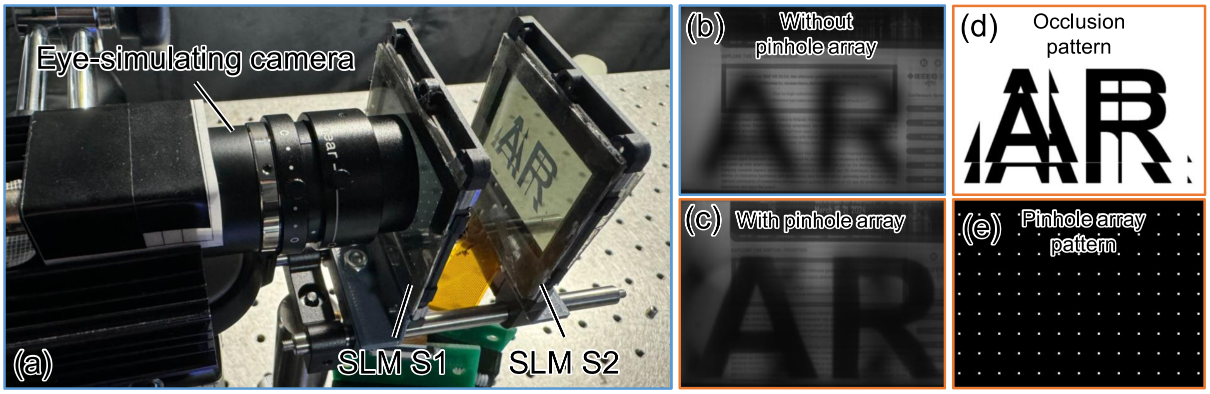

Figure 1: (a) The prototype setup for pinhole occlusion, featuring a spatial light modulator (SLM) S1 to render the pinhole array, SLM S2 for displaying the occlusion pattern, alongside a camera that simulates the human eye. (b) The result of a defocused observation with only the occlusion pattern displayed on S2 and the pinhole array on S1 deactivated. (c) A focused observation achieved by rendering the pinhole array on S1 in conjunction with the occlusion pattern on S2. (d) The occlusion pattern as displayed on S2. (e) The pinhole array pattern as rendered on S1.

Abstract

Systems with occlusion capabilities have gained interest in augmented reality, vision augmentation, and image processing. To address the challenge of creating a precise yet lightweight occlusion system, we introduce a novel architecture to tackle occlusion blurriness due to defocusing. Our approach, utilizing a dynamic pinhole array on a transmissive spatial light modulator (SLM) positioned between the eye and the occlusion layer, offers adaptive pinhole patterns, gaze-contingent functionality, and the potential to reduce visual artifacts. Our preliminary result demonstrates that, with the focal plane at 1.8 m, an occlusion placed at 4 cm can be observed sharply through a 4.3 mm aperture.

Index Terms: Human-centered computing—Human computer interaction (HCI)—Interaction devices—Displays and imagers; Human-centered computing—Human computer interaction (HCI)—Interaction paradigms—Mixed / augmented reality

1 Introduction

In the realms of augmented reality (AR) [3, 4, 6], image processing [5], and vision augmentation [2], occlusion — the selective blocking or alteration of visual elements — is crucial for realistic AR experiences and the enhancement of imaging and visual effects. Occlusion is typically achieved by modulating the intensity of incident light via spatial light modulators (SLMs), with the modulated areas known as the occlusion pattern or mask.

Occlusion systems are typically divided into two types: hard-edge occlusion, which is defined by precise pixel-to-pixel blocking, and soft-edge occlusion, characterized by a blurred occlusion mask due to out-of-focus effects. Hard-edge occlusion systems, noted for their precision, tend to be bulky due to complex optical designs incorporating multiple lenses, resulting in a limited field of view (FOV) [4, 6]. In contrast, soft-edge occlusion systems, which forgo the optical lenses, provide a wider FOV and a more compact form, but at the cost of blurred occlusion masks due to the lack of focused lenses [2, 3].

For the purpose of achieving both precise occlusion blocking and a lightweight, usable design, Zhang et al. developed a hard-edge occlusion system with a folded optical path [6]. This innovation led to an add-on module for standard AR optical see-through head-mounted displays (OST-HMDs). However, the use of 25 mm diameter lenses in this device limits the FOV. Itoh et al. addressed soft-edge occlusion blurriness by overlaying a virtual image onto the blurry mask [3]. However, this method's reliance on integration with a virtual image projector limits its broader application.

In this study, we introduce a novel setup designed to address the blurriness issue commonly associated with soft-edge occlusion. In our approach, the occlusion mask is displayed on a transmissive SLM, and a pinhole array, situated between the occlusion mask and the human eye, is used to extend the depth of field (as shown in Fig. 1(a)). While pinhole arrays have been effectively used in building light field displays [1], their application in occlusion has not yet been explored. Distinctively, our approach does not rely on a static, physical pinhole array; instead, we dynamically render the pinhole array on another transmissive SLM. This dynamic implementation

*e-mail: hu.xiaodan.ht1@is.naist.jp
†e-mail: yan-zh@sjtu.edu.cn
‡e-mail: m.perusquia@is.naist.jp
§e-mail: yutaro.hirao@is.naist.jp
¶e-mail: hideaki.uchiyama@is.naist.jp
‖e-mail: kiyo@is.naist.jp

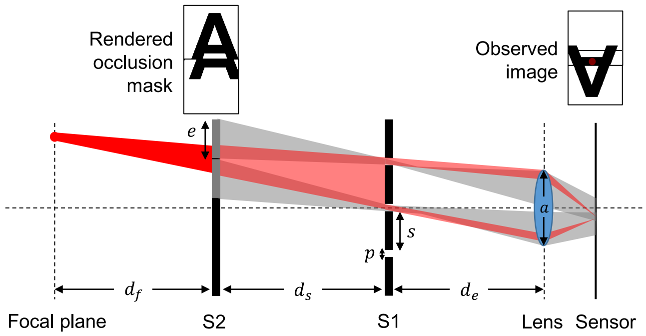

Figure 2: Optical Layout. Illustrating the light path from the rendered occlusion mask through the pinhole array to the sensor.

enables greater flexibility for the entire system, with the benefits of this approach discussed in more detail in Sect. 3.

2 System Layout and Prototyping

The system architecture, illustrated in Fig. 2, is designed to enhance soft-edge occlusion. We detail the implementation of this technique and note its adaptability for OST-HMDs with the addition of a virtual image projection system. As depicted, a red dot on the focal plane symbolizes a point light source, which produces a sharp image on the image sensor. Without SLM S1, responsible for rendering the pinhole array, an occlusion mask on SLM S2 dims the source light but appears blurred on the sensor due to S2's placement off the focal plane [3]. However, when a pinhole array is positioned between S2 and the observer, the point light source still projects a sharp image on the image sensor. Concurrently, light from the mask, now constrained by the pinholes, is directed into a narrow cone of light, resulting in a clear image on the image sensor.

However, the pinholes do more than merely focusing the light into narrower cones; they also separate the rays. This is depicted in Fig. 2, where the occlusion mask on SLM S2, shown in gray, is divided by two pinholes into two beams that then converge on the image sensor. Such separation can cause observers to see multiple, overlapping images. To address this, akin to most light field displays utilizing pinhole arrays, our system transforms the occlusion mask into elemental images, thus preventing the overlapping effect.

Once the parametric design of the entire system has been finalized, we can proceed to generate elemental images. The prototype used for our preliminary experiments is depicted in Fig. 1(a). For the transmissive SLMs S1 and S2, we employ two Sony LCX017 panels (1024 × 768 pixels, 60 Hz, monochrome). To simulate the human eye, a camera equipped with a Sony IMX250 sensor (2448 × 2048 pixels, maximum frame rate of 74.1 fps) and a 6 mm f/1.4 lens is utilized. The aperture size a is set to 4.3 mm, which corresponds to a reasonable pupil size and the maximum aperture of the 6 mm lens. S1 is positioned against the lens to replicate the distance between the eyes and the system, akin to wearing eyewear. The distance between two SLMs d_s is 35 mm. Based on the Rayleigh criterion, the diameter of the pinhole p can be calculated by $p \approx \sqrt{2\lambda d_s}$ [1]. In practice, referring to Akşit et al.'s prototype [1], we set $p = 9$ pixels (pixel pitch of 36 µm), equivalent to 0.324 mm. Using the similarity triangle principle, the pitch distance between pinholes s and the length of the elemental image e can be calculated as:

$$s = \frac{a \cdot d_s}{d_e + d_s}, \qquad e = \frac{(a+p) \cdot d_s}{d_e} + p, \tag{1}$$

where d_e is the distance between the S1 and the lens, approximately equal to 21.6 mm.

Following Akşit et al.'s method [1], we generated elemental images of the occlusion mask, as shown in Fig. 1(d). After accurate parameter computation, each elemental image aligns with its corresponding pinhole, ensuring precise overlap on the sensor and forming a sharp composite image, depicted in Fig. 1(c). Content on the focal plane at a distance, d_f, of 1.8 m remains sharp. In contrast, without the pinhole array, the occlusion mask on S2 appears blurred, as shown in Fig. 1(b).

3 Versatility and Future Potential

Adaptive pitch length Equation 1 shows that pupil size impacts pitch length. However, in typical pinhole array-based light field displays with static panels, the pinhole patterns are fixed, often set to the largest human pupil diameter, around 8 mm [1], which may compromise visual quality. A dynamic pinhole array, in contrast, adjusts the pitch length according to the real-time pupil size, offering a tailored visual experience.

Gaze-contingent feature Taking into account the head and eye movements typical in real-world scenarios, the shifting position of the pupil could result in the pitch continually obstructing light entry. To counteract this, we fixate a pinhole directly in front of the pupil and integrate eye-tracking technology, enabling the entire pinhole array to move in sync with the pupil.

Occlusion-based pinhole array The presence of the pinhole array significantly reduces the overall transmittance of the device. The use of dynamic pinhole arrays enables the selective rendering of pinhole arrays specifically for the occlusion mask. This approach allows for minimal impact on the brightness of the scene, ensuring that the scene's illumination is preserved as much as possible.

Various pinhole array patterns As shown in Fig. 1(c), the observed occlusion mask exhibits pronounced mesh stripes. This effect results from the pitch being out of focus and imaged on the image sensor. By altering the pattern of the pinhole array — such as exploring optimal arrangements, modifying pitch rendering, and adjusting the contrast of the pinhole array — we anticipate that this effect can be significantly reduced. These aspects represent key areas for our future research and exploration.

4 Conclusion

We have introduced an innovative optical setup that advances soft-edge occlusion by dynamically rendering a pinhole array on a transmissive SLM. This approach contrasts with traditional static pinhole systems, offering more flexibility and effectiveness in occlusion optimization. Our system offers a promising basis for future advancements in optimizing soft-edge occlusion, highlighting potential areas for further exploration and improvement.

Acknowledgments

This work was supported by JSPS KAKENHI 22H00539 and JST SPRING Grant Number JPMJSP2140, Japan.

References

[1] K. Akşit, J. Kautz, and D. Luebke. Slim near-eye display using pinhole aperture arrays. *Applied optics*, 54(11):3422–3427, 2015.

[2] X. Hu, Y. Zhang, H. Uchiyama, N. Isoyama, N. Sakata, and K. Kiyokawa. Smart dimming sunglasses for photophobia using spatial light modulator. *Displays*, 81:102611, 2024.

[3] Y. Itoh, T. Hamasaki, and M. Sugimoto. Occlusion leak compensation for optical see-through displays using a single-layer transmissive spatial light modulator. *TVCG*, 23(11):2463–2473, 2017.

[4] K. Kiyokawa, M. Billinghurst, B. Campbell, and E. Woods. An occlusion capable optical see-through head mount display for supporting co-located collaboration. In *Proceedings of ISMAR*, pp. 133–141, 2003.

[5] Nayar and Branzoi. Adaptive dynamic range imaging: Optical control of pixel exposures over space and time. In *Proceedings of ICCV*, pp. 1168–1175, 2003.

[6] Y. Zhang, X. Hu, K. Kiyokawa, and X. Yang. Add-on occlusion: Turning off-the-shelf optical see-through head-mounted displays occlusion-capable. *TVCG*, 29(5):2700–2709, 2023.