Dynamic Lighting for Enhanced Sense of Depth in 2D Mid-air Image

Miyu Fukuoka
The University of Electro-Communications
Tokyo, Japan
fukuoka@media.lab.uec.ac.jp

Naoya Koizumi
The University of Electro-Communications
Tokyo, Japan
koizumi.naoya@uec.ac.jp

***Abstract*—This study aims to enhance the realism of two-dimensional (2D) mid-air images. While a 2D light source is often used to create these images owing to its affordability, high resolution, and ease of implementation, it presents challenges in eliciting a depth perception similar to that of 3D mid-air images. Aligning the lighting environments of real space with mid-air images enhances optical consistency, affecting the perception of presence and shape. Therefore, we introduced dynamic lighting in real space, achieving optical and temporal consistency, which enhances the sense of depth in 2D mid-air images. Additionally, a psychophysical experiment suggested that dynamic lighting increases depth perception, offering an enhanced sense of depth in 2D mid-air images compared to that of traditional static lighting.**

***Index Terms*—depth perception, mid-air image, mixed reality, presence, shadow, shade**

I. Introduction

Mid-air imaging is a mixed-reality technology that displays computer-generated (CG) images in real space. The Mid-air image is a real image formed mid-air by light from a light source through a retroreflective transmission optical system. Mid-air images allow characters and special effects that would not normally exist to be brought into the real world and seamlessly integrated.

To achieve advanced mixed reality, it is essential to ensure geometric, optical, and temporal consistency for a seamless integration of the real and virtual worlds. Geometric consistency involves maintaining the correct relative positions and scales of virtual objects in relation to the real world. Optical consistency refers to accurately representing shadows and reflections of virtual objects by integrating information about light sources from the real environment, whereas temporal consistency refers to the synchronization of virtual elements with real-world changes, including those experienced by users, ensuring that there are no delays in perception.

Although 2D mid-air images formed using 2D light sources offer affordability, high resolution, and ease of implementation, they are difficult to represent in terms of volume, rendering geometric consistency challenging. Nevertheless, a 2D mid-air image that can easily be perceived as three-dimensional may serve as an alternative to three-dimensional mid-air imaging technology. The matching of the lighting environment between real space and virtual objects to achieve optical consistency significantly influences the perception of the position and presence of virtual objects [1]. Incorporating shading and cast shadows as elements of optical consistency is crucial for achieving this sense of depth.

Therefore, we introduce dynamic lighting in real space and present corresponding shadows and cast shadows in real time to enhance the depth perception of 2D mid-air images. By ensuring "optical consistency," correctly rendering shading and cast shadows relative to a virtual light source, and "temporal consistency," updating these effects seamlessly with the moving light source, we achieve depth perception even in 2D mid-air images that lack geometric consistency.

II. Related Work

Mid-air images are formed when light emitted from a light source is reflected or refracted by optical elements, creating an image in mid-air. Optical systems for mid-air images include retroreflective transmissive optical systems such as Micro Mirror Array Plates (MMAP) [2] and Aerial Imaging by Retro-Reflection (AIRR) [3]. Mid-air images can be observed with the naked eye, creating the impression that CG images exist within the real environment.

Displaying a cast shadow beneath a mid-air image is known to influence the perception of the image's position and depth. Kim et al. reported that cast shadows using a projector enhances the understanding of the position of mid-air images [4]. However, the displayed mid-air images in their study were 2D illustrations of characters without shading, and the perception of thickness was not addressed. Yano et al. demonstrated that the length of a cast shadow influences the perception of thickness in mid-air images [5]. However, their study was conducted in an environment where binocular stereopsis enabled 3D perception of objects, and it has not been verified for 2D mid-air images. Furthermore, these studies were limited to static conditions where the light source was positioned above the mid-air images.

These studies did not focus on presenting a sense of depth in 2D mid-air images. While 3D mid-air images can provide volumetric and stereoscopic displays, the systems required are often large-scale and have limitations in their achievable depth range. In contrast, 2D mid-air images are relatively reasonable, easy to implement, and capable of high-resolution displays. Shading and cast shadows are key elements for perceiving 3D information from 2D images. This study aims to present a sense of depth in 2D mid-air images by introducing dynamic lighting and real-time changes in shading and cast shadows, providing an alternative to 3D mid-air images.

III. Proposed Method

Through exploratory prototyping to enhance mid-air image realism, we discovered that providing shading and cast shadows qualitatively improved the sense of presence and depth in 2D mid-air images (Fig. 1).

Typically, CG object lighting in mid-air images is fixed. However, dynamic lighting in images can make them feel more as if they "exist there, " and improve the sense of depth compared to conventional static lighting in 2D mid-air images.

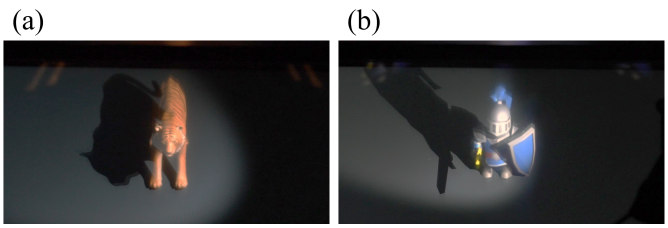

Fig. 1. (a) A mid-air image of a knight with a shield. The shadow of the shield is projected onto the knight. (b) A mid-air image of a tiger. Cast shadows can be seen from the side of the tiger.

We propose a method for rendering the shading and cast shadows of CG objects in mid-air images to match the dynamic light source in real space. The steps to ensure temporal and optical consistency are shown in Fig. 2. The position and angle of the virtual light source in real space are tracked by attaching retroreflective markers to the handheld rod's tip, simulating the light source. This setup is tracked using the optical motion capture system OptiTrack (Fig. 3).

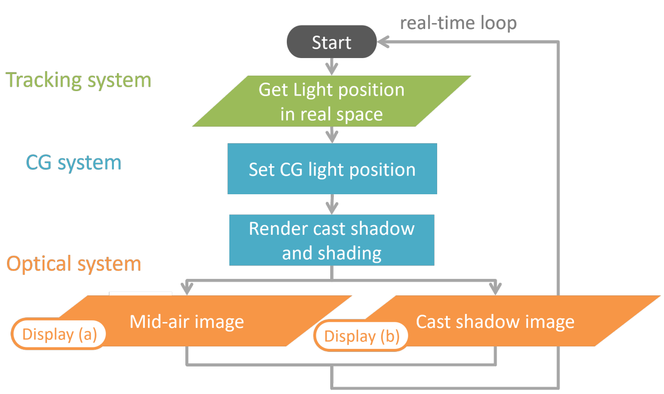

Fig. 2. Flowchart for ensuring temporal and optical consistency.

The optical design for displaying the mid-air image is shown in Fig. 3. We adopted the EnchanTable [6] optical system, which utilizes reflections from the table surface to display upright mid-air images. Moreover, it forms and grounds the mid-air image on the table surface where the cast shadow is displayed. Light from Display (a) passes through the micro-mirror array plates (MMAP), reflects off a half-mirror, and forms the mid-air image on the table surface.

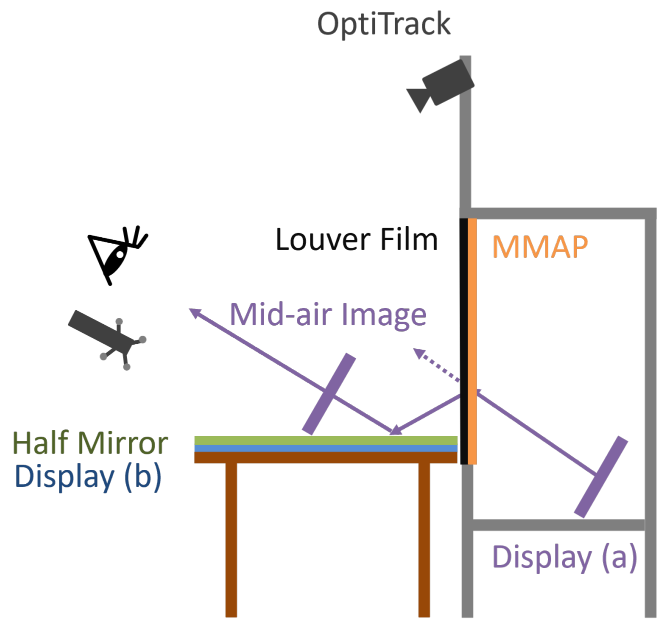

Fig. 3. Optical Design. Light from the display passes through the micro-mirror array plates (MMAP) and is reflected by the half-mirror to form mid-air images. OptiTrack tracks the dynamic light source movement to provide real-time shading information.

The CG object is displayed as a mid-air image. The cast shadows displayed on the Display (b) are rendered in real-time using Unity. In addition, a spotlight for illuminating the virtual object is placed in CG space at the same position as the real-time tracked virtual light source. A plane object is placed beneath the virtual object in the CG space, and the cast shadow projected onto this plane is displayed on Display (b).

IV. Exhibition

We conducted a demonstration at an open lab to qualitatively observe the effects of dynamic lighting in real space on a mid-air image. The participant held a stick with a marker tracked as a virtual light source position and operated the lighting to illuminate the mid-air image. A sphere and an animated character were displayed as mid-air images.

By moving the virtual lighting that illuminates the mid-air CG object, we received responses such as "It looks 3D" and "It seems to be there," indicating an enhanced sense of depth and presence. Additionally, many participants were interested in the interactive experience of changing shadows and highlights by manipulating the lighting in real space. In the experiment in Section V, we focused on sense of depth.

V. Experiment

We studied the influence of shading and cast shadows, depending on whether the lighting is dynamic or static, on the perceived depth of a mid-air image. Based on the research by Inoue et al. [7], we evaluated the perceived depth of 2D mid-air images.

A. Hypothesis

Our main hypothesis based on exploratory prototypes and exhibitions is defined as:

- H1: Dynamic lighting generally makes the sense of depth appear greater.
- H2: Dynamic lighting results in smaller differences in depth perception due to size changes.

Dynamic lighting conditions refer to an environment where the lighting is moved three-dimensionally so that the shading and cast shadows projected in mid-air images change. The specific implementation of dynamic lighting conditions in this experiment is described later. Regarding H1, if dynamic lighting is effective in creating a sense of depth, then the perceived depth should be greater than with static lighting. Regarding H2, if dynamic lighting improves accuracy in depth perception, it should become less susceptible to the effects of size changes, and the difference in depth perception due to size changes should become smaller.

B. Conditions and Stimuli

This experiment was designed with two factors: whether the lighting illuminating the 2D mid-air image is static or dynamic, and whether there is a cast shadow (Fig. 5). The spotlight illuminating the CG object displayed as a mid-air image uses Unity's spotlight object, and the apex angle of the conical lighting area is set to 50 degrees. Under static lighting conditions, the spotlight was positioned 30 cm above the top of the CG object. Under dynamic lighting conditions, as shown in Fig. 4, the spotlight illuminates obliquely downward from a height of 27 cm, always facing the CG object. It reciprocates along a circular arc with a radius of 30 cm centered on the CG object, spanning from -70 degrees to 70 degrees. Moving in one direction takes 3 seconds with an ease-in-out motion, thereby changing the shading and cast shadows of the mid-air image in real-time.

A sphere with a diameter of r cm was presented as a visual stimulus, as shown in Fig. 5. The diameters of the spheres range from 1 cm to 7 cm in 1 cm increments, and seven different sizes of spheres were prepared. Participants observe the 2D mid-air image and the cast shadow from a 30-degree depression angle. The two-dimensional display, which serves as the light source for the mid-air image, is positioned at a 30-degree tilt from upright so that the resulting 2D mid-air image is almost perpendicular to the participants' line of sight. By displaying a 2D mid-air image on a plane almost perpendicular to the line of sight, it became difficult to notice that the image was formed on a flat plane. The image of the sphere displayed on the 2D display was rendered by tilting the CG camera at a 30-degree depression angle, similar to the depression angle when participants observed visual stimuli during the experiment. The central position of the cast shadow under static lighting conditions was arranged to coincide with the imaging position of the mid-air image. Additionally, the mid-air image was formed to be grounded on the display that shows the cast shadow.

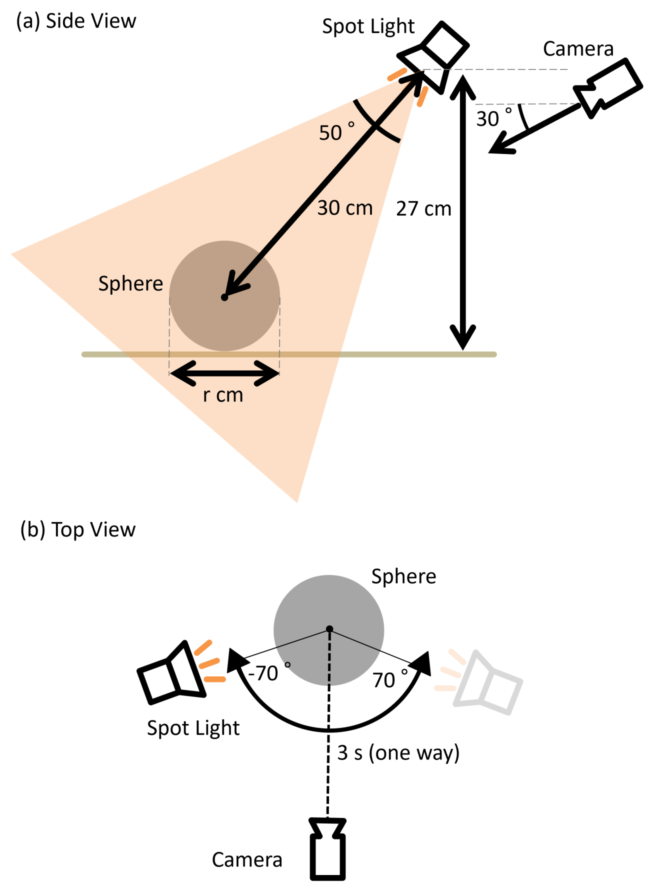

Fig. 4. Dynamic lighting configuration. The spotlight illuminates the CG object from a fixed height, always facing it, and moves along a circular arc.

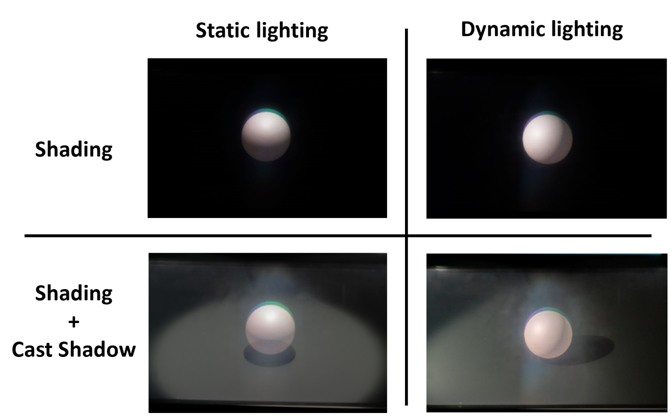

Fig. 5. Visual stimuli under static or dynamic lighting, with or without cast shadows.

C. Procedure

Participants observed mid-air images presented in a dark room. As shown in Fig. 6, the participant's viewpoint was fixed using a head-chin rest to face the mid-air image at a 30-degree depression angle. After explaining the evaluation method for perceived depth using illustrations, the perceived depth of the

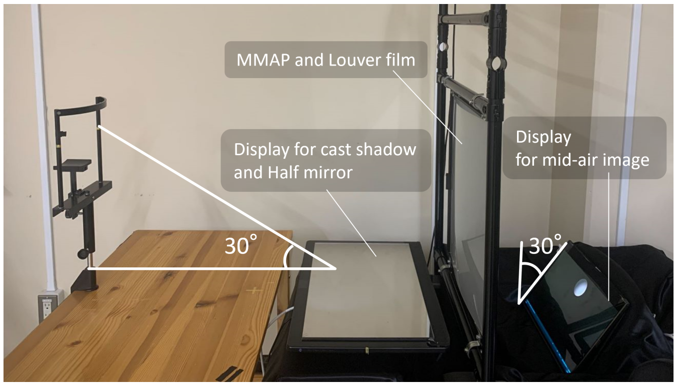

Fig. 6. Setup of the experimental apparatus. Participants fix their viewpoint on a head-chin rest and observe the mid-air image, which is tilted at a 30-degree angle.

visual stimulus was measured using the magnitude estimation method, following the previous study [7]. Specifically, participants were instructed to rate the perceived depth of the visual stimuli, displayed as 2D mid-air images, on a scale from 0 to 100, where 100 represented the convex depth equivalent to the radius of a perceived sphere, and 0 represented a flat plane.

The mid-air image was presented for 6 seconds in each trial, and participants rated the perceived depth after each presentation. For each condition, the presentation order of the visual stimuli was randomized, with 21 trials per condition (7 sphere diameters × 3 repetitions). The effects of the trial order for the lighting environment and the cast shadows were counterbalanced among participants using a Latin square design. The participants were six students (five males and one female, mean age 22.5 years) with normal or corrected-to-normal vision.

D. Result

Fig. 7 is a graph plotting the average perceived depth ratings for each condition and stimulus size. The horizontal axis represents the diameter of the sphere used as the visual stimulus, and the vertical axis shows the perceived depth ratings.

To test H1, the averaged ratings for all stimulus sizes in each condition are shown in Fig. 8. A two-way analysis of variance (ANOVA) was conducted with two factors: 2 (dynamic or static lighting) × 2 (with or without cast shadow). The results showed significant main effects for lighting ($F(1, 5) = 20.19$, $p < 0.01$) and cast shadow ($F(1, 5) = 10.12$, $p < 0.05$). There was no significant interaction between lighting and cast shadow ($F(1, 2) = 0.715$, $p > 0.05$).

To test H2, the comparison of slopes obtained from linear fitting using the least squares method on the data from Fig. 7 is shown in Fig. 9. A two-way ANOVA was conducted with two factors: 2 (dynamic or static lighting) × 2 (with or without cast shadow). The results showed no significant main effects for lighting ($F(1, 5) = 1.352$, $p > 0.05$), cast shadow ($F(1, 5) = 0.172$, $p > 0.05$), or the interaction between lighting and cast shadow ($F(1, 5) = 0.327$, $p > 0.05$).

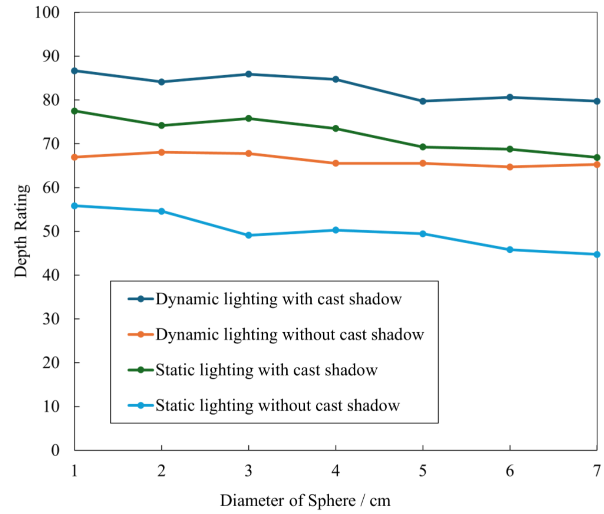

Fig. 7. Perceived depth ratings for each condition and stimulus size.

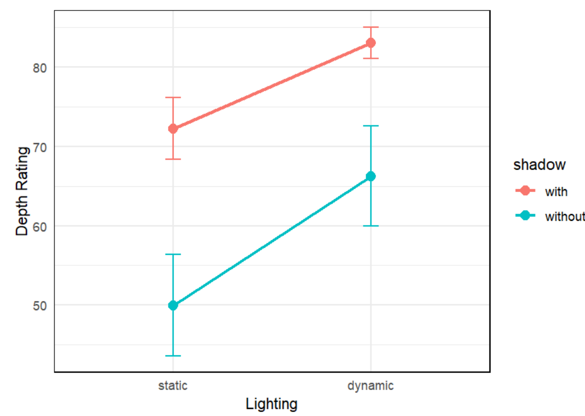

Fig. 8. Mean depth ratings under different lighting conditions and shadow conditions. Error bars represent the standard error of the mean.

VI. Discussion

Based on the result shown in Fig. 8 and the two-way ANOVA, it was suggested that dynamic lighting enhances the sense of depth compared to static lighting. Furthermore, it was suggested that displaying cast shadows beneath the mid-air image further enhances the sense of depth.

From the result of the two-way ANOVA, no differences were observed in the accuracy of depth perception. As shown in Fig. 9, some participants gave higher ratings for smaller visual stimuli (negative slope), while others gave higher ratings for larger visual stimuli (positive slope). Comments from participants after the experiment indicated that smaller stimuli were more challenging to perceive as having depth due to the reduced visibility of shading details as a result of the smaller stimulus size compared to larger stimuli. Conversely, it was noted that smaller stimuli, having less virtual depth, appeared almost as perfect spheres.

In a preliminary experiment, participants rated perceived depth from -100 (concave by the radius) to 100. Under

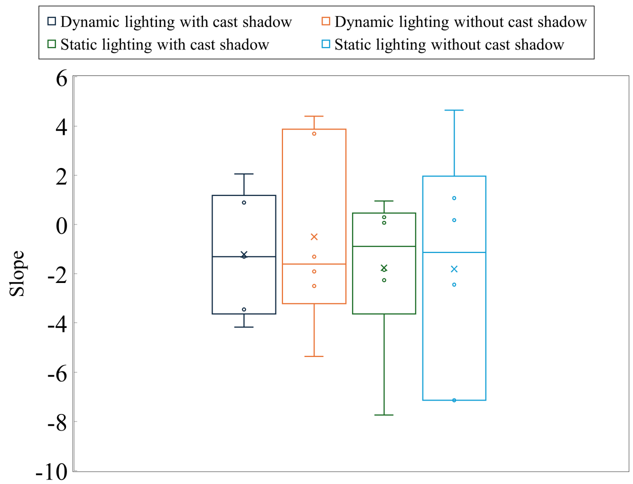

Fig. 9. The distribution of participants' slopes when linear fitting is applied to the plot of depth ratings for each size.

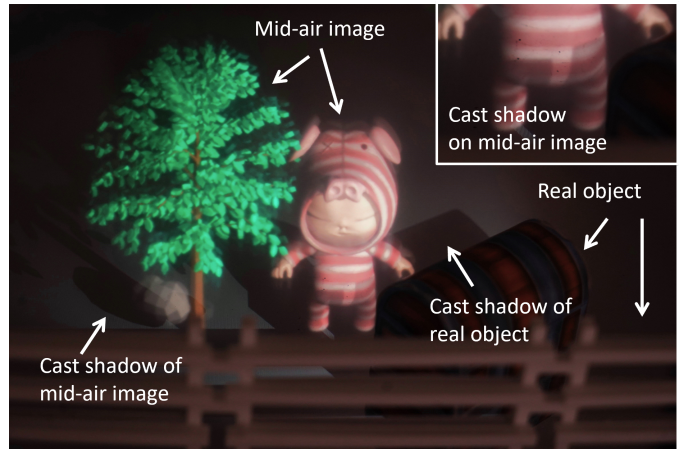

Fig. 10. An application in which the real object and the mid-air image seamlessly integrate through shadow representation. The shadow of the real object is projected onto the mid-air image.

dynamic lighting without cast shadows, negative ratings were observed, indicating concave perceptions. According to the participant, the perceived shape alternated between concave and convex depending on the shading. This phenomenon can be discussed as an ill-posed problem, where a three-dimensional shape cannot be uniquely determined from a two-dimensional image. It is suggested that the participant interpreted the changes in shading as being caused by variations in shape rather than changes in the light source position. The fact that stimuli were consistently perceived as convex under dynamic lighting conditions with the cast shadow suggests that the cast shadow serves as a cue for determining the light source position, thereby influencing the perception of convexity and concavity in the three-dimensional shape.

VII. Application

Mid-air images have the advantage of being able to be placed in the same physical space as real objects, so we developed an application to present the cast shadows of real objects in the same way as mid-air images. A 3D object with the same shape as the real object was placed in a CG space and the cast shadows were rendered in the same way as the CG objects for mid-air images. An example of the application is shown in Fig. 10. A tree and a character were displayed as mid-air images, and real objects such as a fence and a treasure chest were placed in front of the mid-air images.

By displaying the shadows of real objects and mid-air images in the same physical space and ensuring their optical consistency, an enhanced mixed reality was observed during the open campus exhibition. It was suggested that the impression of optical consistency could be exaggerated by dynamic lighting. Additionally, projecting the shadow of real objects onto mid-air images may strengthen the connection between the real object and the mid-air image, facilitate the understanding of their spatial relationship, and enhance the perceived depth of the mid-air image. The ability to display real objects and mid-air images with a similar level of realism can offer new experiences in applications such as product promotion, theme park attractions, and board games.

VIII. Conclusion

We proposed a method to enhance the depth perception of 2D mid-air images by introducing dynamic lighting in real space and presenting corresponding shading and cast shadows in real time. The method ensured optical and temporal consistency between real space and mid-air images. The demonstration and the psychophysical experiment suggest an enhanced sense of depth in the 2D mid-air image when using dynamic lighting compared to that using traditional static lighting.

Acknowledgment

This research was supported by the Japan Science and Technology Agency, Grant Number JPMJFR216L.

References

[1] N. Sugano, H. Kato, and K. tachibana, "The effects of shadow representation of virtual objects in augmented reality," in *The Second IEEE and ACM International Symposium on Mixed and Augmented Reality, 2003. Proceedings.*, 2003, pp. 76–83.

[2] S. Maekawa, K. Nitta, and O. Matoba, "Transmissive optical imaging device with micromirror array," in *Three-dimensional TV, video, and display V*, vol. 6392. SPIE, 2006, pp. 130–137.

[3] H. Yamamoto, Y. Tomiyama, and S. Suyama, "Floating aerial led signage based on aerial imaging by retro-reflection (airr)," *Opt. Express*, vol. 22, no. 22, pp. 26 919–26 924, Nov 2014.

[4] H. Kim, I. Takahashi, H. Yamamoto, S. Maekawa, and T. Naemura, "Mario: Mid-air augmented reality interaction with objects," *Entertainment Computing*, vol. 5, no. 4, pp. 233–241, 2014.

[5] Y. Yano and N. Koizumi, "Mid-air Image's Background Changes the Impression of a Mid-air Image," in *Proceedings of the ICAT-EGVE2023*, 2023, pp. 19–26.

[6] H. Yamamoto, H. Kajita, N. Koizumi, and T. Naemura, "Enchantable: Displaying a vertically standing mid-air image on a table surface using reflection," in *Proceedings of the 2015 International Conference on Interactive Tabletops & Surfaces*, ser. ITS '15. New York, NY, USA: Association for Computing Machinery, 2015, p. 397–400.

[7] H. Inoue and I. Watanabe, "The effect of location of a highlight in a shaded circle on depth perception," *Japanese Psychological Research*, vol. 76, no. 1, pp. 51–56, 2005.