Distortion-free Mid-air Image Inside Refractive Surface and on Reflective Surface

Shunji Kiuchi*
The University of Electro-Communications

Naoya Koizumi†
The University of Electro-Communications

(a) Refractive surface **(b) Reflective surface** **(c) Simulation results**

Figure 1: (a) The IEEEVR logo appears inside the refractive cylinder. (b) An image of a flower formed on a curved reflective surface. The distortion-free mid-air image is observed using the light source image generated using our method for distortion compensation. (c) Simulation results of our method with a flower displayed inside a refractive bunny.

Abstract

In this study, we propose an approach to display a distortion-free mid-air image inside a transparent refractive object and on a curved reflective surface. We compensate for the distortion by generating a light source image that cancels the distortions in the mid-air image caused by refraction and reflection. The light source image is generated via ray-tracing simulation by transmitting the desired view image to the mid-air imaging system, which includes distortive surfaces, and by receiving the transmitted image at the light source position. The proposed methods can be applied to dynamic images using a light source image as a UV map in texture mapping. Finally, we present the results of an evaluation of our method performed in an actual optical system using the generated light source image, which visually demonstrate the effectiveness of the proposed approach.

Index Terms: Computing methodologies—Computer graphics—Graphics systems and interfaces—Mixed / augmented reality; Computing methodologies—Computer graphics—Rendering—Ray tracing

1 Introduction

Display methods that superimpose computer graphics (CG) on views of the real world have played an important role in the visualization of information on real objects and in the entertainment field. Head-mounted displays (HMDs) and handheld augmented reality applications via tablets have been industrially developed as superimposed display methods, which allow users to move around in a real space along with CG. However, they require users to wear special devices on their eyes or hands. By contrast, in digital signage, several display methods that allow users to observe CG without equipment on their body have been proposed, such as the use of multiple directional images inside a refractive surface [8] and non-planar displays using printed optics [5]. However, these methods still have some limitations in terms of image resolution and dynamic full-color representation.

Mid-air image generation is a technique that displays dynamic, full-color images in real space which can be observed with the naked eye. Micromirror array plate (MMAP[1]) is a special optical element used to display mid-air images. An MMAP forms light from a light source as a real image at a plane-symmetrical position with respect to itself. However, the optical device is visible to the viewer, which detracts considerably from the intended visual effect of the display. A possible solution to this problem is to place a reflective or transparent object between the mid-air image and the MMAP. A method using a reflective surface where the mid-air image is displayed on a flat reflective surface, such as a clear-coated table, has been proposed [14]. In Scoopirit [17], a water surface was employed as a reflective surface; however, the mid-air image was distorted owing to the wave motion of the surface of the water. To the best of our knowledge, there is no precedent for the proposed method of forming a mid-air image inside a transparent object, and a solution to compensate for the distortion has not been established. In this study, we define refractive and reflective surfaces that induce distortion in the mid-air image as *distortive surfaces*.

Ray tracing is widely used in the CG field as a computational method to generate photorealistic images. It has also significantly contributed to the optimization of optical systems and advanced the understanding of complex optical behavior, such as the design of lenses and other optical systems, as well as more recently the estimation of materials and geometries via differentiable ray tracing (DRT) [32]. In this study, we leverage a ray-tracing simulation of a mid-air image to address the distortion of the mid-air image.

The proposed approach is designed to compensate for the distortion of mid-air images in optical systems, including the distortive surface. We generate a light source image that cancels the distortion caused by refraction and reflection using a ray-tracing simulation. The key idea of distortion compensation is to solve the inverse problem of generating a light source image that cancels out the distortion of the mid-air image by transporting the desired view image from the viewpoint to the optical system and receiving it at the light source position. To achieve this, we implemented a unique ray-path connection from the viewpoint to the light source, under the assumption that the correspondence between the incoming and outgoing direction of the ray at the MMAP and the distortive surface is uniquely determined. In addition, our method can be extended to dynamic images by using the generated light source image as a mapping function to deform the input image shape, similar to the

*e-mail: kiuchi@media.lab.uec.ac.jp
†e-mail: koizumi.naoya@uec.ac.jp

[1]https://aska3d.com/en/

UV coordinates in the texture mapping technique. In this study, we explain the principle of the compensation method and its practical implementation in simulation and evaluate the effectiveness of the proposed approach using actual mid-air images.

The contributions of this study are summarized as follows.

- We propose a method to compensate for the distortion of an image projected in mid-air via a ray-tracing simulation.
- We implement a ray sampling method that extracts a plane symmetry characteristic of the MMAP to generate a light source image satisfying our principle.
- We provide a demonstration of the proposed approach by applying the compensation method to a dynamic image using a texture mapping technique via the generated UV coordinates image.

2 Related Work

2.1 CG images in physical space

Several display methods have been proposed to make CG images appear inside and outside of physical objects. Papas et al. [21] displayed multiple images that varied according to the position from which they were viewed by optimizing the lens alignment. Hirayama et al. [8] also displayed multiple directional images inside a transparent object by designing the density of tiny bubbles. In works on the use of physical objects as light filters, multi-layered attenuators and refractive surface geometries to yield caustics are optimized to project the image when they are lit from a light source [3,22]. For dynamic expression of CG images, PAPILLON [5] was proposed, which used a non-planar display with printed optics to enable dynamic movement of the eyes of a toy character. In contrast to existing techniques, our proposed method brings a digital image into an actual space without the interference of physical objects. Moreover, it has the potential to enable further exploration of new forms of interface between physical and digital space.

2.2 Mid-air interaction

Several devices have been developed to form mid-air images, such as MMAP, aerial imaging by retro reflection [29], radially arranged dihedral corner reflector arrays [31], and retroreflective mirror arrays [13]. In this study, we use MMAP, which can display a high-luminance mid-air image and is commercially available.

Several mid-air image interactions using MMAP have been proposed. MARIO [11] realized a mid-air image that moved on physical objects by utilizing depth measurements of real space, which enabled users to control the position of a mid-air image in three-dimensions with blocks. HaptoMime [18] implemented haptic feedback in a mid-air image using ultrasound. HaptoClone [16] leveraged ultrasound haptic feedback to develop a telepresence system that transferred the visual and haptic information from other places. Tsuchiya et al. [26] also proposed a telepresence system that enabled communication with others from the perspective of a CG character by superimposing the viewpoint of a camera on a mid-air image. Although these mid-air imaging systems allow users to interact with CG images in physical space, the presence of the MMAP itself is often distracting or intrusive, because the MMAP and mid-air images are aligned with the viewpoint.

A possible solution to reduce the presence of the MMAP is to place a refractive or reflective surface between the MMAP and the mid-air image. Several optical systems employing reflective surfaces have been proposed. These can display upright mid-air images on glossy surfaces, such as marble and a coated desk [14]. Scoopirit [17] leveraged water as a reflective surface that enabled viewers to apprehend projected mid-air images in handfuls of water scooped from a water-based display. However, the mid-air image projected from reflective surfaces can be distorted owing to the surface geometry, especially water waves. In addition, there is no precedent for inserting a refractive surface. Furthermore, although several works have addressed the distortion compensation of displayed images in the context of near-eye displays [7,9], no method has been established in the literature to compensate for the distortion of images in optical systems which include optical elements that differ from conventional lenses and mirrors, such as MMAP. In the present work, we demonstrate a method for compensating for mid-air image distortion by using a CG-based ray-tracing simulation.

2.3 Ray tracing-based simulation

Optical simulation has been used in the design of mid-air imaging systems using MMAP. Otao et al. [20] developed a transparent optical HMD using MMAP by utilizing a light path simulation to determine the lens and imaging position. Choi et al. [6] simulated light paths to determine the position of a light source to eliminate stray lights. Kiuchi et al. [12] demonstrated a CG-based simulation of MMAP by ray tracing which could reproduce the characteristics of MMAP, such as plane symmetry, undesired images, limited visible range, and luminance decay. In this study, we utilize CG-based simulation to perform distortion compensation of the mid-air image.

Ray tracing is a computational optical simulation used to track interactions between rays and objects in the CG field, which is used to generate photorealistic images. Ray tracing has also been widely used to design optical elements and to understand optical behaviors. In lens design, it is used for geometry optimization by evaluation indices such as the point spread function [1, 25]. Recently, DRT has attracted attention for the estimation of the material, shape, and illumination. Our method, which generates a light source to display a distortion-free mid-air image, can be said to be an inverse rendering technique, and is similar to several works on DRT [2,19]. However, our method differs from gradient-based optimization. When gradient-based optimization is applied for distortion compensation, it focuses on the goal of delivering a distortion-free image to the viewpoint, and information in the desired view image may be lost in the generated light source image. Therefore, we simply transport the desired view image from the viewpoint of the light source position passing through the mid-air imaging system.

3 Compensation Method

In this section, we describe the principle and implementation of light source image generation to compensate for the mid-air image distortion caused by refraction and reflection. One possible way to compensate for the distortion is to calculate the homography matrix for the light source shape that cancels out the distortion; however, it may not be applied when the distortion shape is complex. Another possible solution is to determine the correspondence between the mid-air image and the light source pixels by displaying multiple patterns on the light source [23]. However, this would also be affected by the resolution degradation and stray light caused by the MMAP. Further discussion on the structured light method can be found at Section 5.1. We address these difficulties in utilizing the inverse ray-tracing method, which inversely transports view-side information toward the optical system via rays launched from the viewpoint.

3.1 Principle

The optical system is composed of light sources, MMAP, and a distortive surface, as shown in Figure 2. All elements comprising the optical system, the transformation of the MMAP, viewpoint, light source, distortive surface, and geometry of the distortive surface are known. In addition, we assume that the light source emits light uniformly on a hemispherical surface that enters its normal at each pixel. When the refractive surface is not perpendicular to the incident rays of the mid-air image and the reflective surface is not flat, distortion occurs in the mid-air image.

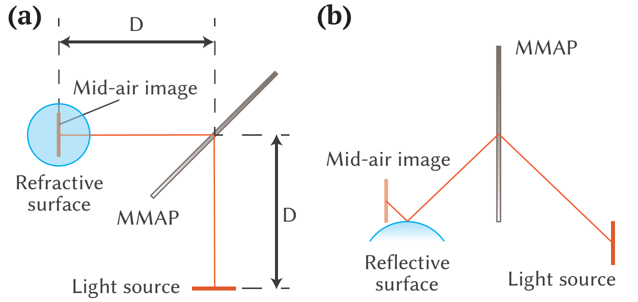


Figure 2: Mid-air images are formed passing through (a) the refractive surface, and (b) the reflective surface.

To address the distortion issue, we leverage a ray tracing-based simulation. Physically based ray tracing, as formulated by Kajiya [10], calculates the radiance at pixel j by integrating the contribution in a path space Ω incident on a viewpoint.

$$I_j(x_0) = \int_{\Omega} f_j(X)\mathrm{dX}, \tag{1}$$

where $X = (x_0, \ldots, x_k) \in \Omega$ represents a ray path of depth k that starts from viewpoint x_0 and ends at x_k on the light source, and $f(X)$ is a radiance from X, which is formulated as follows.

$$f(X) = L_e(x_{k-1} \leftarrow x_k) \prod_{i=1}^{k-1} f_s(x_{i-1} \leftarrow x_i \leftarrow x_{i+1}), \tag{2}$$

where $L_e(x_{k-1} \leftarrow x_k)$ is the radiance emitted from a surface point x_k on a light source to a surface x_{k-1}. $f_s(x_{i-1} \leftarrow x_i \leftarrow x_{i+1})$ is a *bidirectional scattering distribution function* (BSDF) at every interaction surface x_i of the ray path. For our distortion compensation method, the key idea is to find the L_e of the light source embedded in the optical system that cancels out the distortion in the mid-air image.

We generate a light source image that cancels the distortion by inversely transporting a *desired view image*, which is the view image of an undistorted image at the mid-air image position, from a fixed viewpoint to the light source position through the mid-air imaging system. The concept of the proposed method is illustrated in Figure 3.

A ray path, which originates from the viewpoint and ends at the light source, traces each element in the mid-air imaging system in the following order: viewpoint, distortive surface, MMAP, and light source. We assume that the distortive surface is perfectly smooth, and the incidence and outgoing direction at the MMAP and the distortive surface are uniquely determined from the plane-symmetry characteristic of MMAP and Snell's law, respectively. Under this assumption, Equation 1 can be considered as Equation 2 because the light path from the viewpoint to the light source through the optical system is uniquely determined. We denote the radiance of each pixel as I_j^{ref} from the viewpoint of an undistorted mid-air image, and the emittance at the light source as L_e^{ls}. The proposed approach can inversely obtain L_e^{ls}, which results in an undistorted mid-air image by transporting I_j^{ref} from the viewpoint to the light source passing through the distortive surface and MMAP. Ignoring the attenuation due to the BSDF at each surface, the obtained L_e^{ls} replays the light path that delivers I_j^{ref} to the viewpoint; that is, the distortion of the mid-air image is compensated by the obtained L_e^{ls}.

3.2 Unique ray-path connection

Several techniques are required to ensure the assumption described in Section 3.1 that ray paths are uniquely determined. For the distortive surface, we assumed a perfectly smooth surface. However, for the refractive surface, refractive and reflective components simultaneously exist with a certain ratio determined by the *Fresnel reflectance*. In this study, we only focus on the distortion effect caused by refraction; hence, we ignore the ray path of reflection and only track the refraction path determined by *Snell's law*. The original MMAP also violates our assumption. As shown in Figure 4 (a), the MMAP consists of two orthogonal layers of glass-deposited micromirror arrays. When the light incident on the MMAP from the light source is reflected once by each mirror layer, the mid-air image is formed as a real image at the plane-symmetry position. However, the light reflected by either the layer or glass surface becomes a virtual image. Furthermore, the MMAP involves a well-known issue in that it degrades the quality of the mid-air image owing to the microstructure of the mirrors and the slight change in optical path length caused by its thickness [28]. Therefore, MMAP cannot be used directly in our compensation method because the outgoing direction of a ray at the MMAP is not uniquely determined by the incidence ray.

We implemented a ray sampling method that only traces the plane-symmetry paths at the MMAP required to form the mid-air image, as shown in Figure 4 (b). The MMAP creates a mid-air image by re-emitting light rays from the incident direction $\vec{i}$ in the direction $\vec{o}$, which is plane symmetric to the MMAP. Therefore, the goal of tracing the plane-symmetry path is to obtain the output vector $\vec{o}$ from the incident vector $\vec{i}$ and surface normal $\vec{n}$. We define $\vec{h}$ as the vector of the vertical line from the origin of $\vec{i}$ to the surface, and it is the inner product of $\vec{i}$ and $\vec{n}$ multiplied by $\vec{n}$. Therefore, $\vec{o}$ can be expressed using the following equation.

$$\vec{o} = -\vec{i} + 2\vec{h} = -\vec{i} + 2(\vec{i} \cdot \vec{n})\vec{n} \tag{3}$$

3.3 Implementation

The process of generating the light source image is divided into two steps, as shown in Figure 3. The first step is to render the desired view image. In the implementation, a plane that emits light by itself is placed at the mid-air image position and rendered from a fixed camera. In the second step, rays are directed into the optical system, including the distortive surface, from the same viewpoint as in the first step. The pixel value of the desired view image at the pixel position of the launched ray is assigned to the ray. When the ray passes through the distortive surface and the MMAP and arrives at the light source position, the transported pixel value is stored. When the process is completed for all pixels of the desired view image, the pixel values stored at the light source position are output as the light source image.

In our implementation of refractive and reflective surfaces, we assume that the surfaces are completely smooth and that there is no attenuation of ray color due to the surface albedo. Therefore, we assume that there is no attenuation in the color of the desired view image during the process of transmitting it from the viewpoint through the optical system to the light source, and no blur effect due to distortive surfaces occurs. When a ray intersected with the refracting surface, we avoided tracing the reflection path, and instead traced only the refraction path. To consider the condition of total reflection, the refraction path tracing was performed only when the following equation was satisfied.

$$\eta_1 \sin\theta < \eta_2, \tag{4}$$

where the refractive index of the medium at the incident side was η_1, the angle between the normal and the incident ray was θ, and the refractive index of the medium at the exit side was η_2.

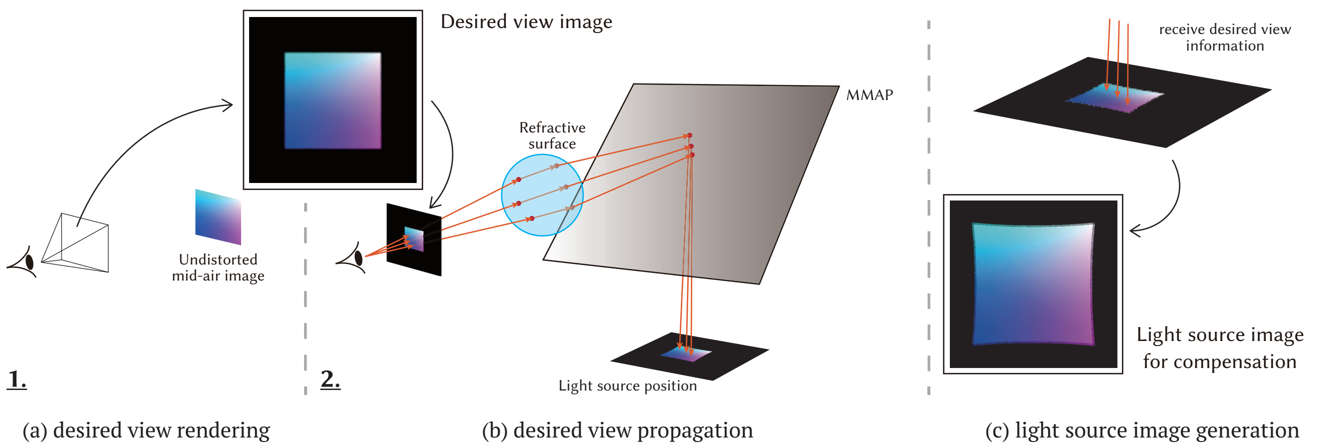


Figure 3: Our method generates a light source image to compensate for the distortion of the mid-air image. (a) The first step is to render the desired view image. In the implementation, the self-emitted plane is placed at the mid-air image position, which is determined by a mid-air imaging system. (b) The second step inversely transports the desired view image from the viewpoint to the light source position through the system, and (c) receives a pixel value of the desired view image at the light source position. The pixel coordinates to store the transported pixel values are determined by multiplying the texture coordinates at the light source with the image resolution.

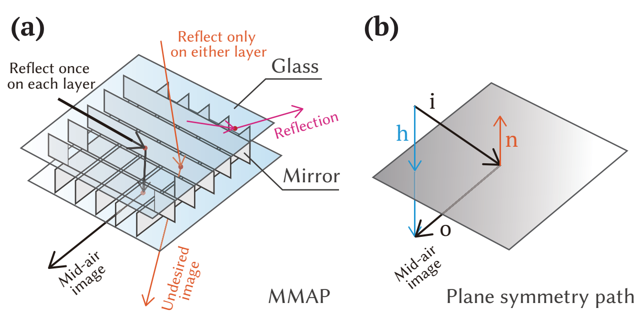


Figure 4: Selective sampling of only the ray path of the mid-air image to improve the efficiency of simulation.

We implemented our algorithm using NVIDIA OptiX 7.2, and ray-tracing computation was performed on a GPU (NVIDIA RTX 3090 with 24 GB RAM). As a benchmark, we measured the computation time using a refractive sphere. The resolution of the desired view image was set to 1024 × 1024 px, and the size of the undistorted mid-air image was set to 2 cm square. The object that distorted the mid-air image was a sphere with a radius of 2.5 cm and a refractive index of 1.49. For the MMAP, a simple planar geometry was used, and the sampling method was attached. The number of samples per pixel was set to 1000 to avoid missing pixel values from the light source, and the maximum number of traces was set to 10. As a result, the computation time was 0.55 s. The generated light source image is shown in Figure3 (c).

3.4 Light source image as mapping function

We present a demonstration in which we performed distortion compensation on dynamic images using the light source image as the UV coordinates for texture mapping. The UV light source image to deform the shape of the input image, as shown in Figure 5, can be generated by assigning UV values of the undistorted mid-air image instead of pixel values when projecting the desired view image to the optical system in the simulation. Because the UV image is generated to produce a distortion-free mid-air image, when the input image is deformed by the UV image, both a static image and a dynamic image without distortion can be displayed as a mid-air image. For texture mapping, the red and green values of the generated UV image are used as UV coordinates to deform the shape of the input video, and the result of the distortion compensation is shown in Figure 5. Distortion compensation was successfully applied to different video frames.

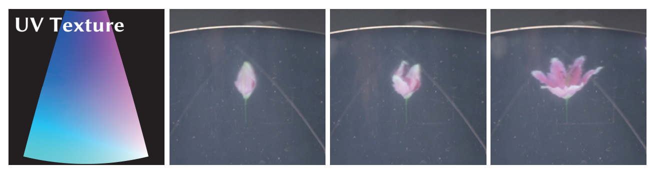


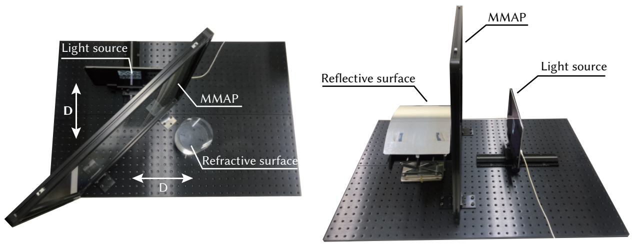
Figure 5: Demonstration of presenting full-color dynamic images using the generated light source as the UV coordinates for texture mapping.

4 Evaluation

4.1 Mid-air image shape

To evaluate our compensation method, we compared the shape of the mid-air image between the reference condition in which the mid-air image was simply displayed and two comparison conditions in which the mid-air image was formed through a distortive surface with and without the use of the generated light source image.

Refractive surface The optical system was assembled on an optical bench, as shown in Figure 6 (a). The MMAP was 48.8 cm on a side and 4 mm thick. The angle between the light source and MMAP was set to 45°, with $D = 15$ cm. The mid-air image to be displayed was a 3 cm square with a yellow lattice of seven rows and seven columns. An acrylic cube of 50 mm per side was used as the refractive surface, and it was rotated vertically by 15.8°. In the simulation, a refractive index of 1.49 was applied to the cube to generate the light source image according to the manufacturer. The mid-air image was captured using a Sony α III camera with a sensor resolution of 6000 × 4000 px and a SEL70300G lens. An iPad Pro (10.5 inch, 2224 × 1668 px) was used as the light source. The camera was moved around the mid-air image with an azimuth of $-15°$ to $15°$ at 5° intervals.



(a) For refractive surface

(b) For reflective surface

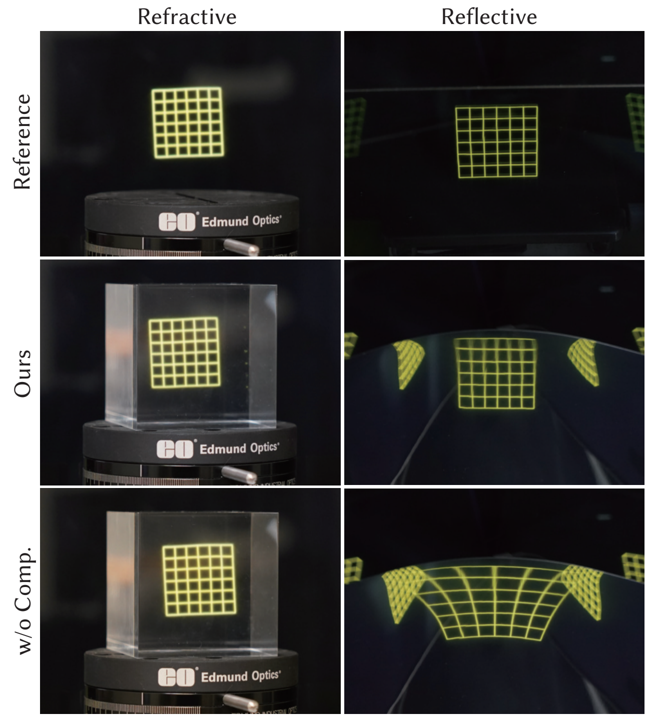
Figure 6: Actual mid-air imaging systems on an optical bench.

The captured images are shown in the left column of Figure 7 (a). The figure shows that the mid-air image inside the refractive cube is almost at the same horizontal position as the reference conditions; however, without compensation, the mid-air image was shifted to the right by refraction.

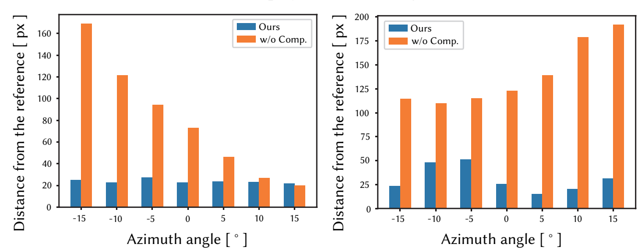
The distances of the intersection points of the displayed lattice between the reference and the two comparison conditions were calculated, as shown in Figure 7 (b). The average distance of the intersection positions from the mid-air image in the reference at $-15°$ to $15°$ was 23.91 px in our study and 78.87 px without compensation. Our method displays the mid-air image close to the desired position by canceling out the distortion effects.

Reflective surface For the reflective surface, we used a bendable polycarbonate mirror that was part of a cylinder with a radius of 25 cm and formed a mid-air image on the top of the mirror. The optical system is shown in Figure 6 (b). The MMAP was placed vertically, and the reflective surface was placed in the ray path between the mid-air image and the MMAP. The mid-air image to be displayed was a 6 cm square with a yellow lattice of seven rows by seven columns. The light source and camera were the same as in the case of a refractive surface. The view angle of the camera was set to $30°$ from the horizontal plane.

The captured mid-air images under the three conditions are shown in the right column of Figure 7 (a). Our method maintained the mid-air image as a square; however, without compensation, the lattice was bent significantly. Figure 7 (c) shows the measurement results for the intersection positions of the lattice. The average distance from the reference at $-15°$ to $15°$ was 30.53 px with compensation and 138.68 px without compensation.



(a) Displayed mid-air images



(b) Refractive surface

(c) Reflective surface

Figure 7: (a) Mid-air images were captured at an azimuth angle of $0°$ to evaluate the effectiveness of our method. (b, c) Distances of the intersection points of the displayed lattice between the reference and the two comparison conditions.

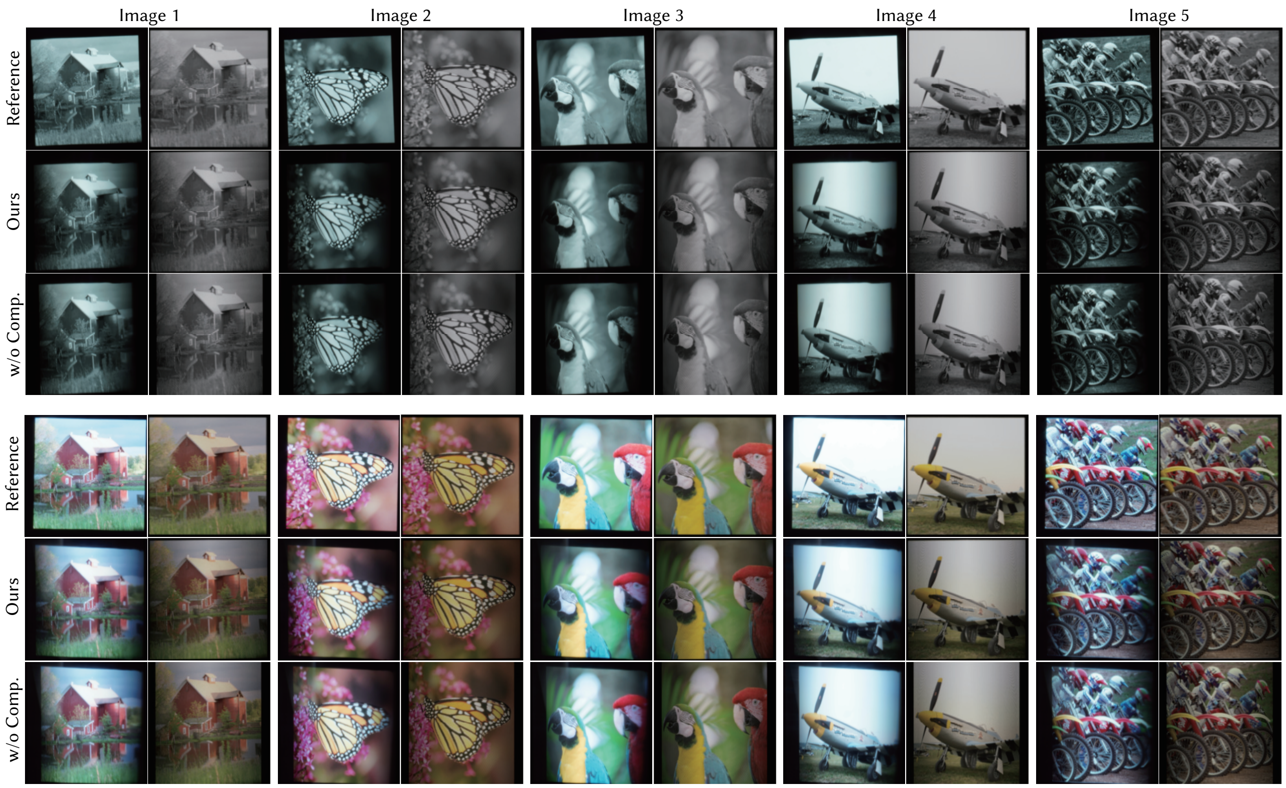
4.2 Image quality

We measured the mid-air image quality using the structural similarity (SSIM) [27] and the perceptual similarity (LPIPS) [33] methods to evaluate not only the shape of the mid-air image but also the similarity with the reference image in actual and simulated mid-air images. The reference image used to evaluate the quality was the image captured when the mid-air image was simply displayed in the same manner as in Section 4.1. The values of SSIM range from 0 to 1, with higher values indicating higher image quality. The values of LPIPS indicate the perceptual loss of a displayed image from the reference image, and lower values are defined as better results. The simulated mid-air images were rendered using Cycles rendered [2] in Blender 2.93, with the MMAP model established by Kiuchi et al. [12]. For the refractive surface, an acrylic cylinder with a radius of 5 cm, height of 10 cm and refractive index of 1.49 was used. For the reflective surface, the same bendable polycarbonate mirror was used in the evaluation described in Section 4.1. The image quality was evaluated using the LIVE1 [24] dataset.

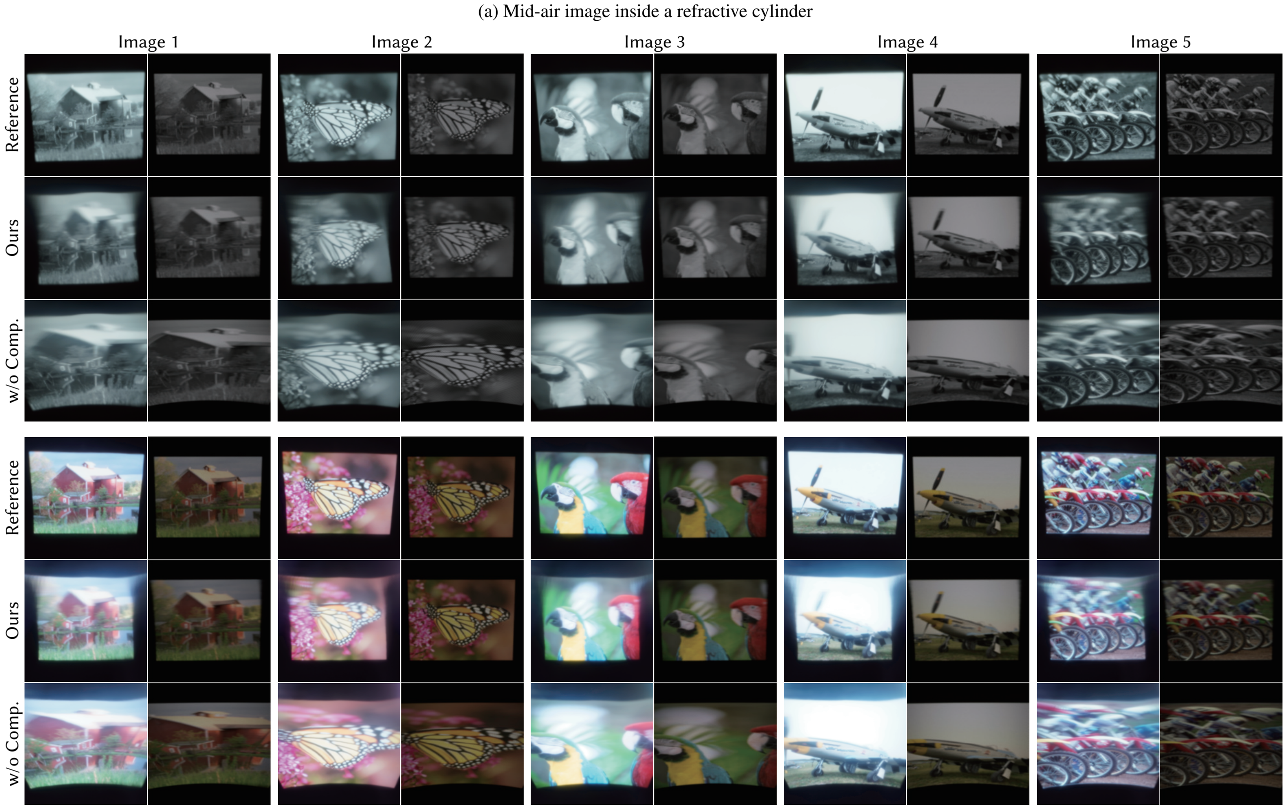
The captured images are shown in Figure 8, and the measured SSIM and LPIPS values are shown in Table1. For the refractive surface, although the width of the mid-air image was narrowed by the cylinder without distortion compensation, our method compensated for the image shape, canceling out the narrowing effect. The average SSIM values in actual mid-air images were 0.620 with our method and 0.594 without compensation, and 0.156 with our method and 0.126 without compensation in simulated images. The average LPIPS values in actual images were 0.441 with our method, 0.457 without compensation, and 0.550 and 0.608 with our method and without compensation in simulated images, respectively. Our method led to a better result, except for the case of Image 1 in the actual image.

For the reflective surface, the mid-air image was significantly stretched owing to the bent mirror; however, our method maintained the image shape as a square as in the reference condition. The average SSIM values in actual mid-air images for were 0.771 with our method, 0.731 without compensation, and 0.478 with our method and 0.374 without compensation in simulated images. The average LPIPS values in the actual images were 0.292 with our method, 0.386 without compensation, whereas they were 0.301 and 0.412 in simulated images with our method and without compensation,

[2] `https://www.cycles-renderer.org/`



(a) Mid-air image inside a refractive cylinder



(b) Mid-air image on a bended mirror

Figure 8: Images 1 to 5 were displayed through distortive surfaces to evaluate the SSIM and the LPIPS. The left column in each image shows photographs taken on an optical bench, and the right column shows the simulated results in Blender using the MMAP model established by Kiuchi et al. [12].

respectively.

The SSIM values in the simulated images were significantly lower than those in the actual images. This result may be attributed to the resolution degradation of the MMAP, as reported by Yahagi et al. [28]. MMAP produces a mirror lattice pattern in the mid-air image owing to its structure, and the pattern is deformed from the original mid-air image by the distortive surface, resulting in a large reduction in the SSIM value. In addition, when the `blur` function of OpenCV was applied to the simulated image with a 5×5 kernel, the average SSIM was 0.742 with our method and 0.616 without compensation for the refractive surface, and 0.903 with our method and 0.735 without compensation in the case of the reflective plane.

5 Discussion and Future Work

5.1 Compensation method by other candidates

In this section, we discuss other possible methods for realizing distortion compensation in comparison with our method. The proposed approach achieves distortion compensation by inversely transporting the desired view image from a fixed viewpoint through a mid-air imaging system to the light source. This technique was designed to solve a specific inverse problem to generate a light source image, resulting in a distortion-free mid-air image. This inverse problem can be solved using gradient-based optimization via DRT. In particular, several works are similar to our method, which inversely transports the view information toward a scene [2, 19]. While these methods can be applied to comprehensive problems other than light source generation, such as optimizing the geometry and surface reflectance, gradient-based methods may cause rough results owing to local optimum and derivative acquisition in pixel space. In this study, we employed a simple method of transportation of the desired view image from a fixed viewpoint to the light source, under the assumption that the ray paths from a viewpoint to a light source are uniquely determined. Although it is not suitable for solving a comprehensive inverse problem, this method has the advantage of obtaining a one-to-one correspondence between the pixel values of the desired view image and the generated light source image, and of avoiding the rough result.

Another possibility is to use a structured light method, which calculates the pixel correspondence between the light source and the mid-air image using gray code pattern. This method has been used in several studies to correct the projector-camera system in projection mapping and other applications [4, 30]. As shown in Figure 9, the gray code pattern is displayed on the light source, and it is possible to obtain the pixel correspondence between the light source and the mid-air image. Therefore, distortion compensation using the structured light method is considered to be possible to some extent. However, because the mid-air imaging optical element generates an obtrusive image called stray light, the structured light method requires a process to separate the mid-air image from the stray light part after obtaining the correspondence in real space. In addition, correspondence cannot be obtained for a part overlapped by stray light. Furthermore, the mid-air image and the stray light parts are not known in advance. To determine the mid-air image part that is separated from the stray light, an optical system designer must manually determine the mid-air image part after capturing images. If the structured light method is used in a simulation that implements the unique path connection in MMAP, as in this study, a clean correspondence map can be obtained without considering stray light. However, the structured light method requires the rendering of a series of mid-air images for the number of gray code patterns, and it requires more computation time than our method.

5.2 Limitation and Future work

Fixed viewpoint One limitation of the proposed approach is that the distortion compensation is performed using a fixed viewpoint; hence, the mid-air image will be distorted when observed from other viewpoints. This can be solved by using a special light source that can selectively present images depending on the viewpoint, such as a light field display, or by developing a system that tracks the user's viewpoint position and dynamically switches the light source image.

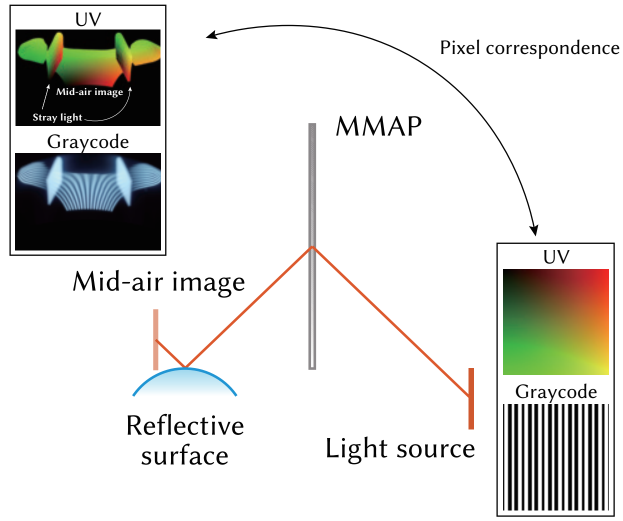


Figure 9: Obtained pixel correspondence between the light source and the displayed mid-air image by the structured light method.

Surface geometries Another limitation is the geometries of the distortive surface. Currently, our method works only with a known geometry of the distortive surface. This could be extended to unknown geometries in combination with prior geometry estimation as implemented by Lyu et al. [15]. Furthermore, especially in the case of complex meshes, light rays may not reach the viewpoint direction, and mid-air images may not be formed. Although this is a limitation of our principle, we plan to implement a preliminary verification phase that explores whether mid-air images are formed under varying conditions through a simulation.

Color compensation For future work, we would like to compensate for the light attenuation that occurs between the light source and viewpoint. In this study, we ignored the attenuation due to the distorted surface and the MMAP in the ray path from the viewpoint of the light source when performing distortion compensation. Therefore, our method cannot consider the darkening effect of the image due to the Fresnel reflectance of the refractive surface, as shown in Fig. 7(a), or the color transformation caused by the surfaces. These attenuations can be compensated for by transforming the color information of the transported desired view image so as to cancel out the attenuation due to the BSDF described in Equation 2.

6 Conclusion

In this study, we have proposed a display method for distortion-free images inside a refractive surface and on a reflective surface. We generated a light source image that canceled out the distortion of an image by transporting the desired view image to the mid-air imaging optical system and received the projected view image at the light source position in a ray-tracing simulation. The compensation method for mid-air image distortion was applied to both reflective and refractive surfaces, and its effectiveness has been verified through experiments with actual optical systems. Furthermore, we have presented a demonstration of dynamic mid-air image distor-

Table 1: SSIM and LPIPS results for each condition, and better results are shown in bold. Our method shows better results than without compensation, except for the SSIM in the refractive surface under actual conditions in Image1.

			Image1		Image2		Image3		Image4		Image5	
Surface	Condition		SSIM	LPIPS	SSIM	LPIPS	SSIM	LPIPS	SSIM	LPIPS	SSIM	LPIPS
Refractive	Actual	Ours	0.618	**0.456**	**0.551**	**0.442**	**0.668**	**0.379**	**0.698**	**0.381**	**0.564**	**0.547**
		w/o Comp.	**0.621**	0.474	0.540	0.457	0.641	0.399	0.652	0.403	0.516	0.554
	Simulation	Ours	**0.146**	**0.553**	**0.162**	**0.540**	**0.155**	**0.484**	**0.159**	**0.519**	**0.160**	**0.653**
		w/o Comp.	0.122	0.607	0.129	0.604	0.132	0.537	0.134	0.573	0.112	0.719
Reflective	Actual	Ours	**0.784**	**0.291**	**0.755**	**0.295**	**0.808**	**0.277**	**0.799**	**0.238**	**0.708**	**0.361**
		w/o Comp.	0.748	0.385	0.711	0.398	0.765	0.392	0.778	0.294	0.651	0.461
	Simulation	Ours	**0.474**	**0.286**	**0.481**	**0.309**	**0.472**	**0.263**	**0.461**	**0.283**	**0.504**	**0.364**
		w/o Comp.	0.376	0.388	0.379	0.420	0.363	0.391	0.367	0.382	0.385	0.481

tion compensation using the generated light source image as UV coordinates to deform the shape of the input image.

Acknowledgments

This work was supported by the Canon Foundation.

References

[1] K. Akşit, P. Chakravarthula, K. Rathinavel, Y. Jeong, R. Albert, H. Fuchs, and D. Luebke. Manufacturing application-driven foveated near-eye displays. *IEEE Transactions on Visualization and Computer Graphics*, 25(5):1928–1939, 2019. doi: 10.1109/TVCG.2019.2898781

[2] D. Azinovic, T. Li, A. Kaplanyan, and M. Nießner. Inverse path tracing for joint material and lighting estimation. In *IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2019, Long Beach, CA, USA, June 16-20, 2019*, pp. 2447–2456. Computer Vision Foundation / IEEE, 2019. doi: 10.1109/CVPR.2019.00255

[3] I. Baran, P. Keller, D. Bradley, S. Coros, W. Jarosz, D. Nowrouzezahrai, and M. H. Gross. Manufacturing layered attenuators for multiple prescribed shadow images. *Comput. Graph. Forum*, 31(2):603–610, 2012. doi: 10.1111/j.1467-8659.2012.03039.x

[4] D. Bergmann. New approach for automatic surface reconstruction with coded light. In T. F. Schenk, ed., *Remote Sensing and Reconstruction for Three-Dimensional Objects and Scenes*, vol. 2572, pp. 2 – 9. International Society for Optics and Photonics, SPIE, 1995. doi: 10.1117/12.216931

[5] E. Brockmeyer, I. Poupyrev, and S. Hudson. Papillon: Designing curved display surfaces with printed optics. In *Proceedings of the 26th Annual ACM Symposium on User Interface Software and Technology*, UIST '13, pp. 457—-462. Association for Computing Machinery, New York, NY, USA, 2013. doi: 10.1145/2501988.2502027

[6] S. Choi, S. Park, and S.-W. Min. Design of ghost-free floating 3d display with narrow thickness using offset lens and dihedral corner reflector arrays. *Opt. Express*, 28(10):15691–15705, May 2020. doi: 10.1364/OE.392036

[7] D. Dunn, C. Tippets, K. Torell, P. Kellnhofer, K. Akşit, P. Didyk, K. Myszkowski, D. Luebke, and H. Fuchs. Wide field of view varifocal near-eye display using see-through deformable membrane mirrors. *IEEE Transactions on Visualization and Computer Graphics*, 23(4):1322–1331, 2017. doi: 10.1109/TVCG.2017.2657058

[8] R. Hirayama, H. Nakayama, A. Shiraki, T. Kakue, T. Shimobaba, and T. Ito. Projection of multiple directional images on a volume structure with refractive surfaces. *Opt. Express*, 27(20):27637–27648, Sep 2019. doi: 10.1364/OE.27.027637

[9] Y. Itoh and G. Klinker. Simultaneous direct and augmented view distortion calibration of optical see-through head-mounted displays. In *2015 IEEE International Symposium on Mixed and Augmented Reality*, pp. 43–48, 2015. doi: 10.1109/ISMAR.2015.14

[10] J. T. Kajiya. The rendering equation. *SIGGRAPH Comput. Graph.*, 20(4):143–150, aug 1986. doi: 10.1145/15886.15902

[11] H. Kim, I. Takahashi, H. Yamamoto, S. Maekawa, and T. Naemura. Mario: Mid-air augmented reality interaction with objects. *Entertainment Computing*, 5(4):233 – 241, 2014. doi: 10.1016/j.entcom.2014.10.008

[12] S. Kiuchi and N. Koizumi. Simulating the appearance of mid-air imaging with micro-mirror array plates. *Computers & Graphics*, 96:14–23, 2021. doi: 10.1016/j.cag.2021.02.007

[13] T. Koike and Y. Onishi. Aerial 3d imaging by retroreflective mirror array. In *Companion Proceedings of the 2018 ACM International Conference on Interactive Surfaces and Spaces*, ISS '18 Companion, pp. 25–29. ACM, New York, NY, USA, 2018. doi: 10.1145/3280295.3281365

[14] N. Koizumi and A. Sano. Optical system to display mid-air images on a glossy plane and remove ground images. *Opt. Express*, 28(18):26750–26763, Aug 2020. doi: 10.1364/OE.400104

[15] J. Lyu, B. Wu, D. Lischinski, D. Cohen-Or, and H. Huang. Differentiable refraction-tracing for mesh reconstruction of transparent objects. *ACM Trans. Graph.*, 39(6), Nov. 2020. doi: 10.1145/3414685.3417815

[16] Y. Makino, Y. Furuyama, S. Inoue, and H. Shinoda. Haptoclone (haptic-optical clone) for mutual tele-environment by real-time 3d image transfer with midair force feedback. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, CHI '16, pp. 1980–1990. ACM, New York, NY, USA, 2016. doi: 10.1145/2858036.2858481

[17] Y. Matsuura and N. Koizumi. Scoopirit: A method of scooping mid-air images on water surface. In *Proceedings of the 2018 ACM International Conference on Interactive Surfaces and Spaces*, ISS '18, pp. 227—-235. Association for Computing Machinery, New York, NY, USA, 2018. doi: 10.1145/3279778.3279796

[18] Y. Monnai, K. Hasegawa, M. Fujiwara, K. Yoshino, S. Inoue, and H. Shinoda. Haptomime: Mid-air haptic interaction with a floating virtual screen. In *Proceedings of the 27th Annual ACM Symposium on User Interface Software and Technology*, UIST '14, pp. 663–667. ACM, New York, NY, USA, 2014. doi: 10.1145/2642918.2647407

[19] M. Nimier-David, S. Speierer, B. Ruiz, and W. Jakob. Radiative backpropagation: An adjoint method for lightning-fast differentiable rendering. *ACM Trans. Graph.*, 39(4), July 2020. doi: 10.1145/3386569.3392406

[20] K. Otao, Y. Itoh, K. Takazawa, H. Osone, and Y. Ochiai. Air mounted eyepiece: Optical see-through hmd design with aerial optical functions. In *Proceedings of the 9th Augmented Human International Conference*, AH '18, pp. 1:1–1:7. ACM, New York, NY, USA, 2018. doi: 10.1145/3174910.3174911

[21] M. Papas, T. Houit, D. Nowrouzezahrai, M. Gross, and W. Jarosz. The magic lens: Refractive steganography. *ACM Trans. Graph.*, 31(6), Nov. 2012. doi: 10.1145/2366145.2366205

[22] M. Papas, W. Jarosz, W. Jakob, S. Rusinkiewicz, W. Matusik, and T. Weyrich. Goal-based Caustics. *Computer Graphics Forum*, 2011.

doi: 10.1111/j.1467-8659.2011.01876.x

[23] J. Salvi, J. Pagès, and J. Batlle. Pattern codification strategies in structured light systems. *Pattern Recognition*, 37(4):827–849, 2004. Agent Based Computer Vision. doi: 10.1016/j.patcog.2003.10.002

[24] H. Sheikh, Z. Wang, L. Cormack, and A. Bovik. Live image quality assessment database release 2, 2005.

[25] Q. Sun, C. Wang, Q. Fu, X. Dun, and W. Heidrich. End-to-end complex lens design with differentiate ray tracing. *ACM Trans. Graph.*, 40(4), July 2021. doi: 10.1145/3450626.3459674

[26] K. Tsuchiya and N. Koizumi. An optical design for avatar-user coaxial viewpoint telepresence. In *2020 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, pp. 108–116, 2020. doi: 10.1109/VR46266.2020.00029

[27] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli. Image quality assessment: From error visibility to structural similarity. *Trans. Img. Proc.*, 13(4):600–612, Apr. 2004. doi: 10.1109/TIP.2003.819861

[28] Y. Yahagi, S. Fukushima, S. Sakaguchi, and T. Naemura. Suppression of floating image degradation using a mechanical vibration of a dihedral corner reflector array. *Opt. Express*, 28(22):33145–33156, Oct 2020. doi: 10.1364/OE.406005

[29] H. Yamamoto, M. Yasui, M. S. Alvissalim, M. Takahashi, Y. Tomiyama, S. Suyama, and M. Ishikawa. Floating display screen formed by airr (aerial imaging by retro-reflection) for interaction in 3d space. In *2014 International Conference on 3D Imaging (IC3D)*, pp. 1–5, Dec 2014. doi: 10.1109/IC3D.2014.7032590

[30] S. Yamazaki, M. Mochimaru, and T. Kanade. Simultaneous self-calibration of a projector and a camera using structured light. In *CVPR 2011 WORKSHOPS*, pp. 60–67, 2011. doi: 10.1109/CVPRW.2011.5981781

[31] Y. Yoshimizu and E. Iwase. Radially arranged dihedral corner reflector array for wide viewing angle of floating image without virtual image. *Opt. Express*, 27(2):918–927, Jan 2019. doi: 10.1364/OE.27.000918

[32] C. Zhang, B. Miller, K. Yan, I. Gkioulekas, and S. Zhao. Path-space differentiable rendering. *ACM Trans. Graph.*, 39(4), July 2020. doi: 10.1145/3386569.3392383

[33] R. Zhang, P. Isola, A. A. Efros, E. Shechtman, and O. Wang. The unreasonable effectiveness of deep features as a perceptual metric. In *CVPR*, 2018.