

# Retinal Homing Display: Head-Tracking Auto-stereoscopic Retinal Projection Display

Hiroto Aoki

The University of Tokyo

Tokyo, Japan

aoki-hiroto633@g.ecc.u-tokyo.ac.jp

Yuichi Hiroi

Cluster Inc., Metaverse Lab

Tokyo, Japan

The University of Tokyo

Tokyo, Japan

y.hiroi@cluster.mu

Yuta Itoh

The University of Tokyo

Tokyo, Japan

yuta.itoh@iii.u-tokyo.ac.jp

Jun Rekimoto

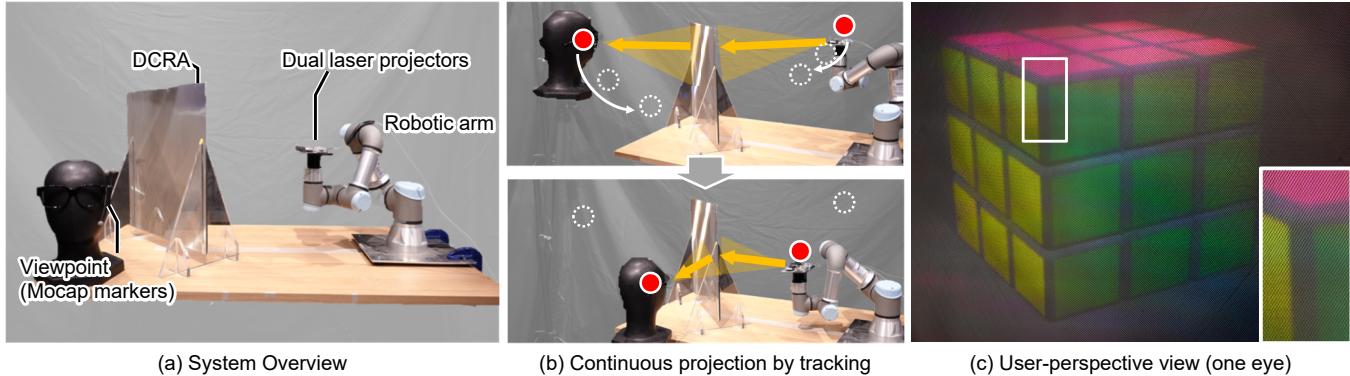
The University of Tokyo

Tokyo, Japan

Sony CSL Kyoto

Kyoto, Japan

rekimoto@acm.org



**Figure 1: Retinal Homing Display.** (a) The system consists of a robotic arm, two laser projectors mounted on the end of the arm, and a DCRA to focus the laser beam to a single point. (b) When the user whose position is being tracked moves, the robotic arm moves accordingly and continues to present images by retinal scanning. (c) An example viewpoint capture of the prototype.

## ABSTRACT

This paper introduces Retinal Homing Display, which presents focus-free stereoscopic images via retinal projection, thus eliminating the need for the user to wear additional equipment. Traditional 3D displays, typically classified as either naked-eye stereoscopic or wearable, present inherent challenges: the former involves a compromise between resolution and accurate depth perception, while the latter imposes an additional burden on the user. Our proposed display employs optical and mechanical mechanisms to converge projector light at the user's pupil center, simultaneously tracking eye movements. This lets the user perceive focus-free, high-resolution stereoscopic images without wearable equipment. We implemented a proof-of-concept system utilizing a robotic arm and a Dihedral Corner Reflector Array (DCRA), subsequently evaluating image quality and its eyebox. Finally, we discuss the limitations

of the current prototype and outline potential directions for future research.

## CCS CONCEPTS

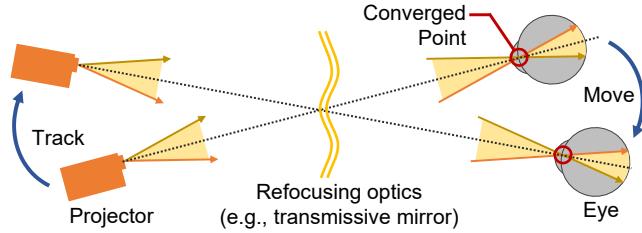
- Hardware → Emerging optical and photonic technologies;
- Computing methodologies → Graphics systems and interfaces.

## KEYWORDS

autostereoscopic display, retinal projection, motion-following display

## ACM Reference Format:

Hiroto Aoki, Yuichi Hiroi, Yuta Itoh, and Jun Rekimoto. 2024. Retinal Homing Display: Head-Tracking Auto-stereoscopic Retinal Projection Display. In *Proceedings of preprint for 29th ACM Symposium on Virtual Reality Software and Technology (preprint for VRST '23)*. ACM, New York, NY, USA, 10 pages. <https://doi.org/10.1145/nnnnnnn.nnnnnnn>



**Figure 2: Retinal Homing Display Concept.** Retinal projection with the naked eye is achieved by tracking the movement of the eyes and adjusting the position of the projector so that the projector light always converges at the center of the pupil via a certain refocusing optics in the scene.

## 1 INTRODUCTION

3D displays [11, 42] allow users to experience visual depth perception in three dimensions, providing spatial awareness and deeper immersion in the content than conventional 2D displays. Applications of 3D displays span numerous areas, including entertainment [24], medical and scientific visualization [9], design [19], and education [17]. 3D displays fall into two major types: naked-eye types and wearable types.

Naked-eye 3D displays commonly produce stereoscopic images or light-field with parallax using various optical methodologies. This display type includes auto-stereoscopic [41], volumetric [36], and holographic displays [38]. Since the naked-eye displays are free of headsets or glasses, they tend to cause less discomfort and fatigue even after prolonged use. However, reproducing the entire light field on these displays requires huge equipment or sacrifices image resolution and field-of-view (FoV), which makes it difficult to reproduce depth perception such as focus blur.

Wearable 3D displays, or head-mounted displays (HMDs), have the advantage of being able to display images directly to the user while allowing the user to be mobile. However, these wearable displays struggle with the design trade-off [13]. They require the integration of optics, batteries, and computing resources into a single headset. None of the modern high-end HMDs are light enough for users to wear for hours at a time, which can lead to user fatigue.

A Retinal Projection Display (RPD) [20] projects images directly onto the retina by converging light from the display onto the center of the pupil. Although RPDs are typically incorporated into wearable displays, their mechanism does not require a screen, allowing for naked-eye viewing [28]. In addition, RPDs have focus-free characteristics, allowing sharp images to be displayed, irrespective of the eye's focal depth. These characteristics can also be used to reproduce depth perception, such as focus blur, through post-rendering [15]. However, RPDs can easily lose images as the eye moves.

We propose a new naked-eye display concept, Retinal Homing Display, which realizes stereoscopic images via retinal projection to the viewer's eyes with a motion-following display source. Our concept aims to be an alternative 3D display method that achieves high-resolution and focus-free image presentation, which are challenges currently faced by naked-eye 3D displays.

Figure 2 shows the schematic mechanism of Retinal Homing Display. Light from the projector converges at the center of the pupil by tracking the user's eyes with optical and mechanical components. Two projectors are positioned to correspond to each eye and present parallax images to achieve stereoscopic viewing. As a result, the user can view focus-free stereoscopic retinal projection images without wearing any additional equipment.

As an initial demonstration of the retinal homing display, we implement a prototype using a robotic arm and a Dihedral Corner Reflector Array (DCRA) [22]. The robotic arm is equipped with two laser projectors. Images are presented by tracking the user's head motion and moving the robotic arm so that each projector's light source always converges on the center of the pupil.

As an initial demonstration of the retinal homing display, a prototype was implemented using a robotic arm and a Dihedral Corner Reflector Array (DCRA) [22]. Rather than in the form of retinal projection, several methods have been proposed to present aerial images while tracking the user's hand using DCRA and a movable monitor in a table-top style display [16, 23, 27]. Our system takes these configurations as a reference to realize RPD by projecting images with movable laser projectors using a robotic arm and DCRA while tracking the position of the eyes. The robotic arm is equipped with two laser projectors. The robot arm is moved so that the light source of each projector always converges on the center of the pupil, and the images are presented by tracking the user's head movement.

Our main contributions include the following:

- Proposing Retinal Homing Display, a naked-eye, stereoscopic RPD by separating the light source and computational resources from the user's head,
- implementing and evaluating prototypes of the Retinal Homing Display by a robot arm and DCRA,
- and discussing the current limitation and applications of our prototype.

## 2 RELATED WORK

### 2.1 Retinal Projection Display

RPD operates based on the principles of the Maxwellian view [40]. In this setup, collimated light from a spatial light modulator is focused at the pupil's center by a lens or another optical element, before being projected onto the retina. Conventional HMDs produce blurry images if the display's focal length doesn't match the eye's lens thickness. However, with RPD, as all light rays pass through the pupil's center, the thickness of the eye lens does not influence the light reaching the retina. Therefore, it consistently presents an image in focus, irrespective of the focal depth.

RPDs allow high-resolution, wide-field of-view images to be viewed directly with the naked eye, without the need for a screen or eyepiece optics, as long as the light converges at the center of the pupil. However, RPDs inherently have a narrow eyebox due to the need to focus a narrow beam of light onto the pupil [34].

The first prototype of wearable RPDs was introduced about 30 years ago [20], and improvements have been made continuously, such as expanding field of view [30, 37] and eyebox [5, 21, 32], or pupil steering [43]. In particular, RPDs based on holographic optical elements (HOEs), which offer a wide field of view and a thin profile,

have been the subject of extensive research [25]. As an example of an RPD combining eye tracking and HOE, Jang et al. [14] combined an HOE-based RPD with a steering mirror to achieve an extended eyebox via eye tracking. Similarly, Kim et al. [18] combined a HOE-based RPD with an LED-based HMD to implement an eye-tracking HMD with foveated rendering. Commercial wearable RPDs such as Retinal 3D (QDLaser Inc.) are also recently emerging.

Ochiai et al. [29] presented the prototype of a naked-eye RPD using a DCRA. In this system, retinal projection is achieved by optically aligning the viewer's eyes with the projector's light source. However, this system does not account for head movement, which can affect the eye box. Our Retinal Homing Display overcomes this limitation by tracking the user's head movement, thereby facilitating retinal projection for the naked eye while mitigating eyebox constraints.

## 2.2 Tracking-based Stereoscopic Display

Several studies have been conducted on stereoscopic displays that present parallax images by tracking the user's head position and orientation. In the CAVE system [6], stereo images are projected on each surface of a room surrounded by a large screen. The user wears polarized or shutter glasses equipped with position and orientation sensors, allowing the system to project stereoscopic images based on the user's eye position and orientation. This method is also used in autostereoscopic displays [31], which use special optics to present stereoscopic images without the need for glasses. For example, the SONY ELF-SR1 uses both motion parallax from eye tracking and binocular parallax from lenticular lenses to provide high-resolution stereoscopic viewing.

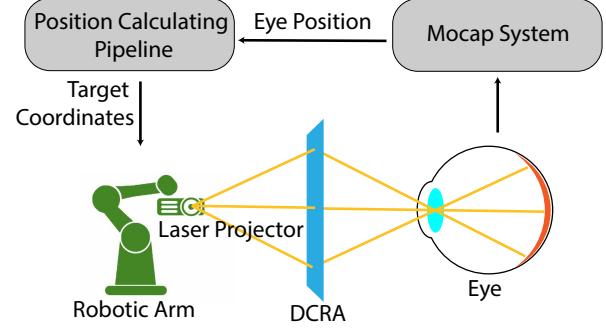
Dynamic Exit Pupil (DEP) tracker techniques [1, 3] combine LCD panels and a spatial light modulator with a Gabor super lens or a Fresnel lens to form auto-stereoscopic images on the user's eyes. Another approach combines head-tracking and pico projectors with a rotating retro-reflective screen in the scene [8]. These approaches tend to be more bulky than our design.

Beaming Display [13] is a concept for HMDs that projects stereoscopic images from environmental projectors onto binocular near-eye screens while tracking the user's head. This approach aims to mitigate the trade-off between the weight of the HMD and image quality by separating the battery, optical system, and computational resources from the user. Recently, a particularly lightweight format using HOE has been proposed [2]. Our Retinal Homing Display can also be seen as an extension of the Beaming Display, as it completely eliminates the need for user equipment through retinal projection.

## 3 SYSTEM OVERVIEW

To realize the concept of the Retinal Homing Display shown in Fig. 2, we propose a system that includes stereo retinal projection by dual laser projectors (Sec. 3.1), retinal projection to the naked eye via DCRA (Sec. 3.2), and eye tracking-based projection by a robotic arm (Sec. 3.3).

Figure 3 illustrates the overall system configuration and flow. A motion capture system tracks the head to estimate the position and orientation of the eyes. Based on this eye position and orientation data, a symmetric position with respect to the DCRA is calculated, corresponding to the optimal position for the laser source to present



**Figure 3: A schematic diagram of the system configuration and processing flow. The system tracks the user's head (eye) position and the robotic arm moves the projectors to align the projected image.**

the image to the user. The robotic arm is then instructed to move its tip to these calculated coordinates. By continuously repeating this process, the image can be presented by retinal projection in response to the user's movements.

### 3.1 Stereo Retinal Projection

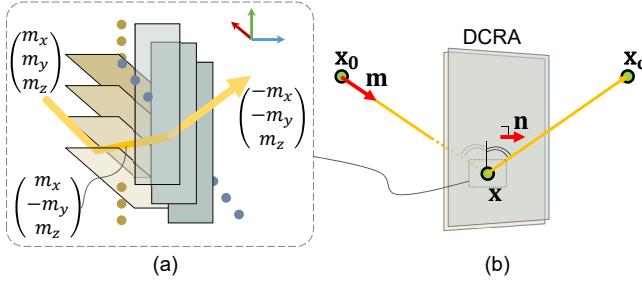
The Retinal Homing Display uses stereo laser projectors. Each laser projector combines color lasers with an optical scanner, such as a galvanometer or MEMS mirror, to create images. The projector scans the laser beam across the display in a raster pattern, illuminating one pixel at a time. After each pixel is illuminated, the angle of the optical scanner is minutely adjusted to display the next pixel. This high-speed process is repeated continuously to produce a 2D image.

The depth of field in RPDs is determined by the thickness of the light beams [7]. When laser projectors are used in RPDs, the small diameter of the laser extends the depth of field and provides the user with a focus-free image. The two laser projectors are horizontally spaced according to the user's predetermined interpupillary distance and mounted on the tip of a robotic arm.

### 3.2 Dihedral Corner Reflector Array

Laser beams produced by laser projectors diverge in different directions. Therefore, when viewed directly, the laser projector produces only a small, blurred, fragmented image on the retina. To produce a retinal scan image, each beam must pass through the pupil. Conventional near-eye displays use eyepiece optics to converge the projected light beams onto the eyeball. In contrast, our system uses a DCRA to focus the laser beam onto the pupil, allowing the user to perceive the image without the need for wearable optics [29].

As shown in Fig. 4, the DCRA behaves as a transmissive mirror [22]. The conceptual fabrication method for a DCRA is as follows: consider a series of mirrored glass (or acrylic) plates that are stacked on top of each other and then cut vertically into thin sheets. Given these two sheets, a DCRA can be obtained by rotating one sheet 90° along an axis perpendicular to its face and then gluing the two sheets together.



**Figure 4: Schematic illustration of DCRA optics.** (a) A magnified view of DCRA. Small mirrors are arranged perpendicular to each other. An incident ray, represented by a directional vector  $(m_x, m_y, m_z)^T$ , passes through the structure, undergoing reflections twice with two orthogonal surfaces, and converges at symmetrical positions. (b) The light emitted from  $x_0$  and passing through the DCRA converges at the symmetrical point  $x_c$  at any incident angle.

The laser light from the projector is expanded by the scanner, passes through the DCRA, and is then focused at a symmetrical position. The relationship between the laser light source and the focal point can be derived using a simple vector equation. As shown in Fig. 4, let  $\mathbf{x}_0$ ,  $\mathbf{x}$  and  $\mathbf{n}$  represent the position vectors of the laser light source, an arbitrary point on the DCRA, and the normal vector of the DCRA, respectively.

From this, the position vector  $\mathbf{x}_c$  of the focal point is derived as:

$$\mathbf{x}_c = \mathbf{x}_0 + 2 \frac{\mathbf{n} \cdot (\mathbf{x} - \mathbf{x}_0)}{|\mathbf{n}|^2} \mathbf{n}. \quad (1)$$

When the center of the user's pupil is at  $\mathbf{x}_c$ , the user can continue to perceive the image by controlling the robot arm to maintain the laser source position at  $\mathbf{x}_0$  according to Eq. (1).

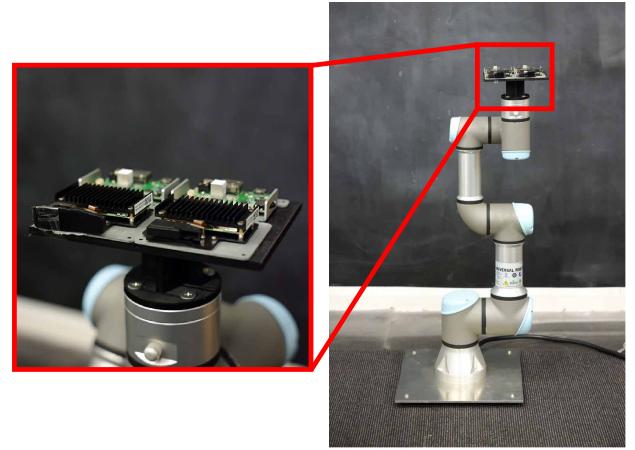
Eq. (1) assumes that the user's head and eyes are facing forward. When they rotate, the symmetry rotation vector with respect to the plane defined by the DCRA must be calculated. In general, given a rotation vector  $\mathbf{r}$ , the symmetry rotation vector  $\mathbf{r}_c$  with respect to a plane defined by the normal  $\mathbf{n}$  can be calculated as

$$\mathbf{r}_c = (\mathbf{I} - 2\mathbf{n}\mathbf{n}^T)\mathbf{r}. \quad (2)$$

The use of DCRA in the Retinal Homing Display has two significant advantages. First, alternative light-focusing methods require complex calculations to determine the focal point and may not converge when the position of the light source changes. Second, the unitary structure of the DCRA makes it easy to increase its size by integrating multiple parts. In the Retinal Homing Display, a larger relay element increases the allowable range of movement of the light source. As a result, users can increase their range of motion while maintaining the visual perception of the displayed image.

### 3.3 Eye-Tracking-based Projection

Retinal Homing Displays require eye-tracking-based projection to align the focal point of the laser beam with the center of the user's pupil. The following are the requirements for the eye-tracking-based projection in our Retinal Homing Displays:



**Figure 5: Projection hardware of our prototype.** Left: Dual laser projectors mounted for stereoscopic viewing. Right: 6-axis robotic arm.

- **Tracking Accuracy:** The diameter of the human pupil varies with brightness, averaging about 3 mm. To maintain focus, the laser position must be aligned within this range, so the tracking accuracy must be compatible.
- **6 degrees-of-freedom (DoF) eye tracking and projection:** Tracking isn't limited to pupil position but extends to pupil orientation. Even if the pupil and focus positions are aligned, any change in the laser's angle of incidence will affect its position on the retina, changing the visual image.
- **Low latency:** Because users are constantly moving, the system must respond immediately to changes in eye position and orientation. Delays in the robotic arm's response can lead to spatial misalignment, which degrades retinal image quality.

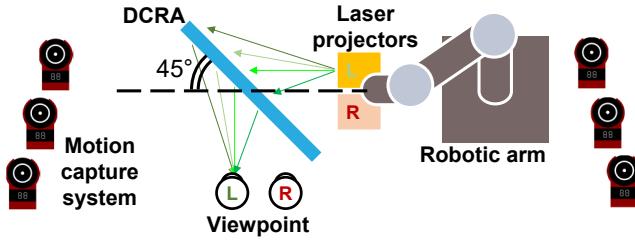
Note that achieving a low-latency system is beyond the scope of this paper, as our initial prototype aims to demonstrate its image quality. Therefore, we have assumed that the user remains stationary at the point where the head moves. Realizing a low latency system is an exciting future challenge that will be discussed in Sec. 6.2.

## 4 IMPLEMENTATION

We have implemented a proof-of-concept prototype of our retinal homing display as shown in Fig. 1. In our prototype, a laser projector is mounted at the tip of the robotic arm, and a motion capture device tracks the user's head, ensuring tracking accuracy and 6-DoF head tracking and projection. After describing the hardware implementation (Sec. 4.1), we describe the details of the tracking and projection methods (Sec. 4.2).

### 4.1 Hardware setup

We use Ultimem HD301D1 as the laser projector. This laser projector integrates three laser light sources to provide a full-color display and uses a MEMS mirror as a scanner. It provides HD resolution with vertical and horizontal viewing angles of 45° and 25°, respectively.



**Figure 6: A schematic illustration of the configuration of the system. At the user’s position, we place the viewpoint cameras for image capture. For a better visualization, we draw rays for the left-eye projector.**

The frame rate is set at 60 frames per second, with a contrast ratio (CR) of 5000 and a brightness of 20 lm. To attenuate the laser light intensity, we placed ND filters in front of the light source, resulting in a brightness level of 2 lm. We designed a fixture to attach the laser projector to the tip of the robotic arm, as shown in Fig. 5. To facilitate stereoscopic vision based on binocular disparity, we used two projectors corresponding to the left and right eyes, with a separation distance equal to the interpupillary distance.

For our robotic arm setup, we chose Universal Robots UR3e. This robotic arm provides position repeatability of  $\pm 0.03$  mm and a working range of approximately 50 cm in radius from the base. We used the dedicated PolyScope interface to program the robot arm controls. TCP/IP communications were used to transmit 3D position information relative to the base of the arm to guide the end of the arm to the desired location.

For the DCRA, we chose a square ASKA3D plate measuring 48 cm on each side and 5.9 mm of thickness. The space between mirrors was 500  $\mu$ m. We made a dedicated stand to hold the DCRA plate upright. The placement of the DCRA and the robotic arm corresponds to the configuration shown in Fig. 6. We defined a specific point as the origin and extended a line from that point to serve as the base for the robotic arm. The line was aligned with the Y-axis direction of the arm. The distance between the origin and the base of the arm was set to 80 cm. The DCRA was positioned at the origin, forming a 45° angle with the aforementioned line.

## 4.2 Tracking and Projection

The user’s head movements are tracked by a motion capture system, and based on this positional data, the robotic arm adjusts the focus of the laser passing through the DCRA to project the image. For this implementation, we used OptiTrack as our tracking system and installed six Flex13 cameras. To track the projection target, markers were placed around the user’s eyes. We also placed markers on the cameras used to capture video. We also placed markers on the DCRA and the base of the robotic arm to measure their positions within the OptiTrack coordinate system. Note that our current proof-of-concept does not account for changes in the user’s gaze direction.

We developed a Python 3 pipeline to control the arm based on the data obtained from OptiTrack. The processing in the pipeline includes obtaining the position of the projection target from OptiTrack, calculating the coordinates of the planar symmetry points

for the DCRA, and sending these coordinates to the robotic arm control program. By moving the arm according to the received coordinate information, the focal point of the laser projector is aligned with the position of the pupil center, resulting in the projection of the visual content.

## 5 EVALUATION

We present the evaluation results of our proof-of-concept system in terms of image quality, eye box, and tracking quality. We will also discuss the potential applications of this system.

### 5.1 Displayed Images

Figure 7 shows the actual projected images. As shown in fig.7 (b), our retinal display can produce high-contrast images. However, when this image is magnified, as shown in fig.7 (c), the uneven light intensity can be observed. This is because some laser beams are not directed to the focal point, depending on the mirror arrangement pattern of the DCRA.

Our prototype can present parallax images to the eye using dual laser projectors. As shown in Fig.7(d, e), there is no overlap or light leakage between the images perceived by the left and right eyes. This indicates that binocular disparity can be accommodated while maintaining the naked eye setup. We have also confirmed that our prototype can consistently project sharp images regardless of the depth of focus of the eye, as shown in Fig.7 (f, g)).

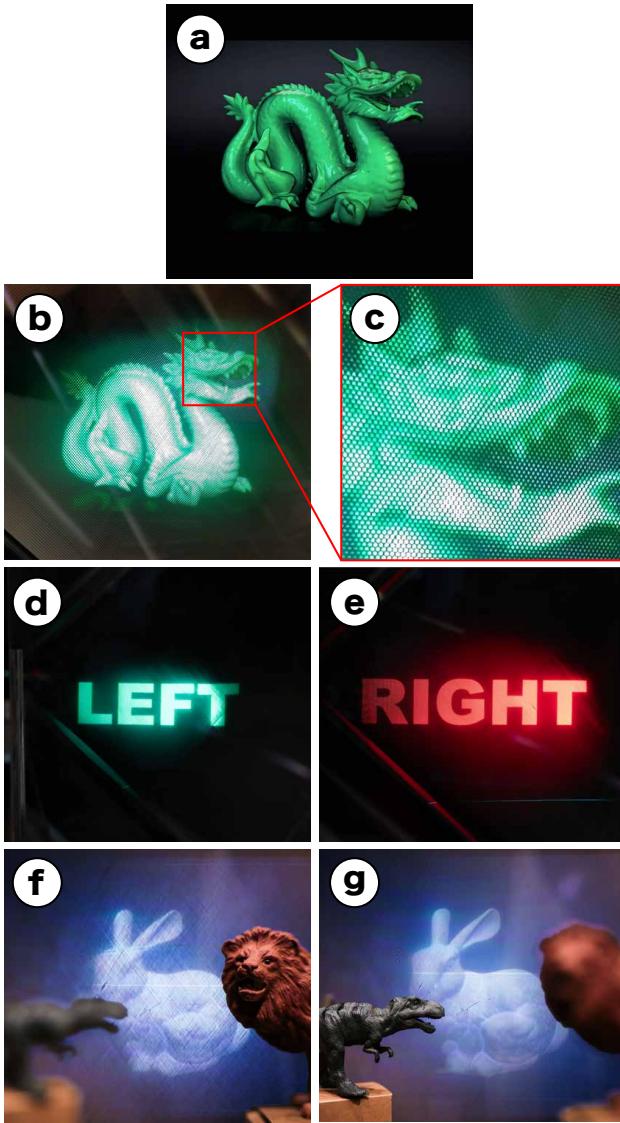
### 5.2 Spatial Resolution

To measure the resolution of the retinal homing display, we used a commonly used method, the slanted edge modulation transfer function (MTF) algorithm [4]. The resolution measurement was performed at the focal point when the incident angle of the laser light on the DCRA, which represents the central pixel of the projected image, was set to 45° and the distance between the incident point and the light source was set to 50 cm. For image acquisition, we obtained a lens modeled after a crystalline lens and an aperture of 3 mm in diameter from the schematic eye optics and then mounted it in front of the Ximea MC023CG-SY-UB camera, which served as an eye-modeling camera.

We performed MTF calculations using the images captured by the slanted edge projected by our prototype. The projected test image is shown in the top row of Fig. 8. This test image contains boxes with slanted edges which we can capture as edge profiles. After capturing projected test image, we crop regions of interest from the captured image that contain slanted edges. An example of such regions of interest are provided in the top row of Fig. 8. The middle row of Fig. 8 shows the variation from black pixels to white pixels observed in the regions of interest, which can be recognized as the edge profiles. Finally, we calculated MTF and plotted the results as shown in Fig. 8. A drop can be observed from 0 to 10 cycles per degree (cpd), which is likely caused by the structure of the DCRA.

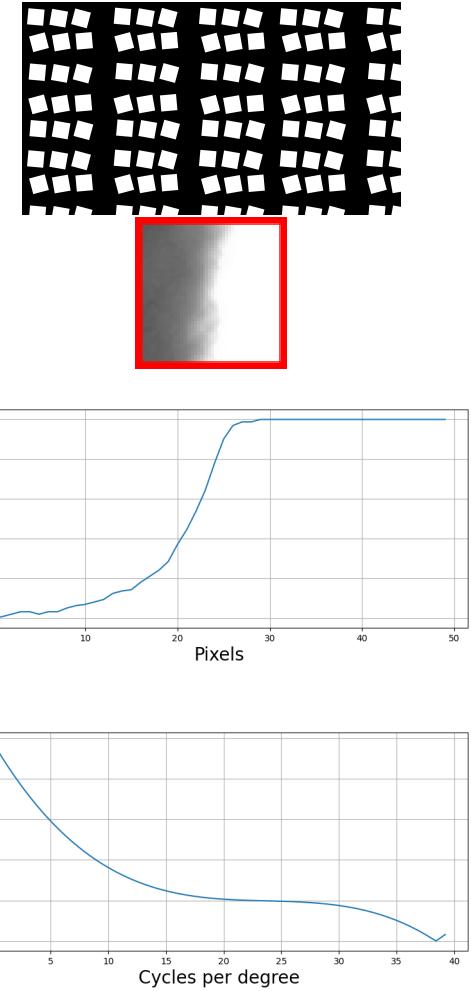
### 5.3 Incident Angle and Image Quality

In the Retinal Homing Display, laser light is focused using a DCRA. However, depending on the orientation of the laser projector, some laser beams may not pass through the focal point. Therefore, in



**Figure 7:** (a) The original image used for projection in (b) and (c). (b) A sample of the projected image using our prototype. (c) An enlarged portion of the image (b). (d,e) Two laser projectors positioned next to each other to match the inter-pupillary distance, projecting images for the left eye (d) and right eye (e). (f, g) The projected image when focused at about 15 cm away (f) and 40 cm away (g). It can be observed that the projected image has a deep depth of field.

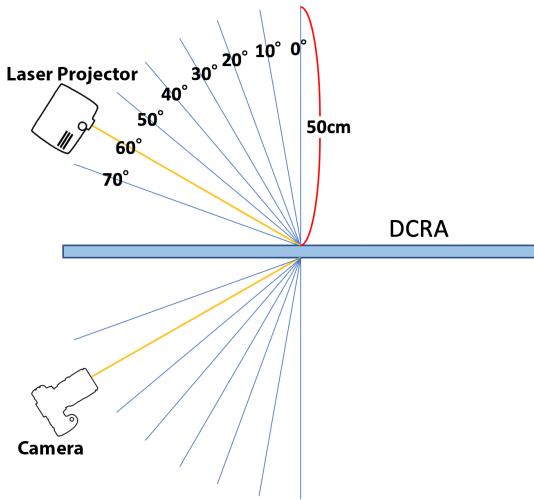
In this study, we investigate the relationship between the angle of the laser projector with respect to the DCRA and the image quality of the displayed content. To investigate the relationship between the angle of the laser projector with respect to the DCRA and the image quality, we projected images onto the DCRA at various angles and compared the captured images at each angle. The DCRA was set up perpendicular to the floor, and the incident angle of the laser beam, which corresponds to the center of the projected image, was



**Figure 8: Resolution characteristics.** Top: Test image used for slanted edge-based modulation transfer function (MTF) analysis of a sample of actual slanted edge profiles taken. Middle: Edge profiles extracted from captured, sample slanted-edge photographs. Bottom: Calculated MTFs as 7.0 cycles per degree (cpd) at half contrast. Higher the cpd, the better the image resolution.

varied from  $0^\circ$  to  $70^\circ$  in increments of  $10^\circ$ , as shown in Fig. 9. The distance between the laser projector and the projection surface on the DCRA was set to 50 cm to ensure that the projected image fit within the field of view of the DCRA, and the displayed content was captured.

The images captured at each angle of incidence are shown in Fig. 10. In addition, to investigate the degree of attenuation in light-converging efficiency as a function of the angle of incidence, we calculated the average luminance of the captured images and plotted the corresponding values, as shown in Fig. 12. From these values, we can see that the DCRA mirror is most efficiently utilized around angles of incidence of  $40^\circ$  to  $50^\circ$ . The luminance decreases



**Figure 9:** The setup for evaluating the relationship between the angle of incidence and light convergence efficiency. Projected images were captured while varying the angle of incidence from  $0^\circ$  to  $70^\circ$  in increments of  $10^\circ$ . The camera position was adjusted to correspond with each change in the angle of incidence.

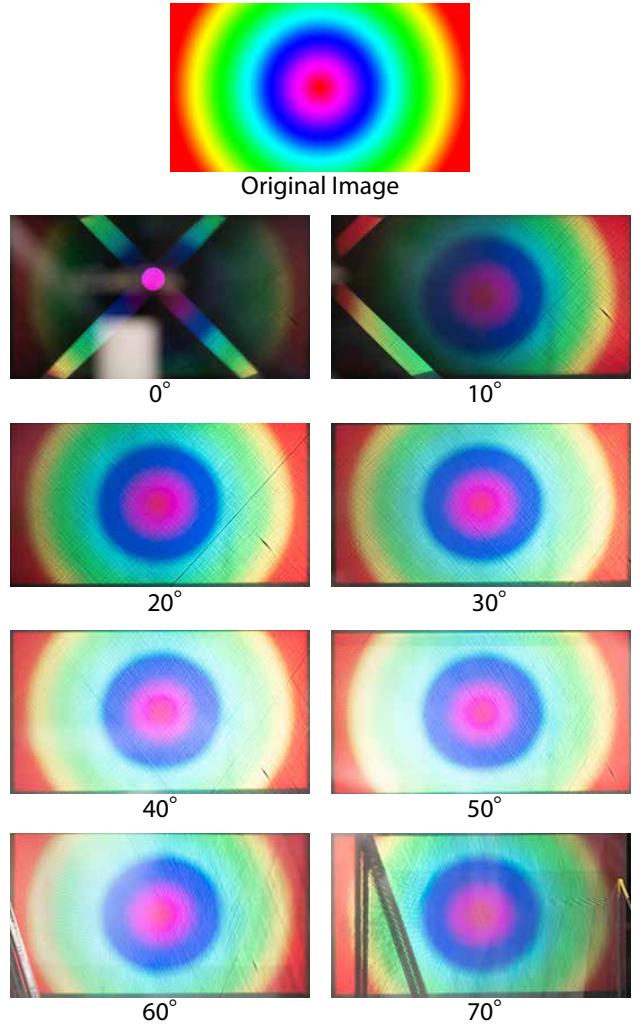
as the angle of incidence approaches  $0^\circ$  or  $90^\circ$ . This phenomenon depends on the structure of the DCRA (Fig. 4). In general, DCRAs reflect light most efficiently at an angle of incidence of  $45^\circ$ , while at  $0^\circ$  or  $90^\circ$ , light rays pass through without being reflected by the mirror.

#### 5.4 Eyebox

The RPD is characterized by a very small eye box. Even in the case of retinal homing displays, if the focal position deviates from the center of the pupil due to control error or latency of the robotic arm, the quality of the image will be degraded. Therefore, we investigated the appearance of the projected image when the focal point is displaced from the center of the pupil. First, we moved the eye modeling camera so that the center of the pupil was displaced perpendicular to the direction of the laser beam, with displacements of 2.5 mm, 5 mm, and 7.5 mm. We then acquired images with the pupil displaced in depth relative to the direction of the laser beam at displacements of 20 mm, 40 mm, and 60 mm.

Figure 11 shows the captured images. In the horizontal direction, when the displacement between the focal point and the center of the pupil was 2.5 mm, the image was almost completely obscured. Since the average diameter of the human pupil is 3 mm, this result indicates that the Retinal Homing Display has an eyebox size similar to that of a conventional RPD when the robot arm is stationary.

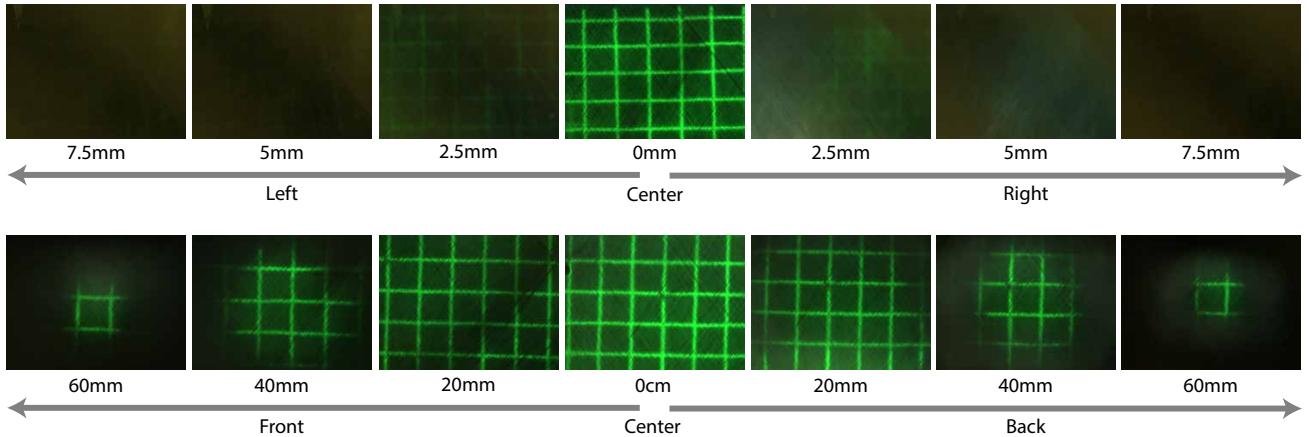
In addition, the image is more robust to depth misalignment than to vertical misalignment, although it gradually fades from the outer edge. These results suggest that image degradation can be mitigated when tracking delays occur by planning trajectories that are accurate in the vertical and horizontal directions while allowing for errors in the depth direction.



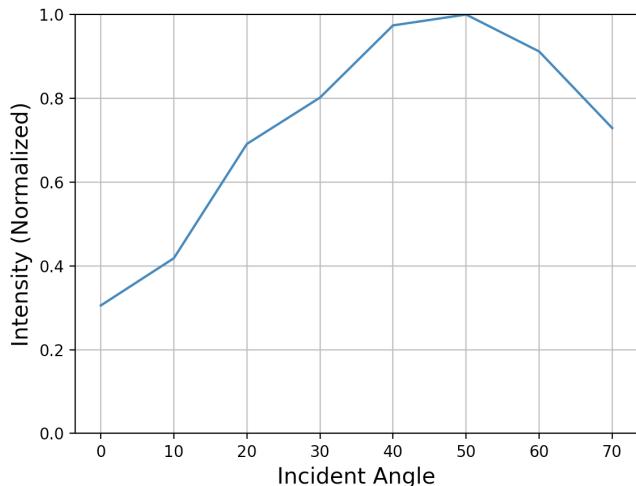
**Figure 10:** Images projected onto the DCRA with incident angles ranging from  $0^\circ$  to  $70^\circ$  at  $10^\circ$  intervals. Images are captured with the identical camera configuration including exposure time, ISO, and F value.

#### 5.5 Tracking quality

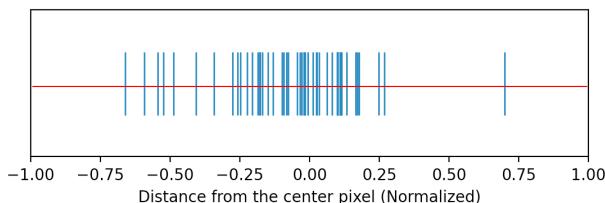
As one of the indicators to measure tracking accuracy, we evaluated the position reproducibility. Firstly, we directed the laser light that corresponds to the center of the projected image to impinge on the DCRA at an angle of 45 degrees. We projected an image with only the center being bright. We installed the aforementioned eye-modeling camera and confirmed that the pixels near the center of the captured image became brighter. Then, we randomly moved within a range of 30cm horizontally around the center, capturing the projected image each time. We performed a total of 50 captures and measured the distance in terms of the number of pixels in the horizontal direction between the brightened pixel and the center pixel. The results are shown in the Fig. 13. In all trials, we were able to keep the target within the sensor of the camera, and 78% of them were located within 25% of the image width from the center.



**Figure 11:** Top: Images taken by moving the eye-modeling camera perpendicular to the direction of laser light propagation. Bottom: Images taken by moving the eye-modeling camera in the same direction (back and forth) as the laser light propagates.



**Figure 12:** Light-converging efficiency analysis. As the incident angle approaches  $0^\circ$ , the laser beam passes through the fine mirror-to-mirror gap in the DCRA without being reflected, resulting in a decrease in the amount of laser light at the convergent point.



**Figure 13:** Horizontal distance between the center pixel of the image and the center of the projected point.

In addition to position repeatability, it is necessary to mention the latency. OptiTrack sends the head position information at 120 Hz to the pipeline that sends instructions to the robotic arm. The pipeline then sends a move instruction to the robotic arm once every 100 ms. The robotic arm (UR3e) can move 100 mm every 1 ms. This means that if the user moves 100 mm within 100 ms, there will be a maximum delay of latency 200 ms. While there is still room to increase the operating speed of position measurement systems and pipelines, the physical limitations of the robot arm's operating speed require attention when designing applications.

## 5.6 Application

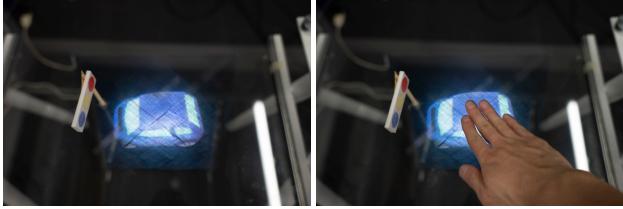
Retinal homing displays provide images with deep DoF and focus-free features, which open up various potential applications. For example, as shown in Fig. 14, if a transparent glass table is placed between the user and the DCRA, the user can work on the table while viewing the projected image. Importantly, if an object obstructs the path of the laser beam, the laser will not reach the retina, thus blocking part of the image. This means that occlusion processing, typically required by many AR displays, is not necessary for this system if the physical object is inside the projected beam frustum. However, occlusion processing is necessary in limited situations where the physical object is outside the projected beam frustum and covers the projected image in the user's field of view.

## 6 DISCUSSION AND FUTURE WORK

In this section, we discuss the limitations of our current prototype and suggest future research directions based on the evaluation.

### 6.1 Tracking Gaze Orientation

Our current prototype assumes that the head and eyeballs direct in the same direction. However, to accurately track the center of the pupil, the direction of gaze must also be tracked [33]. In addition, in the prototype, markers were attached to the eye camera for head tracking, which prevented us from achieving full auto-stereoscopic viewing.



**Figure 14: Left:** Example of projection through a transparent glass table. **Right:** Occlusion observed when the laser beam path is blocked by hand.

To solve this problem, we need to incorporate an external camera into the system that can accurately track both gaze direction and eye position. For example, Tobii Eye Tracker 5, a commercially available installation eye tracker, has an accuracy of  $0.74^\circ$  of vision angle [12], thus placing the eye tracker at a distance of about 20 cm reduces the measurement error to within 3 mm. Another possible method is to use the DCRA characteristics to detect the position of the reflected eye on the robot arm side for tracking, which can alleviate the spatial constraints of eye tracker placement.

## 6.2 Low-Latency System

The latency in our prototype is about 200 ms. This latency has a significant impact on the focal point misalignment when the user moves his/her head quickly. The total latency is the sum of the latency due to head or eye position measurement, the calculation of the target position of the robot arm tip, and the actual robot arm motion. Of these, the robot arm motion is the rate-limiting factor at 100 ms.

Combining a fast-responding beam steering device, such as a galvanometer scanner [26, 39], with a robotic arm is an effective solution to reduce this latency. For example, high-speed motion is tracked by the galvanometer scanner, while global motion is tracked by the robotic arm. Another way to compensate for latency is to predict the future trajectory of the head position, such as using a Kalman filter [10].

## 6.3 Visibility Range

In our Retinal Homing Display, the user's range of motion is limited by the range of motion of the robotic arm and the size of the DCRA surface. Ideally, a sufficiently large DCRA and a robotic arm with a large enough range of motion would increase the user's range of motion. In practice, additional beam steering mechanisms such as galvanometer mirrors can be attached to the tip of the robotic arm, or the DCRA can be moved by a mechanical mechanism to increase the visibility range.

## 6.4 Eyebox Expansion

Our prototype inherits the characteristics of the RPD with a narrow eye box at a single viewpoint, making it vulnerable to misalignment between the pupil center and the convergence point of the image. Since our display is a naked-eye type, it is difficult to apply for eyebox extension by HOEs, which has become mainstream in wearable RPDs in recent years. On the other hand, eyebox multiplexing using spatial light modulators [5, 37] or beam splitter arrays [35] may be

applicable to our prototype because the devices can be placed on the light source side of the DCRA.

## 6.5 Field of View

Since our DCRA-based optics copy rays from the projectors, the FoV of the resulting image is to be identical that of the laser projectors. This allows flexibility to expand the FoV of the system by increasing the scanning angle range of a laser projector,  $45^\circ \times 25^\circ$  in our prototype. However, one would need a bigger DCRA that can cover such diverging rays from the projector.

## 7 CONCLUSION

We proposed the Retinal Homing Display, a naked-eye display concept that offers high-resolution and focus-free image presentation. Our approach achieves stereoscopic imagery through retinal projection by tracking the viewer's eyes with a motion-following display source. The proof-of-concept prototype, equipped with a DCRA and a robotic arm fitted with dual laser projectors, enables a focus-free stereoscopic retinal projection experience without requiring the user to wear any additional equipment.

Our concept presents a novel design for a *focus-free* autostereoscopic display. In addition to providing depth cues, this focus-free feature proves valuable in situations where bifocality and low vision aid design are advantageous, such as in eye clinics and other medical and sports applications. Given the limitations of our current prototype and the direction of future research that we have outlined, we believe that future studies will explore the potential of this retinal autostereoscopic display.

## ACKNOWLEDGMENTS

This work was supported by JST Moonshot R&D Grant Number JPMJMS2012, JST CREST Grant Number JPMJCR17A3, the commissioned research by NICT Japan, The University of Tokyo Human Augmentation Research Initiative, and JSPS KAKENHI Grant Number JP21K19788 and JP20H05958, Japan.

## REFERENCES

- [1] Kaan Akşit, Hadi Baghsiahi, Phil Surman, Selim Ölçer, Eero Willman, David R Selviah, Sally Day, and Hakan Urey. 2013. Dynamic exit pupil trackers for autostereoscopic displays. *Optics express* 21, 12 (2013), 14331–14341.
- [2] Kaan Akşit and Yuta Itoh. 2023. HoloBeam: Paper-Thin Near-Eye Displays. In *2023 IEEE Conference Virtual Reality and 3D User Interfaces (VR)*. IEEE Computer Society, Los Alamitos, CA, USA, 581–591. <https://doi.org/10.1109/VR55154.2023.00073>
- [3] Kaan Akşit, Selim Ölçer, Erdem Erden, VC Kishore, Hakan Urey, Eero Willman, Hadi Baghsiahi, Sally E Day, David R Selviah, F Aníbal Fernández, et al. 2011. Light engine and optics for HELIUM3D auto-stereoscopic laser scanning display. In *2011 3DTV Conference: The True Vision-Capture, Transmission and Display of 3D Video (3DTV-CON)*. IEEE, 1–4.
- [4] Peter Burns. 2000. Slanted-Edge MTF for Digital Camera and Scanner Analysis. *Society for Imaging Science and Technology: Image Processing, Image Quality, Image Capture Systems Conference*, 135–138.
- [5] Chenliang Chang, Wei Cui, Jongchan Park, and Liang Gao. 2019. Computational holographic Maxwellian near-eye display with an expanded eyebox. *Sci. Rep.* 9, 1 (Dec. 2019), 18749.
- [6] Carolina Cruz-Neira, Daniel J Sandin, and Thomas A DeFanti. 1993. Surround-screen projection-based virtual reality: the design and implementation of the CAVE. In *Proceedings of the 20th annual conference on Computer graphics and interactive techniques (Anaheim, CA) (SIGGRAPH '93)*. Association for Computing Machinery, New York, NY, USA, 135–142.
- [7] G. C. de Wit. 1997. Resolution matching in a retinal scanning display. *Appl. Opt.* 36, 22 (Aug 1997), 5587–5593. <https://doi.org/10.1364/AO.36.005587>

- [8] Osman Eldes, Kaan Akşit, and Hakan Urey. 2013. Multi-view autostereoscopic projection display using rotating screen. *Optics Express* 21, 23 (2013), 29043–29054.
- [9] Henry Fuchs, Mark A Livingston, Ramesh Raskar, D'Nardo Colucci, Kurtis Keller, State, Andrei, Jessica R Crawford, Paul Rademacher, Samuel H Drake, and Anthony A Meyer. 1998. Augmented reality visualization for laparoscopic surgery. In *Medical Image Computing and Computer-Assisted Intervention — MICCAI'98*. Springer Berlin Heidelberg, 934–943.
- [10] Serhan Güç, Sebastian Bosse, Dimitri Podborski, Thomas Schierl, and Cornelius Hellege. 2020. Kalman Filter-Based Head Motion Prediction for Cloud-Based Mixed Reality. In *Proceedings of the 28th ACM International Conference on Multimedia*. Association for Computing Machinery, New York, NY, USA, 3632–3641. <https://doi.org/10.1145/3394171.3413699>
- [11] Nicolas S Holliman, Neil A Dodgson, Gregg E Favalora, and Lachlan Pockett. 2011. Three-Dimensional Displays: A Review and Applications Analysis. *IEEE Trans. Broadcast.* 57, 2 (June 2011), 362–371.
- [12] Andrew Housholder, Jonathan Reaban, Aira Peregrino, Georgia Votta, and Tauheed Khan Mohd. 2022. Evaluating Accuracy of the Tobii Eye Tracker 5. In *Intelligent Human Computer Interaction*. Springer International Publishing, 379–390.
- [13] Yuta Itoh, Takumi Kaminokado, and Kaan Aksit. 2021. Beaming Displays. *IEEE Trans. Vis. Comput. Graph.* 27, 5 (May 2021), 2659–2668.
- [14] Changwon Jang, Kiseung Bang, Seokil Moon, Jonghyun Kim, Seungjae Lee, and Byoungcho Lee. 2017. Retinal 3D: Augmented Reality near-Eye Display via Pupil-Tracked Light Field Projection on Retina. *ACM Trans. Graph.* 36, 6, Article 190 (nov 2017), 13 pages. <https://doi.org/10.1145/3130800.3130889>
- [15] Mayu Kaneko, Yuichi Hiroi, and Yuta Itoh. 2021. Focus-Aware Retinal Projection-based Near-Eye Display. In *2021 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct)*, 207–208.
- [16] Hajime Katsumoto, Hajime Kajita, Naoya Koizumi, and Takeshi Naemura. 2016. HoVerTable PONG: Playing Face-to-Face Game on Horizontal Tabletop with Moving Vertical Mid-Air Image. In *Proceedings of the 13th International Conference on Advances in Computer Entertainment Technology (Osaka, Japan) (ACE '16)*. Association for Computing Machinery, New York, NY, USA, Article 50, 6 pages. <https://doi.org/10.1145/3001773.3001820>
- [17] Hannes Kaufmann and Dieter Schmalstieg. 2003. Mathematics and geometry education with collaborative augmented reality. *Comput. Graph.* 27, 3 (June 2003), 339–345.
- [18] Jonghyun Kim, Youngmo Jeong, Michael Stengel, Kaan Akşit, Rachel Albert, Ben Boudaoud, Trey Greer, Jooahn Kim, Ward Lopes, Zander Majercik, et al. 2019. Foveated AR: dynamically-foveated augmented reality display. *ACM TOG* 38, 4, Article 99 (July 2019), 15 pages. <https://doi.org/10.1145/3306346.3322987>
- [19] Kiyoshi Kiyokawa, Haruo Takemura, and Naokazu Yokoya. 2000. SeamlessDesign for 3D object creation. *IEEE Multimedia* 7, 1 (Jan. 2000), 22–33.
- [20] Joel Kollin. 1993. A retinal display for virtual-environment applications. In *SID International Symposium Digest of Technical Papers*, Vol. 24. SOCIETY FOR INFORMATION DISPLAY, 827–827.
- [21] Tiegang Lin, Tao Zhan, Junyu Zou, Fan Fan, and Shin-Tson Wu. 2020. Maxwellian near-eye display with an expanded eyebox. *Opt. Express* 28, 26 (Dec. 2020), 38616–38625.
- [22] Satoshi Maekawa, Kouichi Nitta, and Osamu Matoba. 2006. Transmissive optical imaging device with micromirror array. In *Three-Dimensional TV, Video, and Display V*, Vol. 6392. SPIE, 130–137.
- [23] Sandor Markon, Satoshi Maekawa, and Ahmet Onat. 2012. Robot-Assisted Medical Visualization with Floating Images. In *2012 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining*, 810–814. <https://doi.org/10.1109/ASONAM.2012.147>
- [24] Bernard Mendiburu. 2009. *3D Movie Making* (1st edition ed.). Routledge.
- [25] Lantian Mi, Chao Ping Chen, Yifan Lu, Wenbo Zhang, Jie Chen, and Nizamuddin Maitlo. 2019. Design of lensless retinal scanning display with diffractive optical element. *Opt. Express* 27, 15 (Jul 2019), 20493–20507. <https://doi.org/10.1364/OE.27.020493>
- [26] Yuri Mikawa, Masahiro Fujiwara, Takefumi Hiraki, Yasutoshi Makino, and Hiroyuki Shinoda. 2021. Far-field Aerial Image Presentation of One Point by a Laser Source using Beam Scanning by Two-axis Galvanometer Mirror. In *2021 60th Annual Conference of the Society of Instrument and Control Engineers of Japan (SICE)*, 137–143.
- [27] Daisuke Miyazaki, Noboru Hirano, Yuki Maeda, Siori Yamamoto, Takaaki Mukai, and Satoshi Maekawa. 2013. Floating volumetric image formation using a dihedral corner reflector array device. *Appl. Opt.* 52, 1 (Jan 2013), A281–A289. <https://doi.org/10.1364/AO.52.00A281>
- [28] Yoichi Ochiai, Kazuki Otao, Yuta Itoh, Shouki Imai, Kazuki Takazawa, Hiroyuki Osone, Atsushi Mori, and Ippei Suzuki. 2018. Make Your Own Retinal Projector: Retinal near-Eye Displays via Metamaterials. In *ACM SIGGRAPH 2018 Emerging Technologies* (Vancouver, British Columbia, Canada) (SIGGRAPH '18). Association for Computing Machinery, New York, NY, USA, Article 13, 2 pages. <https://doi.org/10.1145/3214907.3214910>
- [29] Yoichi Ochiai, Kazuki Otao, Yuta Itoh, Shouki Imai, Kazuki Takazawa, Hiroyuki Osone, Atsushi Mori, and Ippei Suzuki. 2018. Make your own retinal projector: retinal near-eye displays via metamaterials. In *ACM SIGGRAPH 2018 Posters*, 1–2.
- [30] Kazuki Otao, Yuta Itoh, Kazuki Takazawa, Hiroyuki Osone, and Yoichi Ochiai. 2018. Air Mounted Eyepiece: Optical See-Through HMD Design with Aerial Optical Functions. In *Proceedings of the 9th Augmented Human International Conference* (Seoul, Republic of Korea) (AH '18, Article 1). Association for Computing Machinery, New York, NY, USA, 1–7.
- [31] Gilbae Park, Jae-Hyun Jung, Keehoon Hong, Yunhee Kim, Young-Hoon Kim, Sung-Wook Min, and Byoungcho Lee. 2009. Multi-viewer tracking integral imaging system and its viewing zone analysis. *Opt. Express* 17, 20 (Sept. 2009), 17895–17908.
- [32] Jae-Hyeung Park and Seong-Bok Kim. 2018. Optical see-through holographic near-eye-display with eyebox steering and depth of field control. *Opt. Express* 26, 21 (Oct. 2018), 27076–27088.
- [33] Bernhard Petersch and Kai Dierkes. 2022. Gaze-angle dependency of pupil-size measurements in head-mounted eye tracking. *Behavior Research Methods* 54, 2 (2022), 763–779.
- [34] Kavitha Ratnam, Robert Konrad, Douglas Lanman, and Marina Zannoli. 2019. Retinal image quality in near-eye pupil-steered systems. *Opt. Express* 27, 26 (Dec. 2019), 38289–38311.
- [35] Pawan K Shrestha, Matt J Pryn, Jia Jia, Jhen-Si Chen, Hector Navarro Fructuoso, Atanas Boev, Qing Zhang, and Daping Chu. 2019. Accommodation-Free Head Mounted Display with Comfortable 3D Perception and an Enlarged Eye-box. *Research* 2019 (Aug. 2019), 9273723.
- [36] D E Smalley, E Nygaard, K Squire, J Van Wagoner, J Rasmussen, S Gneiting, K Qaderi, J Goodsell, W Rogers, M Lindsey, K Costner, A Monk, M Pearson, B Haymore, and J Peatross. 2018. A photophoretic-trap volumetric display. *Nature* 553, 7689 (Jan. 2018), 486–490.
- [37] Yasuhiro Takaki and Naohiro Fujimoto. 2018. Flexible retinal image formation by holographic Maxwellian-view display. *Opt. Express* 26, 18 (Sept. 2018), 22985–22999.
- [38] Koki Wakunami, Po-Yuan Hsieh, Ryutaro Oi, Takanori Senoh, Hisayuki Sasaki, Yasuyuki Ichihashi, Makoto Okui, Yi-Pai Huang, and Kenji Yamamoto. 2016. Projection-type see-through holographic three-dimensional display. *Nat. Commun.* 7 (Oct. 2016), 12954.
- [39] Greg Welch, Gary Bishop, Leandra Vicci, Stephen Brumback, Kurtis Keller, and D'Nardo Colucci. 1999. The HiBall tracker: High-performance wide-area tracking for virtual and augmented environments. In *Proceedings of the ACM symposium on Virtual reality software and technology*, 1–ff.
- [40] Gerald Westheimer. 1966. The maxwellian view. *Vision research* 6, 11–12 (1966), 669–682.
- [41] Douglas F Winnek. 1951. Composite stereography.
- [42] Lin Yang, Haiwei Dong, Abdulhameed Alelaiwi, and Abdulmotaleb El Saddik. 2016. See in 3D: state of the art of 3D display technologies. *Multimedia Tools and Applications* 75 (2016), 17121–17155.
- [43] Junyu Zou, Lingshan Li, and Shin-Tson Wu. 2022. Gaze-matched pupil steering Maxwellian-view augmented reality display with large angle diffractive liquid crystal lenses. *Adv. Photonics Res.* 3, 5 (May 2022), 2100362.