Low-Latency Beaming Display: Implementation of Wearable, 133 μs Motion-to-Photon Latency Near-eye Display

Yuichi Hiroi, Akira Watanabe, Yuri Mikawa, Yuta Itoh

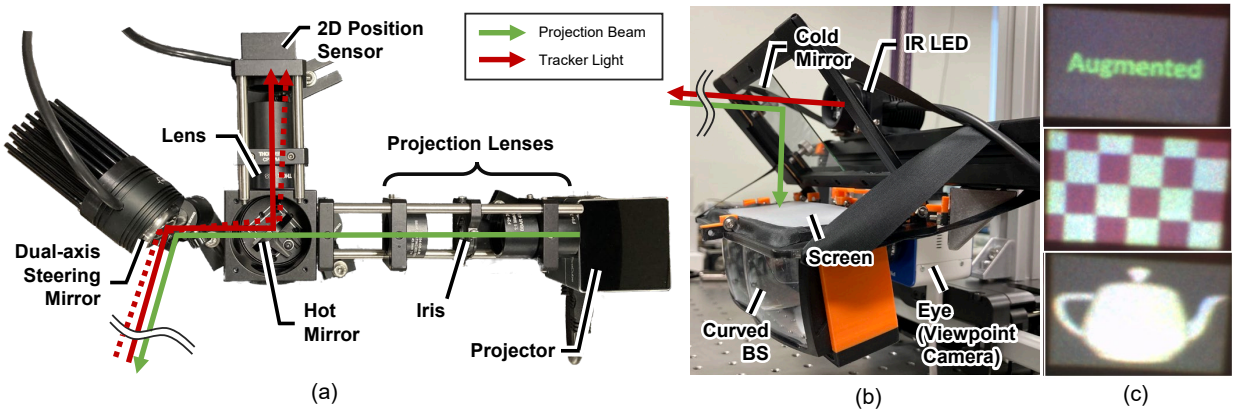

Fig. 1. Low-Latency Beaming Display prototype. (a) The steering projection module. The prototype comprises a projector and a 2D position sensor, which are aligned coaxially using a hot mirror. The light projected by the dual-axis steering mirror (indicated by the green arrow) and the incident light from the emitter on the headset (indicated by the red arrow) share the same optical path. When the headset moves, this displacement is detected by the sensor (marked by the red dotted line), and the steering mirror compensates for the displacement accordingly. (b) The passive wearable headset. The image from the projection module is beamed onto the headset's screen via a cold mirror. The tracker light (an IR LED) is directed towards the steering projector, which aligns coaxially with the screen. (c) User perspective photographs. The photographs from top to bottom depict the word "Augmented," a chessboard, and a teapot, each projected onto the headset's screen. These projected images synchronize with the movements of the headset, as confirmed by our supplemental video.

Abstract— This paper presents a low-latency Beaming Display system with a 133 μs motion-to-photon (M2P) latency, the delay from head motion to the corresponding image motion. The Beaming Display represents a recent near-eye display paradigm that involves a steerable remote projector and a passive wearable headset. This system aims to overcome typical trade-offs of Optical See-Through Head-Mounted Displays (OST-HMDs), such as weight and computational resources. However, since the Beaming Display projects a small image onto a moving, distant viewpoint, M2P latency significantly affects displacement. To reduce M2P latency, we propose a low-latency Beaming Display system that can be modularized without relying on expensive high-speed devices. In our system, a 2D position sensor, which is placed coaxially on the projector, detects the light from the IR-LED on the headset and generates a differential signal for tracking. An analog closed-loop control of the steering mirror based on this signal continuously projects images onto the headset. We have implemented a proof-of-concept prototype, evaluated the latency and the augmented reality experience through a user-perspective camera, and discussed the limitations and potential improvements of the prototype.

Index Terms—Low-Latency Display, Beaming Display, Motion-to-Photon Latency, Lateral-effect Photodiodes

1 Introduction

In augmented and virtual reality (AR/VR) applications, head-mounted displays (HMDs) play a crucial role in superimposing visual content onto the user's field of view. HMDs are designed to integrate cameras, sensors, microdisplays, device-driving circuitry, and optics into a wearable form factor. However, despite the ongoing efforts of both the research community and industry, the development of the "ultimate" HMD [44] remains elusive due to various technological trade-offs [21]. For instance, enhancing certain aspects of the HMD, such as complex optics to increase image realism, leads to a heavier headset. Similarly, while high-performance sensors can improve sensing capabilities, they also introduce more stringent power constraints.

Beaming Displays [17] (Fig. 2) offer a solution to these trade-offs by separating the image-producing electronics and optics from the HMD and positioning them within the environment. In the projection module, the optics first focus light from the projector into a beam, then the camera detects the user's head position, and finally, a steering mirror reflects the beam to the user. The user wears a lightweight, passive light-receiving headset that magnifies the beamed image using the headset's optics and displays it in the user's field of view. Consequently, the Beaming Display reduces the weight of the headset while allowing for a broad range of visual representations that utilize the powerful computing resources placed separately within the environment.

However, Beaming Displays face a significant challenge: motion-to-photon (M2P) latency [48], which refers to the latency between the user's head movement and the image displayed. The Beaming Display imposes more rigorous accuracy requirements for target tracking and

- *Yuichi Hiroi is with the University of Tokyo. E-mail: yuichi.hiroi.1@gmail.com.*
- *Akira Watanabe is with Tokyo Institute of Technology. E-mail: akira.watanabe@ar.c.titech.ac.jp.*
- *Yuri Mikawa is with the University of Tokyo. E-mail: yuri.mikawa@gmail.com.*
- *Yuta Itoh is with the University of Tokyo. E-mail: yuta.itoh@iii.u-tokyo.ac.jp.*

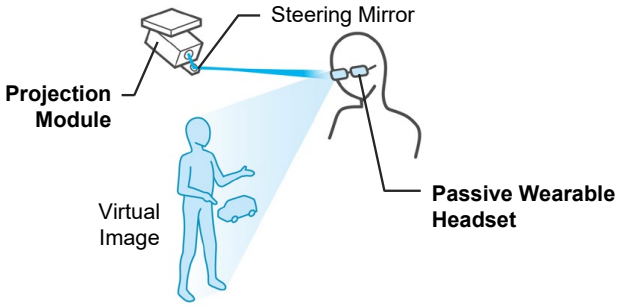

Fig. 2. The Beaming Display concept. Virtual images are projected in front of the user's eyes from an externally mounted projection module onto a passive headset worn by the user.

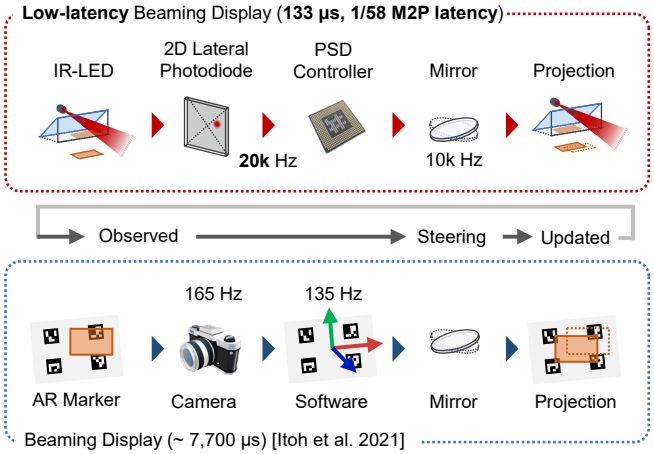

Fig. 3. Comparison of the proposed low-latency Beaming Display (top) and the conventional method [17] (bottom), regarding the process flow from the user's movement to the subsequent adjustment of the projected image. In the conventional method, the head poses estimation, which relies on image processing, acts as the rate-limiting factor and thus increases M2P latency. In contrast, we propose a Beaming Display system that significantly reduces the M2P latency from tracking to projection by incorporating analog closed-loop tracking and mirror control into the projection system.

stricter M2P latency constraints than traditional AR-HMDs because it continuously projects small, high-density images from a distance onto a few-centimeter passive screen mounted on a moving headset. As the optics of the headset magnify the incident light on the screen before the user's eye, even a slight delay in the positioning of the incident light can cause a substantial shift in the image within the field of view.

A previous study using a projected AR system [19] found that most users cannot perceive a presentation delay if the M2P latency is less than 3.2 milliseconds (ms). Therefore, to create a practical beamed display, a low-latency system that tracks and steers images to a distant light-receiving surface must be developed.

Several AR-HMD systems with low M2P latency have been proposed in previous research. Lincoln et al. developed a benchtop optical see-through (OST) HMD system using a projector and a camera, achieving a small M2P latency of 80 μs (microseconds) [22]. However, this system uses pitch and yaw encoders, limiting tracking to head orientation only and necessitating the integration of all components into the headset, resulting in a large form factor. Blate et al. proposed an outside-in tracking system that uses a two-dimensional (2D) lateral photodiode pair for head tracking and achieves a latency of 28 μs [7]. Although this study demonstrated the efficiency of the photodiode for simple tracking like dot markers, it still lacks image display capability. The original Beaming-Display system employs a coaxially aligned projector-camera system but did not account for M2P latency in a tracking module [17].

In this paper, we propose a low-latency Beaming-Display system that incorporates a 2D lateral effect photodiode (LEPD) within a steering projector unit to reduce M2P latency, as shown in Fig. 1. This photodiode detects the center of gravity of the incident light and interprets user motion as displacement. By controlling the steering mirror based on this displacement, the system can project images while tracking the user's head position from a distance. We also present a system design with a light source for low-latency tracking. Our proof-of-concept prototype evaluation shows a sampling rate of 20 kHz and an average M2P latency of 133 μs.

The advantage of our system is that the analog data of detected X/Y positions can be directly used for control, enabling a fast closed-loop process without the need for intervening digitization processes such as image processing (Fig. 3). The Beaming Display originally only requires detecting several points like infrared light point sources and dot markers, for which 2D-LEPD is sufficient, and a camera producing high-quality images is no longer necessary. Since 2D-LEPDs are significantly cheaper than a high-speed camera, they will contribute to the future modularization of projection optics for multi-unit distributed beam displays.

Please note that our current tracking framework is limited to vertical and horizontal 2 degrees of freedom (DoF). However, with the use of multiple projection units and light sources, our system can achieve precise 6 DoF tracking, including head orientation and depth, within a large tracking volume (as discussed in Sec. 6.3). This requires cooperative control of multiple units, which is beyond the scope of this paper, as it involves multi-unit calibration and synchronization, which are distinct from the low-latency topic. The primary focus of this study is to demonstrate the concept of beaming displays with *low-latency* tracking and projection approaches. Our proof-of-concept system identifies areas for improvement and future research directions.

Contributions Our main contributions include the following:

- Providing a Beaming-Display system where tracking, steering, and projection are completed in a low-latency closed-loop control.
- Implementing a proof-of-concept system for a low-latency beaming display that operates in 133 μs M2P latency.
- Discussing the remaining challenges in realizing a more practical low-latency beaming display system and outlining the future research directions.

2 Related Work

2.1 Low-latency System for AR

AR is based on two primary components: tracking the position of the real environment and displaying the virtual content precisely on the tracked position. [40, 46]. M2P latency in AR is known to cause registration errors in virtual images [4, 15], negatively affecting image realism, user comfort, and task performance [24, 30]. Jerald et al. investigated perceptual sensitivity to latency in a projected AR system and found that users can detect latencies ranging from 3.2 ms to 60.5 ms [19]. This study evaluated the M2P latency experienced when a user views an image projected on a wall screen. In contrast, the Beaming Display projects images from a distance onto an approximately 40 mm $\times$ 40 mm square moving screen, resulting in larger registration errors due to M2P latency. Currently, no method exists to measure the latency threshold at which humans perceive registration errors for Beaming Display. Therefore, Beaming Display with the lowest possible M2P latency is needed to investigate such a threshold.

End-to-end M2P latency includes tracking, application, rendering, display, and other latencies [19, 51]. Approaches to reduce temporal registration errors include latency minimization [36] and compensation through image processing [27, 38] and predictive tracking [3, 9, 10, 41]. Since latency compensation can only be applied within a limited time range, we will focus on latency minimization, particularly for tracking latency.

One well-known approach to achieving low-latency AR tracking is to use an imaging system, such as high-speed cameras up to 1,000 fps and event cameras [2] over 10,000 Hz. However, these require time for image capturing, which is the inverse value of the acquisition frequency; it takes 1 ms for a camera of 1,000 fps, which is more than we expect for the Beaming Display. Rolling shutter cameras [5, 6], while more affordable than global shutter cameras, can provide a high acquisition frequency; research reports that they have a latency of 1.5 ms and a sampling rate of 129.6 kHz [8]. However, image distortion due to line scanning can sometimes be problematic for high-speed tracking.

An example of a low-latency, built-in sensor is 3rdTech's HiBall™ tracker [50]. Despite being over two decades old, it remains one of the highest-quality trackers in terms of the sampling rate (750 $\sim$ 2000 Hz) and tracking latency ($\sim$ 3 μs). HiBall detects position by projecting light onto a lateral effect sensor mounted on the ceiling. Blate et al. expanded HiBall by proposing a method to achieve 6 DoF head tracking in 28 μs using two ceiling-mounted 2D lateral sensors [7]. However, they left the image presentation for a future study, and their method limits the trackable spatial range to the extent that light can reach the 2D lateral sensor.

We found that Blate et al.'s approach could be compatible with Beaming Display for three reasons. First, in Beaming Display, the tracker and projector are coaxially aligned; therefore, if the tracker and mirror are controlled with low latency, the displayed image will also follow the mirror with low latency. Also, steering the mirror allows for head tracking over a wider area since the tracker only needs to detect the differential signal. Finally, it is compatible with modularization because it can be implemented with off-the-shelf components.

Complete low-latency, tracking-to-display OST-AR display systems have primarily been coupled with sensor-based head tracking [18, 39]. Lincoln et al. demonstrated a benchtop OST-AR display coupled with rotary encoder head tracking with an average M2P of 80 μs [22]. While these studies track 2 DoF rotational head motion (yaw-pitch), our display is the first OST-HMD to track low-latency 2 DoF translation (vertical-horizontal) and can be expanded to 6 DoF tracking. In addition, our system is highly wearable because the sensors and processor are removed from the headset.

2.2 Dynamic Projection Mapping

The characteristics of the Beaming Display, which projects images onto moving objects, are similar to those of dynamic projection mapping (DPM). DPM systems have matured to manage the entire process from tracking to display at very high frame rates, tracking rotation, motion, and shape changes of objects, thanks to a high-speed steering mirror system of coaxial optical path [35] and a high-speed (1,000 fps) and low-latency (minimum 3 ms) projector [49]. The latest DPM system also includes novel tracking methods like marker-based [26, 29, 42] and IR-based markerless methods [14, 28]. In these studies, the M2P latency criterion is considered to be an average of 6.04 ms [31], where the user perceived no latency in experiments using a high-speed touch display presenting images at the user-touched position.

Beaming Display tracking does not require as much sophistication as DPM, which includes tracking the object's 3D shape, whereas the Beaming Display only needs to detect the orientation of the pose. Like our DPM system, Mikawa et al. proposed an aerial display system that uses two-axis galvanometer mirrors to scan a laser to the user's pupil to realize an aerial display for distant users [25]. This system measures the eye pose with an IR camera, controls the mirror according to the eye pose, and projects the laser spot into the eye with an average M2P latency 1.57 ms. Our configuration replaces the camera in this study with a 2D lateral sensor and uses analog closed-loop mirror control to achieve a lower latency image presentation.

3 Low-Latency Beaming Display

Figure 4 illustrates the system configuration of the proposed low-latency Beaming Display. As in the original [17], the low-latency Beaming Display system consists of a projection module (Fig. 4, left) within the environment and a passive wearable headset that moves around the room (Fig. 4, right). An IR LED attached to the headset emits IR light

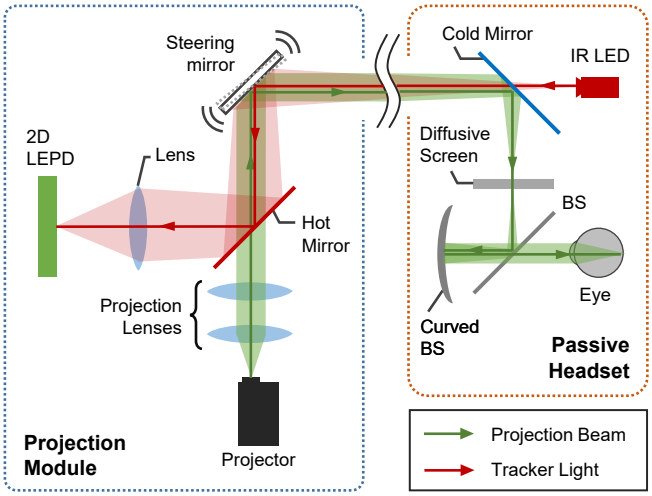

Fig. 4. Schematic diagram of our low-latency Beaming Display. The green arrows indicate the path of the visible projection light, and the red arrows indicate the path of the IR light for tracking. BS refers to the beam splitter.

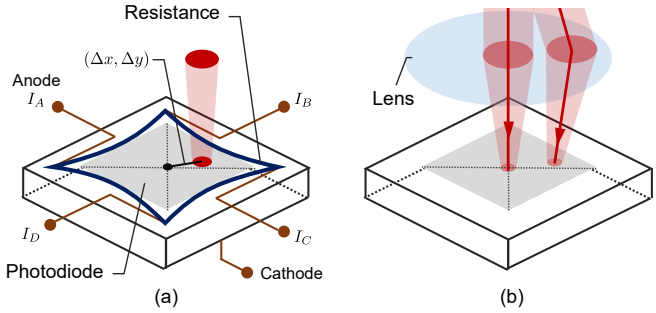

Fig. 5. Schematic diagram of a 2D LEPD. (a) The LEPD used is a tetra-lateral sensor with a layer of common cathode and anode resistors on both sides of the sensing area. By using these resistive elements to partially distribute the photo-current, the LEPD provides position information about the centroid of the spot, regardless of the size of the spot diameter of the incident light. (b) By placing a lens in front of the LEPD, the directional or positional shift of the incident light can be converted into the displacement of a small diameter spot on the photodiode.

onto the projection module. The LEPD detects the displacement of the incident light and controls the steering mirror so that the incident light always enters the center of the LEPD. The projector, placed coaxially with the LEPD, beams the image to the headset according to the mirror's angle.

Note that the current headset system is not strictly "passive," as it requires an IR light source on the headset and a power supply for light projection. We also prototyped a retroreflective low-latency Beaming Display design that includes an IR light source in the projection module and completely eliminates power from the headset, discovering that there are trade-offs to its practical implementation. This limitation will be discussed in Sec. 6.1.

The remainder of this section describes the configuration of the projection module and the headset.

3.1 Projection Module

The projection module consists of a 2D LEPD, a steering mirror, a projector, and a controller that generates difference signals for automatic beam steering. To separate IR light for tracking and visible light for projection, the 2D LEPD and projector are coaxially aligned by a hot mirror. The following will describe the tracking, mirror steering, and

projection methods.

Tracking Figure 5 provides an overview of a 2D LEPD. The 2D LEPD converts light incident on a flat photodiode into an electric current and measures the photocurrent at anodes attached to the four corners. The centroid of the incident light is then calculated from the ratio of the photocurrents at the four points. Finally, the difference between the center of the diode and the calculated centroid is output as a voltage signal (Fig. 5, a). The differential signal generated by the LEPD is sent to the controller. By placing a lens in front of the LEPD, any change in the direction or position of light incidence can be detected as a shift from the center (Fig. 5, b). This lens placement also reduces the spot diameter of the sensing light and improves the sensitivity and accuracy of the LEPD.

Mirror Steering Based on the differential signal from the LEPD, the steering mirror is controlled so that the spot is always in the center of the LEPD. Specifically, the differential voltage signal from the LEPD is converted to a mirror control voltage signal within the controller by PID parameters baked into the controller's firmware.

Projection The light from the projector is collimated by the projection lens and then directed to the headset by the steering mirror. Ideally, the projection light should be perfectly collimated to reach the headset, but this requires precise alignment. The projection lens is currently positioned to focus around the depth at which the headset is moving, partly due to the 2 DoF setup. If 6 DoF tracking becomes feasible (Sec. 6.3), a focus-tunable lens will be added to the projection lens to achieve depth focusing, as in the original [17].

3.2 Passive Wearable Headset

The wearable headset consists of an IR LED for tracking and a diffusive screen for image projection coaxially arranged by a cold mirror. Birdbath optics display the projected image on the diffusive screen in front of the eye. The headset can use any near-eye display optics instead of birdbath optics as long as the IR LED and the screen are coaxial. For example, recently proposed paper-thin holographic optical elements can be used [1].

4 Implementation

This section describes the implementation of a proof-of-concept prototype of the low-latency Beaming Display, divided into a projection module (Sec. 4.1) and a headset (Sec. 4.2). We also provide design considerations for implementing the system, including the theoretical trackable range of the headset and the safety concerns surrounding the light intensity of IR LED.

4.1 Projection Module

Figure 1 (a) shows the implementation of the projection module. The following describes each component that makes up the projection module.

4.1.1 Position Detector and Controller

Position Detector We used Thorlabs PDP90A as the 2D LEPD. This 2D LEPD can handle beam diameters up to 9.0 mm and outputs analog signals with a bandwidth of 150 kHz, depending on the difference between the center of the spot and the center of the photodiode. This LEPD also outputs the sum signal of the light intensity incident on the sensor. We used Thorlabs AC254-045-A-ML (f = 45 mm) as the lens to be placed in front of the LEPD.

Controller We used Thorlabs KPA10 as the controller for the 2D LEPD. This controller takes the difference signal from the LEPD as input and outputs an analog signal to control the mirror. It operates by switching between two modes: open-loop mode and closed-loop mode. In the open-loop mode, it uses the photodetector signal to measure the beam position on the sensor. In the closed-loop mode, the controller outputs a 20 kHz analog signal to control the beam steering element according to the measured beam position. If the sum signal of the light intensity detected by the LEPD falls below a threshold in the closed-loop mode, tracking is considered to have failed. At this point, the controller automatically switches to open-loop mode. During the open-loop mode, the steering mirror is controlled to scan the space so that the position of the IR LED can be captured again. If this switching threshold is high, the tracking performance will degrade for fast movements, but if it is low, stray light is more likely to be falsely detected. We empirically set this threshold to 2 % of the maximum value the controller can set.

PID Gains on Controller The output of the controller is determined by feedback control using a PID control system. PID control uses the proportional, integral, and derivative components of the error between the current system output and the desired setpoint to calculate the input value for the next step. Each component of the PID control system is associated with a gain parameter that can be adjusted to tune the system for optimal control. In this particular setup, we have set the gain values for the proportional, integral, and derivative components to 0.7, 0.02, and 1.00, respectively.

4.1.2 Steering Mirror

We used the Optotune MR-E-2 as our steering mirror. which has an effective aperture of 15 mm, a sampling rate of 10 kHz, and a step resolution of 22 μrads. This mirror was driven by a signal from the controller to ensure that the incoming light always reaches the center of the LEPD. While the mirror can rotate ± 25 ° in a single pan or tilt direction, the amount of rotation is limited when it is rotated in both pan-tilt directions.

In this steering mirror, the tangent of the rotation angle is proportional to the input voltage. Assuming that the upper limit of the voltage applied to the mirror is $V_{\max} = 5$ V and the maximum rotation angle is $\theta_{\max} = 25^{\circ}$, the input voltage V_{in} for desired rotation angle $-\theta_{\max} \leq \theta \leq \theta_{\max}$ can be calculated as $V_{\text{in}} = V_{\max} \cdot \tan\theta \ / \ \tan\theta_{\max}$.

In practice, the steering mirror cannot rotate to the target angle immediately after receiving the input. Therefore, the steering mirror control also uses feedback through PID control. The default gain parameters of the steering mirror for PID control are set as follows: proportional gain to 20, integral gain to 0.03, and derivative gain to 800.

4.1.3 Beaming Projector

A beaming projector consists of a commercial projector (FeliCross Pico Cube X, 1920 $\times$ 1080 pixels), projection lenses (Thorlabs MAP103030-A, f1 = f2 = 30.0 mm, and Thorlabs AC254-045-A-ML, f = 45 mm) to collimate projection beam, and an iris (Thorlabs SP8D). This beaming projector is positioned perpendicular to the 2D LEPD using a hot mirror (Thorlabs M254H45). The projection lenses were positioned so that the beaming image was focused at a distance of 40 $\sim$ 60 cm from the system, forming an image with a deep depth of field.

4.2 Passive Wearable Headset

Figure 1 (b) shows the implementation of our passive wearable headset. We disassemble the Lenovo Mirage AR see-through headset and harvest the birdbath optics. The birdbath optics contains a beam-splitting mirror and a beam combiner for each eye. A rear projection diffusive screen is placed on top of the birdbath optics, where the user can see while wearing the headset. The active area on the diffusive screen is 30 mm $\times$ 20 mm per eye.

A 940 nm IR LED (Thorlabs M940L3) is placed above this diffusive screen for tracking and is connected to a power supply (Thorlabs LEDD1B). This IR LED can output a maximum current of 1 A and a maximum output light intensity of 19.1 μW/mm^2. We powered this light source with a 0.2 A current, thus the output light intensity is 3.82 μW/mm^2. The IR LED and diffusive screen are aligned coaxially by a cold mirror (Thorlabs FM03R, 25 mm $\times$ 36 mm). The cold mirror transmits IR light and reflects visible light. Therefore, the image emitted from the projection module is reflected onto the diffusive screen with $>$ 90% efficiency. The image reaching the diffusive screen passes through the beam splitter and beam combiner of the birdbath optics to the user's eyes. The entire headset, excluding cables, weighs 225 grams, of which 130 grams is the weight of the IR LEDs.

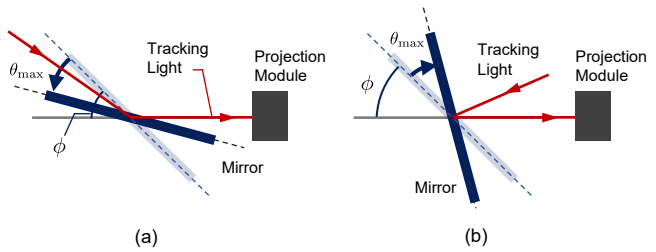

Fig. 6. Schematic of the trackable angle for the mirror's range of motion. (a) The mirror is at a minimum angle, and (b) the mirror is at a maximum angle.

4.3 Trackable Volume

We estimate the tracking volume of the prototype. Figure 6 shows the trackable range with respect to the mirror steering angle $\theta_{\max}$ and its installation angle ϕ. To operate the mirror so that the tracking light always comes to the center of the LEPD, the reflected light from the mirror must enter the projection module perpendicularly. From the figure, the angle ψ at which the tracking light can move along a single axis is

$$2(\phi - \theta_{\max}) \leq \psi \leq 2(\phi + \theta_{\max}). \quad (1)$$

Hence, when the headset is at depth l, the range r over which the headset can move along the axis is

$$\frac{l}{\tan(2(\phi + \theta_{\max}))} \leq r \leq \frac{l}{\tan(2(\phi - \theta_{\max}))}. \quad (2)$$

Note that the steering mirror we used accepts the same maximum input angle in all axes, i.e., θ lies within a circle of radius $\theta_{\max}$. Therefore, although Eq. (2) calculates the range of motion as the movement in one axis, it can be applied in any axis direction.

In our prototype, we set $\theta_{\max} = 25°$ and $\phi = 45°$. Substituting these into Eq. (2) and $l = 1.0$ m, we obtain $|r| \leq 1.192$ m. Hence, the left-right tracking range r' is obtained by $r' = 2r = 2.384\ l$. This tracking volume is significantly larger than the volume reported by the previous low-latency head tracking system (0.12 m $\times$ 0.12 m $\times$ 0.25 m) [7].

The trackable depth range l depends on the light intensity of the IR LED. In general, the higher the light intensity, the larger the trackable depth range, while also causing more hunting at the viewpoint position. For our IR LED (3.82 μW/mm^2) and threshold settings, tracking was possible at about $l = 2.0$ m.

4.4 Safety of Tracking Light Source

The current prototype uses diffuse IR LEDs for the tracking marker. Since this IR light emitted from the headset is invisible, if there are multiple users, the IR light may enter the eyes of the other user without them being aware of it. According to the International Commission on Non-Ionizing Radiation Protection (ICNIRP), conditions hazardous to the human eye should not exceed 10 mW/cm^2 for prolonged exposure ($>$ 1000 s) [37]. The maximum power of the infrared LED used in this system is 0.382 mW/cm^2, well below this threshold. Furthermore, since the IR LED is diffuse, not all of this light intensity enters the eye.

5 Evaluation

5.1 Displayed Image

First, we evaluated the quality of the video displayed on the headset with this system. As shown in Fig. 7, we placed the headset 0.4 m away from the projection module in this evaluation. We installed a user-view camera (Ximea MQ013CG-ON, 1264 $\times$ 1016 pixels) at the eye position to capture these user perspectives. We set the exposure time of the user viewpoint camera to 150 ms (66.7 fps) to approximate the image observed by the eye.

Figure 1 (c) shows the images observed from the user's perspective. From the viewpoint image, it can be confirmed that the AR image is perceptually plausible at rest, even though it is finely oscillating due to hunting (Sec 5.2.3), as confirmed by the supplemental video.

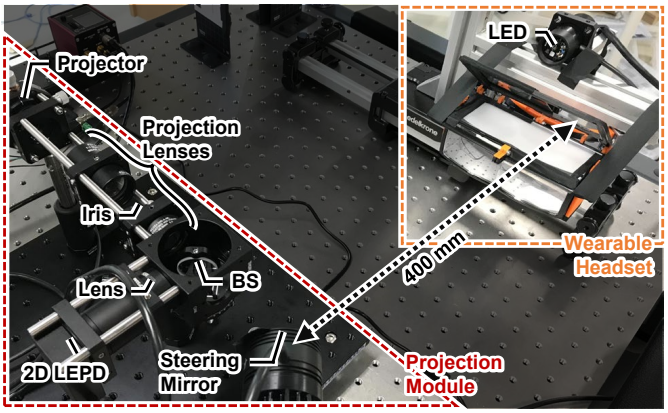

Fig. 7. Relative scale of the projection module and the headset in the experiment in Sec. 5.1.

5.2 M2P Latency

Next, we evaluated the M2P latency of this prototype. In this experiment, the latency was evaluated by analyzing the time offset between the motion of a point projected by the system and a reference point marked on the screen. Specifically, we estimated the latency by calculating the time correlation between the two trajectories of the projected and reference points, following [16]. We describe the latency evaluation setup and metrics, then we discuss the evaluated latency.

5.2.1 Experiment Setup

A schematic of the experimental setup for delay evaluation is shown in Fig. 8, and the actual implementation is depicted in Fig. 9. We replaced the projector of the projection module with a green laser pointer (520 nm). This green laser illustrates the projected point.

An LED is attached to the linear slider (edelkrone SliderONE v2) to transmit a stable horizontal movement of the passive headset to the detector. Furthermore, we placed a blue laser pointer (488 nm) on the opposite side of the LED and mounted it on the slider. Since this blue point acts as a reference point that unaffected by the system, the displacement of the green point relative to the blue point indicates the effect of M2P latency.

Instead of a small screen on the passive headset, we introduced a larger screen that spanned the entire width of the slider to capture these points. The fixed screen was made of A4 paper attached to an acrylic plate and was positioned vertically. The screen was placed 500 mm away from the steering mirror and set 100 mm away from the emitting part of the LED. To capture the two laser spots, we placed a Ximea MQ013CG-ON camera with a 1280 $\times$ 158 pixels region of interest (exposure time 395 μs, 820.23 fps, 1.22 ms per frame) behind the screen. For evaluation, after binarizing the captured video, we calculated the centroid of each spot as the position of the point.

We manually moved this slider rapidly and regularly left and right for approximately 7 seconds and recorded the movement of the spots with the camera. This process was repeated 20 times. Before the trial, we placed a ruler on the screen and measured the correspondence between the camera pixels and the actual size, which was 0.15 mm/pixel. After the trial, we calculated the velocity of the slider from the motion of the blue spot, which was 563 $\pm$ 29 mm/s.

Note that this prototype is theoretically capable of horizontal and vertical 2 DoF tracking. However, due to the symmetric optical design of the system in both dimensions, we evaluated the system with a slider that moves only horizontally.

In the headset, the tracking LED and the screen were aligned coaxially. However, if a similar alignment were applied to this experimental setup, the camera would not detect the green spot because it is in the same position as the LED. To avoid this, the green spot was placed

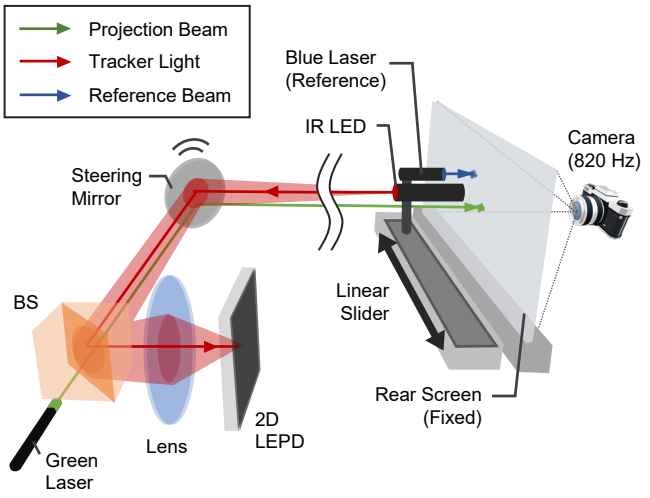

Fig. 8. Schematic diagram of the experimental setup for latency evaluation. Green arrows indicate the path of the green laser that replaces the projector, blue arrows indicate the path of the blue laser that corresponds to the slider movement, and red arrows indicate the path of the IR LED for tracking.

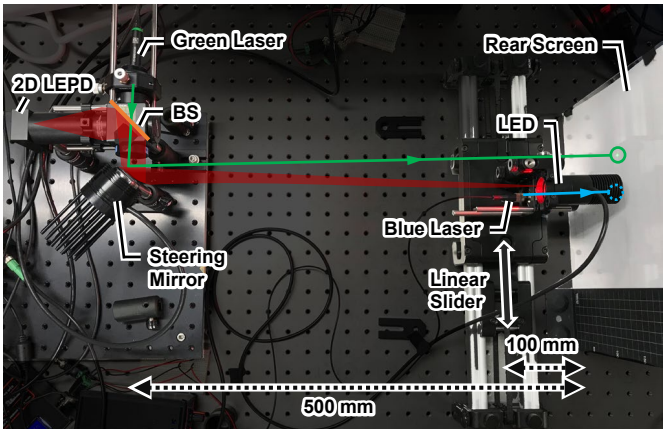

Fig. 9. Hardware implementation for the latency evaluation. The schematic is identical to Fig. 8.

slightly above the blue spot. Since this experiment evaluated only horizontal motion, we could ignore this vertical offset. As the depth of the slider and the screen are different, there was also an offset in the horizontal direction, depending on the position of the LEDs. This horizontal offset can also be ignored when evaluating the latency because we estimate latency from the correlation of the motion trajectories of the two spots.

5.2.2 Latency Estimation Metrics

We used the correlation method of Huber et al. [16] to estimate the time difference between two motion trajectories. The following describes the computation of the latency estimation metrics in this system.

We define the latency of the entire system as Δt. Assume that all devices are co-axially aligned, i.e., the tracking LED (red) and the reference light (blue) occupy the same position. In this situation, the position of the laser pointer controlled by the system (green) would coincide with the position of the reference light at Δt seconds earlier. We express this relationship as

$$x(t+\Delta t) = y(t), \tag{3}$$

where $x(t)$ is the coordinate of the green spot at time t, and $y(t)$ is the coordinate of the blue reference spot at time t.

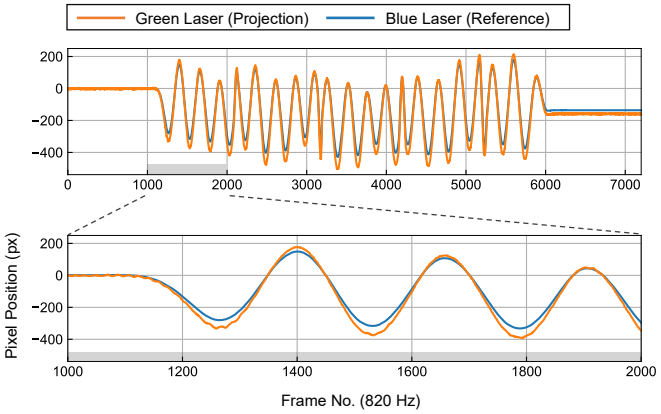

Fig. 10. Changes in pixel positions of blue and green spots as the slider is moved left and right. (top) displacement of spot positions for the entire sequence, (bottom) 1000 frames enlarged.

In practice, projection errors $e(t)$ occur between $x(t+\Delta t)$ and $y(t)$ for several reasons: offset due to non-coaxial optics, changes in the distance between the steering mirror and the LED, and depth difference between the screen and the LED. Thus, Eq. (3) can be replaced by,

$$x(t+\Delta t) = y(t) + e(t). \tag{4}$$

Since time-varying misalignment can cause re-projection errors, $e(t)$ is not constant and challenging to estimate. To estimate the time offsets between two signals containing such spatio-temporal variations, a time delay estimation method has been developed. We use normalized cross-correlation, a common similarity metric to compute the time offset between two motion patterns.

Let $T = \{t_k\}$ be the set of sampling times of the camera. In this case, we define each trajectory as a sequence of points: $X = \{x(t_k)\}, Y = \{y(t_k)\}$. We can then calculate the normalized cross-correlation similarity $S(X,Y)$ as,

$$S(X,Y) = \frac{\mathrm{Cov}(X,Y)}{\sigma_X \sigma_Y} \tag{5}$$

where $\mathrm{Cov}(X,Y)$ is the covariance between X and Y and σ_X, σ_Y are the standard deviations of X and Y.

Our goal is to find the true latency Δt_{true} from the observation. Ideally, the latency Δt in Eq. (3) is constant. However, the observed Δt includes non-constant variation since the camera discretely samples the points. Hence, we search for Δt such that the two trajectories have the highest similarity and treat such Δt as Δt_{true}. Let $X' = \{x(t_k + \Delta t)\}$ be the trajectory on the projector side with delay compensation. We estimate Δt_{true} as,

$$\Delta t_{\mathrm{true}} = \arg\max_{\Delta t} S(X', Y). \tag{6}$$

We compute Eq. (6) using a brute-force approach. First, we apply spline interpolation to trajectory X, Y because each trajectory's sampling interval is larger than the actual latency. Then, we calculate $S(X', Y)$ by increasing Δt in increments of 0.001 ms.

5.2.3 Results

Figure 10 plots the displacement versus time of the green and blue spots for a single trial. From the figure, it can be seen that the green spots track well against the blue spots throughout the entire movement. When we zoom in on the 1000-frame region of the plot, we can also see that the trajectory of the green spot oscillates at a high frequency. This hunting is caused by the PID control parameters and affects the image quality. We will evaluate this hunting in detail in Sec. 5.3.

The M2P latency was calculated for 20 trials by finding the Δt at the maximum correlation coefficient between the two trajectories. In each trial, the correlation coefficient S at Δt_{true} was more than 0.99985. The

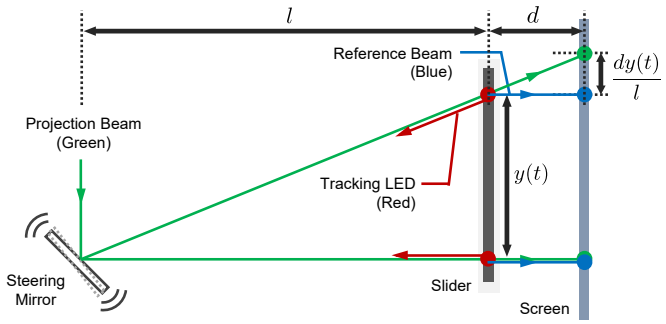

Fig. 11. Schematic of the horizontal offset in our measurement setup depending on the position of the LED.

results show that the average M2P latency was 0.133 ± 0.008 ms. This result confirmed that the M2P latency was reduced to 1/58, and the temporal response of the projection to headset motion was significantly improved.

5.3 Tracking Error

Finally, based on the estimated latency, we evaluated the tracking error of our prototype. While the main contribution of this paper is the system design and its low-latency performance, the analysis of the current tracking error motivates the discussion of future system improvements.

5.3.1 Experiment Setup

We used the same dataset as Sec. 5.2.1 to evaluate tracking errors. When evaluating errors, the horizontal offset created by the depth between the slider and the screen cannot be ignored.

Figure 11 shows a schematic of this horizontal offset. From the figure, if we denote the depth between the slider and the screen as d, we can derive the actual tracking error $e_{\text{track}}(t)$ from the observed error $e(t)$ and the reference (blue spot) trajectory $y(t)$ as

$$e_{\text{track}}(t) = e(t) - \frac{dy(t)}{l}. \tag{7}$$

As a preprocessing step, after we calculated $e(t)$ by Eq. (4), we computed the offset by Eq. (7). In our prototype, since we set $l = 500$ mm and $d = 100$ mm, $e_{\text{track}}(t) = e(t) - y(t)/5$.

5.3.2 Required Accuracy for Image Localization

Prior to the analysis, we computed the error necessary to localize the image at a millimeter scale. In this context, we define the required accuracy as the error within which an image must remain within a single pixel. Using the active range of the diffusion screen (Sec. 4.2) and the resolution of the projector (Sec. 4.1.3), we established the required accuracy per pixel scale as 0.0164 mm/px on the diagonal.

5.3.3 Results

Figure 12 (a) shows $e_{\text{track}}(t)$ from our experiment. According to this figure, we have confirmed that our prototype can track images with an accuracy ranging from -1.68 mm to 2.58 mm. Converting these errors to pixels using the given scale (mm/px), the range is from -102 px to 157 px, equating to up to 14 % (1/12) of the horizontal span of the image. However, this error includes hunting, which could be potentially reduced by fine-tuning the control parameters as discussed in Sec. 6.2.

Figure 12 (b) shows the frequency spectrum of the fast Fourier transform of the result in Fig. 12 (a). From Fig. 12 (b), we confirmed that the frequency component of the hunting begins to appear around 27.5 Hz and peaks around 42.5 Hz.

To obtain a tracking error that eliminated the effects of hunting, we applied a fourth-order Butterworth low-pass filter to remove frequency components below 27.5 Hz from the error trajectory. A comparison of the tracking error through the low-pass filter and the reference light trajectory is represented by the orange line in Fig. 12. The maximum

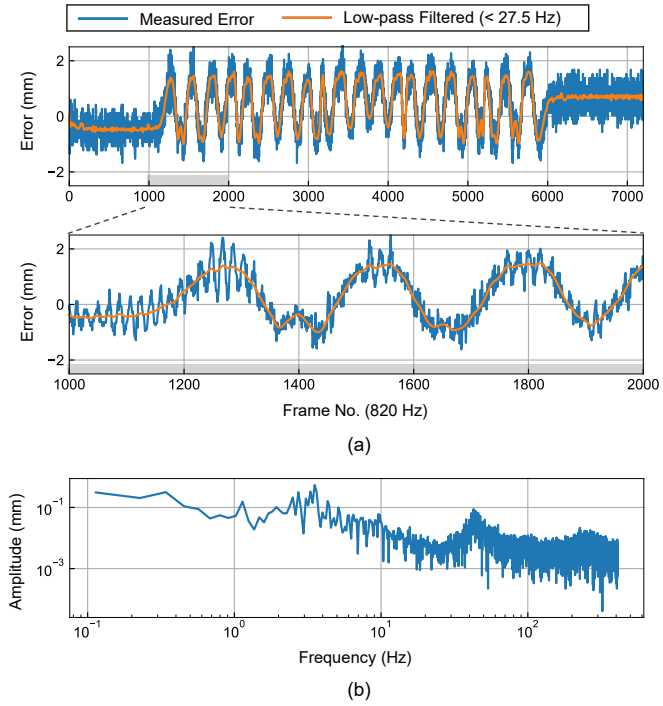

Fig. 12. Tracking error in the data shown in Fig. 10. (a) (top) Error for the entire sequence, and (bottom) 1000 frames enlarged. (b) The frequency spectrum of the error in (a), showing the effect of hunting with a peak at 42.5 Hz.

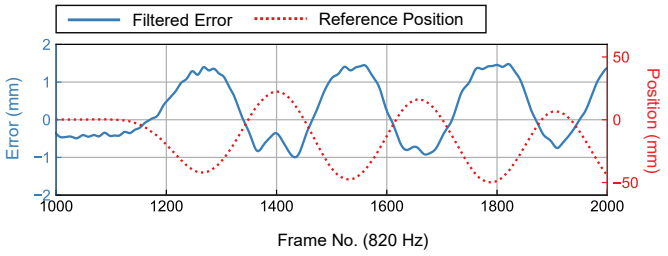

Fig. 13. Comparison of the tracking error through the 27.5 Hz low-pass filter (blue line) and the reference light position (red dashed line). From the entire sequence, we enlarge 1000 frames as in Fig. 12.

discrepancy between the error through this low-pass filter and the error including hunting was 1.29 mm. From this, we deduce that 30 % of the tracking error is due to hunting.

The tracking error after applying the low-pass filter ranges from -1.00 mm to 1.63 mm, or -61 px to 99 px in pixels. This indicates that up to 9 % of the field of view can be shifted horizontally. Note that this result does not include the artifact of the green laser being injected at a slight angle to the mirror, thus the actual error is anticipated to be smaller.

Based on the above results, while the current prototype may be suitable for applications that do not necessitate precise positioning, such as text display, it falls short for AR displays requiring spatial consistency. However, considering that the latency of the original Beaming Display is 7.7 ms, and using the slider speed (563 mm/s) in the experiment, the error introduced by the latency would be up to 4.34 mm. As a result, the original beaming display would add 266 px in pixels to the current tracking error (99 px), which suggests that our low-latency beaming display is more akin to a practical system than the original one.

In PID control, tracking errors can be caused not only by hunting but also by overshooting or undershooting. Figure 13 indicates that the

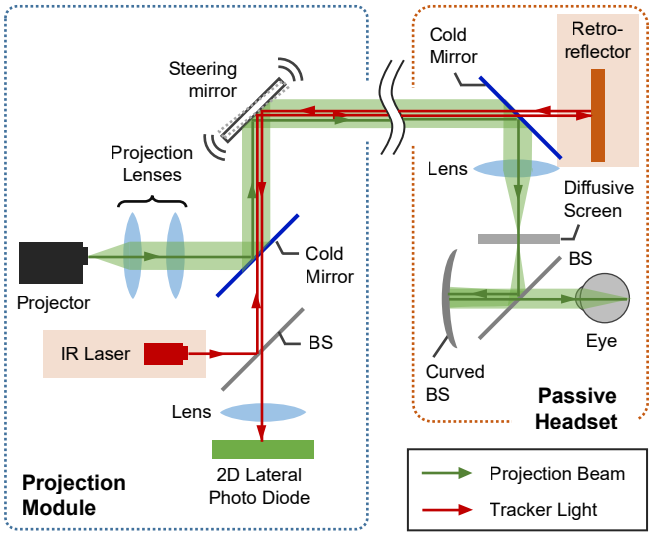

Fig. 14. Schematic of the retro-reflective low-latency Beaming-Display design. The light orange background indicates a change in the design from Fig. 4.

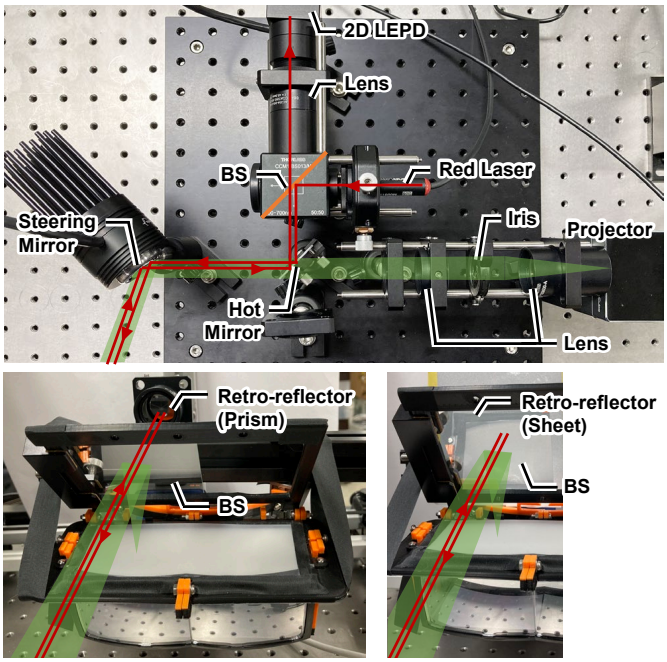

Fig. 15. Hardware implementation of the retro-reflective low-latency Beaming Display. (top) A projection module with the tracking light, (bottom left) a headset with a prismatic retro-reflector (bottom right) and a sheet retro-reflector.

prototype's control undershoots the mirror in a compensatory direction against the motion. Though this paper primarily focuses on reducing latency rather than improving tracking error, as discussed in Sec. 6.2, careful tuning of PID parameters based on response observations could significantly reduce tracking errors.

6 Discussion and Future Works

We have proposed a prototype to reduce the M2P latency of the Beaming Display significantly. This prototype can serve as a foundation for more sophisticated near-eye AR displays and raise interesting research questions for understanding or augmenting human vision using low-latency near-eye AR images in the future. In this section, we discuss improvements to our prototype and identify future research directions.

6.1 Retro-reflective Design for Low-Latency Display

The Beaming Display aims to provide comfortable wearability by leaving only the eyepiece optics in the headset. However, the prototype shown in Fig. 1 has yet to fully achieve this goal, as the IR LED and power supply remain inside the headset.

To solve this issue, we designed an improved low-latency Beaming-Display in which the headset contains only the eyepiece optics by integrating the tracking IR light into the external projection module (Fig. 14). An IR laser is coaxial inside the projection module with the 2D LEPD via a beam splitter. The light emitted from this IR laser is retro-reflected by a retro-reflector on the headset. The reflected light is returned to the projection module and reaches the LEPD. Since there is a latency between the movement of the retro-reflector and the control of the mirror, the displacement is detected on the LEPD [32].

Figure 15 shows our early prototype of the low-latency retro-reflective Beaming Display. Unlike LEDs, IR lasers produce a high amount of light per unit area, raising concerns about eye safety. According to IEC-60825-1, the international standard for the safety of laser products, the light intensity affecting the eye is 2 mW/cm^2 when 800 nm IR light is incident on the eye for more than 10 seconds. Since Class 1 lasers have a maximum intensity of 0.8 mW/cm^2 at 800 nm, the eye is unaffected when using this laser class. In this implementation, we used a Class 1 red laser pointer as a tracking light to visualize the operation.

For the headset, we initially used an internal total reflection prism (Thorlabs PS975, ϕ = 25.4 mm) as a retro-reflector (Fig. 15, bottom left). This prism can reflect red light with 92% efficiency. However, due to the small aperture of the prism, if the mirror overshot or hunting occurred even slightly, or if the headset and mirror moved away from the facing direction, the tracking light did not enter the prism and tracking was lost. In particular, when the headset is far away from the projection module, slight mirror control deviations affect tracking. Similar to Sec. 5.2.1, we placed the headset with the prism on the slider and moved it, confirming that tracking is lost at a speed of about 20 mm/s. To increase the aperture of the retro-reflector, we built another prototype by attaching a commercially available retro-reflective sheet to the top of the headset (Fig. 15, bottom right). However, tracking performance is similarly degraded because such sheets scatter the reflected light, and less tracking light reaches the LEPD.

From the prototype, in addition to the advanced control law (Sec. 6.2), we found the following design space needs to be considered for future retro-reflective low-latency Beaming-Display design:

- Trade-off between efficiency and aperture of retro-reflector(s). Ideally, a laser reflector array with aligned prismatic retro-reflectors, such as those used on lunar landers [43], should be used. It is also important to evaluate the tracking volume as these parameters change and optimize the design for each application.
- Trade-off between light intensity and eye safety. While increasing the amount of light allows the headset to be tracked from a distance, IR lasers have a greater impact on the eye than LEDs. One possible direction of research is to investigate how far directional, eye-safe light sources, such as superluminescent diodes (SLD), can be used. Another direction is to look for novel tracking methods using pulsed lasers.

6.2 Control Law for Stable Tracking and Projection

Consideration of better controller and mirror control laws is crucial for improveing tracking and projection stability. We have already observed the hunting of the projection point with self-luminous LEDs, which degrades the quality of the projected image. Also in Sec. 6.1, we confirmed that a better tracking control law is essential to realize a more practical retro-reflective design.

Hunting and undershooting are primarily caused by the PID gains. In our prototype, the PID gains are determined manually, but there is still room for tuning by observing the response with a high-speed camera or designing appropriate physical models. For example, Nguyen et al. [33]

treats the motion of a retro-reflector on a linear slider as a dynamic friction model and proposes a feed-forward tracking controller using LEPD. Another direction is to model head motion and incorporate it into the controller.

Starting from the two-decades-old HiBall tracker [**?**], there are approaches to predict the motion several steps ahead and use it for tracker control. To implement these prediction-based methods on analog signals, a prediction model such as a Kalman filter must be implementable at the circuit interface [47].

If the degradation of the projected image caused by hunting is so small as to be imperceptible, it may be possible to accept some degree of hunting while increasing temporal responsiveness. Although perceptual metrics exist for temporal and spatial visual sensitivity to wide-field video [23], to our knowledge, no study has defined similar perceptual metrics for DPM. If such studies of perceptual metrics for projection images are developed, it may be possible to control the perception of visual deception.

6.3 6 DoF Tracking

Although the current prototype has achieved low-latency Beaming-Display design, tracking is limited to 2 DoF. For the continuous projection of images to users moving freely in space, 6 DoF tracking is essential. Here, we explore the possibility of extending the current prototype to facilitate 6 DoF tracking.

Blate et al.'s tracking method, which inspired our design, uses four IR emitters and two 2D LEPDs to achieve 6 DoF tracking (Fig. 16, a). In their method, the IR emitters flash sequentially, and each LEPD measures the light spot received. From these measured signals, the 3D poses are calculated on the FPGA using Sutherland's work [45]. Their method requires synchronization signals between the LED driver and multiple 2D LEPDs. Therefore, applying this technique to Beaming Display necessitates wiring between the projection module and the headset, which could limit the user's range of motion and degrade the quality of experience.

A promising approach to achieve 6 DoF tracking while expanding the tracking volume involves placing multiple projection modules in the space (Fig. 16, b). In this approach, the user's 3D pose is estimated by triangulation of multiple modules, similar to motion capture devices. In this case, given the form of the LED mounted on the headset, there is a possibility that the tracking light may not be received by multiple modules, depending on the directional angle of the LED. Therefore, a retro-reflective design of the Beaming Display is desirable for implementation. Moreover, it will be necessary to derive a method to calibrate the relative poses of multiple modules and explore an algorithm and circuit interface to receive the angles of multiple LEPDs and steering mirrors as analog signals and estimate the 3D pose. Such calibration and coordination control among multiple modules will be a topic for future research.

6.4 Image Warping

The current prototype does not compensate for the deformation of the projected image as the viewpoint moves. While existing methods for low-latency OST-HMD [22] also evaluate the M2P latency without compensating for such deformation, image warping is necessary to improve the spatial consistency of the projection further. Lincoln et al. [22] discusses a theoretical staged rendering cascade for handling unrestricted warps, which could be applied to our system.

6.5 Applications

Our low-latency Beaming Display is theoretically capable of presenting high-quality rendered images, using unlimited computational resources, to the eyes of users moving freely in space without latency. This feature could be applied to evaluating the threshold for perceived latency in near-eye displays [20]. It can also be utilized for realistic appearance renderings, such as gloss and transparency, depending on the viewpoint. DPM has matured techniques for rendering such realistic appearances at high frame rates [14, 28, 34]. By implementing these techniques in low-latency projection systems, it is possible to modulate the viewpoint-dependent visual appearance of moving objects, such as clothes and balls, for each user, even in environments where multiple individuals coexist in the same physical space [11]. Lastly, our system is expected to implement deep neural network-based appearance reproduction [12] or vision augmentation [13], which is challenging to achieve with stand-alone OST-HMDs due to computational resource limitations.

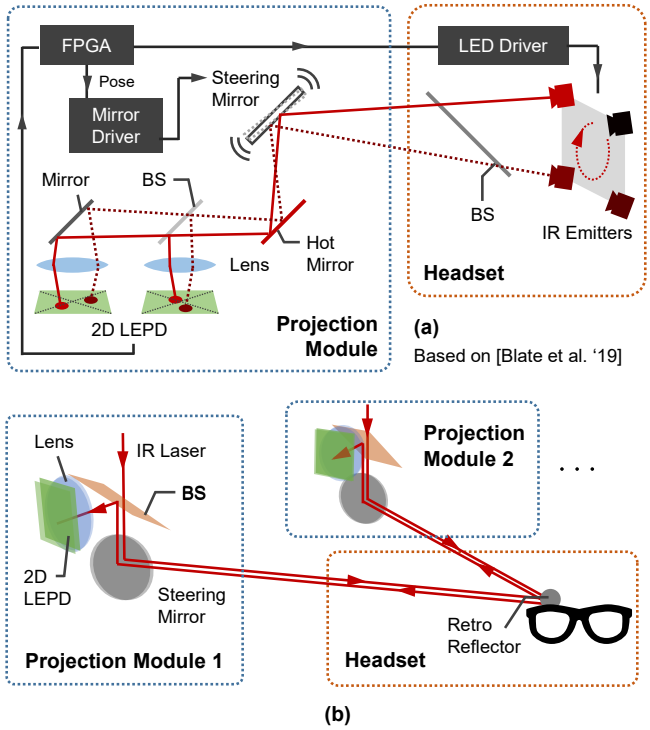

Fig. 16. Examples of future 6 DoF tracking implementations in low-latency beaming displays. (a) 6 DoF tracking with reference to [7]. The projection module and the headset must be wired to synchronize the emitter signals. (b) 6 DoF tracking with multiple projection modules.

7 Conclusion

This paper presented the low-latency Beaming-Display system that addressed the limitations of existing OST-HMDs by using a steerable projector and an analog position detector. The proposed system was able to present users with temporally consistent images by mitigating the M2P latency inherent in projecting images onto a distant, moving wearable headset. The system incorporated a LEPD that detected LED markers on the wearable headset and generated a position difference signal controlling the steering mirror. Our proof-of-concept prototype demonstrated the M2P latency of 133 μs, significantly improving over conventional Beaming-Display designs. Furthermore, we discussed potential improvements and applications for more practical low-latency Beaming displays. We believe that our first attempt has opened up a new direction for realizing temporally natural, all-day wearable AR displays and will inspire further studies in this field.

Acknowledgments

This project was partially supported by JST FOREST Grant Number JPMJFR206E, and JSPS KAKENHI Grant Number JP22J01340, and JP20H05958, Japan.

References

[1] K. Akşit and Y. Itoh. Holobeam: Paper-thin near-eye displays. *arxiv*, Jan. 2023.

[2] A. N. Angelopoulos, J. N. P. Martel, A. P. Kohli, J. Conradt, and G. Wetzstein. Event-Based Near-Eye gaze tracking beyond 10,000 hz. *IEEE Trans. Vis. Comput. Graph.*, 27(5):2577–2586, May 2021.

[3] R. Azuma and G. Bishop. A frequency-domain analysis of head-motion prediction. In *Proceedings of the 22nd annual conference on Computer*

graphics and interactive techniques, SIGGRAPH '95, pp. 401–408. Association for Computing Machinery, New York, NY, USA, Sept. 1995.

[4] R. T. Azuma. A survey of augmented reality. *Presence*, 6(4):355–385, Aug. 1997.

[5] A. Bapat, E. Dunn, and J.-M. Frahm. Towards Kilo-Hertz 6-DoF visual tracking using an egocentric cluster of rolling shutter cameras. *IEEE Trans. Vis. Comput. Graph.*, 22(11):2358–2367, Nov. 2016.

[6] A. Bapat, T. Price, and J.-M. Frahm. Rolling shutter and radial distortion are features for high frame rate multi-camera tracking. In *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 4824–4833. IEEE, June 2018.

[7] A. Blate, M. Whitton, M. Singh, G. Welch, A. State, T. Whitted, and H. Fuchs. Implementation and evaluation of a 50 khz, 28μs Motion-to-Pose latency head tracking instrument. *IEEE Trans. Vis. Comput. Graph.*, 25(5):1970–1980, May 2019.

[8] J. C. Dibene, Y. Maldonado, L. Trujillo, and E. Dunn. Prepare for ludicrous speed: Marker-based instantaneous binocular rolling shutter localization. *IEEE Trans. Vis. Comput. Graph.*, 28(5):2201–2211, May 2022.

[9] Q. Guimard, L. Sassatelli, F. Marchetti, F. Becattini, L. Seidenari, and A. D. Bimbo. Deep variational learning for multiple trajectory prediction of 360° head movements. In *Proceedings of the 13th ACM Multimedia Systems Conference*, MMSys '22, pp. 12–26. Association for Computing Machinery, New York, NY, USA, Aug. 2022.

[10] S. Gül, S. Bosse, D. Podborski, T. Schierl, and C. Hellge. Kalman filter-based head motion prediction for cloud-based mixed reality. In *Proceedings of the 28th ACM International Conference on Multimedia*, MM '20, pp. 3632–3641. Association for Computing Machinery, New York, NY, USA, Oct. 2020.

[11] T. Hamasaki, Y. Itoh, Y. Hiroi, D. Iwai, and M. Sugimoto. HySAR: Hybrid material rendering by an optical See-Through Head-Mounted display with spatial augmented reality projection. *IEEE Trans. Vis. Comput. Graph.*, 24(4):1457–1466, Apr. 2018.

[12] Y. Hiroi, Y. Itoh, and J. Rekimoto. NeARportation: A remote real-time neural rendering framework. In *Proceedings of the 28th ACM Symposium on Virtual Reality Software and Technology*, number Article 23 in VRST '22, pp. 1–5. Association for Computing Machinery, New York, NY, USA, Nov. 2022.

[13] Y. Hiroi, T. Kaminokado, A. Mori, and Y. Itoh. DehazeGlasses: Optical dehazing with an occlusion capable See-Through display. In *Proceedings of the Augmented Humans International Conference*, number Article 3 in AHs '20, pp. 1–11. Association for Computing Machinery, New York, NY, USA, June 2020.

[14] S. Hisaichi, K. Sumino, K. Ueda, H. Kasebe, T. Yamashita, T. Yuasa, U. Lippmann, P. Aswendt, R. Höfling, and Y. Watanabe. Depth-Aware dynamic projection mapping using high-speed RGB and IR projectors. In *SIGGRAPH Asia 2021 Emerging Technologies*, number Article 3 in SA '21 Emerging Technologies, pp. 1–2. Association for Computing Machinery, New York, NY, USA, Dec. 2021.

[15] R. L. Holloway. Registration error analysis for augmented reality. *Presence*, 6(4):413–432, Aug. 1997.

[16] M. Huber, M. Schlegel, and G. Klinker. Application of time-delay estimation to mixed reality multisensor tracking. *Journal of Virtual Reality and Broadcasting*, 11(3), 2014.

[17] Y. Itoh, T. Kaminokado, and K. Aksit. Beaming displays. *IEEE Transactions on Visualization & Computer Graphics*, 27(05):2659–2668, may 2021. doi: 10.1109/TVCG.2021.3067764

[18] Y. Itoh, J. Orlosky, M. Huber, K. Kiyokawa, and G. Klinker. OST rift: Temporally consistent augmented reality with a consumer optical see-through head-mounted display. In *2016 IEEE Virtual Reality (VR)*, pp. 189–190, Mar. 2016.

[19] J. Jerald and M. Whitton. Relating Scene-Motion thresholds to latency thresholds for Head-Mounted displays. In *2009 IEEE Virtual Reality Conference*, VR '09, pp. 211–218. IEEE Computer Society, Washington, DC, USA, Mar. 2009.

[20] R. Jota, A. Ng, P. Dietz, and D. Wigdor. How fast is fast enough? a study of the effects of latency in direct-touch pointing tasks. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '13, pp. 2291–2300. Association for Computing Machinery, New York, NY, USA, Apr. 2013.

[21] B. Kress and T. Starner. A review of head-mounted displays (hmd) technologies and applications for consumer electronics. In *Photonic Applications for Aerospace, Commercial, and Harsh Environments IV*, vol. 8720, p. 87200A. International Society for Optics and Photonics, 2013.

[22] P. Lincoln, A. Blate, M. Singh, T. Whitted, A. State, A. Lastra, and H. Fuchs. From motion to photons in 80 microseconds: Towards minimal latency for virtual and augmented reality. *IEEE transactions on visualization and computer graphics*, 22(4):1367–1376, 2016.

[23] R. K. Mantiuk, G. Denes, A. Chapiro, A. Kaplanyan, G. Rufo, R. Bachy, T. Lian, and A. Patney. FovVideoVDP: a visible difference predictor for wide field-of-view video. *ACM Trans. Graph.*, 40(4):1–19, July 2021.

[24] M. Meehan, S. Razzaque, M. C. Whitton, and F. P. Brooks. Effect of latency on presence in stressful virtual environments. In *IEEE Virtual Reality, 2003. Proceedings.*, pp. 141–148, Mar. 2003.

[25] Y. Mikawa, M. Fujiwara, T. Hiraki, Y. Makino, and H. Shinoda. 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)*, pp. 137–143, Sept. 2021.

[26] Y. Mikawa, T. Sueishi, Y. Watanabe, and M. Ishikawa. Dynamic projection mapping for robust sphere posture tracking using uniform/biased circumferential markers. *IEEE Transactions on Visualization & Computer Graphics*, (01):1–1, 2021.

[27] M. Mine and G. Bishop. Just-In-Time pixels. Technical report, USA, 1995.

[28] L. Miyashita, Y. Watanabe, and M. Ishikawa. MIDAS projection: Markerless and modelless dynamic projection mapping for material representation. *ACM Transactions on Graphics (TOG)*, 37(6):1–12, 2018.

[29] D. Miyazaki and N. Hashimoto. Dynamic projection mapping onto non-rigid objects with dot markers. In *2018 International Workshop on Advanced Image Technology (IWAIT)*, pp. 1–4. IEEE, 2018.

[30] M. Nabiyouni, S. Scerbo, D. A. Bowman, and T. Höllerer. Relative effects of real-world and Virtual-World latency on an augmented reality training task: An AR simulation experiment. *Frontiers in ICT*, 3, 2017.

[31] A. Ng, J. Lepinski, D. Wigdor, S. Sanders, and P. Dietz. Designing for low-latency direct-touch input. In *Proceedings of the 25th annual ACM symposium on User interface software and technology*, pp. 453–464, 2012.

[32] T. T. Nguyen, A. Amthor, and C. Ament. Communication of the multi laser tracker system used as position feedback sensor. In *SPIE Eco-Photonics 2011: Sustainable Design, Manufacturing, and Engineering Workforce Education for a Green Future*, vol. 8065, pp. 325–336. SPIE, Apr. 2011.

[33] T. T. Nguyen, A. Amthor, S. Zschaeck, and S. Lambeck. Model based control of a linear stage using a contactless optical sensor system. In *2011 IEEE International Conference on Control System, Computing and Engineering*, pp. 409–414, Nov. 2011.

[34] T. Nomoto, W. Li, H.-L. Peng, and Y. Watanabe. Dynamic multi-projection mapping based on parallel intensity control. *IEEE Trans. Vis. Comput. Graph.*, 28(5):2125–2134, May 2022.

[35] K. Okumura, H. Oku, and M. Ishikawa. Lumipen: Projection-based mixed reality for dynamic objects. In *2012 IEEE International Conference on Multimedia and Expo*, pp. 699–704. IEEE, 2012.

[36] M. Olano, J. Cohen, M. Mine, and G. Bishop. Combatting rendering latency. In *Proceedings of the 1995 symposium on Interactive 3D graphics*, I3D '95, pp. 19–ff. Association for Computing Machinery, New York, NY, USA, Apr. 1995.

[37] I. C. on Non-Ionizing Radiation Protection et al. Icnirp guidelines on limits of exposure to incoherent visible and infrared radiation. *Health Physics*, 105(1):74–96, 2013.

[38] V. Popescu, J. Eyles, A. Lastra, J. Steinhurst, N. England, and L. Nyland. The WarpEngine: an architecture for the post-polygonal age. In *Proceedings of the 27th annual conference on Computer graphics and interactive techniques*, SIGGRAPH '00, pp. 433–442. ACM Press/Addison-Wesley Publishing Co., USA, July 2000.

[39] M. Regan and G. S. P. Miller. The problem of persistence with rotating displays. *IEEE Trans. Vis. Comput. Graph.*, 23(4):1295–1301, Apr. 2017.

[40] J. P. Rolland, L. D. Davis, and Y. Baillot. A survey of tracking technologies for virtual environments. In *Fundamentals of Wearable Computers and Augmented Reality*, pp. 83–128. CRC Press, 1st edition ed., Jan. 2001.

[41] M. F. R. Rondón, L. Sassatelli, R. Aparicio-Pardo, and F. Precioso. TRACK: A new method from a Re-Examination of deep architectures for head motion prediction in 360° videos. *IEEE Trans. Pattern Anal. Mach. Intell.*, 44(9):5681–5699, Sept. 2022.

[42] T. Sueishi, H. Oku, and M. Ishikawa. Lumipen 2: Dynamic projection mapping with mirror-based robust high-speed tracking against illumination changes. *Presence*, 25(4):299–321, 2016.

[43] X. Sun, D. E. Smith, E. D. Hoffman, S. W. Wake, D. R. Cremons, E. Mazarico, J.-M. Lauenstein, M. T. Zuber, and E. C. Aaron. Small

and lightweight laser retro-reflector arrays for lunar landers. *Appl. Opt.*, 58(33):9259–9266, Nov. 2019.

[44] I. Sutherland. The ultimate display. 1965.

[45] I. E. Sutherland. Three-dimensional data input by tablet. *Proc. IEEE*, 62(4):453–461, Apr. 1974.

[46] T. A. Syed, M. S. Siddiqui, H. B. Abdullah, S. Jan, A. Namoun, A. Alzahrani, A. Nadeem, and A. B. Alkhodre. In-Depth review of augmented reality: Tracking technologies, development tools, AR displays, collaborative AR, and security concerns. *Sensors*, 23(1), Dec. 2022.

[47] C. Wang, E. D. Burnham-Fay, and J. D. Ellis. Real-time FPGA-based kalman filter for constant and non-constant velocity periodic error correction. *Precis. Eng.*, 48:133–143, Apr. 2017.

[48] M. Warburton, M. Mon-Williams, F. Mushtaq, and J. R. Morehead. Measuring motion-to-photon latency for sensorimotor experiments with virtual reality systems. *Behav. Res. Methods*, Oct. 2022.

[49] Y. Watanabe, G. Narita, S. Tatsuno, T. Yuasa, K. Sumino, and M. Ishikawa. High-speed 8-bit image projector at 1,000 fps with 3 ms delay. In *22nd International Display Workshops, IDW 2015*, pp. 1421–1422. International Display Workshops, 2015.

[50] G. Welch, G. Bishop, L. Vicci, S. Brumback, K. Keller, and D. Colucci. 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*, pp. 1–ff, 1999.

[51] F. Zheng, T. Whitted, A. Lastra, P. Lincoln, A. State, A. Maimone, and H. Fuchs. Minimizing latency for augmented reality displays: Frames considered harmful. In *2014 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, pp. 195–200, Sept. 2014.