3D Gaze on Stationary and Moving Visual Targets in Mixed Reality Environments

Kenta Kato, Oky Dicky Ardiansyah Prima
Graduate School of Software and Information Science
Iwate Prefectural University
Takizawa, Japan
e-mail: g236s002@s.iwate-pu.ac.jp, prima@iwate-pu.ac.jp

Abstract— **With the spread of Head-Mounted Displays (HMDs), various simulations have been conducted using Mixed Reality (MR) environments that merge virtual objects in a physical space. Our Three-Dimensional (3D) perception may change as opportunities to have more virtual 3D experiences in such an environment increase. The relationship between differences in depth perception and changes in 3D gaze behavior would be of interest, but such detailed analysis is yet to be conducted. In this study, we developed an Optical See-Through Head-Mounted Display (OST-HMD) and experimentally evaluated the effect of an MR environment on 3D gaze measurement. Our experiments showed that the relative size of the 3D visual targets and the surrounding depth cues had no effect on the accuracy of the 3D gaze. However, it is important to consider the building of a polynomial for projecting the 3D gaze, and there is room for improvement in the polynomial. We have successfully measured the scanpath of 3D gaze on a visual target approaching from 200 cm. This result suggests that our OST-HMD can measure 3D gaze in the personal space, defined by Cutting (1997) as a space up to 1.5 m from the viewer.**

Keywords-3D gaze estimation; Optical See-Through Head-Mounted Display; depth perception; Mixed Reality.

I. INTRODUCTION

Three-Dimensional (3D) visualization of medical images of the human body and their interactive manipulation have been expected to be utilized in the medical field for safe surgery through preoperative scenario planning. In the past decade, the hardware for processing 3D information has improved dramatically, enabling more accurate tracking of the perspective and input of the user. In addition, Augmented Reality (AR) and Mixed Reality (MR) environments, which integrate the virtual world with the real world, have made it possible to manipulate 3D contents naturally. However, 3D contents in these environments may cause operation errors if the contents are perceived incorrectly. Studies on accurate projection of these contents in real environments are still in progress.

Problems related to the position, scale, and posture of 3D contents in virtual and real environments require not only verification based on subjective depth perception, but also objective analysis based on the 3D gaze information of viewers. We studied the relationship between each environment and 3D gaze information by measuring vergence eye movement-based 3D gaze to 3D visual targets in both environments [1]. However, there was no significant difference in the measured 3D gaze between the environments with and without depth cues.

There are nine sources of information for perceiving depth: binocular disparity, convergence, occlusion, relative size, height in the visual field, relative density, aerial perspective, accommodation, and motion parallax [2]. The first two pieces of information are binocular cues, which perceive depth from the disparity caused by the disparity between the eyes, and the remaining are monocular cues caused by the relative position of the observer and the object. Cutting (1997) argued that the effect of depth cues on distance estimation depends on the perceptual areas regarding the observer's position: personal space (~1.5m), action space (1.5m to 30m), and vista space (from 30m) [3]. Many studies on the relationship between vergence and perceived distance have experimented with visual targets of up to 2m [4]-[6].

Binocular disparity-based visual targets are necessary to measure 3D gaze information. Head-Mounted Displays (HMDs) are often used to present these visual targets. HMDs can be classified into two categories: Video See-Through Head-Mounted Display (VST-HMD) and Optical See-Through Head-Mounted Display (OST-HMD). The former projects a 3D object superimposed on a video image of the real world, while the latter projects a 3D object projected on a half-mirror superimposed on the real world. The viewable area in VST-HMD depends on the Field of View (FoV) of the scene camera and the display. In contrast, OST-HMD transparently displays 3D objects on the viewer's FoV, making it suitable for manipulating 3D virtual objects over real objects simultaneously. OST-HMDs are also relatively less susceptible to the VR sickness that occurs with devices that cover the field of view such as the VST-HMD [7][8]. Hence, the adoption of OST-HMD is under positive consideration in the medical field [9].

3D gaze can be measured by analyzing the relationship between the position of an object and the eye's vergence angle when looking at that object [10]-[13]. Öney et al. (2020) used binocular eye tracking of the Microsoft HoloLens to measure 3D gaze in visual search tasks for 3D objects located within 1.25m to 5m from the subject in an MR environment. Their experiments, however, showed a significantly larger measurement error of more than 1 m when the focused object was only 3.5 m away from the viewer [14]. Kato et al. (2018) confirmed the effect of adding patterns to visual targets used in gaze calibration to improve the accuracy of 3D gaze measurement [15].

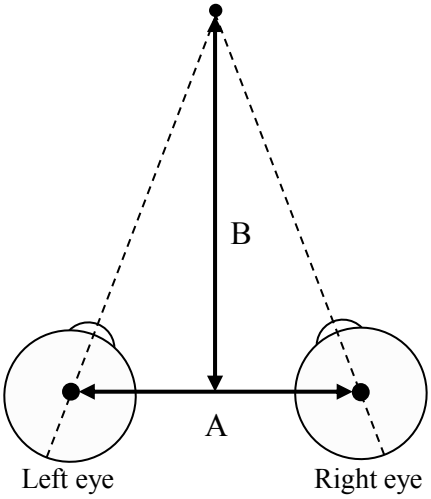


Figure 1. The basic binocular geometry

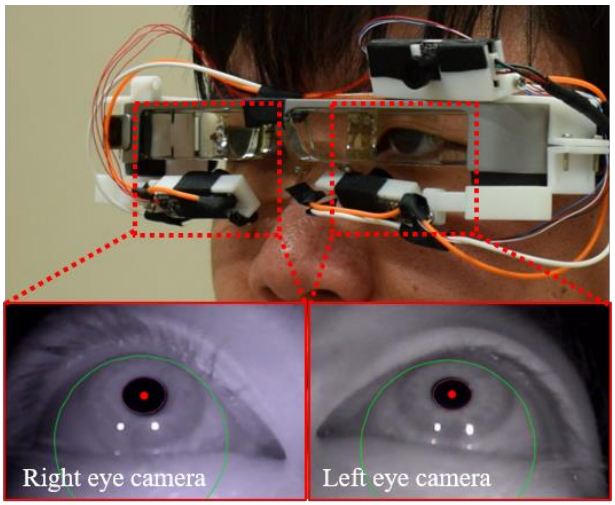


Figure 3. Our 3D eye tracker used in this study.

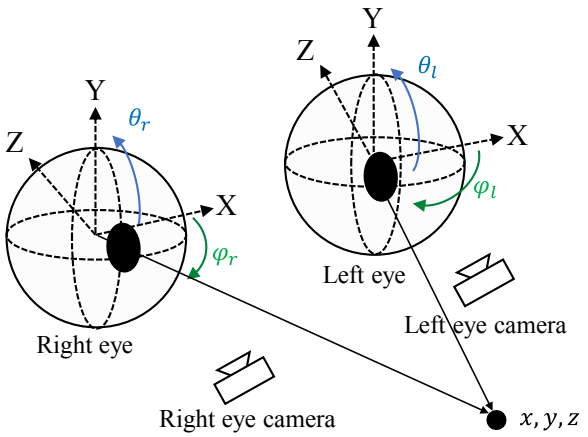


Figure 2. The relationship between the binocular cameras and the estimated the line-of-sight.

In this study, we measure 3D gaze in an MR environment to evaluate the effects of the relative size of the 3D visual target and the surrounding physical environment as a cue for 3D perception on the accuracy of the 3D gaze, and to characterize the scan path of 3D gaze for stationary and moving visual targets. For this study, a 3D eye tracker using OST-HMD intended for MR was developed.

The rest of this paper is organized as follows. Section II describes the calculation of 3D gaze based on eye vergence. Section III describes the 3D eye tracker based on OST-HMD developed for this study. Section IV explains our experiments. Section V describes the results. Finally, Section VI presents our conclusion.

II. 3D Gaze Measurement Based on Eye Vergence

The calculation of the 3D gaze depends on the relative positions of the two eyes to a given visual target. Vergence angle is the angle between the line-of-sight of the two eyes. The change in vergence angle occurs when the target is moved at a certain distance from the observer. Figure 1 shows the basic binocular geometry. A and B are the disparity distance between eye centers and the distance from the visual target to the observer, respectively. A 3D gaze point (x_g, y_g, z_g) can be calculated from each line-of-sight, ($Pitch_L$, Yaw_L) and ($Pitch_R$, Yaw_R). The pitch and yaw angles describe angles between the X axis and Y axis, respectively.

In a binocular eye tracker, the eye cameras are used to determine the lines of sight of the two eyes, but in many cases, the cameras are placed below the eyes to avoid interfering with the field of view. Due to the placement of these cameras, the calculated pitch and yaw angles of the eyes differ from the angles as viewed from the front. Therefore, to obtain the 3D gaze from the eye tracker, the position of the visual target (x, y, z) must be calibrated using the line-of-sight of both eyes ($Pitch_L$, Yaw_L, $Pitch_R$, Yaw_R). Polynomials are generally used for gaze calibration. Figure 2 shows the relationship between the binocular cameras and the estimated pitch and yaw angles. These relationships can be solved by the following nonlinear multiple regression equations.

$$x = f(Pitch_L, Yaw_L, Pitch_R, Yaw_R) + \alpha \quad (1)$$

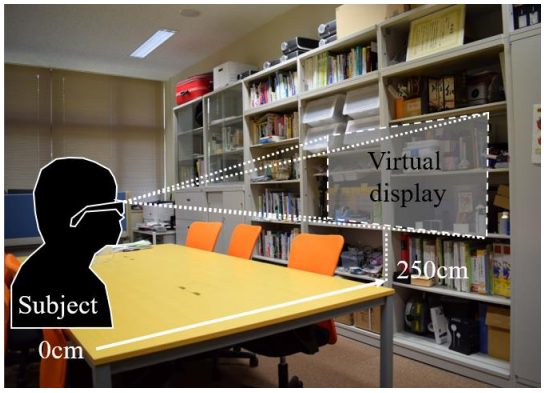
$$y = g(Pitch_L, Yaw_L, Pitch_R, Yaw_R) + \beta \quad (2)$$

$$z = h(Pitch_L, Yaw_L, Pitch_R, Yaw_R) + \gamma \quad (3)$$

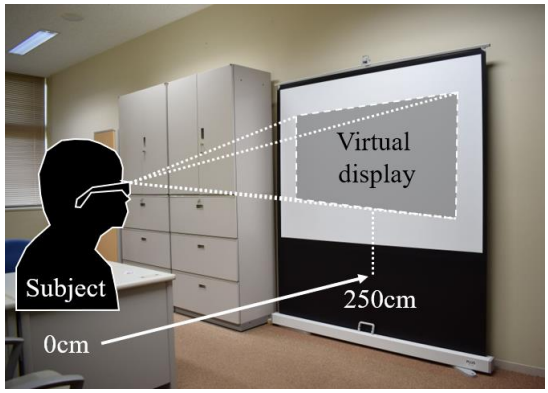
Here, f, g, h are composite functions of line-of-sight. α, β, γ are the residuals of each equation.

III. Our 3D Eye Tracker Based on OST-HMD

A variety of OST-HMDs are available on the market. In this study, we adopted the Moverio (BT-30E, EPSON) to build our OST-HMD 3D eye tracker in MR environment. This device can display images from a PC, making it easy to provide users with 3D visual stimulation. Three simultaneous USB camera modules (KYT-U030-3NF, KAYETON) were used to capture both eyes and the viewer's scene. These

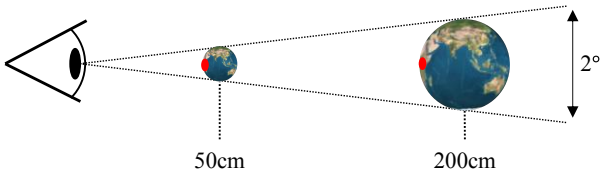


(a) A room with depth cues.

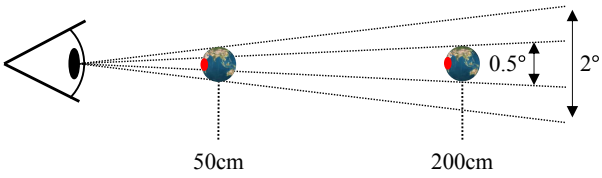


(b) A room without depth cues.

Figure 4. The physical environment for the experiments in this study.



(a) Visual target with a 2° viewing angle in size regardless of distance.



(b) A fixed-size visual target.

Figure 5. Two types of visual targets used in this study.

cameras operate at 60 Hz allowing tracking of eye movements at a rate comparable to low- to mid-end eye trackers, such as the Tobii Pro Nano. For tracking the pupil, we installed a 940nm high-pass filter (FUJIFILM IR-94) in the lens of the camera that captures the eyeball to block visible light while allowing infrared light to pass through. Figure 3 shows our OST-HMD 3D eye tracker. Each camera mount is made with a 3D printer. openFrameworks v0.11.0, an open-source C++ toolkit, was used to visualize the virtual targets and record the measurement data. We used Pupil Capture (Pupil Labs), an open-source eye tracking platform, to calculate the line-of-sight of each eye from the estimated eyeball center [16]. The line-of-sight data measured by the Pupil Capture is sent to openFrameworks using ZeroMQ, an asynchronous messaging library, and 3D gaze is calculated based on the relationship between the visual targets and the line-of-sight data.

The visual targets used for the measurements were created by generating binocular disparity to guide the vergence. The interpupillary distance was fixed at 6.3cm, which is the average interpupillary distance for Japanese [17].

A. MR Environments

In this study, two MR environments were established as shown in Figure 4. An environment with depth cues is a room with a table and chairs and shelves with various objects within the observer's FoV. On the other hand, an environment without depth cues is a room where a projector screen is placed in the observer's FoV. The viewer wearing OST-HMD is seated 250cm away from the shelf and screen projector to perceive a virtual screen equivalent to 40 inches. The lighting in the room is fluorescent with a color temperature of 5500K. The brightness of the room is approximately 150lx.

B. 3D Gaze Estimation

The polynomials in equations (1)-(3) are calculated using the positions of multiple visual targets and the line-of-sight of both eyes when gazing at those targets. Let (θ_l, θ_r) represent pitch angles and (φ_l, φ_r) represent yaw angles of the line-of-sight, respectively. Two types of polynomials were constructed for this study as follows.

1. Type I: polynomial used by Kato and Prima (2021) [1]

This polynomial achieves an average 3D gaze accuracy of less than 25cm, about four times better than the results presented by Öney et al. (2020) [14].

$$G_x = a_1\theta_r^2 + a_2\varphi_r^2 + a_3\theta_l^2 + a_4\varphi_l^2 + a_5\theta_r\varphi_r + a_6\theta_l\varphi_l + a_7\theta_r\theta_l + a_8\theta_r\varphi_l + a_9\theta_l\varphi_r + a_{10}\varphi_r\varphi_l + a_{11}\theta_r + a_{12}\varphi_r + a_{13}\theta_l + a_{14}\varphi_l + a_{15} \quad (4)$$

$$G_y = b_1\theta_r^2 + b_2\varphi_r^2 + b_3\theta_l^2 + b_4\varphi_l^2 + b_5\theta_r\varphi_r + b_6\theta_l\varphi_l + b_7\theta_r\theta_l + b_8\theta_r\varphi_l + b_9\theta_l\varphi_r + b_{10}\varphi_r\varphi_l + b_{11}\theta_r + b_{12}\varphi_r + b_{13}\theta_l + b_{14}\varphi_l + b_{15} \quad (5)$$

$$G_z = c_1\theta_r^2 + c_2\varphi_r^2 + c_3\theta_l^2 + c_4\varphi_l^2 + c_5\theta_r\varphi_r + c_6\theta_l\varphi_l + c_7\theta_r\theta_l + c_8\theta_r\varphi_l + c_9\theta_l\varphi_r + c_{10}\varphi_r\varphi_l + c_{11}\theta_r + c_{12}\varphi_r + c_{13}\theta_l + c_{14}\varphi_l + c_{15}. \quad (6)$$

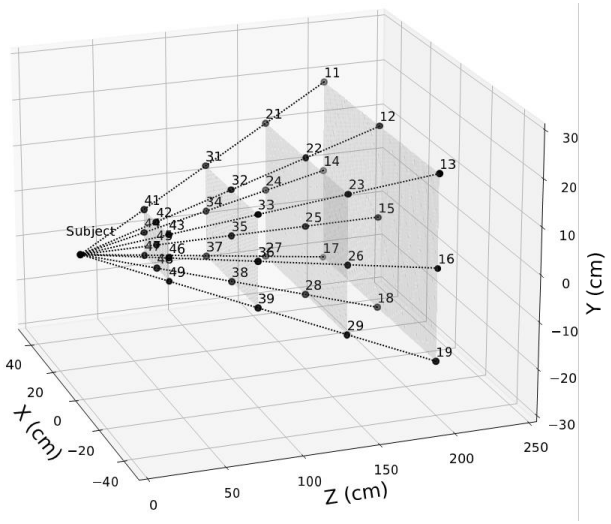


Figure 6. Locations of the visual targets for eye calibration.

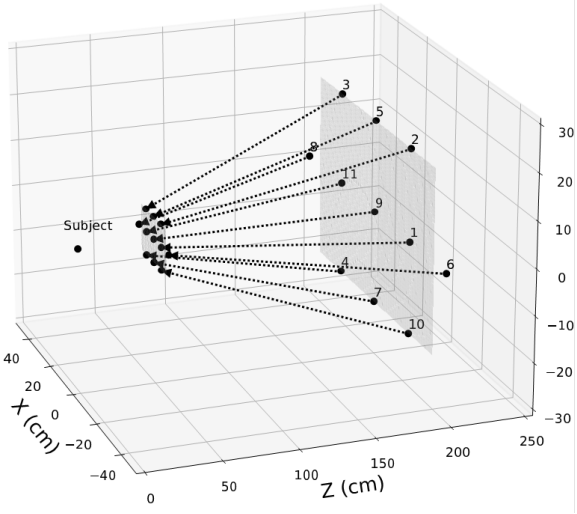


Figure 8. Path of the moving visual targets.

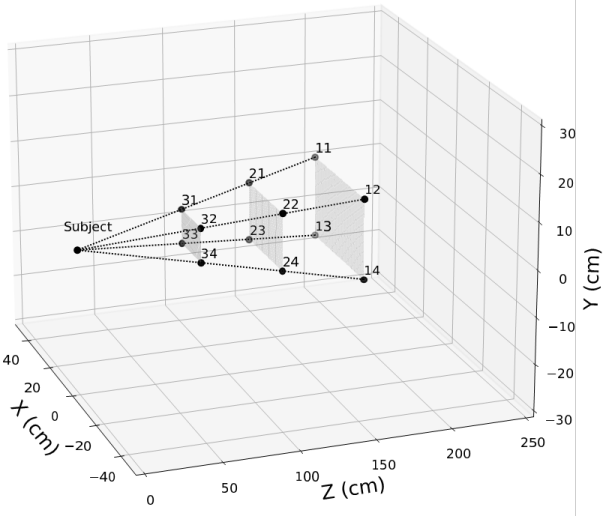


Figure 7. Locations of the stationary visual targets.

Coefficients $(a_1 \sim a_{15}, b_1 \sim b_{15}, c_1 \sim c_{15})$ are calculated by the least-squares method based on the correspondence between the pitch and yaw angles $(\theta_l, \theta_r, \varphi_l, \varphi_r)$ of each eye and the position of the gazing target.

2. Type II: polynomial with constrained pitch angles.

The range of vertical eye movements is shorter than the range of horizontal eye movements. Therefore, we consider it necessary to minimize the effect of vertical eye movements in the x- and y-axes of the 3D gaze. For this purpose, we introduce θ_{rl} as the average of the pitch angles of both eyes.

$$x = a_1\varphi_r^2 + a_2\varphi_l^2 + a_3\varphi_r\varphi_l + a_4\varphi_r + a_5\varphi_l + a_6 \quad (7)$$

$$y = b_1\theta_{rl}^2 + b_2\varphi_r^2 + b_3\varphi_l^2 + b_4\theta_{rl}\varphi_r + b_5\theta_{rl}\varphi_l + b_6\varphi_r\varphi_l + b_7\theta_{rl} + b_8\varphi_r + b_9\varphi_l + b_{10} \quad (8)$$

$$z = c_1\varphi_r^2 + c_2\varphi_l^2 + c_3\varphi_r\varphi_l + c_4\varphi_r + c_5\varphi_l + c_6 \quad (9)$$

Coefficients $(a_1 \sim a_6, b_1 \sim b_{10}, c_1 \sim c_6)$ are calculated by the least-squares method.

IV. EXPERIMENT

The following steps are taken to characterize the scan path of 3D gaze for stationary and moving visual targets as well as to evaluate the impact of the relative size of the 3D visual target and the surrounding physical environment on the accuracy of 3D gaze. The subject wears the OST-HMD and is seated 250cm away from a shelf, an environment with depth cues, and a screen projector, which is without depth cues.

Step 1. Adjust parameters such as blob thresholds and pre-defined pupil sizes to improve pupil detection in the Pupil Capture software program.

Step 2. Present the subject with 36 stationary 3D visual targets placed in a virtual space and measure the line of sight. Each target is presented continuously for 3 seconds.

Step 3. Perform a Two-Dimensional (2D) eye calibration before constructing the polynomial for 3D gaze estimation. After a successful 2D calibration within 2°, display the other 12 stationary visual targets to the subject. If the 2D calibration fails, return to step 1.

Step 4. Ask the subject to track a single visual target approaching from 200cm to 50cm in virtual space. Here, the 200cm was determined in consideration

TABLE I. THE ACCURACIES OF 3D GAZE IN ENVIRONMENTS WITH AND WITHOUT DEPTH CUES, AND WITH AND WITHOUT ADJUSTMENTS OF VISUAL TARGETS (cm)

Subject	With depth cues				Without depth cues			
	Fixed-size		Adaptively-adjusted size		Fixed-size		Adaptively-adjusted size	
	Type I	*Type II*	*Type I*	*Type II*	*Type I*	*Type II*	*Type I*	*Type II*
1	18.65	28.50	21.49	25.79	11.06	25.78	11.25	19.88
2	19.23	24.84	20.27	30.36	21.98	28.71	26.80	32.31
3	13.22	20.29	7.52	10.28	20.22	22.53	12.79	16.61
4	19.00	25.46	22.22	26.03	11.90	13.64	9.91	15.86
5	16.15	21.00	18.36	30.77	25.38	35.41	12.85	18.74
Mean	17.25	24.02	17.97	24.65	18.11	25.21	14.72	20.68
Std. Dev.	2.295	3.026	5.385	7.480	5.669	7.181	6.137	5.992

that the effective range of vergence is within 2m [3]. A total of 11 moving visual targets coming from different directions are displayed. Each target will take 5 seconds to travel. To give subjects a time to prepare for gazing, each target was paused for one second at the beginning and end of its movement.

During Step 2 and Step 3, line-of-sight of both eyes are recorded for 1 second while gazing at the target without blinking. A red dot blinks in the center of the target during the recording. In Step 5, all line-of-sight of both eyes are recorded during the appearance of the target. Post-processing is performed to remove the data when a blink occurs.

Figure 5 shows two types of visual targets in this study. The first type adjusts its size to appear consistently 2° regardless of distance, while the second type keeps the size constant. As a result, the second type of visual target is observed to change its size from 2° in diameter at 50cm to 0.5° at 200cm.

A. Visual Targets for the Eye Calibration

Vertical virtual planes are generated at 50cm intervals from 50cm to 200cm where nine visual targets are placed on each plane. These visual targets were placed at 5° horizontally and 2.5° vertically. Figure 6 shows the locations of the visual targets in each plane. The vector connecting the *n*-th target in each plane intersects the position of the viewpoint.

B. Stationary Visual Targets

Like the visual target for the eye calibration, three stationary vertical virtual planes are generated at 50cm intervals from 75cm to 175cm. Four visual targets are placed on each side, for a total of 12 targets. These visual targets were placed at 2.5° horizontally and 1.25° vertically. Figure 7 shows the locations of the visual targets in each plane.

C. Moving Visual Targets

11 moving visual targets are designed to start moving from a 200cm virtual plane and stop at a 50cm virtual plane. Figure 8 shows the locations of the start and stop locations of the visual targets. These visual targets were placed at 3° horizontally and 2.5° vertically. The dotted lines represent the trajectories of the moving targets.

V. RESULTS

Five subjects (male, mean age 23.6) participated in the experiment. They were tested for visual acuity using a Landolt ring to confirm that their vision achieved 1.0 or better.

TABLE II. ERRORS OF THE 3D GAZE TO THE STATIONARY VISUAL TARGETS PROJECTED BY TYPE I AND II POLYNOMIALS (cm)

	Type 1	Type 2
1	29.65	12.35
2	28.96	30.77
3	52.87	47.27
4	22.84	21.62
5	26.23	22.41
6	35.81	22.40
7	24.53	18.62
8	12.11	10.49
9	34.56	27.50
10	44.84	41.59
11	36.94	25.03
12	27.47	18.41
13	22.73	19.45
14	16.31	11.62
15	71.34	26.74
16	21.96	22.77
17	27.45	20.87
18	23.41	26.82
19	28.14	31.42
20	37.24	41.26
Mean	32.11	26.88
Std. Dev.	9.846	11.744

They were also asked to fill out a questionnaire to confirm that they had no health concerns. However, one of the subjects was not able to track the moving visual targets correctly, so there were no results of gazing at targets only for that subject.

A. Pre-processing

The data with a pupil detection confidence of less than 60% was excluded to avoid the inclusion of various noises in the gaze data. This is done using the features of the Pupil Capture software.

B. Measurement Accuracy

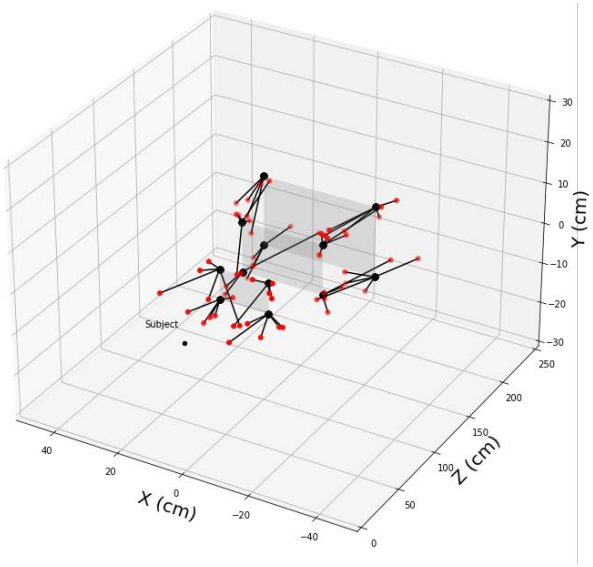
Accuracies for the 3D gaze measurement is measured by

$$Acc = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(T_{xi} - G_{xi})^2 + (T_{yi} - G_{yi})^2 + (T_{zi} - G_{zi})^2}. \quad (10)$$

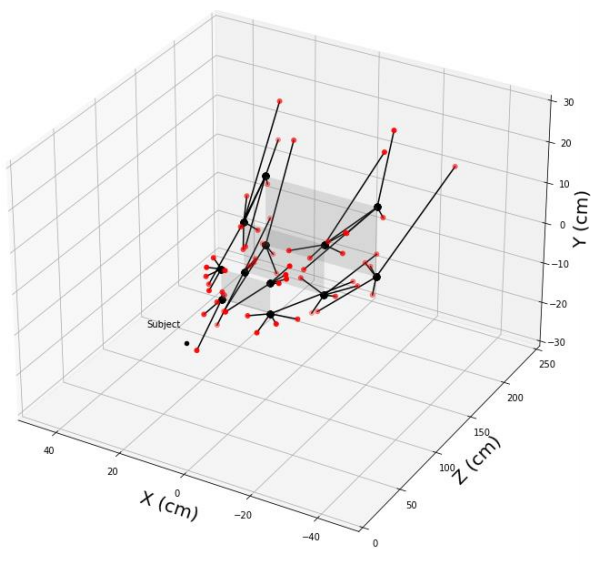
Here, n is the number of targets used for the measurement, T_{xi}, T_{yi}, T_{zi} and G_{xi}, G_{yi}, G_{zi} are the coordinates of the i-th target and the associated eye-gaze points projected by polynomials (Type I or Type II) described in Section III.

TABLE III. ACCURACIES OF THE 3D GAZE TO THE MOVING VISUAL TARGETS (cm)

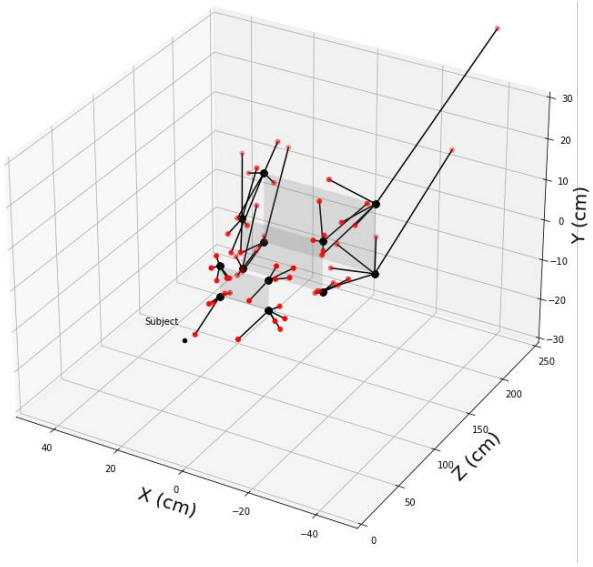
Subject	With depth cues				Without depth cues			
	Fixed-size		Adaptively-adjusted size		Fixed-size		Adaptively-adjusted size	
	Type I	*Type II*	*Type I*	*Type II*	*Type I*	*Type II*	*Type I*	*Type II*
1	33.49	18.31	89.35	17.77	25.87	16.51	43.50	19.34
3	43.13	29.55	15.42	16.33	16.26	14.57	124.74	39.83
4	21.37	28.30	73.70	40.16	24.32	26.48	49.82	56.04
5	27.91	24.36	33.07	26.98	365.92	25.96	99.56	25.68
Mean	31.48	25.13	52.89	25.31	108.09	20.88	79.41	35.22
Std. Dev.	7.980	4.376	29.830	9.500	148.900	5.388	34.008	14.123



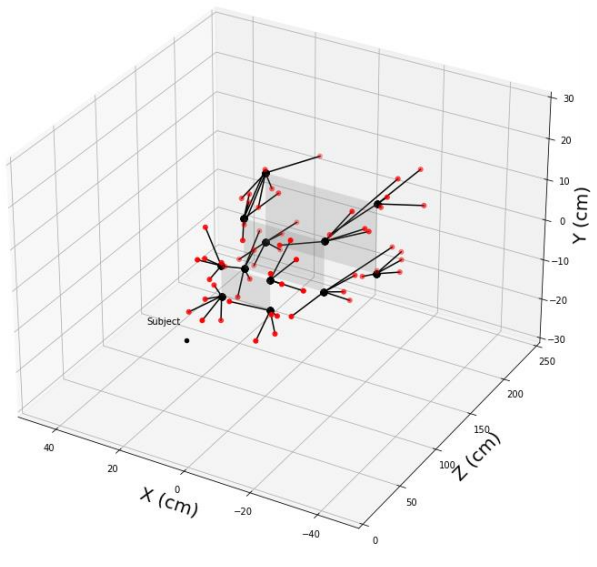
(a) 3D gaze in environments with depth cues and with adjustments of visual targets.



(b) 3D gaze in environments with depth cues and without adjustments of visual targets.

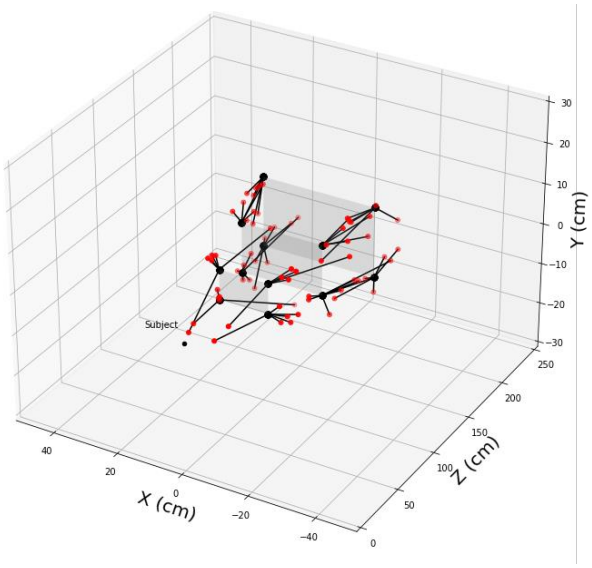


(c) 3D gaze in environments without depth cues and with adjustments of visual targets.

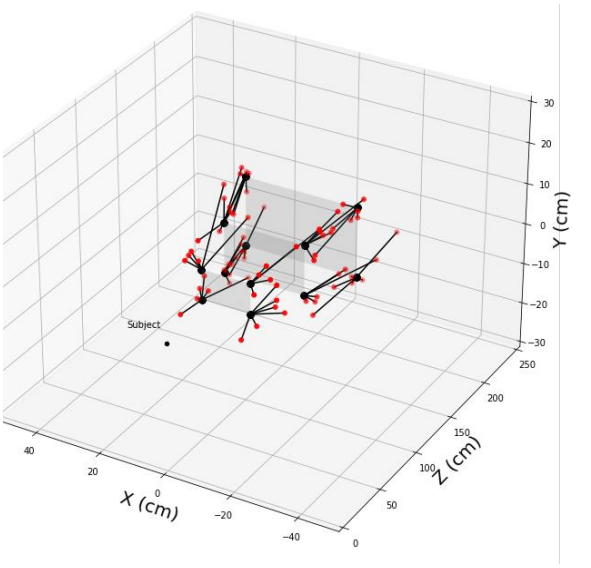


(d) 3D gaze in environments without depth cues and without adjustments of visual targets.

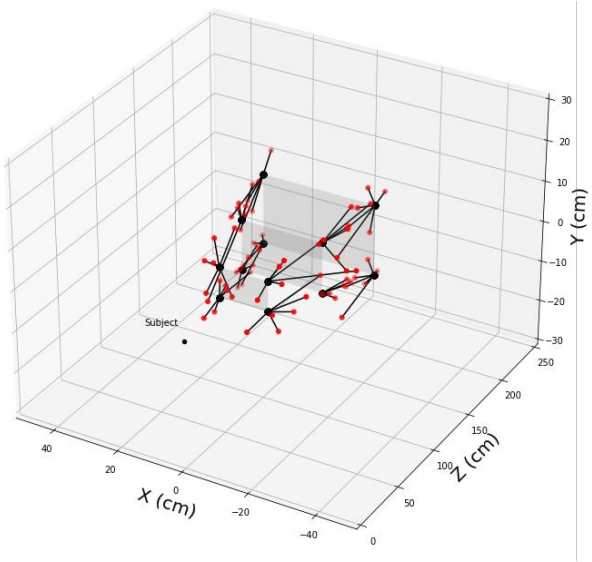
Figure 9. Distribution of 3D gaze to stationary visual targets projected by Type I



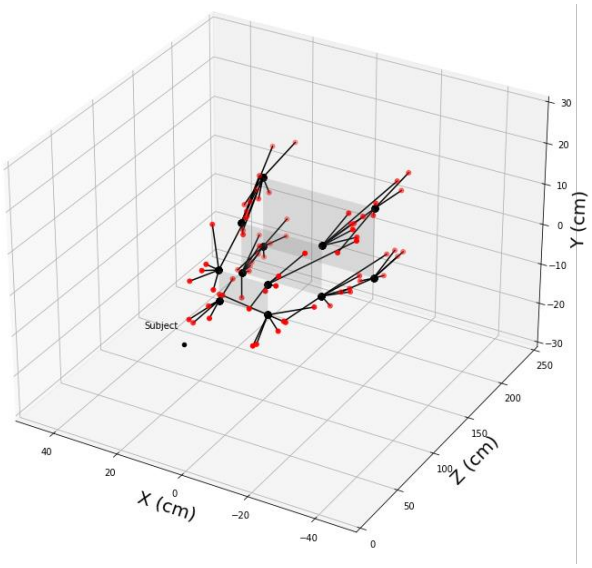
(a) 3D gaze in environments with depth cues and with adjustments of visual targets.



(b) 3D gaze in environments with depth cues and without adjustments of visual targets.



(c) 3D gaze in environments without depth cues and with adjustments of visual targets.

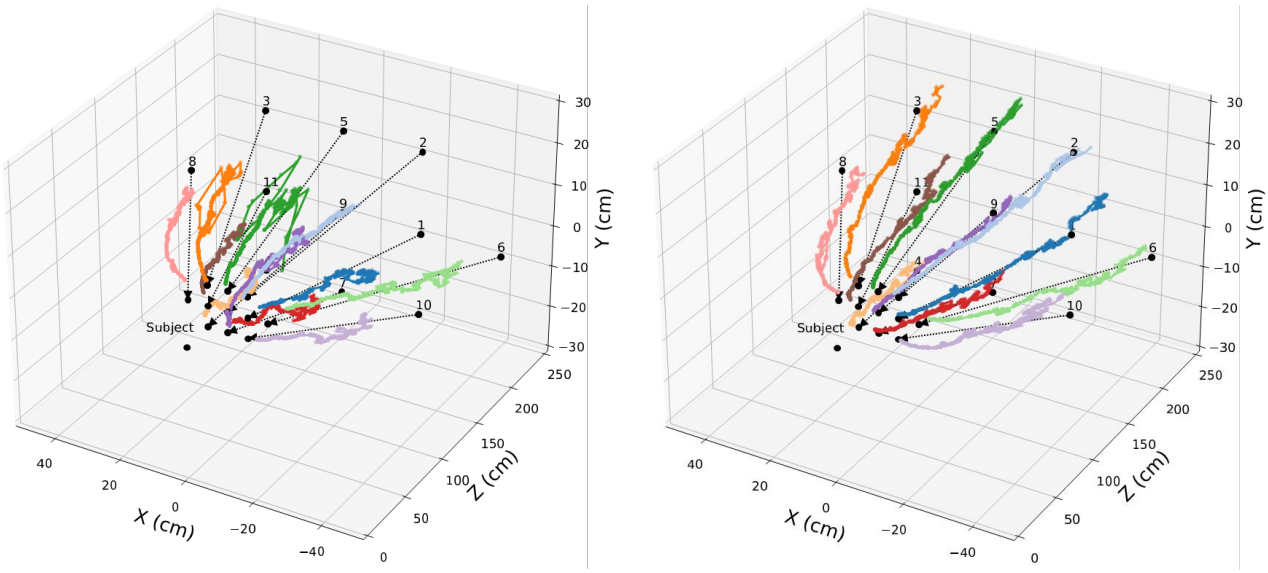


(d) 3D gaze in environments without depth cues and without adjustments of visual targets.

Figure 10. Distribution of 3D gaze to stationary visual targets projected by Type II

Table I shows the accuracies of gaze estimation in environments with and without depth cues, and with and without adjustments of the visual targets. Overall, the accuracy of the 3D gaze calculated by the Type I polynomial is superior regardless of the environments or the visual targets. To confirm what affects the accuracy of 3D gaze, we performed a 2 × 2 × 2, depth cue × target adjustment × polynomial type, three-way analysis of variance on these results. The main effect was found for polynomial type (F (1, 32) = 10.949, p = 0.002), supporting the result that Type I polynomials are better at computing 3D gaze.

Using the polynomial equations of Type I and Type II obtained from the eye calibration, we visualized the distribution of the 3D gaze to stationary visual targets. Figures 9 and 10 show the distribution of the 3D gaze projected by Type I and Type II, respectively. The red lines indicate the relationships between the positions of the projected 3D gaze (red dots) of all subjects and the visual targets (black dots). A longer line segment means a larger error of the 3D gaze. In contrast to the eye calibration results, overall, the errors in the 3D gaze projected by Type II are smaller, regardless of the environment or visual target. Table II shows errors of the 3D



(a) Scanpath of the 3D gaze of users #1.

(b) Scanpath of the 3D gaze of users #3.

Figure 11. Scanpath of the 3D gaze of users #1 and #3 to moving visual targets (without adjustments of visual target) in an environment with depth cues.

gaze projected by Type I and II polynomials. Significant differences were found when the accuracy of the 3D gaze obtained by each polynomial was analyzed by t-test, indicating that Type II can project the 3D gaze with higher accuracy ($t(19) = 2.640$, $p = 0.016$). Type I polynomials have a total of 45 coefficients, while Type II polynomials have only 22. Generally, the higher the number of coefficients produced the better the fitting. This applies to the 3D gaze to the visual targets for the gaze calibration. However, this is not inherently the accuracy of the 3D gaze as fitting of the polynomial to other visual targets is not guaranteed. Therefore, better results for 3D Gaze against the stationary visual targets projected using Type II polynomials are acceptable.

Table III shows the accuracies of the 3D gaze to the moving visual targets. Subject #2 was not able to correctly track the moving visual target hence the results for this subject were excluded. Two-way analysis of variance on these results revealed that no effects were found in both depth cue (F (1, 12) = 0.285, p = .603) and visual target adjustment (F (1, 12) =1.873, p = .196).

Figure 11 shows the scanpath of the 3D gaze of users #1 and #3 to moving visual targets (without adjustment) in an environment with depth cues. The scanpath of subject #1 extends to 1.5m in each direction of the visual target, while the scanpath of subject #3 extends to 2m. This result implies that our OST-HMD can measure 3D gaze in the personal space defined by Cutting (1997) [3]. Figures 12 and 13 show the scan path in Figure 11, grouped by height, and displayed from the top viewpoint. Observation of these scan paths shows that they are curved at locations close to the viewers.

Our experiments showed that the relative size of the 3D visual targets and the surrounding depth cues had no effect on the accuracy of the 3D gaze. However, it is important to consider the building of a polynomial for projecting the 3D gaze, and there is room for improvement in the polynomial.

VI. Conclusion

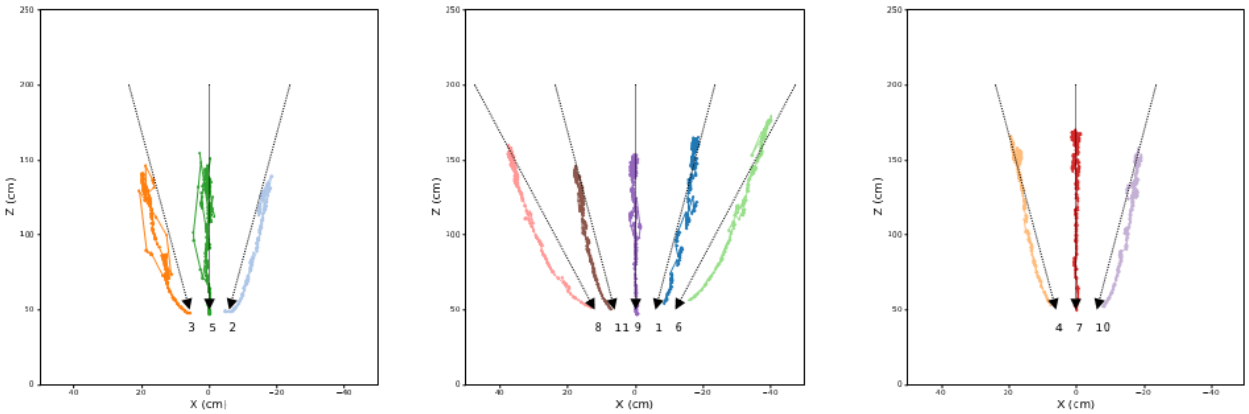
In this study, we conducted 3D gaze measurements in MR environments and confirmed that the relative size of the 3D visual targets and the surrounding physical environments did not affect the accuracy of the 3D gaze. Two types of polynomials were used to estimate the 3D gaze: the polynomial used by Kato and Prima (2021) and a newly constructed polynomial that constrains the pitch angle of the line-of-sight. The former is composed of 45 coefficients, while the latter has only 22 coefficients. Our experiments showed that we achieved higher accuracy in estimating the 3D gaze by using the polynomial consisting of these 22 coefficients. Using higher order polynomials may provide better eye calibration accuracy, but the accuracy achieved by this may not be inherently accurate.

We were concerned about the effect on the user's 3D perception in 3D visualization and interactive manipulation of medical images in the MR environment, however, we were unable to find these concerns in our current experimental results.

References

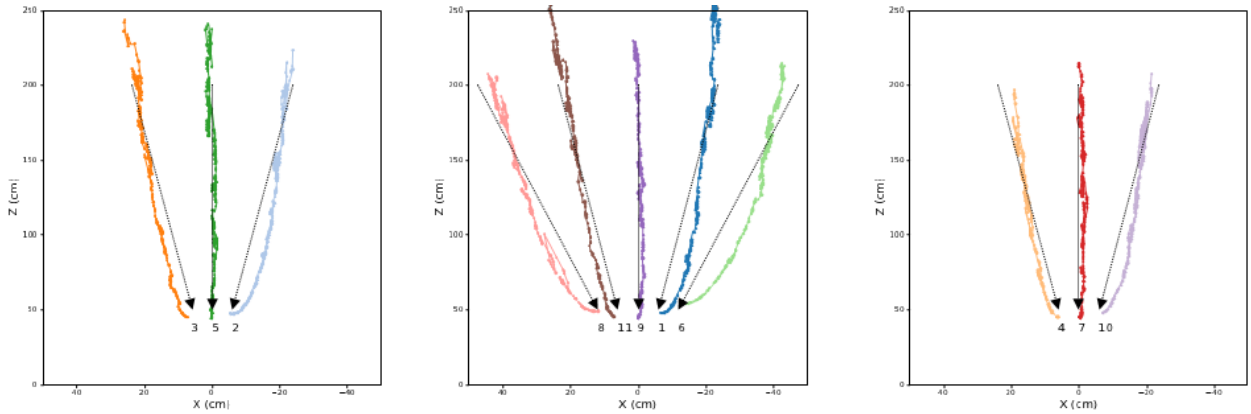
[1] K. Kato and O. D. A. Prima, “3D Gaze Characteristics in Mixed-Reality Environment,” eTELEMED 2021 : The Thirteenth International Conference on eHealth, Telemedicine, and Social Medicine, IARIA, pp.11-15, 2021.

[2] J. E. Cutting and P. M. Vishton, “Perceiving Layout and Knowing Distances : The Integration, Relative Potency, and Contextual Use of Different Information about Depth,” Percept. Sp. Motion, 22(5), pp. 69–117, 1995.



(a) Moving target #2, #3, #5 (b) Moving target #1, #6, #8, #9, #11 (c) Moving target #4, #7, #10

Figure 12. Top view of scanpath of the 3D gaze of users #1 to moving visual targets (without adjustments of visual target) in an environment with depth cues.



(a) Moving target #2, #3, #5 (b) Moving target #1, #6, #8, #9, #11 (c) Moving target #4, #7, #10

Figure 13. Top view of scanpath of the 3D gaze of users #3 to moving visual targets (without adjustments of visual target) in an environment with depth cues.

[3] J. E. Cutting, “How the Eye Measures Reality and Virtual Reality,” Behavior Research Methods, Instruments, & Computers, 29(1), pp. 27–36, 1997.

[4] A. Viguier, G. Clément, and Y. Trotter, “Distance Perception within Near Visual Space,” Perception, 30(1), pp. 115–124, 2001.

[5] S. Lee, X. Hu, and H. Hua, "Effects of Optical Combiner and IPD Change for Convergence on Near-Field Depth Perception in an Optical See-Through HMD," in IEEE Transactions on Visualization & Computer Graphics, 22(5), pp. 1540-1554, 2016.

[6] G. Singh, S. R. Ellis, and J. E. Swan, “The Effect of Focal Distance, Age, and Brightness on Near-Field Augmented Reality Depth Matching,” IEEE Transactions on Visualization Computing Graphics, 26(2), pp. 1385–1398, 2018.

[7] G. Ballestin, F. Solari, and M. Chessa, "Perception and Action in Peripersonal Space: A Comparison Between Video and Optical See-Through Augmented Reality Devices," 2018 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct), pp. 184-189, 2018.

[8] A. Vovk, F. Wild, W. Guest, and T. Kuula, “Simulator Sickness in Augmented Reality Training Using the Microsoft HoloLens,” In Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems (CHI '18), Association for Computing Machinery, 209, pp. 1–9, 2018.

[9] J. P. Rolland and H. Fuchs, “Optical Versus Video See-Through Head-Mounted Displays in Medical Visualization,” Presence Teleoperators Virtual Environment, 9(3), pp. 287–309, 2000.

[10] T. Pfeiffer, M. E. Latoschik, and I. Wachsmuth, “Evaluation of Binocular Eye Trackers and Algorithms for 3D Gaze Interaction in Virtual Reality Environments,” JVRB - Journal of Virtual Reality and Broadcasting, 5(16), 2008.

[11] E. G. Mlot, H. Bahmani, S. Wahl, and E. Kasneci, “3D Gaze Estimation using Eye Vergence,” BIOSTEC, pp. 125-131, 2016.

[12] S. Weber, R. Schubert, S. Vogt, B. M. Velichkovsky, and S. Pannasch, “Gaze3DFix: Detecting 3D fixations with an ellipsoidal bounding volume,” Behavior Research Methods, vol. 50, no. 5, pp. 2004-2015, 2018.

[13] C. Elmadjian, P. Shukla, A. D. Tula, and C. H. Morimoto“3D gaze estimation in the scene volume with a head-mounted eye tracker,” Proceedings of the Workshop on Communication by Gaze Interaction (COGAIN '18), No. 3, pp. 1-9, 2018.

[14] S. Z. Öney et al., "Evaluation of Gaze Depth Estimation from Eye Tracking in Augmented Reality," ACM Symposium on Eye Tracking Research and Applications, pp. 1-5, 2020.

[15] K. Kato, O. D. A. Prima, and H. Ito, "Development of Practical-used 3D Eye Tracker", IPSJ – Research Conference, 80th, 1, pp.133-134, 2018.

[16] M. Kassner, W. Patera, and A. Bulling. "Pupil: an open source platform for pervasive eye tracking and mobile gaze-based interaction," Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct Publication, pp. 1151-1160, 2014.

[17] M. Kouchi and M. Mochimaru, "2008: Anthropometric Database of Japanese Head 2001," National Institute of Advanced Industrial Science and Technology, H16PRO-212.