eTELEMED 2021, The 13th International Conference on eHealth, Telemedicine, and Social Medicine

3D Gaze Characteristics in Mixed-Reality Environment

Kenta Kato and Oky Dicky Ardiansyah Prima

Graduate School of Software and Information Science Iwate Prefectural University (Japan)

eTELEMED 2021 | July 2021

g236s002@s.iwate-pu.ac.jp

■ About Me

KENTA KATO received the master's degrees in Software and Information Science from the Iwate Prefectural University, Japan, in 2018. I am currently a doctoral student at the Iwate Prefectural University developing 3D eye tracker. I am a member of the Information Processing Society of Japan.

Agenda

- Background
- Research Aims
- 3D Eye Tracking System
- Experiments
- Results
- Conclusion

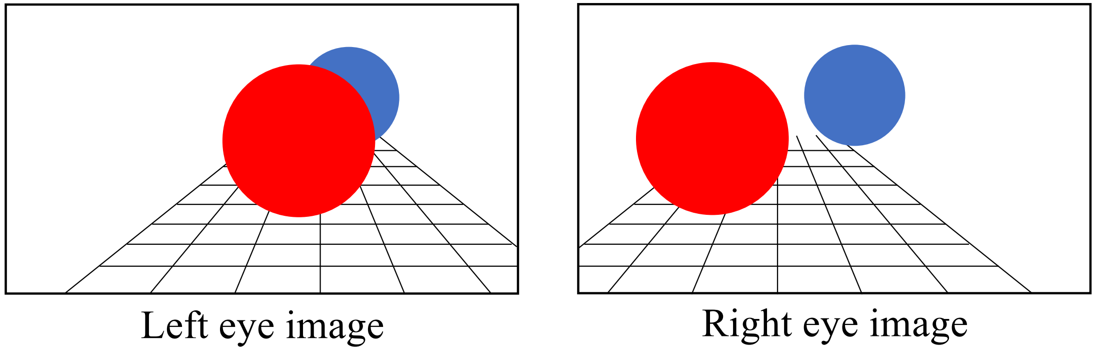
■ Background :: Depth Perception

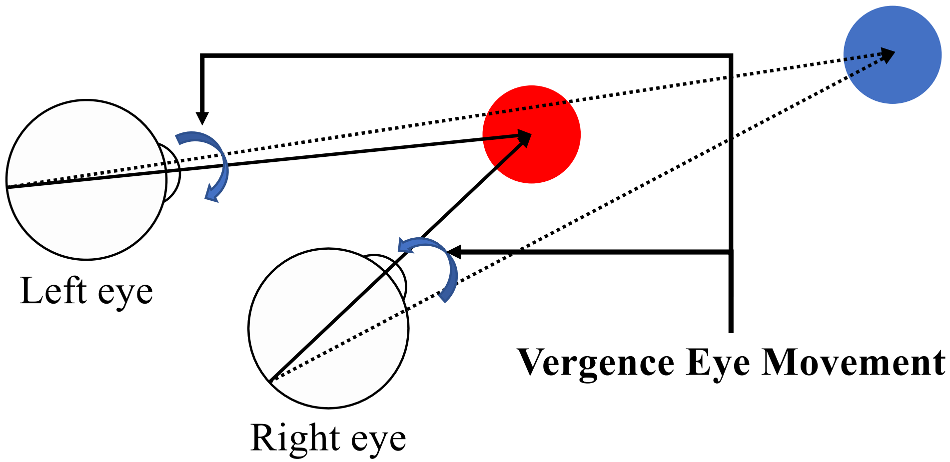
Binocular cues

Binocular cues are the perception of depth from the disparity caused by the different positions of the both eyes.

monocular cues

Monocular cues are the perception of depth through apparent changes caused by the position of objects.





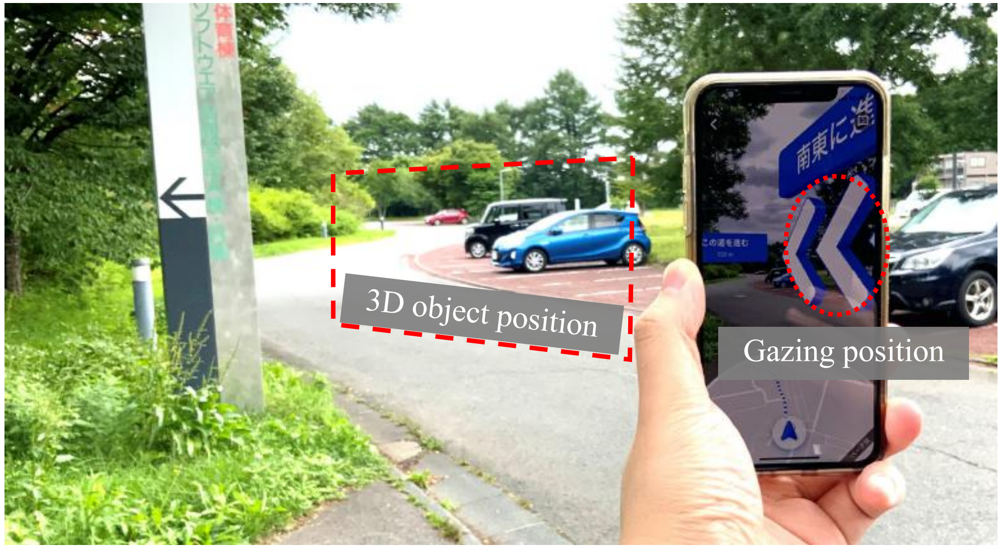
■ Background :: Difference between AR and MR

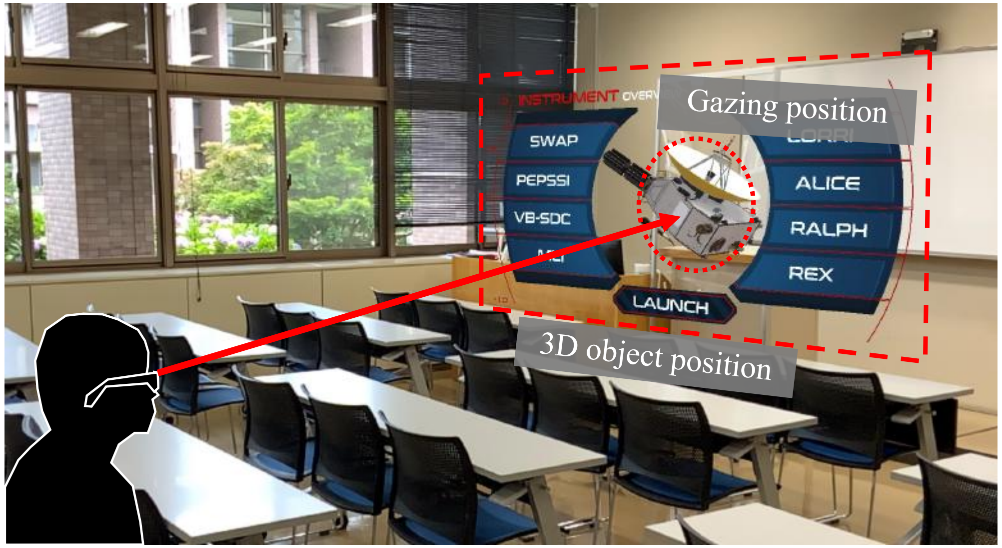
Augmented Reality (AR)

The 3D object position is recognized by monocular cues. It is not same as the gazing position.

Mixed Reality (MR)

The 3D object position coincide with the gazing position by binocular cues.



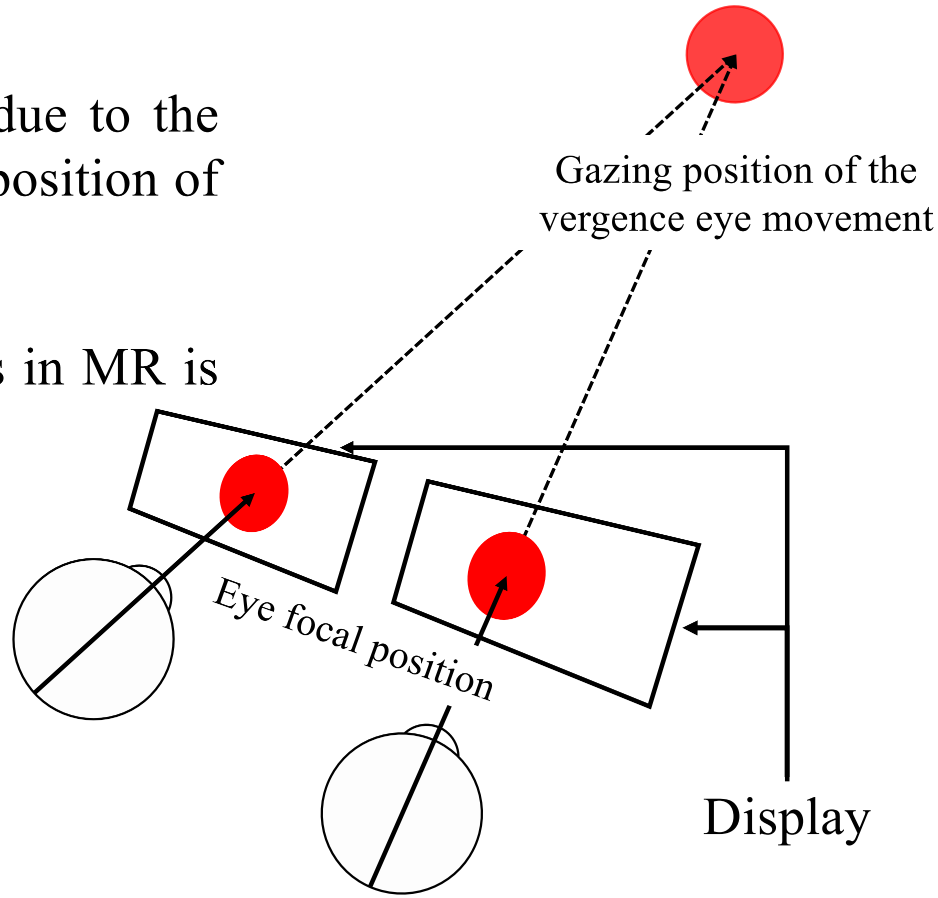


Iwate Prefectural University

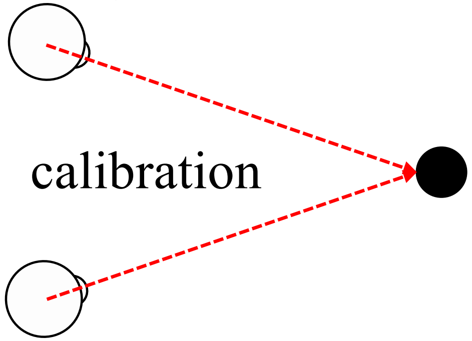
■ Background :: Visual Discomfort and Fatigue

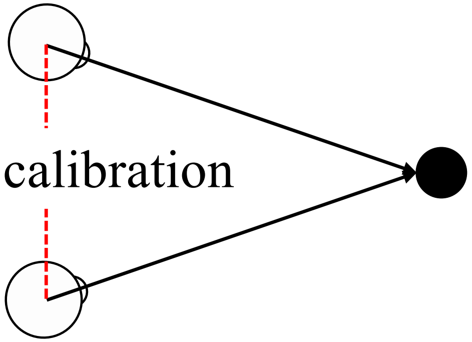
The MR device causes discomfort and fatigue due to the gap between the eye focal position and the gaze position of the vergence eye movement.

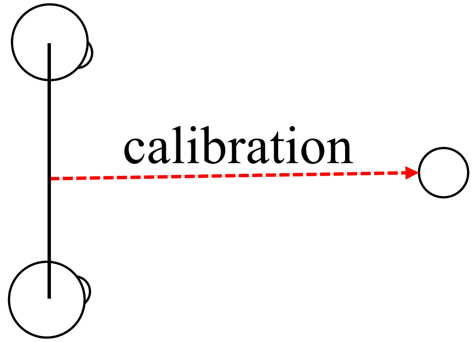
Analyzing the 3D gaze characteristics of viewers in MR is the key to solve these problems.



■ Background :: 3D Gaze Estimation Method





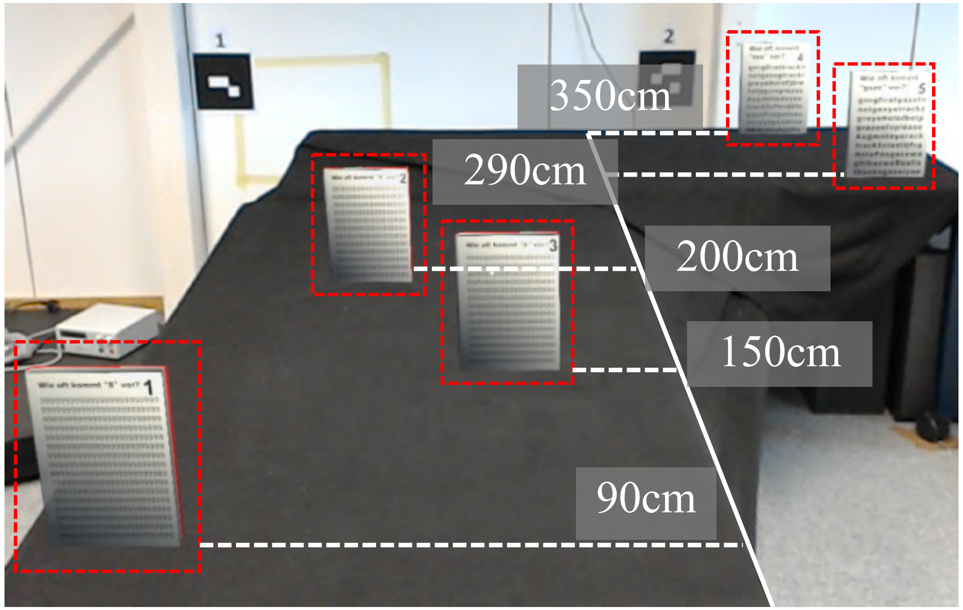


	Elmadjian et al (2018) [1]	Öney et al (2020) [2]	Kapp et al (2021) [3]
Measurement method	The gaze position is calculated by regression analysis of both gaze vectors and gazing targets.	The gaze position is the gaze vector extending from both calibrated eye centers.	The gaze position where the calibrated gaze vector hit the object.
Environment	Physical (Stationary)	MR (Stationary)	MR (Stationary)
Accuracy	Average 26.6cm (Range is 75cm~275cm)	Average 110cm (Range is 90cm~350cm)	0.91~5.03cm (2D) (Range is 50cm~400cm)

■ Background :: Analyzing 3D Gaze in MR environments

Öney et al (2020)

3D gaze information was measured during a visual search task of equivalently positioned 3D and physical objects.
No significantly difference in the resulting 3D gaze accuracies from 3D objects in MR and physical environments.
※The error of the 3D gaze measures more than 1 meter.



■ Research Aims

- To analyze the influence of the surrounding physical environment to 3D gaze measurement in MR.
- To analyze the characteristics of 3D gaze scanpaths of moving targets as well as stationary targets.
- To develop a 3D eye tracking system.

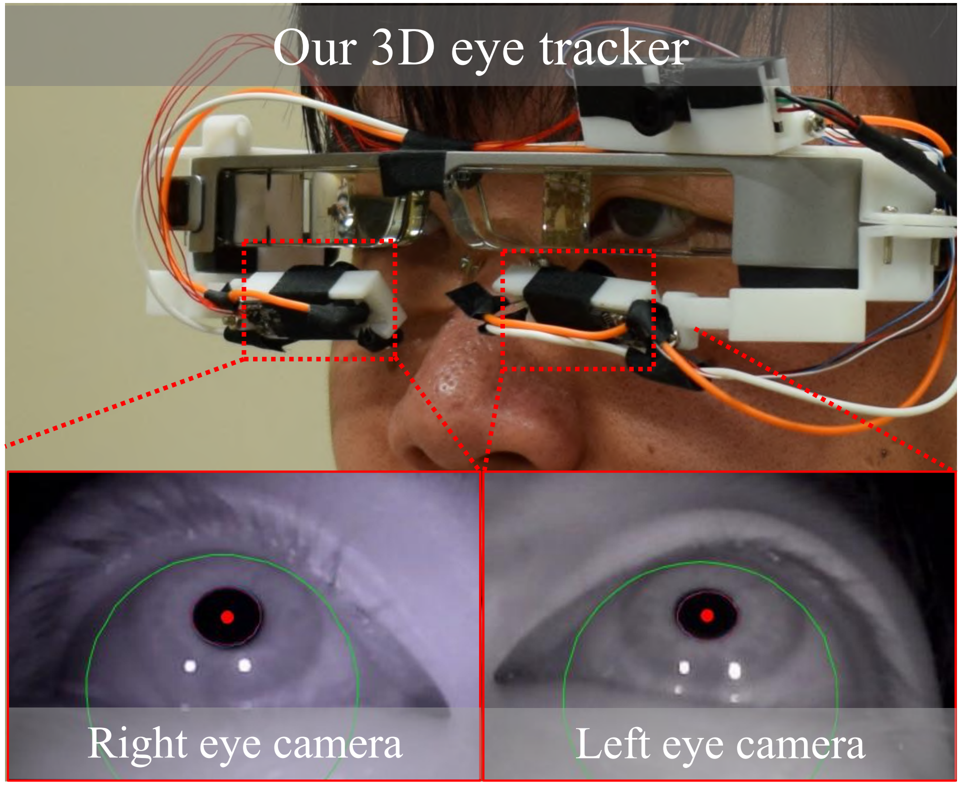
■ 3D Eye Tracking System

MR device (See-Through Head-Mounted Display; ST-HMD)

ST-HMD : Moverio BT-30E
Display resolution : 1280 × 720
Virtual screen size : 40inch (viewing distance of 250cm)

Eye tracking system

Camera : KYT-U030-3NF (KAYETON)
Eye camera resolution : 320 × 240
Capture FPS : 60Hz
Software : Pupil Capture (Pupil Labs) [4]

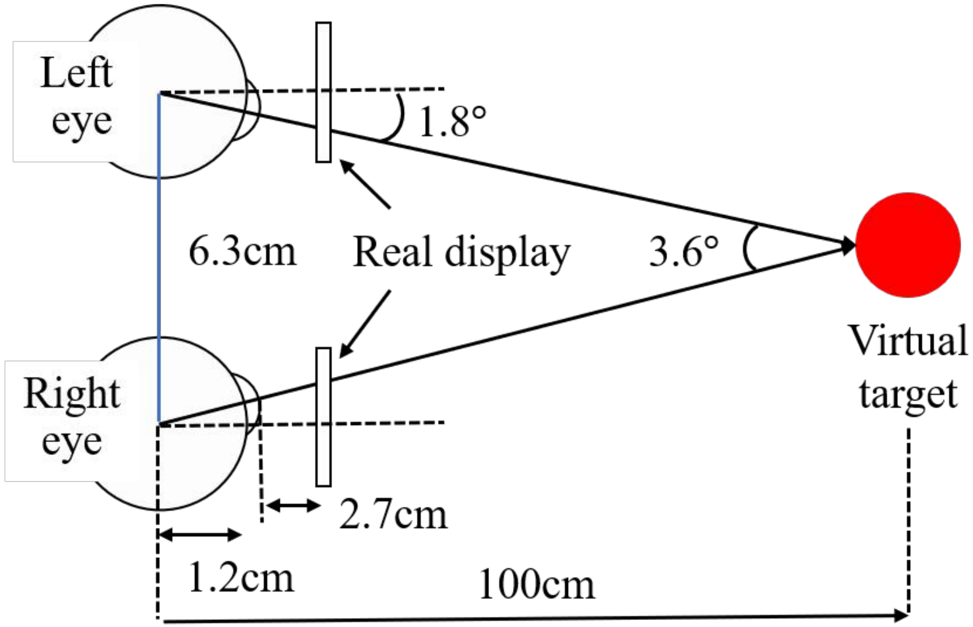


■ 3D Eye Tracking System

3D eye calibration using visual targets

The visual targets are created in the virtual environment to generate binocular disparity to induce vergence.

3D eye calibration is conducted using the visual targets placed in this setting.



The relationship between both eyes and the visual target

※ Interpupillary distance is defined by the Japanese average.
※ No significant differences by gender, race, or age group, the eye radius was defined as the adult's average.

■ 3D Eye Tracking System

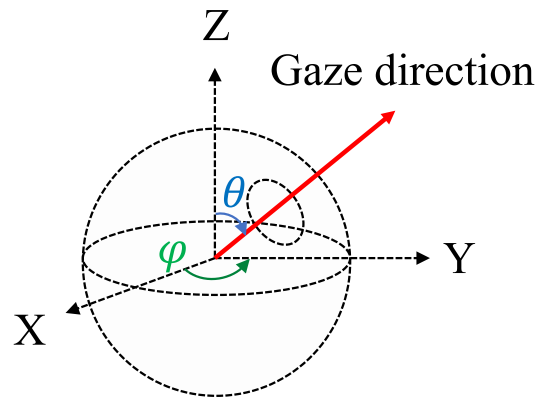
3D Gaze Estimation

We use the following polynomial is used to calculate the 3D gaze.
Coefficients are calculated by the least-squares method.

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

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

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

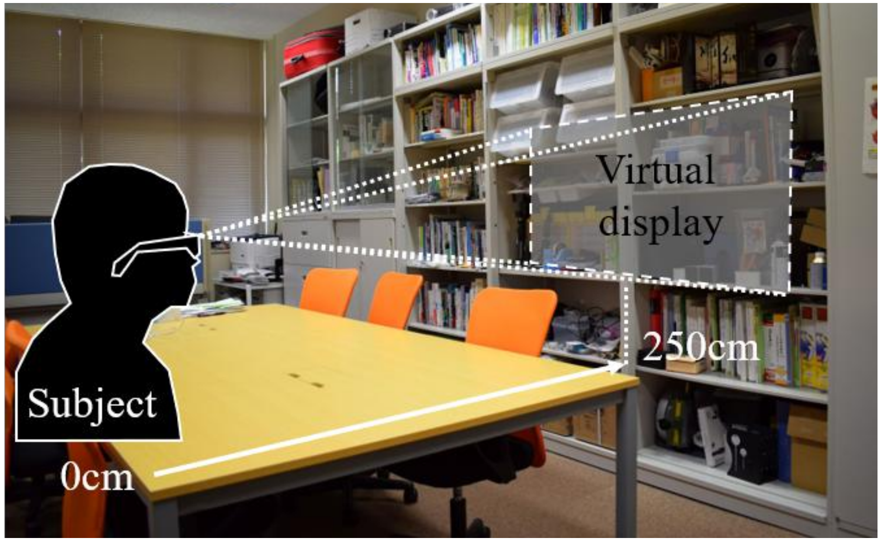


Eyeball model

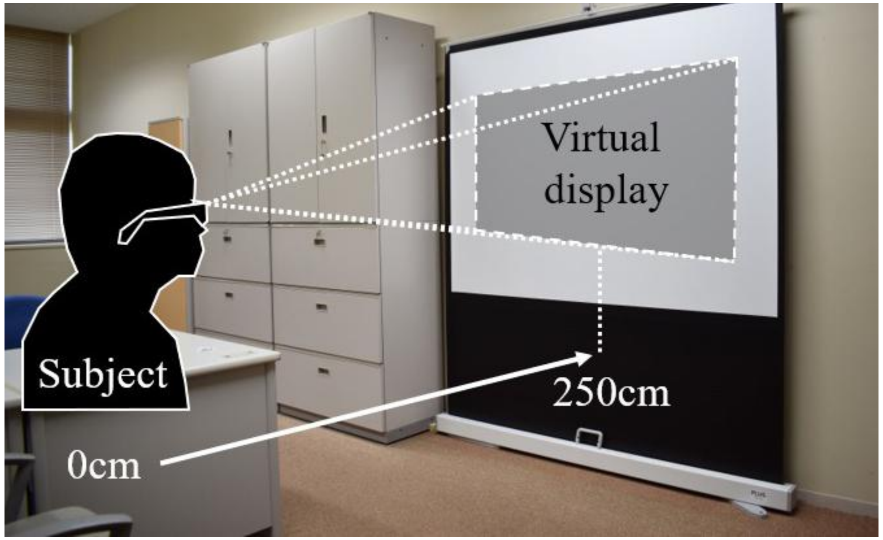
Eyeball pitch angle (θ_l, θ_r)
Eyeball yaw angle (φ_l, φ_r)
Coefficients ($a_1 \sim a_{15}, b_1 \sim b_{15}, c_1 \sim c_{15}$)
3D gaze position (G_x, G_y, G_z)

■ Experiment

We conducted gazing experiments in viewing environments with and without depth cues. Four subjects participated in the experiment (male, mean age 23.3, vision acuity 1.0 or better, no health concerns).



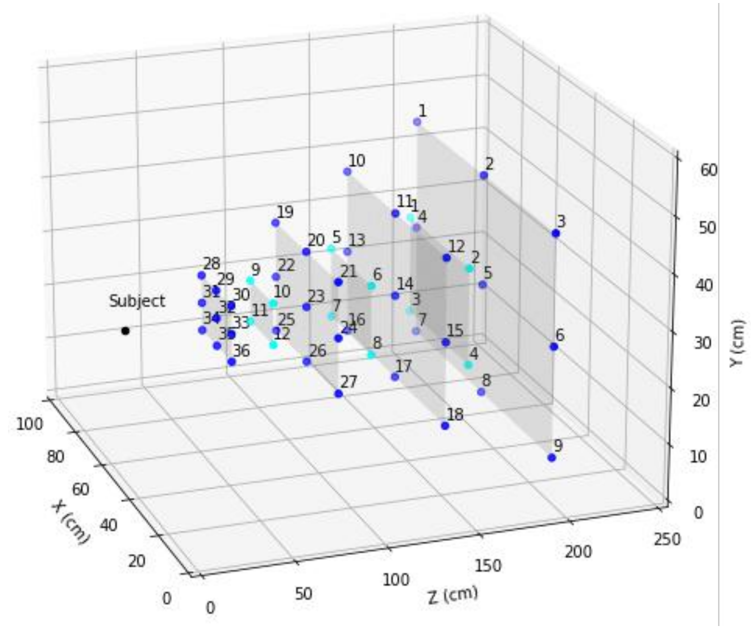
A room **WITH** depth cues



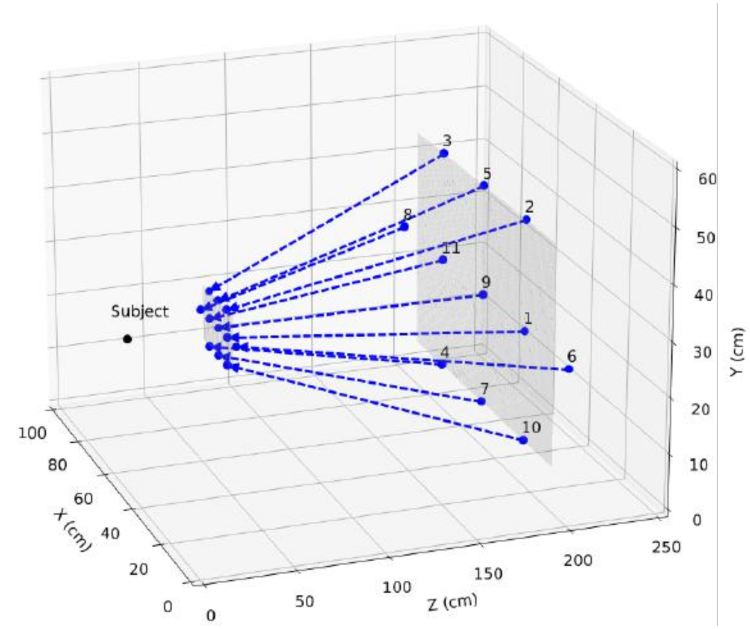
A room **WITHOUT** depth cues

■ Experiment

The experiment consists of the following two tasks.
Task.1 Gazing at 48 **stationary** visual targets (include calibrations and validations task).
Task.2 Gazing at 11 **moving** visual targets.



Visual targets for calibrations and validations



The back-to-front moving visual targets

■ Results

A. 3D gaze accuracy in this study

The 3D gaze measurement accuracy Acc_{3D} is measured by

$$Acc_{3D} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(T_{xi} - G_{xi})^2 + \left(T_{yi} - G_{yi}\right)^2 + (T_{zi} - G_{zi})^2}$$

Visual target position (G_x, G_y, G_z)

3D gaze position (G_x, G_y, G_z)

Accuracies for the 3D gaze measurement (cm)

Subject	With depth cues		Without depth cues	
	Calibration	Validation	Calibration	Validation
1	21.38	24.74	11.48	21.73
2	20.03	23.68	26.65	26.77
3	7.40	11.07	12.85	20.74
4	22.55	34.79	9.78	22.92
Mean	17.84	23.57	15.19	23.04
Std. Dev.	7.037	9.721	7.744	2.643

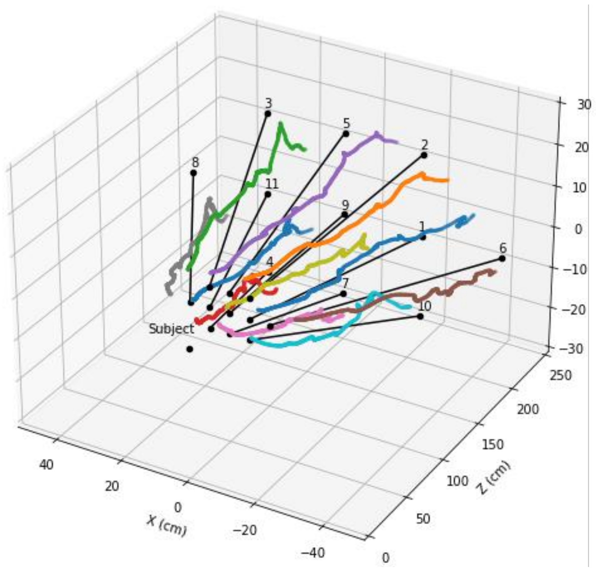
These accuracies were analyzed with two-way Analysis of Variance.
Effects were not found in either the depth cues (F (1, 12) = 0.192, p = .669) or the visual targets (F (1, 12) = 3.497, p = .086).

Iwate Prefectural University

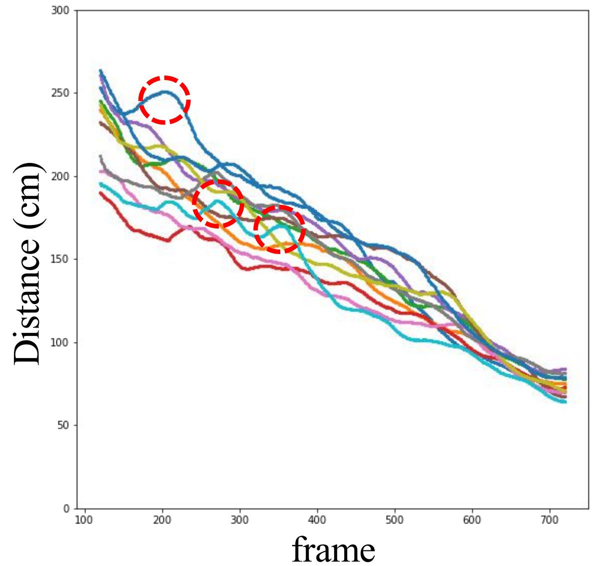
■ Results

B. 3D gaze for moving targets

As the visual target approached the viewer, the resulting 3D scanpaths were expected to follow these approach paths. However, at some distances, the resulting scanpaths did not show a monotonic decrease.



The path where each target moves and its gazing position



Difference in gazing distance with movement of the target

Iwate Prefectural University

■ Conclusion

Achievements

- Achieving an average 3D gaze accuracy within 25cm.
- No significant difference in the measured 3D gaze between rooms with and without depth cues.
- By measuring the 3D gaze scanpath for the back-to-front moving visual targets, we found that, at some distances, the resulting scanpaths did not show a monotonic decrease.

Future works

- We would like to stabilize the calibration accuracy as well as to analyze the factors that caused the accuracy difference between subjects.
- We would like to measure and analyze the scanpaths for other 3D visual target movements to clarify the physiological characteristics.

References

1. 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, doi:10.1145/3206343.3206351
2. 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, doi:10.1145/3379156.3391835.
3. S. Kapp, M. Barz, S. Mukhametov, D. Sonntag, and J. Kuhn, “ARETT: Augmented Reality Eye Tracking Toolkit for Head Mounted Displays,” Sensors, vol. 21, no. 6, 2234, 2021, doi: 10.3390/s21062234
4. 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, doi:10.1145/2638728.2641695.

Iwate Prefectural University