# Gaze Calibration of Eye Trackers for Head-Mounted Displays Using Eye-Frontalization Process

Katsuyoshi Hotta, Oky Dicky Ardiansyah Prima,
Graduate School of Software and Information Science
Iwate Prefectural University
Takizawa, Japan
e-mail: prima@iwate-pu.ac.jp, g236q004@s.iwate-pu.ac.jp

Takashi Imabuchi

Japan Atomic Energy Agency
Fukushima, Japan
e-mail: imabuchi.takashi@jaea.go.jp

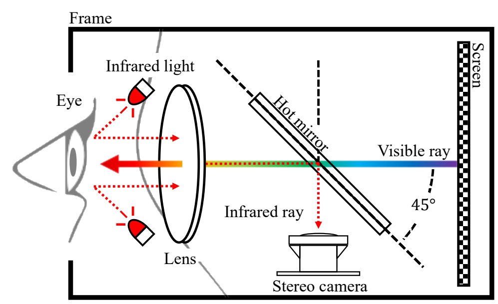
***Abstract*— In order to simplify the gaze calibration required for gaze measurement in Virtual Reality Head-Mounted Displays (VR-HMDs), we proposed a new gaze calibration method that combines eye-frontalization and single-point calibration. Deep Convolutional Neural Network was used for the frontalization of the eyes. Our method enables to estimate the gaze coordinates on the screen as soon as the user puts on the VR-HMD and to improve the accuracy of gaze measurement due to slight changes in the positional relationship between the VR-HMD and the face by updating the parameters of the gaze calibration. The accuracy of the proposed method was about 5 degrees for both eyes.**

***Keywords-VR; EyeTracking; Gaze Calibration; Gaze redirection.***

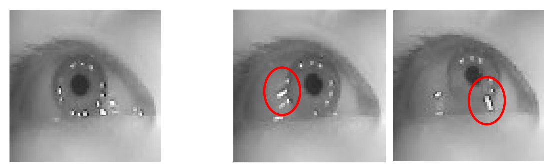
## I. INTRODUCTION

Gaze information is essential in research and technology using visual processing. This information provides insight into the characteristics of the object of interest, gazing order, and eye movements. Although we can control the gaze position arbitrarily, eye movements associated with changes in the gaze position are involuntary [1]. Therefore, eye movements are used as response indices to stimuli presented on a screen [2]. In the medical field, these indices are used to determine visual acuity for visual field testing [3] and as input interfaces for Amyotrophic Lateral Sclerosis (ALS) patients [4]. The information of eye movements also has been applied to rendering processing [5] and industrial robot control [6].

Recently, there are Virtual Reality Head-Mounted Displays (VR-HMDs) equipped with eye trackers, such as HTC Vive Pro and Fove, which enable eye tracking in Virtual Reality (VR). These devices have a wider viewing angle of the stimulus and provide more stable eye tracking than non-contact eye trackers. However, if the position of the VR-HMD during the initial gaze calibration and the subsequent position of the VR-HMD relative to the face change, the accuracy of the gaze measurement will be degraded. This problem is caused by its inability to compensate for the gaze point using the corneal reflection, as in non-contact eye trackers. Figure 1 shows the eye tracker mechanism in the VR-HMD and the corneal reflection image indicating that the corneal reflection image cannot be captured correctly when gazing at a wide viewing angle.




(a) The structure of the VR-HMD eye tracker.



Looking at the front. Looking at an angle.

(b) Corneal reflection images

Figure 1. Eye tracker mounted on the VR-HMD and corneal reflection images captured by the tracker's camera.

In this study, we propose a new gaze calibration method that can compensate the accuracy of gaze measurement for changes in the position of the VR-HMD relative to the face. The proposed method consists of eye-frontalization and single-point calibration. The former involves inferring the image of an eye gazing in the frontal direction from the image of an eye gazing in an arbitrary direction. This method was inspired by the study of gaze redirection [7]. The latter is a higher-order polynomial that connects the gaze direction of the frontal gaze to the direction of the surroundings. This polynomial allows to estimate the gaze point when gazing at the screen (Point of Regard; PoR) without gaze calibration if the image of the eye while looking at the center of the screen can be obtained [8].

Our proposed method will have the following advantages. First, since the single-point calibration can be performed automatically by frontalizing the eyes, the PoR can be obtained instantly after the user wears the VR-HMD. Secondly, the gaze accuracy will not be degraded even if the

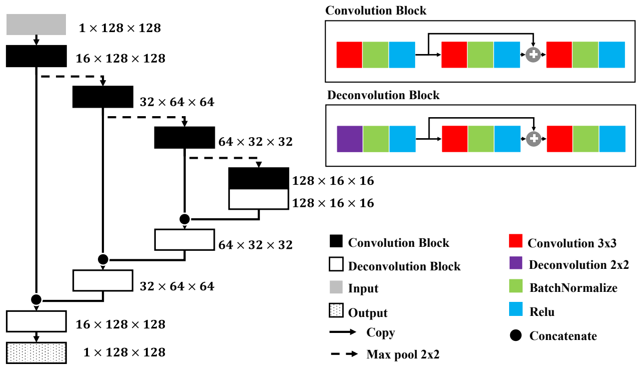



Figure 2. The convolution neural network for the eye-frontalization.

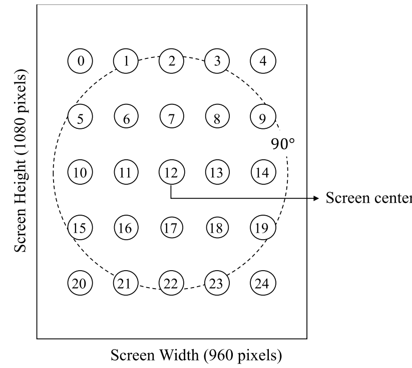



Figure 3. 25 visual targets for data collection in this study.

positional relationship between the VR-HMD and the face changes, thus enabling gaze measurement in long-term.

This paper is organized as follows. In Section II, we describe related work on the frontalization process of eye images. In Section III, we describe the frontalization process and how to obtain the gaze position. In Section IV, we present the validation methods and results on reconstruction error and gaze accuracy. Finally, the results of this research and future perspectives are presented.

## II. Related Work

Gaze redirection is a method of modifying the gaze in any direction for a given eye image. This method is an emerging research topic in computer vision and computer graphics, especially in applications such as the generation of eye contact in video conferencing to improve communication.

The initial work on gaze redirection was conducted by Wolf et al. [7]. Their method detects the position of an eye and replaces it with the image of the most similar eye in the dataset, looking straight at the camera. A simple deformable eye model was used to adjust the size and orientation of the eye area. The limitations of this method are that the results depend on the accuracy of the eye region detection and the inability to handle large head movements. Wood et al. [9] proposed a 3D model of a large eye region covering the eyelids and lower eye area to achieve a more natural eye resemblance. The fitting of the 3D model to the eye region was done by minimizing the energy in the similarity of both images and landmarks of the model and the target eye. The disadvantages of this method are the high computational cost that makes it impractical to run in real-time. Instead of using a large 3D model, Isikdogan et al. [10] used the Eye Correction Model Network (ECC-NET) to guide the gaze from the eye region cropped by the facial landmark detector and applied the Generative Adversarial Network (GAN) to the U-Net based network to generate natural-synthetic eye images. Moreover, in order to stabilize the output image, temporal filtering was performed before generating the final output.

TABLE I. Specifications of Experimental Equipment

| Device | Element | Specification |
|---|---|---|
| VR-HMD Eyetracker | FoV | 90° |
| | Resolution(Mono) | 960 × 1080 |
| | Luminance | 250cd/m2 |
| | Sampling Rate | 240Hz |
| Desktop PC | CPU | i7-9700K(3.6GHz) |
| | GPU | RTX2080(8GB) |
| | Memory | 16GB |
| | OS | Windows 10 64bit(Ver.2004) |

## III. Proposed Gaze Calibration Method

The proposed gaze calibration method consists of three processes: eye-frontalization, pupil center localization, and PoR calculation.

### A. *Eye frontalization*

We build a convolutional neural network (CNN) that performs eye-frontalization by referring to ECC-Net, which is based on U-Net. This network uses a Pre-Activation Module (PAM) [11], which makes it easier to inherit local features. This not only lowers the computational cost but also allows for a better output of eye images. Figure 2 shows the CNN build for this study. The input eye image is set to 128x128 pixels. L1Loss (absolute error) is used for the loss function of the network, as in previous studies.

### B. *Getting the pupil center coordinates*

Blob searching is commonly used to localize the pupil center coordinates from an eye image. By approximating the shape of the blob with ellipse, the center of the resulting ellipse is treated as the center of the pupil. The brightness of the pupil region, however, is not constant due to eyelashes, eyelid shadows, and Purkinje images. Hotta et al. [8] used Semantic Segmentation (SS) to generate a blob image from an eye image, and then used the Starburst algorithm to find the center of the blob [12]. This method provides a stable

 ISBN: 978-1-61208-872-3 

(a) Input eye image (b) Eye-frontalized image (c) The actual eye image gazing the center screen

Figure 4. The input eye image, the resulting frontalized image, and the image of the eye gazing at the center screen.

result of the center of the pupil. In this study, we adopted the method of Hotta et al. [8] to determine the coordinates of the pupil center.

### C. PoR calculation

In this study, the PoR is obtained by applying single-point calibration [8] to the coordinates of the pupil center. The relationship between the pupil center coordinates $(x, y)$ after the eye-frontalization and the screen coordinates $(u, v)$ is computed using the following pre-determined $n$-order polynomial obtained by the single-point calibration experiment.

$$u = \sum_{j=0}^{n}\sum_{k=0}^{j} a_{j,(j-k)} x^j y^{j-k}, \quad v = \sum_{j=0}^{n}\sum_{k=0}^{j} b_{j,(j-k)} x^j y^{j-k} \tag{1}$$

The coefficients $a_{j,(j-k)}$ and $b_{j,(j-k)}$ are obtained by the least-squares method. Finally, the current gazing point $(u', v')$ on the screen (PoR) is calculated by

$$\begin{bmatrix} u' \\ v' \end{bmatrix} = \begin{bmatrix} u' \\ v' \end{bmatrix} + \begin{bmatrix} t_x \\ t_y \end{bmatrix}. \tag{2}$$

Here, $(t_x, t_y)$ represents the horizontal and vertical difference between $(u, v)$ and the screen center coordinate.

## IV. Experiment and Result

36 subjects participated in this experiment. All of the subjects had no visual acuity problems. The experimental setting, data collection, and accuracies of the eye-frontalization and PoR are described as follows.

### A. Environmental setting

Table I shows the specifications of VR-HMD and the desktop PC used in the experiment. The software application programs for frontalizing the eye and displaying the stimuli during the experiment were implemented using openFrameworks (C/C++).

### B. Data collection and training the CNN

To train the CNN for the frontalization process, we collected images of eyes gazing at the center screen and other locations on the screen. We collected images of all subjects' eyes gazing at 25 visual targets on the screen, as shown in Figure 3. Each subject was requested to fixate on 25 targets three to four times, resulting 4,000 eye images. These images were used for training. Figure 4 shows the input eye image, the resulting frontalized image, and the image of the eye gazing at the center screen.

### C. Eye-frontalization accuracy

The accuracy of the frontalized eye image is not determined by the difference in its appearance relative to the image of the eye gazing at the center screen. Instead, both pupil centers are extracted, and the difference in their distances (deviations) is calculated as a measure of eye-frontalization accuracy. Figure 5 shows the distribution of horizontal and vertical deviations for the frontalized left and right eyes. The mean deviation for this experiment was 1.67 ± 0.98 pixels and 1.74 ± 1.04 pixels for the left and right eyes, respectively.

### D. Accuracy of PoR

We calculated the PoRs using equation (2) from the images of the eyes when gazing at 24 targets, excluding the center of the screen. The accuracy of PoRs is calculated by

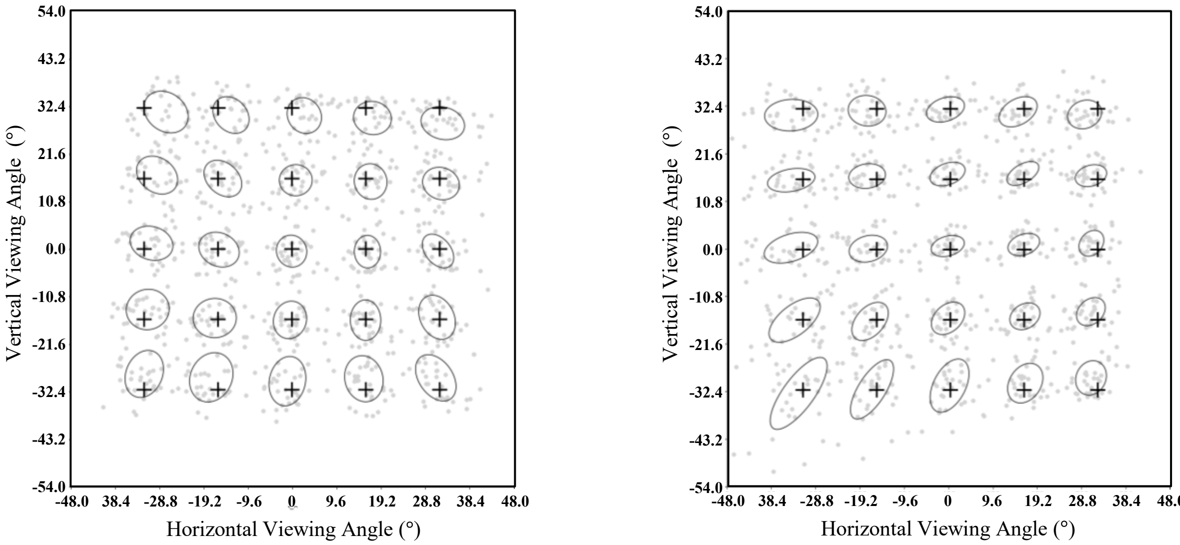
$$Accuracy = \frac{1}{N}\sum_{i=1}^{N}\sqrt{\left(Target_{x_i} - PoR_{x_{i,j}}\right)^2 + \left(Target_{y_i} - PoR_{y_{i,j}}\right)^2} \tag{3}$$

Here, $Target_{x_i}$ and $Target_{y_i}$ are the abscissa and the ordinate of the $i$-th target, respectively. Likewise, $PoR_{x_{i,j}}$ and $PoR_{y_{i,j}}$ are the abscissa and the ordinate of the $i$-th PoR, respectively. $N$ is the total number of PoRs.

The accuracy of PoR was 5.07±3.30° and 5.50±3.25° for the left and right eyes, respectively. Figure 6 shows the distribution of PoRs for all subjects. As the accuracy of the typical eye Tracker is approximately 1°, we can say that the accuracy of the proposed method is insufficient. However, it is acceptable for the first attempt of automatic gaze calibration.

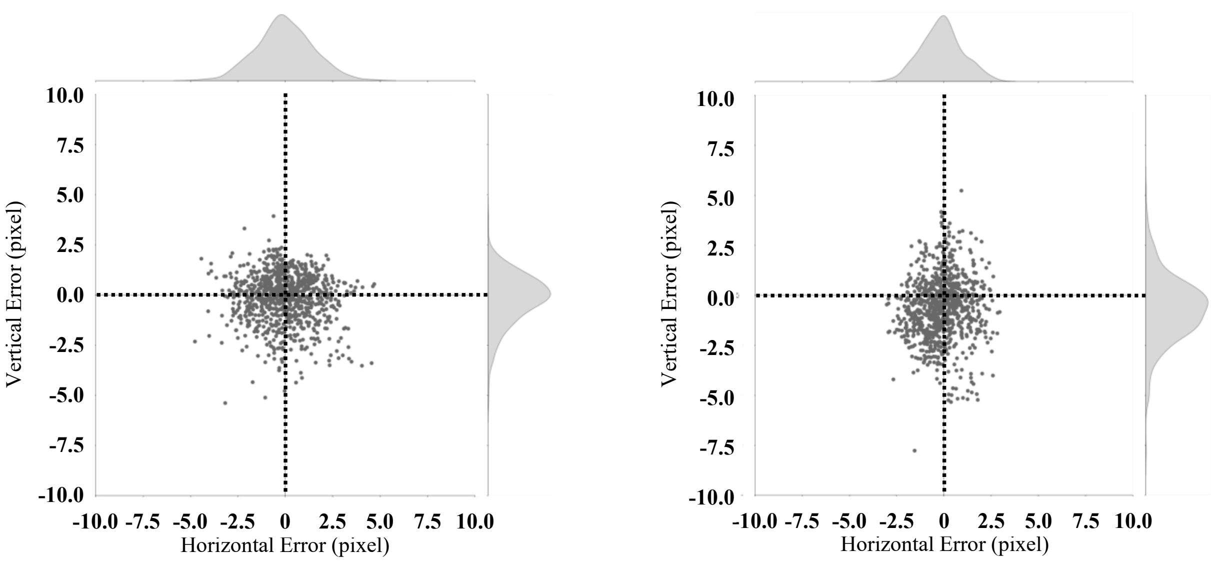
## V. Conclusion and Future Work

In this study, we have proposed a new gaze calibration method that combines eye-frontalization and single-point




(a) Left eye (b) Right eye

Figure 6. The distribution of PoRs for all subjects calculated by the proposed method.




(a) Left eye (b) Right eye

Figure 5. the distribution of horizontal and vertical deviations for the frontalized left and right eyes.

calibration. Since the accuracy of a typical eye tracker is about 1°, the accuracy of the proposed method can be considered low, but it is within the acceptable range for the first attempt of automatic gaze calibration. The proposed gaze calibration method can be used to compensate for the accuracy of gaze measurement even when the position of the VR-HMD relative to the face changes.

In the future, we will continue to improve the calibration method to make the eye tracking system for head-mounted displays easier to use and more accurate.

## REFERENCES

[1] R. Monty and J. W. Senders, Eye Movements and Psychological Processes. Routledge, 2019.

[2] A. Talk, I. Ant-Méndez, and B. Pennefather, "Graded expression of source memory revealed by analysis of gaze direction," PLoS One, vol. 12, pp. e0188727, 2017, doi:10.1371/journal.pone.0188727.

[3] I. Murray et al., "Saccadic Vector Optokinetic Perimetry (SVOP): a novel technique for automated static perimetry in children using eye tracking," Annu Int Conf IEEE Eng Med Biol Soc, pp.3186-3189, 2013, doi:10.1109/EMBC.2013.6610218.

[4] A. Calvo et al., “Eye Tracking Impact on Quality-of-Life of ALS Patients,” 11th International Conference on Computers Helping People with Special Needs (ICCHP 2008), Jul. 2008, pp. 70-77.

[5] P. Anjul et al., “Towards foveated rendering for gaze-tracked virtual reality,” Association for Computing Machinery, vol. 35, no. 6, pp. 1–12, 2016, doi:10.1145/2980179.2980246.

[6] R. Alessandro, P. Ugo, M. Giorgio, and N. Lorenzo, “A Cartesian 6-DoF Gaze Controller for Humanoid Robots,” Proceedings of Robotics: Science and Systems, Jun. 2016, doi: 10.15607/RSS.2016.XII.022.

[7] L. Wolf, Z. Freund, and S. Avidan, “An eye for an eye: A single camera gaze-replacement method,” Computer Vision and Pattern Recognition (CVPR 2010), Jun. 2010, pp. 817–824, doi: 10.1109/CVPR.2010.5540133.

[8] K. Hotta, O. D. A. Prima, T. Imabuchi, and M. Kameda, “Development of High-Performance Visual Field Tester Based on The Eye Movement,” The Institute of Image Electronics Engineers of Japan, vol. 50, no. 3, 2021.

[9] E. Wood, T. Baltrusaitis, L. P. Morency, P. Robinson, and A. Bulling, “GazeDirector: Fully Articulated Eye Gaze Redirection in Video,” Computer Graphics Forum, vol.37, no.2, pp. 217-225, 2018, doi: 10.1111/cgf.13355

[10] I. F. Isikdogan, G. Timo, and M. Gilad, “Eye contact correction using deep neural networks,” Winter Conference on Applications of Computer Vision (WACV, 2020), Mar. 2020, doi:10.1109/wacv45572.2020.9093554.

[11] K. He, X. Zhang, S. Ren, and J. Sun ,“Identity Mappings in Deep Residual Networks,” European Conference on Computer Vision (ECCV 2016), Mar. 16, vol.9908, ISBN : 978-3-319-46492-3.

[12] D. Li, D. Winfield, and D. J. Parkhurst, “Starburst: A hybrid algorithm for video-based eye tracking combining feature-based and model-based approaches,” Computer Society Conference on Computer Vision and Pattern Recognition (CVPR 2005) - Workshops, Sep. 2005, pp. 79–79, doi:10.1109/CVPR.2005.531.