Simple calibration method for low-cost eye-tracker

Takehito Kikuchi, Manyo Takenaka, and Yusuke Fujii
Faculty of Science and Engineering, Oita University
Dannoharu 700
Oita Japan
+81-97-554-7771
t-kikuchi@oita-u.ac.jp

ABSTRACT

Eye-tracking devices generally require calibration processes in which users gaze four or more points displayed on a monitor. However, it is hard to accomplish such complicated processes for infants or patients who have severe mental disabilities. In order to reduce such difficulty, we suggested a simple calibration method for a low-cost eye-tracking device. In this method, we use a set of calibration data previously made by a model person. As a post process, users are required to perform a simple calibration process in which they gaze only one point displayed on the center of a monitor. The EyeTribe was used in this study. This paper describes the theory and evaluation results of this method. The experimental results show the suggested method reduced the error in the gaze points for male adult subjects.

Categories and Subject Descriptors

B.4.2 [**Hardware**]: Input / output and data communications – *Input / output devices.*

General Terms

Measurement

Keywords

Eye-tracking, gaze point, calibration method

1. INTRODUCTION

Eye-tracking technologies [1] are useful not only for analyses of infant behaviors [2] and autism [3,4], but also for computer interfaces for patients who have severe physical disabilities [5]. Up until a decade ago, high cost of the eye-tracking devices had prevented general uses. However, low-cost eye-tracking devices have been recently released from several companies, and it drives the general uses. In particular, the EyeTribe [6] costs only around 200 USD, and is ready for the open-source development. Therefore, several researchers have experimentally investigated its performances and developed application software [7, 8].

Generally available eye-tracking devices often utilize a corneal reflection method [1] as their basic principles. This method requires calibration processes in which the users gaze four or more points displayed on a monitor [6], in order to associate the optical properties of pupil images with the coordinates of the gaze points on a monitor. However, it is hard to accomplish such complicated processes for infants and mentally disabled users.

To solve this problem, we suggested a simple calibration method for eye-tracking devices. The EyeTribe was used in this study. This paper describes the theory and evaluation results of this method.

2. EyeTribe

Table 1 shows types of measureable information with the EyeTribe [6]. This sensor originally acquire an infrared image of user's eyes in real time. After a normal calibration process (nine-points gazing), the sensor also calculate the coordinates of a gaze point and center of pupils.

The coordination system of the gaze point is located on the monitor or screen used in the calibration process. Their units are pixel. On the other hand, the coordination system of the pupil position is located on a standard plane in 65 cm (L_s) front of the sensor. The sizes of the standard plane are 40 cm in width (W_s) and 30 cm in height (H_s). The coordinates of each pupil are normalized with these width and height.

In this paper, the unit of all distances are "cm". In addition, the sensor looks at the user's eyes straight ahead. According to our previous experimental results, the measurable range of the sensor is 50-85 cm in horizontal distance (L_h), and 0-15 degree in elevation angle (θ) of Fig.1.

Table 1. Measureable information of EyeTribe

(MyGaze Class [6])

Items	Meanings
Gaze (X, Y)	Coordinates of gaze point on a monitor Unit: pixel
Center of Pupils (X, Y)	Normalized coordinates of each pupil (left and right) on a standard plane Unit: none (0-1)

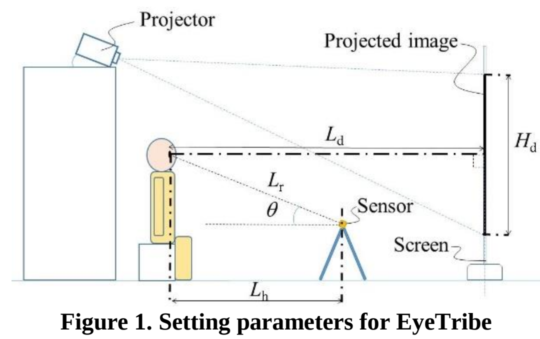


Figure 1. Setting parameters for EyeTribe

3. Proposal of a simple calibration method

3.1 Estimation of sensor/display setting

The real sensor distance (L_r) in Fig.1 is calculable from the data of the EyeTribe (Table 1), if we previously know the real distance between the left and right pupil (δ_r) of a subject. We can measure this value even for infants by using a captured image of his/her face with a scale.

The distance between the left and right pupil on the standard plane mentioned in the section 2 (δ_s) is calculable as follows;

$$\delta_s = \frac{L_s}{L_r}\delta_r, \tag{1}$$

where L_s is 65 cm as mentioned in the section 2. The normalized distance between both pupils (δ) is directly measureable with the EyeTribe (Table 1) and also defined as follows;

$$\delta = \frac{\delta_s}{W_s}. \tag{2}$$

where W_s is 40 cm as mentioned in the section 2. Here, we assumed that a line connecting both pupils is parallel to the X-axis of the coordination system. From Eq.(1) and (2), the sensor distance is calculable as follows;

$$L_r = \frac{L_s}{W_s}\cdot\frac{\delta_r}{\delta} = \frac{60}{45}\cdot\frac{\delta_r}{\delta}. \tag{3}$$

We conducted a set of experiments for a male adult (δ_r =6.4 cm) with a setup shown in Fig.1, where the distance was 55-80 cm and the angle was 5-15 degree. As results, the estimated distance with Eq.(3) included a systematic error of around 6 cm (Fig.2). Therefore, we modified the estimated distance as follows;

$$L'_r = 1.04L_r - 6.23, \tag{4}$$

where L'_s is the modified estimated distance.

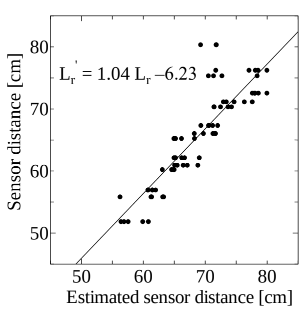


Figure 2. Real sensor distance vs. estimated distance

The elevation angle (θ) of Fig.1 is not calculable in the situation where the sensor and display are separately located. In almost all of real situations, the sensor is attached on the vertical centerline and lower edge of the monitor. In this situation, the distance between a subject and the display (L_d) is equal to L_h and, the elevation angle is calculable as follows;

$$\theta = \sin^{-1}\left(\frac{H_d/2}{L'_r}\right), \tag{5}$$

where H_d is a height of the display. In addition,

$$L_d = L'_r \cos(\theta). \tag{6}$$

3.2 Modification on the display size

Once the calibration process was done, the acquired gaze data are based on the original coordination system (a display used in the calibration process). When we use a different monitor with the same calibration data, we can convert the correct gaze data on the real monitor by using geometrical information of the monitors. Here we assumed the real monitor is smaller than the calibration one.

We defined L_c, W_c, and H_c as a distance, width, and height of the monitor used in the calibration process, respectively. In addition, L_d, W_d, and H_d in a real situation. The width and height of the real monitor have relatively the following sizes at the same distance of the calibration monitor;

$$W'_d = \frac{L_c}{L_d}W_d, \tag{7}$$

$$H'_d = \frac{L_c}{L_d}H_d. \tag{8}$$

Here, we can calculate the coordinates of a gaze point on the real monitor (X_d, Y_d) [pixel] from those on the calibration monitor (X_c, Y_c) [pixel] as follows;

$$X_d = \left(X_c - \frac{\Delta_W}{W_c}W_{pc}\right)\cdot\frac{W_{pd}W_c}{W_{pc}W'_d}, \tag{9}$$

$$Y_d = \left(Y_c - \frac{\Delta_H}{H_c}H_{pc}\right)\cdot\frac{H_{pd}H_c}{H_{pc}H'_d}, \tag{10}$$

where, H_{pc} and W_{pc} are resolution [pixel] of the calibration monitor, and H_{pd} and W_{pd} are those of the real monitor. Moreover,

$$\Delta_W = \frac{W_c - W'_d}{2}, \tag{11}$$

$$\Delta_H = \frac{H_c - H'_d}{2}. \tag{12}$$

3.3 Modeling of errors

The original gaze data (X, Y) include errors against the correct data on the calibration monitor (X_c, Y_c) when the sensor settings (L_r and θ) are different from the calibration conditions. To model the error tendency, we use a liner equation and investigate the relationship between the sensor settings;

$$X_c = a_x X + b_x, \tag{13}$$

$$Y_c = a_y Y + b_y, \tag{14}$$

where a_x, b_x, a_y, and b_y are modeling coefficients.

We conducted a set of experiments for a male adult (δ_r =6.4 cm) with a setup shown in Fig.1, where the distance was 55-80 cm and the angle was 5-15 degree. Previously the subject did a nine-point calibration on a displayed image whose L_c, W_c, and H_c were 120, 121, and 68 cm, respectively. The sensor settings L_r and θ in the calibration were 50 cm and 15°, respectively.

As the experimental results, a_x has no tendency with the sensor settings and almost 1 in all cases (Fig.3). Therefore, we assumed 1 in all cases. For a_y , it has a linear tendency in the elevation angle of the sensor (Fig.4). Then, we modeled as follows;

$$a_y = 0.0382\theta + 0.613. \tag{15}$$

For b_x and b_y, we decided to calibrate with a simple calibration method in the next section.

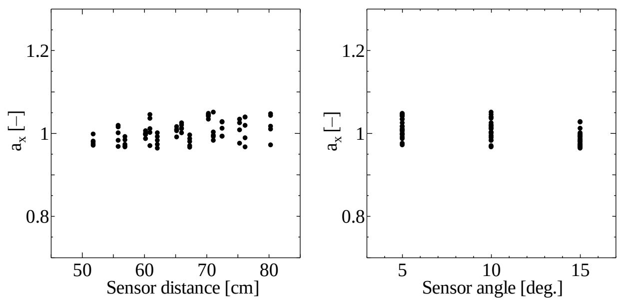


Figure 3. Tendency in a_x

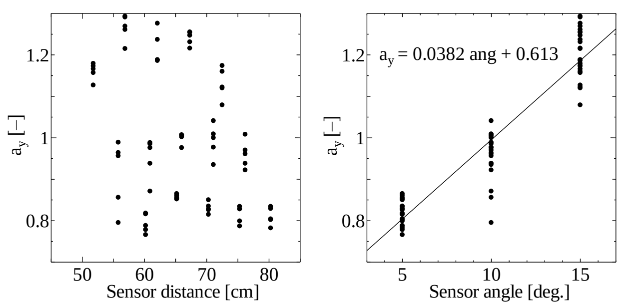


Figure 4. Tendency in a_y

3.4 Single point calibration

As a simple calibration method for the EyeTribe sensor, we propose to use a set of calibration data previously made by a model person (model calibration data). As a post process, users are only required to perform a simple calibration process in which they gaze only one point displayed on the center of a monitor.

As mentioned in the section 3.3, the acquired gaze data includes errors in the modeling coefficients a_x, b_x, a_y, and b_y when we use the model calibration data for different sensor / monitor settings.

The modeling coefficient a_y are calculable with Eq.(15) if we know the elevation angle. We can get this value from the single point gazing on the center point of the real monitor, and with Eq. (3), (4), and (5).

When a user gazes the center point of the monitor, the original gaze data acquired with the EyeTribe (X_o, Y_o) can be modified with Eq. (13) and (14) as follows;

$$X_{c0} = a_x X_0 + b_x, \tag{16}$$

$$Y_{c0} = a_y Y_0 + b_y. \tag{17}$$

In addition, they can be converted to the values on the real monitor with Eq. (9) and (10) as follows;

$$X_{d0} = \left(X_{c0} - \frac{\Delta_w}{W_c} W_{pc}\right) \cdot \frac{W_{pd} W_c}{W_{pc} W_d'}, \tag{18}$$

$$Y_{d0} = \left(Y_{c0} - \frac{\Delta_H}{H_c} H_{pc}\right) \cdot \frac{H_{pd} H_c}{H_{pc} H_d'}. \tag{19}$$

Here, when the target point is the center of the monitor, (X_{d0}, Y_{d0}) = (W_{pd} / 2, H_{pd} / 2), then,

$$b_x = \frac{W_{pc} W_d'}{2W_c} + \frac{\Delta_w}{W_c} W_{pc} - a_x X_0. \tag{20}$$

$$b_y = \frac{H_{pc} H_d'}{2H_c} + \frac{\Delta_H}{H_c} H_{pc} - a_y Y_0. \tag{21}$$

4. EXPERIMENT

4.1 Method

To evaluate the accuracy of the simple calibration method, we conducted a set of the gaze tracking for different sensor distance 50, 60 and 70 cm for 9 male adult subjects (22–26 years old, 6.3-6.8 cm in δ_r). The same model calibration data used in the section 3.3 was used. An LED monitor (1920×1080 pixel in resolution, 52.7×29.7 cm in real size) is put in front of the subjects who lie on a bed. The EyeTribe sensor was attached on the vertical centerline and lower edge of the monitor. The elevation angle of the sensor was adjusted to look straightly at the subject eyes for each distance.

The subjects watched the nine-point calibration view (Fig.6). They are instructed to look at the target mark (smile mark) which moves from 0-10 in the figure. The mark stops for 2 seconds at each point and moves to the next point in 2 seconds. The data when the target mark was at the center point were used for the simple calibration method mentioned in the section 3.4. The data when the target mark stopped at any points were used for evaluations.

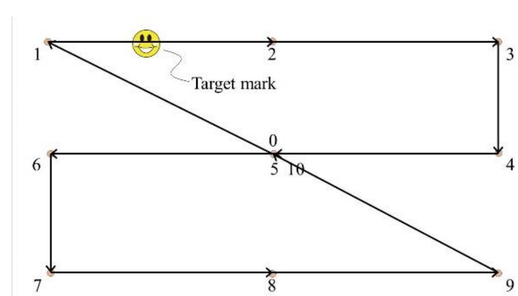


Figure 6. Target mark and its motion

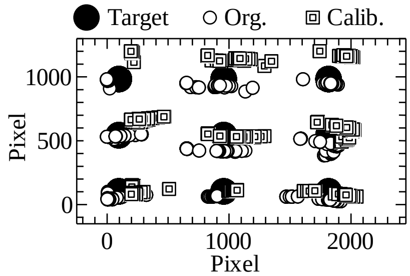


Figure 7. Result of L50 cm for subject 1

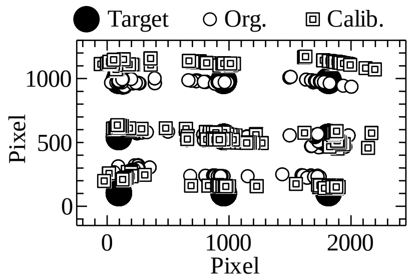


Figure 8. Result of L60 cm for subject 1

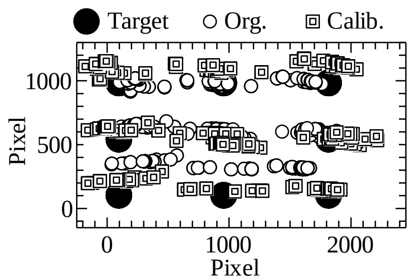
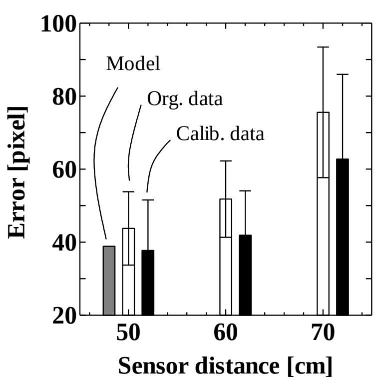


Figure 9. Result of L70 cm for subject 1



3

Figure 10. Errors between target positions and gaze points

4.2 Results and discussions

Figure 7-9 show the experimental results for the subject 1 for each monitor distance. In these figures, the vertical and horizontal axes denote the coordinates on the monitor. The black circles show the positions of the target mark. The white circles and squares denote the calibrated gaze data only using the model calibration data (original data) and the data using the proposed method (calibrated data), respectively.

Figures 10 shows the error between the target positions and the acquired gaze points. The gray bar shows an average error for a model person for the same condition of the original data. The white and black bars show the average error of the original and calibrated data for all subjects. The error bars show the standard deviations.

The average error for the model person was 39 pixel. This value represents the eye motion in fixed gaze (standard error). The error of the original data include 13 % 33 %, and 95 % for 50, 60, and 70 cm against the standard error, respectively. On the other hand, the error of the calibrated data include -3 %, 8 %, and 62 % for 50, 60, and 70 cm, respectively. The proposed calibration method shows enough accuracy for 50-60 cm distances. However, the error in 70 cm distance is not enough decreased. For more long distances than 70 cm, we should improve the method in the future.

5. CONCLUSION

In this article, we suggested a simple calibration method for the EyeTribe sensor in order to reduce the efforts for infants and mentally disabled users. In the proposed method, we used a model calibration data which previously made by a model person. As a post process, users are only required to perform a simple calibration process in which they gaze only one point displayed on the center of a monitor. The proposed method successfully reduced the calibration error for 50-60 cm of the sensor distances.

6. ACKNOWLEDGMENTS

This work was supported in part by Grant-in-Aid for Scientific Research of Japan (JP16K01555), and JGC-S Scholarship Foundation.

7. REFERENCES

[1] Horimoto C. H. and Mimica M. R. M. 2004. Eye gaze tracking techniques for interactive applications. *Computer Vision and Image Understanding*. 98, 4-24.

[2] Corbetta D., Guan Y. and Williams J. L. 2012. Infant eye-tracking in the context of goal-directed actions. *Infancy*. 17 (1), 102-125.

[3] Boraston Z. and Blakemore, S. J. 2007. The Application of eye-tracking technology in the study of autism. Journal of Physiology, 581, 893-898.

[4] Nakano T., Tanaka K., Endo Y., Yamane Y., Yamamoto T., Nakano Y., Ohta H., Kato N. and Kitazawa S. 2010. A typical gaze patterns in children and adults with autism spectrum disorders dissociated from developmental changes in gaze behaviour. Proceedings of The Royal Society Biological Science, 277, 2935-2943.

[5] Santis A. and Iacoviello D. 2009. Robust real time eye tracking for computer interface for disabled people. *Computer Methods and Programs in Biomedicine*. 96, 1-11. Doi: 10.1016/j.cmpb.2009.03.010

[6] EyeTribe: *https://s3.eu-central-1.amazonaws.com/theeyetribe.com/theeyetribe.com/dev/general/index.html*

[7] Ooms K., Dupout L., Lapon L. and Popelka S. 2015. Accuracy and precision of fixation locations recorded with the low-cost EyeTribe Tracker in different experimental setups. *Journal of Eye Movement Research*. 8 (1):5, 1-24.

[8] Dalmaijer E. 2014. Is the low-cost EyeTribe eye tracker any good for research? *PeerJ PrePrints*.2:e585v1 https://doi.org/10.7287/peerj.preprints.585v1