Proposal for Determining the Calibration Fixation Points for Achieving High-Precision Gaze Estimation with a Pupil Imaging Camera Installed at the Side of an Eye

Atsumu Naito and Kiyoshi Hoshino
University of Tsukuba, Tsukuba, Ibaraki, Japan
Email: naito.atsumu.ss@alumni.tsukuba.ac.jp, hoshino@esys.tsukuba.ac.jp

Abstract—**When an eye is captured from the side, the position of the pupil on the surface of the eyeball in spherical form is recorded in a two-dimensional camera image. As a result, the visual changes corresponding to eyeball movements form a nonlinear shape. Consequently, the calibration method where 9 points in a grid form are gazed at in advance in the same way as for the conventional gaze estimation inhibits correct mapping from the two-dimensional imaging to the three-dimensional eyeball movements. In this study, the authors examined the method of gaze estimation. In this method, an examinee is shown 7 points as calibration fixation points and correction is performed by using the projection transformation matrix that is produced from the 6 points. The result confirmed that satisfactory estimation can be obtained by instructing an examinee to gaze at 7 calibration fixation points that are determined from the total of evaluation points in the preliminary test and also gaze at 6 points for obtaining a projective transformation matrix even if the pupil imaging camera is installed at the side of the eye.**

Index Terms—**calibration fixation points, the side of an eye, gaze estimation, pupil, camera, eyeball movements**

I. Introduction

In general, in the gaze estimation using a camera, the camera needs to be installed in the front to take a photograph of a pupil [1], [2]. However, in such a case, the camera blocks the field of vision of the user. In some cases, this leads to some optical restrictions; for instance, when gaze estimation is to be conducted by installing a small camera in a narrow space such as inside of a head guard of a face protector of Kendo, a camera can only be installed at the side of the face protector [3]. When an eye is captured from the side, the position of the pupil on the surface of the eyeball, which is in a spherical form, is captured in a two-dimensional camera image [4], [5]. As a result, the visual changes corresponding to eyeball movements present a nonlinear form. Consequently, the calibration method where 9 points in a grid form are gazed at in advance in the same way as for the conventional gaze estimation inhibits correct mapping from the two-dimensional imaging to the three-dimensional eyeball movements [6]-[9].

This study proposes a method of presenting a calibration fixation point to achieve high-precision gaze estimation in conducting gaze estimation with a pupil imaging camera that is installed at the side of an eye.

II. System Conuration

A. Hardware Configuration

Fig. 1 shows the appearance of the ocular counter-roll instrument used for estimating a sight direction and Fig. 2 shows the schematic diagram. As for the instruments, a spectacles-shaped monochrome camera (UI-3271LE-M-GL of IDS) and near infrared radiation LED (OSI5LA5A33A-B of OptoSupply) are installed by using a flexible wire at the outer canthus of the eye to allow for shifting according to the individual. A pupil imaging camera is installed on the side of the eye, at the 45 degrees of the outer canthus of the eye in the horizontal direction and 45 degrees in the vertical downward direction and images are taken at the resolution of 808 x 608 [pix.].

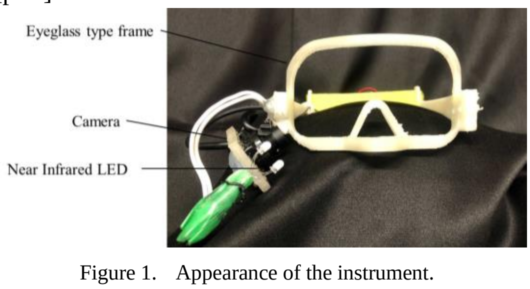

Figure 1. Appearance of the instrument.

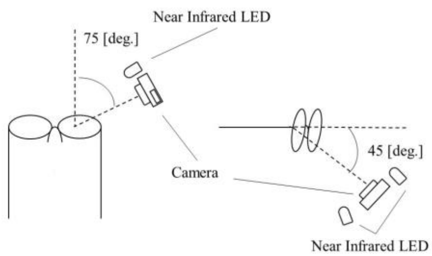

Figure 2. Structure of the instrument.

Manuscript received February 14, 2022; revised June 24, 2022.

doi: 10.12720/jait.13.5.462-469

B. Software Configuration

1) Calculation of pupil center coordinates

Pupil center coordinates necessary for estimation of eyeball rotational angles are obtained by processing the images taken from the eye. A dark pupil method is to detect pupil center coordinates causing the pupil to appear to be darker than the iris and sclera by irradiating with near infrared rays. Fig. 3(a) shows the image of the eyeball without a near infrared ray irradiated and Fig. 3(b) shows an image taken by irradiating with a near infrared ray. As shown in Fig. 3(a) and Fig. 3(b), a pupil can be captured by increasing the contrast of the pupil darker than the surrounding area. To separate the pupil from the surrounding area, binarization processing is performed and then contour extraction is performed for the binarized image. Elliptic approximation is performed for the group of points created through the contour extraction. The fitting of the elliptic approximation to the pupil is visually checked on each image and if ellipse approximation has targeted eyelashes or so on, the approximation is performed again by changing the parameters. The centroid of the approximated ellipse is determined as the pupil center coordinates.

(a) (b)

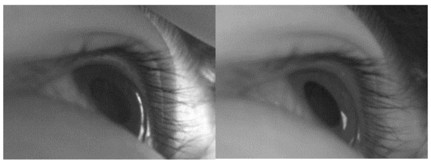

Figure 3. Examples of eyeball images produced in the dark pupil method (a) Without near infrared ray irradiated, (b) With near infrared ray irradiated.

2) Caliburation

To transform pupil center coordinates to an eyeball rotational angle based on the pupil center coordinates that are obtained, the following commonly used co-quadratic equations (1) and (2) are used. Equations (1) and (2) contain parameters to be calibrated. Parameters a_0~a_5 and b_0~b_5 are calibrated by applying the least squares method based on the indicator presentation position of the calibration fixation point and the pupil center coordinate data that is described in the previous section.

$$\theta_x = a_0x^2 + a_1y^2 + a_2xy + a_3x + a_4y + a_5 \quad (1)$$

$$\theta_y = b_0x^2 + b_1y^2 + b_2xy + b_3x + b_4y + b_5 \quad (2)$$

where,

θx: Eyeball rotation angle in the horizontal direction,

θy: Eyeball rotation angle in the vertical direction,

x: Pupil center coordinate in the horizontal direction in the image,

y: Pupil center coordinate in the vertical direction in the image,

$a_0 \sim a_5, b_0 \sim b_5$: Parameters to be calibrated.

3) Correction by using projective transformation

Projective transformation matrix M is obtained based on the difference between the estimation result of equations (1) and (2) performed on 6 points and the fixation point target value that is presented. The eyeball rotation angle that is estimated by equations (1) and (2) is corrected by using projective transformation.

$$\begin{pmatrix} \theta'_x \\ \theta'_y \\ 1 \end{pmatrix} = M \begin{pmatrix} \theta_x \\ \theta_y \\ 1 \end{pmatrix} \quad (3)$$

where,

M: Projective transformation matrix,

$\theta'x$: Eyeball rotation angle in the horizontal direction processed by projective transformation,

$\theta'y$: Eyeball rotation angle in the vertical direction processed by projective transformation,

θx: Estimated eyeball rotation angle in the horizontal direction,

θy: Estimated eyeball rotation angle in the vertical direction.

III. Preliminary Test

A. Overview

A pupil imaging camera is installed at a position 75 degrees in the horizontal direction of the outer canthus and 45 degrees in the vertical downward direction. Gaze estimation is performed for all the combinations for selecting 6 points from calibration fixation points and evaluation values according to the average errors of the estimation results are assigned to the calibration points that are used. An evaluation value at each calibration point is aggregated and the seven points of the highest total values are determined as the calibration fixation points. By using the same procedure, 6 projective transformation fixation points are determined for calibrating the projective transformation matrix.

B. Technique

For the measuring instrument, the sight line detector with a camera installed in a spectacles-type frame is used as shown in Section II. Fig. 4 shows the test environment. An examinee wears a spectacles-type ocular counter-roll instrument and the head movements are restricted by placing the head on the chin support. The chin support is installed at the position where the right eye of the examinee is at the distance of 1[m] from the calibration board. Fig. 5 shows the calibration fixation points and Fig. 6 shows the fixation points for evaluation as the indicators for the examinee to gaze at. Calibration fixation points are set at positions from 0[deg.] to 40 [deg.] in the horizontal direction and from -30[deg.] to 10[deg.] in the vertical downward direction, in the unit of 10[deg.] for the eyeball rotation angles on the right eye. Evaluation fixation points are set in the positions from 5[deg.] to +35[deg.] in the horizontal direction and from -25[deg.] to +15[deg.] in the vertical downward direction in the unit of 10[deg.]. Fig. 7 outlines the form of the fixation points. The fixation points present a cross shape of 10cm vertically and horizontally and an examinee is instructed to gaze at the grid section. An examinee should be an adult male without any eyeball movement defects. As the fixation point gazing method, an examinee is to gaze at the target fixation point, then gaze at the fixation point at the center (0[deg.] in the horizontal direction and 0[deg.] in the vertical direction), and move on to the next target fixation point. This procedure is repeated. Fixation points are to be followed in the order from the top left point to the point on the right side up to the rightmost point. Then, the examinee restarts gazing from the left end of the one row below. The examinee gazes at the indicator in his own time and moves on to the next fixation point after photographing the eyeball by pressing the switch. This procedure is repeated. The examinee removes the measuring instrument at each trial and leaves the seat.

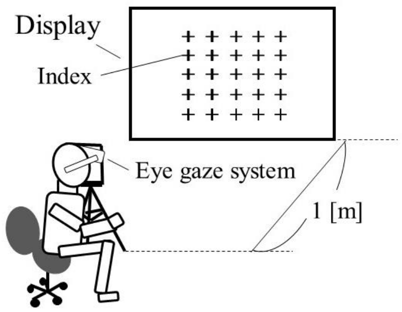

Figure 4. Appearance of the instrument.

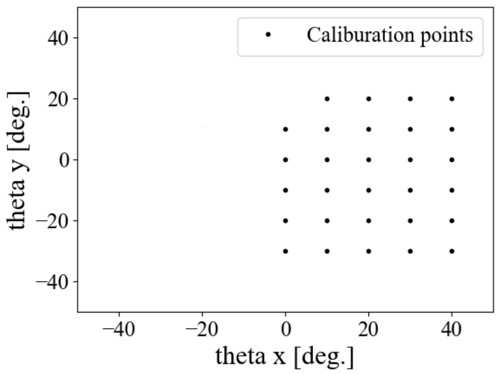

Figure 5. Example of calibration fixation points.

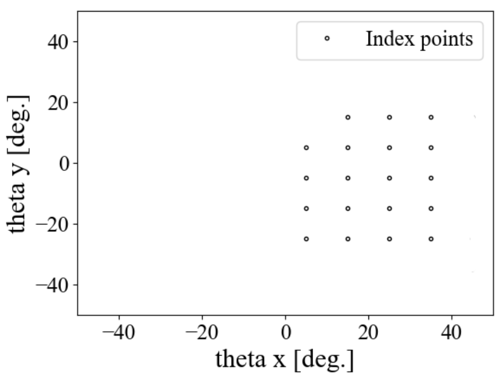

Figure 6. Example of fixation points for evaluation.

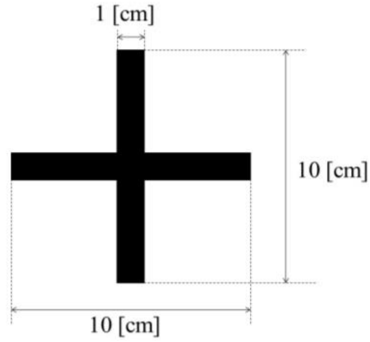

Figure 7. Appearance of the instrument.

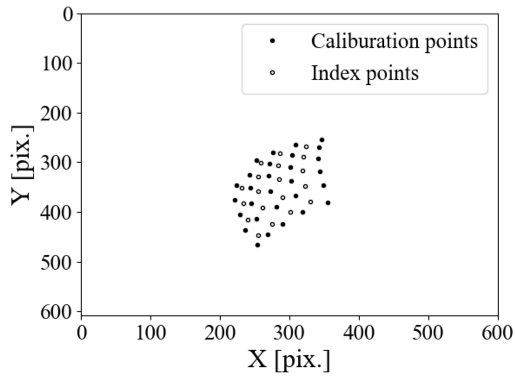

Figure 8. Example of distribution of pupil center coordinates.

C. Calibration Fixation Point Determination Method

1) Estimation results in all the combinations of six points

After gazing at calibration fixation points that are shown in Fig. 5, an examinee repeats the procedure ten times for gazing at evaluation fixation points as shown in Fig. 6. The pupil center positions are calculated for the eyeball images that are taken through image processing. Fig. 8 shows the pupil center coordinates that are calculated. Equations (1) and (2) that are shown in II.*B.(2* are used for transformation from the pupil center coordinates to the eyeball rotation angles. Parameters a_0~a_5 and b_0~b_5, which are used in equations (1) and (2) are calibrated by using the least squares method in each combination for selecting six points from 29 calibration fixation points and gaze estimation is performed by using the 19 evaluation fixation points. Equation (4) is used for calculating average errors.

$$E = \frac{1}{19}\sum_{i=0}^{19}\sqrt{((\bar{\theta}_{xi} - \hat{\theta}_{xi})^2 + (\bar{\theta}_{yi} - \hat{\theta}_{yi})^2)} \tag{4}$$

where,

E: Estimated average error,

$\bar{\theta}xi$: Eyeball rotation angle in the horizontal direction as the target of the gaze point,

$\bar{\theta}yi$: Eyeball rotation angle in the vertical direction as the target of the gaze point,

$\hat{\theta}xi$: Estimated eyeball rotation angle in the horizontal direction,
$\hat{\theta}yi$: Estimated eyeball rotation angle in the vertical direction.

Fig. 9 shows the histogram of the estimation result average errors and number of combinations of 6 points.

2) Total evaluation value calculated from average errors

Evaluation values of average errors that are calculated from all the combinations of 6 points are calculated by using equation (5) and the evaluation value of each point at each point used for calibration is aggregated by using equation (6).

$$V_j = a * \coth(E_j - b) - a \tag{5}$$

$$SUM\_V_k = \sum_{j=0}^{M} flag_{jk} * V_j \tag{6}$$

where,
V_j: Evaluation value of the j[th] combination of 6 points,
a, *b*: Any parameter,
E_j: Average error of the j[th] combination of 6 points,
SUM_V_k: Total evaluation value of the kth calibration points,
M: Number of combinations of 6 points,
$flag_{jk}$: k[th] calibration points are used in the j[th] combination.

Calibration fixation points are determined by using the 7 points of highest evaluation values. For parameters a and b, a is changed from 1 to 100 in the unit of 1 and b is changed from 0 up to the lowest value of E of 10 trials in the unit of 0.01. The values at which the total estimation result is lowest among the 10 trials are applied. The results are a=1 and b=0.66. Fig. 10 shows the changes of the total average errors and evaluation values of 10 trials and Fig. 11 shows the results of the aggregation of evaluation values of the calibration fixation points.

D. Fixation Point Determination Method for a Projective Transformation Matrix

1) Fixation point presentation method

Fixation points for determination of a projective transformation matrix is examined. The same test environment as that in III.*B* was used. Four trials were conducted. Seven calibration fixation points that were determined in III.*C.(2* as shown in Fig. 12 are presented and after an examinee gazed at these points, the fixation points are presented for creating a projective transformation matrix as shown in Fig. 13 and the evaluation fixation points are presented as shown in Fig. 14.

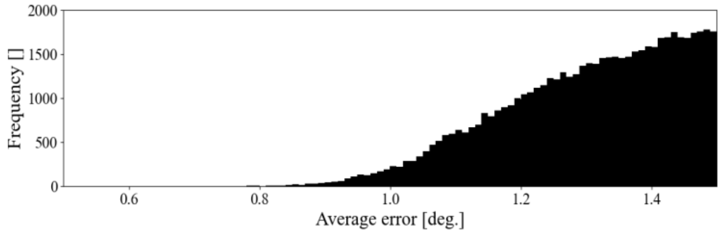

Figure 9. Histogram of estimation result average errors and the number of combinations of 6 points.

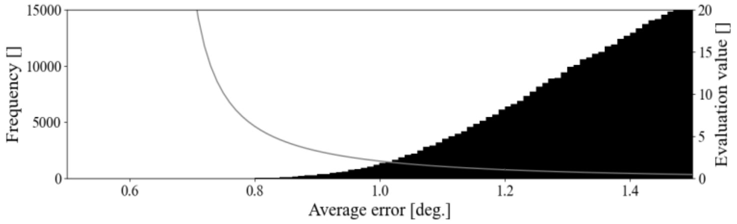

Figure 10. Changes of the total of average errors and evaluation values of 10 trial.

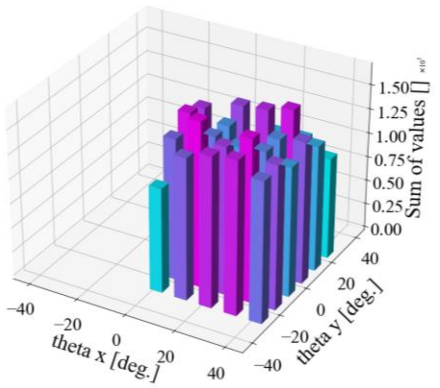

Figure 11. Total of evaluation values at each calibration point.

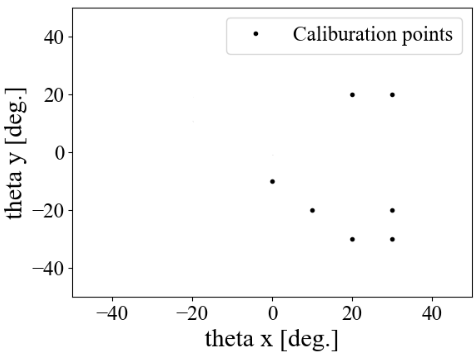

Figure 12. Example of calibration fixation points.

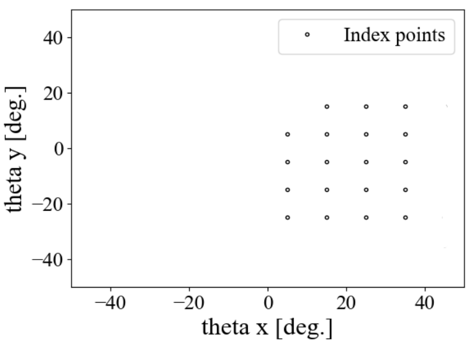

Figure 13. Example of projective transformation fixation points.

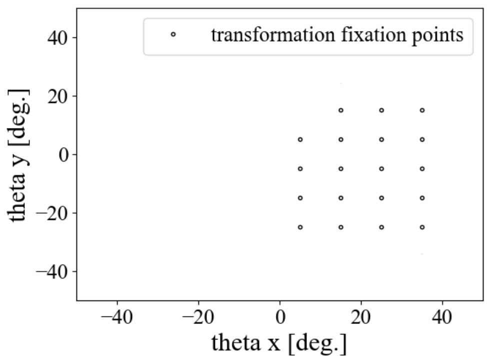

Figure 14. Example of evaluation fixation points.

2) Projective transformation matrix of four points

Parameters for equations (1) and (2) are established by using 7 calibration fixation points and a projective transformation matrix is determined by using all the combinations for selecting 6 points from the 19 points that are presented as the projective transformation fixation points matrix creation. Estimation is performed by using the evaluation fixation points and the average estimation error after projective transformation by using equation (7).

$$E' = \frac{1}{19}\sum_{i=0}^{19}\sqrt{((\bar{\theta}_{xi} - \widehat{\theta}'_{xi})^2 + (\bar{\theta}_{yi} - \widehat{\theta}'_{yi})^2)} \qquad (7)$$

where,

E': Average estimation error as a result of projective transformation,

$\bar{\theta}_{xi}$: Target eyeball rotation angle in the horizontal direction at the gazed point,

$\bar{\theta}_{yi}$: Target eyeball rotation angle in the vertical direction at the gazed point,

$\hat{\theta}'_{xi}$: Eyeball rotation angle in the horizontal direction after projective transformation,

$\hat{\theta}'_{yi}$: Eyeball rotation angle in the vertical direction after projective transformation.

Fig. 15 shows histogram of the estimate result average errors of one trial and the number of combinations for selecting 6 points for a projective transformation matrix.

3) Total of the evaluation values calculated from average errors after projective transformation

By using equation (8), evaluation values are calculated from the average errors that are calculated from all the combinations of 6 points. By using equation (9), an evaluation value is aggregated for each point that is used for determining the projective transformation matrix.

$$V'_j = a' * \coth(E'_j - b') - a' \qquad (8)$$

$$SUM\_V'_k = \sum_{j=0}^{M} flag_{jk} * V_j' \qquad (9)$$

where,

V'_j: Evaluation value of the j^th^ combination of 6 points for projective transformation,

a', b': Any parameter,

E'_j: Average error of the j^th^ combination of 6 points after projective transformation,

$SUM\_V'_k$: Total evaluation value of the k^th^ projective transformation fixation point,

M: Number of combinations of 6 points,

$flag'_{jk}$: The k^th^ projective transformation fixation point was used in the j^th^ combination.

Projective transformation fixation points are determined by using the 6 points of the highest total evaluation values. For parameters a and b, a' is changed from 1 to 100 in the unit of 1 and b' is changed from 0 to the lowest value of E' of 4 trials in the unit of 0.01. The values at which the total estimation result is the lowest among the 10 trials are determined. The results are a=1 and b=0.06. Fig. 16 shows the changes of the total average errors and evaluation values and Fig. 17 shows the results of the aggregation of evaluation values at the projection transformation fixation points.

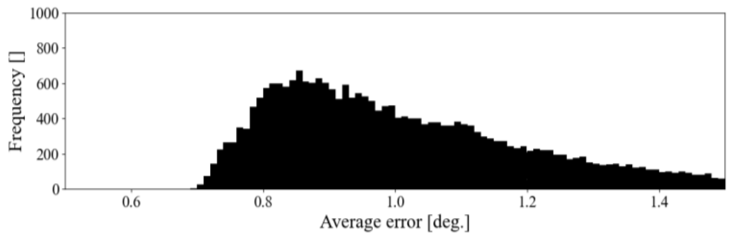

Figure 15. Example of a histogram of the average errors of estimation results and the number of combinations for selecting 6 points for a projective transformation matrix.

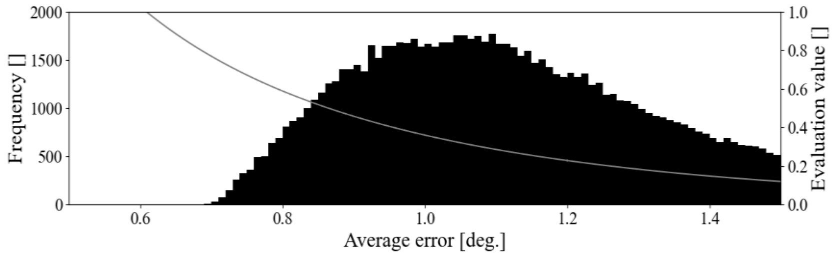

Figure 16. Example of average errors of 4 trials and changes of combination frequencies and evaluation values.

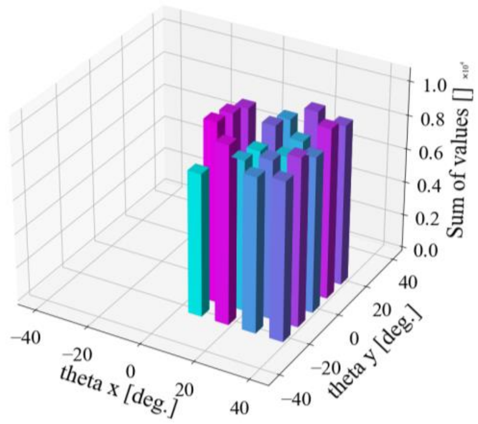

Figure 17. Example of a total evaluation value at each projective transformation fixation point.

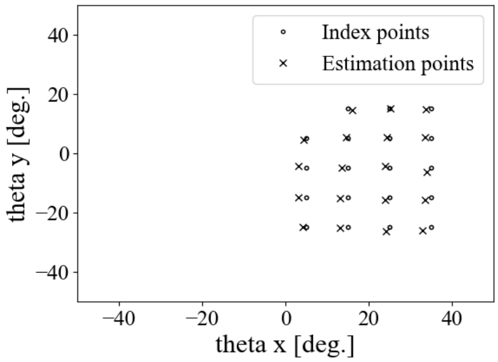

Figure 18. Example of estimation results without projective transformation performed.

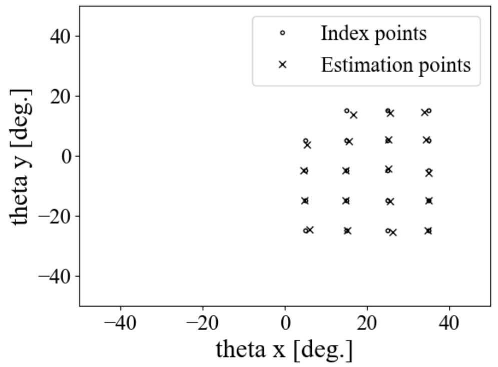

Figure 19. Example of estimation results of with projective transformation performed.

Fig. 18 shows the estimation results prior to projective transformation by using the highest 6 total evaluation values and Fig. 19 shows the estimation results after projective transformation.

IV. Evaluation Test

A. Test Technique

The same test environment as described in Section III is applied. The calibration fixation points are presented as shown in Fig. 20. The fixation points are presented for creating a projective transformation matrix as shown in Fig. 21. After gazing at 7 calibration fixation points that were determined in Section III, an examinee gazes at 6 projective transformation fixation points and then evaluation fixation points that are shown in Fig. 22. The examinee should be one adult male in the twenties without eyeball movement defects and who can perform 6 sets of tests.

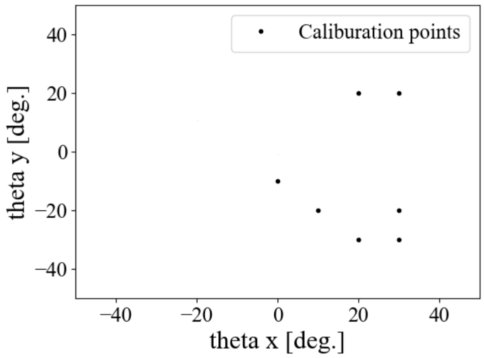

Figure 20. Example of calibration fixation points.

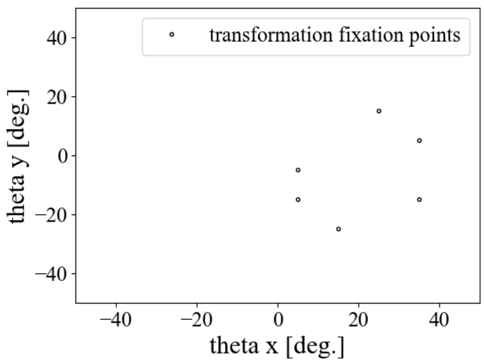

Figure 21. Example of projective transformation fixation points.

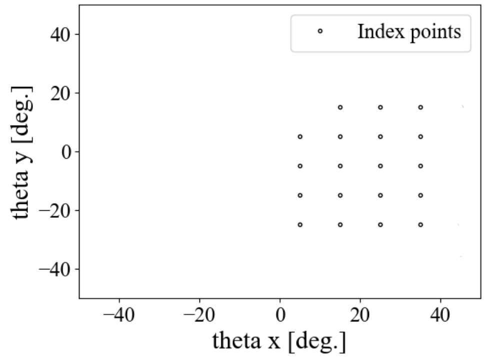

Figure 22. Example of evaluation fixation points.

B. Test Results

Fig. 23 shows the estimation results that were obtained through the tests. The lowest average error from the tests is 1.01[deg.] and highest average error is 1.33[deg.]. The average estimation error from 6 tests is 1.18[deg.].

V. Summary

When an eye is captured from the side, the position of the pupil on the surface of the eyeball in spherical form is recorded in a two-dimensional camera image. As a result, the visual changes corresponding to eyeball movements form a nonlinear shape. Consequently, the calibration method where 9 points in a grid form are gazed at in advance in the same way as for the conventional gaze estimation inhibits correct mapping from the two-dimensional imaging to the three-dimensional eyeball movements. In this study, the authors examined the method of gaze estimation. In this method, an examinee is shown 7 points as calibration fixation points and correction is performed by using the projection transformation matrix that is produced from the 6 points. The result confirmed that satisfactory estimation can be obtained by instructing an examinee to gaze at 7 calibration fixation points that are determined from the total of evaluation points in the preliminary test and also gaze at 6 points for obtaining a projective transformation matrix even if the pupil imaging camera is installed at the side of the eye.

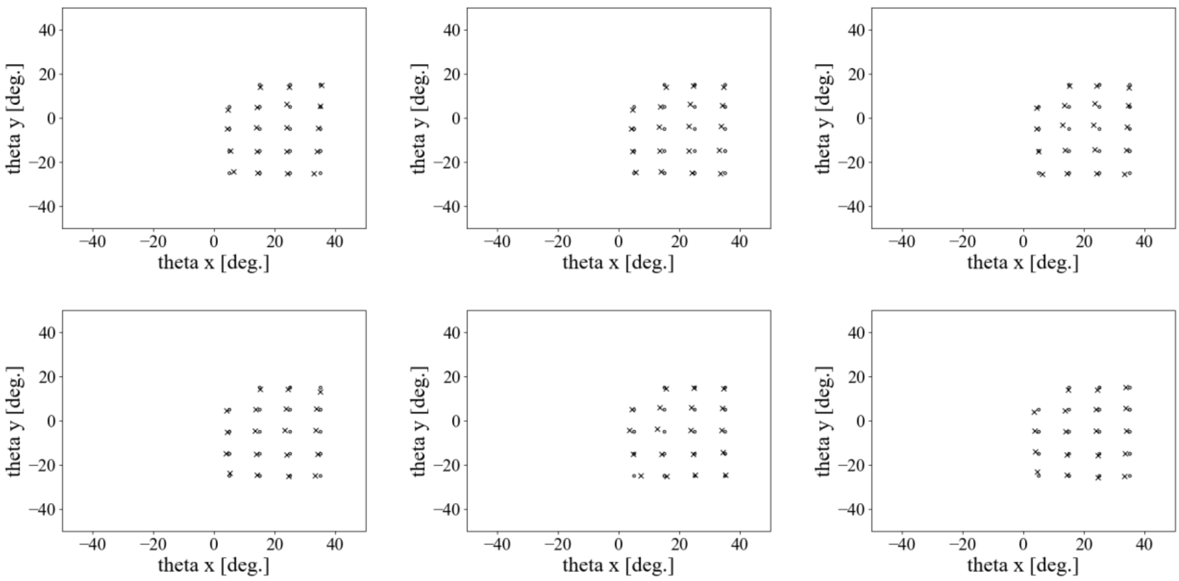

Figure 23. Example of evaluation fixation points.

Conflict of Interest

The authors declare no conflicts of interest associated with this manuscript.

Author Contributions

A. Naito executed the experiments, was involved in the data analysis, and wrote the manuscript. K. Hoshino is a supervisor, designed the experiments, and edited the manuscript. All authors critically revised the report, commented on drafts of the manuscript, and approved the final report.

Acknowledgment

A part of this work was supported by Sciences Research Grants of the Ministry of Education, Culture, Sports, Science and Technology (grant numbers: 16K12199 and 17K01071). A part of this work was also commissioned by the Ministry of Internal Affairs and Communications under SCOPE (reference number: 181603016). The authors are grateful for these supports.

References

[1] S. Kohlbecher, S. Bardinst, K. Bartl, E. Schneider, T. Poitschke, and M. Ablassmeier, "Calibration-Free eye tracking by reconstruction of the pupil ellipse in 3D space," in *Proc. Symposium on Eye Tracking Research & Applications*, 2008, pp. 135-138.

[2] X. Long, O. K. Tonguz, and A. Kiderman, "High speed eye tracking system with robust pupil center estimation algorithm," in *Proc. 29th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, 2007, pp. 22-26.

[3] S. Shimanoe and K. Hoshino, "A wearable imaging system for recording the game tactics in martial arts using swords," in *Proc.*

IEEE 1st Global Conference on Life Sciences and Technologies, 2019, pp. 139-140.

[4] K. Hoshino, K. Kawamoto, and Y. Nakai, "Calibration fixation points that enable accurate gaze estimation with pupil image capturing camera installed almost next to eye," in *Proc. Intl. Conf. on Image, Video and Signal Processing*, 2021, pp. 38-45.

[5] K. Hoshino, Y. Noguchi, and Y. Nakai, "Gaze estimation with easy calibration method," in *Proc. the 5th International Conference on Intelligent Information Technology*, 2020, pp. 102-106.

[6] K. Hoshino, Y. Noguchi, and N. Ono, "Eye tracking and measurement of eye rotation using a small camera installed roughly next to the eye," in *Proc. 13th International Conference on Human System Interaction*, 2020, p. 6.

[7] K. Hoshino, Y. Nakai, Y. Noguchi, and N. Ono, "Estimation of rotational eye movement based on intensity gradients in the eye images," in *Proc. the 5th International Conference on Intelligent Information Technology*, 2020, pp. 121-126.

[8] K. Hoshino, S. Shimanoe, Y. Nakai, Y. Noguchi, and M. Nakamura, "Estimation of the line of sight from eye images with eyelashes," in *Proc. the 5th International Conference on Intelligent Information Technology*, 2020, pp. 116-120.

[9] K. Hoshino and N. Ono, "A compact wearable eye movement measurement system for support of safe driving," in *Proc. IEEE International Conference on Mechatronics*, 2019, pp. 225-231.

Kiyoshi Hoshino received two doctor's degrees; one in Medical Science in 1993, and the other in Engineering in 1996, from the University of Tokyo respectively. He is now a full professor at the Biological Cybernetics Lab of University of Tsukuba.
Professor Hoshino served as a member of the "cultivation of human resources in the information science field" WG, Special Coordination Funds for the Promotion of Science and Technology, MEXT, a member of "Committee for Comport 3D Fundamental Technology Promotion", JEITA, a councilor and director of the Ibaraki Sports Association, and the General Conference Chair of the 43rd Annual Meeting of Japanese Society of Biofeedback Research, ICBBE 2018 to 2020, and DMIP 2018 to 2021. He received Laval Virtual Awards in 2009, 2013, and 2014.