

Improvement of Accuracy in Remote Gaze Detection for User Wearing Eyeglasses Using Relative Position Between Centers of Pupil and Corneal Sphere

Kiyotaka Fukumoto^(✉), Takumi Tsuzuki, and Yoshinobu Ebisawa

Graduate School of Engineering, Shizuoka University, Shizuoka 432-8561,
Japan

{fukumoto.kiyotaka, ebisawa.yoshinobu}@shizuoka.ac.jp

Abstract. One of the general problems of the pupil-corneal reflection-based gaze detection systems is that the frames and lens of eyeglasses produce reflection images of the light sources in the camera image when a user wears eyeglasses. The glass reflections tend to be misdetected as the pupil and corneal reflections. In the present paper, we propose a novel geometrical methodology based on the optical structure of the eyeball to detect a true pair of the pupil and corneal reflection. The experimental results show that the proposed method improved the precision of gaze detection when the subjects wore glasses or when disturbance light sources existed.

Keywords: Pupil · Corneal reflection · Corneal sphere center · Gaze detection

1 Introduction

Video-based gaze detection systems are about to be used in various fields such as the entertainment [1], medicine [2], and safety driving support [3]. In our previous study, we have developed a pupil-corneal reflection-based robust and precise gaze detection system using the two light sources and the image difference method, which allows large head movements and easy user calibration [4, 5]. In this system, an optical system for detecting the pupils and corneal reflections consists of a camera and two concentric near-infrared LED rings (inner and outer rings) light source attached to the camera. The inner and outer LED rings generate bright and dark pupil images, respectively. The pupils are detected from a difference image created by subtracting the bright and dark pupil images. In the difference image, a threshold for binarization to detect the pupils is easily determined automatically because the pupils are embossed from the relatively flat background image. However, when the users move their head, the pupil position differs between the bright and dark pupil images because of the acquisition time difference of both pupil images. As a result, the pupil position is not detected accurately. Therefore, in our system, the image difference processing is performed after shifting the small areas (small windows) including each pupil in the dark pupil image so that the corneal reflection in this dark pupil image may coincide with that in the bright pupil image. We call this method *the image difference method with positional*

compensation based on the corneal reflection (the positionally compensated image difference (PCID) method) [6]. In addition, we proposed the easy gaze calibration methods: the automatic, one-point, and two-point calibration methods [4]. In the one-point calibration method, the user has only to fixate on one target having known coordinates presented at the center of the PC screen. By this procedure, the gaze points on the whole of the PC screen can be detected almost exactly.

However, when the user wears eyeglasses, their frames and lens produce various size, shape and intense of areas in the camera image as so-called glass reflections. The reflections often show image features similar to the pupil and the corneal reflection and tend to be misdetected as the pupil or the corneal reflection. The reflections of tears and disturbance light sources also cause the misdetection. In the present paper, we propose a novel geometrical methodology based on the optical structure of the eyeball to detect a true pair of the pupil and corneal reflection for accurate gaze detection even if the user wears glasses.

2 Our Gaze Detection System

2.1 System Configuration

Figure 1(a) shows an overview of the gaze detection system which we developed. This system has two optical systems (Fig. 1(b)), each of which consists of a digital video camera having near-infrared sensitivity, a 16-mm lens, an infrared filter (IR80), and a light source. Each of the two optical systems was placed under a 19-in. liquid crystal display (screen size: 376.3×301.1 mm, $1,280 \times 1,024$ pixels). The light source consisting of near-infrared 3ϕ LEDs which are arranged in a double concentric circle ring form is attached to the camera. The wavelengths of the inner and outer rings were 850 and 940 nm, respectively. The pupil becomes brighter in the 850 nm ring than the 940 nm ring because the transmissivity of the eyeball medium is different. The distance between the LEDs and the aperture of the camera also varies the pupil brightness. The combined effects of the differences of the distance and transmissivity were applied to the light source. In order to reduce the effect of an ambient light, it is desirable that the LEDs irradiation power on the user's face becomes as strong as possible compared with the ambient light. Therefore, the LEDs were flashed while the camera shutter opened (shutter speed $500 \mu\text{s}$). The current flow was approximately one ampere during LED flashing. The two cameras were driven with a slight synchronization difference ($670 \mu\text{s}$) for avoiding mutual light interference of the optical systems. By this, basically, only one corneal reflection appears for each eye in an image.

An 8-bit gray scale image (640×480 pixels) of the user's face was input into a personal computer (PC, Intel Core i7 3.20 GHz CPU and 12 GB RAM) at 60 fps.

2.2 Detection of Centers of Pupils and Corneal

Our Conventional Method for Detection of Pupils and Corneal Reflections. First, the pupils are searched for and detected from the difference image generated from the

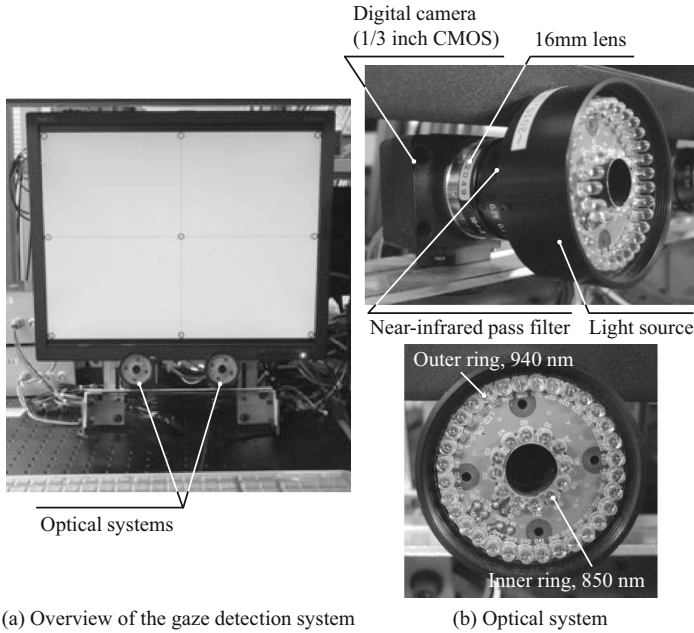


Fig. 1. (a) Our gaze detection system. (b) Optical system for detecting pupil and corneal reflection.

bright and dark pupil images. The image is processed in the following order: binarization, removal of isolated pixels, noise reduction using mathematical morphology operations, and labeling. The largest and second largest labeled regions are detected as the two pupils.

When the pupil is undetected in the prior difference image (e.g., when the pupil is covered with the glass reflection), the pupils are searched for in the whole of the current difference image again. When two consecutive pupil detections occur, in order to perform *the PCID method*, the pupil positions in the current images are predicted using the linear Kalman filter, and the small window (70×70 pixels) is then applied around the predicted pupil positions, respectively. The image within the small window is transformed into the double-resolution (DR) image (140×140 pixels). An intense and tiny region closest to a center of the DR images is extracted and then the center of gravity considering the values of the pixels in the region is determined as the center of the corneal reflection. As described before, when the user's head is moving, the pupils cannot be obtained correctly because the pupil position differs between the bright and dark pupil images. Therefore, the DR difference image is generated after shifting the DR dark pupil image so that the corneal reflection in the DR dark pupil image coincides with that of the DR bright pupil image (*the PCID method*). When the corneal reflection is not detected in either one of the bright and dark pupil image, the difference image is generated without *the positional compensation*. After the image areas whose image feature similar to the pupil are labeled in the binarized DR difference image, the nearest

area to the predicted pupil position is determined as the pupil in the current DR difference image. In this image, the ellipse-fitting of the contour of the pupil is performed. The center of the ellipse is determined as the pupil center.

Proposal Geometrical Method for Determining a True Pair of the Pupil and the Corneal Reflection. When the user wears glasses, the glass reflections are tend to be misdeteected as the pupils and corneal reflections. In addition, the false images of disturbance light sources may be misdeteected as the true corneal reflection of the light sources of the system. Therefore, we propose the geometrical method for detecting a true pair of the pupil and corneal reflection. Assuming the corneal surface to be a sphere, the corneal sphere center is determined as shown in Fig. 2(a). We use the pinhole camera model and assume that the light source is located at the same position as the pinhole. Therefore, the corneal sphere center exists on the line connecting the pinhole and the corneal reflection in the image sensor. The 3D position of the corneal sphere center can be determined by stereo-matching the corneal reflections obtained from the two cameras. However, when one or both corneal reflections (Fig. 2(b) and (c)) obtained from the two cameras are misdeteected due to the glass reflection or the disturbance light source, the corneal sphere center is deteected at wrong positions.

In the proposed method, m and n corneal reflection candidates for each camera and each eye are extracted from the bright and dark pupil images, respectively, which include the true and false corneal reflections. The *PCID method* is performed for all $m \times n$ combinations of the candidates. When the pupil is not deteected, it is judged that at least one of the paired two candidates used for performing the method was not the true corneal reflection. When the pupil is deteected, the pupil and corneal reflection pair used in the method is retained as one of the pair candidates. The 3D positions of the pupil and corneal sphere centers are deteected by stereo-matching the pupils and the corneal reflections, respectively, of the remaining pair candidates obtained from two cameras. One true pair of the pupil and corneal sphere centers is chosen by the following two conditions:

Condition I: the angle between the vector from the corneal sphere center to the pupil center and the vector from the pupil center to the middle point between the two cameras is within 40° . This is because we thought that the gaze deteetection system suffices gaze deteetection of only PC screen area. So, considering the unknown angle difference between the visual and optical axes, we gave 40° so as to include the screen area.

Condition II: the distance d between the corneal sphere center and the pupil center satisfies $D_{C-P} - 1.5 \text{ [mm]} < d < D_{C-P} + 1.5 \text{ [mm]}$, where D_{C-P} is the distance obtained beforehand from the individual users.

Based on the chosen pair, the 3D pupil position and the coordinates of the pupil and corneal reflection in the camera image are obtained and are used for the gaze deteetection [5].

2.3 Gaze Deteetection Theory and Calibration Method [5]

In Fig. 3, O'_1 and O'_2 indicate the pinholes of the two camera-calibrated cameras. The 3D pupil position P is obtained by stereo-matching. As mentioned before, we assume

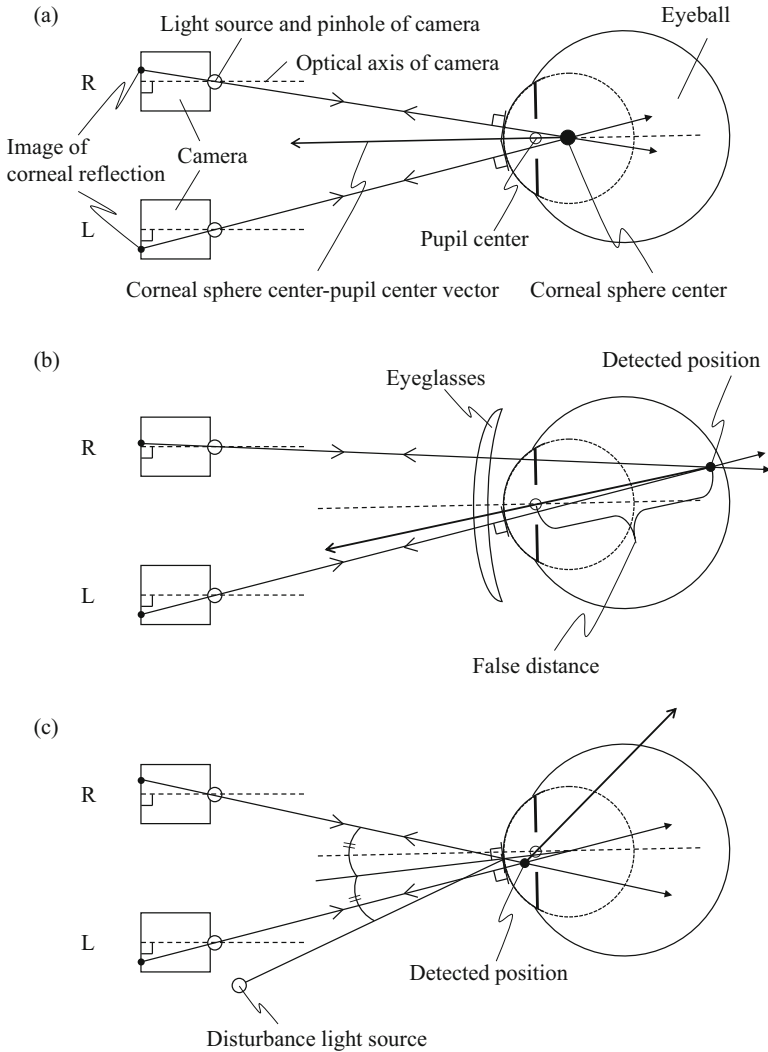


Fig. 2. (a) When two cameras detect a true corneal reflection of the light source attached to the corresponding camera, respectively. (b) When the left and right cameras detect a true corneal reflection and a false reflection of a glass reflection, respectively. (c) When the left camera detects a true corneal reflections produced by the light source attached to the left camera and the right camera detects a false reflection of the disturbance light source, respectively.

that the light source attached to each camera is located at the same position as the corresponding camera. The line of sight (visual axis of the eyeball) passes through the fovea on the retina, the pupil center P and gaze point Q on the screen plane of the PC display. Now we define the virtual gaze planes H_1 and H_2 of the cameras for one eyeball. These planes are perpendicular to the line passing through P and O'_1 and the

line passing through P and O_2 , respectively, and they include O'_1 and O'_2 , respectively. The X-axis (X_1 and X_2) of planes H_1 and H_2 is the intersection between the corresponding plane and the horizontal plane in the global coordinate system ($x - y - z$). H_1 and H_2 rotate according to the displacements of the pupil position.

Next, we define the virtual gaze sphere S whose center is P and whose radius is arbitrary. The visual axis PQ has intersection points with sphere S and planes H_1 and H_2 . The intersection points are denoted as G , G'_1 and G'_2 , respectively. Here, we define angular vectors θ_1 and θ_2 on sphere S as the projections of ordinary vectors $\overrightarrow{O'_1G'_1}$ and $\overrightarrow{O'_2G'_2}$ on planes H_1 and H_2 to sphere S . By projecting the horizontal axes X_1 and X_2 on planes H_1 and H_2 to sphere S , orientations ϕ_1 and ϕ_2 of vectors $\overrightarrow{O'_1G'_1}$ and $\overrightarrow{O'_2G'_2}$ can be also projected to sphere S and can be defined. According to these projections, you can see that angular vectors θ_1 and θ_2 can be determined by using ϕ_1 and $\angle O_1PG$ and by using ϕ_2 and $\angle O_2PG$, respectively. Here, angular vector $\overrightarrow{O_1O_2}$ ($|\overrightarrow{O_1O_2}| = \angle O_1PO_2$) is expressed as follows:

$$\overrightarrow{O_1O_2} = \theta_1 - \theta_2 \quad (1)$$

We assume a linear relationship between the actual size vector \mathbf{r} from the corneal reflection to the pupil center and the angle θ between the visual axis of the eyeball and the line connecting the pupil and the camera as follows:

$$\theta = k\mathbf{r} \quad (2)$$

where \mathbf{r} is converted from the vector from the corneal reflection center to the pupil center, which is obtained from the camera image, using the pinhole model. k is a constant. Actually, in general, there is a difference between the optical and visual axes of the eyeball. So, \mathbf{r} is calculated by compensating a measured vector \mathbf{r}' using an offset vector \mathbf{r}_0 as the following equation:

$$\mathbf{r} = \mathbf{r}' - \mathbf{r}_0 \quad (3)$$

From Eqs. (2) and (3), the following equations are given for cameras 1 and 2.

$$\theta_1 = k\mathbf{r}_1 = k(\mathbf{r}'_1 - \mathbf{r}_0) \quad (4)$$

$$\theta_2 = k\mathbf{r}_2 = k(\mathbf{r}'_2 - \mathbf{r}_0) \quad (5)$$

From the above equations, k is calculated by the following equation:

$$k = \frac{|\theta_1 - \theta_2|}{|\mathbf{r}'_1 - \mathbf{r}'_2|} = \frac{\angle O_1PO_2}{|\mathbf{r}'_1 - \mathbf{r}'_2|} \quad (6)$$

Using the value of k , \mathbf{r}_0 is determined from Eqs. (4) and (5). Determining k and \mathbf{r}_0 mean the user calibration.

In the gaze detection procedure, first, the pupil-corneal reflection vectors r'_1 and r'_2 are obtained from the images of the two cameras. By using Eqs. (4) and (5), θ_1 and θ_2 are calculated. Next, the visual axis is determined for each eye from pupil position P , θ_1 and θ_2 . Finally, the gaze point on the screen is estimated as the intersection point between the screen plane and the visual axis.

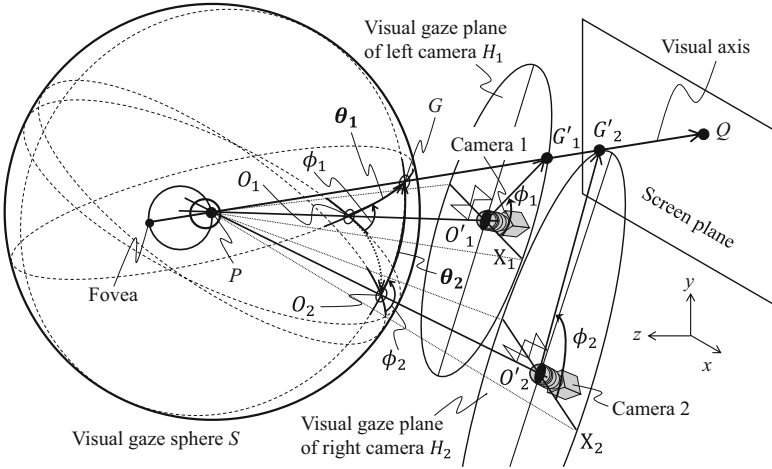


Fig. 3. Gaze detection theory using visual gaze sphere

3 Experiments

3.1 Experiment 1: Measurement of Distance Between Corneal Sphere Center and Pupil Center

Method. In order to examine and determine the distance D_{C-P} shown in Condition II, the 3D corneal sphere and pupil center positions of three university students who did not wear glasses were measured. In the calibration procedure, the subjects were asked to fixate on a calibration target presented at the center of the PC screen (the one-point calibration method). The distance between the eyes and the screen was approximately 80 cm. After the calibration procedure, the subjects fixated on a stationary target presented at the center of the screen and a slowly moving target between the right and left edges of the screen. Using a chinrest stand, the subjects' heads were positioned at the following five positions: approximately 75, 80, and 85 cm from the PC screen, and 5 cm to the left and 5 cm to the right at 80 cm. In addition, subject A wore glasses and participated in the same experiment again.

Results. Figure 4(a) and (b) show the averages and SDs of the distance d between the pupil and corneal sphere centers at the five head positions when the subjects fixated on the stationary target and the moving target, respectively. Although the distance d was different among the subjects, it did not depend on the head positions and the gaze

directions for each subject. Figure 5 shows the results when subject A wore and did not wear glasses, respectively. Almost the same values were obtained by whether the subject wore glasses or not.

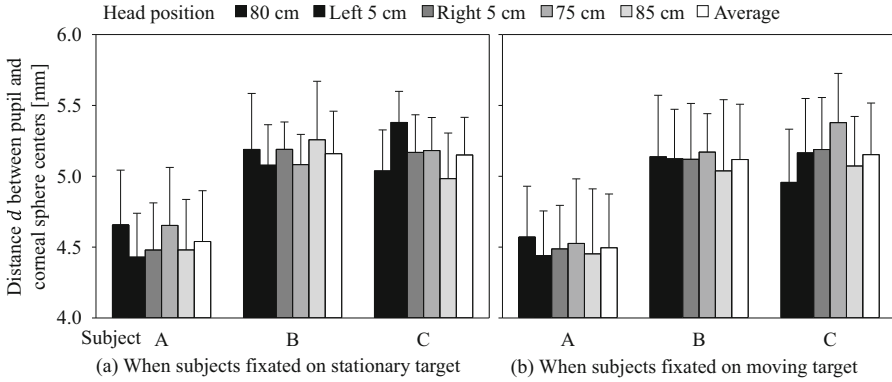


Fig. 4. Averages and SDs of the distance d between the pupil and corneal sphere centers at the five head positions for each subject.

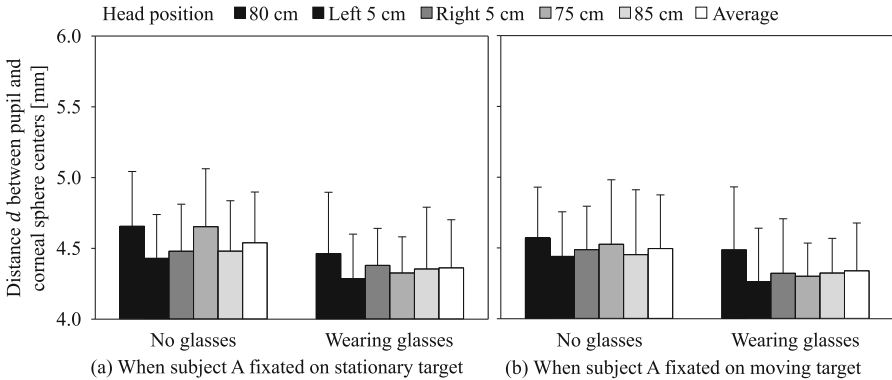


Fig. 5. Averages and SDs of the distance d between the pupil and corneal sphere centers at the five head positions when subject A wore glasses and when he did not wear glasses.

3.2 Experiment 2: Gaze Detection When Subjects Wear Glasses

Method. This experiment was conducted in order to compare the precision of gaze detection between the proposed and our previous methods when subjects wore glasses. In the previous method, the corneal reflection nearest to the predicted pupil center was chosen for the *PCID method*. The values of m and n were both three. The subjects were three university students. In the one-point calibration procedure, the head direction of the subjects was adjusted so that the lens reflections did not appear in the camera

image. After the procedure, the subjects wearing glasses fixated on 25 (5 by 5) visual targets equally arranged on the PC screen one by one. The values of D_{C-P} were obtained from each subject when they did not wear glasses before this experiment.

Results and Discussion. Figure 6(a) and (b) show the gaze point distributions of the left eye for the subject A in the proposed and our previous methods, respectively. In the previous method, the dispersion of the gaze points was large compared to the proposed method, especially when the subject fixated on the lower targets. This was caused by misdetection of the pupil and/or the corneal reflection due to the glass reflections. Especially, the subject fixated on the targets 17 and 22, the glass reflection had covered the left pupil. In the proposed method, no gaze point was detected. These results mean that the pupil and/or the corneal reflection were misdetections in the previous method, whereas the proposed method prevented these misdetections. Furthermore, 1.0 % gaze points outside of the region presented in Fig. 6 existed in the previous method while 0 % in the proposed method. The average and SD of the gaze error in visual angle for the subject A were 1.24 ± 1.61 [deg] in the previous method, whereas those of the proposed method were 1.08 ± 1.23 [deg]. The results of the other two subjects showed the similar results. The average gaze error for all the three subjects was 1.26 ± 1.62 [deg] in the previous method and 1.14 ± 1.99 [deg] in the proposed method. These results indicate that the proposed method functioned to prevent the misdetection of the pupil and corneal reflection and to select a true pair of the pupil and corneal reflection.

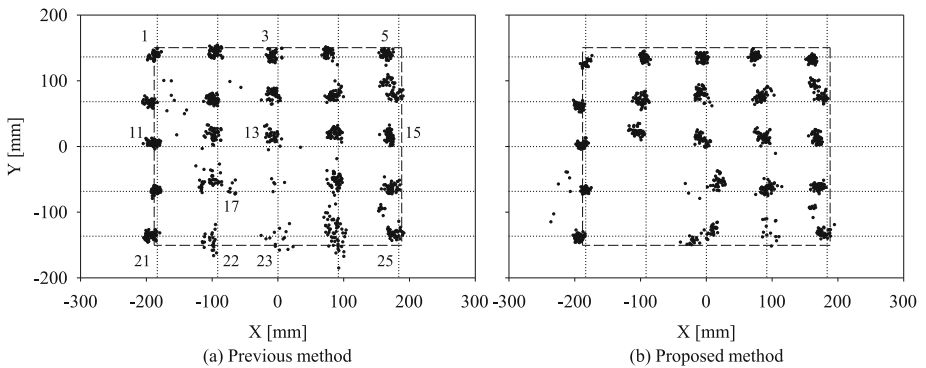


Fig. 6. Detected gaze point distributions in the previous and proposed methods when subject A wore glasses. Dots and intersections of dotted lines indicate the gaze points and visual target positions, respectively. The rectangular area enclosed by the broken lines indicates the PC screen.

3.3 Experiment 3: Gaze Detection When Disturbance Light Sources Generated False Corneal Reflection

Method. Four small disturbance light sources were installed at the four corners of the PC screen, respectively, and they generated the false corneal reflections. Subjects were two university students who wore glasses. The calibration and gaze detection

procedures were the same as in experiment 2, where the distance between the eyes and the screen was approximately 80 cm.

Results and Discussion. Figure 7(a) and (b) compare the averaged gaze points of the left and right eyes for subject A between the previous and proposed methods. In the previous method, the gaze point dispersions were large for many of the 25 targets. Whereas the proposed method showed the smaller dispersion for almost all targets. The gaze error in the previous method for subject A was 3.08 ± 5.62 [deg], whereas that of the proposed method was 1.23 ± 2.55 [deg]. Subject B showed the error of 5.31 ± 10.61 [deg] in the previous method and 2.43 ± 5.52 [deg] in the proposed method, respectively. These results indicate that the proposed method functioned to prevent the misdetection of the false corneal reflections produced by the disturbance light sources.

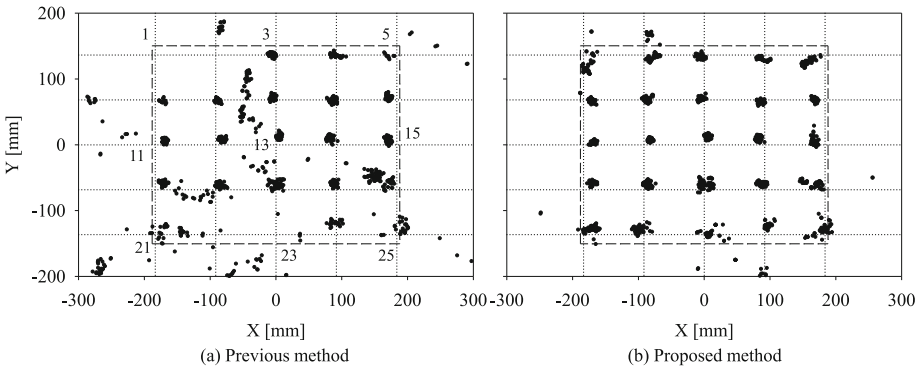


Fig. 7. Detected gaze points (average of right and left eyes) in the previous and proposed methods when four disturbance light sources were installed at four corners of PC screen and generated false corneal reflections.

4 Conclusions

In our remote gaze detection system, in order to prevent the misdetection of the pupil and corneal reflection when a user wears glasses and/or when the disturbance light sources exist, the novel geometrical method based on the optical structure of the eyeball was proposed. The experimental results showed that the proposed method detects a true pair of the pupil and corneal reflection and improves the accuracy of the gaze detection when the glass reflections or the false corneal reflections of the disturbance light sources appear in the camera image. The proposed method would function well also in the other pupil-corneal reflection-based gaze detection systems.

References

1. Tobii Technology. <http://www.tobii.com/>
2. Bakker, N.M., Lenseigne, B.A.J., Schutte, S., Geukers, E.B.M., Jonker, P.P., van der Helm, F.C.T., Simonsz, H.J.: Accurate gaze direction measurements with free head movement for strabismus angle estimation. *IEEE Trans. Biomed. Eng.* **60**(11), 3028–3035 (2013)
3. Tawari, A., Martin, S., Trivedi, M.M.: Continuous head movement estimator for driver assistance: issues, algorithms, and on-road evaluations. *IEEE Trans. Intell. Transp. Syst.* **15**(2), 818–830 (2014)
4. Ebisawa, Y., Fukumoto, K.: Head-free remote eye-gaze detection system based on pupil-corneal reflection method with easy calibration using two stereo-calibrated video cameras. *IEEE Trans. Biomed. Eng.* **60**(10), 2952–2960 (2013)
5. Ebisawa, Y., Fukumoto, K.: Head-free, remote gaze detection system based on pupil-corneal reflection method with using two video cameras – one-point and nonlinear Calibrations. In: Kurosu, M. (ed.) *HCI/HCI 2013, Part IV. LNCS*, vol. 8007, pp. 205–214. Springer, Heidelberg (2013)
6. Nakashima, A., Ebisawa, Y., Nurikabe, Y.: Pupil detection using light sources of different wavelengths. *J. Inst. Image Inf. Telev. Eng.* **60**(12), 2019–2025 (2006)