Head-Free, Remote Eye-Gaze Detection System with Easy Calibration Using Stereo-Calibrated Two Video Cameras

Yoshinobu Ebisawa, Kazuki Abo, and Kiyotaka Fukumoto

Faculty of Engineering, Shizuoka University, Johoku 3-5-7, Hamamatsu, 432-8561 Japan
ebisawa@sys.eng.shizuoka.ac.jp

Abstract. The video-based, head-free, remote eye-gaze detection system based on detection of the pupil and the corneal reflection was developed using stereo-calibrated two cameras. The gaze detection theory assumed the linear relationship; $\theta=k|\mathbf{r'}|$. Here, θ is the angle between the line of sight and the line connecting between the pupil and the camera, and $|\mathbf{r'}|$ indicates the size of the corneal reflection - pupil center vector. Three novel easy calibration methods were proposed; 'automatic', 'one-point', and 'two-point'. In the 'automatic', the user does not have to fixate the specific location in the PC screen. In the 'one-point', the angular difference between the optical and visual axes of the eye was determined and used for compensation. The 'two-point' was proposed to compensate the nonlinear relationship between $|\mathbf{r'}|$ and θ, which occurs when θ is large. The precision of gaze detection was compared among the three methods using the developed gaze detection system.

Keywords: eye-gaze detection, pupil, corneal reflection, user calibration, head-free, infant.

1 Introduction

So far, several non-contact, remote, eye-gaze detection systems allowing head displacements have been developed for human monitoring or human interface [1][2]. However, the current systems require some effort by the user that he or she must fixate several points on a PC screen for gaze calibration [3][4]. Moreover, in many of the other systems, a movable range of the user's head is narrow.

We have already developed a precise head-free eye-gaze detection system with three cameras [5]. To detect pupil easily and to generate the corneal reflection, which are necessary to estimate the eye-gaze point, the near-infrared LEDs were installed in each camera. Two of the cameras are stereo-calibrated wide view cameras, which detect 3-D positions of the pupils. Another one is a narrow view camera with a pan-tilt drive unit. The latter camera tracks one of the pupils and detects the angle of the line of sight by high resolution. The angle was determined using the corneal reflection - pupil center vector.

C. Stephanidis (Ed.): Posters, Part II, HCII 2011, CCIS 174, pp. 151–155, 2011.

Although the resolution of eye-gaze detection of this system was high, the system can detect only one eye. Moreover, the pan-tilt unit may obstruct the progress of development of the system by its high cost and dynamic mechanism, and so on. So, in the present study, we develop two camera systems in which the function of the narrow camera is installed in those of the stereo-calibrated two cameras.

In addition, in the three camera system, we proposed two easy calibration methods: the moving target and two-point calibration methods [5]. Both methods used only the inclination of the corneal reflection - pupil center vector for calibration. Accordingly, by using the size as well as the inclination of the vector, the improvement of the precision is expected. Moreover, in the two-camera system, four points of gaze on the screen can be detected. Averaging them may increase the resolution of gaze detection.

Recently, the recent research showed that the gaze distribution of the infant with autism is peculiar [5], and the application of the eye-gaze detection system to the autism diagnosis support system is expected. In addition, the medical screening at the stage of the infant is hoped for. However, it is difficult to make infants fixate several calibration points. So we propose three easy calibration methods and compare them in the developed two-camera system experimentally.

2 System Configuration and Eye-Gaze Detection Theory

Each optical system of the developed two-camera system consists of a NTSC B/W video camera having a near-infrared sensitivity, a 25 mm lens, an infrared pass filter (IR80), and light sources having many infrared LEDs (Fig. 1(a)). The two cameras are driven with slightly shifting synchronization. The LEDs are lit synchronized with the exposure period (500μs). A PC captures the images of the user's face via an image grabbing board and then detects the centers of a pupil and a corneal reflection for each are detected.

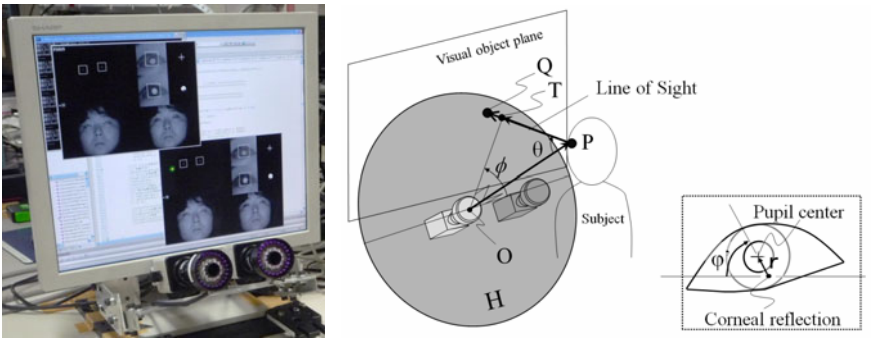
The two cameras are separately set under a 19 inch display. By camera calibration, the camera position, O, comes to be known in the world coordinate system (Fig. 1(b)). By stereo matching, the 3-D pupil center position, P, is detected. Let a gaze point, Q, be the intersection between the line of sight and a visual object plane (PC screen, known), the line of sight and the position of Q are determined by the two angles (θ and ϕ) and the position of P. To understand this theory easily, we define a virtual gaze plane, H, which is perpendicular to line OP and passes the camera position, O. The plane rotates along with the pupil position. The intersection of plane H and line PQ is defined as point T. The position of T in the coordinates system of H is expressed by θ and ϕ.

In the gaze detection theory, the angels, θ and ϕ, are related to the corneal reflection - pupil center vector, **r**, (Fig.1 (b)). Here, actually the vector is calculated as the corresponding 3-D vector, **r'**, and is used. The direction of **r'**, ϕ', is defined as the angle from the horizontal axis in the camera image. Actually ϕ' is compensated by the pose of the camera. First, the theory assumes $\phi'=\phi$. Here, ϕ is angle between line OT and the horizontal axis on the plane H. This horizontal axis is defined as the nodal line between plane H and the horizontal plane in the world coordinate system. Second, the theory assumes a linear relationship between $|\mathbf{r'}|$ and θ, as follows.

$$\theta = k\,|\mathbf{r'}| \tag{1}$$

Therefore, obtaining the value of k in eq. (1) means the gaze calibration. However, in general, there is the several-degree difference between the visual axis and the optical axis of the eye. This difference produces the errors in gaze point detection. To determine the point precisely, the user must fixate one calibration point at least. In addition, when θ is large, the relationship between $|\mathbf{r'}|$ and θ may become nonlinear. It depends on the user individual.

Since the system in this study has the two cameras, the gaze detection theory is expanded as shown in Fig. 2.



(a) (b)

Fig. 1. (a) System appearance and (b) gaze detection theory. (a) The light source arranging the infrared LEDs having two different wavelength in two concentric rings (inner: 850 nm, outer: 940 nm) is attached to each camera. When the inner LEDs turn on, the pupils become bright. When the outer LEDs turn on, they become dark. These two groups of LEDs are alternately turned on synchronously with the field of the camera. By differentiating the consecutive bright and dark pupil images, the pupils become conspicuous in the difference image due to cancellation of the background images except for the pupils. The pupils are detected in the difference image. (b) See in the text.

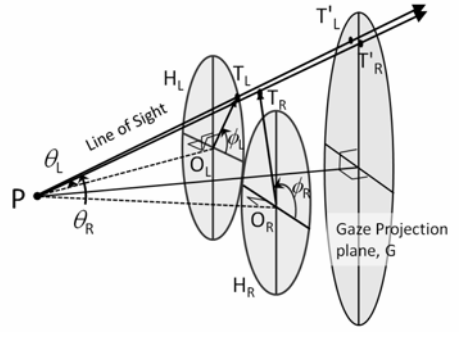


Fig. 2. The expanded gaze detection theory. P indicates one of the two eyes. O_L and O_R indicate the positions of the two cameras. If the calibration value, k, is determined, the two lines of sight can be calculated for one eye. For another eye, the two lines of sight can be also determined.

3 Eye-Gaze Calibration Methods

We tried the following three calibration methods. (a) 'Automatic' calibration method: The user does not need to fixate any known point. In Fig. 2, the vectors O_LT_L and O_RT_R are determined by k, $|\mathbf{r'_L}|$, $|\mathbf{r'}_R|$, ϕ_L' and ϕ_R'. T_L and T_R are projected onto the gaze projection plane, G. To minimize the distance between the two projected points (T_L' and T_R') on plane G, the value of k was determined. This method assumes to detect the optical axis of the eyeball. (b) 'One-point' calibration method: The user fixates one point presented at the center of the PC screen. The horizontal and vertical angular differences between the optical and visual axes (ΔQ_x, ΔQ_y) are recorded. The value of k is also determined by the same method as method (a). They are used for gaze detection. (c) 'Two-point' calibration method: The user fixates not only the calibration point presented at the center but also another point presented at the top of the screen. This method is proposed to compensate the nonlinear relationship between θ and |r|; this may occur when θ is large. The following equations are used.

$$\theta = \begin{cases} k_1|\mathbf{r}'| & k_1|\mathbf{r}'| \le \theta_B \\ \theta_B + k_2(|\mathbf{r}'| - |\mathbf{r'_B}|) & k_1|\mathbf{r}'| > \theta_B \end{cases}, \tag{2}$$

where θ_B and $|\mathbf{r'_B}|$ indicate the averages of θ and $|\mathbf{r'}|$ when the user fixated the center calibration point. The values of k and (ΔQ_x, ΔQ_y) obtained by method (b) are used for gaze detection. The value of k is used as k_1.

4 Experiments

Three university students participated in the experiments. In the calibration procedure, the subject fixated the two calibration points; the center and top of the screen. After calibration procedure, the subject fixated nine targets evenly arranged on the PC screen one by one. 30 gaze points (1 sec) were acquired for each target. In both procedures, the distance between the subject's face and the PC screen was approximately 80 cm. Fig. 3 (a)-(c) compared the gaze point distributions among the three calibration methods for subject KY. Each dot shows the average of both eyes' gaze points. Fig. 3 (d) compares the average gaze detection errors in visual angle among the three calibration methods. The three subjects showed the average errors of 2.16 deg in the 'automatic', 1.37 deg in the 'one-point', and 0.94 deg in the 'two-point' calibration method.

In another experiment, the same three subjects moved back and forth by ±10 cm and right and left by ±5 cm. At seven positions within the range, they fixated the nine targets. Here, either was selected from among the one-point and two-point calibration methods depending on subjects. The average and SD of the gaze error for the three subjects was 0.92±0.40 deg.

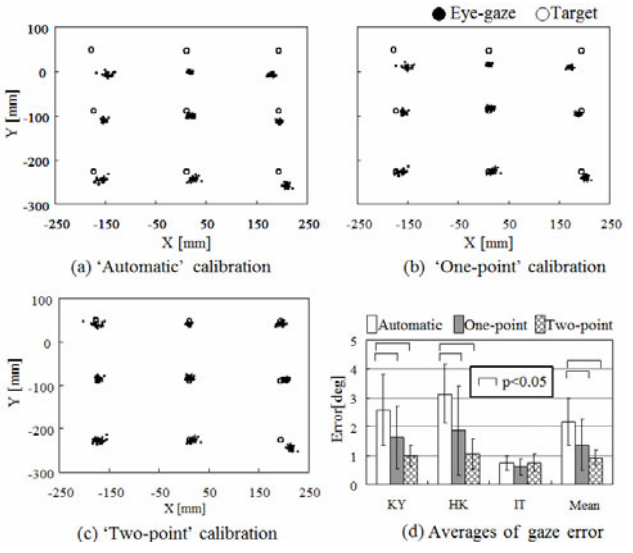


Fig. 3. Comparison of results among three easy calibration methods

5 Conclusions

The eye-gaze detection system using the stereo-calibrated two cameras was developed. The three novel easy calibration methods were proposed. The automatic and one-point calibration methods may be useful for the infant. If calibrated well, the developed gaze detection system experimentally showed the average gaze error less than 0.92±0.40 deg for 20 cm-back and forth, and 10 cm-right and left head movements.

References

1. Guestrin, E.D., Eizenman, M.: General theory of Remove Gaze Estimation Using the Pupil Center and Corneal Reflections. IEEE Trans. on Bioml. Eng. 53(6), 1124–1133 (2006)
2. Zhu, Z., Ji, Q.: Novel Eye Gaze Tracking Techniques under Natural Head Movement. IEEE Trans. on Bioml. Eng. 54(12), 2246–2260 (2007)
3. Tobii Technology, http://www.tobii.com/
4. Eye Tech Digital Systems, http://www.eyetechds.com/
5. Kondou, Y., Ebisawa, Y.: Easy Eye-Gaze Calibration using a Moving Visual Target in the Head-Free Remote Eye-Gaze Detection System. In: Proc. of VECIMS 2008, pp. 145–150 (2008)
6. Jones, W., Carr, K., Klin, A.: Absence of Preferential Looking to the Eyes of Approaching Adults Predicts level of Social Disability in 2-Year-Olds Toddlers with Autism Spectrum Disorder. Archives of General Psychiatry 65(8), 946–954 (2008)