An Eye-Gaze Input System Using Information on Eye Movement History

Kiyohiko Abe[1], Shoichi Ohi[2], and Minoru Ohyama[3]

[1] College of Engineering, Kanto Gakuin University, 1-50-1 Mutsuura-higashi, Kanazawa-ku, Yokohama, Kanagawa 236-8501, Japan
abe@kanto-gakuin.ac.jp
[2] School of Engineering, Tokyo Denki University, 2-2 Kandanishiki-cho, Chiyoda-ku, Tokyo 101-8457, Japan
ohi@d.dendai.ac.jp
[3] School of Information Environment, Tokyo Denki University, 2-1200 Muzaigakuendai, Inzai-shi, Chiba 270-1382, Japan
ohyama@sie.dendai.ac.jp

Abstract. We have developed an eye-gaze input system for people with severe physical disabilities such as amyotrophic lateral sclerosis. The system utilizes a personal computer and a home video camera to detect eye gaze under natural light. It also compensates for measurement errors caused by head movements; in other words, it can detect the eye gaze with a high degree of accuracy. We have also developed a new gaze selection method based on the eye movement history of a user. Using this method, users can rapidly input text using eye gazes.

Keywords: Eye-gaze input, Communication aid, Natural light, Gaze selection, Welfare device.

1 Introduction

Recently, the eye-gaze input system was reported as a novel human-machine interface [1,2,3,4,5,6,7]. The operation of this system only requires the user eye movements. Using this system, many communication aid systems have been developed for people with severe physical handicaps such as amyotrophic lateral sclerosis.

Eye-gaze input systems commonly employ a non-contact-type eye-gaze detection method. Infrared or natural light can be used as the light source for eye-gaze detection. The detection method that uses infrared light can detect an eye gaze with a high degree of accuracy [1,2,3]. However, it requires an expensive device. The detection method that uses natural light requires an ordinary device such as a home video camera or a personal computer; therefore, a system using this method is cost-effective [4]. However, the systems that operate under natural light often have a low accuracy. Therefore, they are capable of classifying only a few indicators. This problem can be resolved by using multi-camera systems [5].

C. Stephanidis (Ed.): Universal Access in HCI, Part II, HCII 2007, LNCS 4555, pp. 721–729, 2007.

We have developed a new eye-gaze input system with multi-indicators. This system utilizes a personal computer and a home video camera to detect eye gaze under natural light [7]. The system is not only inexpensive but also user-friendly, thereby making it fit for personal use. A new eye-gaze input system, which employs multi-indicators (12 indicators), was developed for text input. It can input English or Japanese text at a faster rate by using the information obtained from the eye movement history of a user.

This paper presents the evaluation experiments of the proposed eye-gaze text input system.

2 Eye-Gaze Detection Methods by Image Analysis

The eye-gaze input has to detect the eye gaze of users. Several eye-gaze detection methods have been studied [1.2.3.4.5.6]. The method of iris tracking is the most popular detection method under natural light by image analysis [4.5.6]. However, it is difficult to distinguish the iris from the sclera by image analysis. Therefore, the degree of accuracy of this method is not very high. We propose a new method for detecting the eye gaze in both the horizontal and vertical directions by image analysis. The proposed method can detect the eye-gaze of users from a single eye image captured using one home video camera and does not require special devices such as infrared lights. The horizontal eye-gaze is detected by the limbus tracking method that is extended by image analysis [7], and the vertical eye-gaze is detected by the change in the light intensity distributions in the eye image associated with eye movement.

We now describe our eye-gaze input detection method in detail.

2.1 Horizontal Eye-Gaze Detection

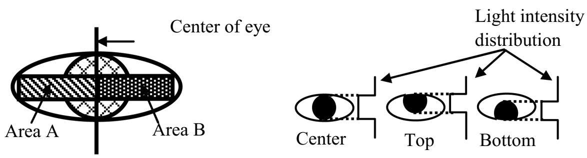
The overview of the proposed horizontal eye-gaze detection method is shown in Fig. 1 (a). This method detects a horizontal eye-gaze by image analysis based on the limbus tracking method. The difference in the reflectance between the iris and the sclera is used for detecting the horizontal eye-gaze; in other words, the eye gaze is estimated by a difference between the integral values of the light intensity on areas A and B, as shown in Fig. 1 (a).

We define this differential value as the eye-gaze value, and it indicates the value of the horizontal eye-gaze. The eye-gaze value and the angle of sight are nearly proportional to each other. Therefore, if the eye-gaze input system is calibrated using this relation, the horizontal eye-gaze of the users can be estimated. This method is simple and it is robust against the noise in an image of the eye such as the fluctuation of light. Therefore, the system using this method can rapidly detect the horizontal eye-gaze at a high sampling rate.

2.2 Vertical Eye-Gaze Detection

Vertical eye-gaze detection requires a different process as compared to that used in horizontal detection, since a major part of the sclera is hidden by the change in the eyelid shape with eye movement. We focus on the vertical movement of the iris for the detection of the vertical eye-gaze. The light intensity distribution from the eye image changes with the iris movement. The vertical eye-gaze can be detected by using this change.

Concretely speaking, the system stores eye images gazing on the indicators aligned vertical. The light intensity distributions (the results of one-dimensional projection) are calculated using these eye images as the reference data. The vertical eye-gaze of the users can be detected by pattern matching using these reference data. The overview of the proposed vertical eye-gaze detection method is shown in Fig. 1 (b). This figure describes a detection method for the three eye-gaze directions, namely, top, center, and bottom.



(a) Horizontal detection (b) Vertical detection

Fig. 1. The overview of eye-gaze detection

3 Proposed Eye-Gaze Input System

3.1 System Configuration

The proposed eye-gaze input system comprises a personal computer, a home video camera, and an IEEE1394 interface for image capture from the video camera. The personal computer runs an image analysis software for eye-gaze detection in the Windows XP platform. This system does not require an exclusive image processing device.

The video camera used in this system records the images of the user's eye from a distant location (distance between the user and camera is approximately 70 cm) and then this image is enlarged. The head movements of the user induce a large error in the measurements. Therefore, this system must compensate for the user head movement. Methods such as tracing a marker on glasses have been developed for the compensation of head movement [6]. We compensated for the head movement by tracing the location of an inner corner within the eye.

The proposed system executes the eye-gaze detection process every 200 ms (CPU: Pentium 4; clock frequency: 2.2 GHz). The hardware configuration of this system is shown in Fig. 2. A screenshot of the proposed system is shown in Fig. 3 (a). The center of the display is assigned to the workspace. The workspace is used for displaying an application software window. Therefore, this system can input text while the application software is running.

Users must calibrate the system before using it as described above. After adjusting the camera location, the calibration begins. While the calibration is being performed, users gaze at each of the five indicators shown in Fig. 3 (b). If the calibration terminates successfully, the users can select each indicator by an eye gaze.

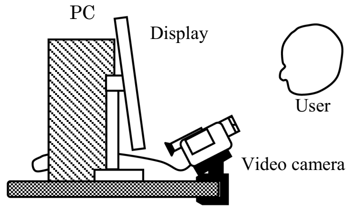
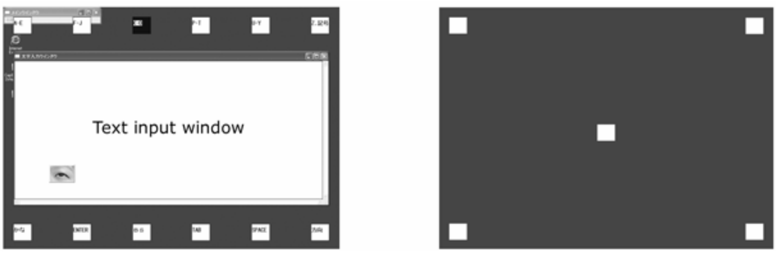


Fig. 2. Hardware configuration



(a) (b)

Fig. 3. Screenshot (a) and indicators for calibration (b)

3.2 Text Input by Eye Gaze

The text input method by using the above system is described below. Our proposed eye-gaze detector is capable of classifying 12 indicators on a display. However, around 60 indicators are required to input Japanese text (our mother tongue). Similarly, 12 indicators are insufficient for English text input. This is because the English language contains upper- and lowercase letters and symbols. Moreover, control keys are required for text input.

We designed an interface for Japanese and English text input, as shown in Fig. 4. This interface requires two selections: one for character input and another for character group selection, for example, “groups A to E.” Fig. 4 shows the text input method used in this system. Alphabets and symbols (“Etc.” in Fig. 4) are selected by using two selections; however, commonly used characters (for example, “Space”) are selected by using one selection.

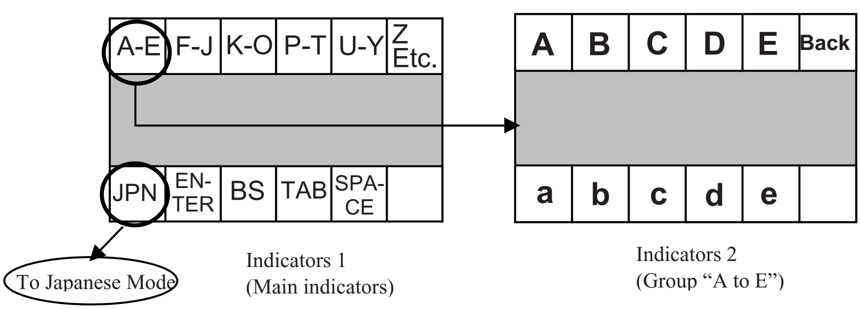


Fig. 4. Interface for English text input

4 Eye-Gaze Input Method Using Eye Movement Information

The eye-gaze input decision requires the detection of not only the location of a user's gazing point but also the user's command for an indicator (assigned character) selection. An input decision can be made either by blinking on gazing or by using eye fixations (measuring the time for which the eye fixates on a target such as the indicators). Blinking occurs as a physiological phenomenon. Therefore, it is difficult to classify the conscious blinking of a user. Hence, we consider the method involving the use of eye fixations for the eye-gaze input decision. In this method, if the user gazes at an indicator for more than a defined period of time, the state of the user's eye gaze is classified as "gazing." Then, the gazed indicator is selected for input.

However, we confirmed that many input errors occur by directly using this method. When the user gazes at an indicator, the eye gaze of the user drifts around a gazed indicator. If the speed of eye movement is very slow during this drifting, the eye-gaze input system makes an erroneous decision for estimating the gazed indicator. In order to resolve this problem, we have developed methods for eye fixation detection. This method detects eye fixations using the information on the eye movement history (or gaze state history) of users. We now describe the methods we have developed for eye fixation detection. One of these methods has already been reported, while the other is a completely new method for eye fixation detection.

4.1 Method 1 (Our Conventional Method)

We have reported the eye-gaze input method that uses the gaze state history of users. In this method, the gaze state is defined by including two states such as the initial and continuous states. Hence, if the users gaze at an indicator for more than a defined period of time, we define this state as the initial state. After the initial state is completed, the user's state changes to the continuous state. Although the users gaze at the same indicator in these states, eye-gaze measurement errors may occur due to unexpected eye movements such as those associated with blinking. Therefore, the gazed indicator is determined based on the entire continuous state; the majority number of gazed indicator is extracted as the user's decision. A schematic illustration of this input decision is shown in Fig. 5.

In this figure, the small rectangles indicate a character assigned to an indicator. The measurement ranges of the initial and continuous gaze states are five and 10 samples, respectively.

By using method 1, we confirmed that five subjects could input text at an average rate of 6.9 characters/min. However, this method should count the same indicators (e.g., five samples) for classifying the initial gaze state. Therefore, if the user's gaze drifts out of the gazed indicator, the system resets the process of classifying the initial gaze state and restarts the process from the beginning. Further, method 1 does not confirm whether the indicators of a candidate input at the initial and continuous gaze states are equal or not. Therefore, method 1 has the disadvantage of causing input errors when the indicators of the candidate input are not equal.

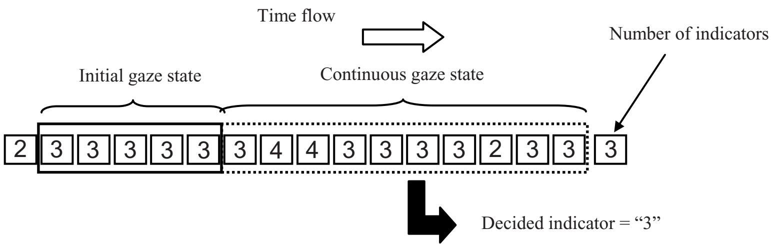


Fig. 5. Input decision based on eye movement history. (Method 1)

4.2 Method 2 (Our Proposed Method)

To eliminate the disadvantages of method 1, we have developed a completely new eye-gaze input method using the gaze state of users. By using this method (method 2), users can input text at higher speed and with greater accuracy as compared to that using method 1. Method 2 measures the gaze state history using two gaze states as effectively as method 1. The majority numbers of gazed indicator in each state are extracted as the candidate inputs. If these two indicators are equal, the candidate input is selected. By using method 2, the input decision process is not reset even if the user blinks or a drifting eye movement occurs.

In the state where the user is tracing an indicator, adjacent gaze selection data are random; in other words, the eye-gaze input is not predetermined because the two candidate input indicators are not equal. A schematic illustration of this input decision is shown in Fig. 6. In this figure, the small rectangles indicate a character assigned to an indicator. In addition, a user's gaze is fixed on a selected indicator after an eye-gaze input decision. This gaze state was investigated in an experiment. The length of this state depends on the user. This eye fixation induces input errors. Users can set the measurement ranges of the two states for the eye-gaze input decision. Therefore, users should appropriately adjust these ranges to decrease input errors.

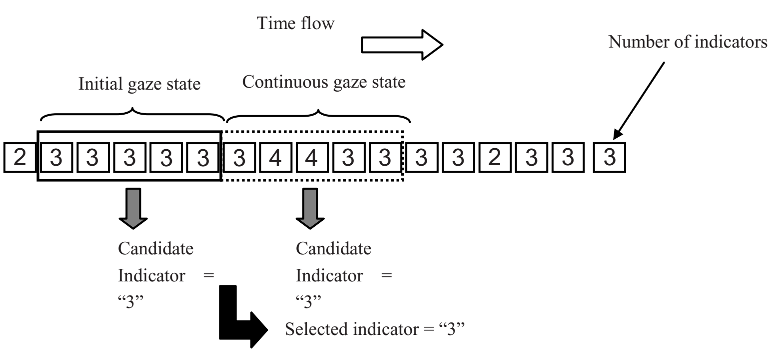


Fig. 6. Input decision based on eye movement history. (Method 2)

5 Evaluation Experiment for Proposed System

The evaluation experiments were conducted with five subjects using method 2 as described above. In these experiments, the measurement range for the eye movement history is varied. It changes to 12, 10, 8, 6, or 4 samples. The experiment includes two measurement steps (initial and continuous gaze states), as shown in Fig. 6. We confirm the following experimental parameters: rates of gaze selection, rates of text input, and errors in gaze selection. Prior to the experiments, the subjects memorized the input text, which was as follows: "I am using our eye-gaze input system." It contains 37 characters and requires 68 gaze selections for input. The experimental system is calibrated at the beginning of the experiments.

When a user commits an error during the text input, it can be corrected by the following two procedures. In case the input error occurs during group selection

Table 1. Rate of gaze selection [selections/min]

Subject	Measurement range [samples]				
	12	10	8	6	4
A	19.4	23.5	25.3	32.7	34.6
B	20.1	23.4	25.9	34.3	33.0
C	20.4	24.0	28.4	30.1	22.9
D	16.2	19.5	18.6	29.2	24.1
E	17.5	22.7	23.8	22.6	34.5
Average	18.7	22.6	24.4	29.8	29.8

Table 2. Rate of text input [characters/min]

Subject	Measurement range [samples]				
	12	10	8	6	4
A	10.6	12.8	13.8	17.8	18.8
B	11.0	12.7	14.1	18.7	18.0
C	11.1	13.0	15.5	16.4	12.4
D	8.8	10.6	10.1	15.9	13.1
E	9.5	12.4	13.0	12.3	18.7
Average	10.2	12.3	13.3	16.2	16.2

Table 3. Errors in gaze selection [counts]

Subject	Measurement range [samples]				
	12	10	8	6	4
A	0	4	8	2	2
B	0	4	2	0	3
C	2	0	0	10	37
D	0	8	12	9	35
E	0	0	5	28	9
Average	0.4	3.2	5.4	9.8	17.2

(indicators 1, as shown in Fig. 4), the subject must gaze at “Back” (on indicators 2) in order to return to indicators 1. If an input error is noticed after text input, it can be corrected by gazing at “BS” (backspace) on indicators 1, following which the subject can input the correct character.

The gaze selection and text input rates are shown in Table 1 and Table 2, respectively. These rates include the time required for error correction. We define over-selection counts as errors in the text input. The errors in gaze selection are shown in Table 3.

6 Discussion

We summarize the obtained results as follows: From Tables 1 and 2, it is evident that the rate of text input by a subject increases with the narrowing of the measurement range of the eye movement history. However, it appears that there is a decrease in the rates of gaze selection and text input by subjects B, C, and D when the measurement range is reduced to four samples. This is due to the increase in the errors when the measurement range of the eye movement history is narrowed, as shown in Table 3. In other words, there is a decrease in the required gazing (dwell) time. We surveyed the comments about these errors. Subjects B, C, and D say that many errors occur on selecting the correction indicator and therefore input correction is unsuccessful.

As shown by the above experimental results, the average input rate is the highest when the measurement range of the eye movement history is either four or six samples. In these cases, subjects can input text at an average rate of 16.2 characters/min. Hence, our system can input text at a rate that is comparable to that of a conventional infrared eye-gaze input system without requiring word prediction [3]. The proposed system analyzes the image of an eye taken under natural light, and it does not employ an exclusive device for this purpose; therefore, it has a lower sampling rate than most conventional systems that employ infrared light [3]. The sampling time of the proposed system is approximately 200 ms; under such conditions, the user can input English text at a comfortable rate.

Apart from the reported results, we also observed that subject A was able to input text at the largest rate when the measurement range of the eye movement history was four samples. This result shows that the optimum parameters are different for individual subjects. Therefore, it is desirable to vary the measurement range when this system is utilized.

7 Conclusions

We present a new eye-gaze input system based on the user’s eye movement history, which is suitable for personal use. This system is also compact and easy to use since it does not require exclusive image processing units or sensors.

In the evaluation experiments, the measurement range of the eye movement history is varied. The results for five subjects show that the optimum parameter for text input is approximately six samples (1.2[sec.]), and the average rate of English text input is approximately 16.2 [characters/min]. This result shows that the text input rates are approximately 2.5 times greater than that in our previous system. The results also indicate that the rate of text input using the proposed system is comparable to that of a conventional system that employs infrared light.

We will attempt to conduct additional evaluation experiments in the future. The change in the error rate will be researched in a prolonged experiment. We will also develop a new eye-gaze input system that is more comfortable to use and inputs at a greater rate. The new system will be capable of dynamically altering the arrangements of indicators by utilizing the character occurrence rate [2].

References

1. Huchinson, T.E., White Jr., K.P., Martin, W.N., Reichert, K.C., Frey, L.A.: Human-computer Interaction using Eye-gaze Input. IEEE Trans. Systems, Man, and Cybernetics 19,7, 1527–1534 (1989)
2. Ward, D.J., MacKay, D.J.C.: Fast Hands-free Writing by Gaze Direction. Nature 418, 838 (2002)
3. Hansen, J.P., Torning, K., Johansen, A.S., Itoh, K., Aoki, H.: Gaze Typing Compared with Input by Head and Hand. In: Proc. of the Eye Tracking Research and Applications Symposium on Eye Tracking Research and Applications, pp. 131–138 (2004)
4. Corno, F., Farinetti, L., Signorile, I.: A Cost-effective Solution for Eye-gaze Assistive Technology. In: Proc. IEEE International Conf. on Multimedia and Expo, 2, 433–436 (2002)
5. Wang, J.G., Sung, E.: Study on Eye gaze Estimation. IEEE Trans. on Systems, Man and Cybernetics 32,3, 332–350 (2002)
6. Kim, K.N., Ramakrishna, R.S.: Vision-based Eye-gaze Tracking for Human Computer Interface. In: Proc. IEEE International Conf. on Systems, Man and Cybernetics, 2, pp. 324–329 (1999)
7. Abe, K., Ohi, S., Ohyama, M.: An Eye-gaze Input System based on the Limbus Tracking Method by Image Analysis for Seriously Physically Handicapped People. In: Adjunct Proc. 7th ERCIM Workshop "User Interface for All", pp. 185–186 (2002)