Gaze Tracking System for Gaze-Based Human-Computer Interaction

Takehiko Ohno†*, Naoki Mukawa, and Atsushi Yoshikawa*

Abstract

We introduce a novel gaze-tracking system called FreeGaze, which is designed to support gaze-based human-computer interaction (HCI). Existing gaze-tracking systems require complicated and burdensome calibration, which prevents gaze being used to operate computers. To simplify the calibration process, FreeGaze corrects the refraction at the surface of the cornea. Unlike existing systems, this calibration procedure requires each user to look at only two points on the display. After the initial calibration, our system needs no further calibration for later measurement sessions. A user study shows that its gaze detection accuracy is about 1.06° (view angle), which is sufficient for gaze-based HCI.

1. Introduction

Computers are now widely used in our daily lives, and it is difficult to live without them in our offices and homes. The use of computers is also changing. Thirty years ago, they were enshrined in an air-conditioned room and shared by many users. Computers then came to our desks and we have been using them as a convenient tool like a word processor and a spreadsheet. They are now breaking free from the desks and a wide diversity of computer styles is emerging, *e.g.*, personal digital assistants, cellular phones, and information appliances. In the near future, computers will be integrated and embedded into everything, so the term "using a computer" will become redundant.

The diversity of computers requires improvements in computer operation methods. Typical peripherals for operating a computer, like a keyboard and mouse, provide efficient and effective access to the computer. However, they are not suitable for the emerging computer styles. For example, who can imagine that the numerous computers in a home will each have a keyboard and a mouse? Such peripherals also create an information barrier: users have to spend a lot of time acquiring the operation skills needed. For those people who cannot use them, computers are simply inconvenient tools.

Gaze-based human-computer interaction (HCI), which enables users to operate computers by means of eye movement, is a prime candidate to solve the problem. Many interactive systems integrated with a gaze-tracking system have been proposed. One features selection by gaze. Without using a mouse or a keyboard, users can select an icon or menu item by just looking at it for a while [1]-[3]. Continuous gaze can be replaced by separating the selection area from the menu, which enables very fast selection by gaze [4]. A continuous zooming selection method called Dasher, which achieves fast text input, has been proposed [5]. Interaction between gaze and other input devices like a mouse also accelerates selection [6]. Another way of using the eye is to extract the user's attention and intention. Starker *et al.* proposed a method of detecting the user's attention from his eye movement pattern while he is looking at objects on the screen (Starker *et al.*, 1990). In their system, a narration about the object of interest is given. In the near future, it will lead to various helpful methods of operating computers.

Any gaze-tracking system for gaze-based HCI must be comfortable to use. However, current systems are far from comfortable and are not at the same level as traditional input devices like the mouse and keyboard. In this paper, we describe a gaze-tracking system called *FreeGaze*, which is designed to achieve

† NTT Communication Science Laboratories
Atsugi-shi, 243-0198 Japan
E-mail: takehiko@brl.ntt.co.jp

truly effective gaze-based HCI. Unlike existing systems, this novel system drastically reduces the effort of initial setup. It only requires the user to look at two points on the display for calibration.

In the rest of this article, first we describe the requirements of gaze-tracking systems and then present our gaze-tracking method. Next, we describe our prototype gaze-tracking system and discuss its performance and remaining problems.

2. Requirements for gaze-tracking system for gaze-based HCI

To use the user's gaze for HCI, it is important that gaze tracking should not disrupt the user's activities. Traditional input devices do not demand much effort to operate them. We can use them as soon as we want to operate the computer, and there are few restrictions on our behavior during these activities. Gaze-based HCI should be more comfortable, or at least, as comfortable as other methods currently in use. Therefore, gaze-tracking systems will play an important role in achieving the required comfortable operation. The structure of the eyeball is shown in Fig. 1. To use gaze for HCI, the gaze-tracking system should meet the following requirements.

1) Little or no configuration setup, including personal calibration. Existing gaze-tracking systems require personal calibration for accurate gaze detection, where the user looks at five to twenty calibration markers on the display (Fig. 2(a)). This is a time consuming and burdensome task. Compared with other operation methods, gaze-tracking is not carefree.
2) Free movement of the user's head position. Many gaze-tracking systems prevent the user from moving his head freely because a fixed camera is used to detect the user's eye. This limitation greatly hinders comfortable computer operation.
3) No eye-detection camera or other attachments near the user's eye; otherwise, continuous use is tiring.

The challenge of developing a gaze-tracking system is to satisfy all of the above requirements. To satisfy the first one, Shih *et al.* proposed a calibration-free gaze-tracking technique [7] in which multiple cameras and multiple point light sources are used to estimate gaze direction. They performed computer simulations to confirm their method. However, since the user's visual axis is different from the estimated gaze direction, which equals the optical axis of the eyeball, some measurement error remains.

Some gaze-tracking systems use a pan-tilt camera

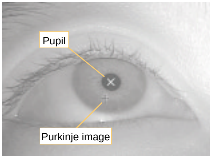

Fig. 1. Pupil and Purkinje image.

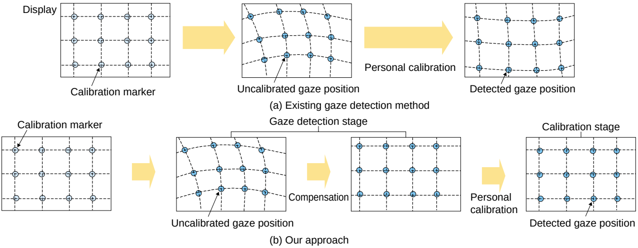

Fig. 2. Personal calibration method.

or movable mirror to track the user's head movement. However, they do not satisfy the third requirement because they require the user to wear special eyeglasses or a magnetic positioning sensor to detect the head's position.

3. Our approach: gaze-tracking method with easy personal calibration

To develop a gaze-tracking system for gaze-based HCI, we mainly focused on the first and third requirements because they are important in making a gaze-based HCI that can be used in our everyday lives. To satisfy the third requirement, we place a camera in front of the user to detect his gaze. To reduce the burden of calibration (first requirement), we propose a novel gaze-tracking method that requires only two points on the display for personal calibration. Once the initial calibration is finished, the system reuses the calibration data, so there is no need to repeat the calibration step.

The gaze-tracking method consists of two stages (Fig. 2(b)): the rough gaze detection stage and the calibration stage [8]. In the first stage, the center of the pupil and the Purkinje images (Fig. 3) are detected to calculate the rough gaze direction. For this calculation, a geometric eyeball model is used to correct the major factors of gaze detection error caused by refraction at the surface of the cornea (Fig. 4). In the model, we prepare an edge set of the observed virtual pupil image. The real position of the each edge point is then estimated, and the center of the real pupil is derived from the set of real pupil edges. The center of the cornea curvature is estimated from the position of the Purkinje image, and the gaze direction is defined as the difference between two points: the center of the pupil and the center of the corneal curvature. If we use the center of the observed pupil (virtual pupil image in Fig. 4) to calculate the gaze direction, then the derived direction differs from the real direction, which causes gaze direction error.

In the second stage, residual error is compensated for by a two-point calibration. In our approach, when the user looks at the markers, which are positioned as a lattice-like array, the output of the first stage is positioned in an equally spaced orthogonal array. Therefore, linear transformation by two calibration markers can suppress the residual error.

4. Implemented system: FreeGaze

We implemented a prototype system called FreeGaze to check the accuracy of our method. Figure 5 shows a photograph of FreeGaze. It has a CCD (charge-coupled device) camera that is sensitive to near-infrared light to detect the user's eye. Below the camera, there is an infrared LED (light-emitting diode) array. The combination of CCD camera and infrared LED enables us to detect the pupil and the Purkinje image from the user's eye. The eyeball position is measured by the depth-from-focus method. Then the gaze direction is derived from the pupil and the Purkinje image using the gaze detection method described in section 3. The user's gaze position is obtained as the intersection of the gaze direction and the display (Fig. 6).

FreeGaze is controlled by a Windows 2000 or Linux personal computer, where the gaze direction is derived from the captured eye image in real time. The

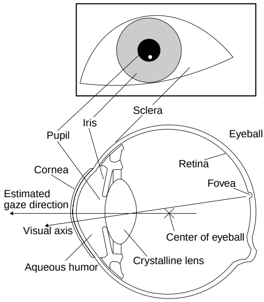

Fig. 3. Structure of the eyeball.

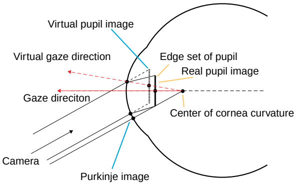

Fig. 4. Geometric eyeball model.

Fig. 5. Gaze Tracking System *FreeGaze*.

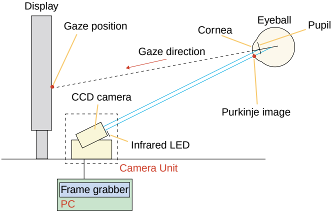

Fig. 6. Overview of gaze detection method.

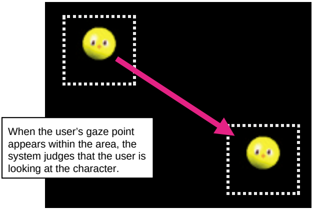

Fig. 7. Calibration with no additional devices.

computer also controls the LED brightness, CCD gain, and camera focal distance. The sampling rate is 30 frames per second. The current prototype restricts the user's eye position to a volume 4 cm high by 4 cm wide by 2 cm deep, because the camera direction is fixed.

A user study showed that the gaze detection accuracy is about 1.06° (view angle) for naked-eye users in the head-free condition, which is sufficient for gaze-based interaction. This system works well even if the user is wearing spectacles.

5. Discussion

Our method does not completely satisfy the first requirement because it does require a simple calibration. However, the calibration procedure takes only a few seconds and is necessary only at the beginning of use: the calibration parameters can be reused for the same user. We believe that it is a reasonable solution to accurate gaze detection.

Personal calibration requires an additional input device to detect when the user is looking at the calibration marker. For personal use, it is possible to use regular devices like the keyboard and mouse. However, for public use, like an information terminal in a public space, it is sometimes difficult to provide extra devices. In those cases, personal calibration with no additional selection method is necessary. Our gaze-tracking system suggests a way to meet this requirement. In the calibration procedure, a small agent-like character appears in the top-left corner of the display (Fig. 7) and jiggles about to attract the user's attention. After the user looks at the character, it moves to the opposite corner. Once the user looks at this position, personal calibration is complete. This method is possible because FreeGaze can roughly estimate the user's gaze position even if personal calibration has not yet been performed, and only two points are required for calibration. FreeGaze and its simple personal calibration technique enable the user to control the computer by gaze even if there are no additional computer operation methods.

The second requirement (free movement of the user's head position) is a remaining problem. In our efforts to date, we have developed a head-free system [9] that has a stereo-camera unit to detect the user's eye position integrated with a gaze-tracking unit placed on a pan-tilt stand.

6. Conclusion

We have designed a novel gaze-tracking system for gaze-based HCI. It requires only two points for the initial personal calibration, and no recalibration is required thereafter. We are planning to develop a more precise geometric eyeball model for more accurate gaze detection. We also plan to apply FreeGaze to various gaze-based HCI applications that support our everyday lives.

References

[1] J. P. Hansen, A. W. Anderson, and P. Roed, "Eye-Gaze Control of Multimedia Systems", Y. Anzai, K. Ogawa and H. Mori (Eds.), In Symbiosis of Human and Artifact. Proceedings of the 6th international conference on human computer interaction, Vol. 20A, Elsevier Science Publisher, pp. 37-42, 1995.

[2] R. J. K. Jacob, "What You Look At is What You Get: Eye Movement-Based Interaction Techniques," In Proceedings of CHI'90, ACM Press, pp. 11-18, 1990.

[3] R. J. K. Jacob, J. J. Leggett, B. A. Myers, and R. Pausch, "Interaction Styles and Input/Output Device," Behaviour and Information Technology, Vol. 12, No. 2, pp. 69-79, 1993.

[4] T. Ohno, "Features of Eye Gaze Interface for Selection Tasks," The Third Asia Pacific Computer Human Interaction -APCHI'98, IEEE Computer Society, pp. 176-182, 1998.

[5] D. J. Ward and D. J. C. MacKay, "Fast Hands-Free Writing by Gaze Direction," Nature, Vol. 418, 838, 2002.

[6] S. Zhai, C. Morimoto, and S. Ihde, "Manual and Gaze Input Cascaded (MAGIC) Pointing," In Proceedings of CHI'99, pp. 246-253, 1999.

[7] S. W. Shih, Y. T. Wu, and J. Liu, "A Calibration-Free Gaze Tracking Technique," In Proceedings of the International Conference on Pattern Recognition (ICPR'00), pp. 201-204, 2000.

[8] T. Ohno, N. Mukawa, and A. Yoshikawa, "FreeGaze: A Gaze Tracking System for Everyday Gaze Interaction," In Proceedings of Eye Tracking Research & Applications Symposium 2002, pp. 125-132, 2002.

[9] T. Ohno, N. Mukawa, and S. Kawato, "Just Blink Your Eyes: A Head-Free Gaze Tracking System," In CHI 2003 Extended Abstracts, to appear, 2003.

[10] C. Morimoto and M. Flickner, "Real-Time Multiple Face Detection Using Active Illumination," In Fourth IEEE International Conference on Automatic Face and Gesture Recognition, pp. 8-13, 2000.

[11] I. Starker and R. A. Bolt, "A Gaze-Responsive Self-Disclosing Display," In Proceedings of CHI'90, 3-9, ACM Press, 1990.

Takehiko Ohno

Research Scientist, Communication Scene Analysis Research Group, NTT Communication Science Laboratories.

He received the B.S. and M.S. degrees in information science from Tokyo Institute of Technology, in 1992 and 1994, respectively. In 1994, he joined the Basic Research Laboratories. His research areas include gaze-based human-computer interaction, gaze tracking methods, cognitive modeling and communication analysis. He is a member of ACM, The Information Processing Society of Japan, and Japanese Cognitive Science Society.

Naoki Mukawa

Senior Research Scientist, Supervisor, Communication Scene Analysis Research Group, NTT Communication Science Laboratories.

He received the B.E., M.E., and Ph.D. degrees from Waseda University in 1974, 1976, and 1994, respectively. In 1976 he joined NTT. Since then, he has been engaged in research on image processing, computer vision, pattern recognition, image analysis and nonverbal communication. He also worked on developing image coding systems and visual telephony standards. From 1994 to 2000 he worked for NTT DATA Corporation and contributed to the establishment of an image processing research group. Since 2000 he has been with NTT Communication Science Laboratories.

Atsushi Yoshikawa

Senior Researcher and Manager, NTT DATA Business Incubation Center, and Research and Development Headquarters.

He received a Ph.D. degree in science and engineering from Keio University, in 1991. In 1991, he joined the Software Laboratories, NTT. In 1993, he moved to Basic Research Laboratories. In 2000, he joined NTT Data Corporation. His research area is the analysis of skill acquisition in the game of go, which includes gaze analysis, gaze tracking, and protocol analysis. He is a member of the Japanese Cognitive Science Society, The Information Processing Society of Japan and the Japan Society for Science Education.