Development of a Mobile Tablet PC with Gaze-Tracking Function

Michiya Yamamoto[1], Hironobu Nakagawa[1], Koichi Egawa[1], and Takashi Nagamatsu[2]

[1] 2-1 Gakuen Sanda Hyogo 669-1137, Japan
[2] 5-1-1 Fukae-minami Higashi-nada Kobe Hyogo 658-0022, Japan
{anf73165,egawa,michiya.yamamoto}@kwansei.ac.jp,
nagamatu@kobe-u.ac.jp

Abstract. In the near future, interfaces for personal information devices with large touch screens that are capable of processing different types of information in a more intuitive manner will become indispensable. In this study, we developed "MobiGaze.PC," a system that can achieve stable gaze tracking on a mobile tablet PC. Users can interactively acquire information using both the touch screen and gaze tracking on a mobile device. First, we created the hardware setup using a tablet PC, cameras, and other apparatus. Next, we developed a method of detecting the eye area using a Purkinje image, and the position of the center of the pupil and the Purkinje image in low resolution. We then performed experiments to evaluate the accuracy of these methods. Finally, we developed a number of multimodal applications of the proposed system.

Keywords: gaze-tracking, mobile device.

1 Introduction

Various personal information devices with large touch screens, such as the Apple iPad, are now widely used. These devices allow for intuitive processing of different types of information. In the near future, such personal information devices will become more multifaceted, and hence, interfaces that can process large amounts of information more intuitively will become indispensable.

Some research has already been conducted in the field of eye tracking using mobile devices. For example, eyeLook was developed by integrating a mobile device and an eye-detection system [1]. Recently, EyePhone was developed, using a mobile phone as the concept model [2]. Further, Holland proposed a system that makes eye tracking possible on an iPad [3]. These approaches, however, do not make it possible to accurately track gaze position. Nevertheless, for precise analysis using a mobile device, one research approach utilized a head-mounted eye-tracker in conjunction with a mobile device [4], while another approach used a desktop eyetracker that was tested by mounting a mobile device on its display [5]. Recently, an option for a portable device was made commercially available by Tobii [6]. These approaches,

S. Yamamoto (Ed.): HIMI/HCII 2013, Part II, LNCS 8017, pp. 421–429, 2013.

however, are designed only for the purpose of analysis. NTT docomo in Japan have performed a demonstration of Android tablet with eye tracking function [7].

Much research on gaze interaction, especially for disabled persons, has been conducted for some years now. However, they have not found an alternative approach yet. On the other hand, the authors have focused on the importance of the integration of gaze and touch in practical interaction [8]. A few studies have developed the concept of gaze and touch interaction. For example, Kumar's system uses the gaze and keyboard input for desktop PCs [9], and Stellmach uses an iPod as an input device [10]. The authors have focused on the importance of the integration of gaze and touch in interactions, and have developed an "Eye-Tracking Pen Display," which can achieve precise eye tracking within about 1.0 ° by using an instinctive pen display on a desktop [11][12]. In addition, they have developed "MobiGaze," an interface that allows a user to interact with a personal information device through both touch and gaze [13]. To do so, however, the user must track his/her left eye precisely in order to track his/her gaze on MobiGaze, and must also carry a notebook PC for image processing, and batteries for the auxiliary devices.

In this study, we developed a tablet PC "MobiGaze.PC," an all-in-one system for interactive information acquisition using the instability of camera by using both a touch screen and high precision eye-tracking. First, we developed novel methods of realizing high-speed and precise detection of the pupil and the Purkinje image. Further, we evaluated this system, which is capable of stable eye-tracking on a tablet PC. In addition, we developed some applications.

2 MobiGaze.PC

2.1 Hardware Configuration

We fabricated a prototype of the MobiGaze.PC system, which comprises a tablet PC (HP, EliteBook 2740p, 12.1 inch) and the developed gaze estimation system with stereo cameras (PointGray, Firefly-MV03), as shown in Fig. 1.

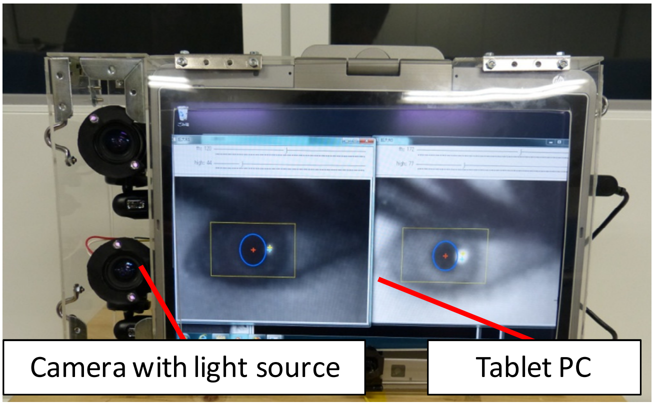

Fig. 1. MobiGaze.PC layout

The cameras must avoid the noise generated by the eyelids and eyelashes. They must also prevent any failure that could be caused by the movement of the arms and hands during touch interaction. For these reasons, we placed the cameras at the left side of the screen. The cameras could capture images with a size of 720 × 480 pixels using an 8 mm lens. With this arrangement, the gaze measurement range could cover the entire area of the screen being used. This setting is suitable for only a right-handed person, and it needs to be mirror-reversed for a left-handed person. The system uses OpenCV1.0 and Intel Integrated Performance Primitives for image processing.

2.2 High-Speed Detection of Purkinje Image

When tracking the gaze using a mobile device, the position of the eye relative to the cameras changes. To track the eye position, we generally detect the face of the user, determine the surrounding area of the eyes, and calculate the position of the pupil [14]. Recently, a number of methods for face detection using a graphics processing unit (GPU) have been introduced, but mobile devices require a high-speed method with high precision [15]. In this study, we developed a novel method that extracts the eye area from a face image using a high-brightness Purkinje image.

There are two phases in this method. First, the method binarizes the face image and detects its edges to estimate an area for extraction of a Purkinje image. The threshold value is determined using Otsu's method, which decreases the variance among classes and increases the variance within a class. Then, after noise elimination, the binarized image is fitted to ellipses. An ellipse resembling the shape of a human face is selected, and its parameters are obtained for area estimation.

Next, the method binarizes the face image to extract a Purkinje image. The threshold value is determined from the parameters of the nearest pixels. The binarized image is fitted to ellipses, and an ellipse resembling the shape of the Purkinje image is selected from an area. This process is shown in Fig. 2.

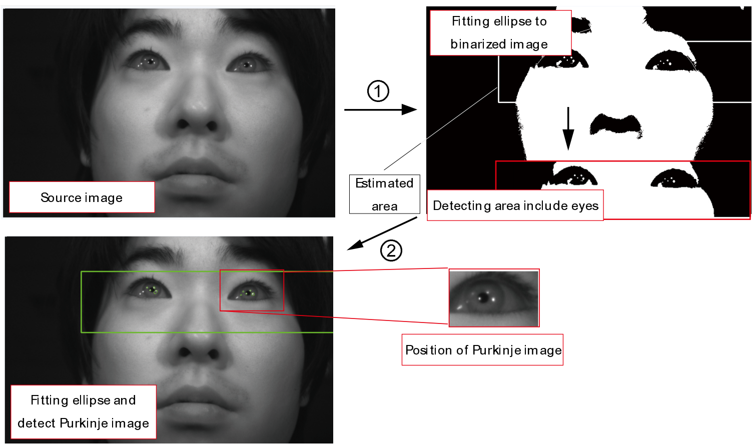

Fig. 2. Detecting position of Purkinje image

2.3 Precise Detection of Pupil and Purkinje Image in Low Resolution

When using low-resolution images, a slight error in estimating the pupil and the Purkinje image has a considerable influence on the accuracy of the gaze. Hence, it is important to detect such errors. One such method is Droege's method of image processing [16]; in this case, however, the processing speed is important. By simplifying image processing, we were able to develop a faster, more stable algorithm.

In this case, the contour was significantly affected by the high brightness of the Purkinje images, which produces strains, as seen in Fig. 3. In order to circumvent this effect arising from the Purkinje image, first, we binarized the low-resolution image and estimated the pupil contour. Next, we performed a convex closure of the contour, which is composed of points. We then determined the center of the pupil by fitting an ellipse to the points that comprise the convex closure contour. As shown in Fig. 4, we can determine the precise ellipse from the contour.

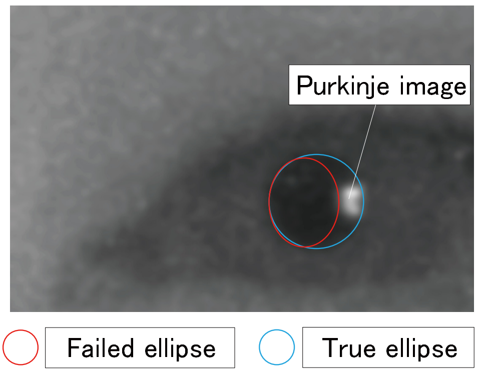

Fig. 3. Unsuccessful pupil extraction

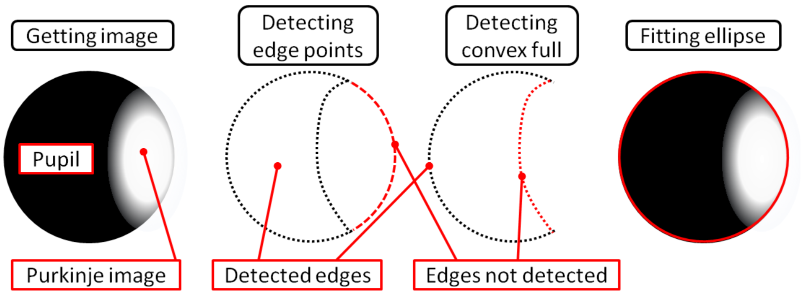

Fig. 4. Process of detecting pupil center

Next, we detected the coordinates of the Purkinje images through binarization. In this case, the Purkinje images resemble a long vertical ellipse. We split the ellipse into two parts, and assumed that the centers of these coordinates are those of the Purkinje image (Fig. 5). After this, we computed the visual axis following one-point calibration to the optical axis [17].

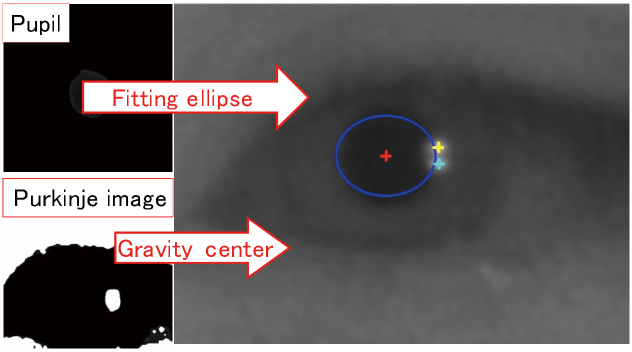

Fig. 5. Image processing for low resolution

3 Evaluation

3.1 Methodology

We then evaluated the MobiGaze.PC system. Fig. 6 shows the experimental setup. The minimum distance between the subject and the display was 30 cm, and the subject was seated in front of the display. In the experiment, to perform the calibration, we requested the user to gaze at the marker on the left side of the display. Next, we showed 12 white crosses on the display in order, and asked the user to gaze at the center of the white cross for over 50 frames. Three students participated in the experiment.

Fig.7 shows the experimental setup for a handheld device. Similar to the previous experiment, we showed 9 white crosses, requested the user to gaze as straight ahead as possible, following which the user moved the MobiGaze.PC system closer to his/her point of gaze. Furthermore, we asked the user to move the device forward and backward by approximately 10 cm. Because the device weighs approximately 2.9 kg, it was supported in order to decrease the burden on the user.

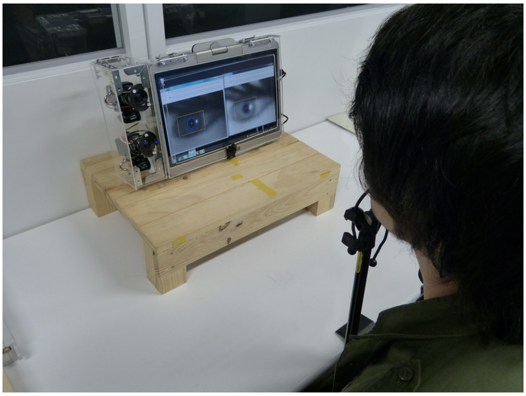

Fig. 6. Experimental setup (handheld)

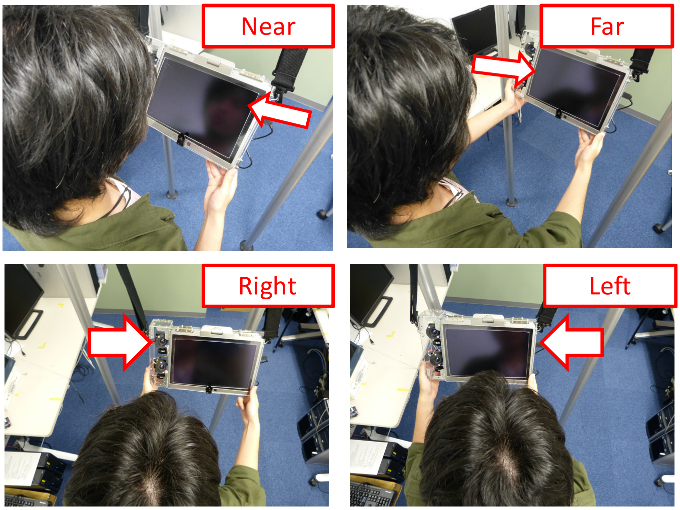

Fig. 7. Experimental setup (handheld)

3.2 Results

Fig.8 shows the results obtained with the device in a fixed position. The average accuracy was about 1.9 °. Fig. 9 shows the results obtained in the experiment with a handheld device. The average accuracy in this case was about 1.6 °. These results confirm that our method can achieve accuracy in the construction of effective mobile-device interfaces. Moreover, our method is flexible, so this system can be used in some circumstances with an editing camera setting. At some points, however, a low degree of accuracy was observed.

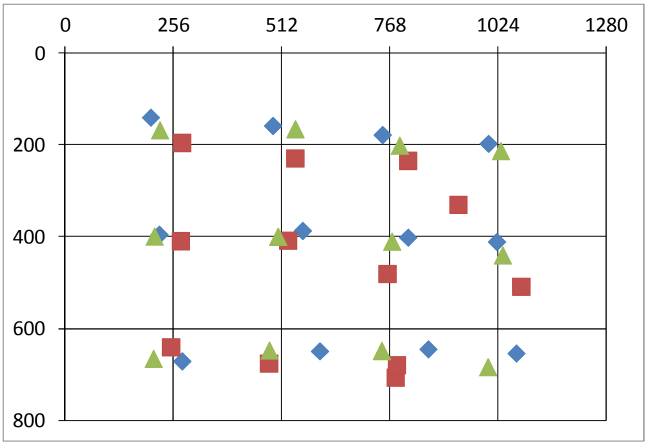

Fig. 8. Experimental results (fixed)

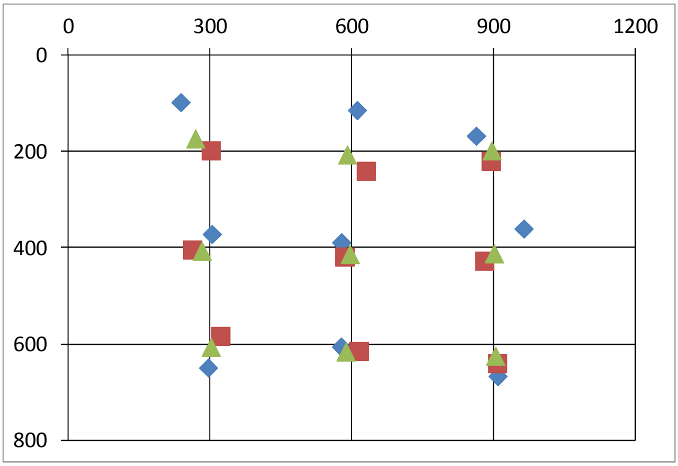

Fig. 9. Experimental results (handheld)

Fig. 10 shows the results of the transition accuracy before and after we changed the distance between the user and the device. The average accuracy was 2.4 °and 1.0 °. These results prove that our method is robust to a certain extent of change in the relative distance between the camera and the eye.

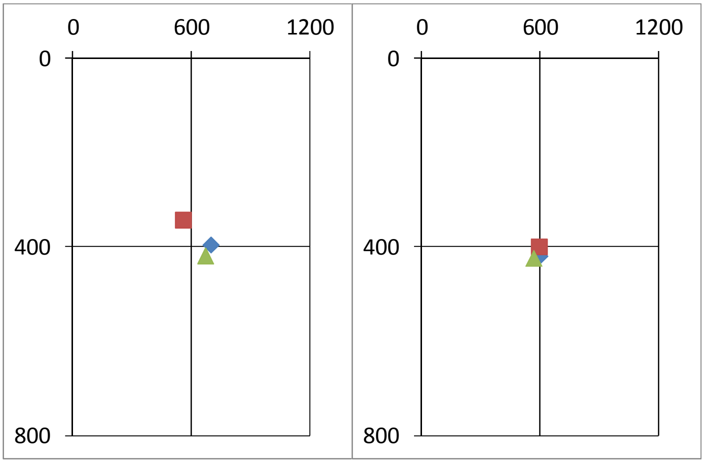

Fig. 10. Experimental results (far and near)

4 Applications

We can propose various multimodal interactions that can be performed using MobiGaze.PC. One is a gaze-and-one hand interaction that can be an alternative to and extension of both hands. For example, when we hold a large-screen mobile device in one hand, the area on which we can physically interact is limited according to the length of our fingers, the position of our hand, etc. Fig. 11 shows extensions of such typical interaction. The other novel form of interaction is direct operation of the gaze point via touch interaction, which can be realized by the all-in-one MobiGaze.PC. This can be applied to various software such as map viewers, picture editors, photo viewers, and three-dimensional maps (Fig. 12). These approaches can not only provide a solution to the Midas Touch Problem but also generate a new operational feeling based upon gaze and touch.

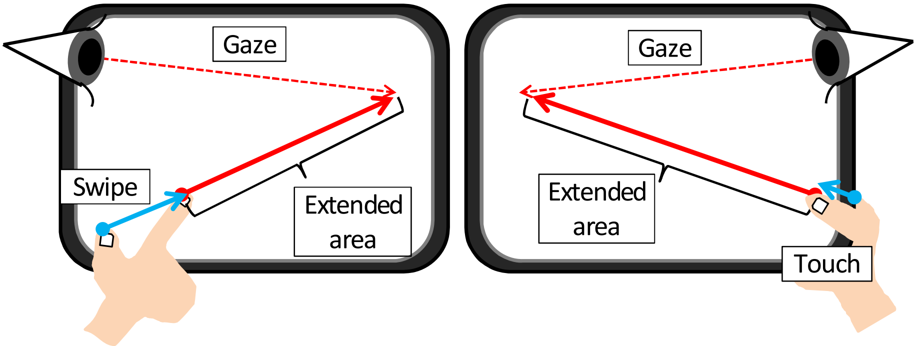

Fig. 11. Extended areas of swipe and touch

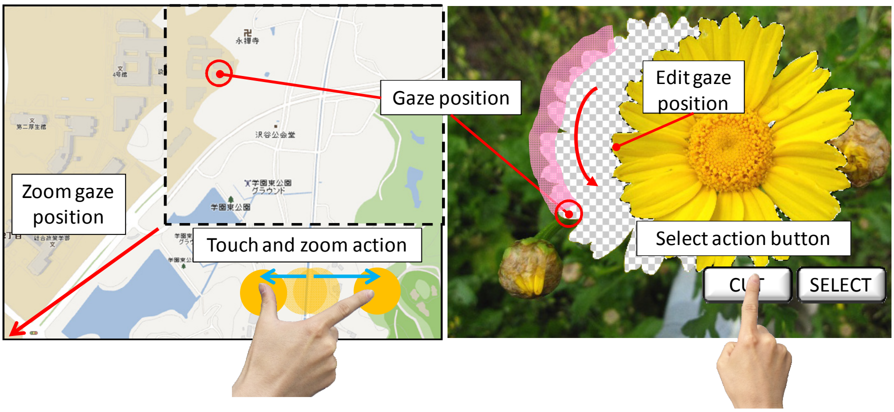

Fig. 12. Applications

5 Conclusion

In this paper, we have described a prototype of MobiGaze.PC, a mobile device that uses gaze estimation in combination with touch input. First, we developed the hardware, which comprises a tablet PC, cameras, and other auxiliary apparatus. Next, we defined a method for high-speed detection of the Purkinje image. We then developed an algorithm to detect the position of the center of the pupil and the

Purkinje image in low resolution. Finally, we performed experiments to evaluate the accuracy of these methods. Two sets of experiments were conducted: one with the device held fixed on a desk, and the other with the device held by hand. The results displayed an average accuracy of 1.9 °and 1.6 °, respectively. Finally, we proposed a number of useful gaze and touch applications for the novel proposed system.

Acknowledgements. A part of this work was supported by KAKENHI (23300047).

References

1. Dickie, C., Vertegaal, R., Sohn, C., Cheng, D.: EyeLook: Using Attention to Facilitate Mobile Media Consumption. In: Proc. UIST 2005, pp. 103–106 (2005)
2. Miluzzo, E., Wang, T., Campbel, A.T.: EyePhone: activating mobile phones with your eyes. In: Proc. MobiHeld 2010, pp.15–20 (2010)
3. Holland, C., Komogortsev, O.: Eye Tracking on Unmodified Common Tablets: Challenges andSolutions. In: Proc. ETRA 2012, pp. 277–280 (2012)
4. Cheng, S.: The Research Framework of Eye-Tracking Based Mobile Device Usability Evaluation. In: Proc. PETMEI 2011, pp. 21–26 (2011)
5. Lukander, K.: Measuring Gaze Point on Handheld Mobile Devices. In: Ext. Abstracts CHI 2004, p. 1556 (2004)
6. Tobii: Mobile Device Stand, `http://www.tobii.com/en/eyetracking-research/global/products/hardwareaccessories/tobii-mobile-device-stand/`
7. NTT DoCoMO: ibeam, `http://docomo-exhibition.jp/cj2012/pc/index.html`
8. Bulling, A., Gellersen, H.: Toward Mobile Eye-Based Human-Computer Interaction. IEEE Pervasive Computing 9(4), 8–12 (2010)
9. Kumar, M., Paepcke, A., Winograd, T.: EyePoint: practical pointing and selection using gaze and keyboard. In: Proc. CHI 2007, pp. 421–430 (2007)
10. Stellmach, S., Stober, S., Nürnberger, A., Dachselt, R.: Designing gaze-supported multimodal interactions for the exploration of large image collections. In: Proc. NGCA 2011, p. 1 (2011)
11. Yamamoto, M., Nagamatsu, T., Watanabe, T.: Development of Eye-Tracking Pen Display Based on Stereo Bright Pupil Technique. In: Proc. ETRA 2010, pp. 165–168 (2010)
12. Yamamoto, M., Sato, H., Yoshida, K., Nagamatsu, T., Watanabe, T.: Development of an eye-tracking pen display for analyzing embodied interaction. In: Proc. HI 2011, vol. 1, pp. 651–658 (2011)
13. Nagamatsu, T., Yamamoto, M., Sato, H.: MobiGaze: development of a gaze interface for handheld mobile devices. In: Proc. CHI EA 2010, pp. 3349–3354 (2010)
14. Yamazoe, H., Utsumi, A., Yonezawa, T., Abe, S.: Remote gaze estimation with a single camera based on facial-feature tracking without special calibration actions. In: Proc. ETRA 2008, pp. 245–250 (2008)
15. Lozano, O.M., Otsuka, K.: Real-time visual tracker by stream processing. Journal of Signal Processing Systems 57(2) (2009)
16. Droege, D., Paulus, D.: Pupil Center Detection in Low Resolution Images. In: Proc. ETRA 2010, pp. 169–172 (2010)
17. Nagamatsu, T., Iwamoto, Y., Sato, H.: Gaze Estimation Method based on an Aspherical Model ofthe Cornea: Surface of Revolution about Optical Axis of the Eye. In: Proc. ETRA 2010, pp. 255–258 (2010)