



One-point Calibration Gaze Tracking Based on Eyeball Kinematics Using Stereo Cameras

Takashi Nagamatsu*
Kobe University

Junzo Kamahara†
Kobe University

Takumi Iko‡
Kobe University

Naoki Tanaka§
Kobe University

Abstract

This paper presents a one-point calibration gaze tracking method based on eyeball kinematics using stereo cameras. By using two cameras and two light sources, the optic axis of the eye can be estimated. One-point calibration is required to estimate the angle of the visual axis from the optic axis. The eyeball rotates with optic and visual axes based on the eyeball kinematics (Listing's law). Therefore, we introduced eyeball kinematics to the one-point calibration process in order to properly estimate the visual axis. The prototype system was developed and it was found that the accuracy was under 1° around the center and bottom of the display.

CR Categories: H.5.2 [Information Interfaces and Presentation]: User Interfaces—Ergonomics; I.4.9 [Image Processing and Computer Vision]: Applications

Keywords: eye tracking, calibration, stereo camera, eyeball kinematics

1 Introduction

Gaze tracking technology is being used as a human-machine interface [Jacob 1991; Duchowski 2007; Nagamatsu et al. 2007]. However, most accurate gaze tracking systems require personal calibration process before the system can be used. Several works have attempted to reduce the calibration effort by using a model-based approach [Ohno and Mukawa 2004; Shih and Liu 2004; Ohno 2006; Guestrin and Eizenman 2006; Guestrin and Eizenman 2007; Villanueva and Cabeza 2007]. For example, Shih and Liu [2004], and Guestrin and Eizenman [2006; 2007] have presented methods to reconstruct the optic axis of the eye using stereo cameras without actually knowing the personal eye parameters. Since the optic axis is an approximation of the visual axis with an accuracy of approximately 5° , after the reconstruction of the optic axis, the visual axis is estimated by at least a one-point calibration procedure in their work.

However, most previous works did not consider the eyeball kinematics when estimating the visual axis; in contrast, Villanueva and Cabeza [2007] pointed out the importance of eyeball kinematics. In this paper, we describe one-point calibration gaze tracking based on eyeball kinematics using stereo cameras.

*e-mail: nagamatsu@kobe-u.ac.jp

†e-mail: kamahara@maritime.kobe-u.ac.jp

‡e-mail: 071w101w@stu.kobe-u.ac.jp

§e-mail: ntanaka@maritime.kobe-u.ac.jp

Copyright © 2008 by the Association for Computing Machinery, Inc. Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, to republish, to post on servers, or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from Permissions Dept, ACM Inc., fax +1 (212) 869-0481 or e-mail permissions@acm.org.

ETRA 2008, Savannah, Georgia, March 26–28, 2008.

© 2008 ACM 978-1-59593-982-1/08/0003 \$5.00

2 Estimation of optic axis

In this section, we describe a method for the estimation of the optic axis. The cameras are modeled as pinhole cameras and the light sources are modeled as point sources. The position of the light sources must be measured and the intrinsic and extrinsic camera parameters must be determined in advance by the camera calibration.

2.1 Image processing of eye image

The eye images are captured by two cameras. From the images captured by camera 0, the center of the pupil, B'_0 , is detected by ellipse fitting, and the centers of the first Purkinje images (the reflection of light sources (L_0 and L_1) from the outer surface of the cornea), P'_{00} and P'_{01} , are detected by seeking near the center of the pupil. Similarly, from the image captured by camera 1, B'_1 , P'_{10} , and P'_{11} are detected, where B'_1 is the center of the pupil and P'_{10} and P'_{11} are the centers of the first Purkinje images of L_0 and L_1 captured by camera 1.

2.2 Center of corneal curvature

We estimate the position of the center of the corneal curvature, A . Figure 1 shows a cross section of the eyeball that includes A , light source 0, L_0 , the nodal point of camera 0, C_0 . L_0 and C_0 are known. A ray from L_0 reflects at a point P_{00} on the corneal surface such that the reflected ray passes through C_0 and intersects the camera image plane at a point P'_{00} . C_0 , P_{00} , L_0 , P'_{00} , and A are coplanar, and the normal vector of the plane, π_{00} , is $(P'_{00} - C_0) \times (L_0 - C_0)$. The plane can be expressed as

$$(P'_{00} - C_0) \times (L_0 - C_0) \cdot (x - C_0) = 0 \quad (1)$$

where $x = (x, y, z)$ is the point on the plane.

Similarly, with regard to the relation among the cameras, C_0 and C_1 , the light sources, L_0 and L_1 , and the Purkinje images, P'_{01} , P'_{10} , and P'_{11} , three other planes, π_{01} , π_{10} , and π_{11} can be expressed respectively as

$$(P'_{01} - C_0) \times (L_1 - C_0) \cdot (x - C_0) = 0 \quad (2)$$

$$(P'_{10} - C_1) \times (L_0 - C_1) \cdot (x - C_1) = 0 \quad (3)$$

$$(P'_{11} - C_1) \times (L_1 - C_1) \cdot (x - C_1) = 0. \quad (4)$$

All planes include A , and A can be known if three planes are given. Two LEDs were installed to make an angle of 90° around a camera in order to improve its accuracy.

2.3 Center of pupil

Radius of cornea In order to estimate the radius of the cornea, we will estimate the reflected point, P_{00} , that is on the middle line between L_0A and C_0A , as shown in figure 1. The line AP_{00} can be expressed in a parametric form as

$$x = A + t \left(\frac{L_0 - A}{|L_0 - A|} + \frac{C_0 - A}{|C_0 - A|} \right) \quad (5)$$

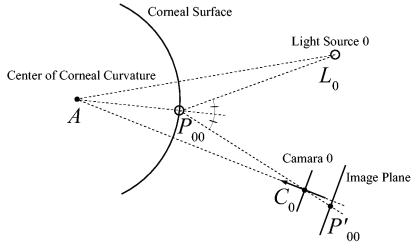


Figure 1: Center of the Corneal Curvature.

where t is a parameter.

On the other hand, the line from C_0 to P_{00} can be expressed in parametric form as

$$\mathbf{x} = \mathbf{C}_0 + t(\mathbf{C}_0 - \mathbf{P}'_{00}). \quad (6)$$

The position of P_{00} can be estimated as the intersection point of the above two lines. Therefore, the radius of the cornea, r , is determined as

$$r = |\mathbf{P}_{00} - \mathbf{A}|. \quad (7)$$

Since r can be determined using three other combinations $(\mathbf{C}_0, \mathbf{L}_1, \mathbf{P}'_{01})$, $(\mathbf{C}_1, \mathbf{L}_0, \mathbf{P}'_{10})$, and $(\mathbf{C}_1, \mathbf{L}_1, \mathbf{P}'_{11})$, the average is used for improving the robustness.

Refraction As shown in figure 2, a ray that originates from the center of the pupil, B , refracts at the point B''_0 , passes through the nodal point of camera 0, C_0 , and intersects the camera image plane at a point B'_0 . B''_0 can be determined to solve the equations given below.

$$\mathbf{x} = \mathbf{C}_0 + t(\mathbf{C}_0 - \mathbf{B}'_0) \quad (8)$$

$$r = |\mathbf{x} - \mathbf{A}| \quad (9)$$

t is solved as follows:

$$t = \frac{-b - \sqrt{b^2 - ac}}{a} \quad (10)$$

where $a = (C_{0x} - B'_{0x})^2 + (C_{0y} - B'_{0y})^2 + (C_{0z} - B'_{0z})^2$, $b = (C_{0x} - A_{0x})(C_{0x} - B'_{0x}) + (C_{0y} - A_{0y})(C_{0y} - B'_{0y}) + (C_{0z} - A_{0z})(C_{0z} - B'_{0z})$, and $c = (C_{0x} - A_{0x})^2 + (C_{0y} - A_{0y})^2 + (C_{0z} - A_{0z})^2$.

The refracted vector at B''_0 , $\mathbf{t}_0$, can be calculated by using Snell's law as follows:

$$\mathbf{t}_0 = \left(-\rho \mathbf{n}_0 \cdot \mathbf{v}_0 - \sqrt{1 - \rho^2 (1 - (\mathbf{n}_0 \cdot \mathbf{v}_0)^2)} \right) \mathbf{n}_0 + \rho \mathbf{v}_0 \quad (11)$$

where the incident vector $\mathbf{v}_0 = (\mathbf{C}_0 - \mathbf{B}'_0) / |\mathbf{C}_0 - \mathbf{B}'_0|$, normal vector at the point of refraction $\mathbf{n}_0 = (\mathbf{B}''_0 - \mathbf{A}) / |\mathbf{B}''_0 - \mathbf{A}|$, and $\rho = n_1/n_2$ (n_1 : the index of air ≈ 1 ; n_2 : the effective refractive index ≈ 1.3375).

Center of pupil The center of the pupil, B , can be obtained from the intersection of two rays from the two cameras as follows:

$$\mathbf{x} = \mathbf{B}''_0 + t\mathbf{t}_0 \quad (12)$$

$$\mathbf{x} = \mathbf{B}''_1 + t\mathbf{t}_1. \quad (13)$$

where $\mathbf{B}''_1$ and $\mathbf{t}_1$ are the refracted point and the refracted vector, respectively (using camera 1).

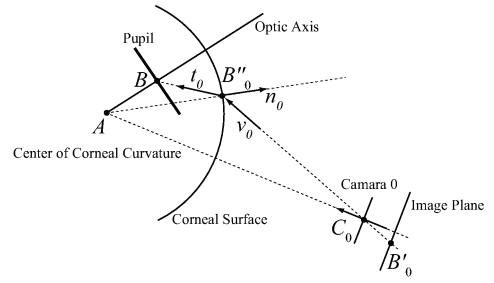


Figure 2: Refraction.

2.4 Optic Axis

The optic axis can be written in parametric form as follows:

$$\mathbf{x} = \mathbf{A} + t(\mathbf{B} - \mathbf{A}). \quad (14)$$

3 One-Point Calibration

This section describes the calibration process that determines a unit direction vector of an optic axis, $\mathbf{b}$, at the primary position, which is the position of the eye relative to the head when looking straight ahead at an object at eye level. In this paper, the unit direction vector of the visual axis at the primary position is $(0, 0, -1)^T (= \mathbf{a})$.

Figure 3 shows the eyeball model of rotation and translation. $\mathbf{E}_0$ and $\mathbf{A}_0$ are the rotation center of the eyeball and the center of the corneal curvature at the primary position, respectively. $\mathbf{E}$, $\mathbf{d}$, and $\mathbf{c}$ are the rotation center of the eyeball, a unit direction vector of the optic axis, and, a unit direction vector of the visual axis after the eye movement, respectively.

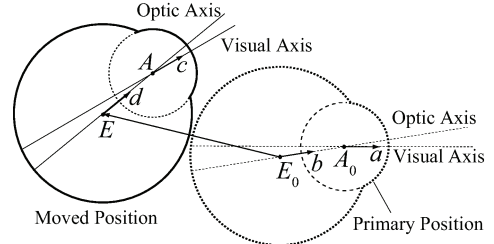


Figure 3: Model of an eyeball.

3.1 Movement of eyeball

The visual axis at the primary position can be expressed as $\mathbf{x} = \mathbf{A}_0 + t\mathbf{a}$, and the new visual axis is expressed as $\mathbf{x} = \mathbf{A} + t\mathbf{c}$. If the eye moves, the visual axis rotates around $\mathbf{E}_0$ by the rotation matrix, $\mathbf{R}$, and $\mathbf{E}_0$ is translated to $\mathbf{E}$ by the translation matrix, $\mathbf{T}$. The relation can be written as

$$\mathbf{E} + ((\mathbf{A} - \mathbf{E}) + t\mathbf{c}) = \mathbf{T}\mathbf{E}_0 + \mathbf{R}((\mathbf{A}_0 - \mathbf{E}_0) + t\mathbf{a}) \quad (15)$$

Since $\mathbf{E} = \mathbf{T}\mathbf{E}_0$ and $(\mathbf{A} - \mathbf{E}) = \mathbf{R}(\mathbf{A}_0 - \mathbf{E}_0)$, the relation between $\mathbf{a}$ and $\mathbf{c}$ can be written as

$$\mathbf{c} = \mathbf{R}\mathbf{a}. \quad (16)$$

Therefore, in order to analyze the movement of the visual axis, we only have to analyze the movement of the direction vector of the visual axis.

3.2 Listing's law

Listing's law states that any eye position can be reached by a single rotation from the primary position and any rotation axis lies in a plane (Listing's plane) [Zatsiorsky 1999]. In this paper, the rotation axis of the eyeball, $\mathbf{l}$, is assumed to be parallel to the xy plane. We will calculate the eyeball rotation based on Listing's law.

3.3 Visual axis

$\mathbf{D}$ is assumed to be a center on the display. When a user is gazing at $\mathbf{D}$, the vector of the visual axis can be written as $\mathbf{D} - \mathbf{A}$. Figure 4 shows the relation among $\mathbf{a}$, $(\mathbf{D} - \mathbf{A})/|\mathbf{D} - \mathbf{A}|$, and $\mathbf{l}$. The rotation from $\mathbf{a}$ to $(\mathbf{D} - \mathbf{A})/|\mathbf{D} - \mathbf{A}|$ by ψ can be written as a product of three rotation matrices as

$$\frac{\mathbf{D} - \mathbf{A}}{|\mathbf{D} - \mathbf{A}|} = \mathbf{R}(\mathbf{z}, \phi) \mathbf{R}(\mathbf{x}, \psi) \mathbf{R}(\mathbf{z}, -\phi) \mathbf{a} \quad (17)$$

$$= \mathbf{M} \mathbf{a} \quad (18)$$

where $\mathbf{R}(\mathbf{z}, \phi)$ is the rotation matrix around the z-axis by ϕ , $\mathbf{R}(\mathbf{x}, \psi)$ is the rotation matrix around the x-axis by ψ , $\mathbf{R}(\mathbf{z}, -\phi)$ is the rotation matrix around the z-axis by $-\phi$, and $\mathbf{M}$ is a combined matrix of all the matrices.

The unit vector of the axis of rotation, $\mathbf{l}$, and the angle, ψ , can be written as

$$\mathbf{l} = \frac{\mathbf{a} \times (\mathbf{D} - \mathbf{A})}{|\mathbf{a} \times (\mathbf{D} - \mathbf{A})|} \quad (19)$$

$$\psi = \arccos \left(\frac{\mathbf{a} \cdot (\mathbf{D} - \mathbf{A})}{|\mathbf{a}| |\mathbf{D} - \mathbf{A}|} \right). \quad (20)$$

Since ϕ is made by $\mathbf{l}$ and $(1, 0, 0)^T (= \mathbf{e})$, it can be obtained as follows:

$$\phi = \arccos \left(\frac{\mathbf{e} \cdot \mathbf{l}}{|\mathbf{e}| |\mathbf{l}|} \right). \quad (21)$$

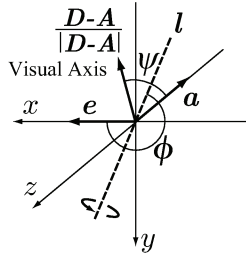


Figure 4: Rotation about the visual axis.

3.4 Estimation of visual axis at primary position

If $\mathbf{b}$ is the optic axis at the primary position, then the relation between $(\mathbf{B} - \mathbf{A})$ and $\mathbf{b}$ can be written using $\mathbf{M}$ as follows:

$$\frac{\mathbf{B} - \mathbf{A}}{|\mathbf{B} - \mathbf{A}|} = \mathbf{M} \mathbf{b}. \quad (22)$$

$\mathbf{b}$ can be obtained by using ϕ and ψ as

$$\mathbf{b} = \mathbf{M}^{-1} \frac{\mathbf{B} - \mathbf{A}}{|\mathbf{B} - \mathbf{A}|} \quad (23)$$

$$\begin{aligned} &= (\mathbf{R}(\mathbf{z}, \phi) \mathbf{R}(\mathbf{x}, \psi) \mathbf{R}(\mathbf{z}, -\phi))^{-1} \frac{\mathbf{B} - \mathbf{A}}{|\mathbf{B} - \mathbf{A}|} \\ &= \mathbf{R}(\mathbf{z}, \phi) \mathbf{R}(\mathbf{x}, -\psi) \mathbf{R}(\mathbf{z}, -\phi) \frac{\mathbf{B} - \mathbf{A}}{|\mathbf{B} - \mathbf{A}|}. \end{aligned} \quad (24)$$

4 Estimation of visual axis

In this section, a method to estimate the visual axis using $\mathbf{b}$, which was determined by calibration, is described.

4.1 Calculation of rotation axis and angle

$\mathbf{d}$ is a unit vector of the optic axis and it is given as $\mathbf{d} = (\mathbf{B} - \mathbf{A})/|\mathbf{B} - \mathbf{A}|$. $\mathbf{l}'$ is a unit direction vector of the rotation axis of the eye, which is parallel to xy plane. Figure 5 shows the relation among $\mathbf{b}$, $\mathbf{d}$, and $\mathbf{l}'$. $\mathbf{l}'$ is included in the xy plane $\mathbf{a} \cdot \mathbf{x} = 0$ in the figure. Since $\mathbf{b}$ rotates to $\mathbf{d}$ around $\mathbf{l}'$, $\mathbf{l}'$ is also included in the plane $(\mathbf{d} - \mathbf{b}) \cdot \mathbf{x} = 0$. Therefore the unit direction vector of the axis is given as

$$\mathbf{l}' = \frac{\mathbf{a} \times (\mathbf{d} - \mathbf{b})}{|\mathbf{a} \times (\mathbf{d} - \mathbf{b})|}. \quad (25)$$

The rotation axis can be expressed as $\mathbf{x} = t\mathbf{l}'$. The intersection point between $\mathbf{x} = t\mathbf{l}'$ and the plane perpendicular to $\mathbf{x} = t\mathbf{l}'$ including $(\mathbf{d} - \mathbf{b})$ must be determined in order to estimate ψ . t can be determined by the inner product of $\mathbf{l}'$ and $\mathbf{b}$, which is given as $t = \mathbf{l}' \cdot \mathbf{b}$. Therefore the center of rotation of $\mathbf{b}$ is $(\mathbf{l}' \cdot \mathbf{b})\mathbf{l}'$. The rotation angle of the eyeball is estimated as

$$\psi = \arccos \left(\frac{((\mathbf{l}' \cdot \mathbf{b})\mathbf{l}' - \mathbf{b}) \cdot ((\mathbf{l}' \cdot \mathbf{b})\mathbf{l}' - \mathbf{d})}{|((\mathbf{l}' \cdot \mathbf{b})\mathbf{l}' - \mathbf{b})| |((\mathbf{l}' \cdot \mathbf{b})\mathbf{l}' - \mathbf{d})|} \right). \quad (26)$$

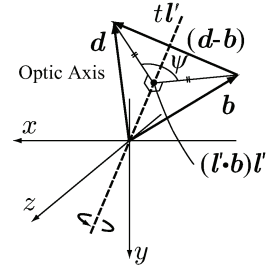


Figure 5: Rotation about the optic axis.

4.2 Calculation of visual axis

The visual axis can be calculated by rotating $\mathbf{a}$ around $\mathbf{l}'$ by an angle ψ . A quaternion is convenient to express the rotation of a point around a certain axis. The elements of $\mathbf{a}$ and $\mathbf{l}'$ are written as $\mathbf{a} = (a_x, a_y, a_z)$ and $\mathbf{l}' = (l'_x, l'_y, l'_z)$, respectively.

$\mathbf{a}$ is expressed in quaternion form as $\mathbf{P}_a = (0; a_x, a_y, a_z)$ and the rotation is expressed as $\mathbf{Q} = (\cos(\psi/2); l'_x \sin(\psi/2), l'_y \sin(\psi/2), l'_z \sin(\psi/2))$. Then, the visual axis vector, $\mathbf{P}_e$, in quaternion form can be obtained as follows:

$$\mathbf{P}_e = (0; c_x, c_y, c_z) = \mathbf{Q} \mathbf{P}_a \mathbf{Q}^{-1}. \quad (27)$$

Consequently we can obtain the vector of the visual axis, $\mathbf{c} = (c_x, c_y, c_z)$.

4.3 Estimating gazing position on display

The display can be expressed by four points: D_{TL} , D_{TR} , D_{BL} , and D_{BR} . The normal vector of the display, n_d , is $(D_{TL} - D_{BR}) \times (D_{TR} - D_{BL})$; therefore, the display can be expressed as $n_d \cdot (x - D_{TL}) = 0$. Since the visual axis is written as $x = A + tc$, the intersection point is the point of gazing (POG) on the display. After estimating the position of the intersection point, it is converted to graphical coordinates on the screen.

5 Implementation and experimental results

5.1 Implementation

A prototype system was implemented, as shown in figure 6. This system uses two synchronized monochrome IEEE-1394 digital cameras using a 1/3" CMOS image sensor (Firefly MV, Point Grey Research Inc.) with 50-mm lens and an IR filter, two infrared light sources attached to a 17" LCD, and a Windows-platform-based PC. Software was developed using OpenCV[Intel]. If a user clicks on the mouse, a fixation point is displayed. When the user gazing at the fixation point releases the mouse button, the calibration is performed.

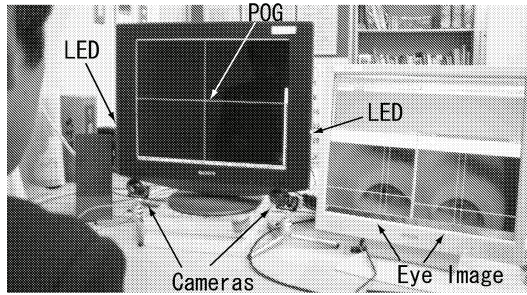


Figure 6: System.

5.2 Experimental results

The prototype system was evaluated experimentally with one adult subject. The subject was asked to fixate on 25 points on the display. The area within which the user could move was 3 cm laterally, 2.5 cm vertically, and 5 cm backward/forward, and was approximately 57.5 cm from the display. Figure 7 shows the result of the experiment. It was found that the accuracy was under 1°, except that around the top left and top right, the measurement had a low accuracy or was impossible to determine. This may be because the penumbra of the cornea is not like a sphere or an LED that reflects on the sclera. This may be attributed to the limitation of the two-light-source and two-camera arrangement. The next step will be the improvement of the arrangement of light sources and cameras in order to observe Purkinje images on the corneal surface near the optic axis.

6 Conclusion

This paper presented a one-point calibration gaze tracking method based on eyeball kinematics using stereo cameras. A prototype system was developed based on this method and the accuracy was under 1° around the center and bottom of the display.

Acknowledgements

This research was partially supported by the Ministry of Education, Culture, Sports, Science and Technology, Grant-

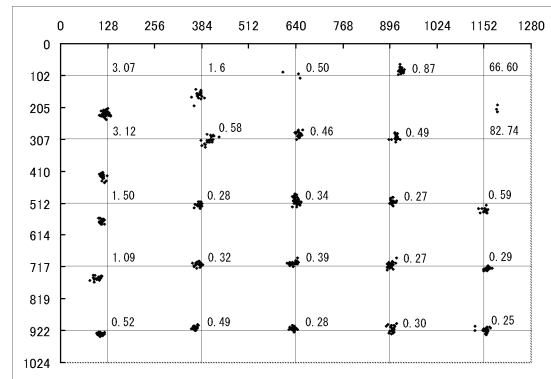


Figure 7: Experimental results: the dotted points indicate the estimates of the POG, the 25 grid points represent the intended fixation points, and the numbers represent errors in view angle (degree).

in-Aid for Young Scientists (B), 18700645, 2006.

References

- DUCHOWSKI, A. T. 2007. *Eye Tracking Methodology: Theory and Practice*, 2nd ed. Springer-Verlag.
- GUESTIN, E. D., AND EIZENMAN, M. 2006. General theory of remote gaze estimation using the pupil center and corneal reflections. *IEEE Transactions on Biomedical Engineering* 53, 6, 1124–1133.
- GUESTIN, E. D., AND EIZENMAN, M. 2007. Remote point-of-gaze estimation with free head movements requiring a single-point calibration. In *Proceedings of the 29th Annual International Conference of the IEEE EMBS*, 4556–4560.
- INTEL. Open source computer vision library. <http://www.intel.com/technology/computing/opencv/index.htm>.
- JACOB, R. J. K. 1991. The use of eye movements in human-computer interaction techniques: what you look at is what you get. *ACM Transactions on Information Systems* 9, 2, 152–169.
- NAGAMATSU, T., KAIEDA, Y., KAMAHARA, J., AND SHIMADA, H. 2007. Development of a skill acquisition support system using expert's eye movement. In *Proceedings of HCI International 2007*, vol. 9, 430–439.
- OHNO, T., AND MUKAWA, N. 2004. A free-head, simple calibration, gaze tracking system that enables gaze-based interaction. In *Proceedings of the 2004 symposium on Eye tracking research & applications*, 115–122.
- OHNO, T. 2006. One-point calibration gaze tracking method. In *Proceedings of the 2006 symposium on Eye tracking research & applications*, 34–34.
- SHIH, S.-W., AND LIU, J. 2004. A novel approach to 3-d gaze tracking using stereo cameras. *IEEE Transactions on Systems, Man, and Cybernetics, Part B* 34, 1, 234–245.
- VILLANUEVA, A., AND CABEZA, R. 2007. Models for gaze tracking systems. *EURASIP Journal on Image and Video Processing* 2007, Article ID 23570.
- ZATSIORSKY, V. M. 1999. *Kinematics of Human Motion (Japanese Edition)*. NAP Limited.