Automatic User Calibration for Gaze-Tracking Systems by Looking into the Distance

Takashi Nagamatsu
Kobe University

Tatsuhiko Ueki
Kobe University

Junzo Kamahara
Kobe University

Commercially available gaze tracking system can be used after user calibration in which the user gazes at several points. Several studies proposed one-point calibration method, that estimates the optical axis of the eye without any user calibration procedure and decide the offset between the optical and visual axes of the eye using a single calibration point. To implement gaze tracking without user calibration, the offsets between the optical and visual axes of the eyes must be estimated. More advanced studies propose automatic user calibration methods where the user gazes at a computer display, i.e., the user do not have to gaze at a known point but gazes at unknown points on the display for a while. In this paper, we propose an automatic user-calibration method where the user looks into the distance, which is suitable for drivers of cars, trains, and ships, who spend much of the time looking into the distance. We calculate the offsets between the optical and visual axes of the eyes based on the fact that the visual axes of both eyes are always parallel when the user looks afar, irrespective of the direction of the gaze. We evaluated the proposed method by computer simulation and a user experiment using an experimental system.

Keywords: Gaze tracking, user calibration, personal calibration, eye model

Introduction

Gaze-tracking technology is widely used in a variety of areas, such as psychological experiments (analysis of driver behavior, advertisements, etc.) and user interfaces (e.g. computer input). However, personal calibration is one of the biggest problems in the use of gaze trackers (Hansen & Ji, 2010).

Commercially available gaze tracking system can be used after personal calibration in which the user gazes at several points. Several studies (Guestrin & Eizenman, 2006; Nagamatsu, Kamahara, & Tanaka, 2008; Shih & Liu, 2004; Villanueva & Cabeza, 2008) proposed one-point calibration method, that estimates the optical axis of the eye without any personal calibration procedure and decide the offset between the optical and visual axes of the eye using a single calibration point. To implement gaze tracking without user calibration, the offsets between the optical and visual axes of the eyes are estimated. More advanced studies (Model & Eizenman, 2010; Nagamatsu, Kamahara, & Tanaka, 2009) propose automatic user-calibration methods where the user gazes at a computer display, i.e., the user do not have to gaze at a known point but gazes at unknown points on the display . The offsets between the optical and visual axes of the eyes are estimated automatically while the user looking at the display.

In this paper, we propose an automatic user-calibration method for gaze trackers that simply requires the user to look into the distance. This gaze tracker is suitable for driving situations, such as in a car, a train, or a ship, in which the driver spends most of their time looking into the distance. If we constantly measure the gaze of the driver, we can use the information in active safety systems, such as alarms or vehicle stopping devices.

In this stage of research, the prerequisite of our proposed method is that the user always looks into the distance. However, we discuss how to detect the user looks distant or nearby objects.

Related Works

Existing studies (Guestrin & Eizenman, 2006; Nagamatsu et al., 2008; Shih & Liu, 2004; Villanueva & Cabeza, 2008) constructed 3D gaze-tracking systems with one-point user calibration. In these studies, the optical axis of the eye is estimated without user calibration by using two calibrated cameras and two light sources.

Based on their method, other studies have been conducted on gaze-tracking systems that do not require intentional user participation for calibration. Nagamatsu et al. reported a gaze-tracking system that calculates the point-of-gaze (POG) as the midpoint of the optical axes of both eyes when the user is gazing at the display (Nagamatsu et al., 2009). Model and Eizenman reported an automatic personal-calibration technique using the optical axes of the eyes (Model & Eizenman, 2010). Their technique calculates the offsets between the optical and visual axes of both eyes by minimizing the distance between the intersections of the estimated visual axes of both eyes with the display when the user is gazing at the display. These studies treat the situation of the user looking at a display. That is, their methods are supposed to be used in situations where the user is gazing at nearby objects, whose shapes and positions are known.

On the other hand, drivers look into the distance most of the time, although they sometimes look at the rearview mirror or at nearby objects. Therefore, it is more appropriate for the system to be calibrated when the user is looking into the distance. In this paper, we focus on such a calibration method.

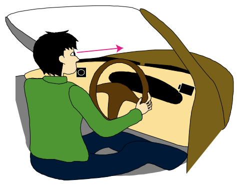
On-board Gaze Tracking of Drivers

Gaze tracking is needed while driving for safety reasons, and conventional gaze-tracking systems require some user calibration. Although some studies proposed systems that require user calibration only once (Guestrin & Eizenman, 2006; Nagamatsu et al., 2008; Shih & Liu, 2004; Villanueva & Cabeza, 2008), these systems still require an initial user calibration and user authentication.

Figure 1 (a) shows the situation of a driver looking into the distance while driving a car. Not everyone owns a car, and cars do not always have the same driver. An on-board gaze-tracking system must therefore be calibrated every time the driver changes, e.g., if the car is driven by a family member or friend.

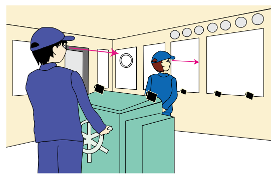
Figure 1 (b) shows the situation of a train operator. The train operator changes after each shift.

On a ship, several people could be on the bridge, as shown in Figure 1 (c). The captain or navigation officers move around; therefore, the user who must be tracked could change. This is a very difficult situation for gaze tracking.



(a) car

(b) train



(c) ship

Figure 1. Gaze tracking on a car, a train, and a ship.

Eye Rotation Based on Listing's Law

Before discussing the proposed method, we explain the calculation of the rotation of the eye based on Listing's law, which is the basis of our study. Listing's law (Wong, 2004) states that there exists an eye position, called the primary position, from which any eye position can be reached by a single rotation, and all rotation axes lie in a single plane (Listing's plane). A calculation method that applies Listing's law to a gaze-tracking system is described in (Nagamatsu et al., 2008).

Figure 2 shows the optical and visual axes of the eye when it rotates from the primary position. The unit direction vectors of the optical and visual axes of the eye at the primary position are denoted by **b** and **a**, respectively. The offset between the optical and visual axes of one eye is expressed by the parameters α (horizontal) and β (vertical), so **b** is described as follows:

$$\mathbf{b} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\beta & -\sin\alpha \\ 0 & \sin\beta & \cos\beta \end{pmatrix} \begin{pmatrix} \cos\alpha & 0 & \sin\alpha \\ 0 & 1 & 0 \\ -\sin\alpha & 0 & \cos\alpha \end{pmatrix} \mathbf{a}. \qquad (1)$$

The unit direction vectors of the optical and visual axes of the eye after the eye movement are denoted by **d** and **c**, respectively. **A** is the center of corneal curvature after the eye movement.

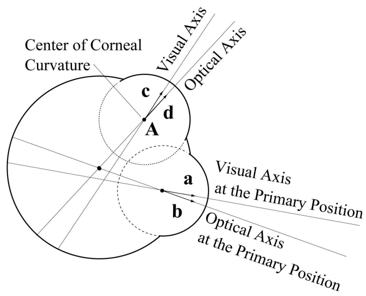


Figure 2. Model of eyeball rotation.

Figure 3 expresses the relations among **a**, **b**, **c**, and **d**, whose origins are moved to the point **O**. If **l** is a unit direction vector along the rotation axis of the eye, Listing's law is expressed as follows:

$$\mathbf{a} \cdot \mathbf{l} = \mathbf{c} \cdot \mathbf{l} = 0 \quad (2)$$

If **a**, **b**, and **d** are given, then **l** and the rotation angle of the eyeball, ψ, are estimated by the method described in (Nagamatsu et al., 2008) as

$$\mathbf{l} = \frac{\mathbf{a} \times (\mathbf{d} - \mathbf{b})}{\|\mathbf{a} \times (\mathbf{d} - \mathbf{b})\|}, \quad (3)$$

$$\psi = \arccos\left(\frac{(\mathbf{b} - (\mathbf{l} \cdot \mathbf{b})\mathbf{l}) \cdot (\mathbf{d} - (\mathbf{l} \cdot \mathbf{b})\mathbf{l})}{\|\mathbf{b} - (\mathbf{l} \cdot \mathbf{b})\mathbf{l}\| \|\mathbf{d} - (\mathbf{l} \cdot \mathbf{b})\mathbf{l}\|} \right), \quad (4)$$

respectively.

Therefore, **c** can be calculated by rotating **a**, as

$$\mathbf{c} = \mathbf{R}(\psi, \mathbf{l})\mathbf{a}, \quad (5)$$

where $\mathbf{R}(\psi, \mathbf{l})$ is the matrix of rotation around **l** by an angle ψ.

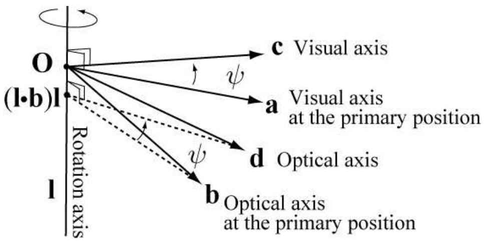


Figure 3. Rotation of optical and visual axes of the eye.

User Calibration

User calibration (also known as personal calibration) is a process that estimates user-dependent parameters. In our method, these parameters are the offset between the optical and visual axes of the eye, α and β. Therefore, calibration is a process that estimates the relation between the unit direction vectors **a** and **b**.

The visual axis of the eye at the primary position, **a**, is approximately relative to the head as the direction of the face. The optical axis of the eye, **d**, is estimated using two calibrated cameras, two light sources, and the method described in (Nagamatsu et al., 2008). Therefore, **a** and **d** are known.

In order to determine α and β, we must determine **b** for each eye; therefore, we add an additional condition for **c** that requires the user to look into the distance. Then, we determine **b** by using the relations among **a**, **b**, **c**, and **d**.

Relation between the Optical and Visual Axes of Both Eyes When the User Looks into the Distance

Figure 4 depicts the relation between the optical and visual axes of both eyes when the user looks into the distance. The key point is that the visual axes of the eyes are parallel when the user looks into the distance (infinity). Here, the offsets for both eyes are expressed by four parameters: α_L, β_L, α_R, and β_R.

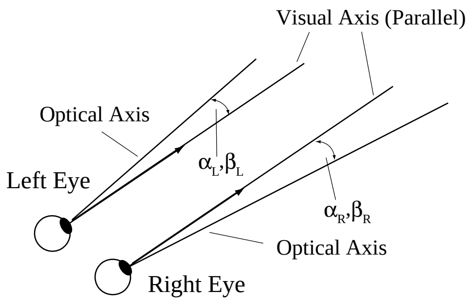


Figure 4. Relations among the optical and visual axes of both the eyes when the user looks into the distance.

Automatic User Calibration While the User Is Looking into the Distance

The estimation of α_L, β_L, α_R, and β_R is based on the fact that the visual axes of both eyes are always parallel as long as the user is looking into the distance.

$\mathbf{c}_L$ and $\mathbf{c}_R$ are the visual axes of the left and right eye, respectively, and are calculated from {$\mathbf{a}_L$, $\mathbf{b}_L$, and $\mathbf{d}_L$} and {$\mathbf{a}_R$, $\mathbf{b}_R$, and $\mathbf{d}_R$}, respectively, by equations (1), (3), (4), and (5). Therefore, $\mathbf{c}_L$ and $\mathbf{c}_R$ are expressed by the functions of {α_L and β_L} and {α_R and β_R}, respectively.

$\mathbf{c}_L$ and $\mathbf{c}_R$ are parallel unit vectors, and, therefore, the inner product of $\mathbf{c}_L$ and $\mathbf{c}_R$ must be 1:

$$\mathbf{c}_L \cdot \mathbf{c}_R = 1. \quad (6)$$

We have four unknown parameters. If the user looks in four directions, we can obtain four equations from (6) and can thus calculate α_L, β_L, α_R, and β_R.

Simulation of Estimation of the Offset between Optical and Visual Axes under Ideal Conditions

We simulated the eye rotation based on Listing's law and calculated the optical and visual axes of the eyes. We suppose that the user is looking in four directions: top left (20° left, 20° up), top right (20° right, 20° up), bottom left (20° left, 20° down), and bottom right (20° right, 20° down). For example, the true values of α_L, β_L, α_R, and β_R were set to -5.0°, 1.0°, 5.0°, and 1.0°, respectively.

Our objective was to estimate α_L, β_L, α_R, and β_R using only the simulated optical axes of the eye. Then, we obtained the following simultaneous equations:

$$\begin{cases} \mathbf{c}_{L0}(\alpha_L, \beta_L) \cdot \mathbf{c}_{R0}(\alpha_R, \beta_R) = 1 \\ \mathbf{c}_{L1}(\alpha_L, \beta_L) \cdot \mathbf{c}_{R1}(\alpha_R, \beta_R) = 1 \\ \mathbf{c}_{L2}(\alpha_L, \beta_L) \cdot \mathbf{c}_{R2}(\alpha_R, \beta_R) = 1 \\ \mathbf{c}_{L3}(\alpha_L, \beta_L) \cdot \mathbf{c}_{R3}(\alpha_R, \beta_R) = 1 \end{cases}. \quad (7)$$

We solved these equations numerically using Mathematica (Wolfram Research, Inc.) to find α_L = -4.99993°, β_L = 0.992952°, α_R = 5.00004°, and β_R = 0.99205°. Thus, we obtained good results with a reasonable degree of certainty.

Numerical Simulations with Noise

If the measurements of the optical axes of the eye include noise, more than four directions of the optical axes of the eyes are required for a stable solution.

From a preliminary experiment that estimated the optical axis of the eye using the system described in (Nagamatsu et al., 2009), the horizontal and vertical root-mean-square errors (RMSEs) of estimating the optical axis of an artificial eye were 0.21° and 0.18°, respectively.

Therefore, we simulated the noise that may occur when taking real eye measurements by adding a signal with a standard deviation of 0.3° from the direction of the optical axis of the eye. Furthermore, we simulated 25 directions by shifting 10° horizontally (from -20° to 20°) and vertically (from -20° to 20°). For each direction, 60 sets of data were generated.

We estimated α_L, β_L, α_R, and β_R from the simulated optical axes of the eye. First, we calculated the median value of the 60 datasets as the representative direction of the simulated optical axes of the eye for each direction. Then, we estimated α_L, β_L, α_R, and β_R by minimizing the following objective function:

$$F(\alpha_L, \beta_L, \alpha_R, \beta_R) = \sum_i \left\| \mathbf{c}_{Li}(\alpha_L, \beta_L) \cdot \mathbf{c}_{Ri}(\alpha_R, \beta_R) - 1 \right\|^2. \quad (8)$$

Table 1
Simulation result.

Case		Degrees			
		α_L	β_L	α_R	β_R
1	True value	-5.00	1.00	5.00	1.00
	Our method (ave.)	-5.08	1.16	4.92	1.16
	Our method (SD)	0.25	0.38	0.25	0.37
2	True value	-5.00	-1.00	5.00	-1.00
	Our method (ave.)	-5.00	-0.97	5.00	-0.97
	Our method (SD)	0.32	0.22	0.31	0.22
3	True value	-3.00	1.00	3.00	1.00
	Our method (ave.)	-3.39	0.66	2.61	0.67
	Our method (SD)	0.33	0.56	0.33	0.56
4	True value	-3.00	-1.00	3.00	-1.00
	Our method (ave.)	-3.01	-1.08	2.98	-1.08
	Our method (SD)	0.33	0.51	0.33	0.50
5	True value	-1.00	0.50	1.00	0.50
	Our method (ave.)	-0.77	-0.01	1.23	-0.01
	Our method (SD)	1.75	1.05	1.74	1.05
6	True value	-1.00	-0.50	1.00	-0.50
	Our method (ave.)	0.01	-0.88	2.01	-0.87
	Our method (SD)	1.07	1.75	1.06	1.75

We calculated five times for each set of α_L, β_L, α_R, and β_R values. The results are shown in Table 1. The table shows the true values of α_L, β_L, α_R, and β_R, the average of the five estimated values for each true value, and the standard deviation (SD).

We can see that, when $|\alpha|$ is large, our method gives accurate estimates and relatively small SDs. However, for smaller values of $|\alpha|$, the estimates were not good, and the SDs were large.

This suggests that, if the SD is large, the offset between the optical and visual axes of the eye is small. In such cases, the optical axis can be regarded to be almost the same as the visual axis of the eye.

Implementation and Experiments

Implementation

We developed an experimental system based on a calibration-free gaze tracking system for a display described in (Nagamatsu et al., 2009). We replaced the display with transparent glass, so that the user can look into the distance. LEDs for illumination were attached to the glass.

As shown in Figure 5, the experimental system consisted of four synchronized monochrome USB 2.0 digital cameras with a 1/3" CMOS image sensor (Firefly MV, Point Grey Research Inc.) that has a 50-mm lens and an infrared (IR) filter, two IR light sources attached to the glass, and a Windows-based PC (Windows 7, Intel Core i7).

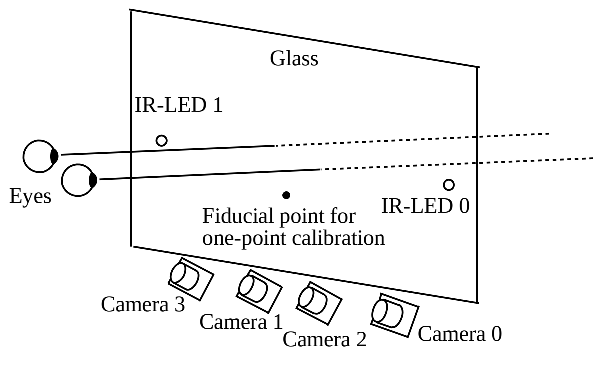


Figure 5. Experimental system.

The software was developed using OpenCV (http://code.opencv.org/). A pair of cameras (Cameras 0 and 1) was used for the estimation of the optical axis of the left eye, and the other pair (Cameras 2 and 3) was used for the right eye.

Experiments

Method. We evaluated the experimental system in a laboratory with three adult participants (one man and two women); only the participant 3 wore soft contact lenses.

The head was supported by a chin rest to prevent from it from being out of focus or out of the field of view of the cameras in this experiment. The eyes were approximately 600 mm from the cameras.

The participants were asked to fixate on 9 points that were arranged within a 3 m (horizontally) by 2 m (vertically) area on the opposite wall of the room as shown in Figure 6, which is a limitation size of our laboratory and we assumed to be far enough from the participants in this stage of experiment. We wanted participants to look at all directions uniformly.

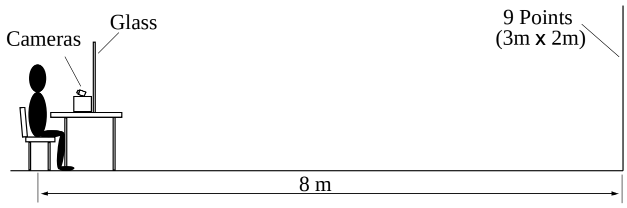


Figure 6. Experimental setting.

Data were recorded when the optical axes of both eyes were detected. We recorded 30 data points when the participant gazed at each point.

In order to get the true values of α_L, β_L, α_R, and β_R, we conducted one-point calibration (Nagamatsu et al., 2008) using the fiducial point on the glass, as shown in Figure 5, which is located in front of the participant.

Results. We estimated the offsets as shown in Table 2 when α_L is varied from -7.0 to 0.0°; β_L from -2.0 to 2.0°; α_R from 0.0 to 7.0°; and β_R from -2.0 to 2.0°. The offsets estimated by the one-point calibration method are listed in Table 2.

The results show that the estimations by our method are similar to those achieved by the one-point calibration method.

Table 2
Estimated offsets.

Participants		Degrees			
		α_L	β_L	α_R	β_R
1	One-point	-1.94	-1.93	2.13	0.86
	Our method	-1.75	-1.98	1.73	2.00
2	One-point	-2.13	1.14	2.51	1.93
	Our method	-2.03	1.36	2.20	2.00
3	One-point	-0.80	0.65	0.64	0.80
	Our method	0.0	0.42	0.50	0.17

Discussion

In this stage of research, the prerequisite of our proposed method is that the user always looks into the distance. There may be several ways how to detect the user looks distant or nearby objects. One way uses the fact that the optical axis of the eye can be estimated without user calibration. The estimated optical axis of the eye differs from the visual axis of the eye by up to 7 degrees. For example, when the user drive a car, we only use the data of optical axis of the eye that are 7 degrees or more inside the window in a user calibration phase.

After a personal calibration, all the user dependent parameters are decided, so the system can estimate gazes when the user gazes at nearby objects such as meters and navigation system.

The experiment was primary step. Therefore, in a future work, we change the distance of targets that the user looks at, and conduct experiments for a large population of participants.

Conclusion

We proposed an automatic user-calibration method for gaze trackers that operates when users look into the distance. This system is suitable for drivers of cars, trains, ships, etc., who spend most of their time looking afar. We calculated the offset parameter values based on the fact that the visual axes of both eyes are always parallel when the user looks into the distance, irrespective of the direction of gaze. We evaluated the proposed method by computer simulation. The method works well when the offset between the optical and visual axes of the eye is large. Furthermore, we developed and evaluated an experimental system.

Acknowledgements

This work was partially supported by KAKENHI (23300047). I wish to thank Prof. Dan Witzner Hansen for helpful suggestions.

References

Guestrin, E. D., & 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.

Hansen, D. W., & Ji, Q. (2010). In the Eye of the Beholder: A Survey of Models for Eyes and Gaze. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 32*(3), 478-500.

Model, D., & Eizenman, M. (2010). *User-calibration-free remote gaze estimation system.* Paper presented at the 2010 Symposium on Eye-Tracking Research & Applications, Austin, Texas (pp. 29-36).

Nagamatsu, T., Kamahara, J., & Tanaka, N. (2008). *3D Gaze Tracking with Easy Calibration Using stereo Cameras for Robot and Human Communication.* Paper presented at the 17th International Symposium on Robot and Human Interactive Communication (IEEE RO-MAN) 2008, Munich, Germany (pp. 59-64).

Nagamatsu, T., Kamahara, J., & Tanaka, N. (2009). *Calibration-free gaze tracking using a binocular 3D eye model.* Paper presented at the 27th international conference extended abstracts on Human factors in computing systems, Boston, MA, USA. (pp. 3613-3618).

Shih, S.-W., & 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., & Cabeza, R. (2008). A Novel Gaze Estimation System With One Calibration Point. *IEEE Transactions on Systems, Man, and Cybernetics, Part B, 38*(4), 1123 - 1138.

Wong, A. M. F. (2004). Listing's law: clinical significance and implications for neural control. *Survey of Ophthalmology, 49*(6), 563-575.