Automatic Calibration of an In-vehicle Gaze Tracking System Using Driver's Typical Gaze Behavior

Kenji Yamashiro[1], Daisuke Deguchi[1], Tomokazu Takahashi[1,2],
Ichiro Ide[1], Hiroshi Murase[1], Kazunori Higuchi[3], and Takashi Naito[3]

[1] Graduate School of Information Science, Nagoya University, Japan
[2] Faculty of Economics and Information, Gifu Shotoku Gakuen University, Japan
[3] Toyota Central Research & Development Laboratories, Inc., Japan
E-Mail: yamashiro@murase.m.is.nagoya-u.ac.jp, ttakahashi@gifu.shotoku.ac.jp,
{ddeguchi, ide, murase}@is.nagoya-u.ac.jp

***Abstract*—Many research works have been carried out to measure and use a driver's gaze directions to prevent traffic accidents caused by inattentive driving, neglect to confirm safe conditions, and other driver errors. A calibration process is needed to measure correct gaze directions for a gaze tracking system. However, existing calibration methods require a driver to gaze at specified points before driving. In this paper, we propose a method for automatic calibration of an in-vehicle gaze tracking system by analyzing the driver's typical gaze behavior. The proposed method uses the rear-view and the side-view mirror positions as reference points. The effectiveness of the proposed method is demonstrated by experiments on measuring gaze directions in actual road environments.**

I. Introduction

In recent years, fatalities due to traffic accidents have decreased in Japan, reaching as few as 6,000 deaths in 2007. However, about 830,000 traffic accidents still occur every year, and 1.04 million people are injured [1]. Most of these traffic accidents occur when a driver looks aside or otherwise engages in careless driving. Some of these might be avoided by warning drivers based on their gaze direction. For example, if they pass through an intersection without noticing the traffic lights, it could be worthwhile to warn them by monitoring their gaze directions and traffic light recognition [2]. Therefore, it is important to measure a driver's gaze directions to support safe driving.

To measure gaze directions accurately, calibration is one of the most important functions required for a gaze tracking system. Calibration compensates the difference between the actual gaze direction of a driver and the gaze direction measured by the gaze tracking system. The difference is caused by individual differences in eyeball shape, facial features, etc. Moreover, even with the same driver, lighting conditions and the driver's head position can cause a difference.

Usually, such calibration requires several measurements of gaze directions by making the driver gaze at reference points. Then, calibration parameters are calculated by comparing the actual and measured gaze directions. However, this task is very time-consuming and troublesome for users. It also has to be done before actually measuring gaze direction, but it is not realistic to expect users to carry out calibration each time before driving.

In this paper, we propose a method for automatic calibration of an in-vehicle gaze tracking system by analyzing the driver's gaze behavior. By obtaining the gaze directions to two reference points whose exact positions are known, we can calibrate the system for rotation, scaling, and translation. The method adopts the positions of the rear-view mirror and the side-view mirror as reference points. It consists of 4 steps: (1) selection of gaze direction data in lane-changing situations, (2) extraction of stationary gaze directions, (3) estimation of mirror positions by the EM algorithm, and (4) calculation of calibration parameters. Instead of the general time-consuming and troublesome calibration process, the proposed method completes the calibration process implicitly as the driver drives.

Section II introduces previous works related to this research. Section III describes the detailed procedures of the proposed method. Section IV demonstrates the effectiveness of the proposed method by experiments on measuring gaze directions in actual road environments. Finally, we summarize this paper in Section V.

II. Related works

In general, a calibration process requires two to twenty reference points, and users have to gaze at these points one by one, which is a very time-consuming task. Consequently, two types of methods have been developed that facilitate this calibration process.

Ohno et al. proposed a method for reducing the number of reference points required for calibration called FreeGaze [3]. This system uses an eyeball model, which reduces the number of fixed reference points for calibration.

Yabuuchi et al. proposed a method that uses a moving reference point for calibration to reduce the burden on users [4].

Gazing at a moving point is more natural for users than gazing at fixed points. They display a moving point that scans over a screen and have users gaze at it continuously. This method achieved higher accuracy than approaches using fixed reference points.

In the case of a car-mounted system, however, the driver cannot gaze at specified points required for calibration for safety reasons. Therefore, these methods have to be conducted each time before driving, which is not a realistic expectation. As a solution to this problem, we propose a method for automatic calibration of an in-vehicle gaze tracking system by using the driver's typical gaze behavior practiced in actual driving.

III. Automatic Calibration of an In-vehicle Gaze Tracking System Using Driver's Gaze Behavior

A. Overview

This paper proposes a method for automatic calibration of an in-vehicle gaze tracking system by using a driver's typical gaze behavior. The positions of the rear-view and the side-view mirrors are used as reference points for the calibration process. The number of gaze directions corresponding to rear-view and side-view mirrors increases when the driver changes the driving lanes [5]. Accordingly, the proposed method exploits the gaze directions naturally measured in lane-changing situations. From the gaze directions in lane-changing situations, the proposed method extracts and uses the driver's gazes at the rear-view mirror or the right side-view mirror for calibration.

The proposed method uses a noncontact gaze tracking system for obtaining driver's gaze directions. The gaze directions measured by an uncalibrated system are used as inputs for the proposed method. The obtained gaze directions are represented by horizontal and vertical angles.

Fig. 1 shows a flow chart of the proposed method. First, the gaze directions corresponding to the rear-view and the right side-view mirrors and the front view are extracted by detecting stationary gaze directions. Next, assuming that the distribution of the gaze directions is a mixture of three normal distributions corresponding to the gaze directions of the rear-view and the right side-view mirrors and the front view, the method estimates these distributions by the EM algorithm [6]. Finally, the calibration parameters are calculated by comparing the actual and the estimated gaze directions.

B. Selection of gaze direction data in lane-changing situations

When a driver changes the driving lanes from left to right †, the driver frequently gazes at both the rear-view and the right side-view mirrors [5]. The proposed method makes use of this behavior for the calibration of the gaze tracking system. Fig. 2 shows a distribution of gaze directions measured over an hour, and Fig. 3 shows the distribution of only those gaze directions obtained when the driver changes the driving lanes. As can be seen from Fig. 3, we can clearly distinguish three distributions. Actually, they correspond to the rear-view and the right side-view mirrors and the front view.

†In Japan, drivers keep to the left side of the road.

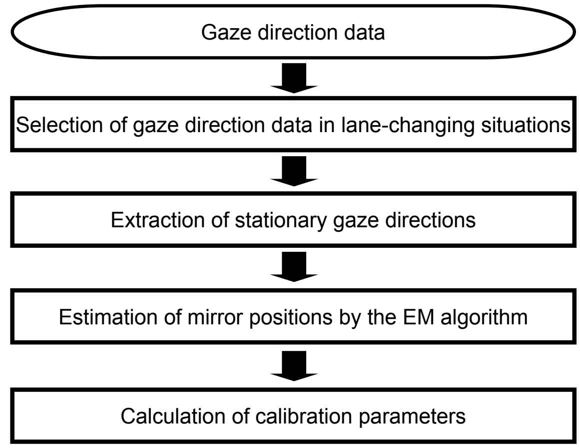


Fig. 1. Flow chart of proposed method.

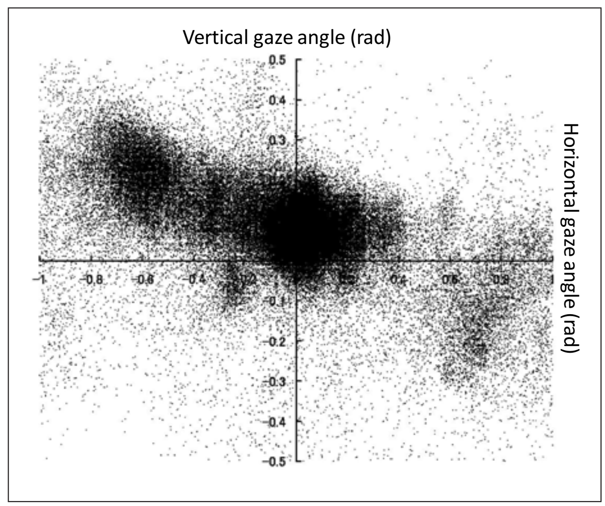


Fig. 2. Distribution of gaze directions over an hour.

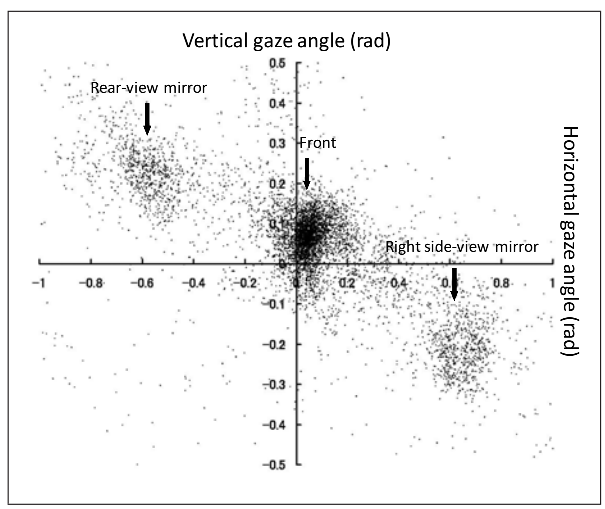


Fig. 3. Distribution of gaze directions during lane changing.

C. Extraction of stationary gaze directions

There are directions that a driver gazes at for a short but sustained period. We call these directions "stationary gaze directions". On the other hand, gaze directions include not only directions to mirrors but also directions moving between the mirrors and the front. Since the proposed method uses mirror positions for calibration, these unnecessary gaze directions may affect the estimation of correct mirror positions. Therefore, the proposed method tries to eliminate these unnecessary gaze directions by evaluating whether gaze directions remain stationary for a certain period of time.

The extraction method for stationary gaze directions is as follows. We define a stationary gaze direction so that the angle between any two of N successive samples $\boldsymbol{x}_i(i = n, n+1, ..., n+N)$ should be less than or equal to a threshold C. Their definition is formulated as

$$\max ||\boldsymbol{x}_i - \boldsymbol{x}_j|| \leq C, \quad (i, j = n, n+1, \ldots, n+N) \tag{1}$$

where $||\boldsymbol{x}||$ is the norm of a vector $\boldsymbol{x}$.

Fig. 4 shows stationary gaze directions extracted by applying Eq. (1) to the gaze direction data in a lane-changing situation ($N = 10$, $C = 0.15$). As seen in Fig. 4, it is possible to eliminate unnecessary gaze directions.

D. Estimation of mirror positions by the EM algorithm

The proposed method assumes that the distribution of the gaze directions is a mixture of three distributions corresponding to gaze directions to the rear-view and the right side-view mirrors and the front view. As seen in Fig. 4, these three distributions are clearly observed during lane-changing. Each distribution is approximated as a Gaussian distribution, and gaze directions measured by the system are considered a mixture of the three Gaussian distributions. A mixture of three Gaussian distributions could be described as

$$p(\boldsymbol{x}) = \sum_{i=1}^{3} \omega_i \phi(\boldsymbol{x}|\boldsymbol{\mu}_i, \boldsymbol{\Sigma}_i), \tag{2}$$

where ω_i, $\boldsymbol{\mu}_i$, and $\boldsymbol{\Sigma}_i$ represent the mixture ratio, the mean vector, and the covariance matrix, respectively, of the i th distribution, and ϕ represents the Gaussian distribution.

These parameters are estimated by applying the EM algorithm [6] to the stationary gaze direction data.

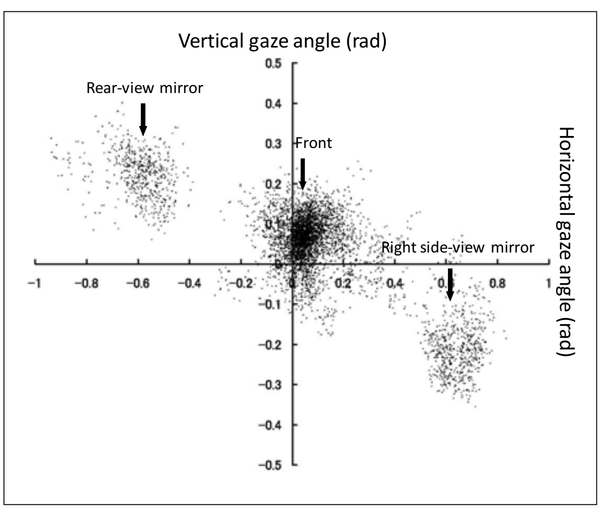


Fig. 4. Distribution of stationary gaze directions.

E. Compensation of gaze directions

This step compensates the difference between the actual and the measured gaze directions. We assumed that the differences can be approximated by a combination of translation, rotation, and scaling. The actual gaze direction (x', y') can be calculated by the measured gaze direction (x, y) by

$$\begin{pmatrix} x' \\ y' \\ 1 \end{pmatrix} = \mathbf{H} \begin{pmatrix} x \\ y \\ 1 \end{pmatrix}, \tag{3}$$

where $\mathbf{H}$ is represented as

$$\mathbf{H} = \begin{pmatrix} w_1 & w_2 & w_3 \\ -w_2 & w_1 & w_4 \\ 0 & 0 & 1 \end{pmatrix}. \tag{4}$$

To solve four unknown parameters $w_1, ..., w_4$ in matrix $\mathbf{H}$, four simultaneous equations are required. Since the proposed method refers to two points corresponding to the rear-view and the right side-view mirrors, Eq. (3) can be rewritten as

$$\begin{pmatrix} x'_b & x'_r \\ y'_b & y'_r \\ 1 & 1 \end{pmatrix} = \widehat{\mathbf{H}} \begin{pmatrix} x_b & x_r \\ y_b & y_r \\ 1 & 1 \end{pmatrix}, \tag{5}$$

where (x_b, y_b),(x_r, y_r) are the estimated gaze directions toward the rear-view and the right side-view mirrors, respectively, and (x'_b, y'_b),(x'_r, y'_r) are the corresponding actual gaze directions. Hence, the matrix $\widehat{\mathbf{H}}$ can be easily solved.

Since the proposed method uses three reference points, corresponding to the two mirrors and the front view, the calibration parameters are solved as a least squares problem.

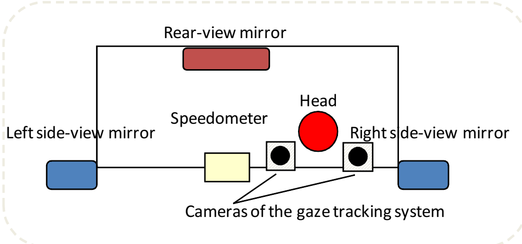
Finally, the gaze directions are corrected using the calculated calibration parameters by

$$\begin{pmatrix} \hat{x} \\ \hat{y} \\ 1 \end{pmatrix} = \widehat{\mathbf{H}} \begin{pmatrix} x \\ y \\ 1 \end{pmatrix}, \tag{6}$$

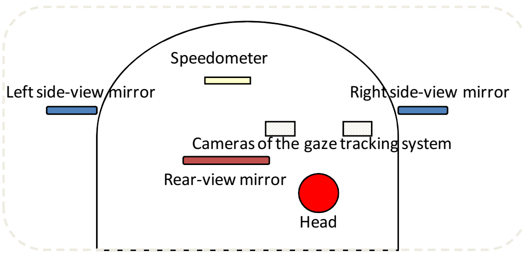
where (x, y) are the measured gaze directions, and $(\hat{x}, \hat{y})$ are the gaze directions after calibration.

IV. EXPERIMENTS

We evaluated the proposed method by applying it to gaze directions measured in actual road environments. The evaluation was conducted through two experiments: (1) evaluation of accuracy in estimating gaze directions corresponding to mirrors, and (2) evaluation of accuracy by increasing the number of lane-changes. The following sections describe the experimental conditions and the results.



(a) Mirror positions observed from behind the driver.



(b) Mirror positions observed from above the driver.

Fig. 5. Positions of the driver, cameras used by the gaze tracking system, and the mirrors.

A. CONDITIONS

In this experiment, gaze directions were measured for over one hour while driving a vehicle, at a rate of 60 measurements per second from an in-vehicle noncontact gaze tracking system "FaceLAB" [7]. Each gaze direction consists of horizontal and vertical gaze angles. The cameras of the FaceLAB system were set up on the dashboard of a vehicle. Fig. 5 shows the positions of the driver, the cameras used by the gaze tracking system, and the rear-view and the side-view mirrors.

To obtain the ground truth data of gaze directions corresponding to the mirrors, we manually labeled the stationary gaze directions as the rear-view mirror or the right side-view mirror.

B. Evaluation of accuracy in estimating gaze directions corresponding to the mirrors

In this experiment, we evaluated the accuracy of estimating the gaze directions corresponding to the rear-view and the right side-view mirrors.

1) Method: Fourteen sets of gaze direction data were prepared for this evaluation. Each lane-changing segment was defined as the five seconds before turning the blinker on or off. Then, the stationary gaze directions were extracted by applying the method described in Section III-C with parameters $N = 10$ and $C = 0.15$. Finally, each mirror direction is estimated by applying the EM algorithm as described in Section III-D. Table I shows the initial values used for this step.

We evaluated the estimation accuracy by using two types of datasets. The first one consists of stationary gaze directions extracted by the proposed method, and the other consists of all gaze directions in the lane-changing segment.

TABLE I
INITIAL VALUES OF THE EM ALGORITHM.

	Mean vector	Covariance matrix	Mixture ratio
Rear-view mirror	$(-0.67, 0.14)$	$\begin{pmatrix} 0.01 & 0 \\ 0 & 0.01 \end{pmatrix}$	0.80
Right side-view mirror	$(0.65, -0.40)$	$\begin{pmatrix} 0.01 & 0 \\ 0 & 0.01 \end{pmatrix}$	0.10
Front view	$(0.00, 0.00)$	$\begin{pmatrix} 0.01 & 0 \\ 0 & 0.01 \end{pmatrix}$	0.10

TABLE II
RESULT: ANGLES BETWEEN THE ESTIMATED MIRROR DIRECTIONS AND THE GROUND TRUTH.

Experimental data	Rear-view mirror	Right side-view mirror
All gaze directions	0.105 [rad]	0.105 [rad]
Stationary gaze directions	0.009 [rad]	0.041 [rad]

To obtain the ground truth of gaze directions corresponding to the rear-view and the right side-view mirrors, we manually labeled 54 and 32 sequences of gaze direction data, respectively. Their average positions were used as the ground truth.

2) Results: Table II shows the angles between the estimated mirror directions and the ground truth.

3) Discussion: From Table II, we can confirm that the estimation accuracy using only the stationary gaze directions is much higher than that using all gaze directions. When all gaze directions are used for the estimation, the estimated directions to both mirrors were affected by inappropriate gaze direction data, especially the data captured when the driver's eyes drifted between the mirrors and the front. In contrast, by extracting the stationary gaze direction data, the estimation accuracy improved significantly.

In a typical in-vehicle environment, the distances between the head of the driver and the rear-view and the right side-view mirrors are 50 cm and 70 cm, respectively. In this case, the error of 0.009 rad corresponds to about 0.5 cm on the rear-view mirror, while that of 0.041 rad corresponds to about 2.8 cm on the right side-view mirror. Since 0.5 cm on the rear-view mirror and 2.8 cm on the right side-view mirror are sufficiently smaller than the sizes of the mirrors themselves, we confirmed that the proposed method could estimate the directions to the mirrors with sufficient accuracy.

C. Evaluation of accuracy by increasing the number of lane-changes

In this experiment, we observed how an increase in lane-changes contributes to the accuracy of the proposed method.

1) Method: The estimation accuracy was evaluated by increasing the number of lane-changes from 1 to 14. For each evaluation, twenty trials were conducted by randomly selecting and ordering the lane-changing segments.

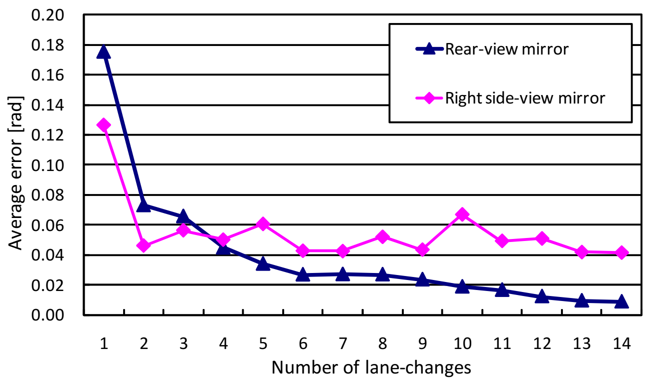


Fig. 6. Estimation error when increasing the number of lane-changes.

2) Results: Fig. 6 shows the average estimation error when increasing the number of lane-changes.

3) Discussion: From Fig. 6, we can confirm that the estimation error tends to decrease gradually by increasing the number of lane-changes. In our experiments, the driver sometimes did not gaze at the mirrors when changing the driving lanes. In this case, the proposed method cannot estimate the gaze directions to the mirrors. The driver also sometimes gazed at the speedometer, which was not considered as a reference object in the experiment. This strongly affected the estimation accuracy. However, the ratio of these inappropriate gaze directions decreased by increasing the number of lane-changes, since the driver gazed at the mirrors more frequently than at the speedometer while changing driving lanes. Therefore, the proposed method can reduce the estimation error with an increase in lane-changes, even if inappropriate gaze directions occur.

D. Example of calibration

This section shows an example of the proposed method's calibration of gaze directions by applying it to intentionally misaligned gaze direction data.

1) Method: The following four types of data were prepared for the experiment.

(i) Ground truth of the gaze directions to the mirrors and to the front
  - The gaze directions to the mirrors and the front estimated in Section IV-B were considered the ground truth

(ii) Input data: Uncalibrated gaze direction data including some misalignments
  - Data synthesized by rotating data used in Section IV-A 10 degrees clockwise and then magnifying them 1.2 times and finally translating them 0.2 rad in both horizontal and vertical directions by the following equation:

$$\begin{pmatrix} x' \\ y' \\ 1 \end{pmatrix} = \mathbf{R}\ \mathbf{S}\ \mathbf{T} \begin{pmatrix} x \\ y \\ 1 \end{pmatrix}, \tag{7}$$

where $\mathbf{R} = \begin{pmatrix} \cos(\frac{\pi}{18}) & \sin(\frac{\pi}{18}) & 0 \\ -\sin(\frac{\pi}{18}) & \cos(\frac{\pi}{18}) & 0 \\ 0 & 0 & 1 \end{pmatrix}$,

$\mathbf{S} = \begin{pmatrix} 1.2 & 0 & 0 \\ 0 & 1.2 & 0 \\ 0 & 0 & 1 \end{pmatrix}$, $\mathbf{T} = \begin{pmatrix} 1 & 0 & 0.2 \\ 0 & 1 & 0.2 \\ 0 & 0 & 1 \end{pmatrix}$,

and (x, y) represents the appropriately calibrated gaze directions and (x', y') the synthesized data.

(iii) Ground truth of the evaluation data
  - Data for the evaluation of the calibration accuracy at fixed grids set at 0.1 rad intervals in the range from -1.5 rad to 1.5 rad in the horizontal direction and from -1.0 rad to 1.0 rad in the vertical direction.

(iv) Uncalibrated evaluation data including some misalignments
  - Data synthesized from data (iii) by Eq. (7)

First, we calculated the calibration parameters by comparing the directions estimated by using data (ii) and the ground truths data (i). Then, the evaluation points data (iv) were calibrated by using the obtained parameters, which were compared visually with data (iii).

The experiments were conducted for the case of referring to two points, corresponding to the rear-view and the right side-view mirrors, and also for the case of referring to three reference points, corresponding to the two mirrors and the front view.

2) Results: Fig. 7 shows the calibrated results by using three reference points that correspond to the rear-view mirror, the right side-view mirror, and the front view. Table III shows the average and the standard deviation of the angles between the calibrated evaluation points and the ground truth.

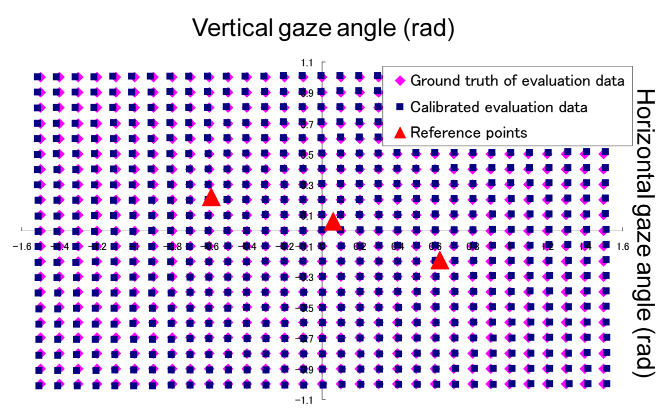


Fig. 7. Results of calibration using three reference points.

As an example of the calibrated results, Fig. 8 shows the shift from uncalibrated to calibrated gaze directions when a driver gazed at a traffic signal.

3) Discussion: If we assume a point 1.0 m away from the head of a driver, the error of 0.0098 rad by the calibration from two reference points corresponds to approximately 1.0 cm in a real environment. In general, the more reference points that

TABLE III
ANGLES BETWEEN THE CALIBRATED DATA AND THE GROUND TRUTH [RAD].

Number of reference points	Average	Standard deviation
Two	0.0098	0.0039
Three	0.0096	0.0038

Fig. 8. Example of calibrated result when driver gazed at a traffic signal.

are used, the higher calibration accuracy should be obtained. We confirmed this from the result that calibration by three reference points slightly improved that by two reference points.

Another observation was that the further from the reference points, the larger the error was. This implies that the error caused by the estimation of the direction to the reference points has a great influence on the calibrated result far from the reference points.

Increasing the number of reference points would be an effective way to improve the calibration accuracy and to decrease the error over a large range. However, since a driver does not gaze the front only but also some objects close to the front, such as traffic signal or signs, gaze direction to the front view are not so stable as gaze directions to the mirrors. Therefore we should evaluate another reference points instead of the front view. In this experiment, the gaze direction to the left side-view mirror was not used because the current camera position cannot measure gaze directions corresponding to both right side-view and left side-view mirrors due to the measurement range. In the future, we will try using the left side-view mirror as a reference point by adequately arranging the camera positions. We will also try to use the speedometer as a reference point because we noticed throughout the experiments that the drivers gazed at it as much as at the mirrors.

V. CONCLUSION

This paper proposed a method for automatic calibration of an in-vehicle gaze tracking system using the driver's typical gaze behavior. The proposed method uses the rear-view and the side-view mirror positions as reference points for the calibration process and estimates the gaze directions to the mirrors using the gaze direction data measured during lane-changing situations. Finally, the calibration parameters are calculated using the estimated mirror directions.

In the experiments, the proposed method was applied to gaze directions measured in actual road environments. The experimental results showed that the estimation error of mirror positions could be reduced to less than 0.05 rad. Furthermore, the calibration error could be decreased to 0.01 rad on average.

Future works include

- Improvement of the extraction of stationary gaze directions
- Calibration using more than three reference points (e.g. left side-view mirror, speedometer)
- Experiments on many more data (e.g. more drivers, different driver's head position)

ACKNOWLEDGMENTS

The authors would like to thank their colleagues for useful discussions. The proposed method was implemented using the MIST library (http://mist.murase.m.is.nagoya-u.ac.jp/).

REFERENCES

[1] National Police Agency, http://www.npa.go.jp/english/index.htm
[2] F. Lindner, U. Kressel, and S. Kaelberer, "Robust Recognition of Traffic Signals," Proc. IEEE Intelligent Vehicles Symposium 2004, pp. 49–53, June 2004.
[3] Takehiko Ohno, Naoki Mukawa, and Atsushi Yoshikawa, "FreeGaze: a gaze tracking system for everyday gaze interaction," Proceedings of the Symposium on Eye Tracking Research & Applications 2002, pp. 125–132, March 2002.
[4] Tsutomu Yabuuchi, Takanori Miyamoto, Tetsuya Yamamoto, Norifumi Katafuchi, Joji Nakayama, and Ken-ichiro Shimokura, "Automatic Calibration of Gaze Tracking with FreeGaze (in Japanese)," Technical Report of IEICE HIP-2004-11-23, July 2004.
[5] Dario D. Salvucci and Andrew Liu, "The time course of a lane change: Driver control and eye-movement behavior,"Transportation Research, Part F, Vol. 5, No.2, pp. 123–132, June 2002.
[6] Richard O. Duda, Peter E. Hart, and David G. Stork, "Pattern Classification Second Edition", Wiley Interscience, 2000.
[7] Seeing Machines, http://www.seeingmachines.com/facelab.htm