



Driver's eye-based gaze tracking system by one-point calibration

Hyo Sik Yoon¹ · Hyung Gil Hong¹ · Dong Eun Lee¹ · Kang Ryoung Park¹

Received: 22 November 2017 / Revised: 7 June 2018 / Accepted: 2 August 2018 /

Published online: 7 August 2018

© Springer Science+Business Media, LLC, part of Springer Nature 2018

Abstract

The accuracies of driver's gaze detection by previous researches are affected by the various sitting positions and heights of drivers in case that initial calibration of driver is not performed. By using dual cameras, the driver's calibration can be omitted, but processing time with complexity is increased. In addition, the problem of disappearing corneal specular reflection (SR) in the eye image as the driver severely turns his/her head has not been dealt in previous researches. To consider these issues, we propose a gaze tracking method based on driver's one-point calibration using both corneal SR and medial canthus (MC) based on maximum entropy criterion. An experiment with collected data from 26 subjects (wearing nothing, glasses, sunglasses, hat, or taking various hand pose) in a vehicle, showed that the accuracy of the proposed method is higher than that of other gaze tracking methods. In addition, we showed the effectiveness of our method in the real driving environment.

Keywords Driver's gaze detection · NIR camera and NIR illuminator · Pupil, corneal specular reflection and medial canthus · Initial calibration of driver · Maximum entropy criterion

1 Introduction

Gaze detection determines the areas at which a user is looking by using cameras and other sensors [16, 18, 67]. With advances in technology in recent years, driver assistance systems have also

✉ Kang Ryoung Park
parkgr@dongguk.edu

Hyo Sik Yoon
yoonhs@dongguk.edu

Hyung Gil Hong
hell@dongguk.edu

Dong Eun Lee
leede@dongguk.edu

¹ Division of Electronics and Electrical Engineering, Dongguk University, 30 Pil-dong-ro 1-gil, Jung-gu, Seoul 100-715, South Korea

undergone a number of developments. Examples include the frontal collision warning system, the lane departure warning system, and the blind spot detection system. Car accidents are mainly caused by human error, through occurrences such as distraction and drowsiness [37–39]. Driver assistance systems mostly use sensors to detect external dangers such as the frontal collision and lane departure, etc. Therefore, the distracted driving such as not looking at frontal view or driver's drowsiness is not monitored by these sensors. There have been various methods to detect the status of driver [2, 14, 33, 55], and gaze detection is one of these methods. Most of previous gaze detection researches were performed in indoor desktop environments [9, 19, 23, 30, 35, 48, 51, 56, 63, 68], and even the previous studies of driver's gaze detection considered only head movement without eye tracking [20, 43]. Although there have been previous researches on driver's gaze detection based on eye-tracking techniques [4, 20, 57, 60], their accuracies are affected by the various sitting positions and heights of drivers in case that initial calibration of driver is not performed. In general, for single camera-based gaze detection, initial user calibration is needed. By having the user observe a few designated spots on a display used for gaze detection at the outset, calibration maps the relationship between the location of the display and that of the center of the pupil, which is compensated to account for corneal specular reflection (SR) in the captured image of the eye. User calibration is also needed to compensate for the kappa angle between the location of the user's gaze and that of the pupil, and for the fact that people's eyeballs vary in size. In a vehicular environment, however, drivers find it difficult to perform initial calibration because the spots to observe to this end cannot be marked inside the vehicle. Therefore, past research on gaze detection in vehicular environments have usually skipped the calibration step, and focused on tracking a driver's gaze based on the position of his/her head, which in turn is determined using the location of the iris center (instead of pupil center) or a statistical 3D model of the human head, which has the limitation of accuracy enhancement on gaze tracking in car environment. By using dual cameras, the driver's calibration can be omitted, but processing time with complexity is increased by using the images from two cameras. In addition, the problem of disappearing corneal SR in the eye image as the driver severely turns his/her head to look at the side-view mirror has not been dealt in previous researches. To solve this problem, we propose the single camera-based driver's gaze tracking method in actual vehicle environment based on one-point calibration, and using medial canthus (MC) in addition to pupil center to cope with the severe rotation of driver's head. Our research is novel in three ways compared to past researches.

- The driver's gaze is calculated by combining the prior knowledge of average calibration in desktop monitor with the offset of each driver's gazing at near-infrared (NIR) illuminator at the initial step (one-point calibration).
- The robust gaze tracking is possible by tracking both corneal SR and MC to solve the problem of corneal SR being missed in eye image. And, the accurate parameters for calculating gaze position are obtained based on maximum entropy criterion.
- Because there is few open database for NIR light-based gaze tracking in car environment, we made our collected database open through [15] in order to enable other researchers to compare their performance.

2 Related works

Previous gaze detection studies in vehicular environments can be largely sorted into dual cameras-based methods [3, 43, 61] and single camera-based methods. The latter can be also

classified into 3D [57, 58, 60, 65] and 2D methods. The 2D gaze detection methods either use Purkinje images [4, 50] or detect facial feature points [20] to estimate gaze. One method that used Purkinje images employed the first image, the reflection from the outer surface of the cornea, and the fourth Purkinje image, the reflection from the inner surface of the lens, of the four Purkinje images to track gaze [4]. Another study [20] used facial feature points to find the iris and binarized it to estimate the area being gazed at. Regression-based methods have been researched, which includes appearance-based gaze estimation via uncalibrated gaze pattern recovery, adaptive linear regression for appearance-based gaze estimation [27, 31]. In [25], Gosh et al. proposed the method for eye detection and tracking for monitoring driver's vigilance. In [24], they proposed a non-intrusive approach for drowsiness detection, and their method just recognized the drowsiness status of driver by recognizing open or closed eye. In [42], authors proposed *SafeDrive* system which could automatically determine driver phone use leveraging built-in smartphone sensors sensing driving conditions. In [41], they proposed *Health Driving*, a smartphone-based system for detecting driving events and road conditions solely with a built-in smartphone acceleration sensor. In [40], authors proposed *Safe Walking*, an Android smartphone-based system that can detect the walking behavior of pedestrians by leveraging the sensors and front camera on smartphones. In previous research [45], they proposed the method for generic human motion tracking based on annealed particle filtering and Gaussian process dynamical model. In [11], they also proposed a fusion formulation which combines low- and high-dimensional tracking methods into one framework. In previous research [46], they proposed the method for complicated activity recognition being composed of two components such as temporal pattern mining and adaptive multi-task learning. In [47], they also proposed the method to identify temporal patterns among actions and use the identified patterns to represent activities for automated recognition based on support vector machine (SVM) or k-nearest neighborhood (kNN) classifier. In [44], they defined an atomic activity-based probabilistic framework that uses Allen's interval relations. However, all of these previous researches did not study the driver's gaze detection. Table 1 shows a comparative analysis of proposed and previous studies.

3 Proposed gaze detection method

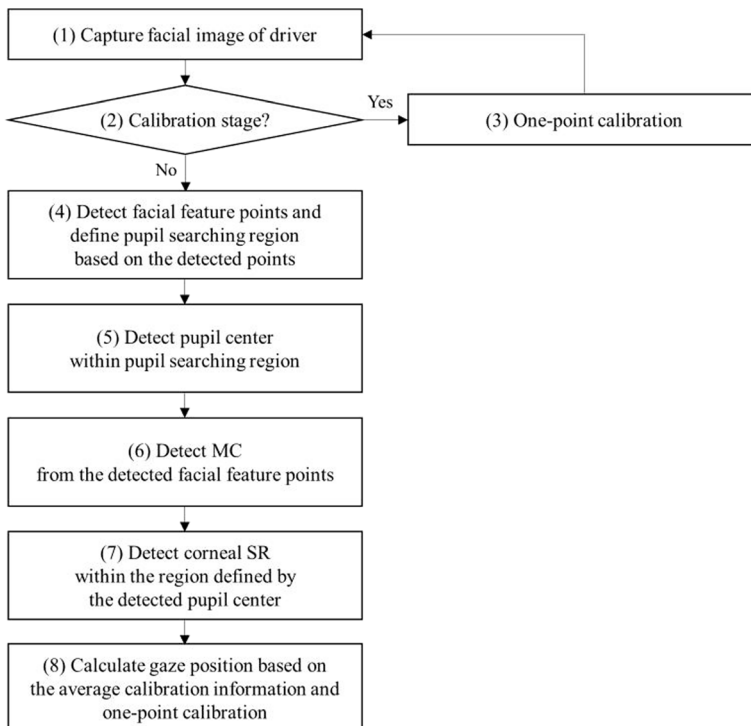
3.1 Overview of proposed method

Figure 1 shows the flowchart of the proposed method. In step (1), a gaze detection system (see Section 3.2) that combines an NIR camera and an NIR illuminator is used to capture the facial image of a driver.

If the driver has yet to go through calibration, he or she looks once at the NIR illuminator in the gaze detection system to perform one-point calibration, as shown in step (3) (see Section 3.5). If the driver has already performed one-point calibration, facial feature points are detected from the captured facial image, and the pupil search region is defined based on the detected eye feature points, as shown in step (4) (see Section 3.3). Then, the center of the pupil is detected within the search region and the MC points are detected from the facial feature points (steps (5) and (6)) (see Section 3.4). The SR search region is defined to find corneal SR based on the detected pupil center, and the corneal SR is detected within this region (step (7)) (see Section 3.4). In the final step, the driver's gaze is calculated by consolidating the direction of the gaze of the driver obtained through one-point calibration in step (3) and the average

Table 1 Comparison between the proposed method and past research

Category	Methods	Advantage	Disadvantage
By dual cameras [3, 43, 61]	Tracks drivers' gaze using dual cameras	Enhanced reliability of tracking gaze	- Expensive gaze detection system and complicated calibration [3, 43] - Difficult to use with bespectacled users [43]
By single camera	3D [57, 58, 60, 65]	3D head or eyeball model by eye tracking based on optical flow [5, 57, 58]	Lower complexity of calculation
	2D	Purkinje image-based [4], and iris center-based [20]	- No measurement of accuracy in car [4] - Lower accuracy due to iris center-based gaze detection [20]
	Regression-based [27, 31] One-point calibration & MC tracking-based (Proposed method)	Continuous gaze position can be obtained - Enhanced accuracy based on one-point calibration - Robust gaze detection by both corneal SR and MC	Training for regressor is required Less convenient compared to calibration-free method

**Fig. 1** Flowchart of proposed method

calibration information of multiple subjects acquired beforehand using the detected pupil center, the MC, and the corneal SR (step (8)) (see Section 3.5).

3.2 Proposed gaze detection system

The gaze detection system for vehicles developed for this study consisted of an NIR camera and an NIR light-emitting diode (LED) illuminator, as shown in the upper part of Fig. 2; it was small in size ($8.8 \times 4.3 \times 4.0 \text{ cm}^3$ (in width \times height \times depth)). Because it was small, it could be installed in the vicinity of the dashboard, as shown in the lower part of Fig. 2, and can continuously track the driver's gaze without obscuring the dashboard. Only in the case that the driving wheel is severely rotated (more than 60 or 70 degrees), the NIR illuminator or camera can be occluded by the driving wheel. However, this seldom occurs while actual driving. The NIR illuminator of 6 NIR-LEDs in the gaze detection system was placed to the left of the camera, and helped capture the driver's facial image without being influenced by the changing light at night and day. Using the NIR LEDs at a wavelength of 850 nm prevented uncomfortable dazzling to driver. A universal serial bus (USB) Web camera [17] was used to capture NIR images of 1600×1200 pixels, and a zoom lens (focal length of 9 mm) was attached to the camera. An 850 nm band pass filter (BPF) was also mounted on the camera's lens to minimize interference due to sunlight [1]. Power was supplied by a laptop computer using two USB lines, one connecting the camera and the other the illuminator. All the data acquisition and testing were performed on a laptop computer with 2.80 GHz CPU (Intel® Core™ i5-4200H) and 8 GB of RAM.

3.3 Detecting facial feature points and defining pupil search region

Previous gaze detection studies used the bright and dark pupil method to detect the center of the user's pupil [18, 68]. For this to work, however, two separate illuminator groups need to be turned on and off in synchrony with the moments captured by the camera images to generate

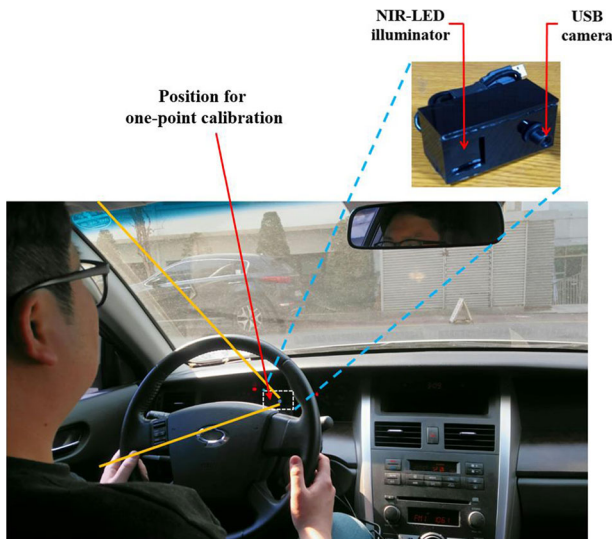


Fig. 2 Proposed gaze detection system in vehicular environment

both bright and dark pupils, which makes the system larger and therefore more difficult to mount on a vehicle. This method was primarily used in an indoor desktop gaze detection environment; hence, in a vehicular environment where external sunlight flows in freely, bright pupils cannot be properly generated. Other studies have used a method to detect bright pixels within the captured images to detect the user's corneal SR [9, 30], but it is difficult to use such methods in vehicular environments because brighter pixels by sunlight exist than in the corneal SR. For these, the proposed method first detects facial feature points, and then designates the pupil search region based on these points to detect the pupil center (step (4) of Fig. 1). Using the dlib facial feature point tracking method [13, 32], 68 feature points were automatically detected as shown in Fig. 3. The pupil search regions were defined as in Fig. 3b (green boxes) based on detected points 36 to 41 and points 42 to 47. In our research, the rough search region is defined and accurate pupil center is again detected based on the method of Fig. 4. Therefore, the incorrect search region caused by the incorrect detection of dlib tracking method does not affect the accurate detection of pupil center. However, if the rough search region is defined in completely different area of facial feature such as nostril caused by the incorrect detection by dlib method, it can cause the error of pupil detection. To solve this problem, our system successively compares the positions of eye centers detected by the dlib method in previous and current frames. As explained in Section 4.5, the proposed method can operate at a speed of approximately 35 frames per second (1000/28.6). Considering the time difference of 28.6 ms, the difference between the previous and current positions of eye centers cannot be large. Therefore, if there exists larger difference than threshold, our system determines that the positions of eye centers in current frame are not correct due to the failure of dlib method, and it uses the positions in previous frame for defining the rough pupil search region in current frame.

3.4 Pupil, MC, and Corneal SR Detection

To estimate the area at which the driver is gazing using the pupil center corneal reflection (PCCR)-based gaze detection method, the center of pupil and the geometric center [6]

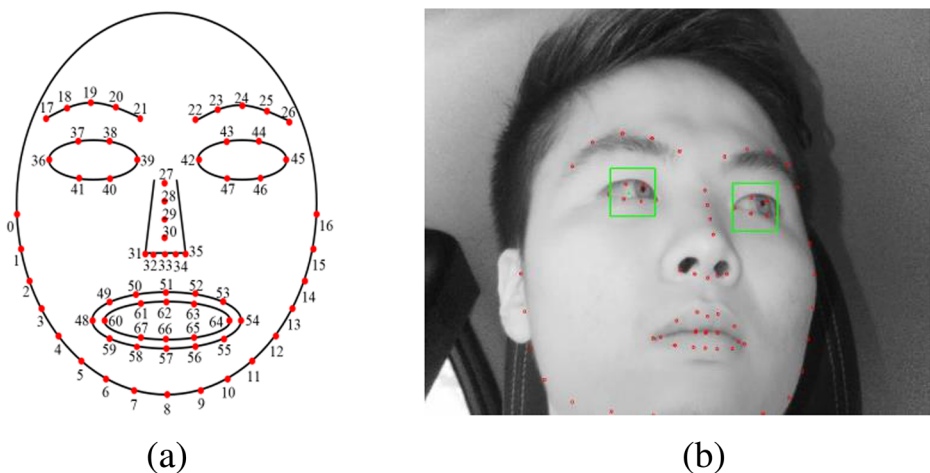


Fig. 3 Examples of **a** detected facial feature points and **b** the defined pupil search region in an image in a vehicular environment

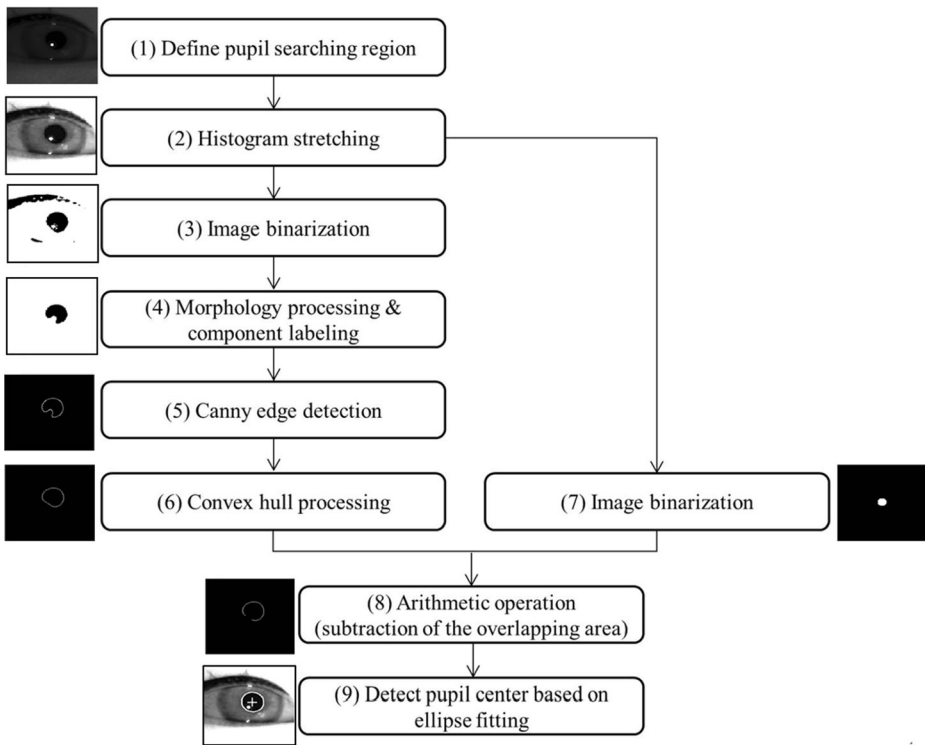


Fig. 4 Flowchart of proposed method used to detect the center of pupil

of corneal SR have to be accurately detected. Figure 4 shows a flowchart of the pupil detection method based on image processing algorithms [26] and detail explanations of this method can be referred in [30]. As stated above, unlike gaze detection in a general indoor desktop monitor environment, where subjects stare into a small monitor, the area drivers look at in a vehicular environment, including both side mirrors, is far more extensive and results in greater head and eye movements. This is why corneal SR is sometimes detected in the bright areas of the sclera or the skin instead of the dark regions of the pupil or the iris, or is lost entirely and is undetectable. Corneal SR is generally used in the PCCR method as a reference point to compensate for head movements when estimating the driver's gaze [56, 68]. As already mentioned, however, the corneal SR, or the reference point, can be lost when there is excessive head or eye movement in a vehicular environment; therefore, this study uses the MC as another reference point. As described in Section 3.3, points 39 and 42 of the detected facial feature points were designated as the location of the MC to be used for gaze detection. These points were used instead of points 36 and 45 because when the driver turns his or her head excessively, it is more likely that one of points 36 and 45 is rendered unobserved in the captured image than one of points 39 and 42. In [29], they proposed the method of tracking user's gaze position on frontal viewing home appliance as the interface of disabled people in indoor environment. For eye-tracking, they used the simple method of detecting pupil center and corneal specular reflection (SR). However, our system is used in car environment of outdoor, and more illumination changes exist than that in [29]. Therefore, the region of interest of eye is firstly defined by the detected eye features by dlib facial feature point tracking method

as shown in Fig. 3, and pupil center and corneal SR center are detected within this region. The performance of dlib facial feature point tracking method is not affected by the illumination changes in an external environment as shown in Table 2. In addition, as shown in Fig. 4, our eye-tracking method includes sophisticated procedure of reducing the detection error of pupil boundary caused by specular reflection (SR) hiding the boundary. For that, the steps (6) ~ (8) are newly proposed in our eye-tracking algorithm compared to [29]. In addition, different from [36], our method also detect medial canthus in order to cope with the case of severe head rotation where corneal SR disappears in the captured image.

3.5 Calculating gaze position by combining average calibration information and one-point calibration based on maximum entropy criterion

User calibration is the process of identifying the relationship between the monitor display and the region of the movements of the user's pupils obtained by having users look at certain spots on the monitor display (e.g., spots 4, 6, and 9). User calibration is also needed to compensate for the kappa angle between where the user is actually looking and where the pupil is directed, and for the fact that people have eyeballs of varying size. In a vehicular environment, however, drivers find it difficult to perform initial calibration because the spots to look at for the purpose of calibration cannot be marked inside the vehicle. Therefore, this study uses the average of calibration information obtained beforehand from 10 users, each of whom was asked to stare at nine pre-designated spots on a monitor ($M_1 - M_9$ in Fig. 5) in an indoor desktop monitor environment. For fair comparisons, these 10 people's data are not included in the data of 26 people which are used for performance evaluation in Section 4. The details of user calibration are as below.

As described in Section 3.4, the PCCR vector (the location of the center of the pupil is compensated for by that of the corneal SR) is calculated based on the detected location of the center and the corneal SR; then, based on this, four geometric transform matrices (Matrix 1–4) are found using nine pupil centers ($P_1 - P_9$) (compensated for by the location of the corneal SR), each acquired when the subjects stared at the nine spots ($M_1 - M_9$) on the monitor plane, as illustrated in Fig. 5 and Eq. (1) [30, 35].

However, the relationship between pupil subregion and monitor subregion is actually nonlinear due to the nonlinear movement of pupil center on the 3D eyeball, camera parameters, and user's position. Therefore, this relationship cannot be represent by the simple linear matrix of Eq. (1), and accurate parameters for this relationship is calculated based on maximum entropy criterion as shown in Eqs. (2) ~ (6). Detailed explanations are as follows. Figure 5 and Eq. (1) show an example of how such a geometric transform matrix is calculated:

$$\begin{bmatrix} M_{x1} & M_{x2} & M_{x3} & M_{x4} \\ M_{y1} & M_{y2} & M_{y3} & M_{y4} \end{bmatrix} = \begin{bmatrix} a & b & c & d \\ e & f & g & h \end{bmatrix} \begin{bmatrix} P_{x1} & P_{x2} & P_{x3} & P_{x4} \\ P_{y1} & P_{y2} & P_{y3} & P_{y4} \\ P_{x1}P_{y1} & P_{x2}P_{y2} & P_{x3}P_{y3} & P_{x4}P_{y4} \\ 1 & 1 & 1 & 1 \end{bmatrix} \quad (1)$$

Table 2 Error in the detection of center of pupil, corneal SR, and MC (unit: pixel)

	Pupil Center	Corneal SR	MC
Average Euclidean distance	3.2	3.5	3.1

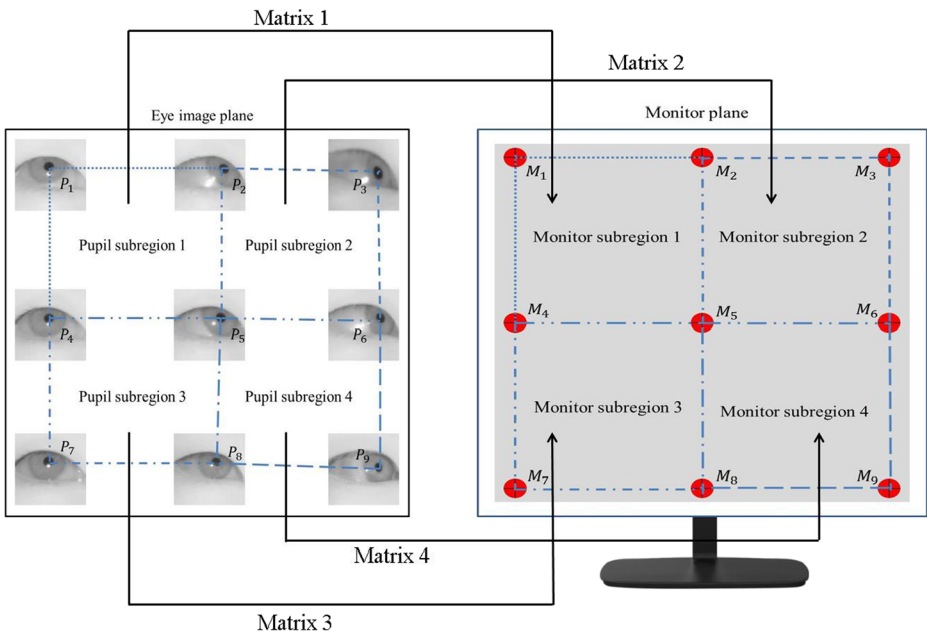


Fig. 5 Relationship between each pupil subregion and the monitor subregion

Where (M_{xi}, M_{yi}) and (P_{xi}, P_{yi}) are the x and y coordinates of M_i and P_i , respectively. Then, we can obtain the following two functions.

$$J(a, b, c, d) = 1 / \left[(Mx1 - (aPx1 + bPy1 + cPx1Py1 + d))^2 + (Mx2 - (aPx2 + bPy2 + cPx2Py2 + d))^2 + (Mx3 - (aPx3 + bPy3 + cPx3Py3 + d))^2 + (Mx4 - (aPx4 + bPy4 + cPx4Py4 + d))^2 + offset1 \right] \tag{2}$$

$$J(e, f, g, h) = 1 / \left[(My1 - (ePx1 + fPy1 + gPx1Py1 + h))^2 + (My2 - (ePx2 + fPy2 + gPx2Py2 + h))^2 + (My3 - (ePx3 + fPy3 + gPx3Py3 + h))^2 + (My4 - (ePx4 + fPy4 + gPx4Py4 + h))^2 + offset2 \right] \tag{3}$$

The goal of our research is to obtain the optimal parameters (a, b, c, d) and (e, f, g, h) , with which $J(a, b, c, d)$ and $J(e, f, g, h)$ are maximized, respectively. In Eqs. (2) and (3), *offset1* and *offset2* are the terms which make the denominator non-zero. From the eqs. (2) and (3), we can obtain the probability density function (PDF) of $J(a, b, c, d)$ and $J(e, f, g, h)$ as follows.

$$P(a, b, c, d) = J(a, b, c, d) / (J(a, b, c, d) + J(e, f, g, h)) \tag{4}$$

$$P(e, f, g, h) = J(e, f, g, h) / (J(a, b, c, d) + J(e, f, g, h)) \tag{5}$$

Then, based on the maximum entropy criterion [7], the goal of our research is to obtain the optimal parameters (a, b, c, d) and (e, f, g, h) ($P(a, b, c, d)$ and $P(e, f, g, h)$ are not biased but evenly fitted), with which $H(a, b, c, d, e, f, g, h)$ are maximized, as follows.

$$H(a, b, c, d, e, f, g, h) = -P(a, b, c, d) \log(P(a, b, c, d)) - P(e, f, g, h) \log(P(e, f, g, h)) \tag{6}$$

From that, we obtain the parameters of (a, b, c, d) and (e, f, g, h) , and user's gaze position (G_x, G_y) can be obtained the following equation with the extracted pupil center position (P'_x, P'_y) .

$$\begin{bmatrix} G_x \\ G_y \end{bmatrix} = \begin{bmatrix} a & b & c & d \\ e & f & g & h \end{bmatrix} \begin{bmatrix} P'_x \\ P'_y \\ P'_x P'_y \\ 1 \end{bmatrix} \quad (7)$$

The gaze (G_x, G_y) from the pupil center (P'_x, P'_y), compensated for by the location of the corneal SR, in the given image can be calculated. Furthermore, when the corneal SR cannot be found in the captured image due to excessive head rotation, gaze (G_x, G_y) is calculated from the pupil center (P'_x, P'_y) that is compensated for by the location of the MC. Based on this, if the center of the pupil (compensated for by the location of the corneal SR or the position of MC) is in pupil subregion 2 of Fig. 5, the parameters of Eq. (7) based on Matrix 2 are used to estimate gaze; if the center is in pupil subregion 4, those based on Matrix 4 is used instead. The gaze value for each direction is determined from one image frame. For example, ten gaze values are determined from ten image frames. As stated above, the proposed tracks MC points (points 39 and 42 in Fig. 3a) at the same time to be able to calculate the gaze even when the corneal SR cannot be found in the captured image due to excessive head rotation; it also calculates the four sets of parameters of Eq. (7) based on four geometric transform matrices by using nine additional pupil centers (compensated for by the location of the MC), acquired when the test subjects were asked to stare at nine spots ($M_1 \sim M_9$) on the monitor plane illustrated in Fig. 5. Instead of having subjects stare at the nine spots on the monitor again, the nine pupil centers (compensated for by the location of the MC) are calculated from the images captured when the subjects look at the nine spots on the monitor to calculate the PCCR vector (the location of the pupil center is compensated for by that of the corneal SR). In other words, user calibration is accomplished with each user staring once at the nine spots on an indoor desktop monitor. Hence, the method uses the four sets of parameters of Eq. (7) based on four geometric transform matrices based on the location of the corneal SR and additional four sets of parameters of Eq. (7) based on four matrices based on that of the MC, for a total of eight sets of parameters of Eq. (7). The points of gaze calculated using the pupil center compensated for by the location of the corneal SR or the MC are less influenced by head movements. As mentioned above, the proposed method calculates the average of the pupil location (each compensated for by the locations of the corneal SR or the MCs) for each of the nine pre-designated spots ($M_1 \sim M_9$ of Fig. 5) that are stared at by 10 users in an indoor desktop monitor environment beforehand, and uses it as user calibration information in a vehicular environment. Because pupil information was obtained from general users in an indoor desktop monitor environment, however, the margin of error increases when this information is used to calculate the gaze of the given vehicle drivers. This is due to differences in each driver's kappa angle and sitting position. To resolve this issue, the 10 users (whose data are not included in the data of 26 people which are used for performance evaluation in Section 4 for fair comparisons) also look into the NIR illuminator in the gaze detection system, installed at the bottom of the indoor desktop monitor, in advance. From that, we can obtain the average gaze position ($G_{x_cal_indoor}, G_{y_cal_indoor}$) of Eq. (7) of the 10 users. In addition, during the testing (see the details in Section 4.1), each people of 21 participant look once (at the initial step) at the NIR illuminator in the gaze detection system (one-point calibration),

attached to the dashboard as shown in Fig. 2. From that, we can obtain the gaze position $(G_{x_cal_vehicle}, G_{y_cal_vehicle})$ of Eq. (7) when each people performs one-point calibration. Our system then finds the difference of between these two points $((G_{x_cal_indoor}, G_{y_cal_indoor})$ and $(G_{x_cal_vehicle}, G_{y_cal_vehicle}))$ to compensate for the gaze position (G_x, G_y) of Eq. (7) when each people looks at one of the fifteen positions of vehicle of Fig. 6. This reduces the gap in the direction of the gaze due to different kappa angles and sitting positions of drivers, resulting in more accurate gaze detection in vehicles. Nevertheless, this gap cannot be completely removed. However, the distance between gazing positions (as shown in Fig. 6) are much larger than that in desktop monitor because of large viewing area in car, which has the effect on reducing the consequent gazing error in our experimental environment. The above-mentioned NIR illuminator uses 850 nm NIR LEDs, which emit weak light, rendering it appropriate for indicator in one-point calibration.

4 Experimental results

4.1 Experimental setups

Although there have been previous researches on gaze detection in car environments [3, 4, 10, 20, 43, 57, 58, 61, 65], there is few open database for NIR light-based gaze tracking of drivers in car environment. Therefore, we collected our database for NIR light-based gaze tracking of drivers in car environment, and we made our collected database open through [15]. Our gaze tracking device aforementioned in Section 3.2 was used for collecting database. As shown in Fig. 7, the database was collected from a total of 26 participants: 10 wearing nothing, 8 wearing only four kinds of glasses, 5 wearing only two kinds of sunglasses, and 3 wearing only hat. In addition, even the people

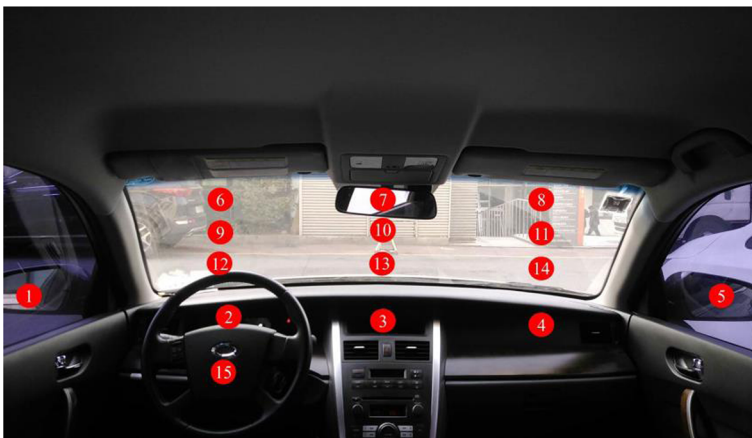


Fig. 6 Fifteen positions in the vehicle for measuring the accuracy of gaze detection

(wearing nothing) took various pose including putting one hand to cheek or using mobile phone as shown in Fig. 7 (e). Fifteen spots were designated to gaze at for the experiment, as shown in Fig. 6, and each participant stared at each spot during about three seconds and this procedure was iterated five times. Between each iteration, participant took a rest of about 2 min. A total of about 12 or 13 min were taken for each participant. When the participants were staring at each spot, they were told to act normally, as if they were actually driving and were not restrained to one position or given any special instructions to act in an unnatural manner. Due to the risks of car accidents, it is difficult to motivate the participants to stare at the 15 designated spots while actually driving for the experiment. Therefore, this study obtained images from various locations (from roads in daylight to a parking garage) in a real vehicle (model name of SM5 New Impression by Renault Samsung [54]) with its power on, but in park in order to create an environment most similar to when it is being driven (including factors like car vibration and external light). Moreover, to understand the influence of various kinds of external light on driver gaze detection, test data were acquired at different times of the day: in the morning, the afternoon, and at night. All algorithms of our method were implemented on Microsoft Visual Studio 2013 C++, and OpenCV (version 2.4.5) [49] library and Boost (version 1.55.0) library were used. Figure 7 shows some examples of images captured with the gaze detection system developed for this study in a vehicular environment. The research [36] uses more positions (18 positions) than our research (15 positions), and the

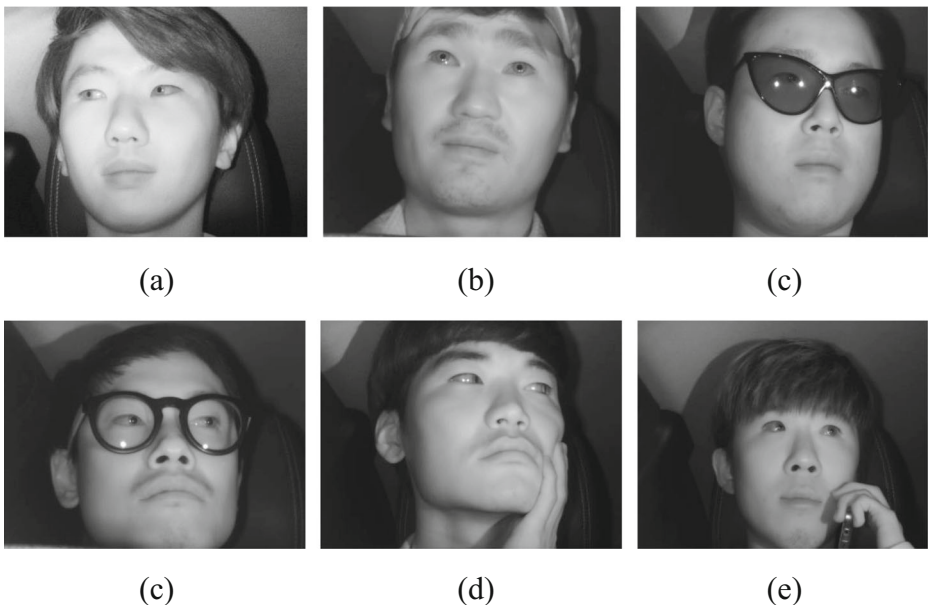


Fig. 7 Examples of images captured by gaze detection system in vehicular environment. Driver **a** wearing nothing, **b** wearing hat, **c** wearing sunglasses, **d** wearing glasses, and **e** putting one hand to cheek or using mobile phone

three additional positions are respectively the left position of region 1, the right position of region 5, and the upper position of region 6 of Fig. 6. The camera resolution of their system [36] is low with low illuminator power, and the driver's pupil center cannot be detected. Therefore, they tracked the driver's gaze position only by measuring head rotation (not eye rotation), and got the experimental data where drivers were instructed to rotate his or her head intentionally and sufficiently. In case that driver only moves eye to gaze at some position in car without rotating head (it is more often the case with driver while driving), their method cannot detect driver's gaze position. Different from them, our NIR solution can provide the clear pupil image for gaze detection by tracking eye movement in car environment while allowing the natural movement of driver's head. In addition, the cases of gazing at the additional three positions (the left position of region 1, the right position of region 5, and the upper position of region 6 of Fig. 6) do not frequently occur while driving [28, 52, 53, 64]. Therefore, in previous researches, they did not use these three positions for experiments, either [10, 21, 22, 60, 66]. Based on that, we did not use these three positions, and performed the experiments with the 15 positions of Fig. 6.

4.2 Accuracy of detecting pupil center, corneal SR, and MC

Figure 8 below shows examples of the driver's pupil center, corneal SR, and MC detected using the proposed method. The pupil center, corneal SR, and MC were accurately detected without the influence of external light, even though the drivers were moving their heads in various manners. Even when the corneal SR could not be found in the captured image due to a driver's excessive head rotation, the pupil center and the MC (points 39 and 42 in Fig. 3a) were accurately detected as shown in Fig. 8b.

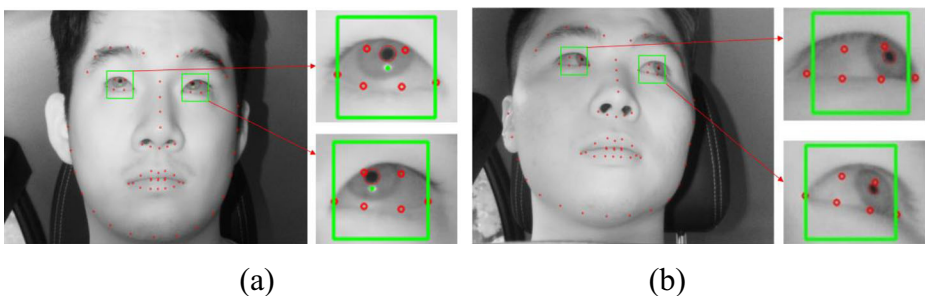


Fig. 8 Examples of detected points of pupil center, corneal SR, and MC **a** When corneal SR exists in image **b** When corneal SR disappears due to excessive rotation of driver's head. Red small circles, red big circle, and green point show the detected facial feature points, pupil boundary, and corneal SR, respectively. In our experiments, there was no case that pupil center, corneal SR center, and MC were not detected in all the experimental images. Therefore, it is difficult to measure the false recognition rate because there is no ground criterion of correct and incorrect detection of pupil. However, their detected positions show error compared to accurate positions. Therefore, we measured the accuracies based on the following method. Table 2 below shows the error in detecting the pupil center, the corneal SR, and the MC using the method proposed in this study. The error was below 3.5 pixels, indicating high accuracy of detection. The detection error is shown as the average Euclidean distance between the manually marked locations of the pupil center, the corneal SR, and the MC, and the locations detected using this study's algorithm

4.3 Accuracy of gaze detection

The subsequent experiment estimated gaze detection accuracy of the 15 points shown in Fig. 6. Because no pixel location exists in the vehicle in Fig. 6, unlike on a desktop monitor, the following method was used to measure the accuracy of gaze detection. As the 1st fold evaluation, among the data of 26 participants, the data of 5 people were randomly selected for training, and the remained data of 21 people were used for testing. Then, in the 2nd fold evaluation, among the data of 21 people used for testing in the 1st fold evaluation, the data of 5 people were randomly selected for training, and the remained data of 21 people were used for testing. From these evaluations of two-fold cross-validation, we obtained the average accuracy for performance evaluation. Same scheme were also used for measuring the error in the detection of center of pupil, corneal SR, and MC of Table 2. With the trained data of 5 people, the average gaze value ((G_x, G_y) of Eq. (7) compensated by one-point calibration) of each gaze position of the 15 positions of Fig. 6 was determined as the position's reference gaze value. Using this method, 15 reference gaze values ($(G_{x1}, G_{y1}), (G_{x2}, G_{y2}), \dots (G_{x15}, G_{y15})$ of Eq. (7) compensated by one-point calibration) were calculated for all the 15 gaze positions of Fig. 6. Then, using the Euclidean distance between the gaze calculated from the testing data and the coordinates of these 15 reference gaze values, one gaze position with the minimum distance was determined as the driver's gaze, of the 15 positions in Fig. 6. Accuracy was measured using the strictly correct estimation rate (SCER) and the loosely correct estimation rate (LCER). The SCER indicates the accuracy of the corresponding spot (gaze position) while the LCER specifies the accuracy of the corresponding spot as well as spots in the vicinity. Table 3 shows the accuracies of gaze detection with all the users. As seen in Table 3, the average SCER was approximately 79.3% while the average LCER was approximately 96%. Table 4 presents the accuracies in case of users wearing nothing, glasses, sunglasses, and hat, respectively. As shown in Table 4, the accuracies in case of users wearing nothing are highest whereas those in case of users wearing sunglasses are lowest. The reason why the accuracies in case of users wearing sunglasses are lowest is that the visibility of eye region with some sunglasses reduces because the sunglasses reduces the NIR light of our gaze detection system to eye as shown in Fig. 7c, which degrades the accuracy of detection of

Table 3 SCER and LCER of proposed method with all the users

Target Zone	Neighbors	SCER (%)	LCER (%)
1	2, 12, 15	97.34	99.73
2	1, 3, 12, 13, 15	95.79	99.84
3	2, 4, 12, 13, 14, 15	91.37	98.58
4	3, 5, 13, 14	73.27	92.6
5	4, 14	80.21	96.01
6	7, 9, 10	94.67	98.58
7	6, 8, 9, 10, 11	78.24	98.75
8	7, 10, 11	68.45	90.92
9	6, 7, 10, 12, 13	91.54	97.32
10	6, 7, 8, 9, 11, 12, 13, 14	79.46	98.65
11	7, 8, 10, 13, 14	72.31	91.43
12	1, 2, 3, 9, 10, 13	86.55	96.8
13	2, 3, 4, 9, 10, 11, 12, 14	71.29	95.05
14	3, 4, 5, 10, 11, 13	60.61	93.21
15	1, 2, 3	48.13	92.9
<i>Average</i>		79.28	96.02

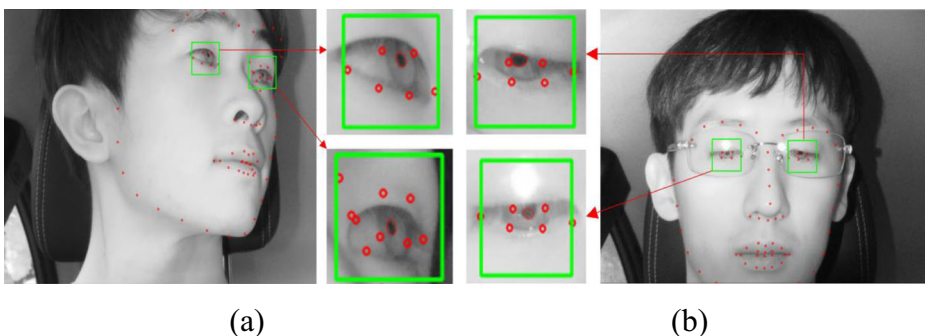
Table 4 SCER and LCER of proposed method in case of users wearing nothing, glasses, sunglasses, and hat, respectively

Target Zone	Neighbors	Wearing nothing		Wearing glasses		Wearing sunglasses		Wearing hat	
		SCER (%)	LCER (%)	SCER (%)	LCER (%)	SCER (%)	LCER (%)	SCER (%)	LCER (%)
1	2, 12, 15	99.64	100	94.7	98.91	95	100	100	100
2	1, 3, 12, 13, 15	100	100	95.76	99.35	90.08	100	97.31	100
3	2, 4, 12, 13, 14, 15	100	100	86.19	97.62	84.36	96.7	94.93	100
4	3, 5, 13, 14	83.79	92.62	76.73	94.5	51.91	93.61	80.66	89.67
5	4, 14	79.03	93.86	77.9	98.82	63.9	91.3	100	100
6	7, 9, 10	100	100	94.19	100	84.49	99.8	100	100
7	6, 8, 9, 10, 11	83.83	100	81.64	95.02	62.5	100	85	100
8	7, 10, 11	77.96	91.39	76.96	93.33	58.86	88.96	60	90
9	6, 7, 10, 12, 13	94.16	100	94.79	100	83.5	91.33	93.73	100
10	6, 7, 8, 9, 11, 12, 13, 14	86.67	98.16	76.91	99.16	69.65	97.29	84.6	100
11	7, 8, 10, 13, 14	76.44	96	70.51	89.79	64.37	91.77	77.91	88.14
12	1, 2, 3, 9, 10, 13	91.32	98.97	86.61	100	78.89	91.64	89.36	96.6
13	2, 3, 4, 9, 10, 11, 12, 14	84.77	99.61	69.63	93.79	54.7	90.32	80.75	96.47
14	3, 4, 5, 10, 11, 13	77.25	93.28	48.61	92.6	43.96	89.58	73.65	97.39
15	1, 2, 3	59.61	100	41.07	90.44	40	89.48	51.83	91.67
<i>Average</i>		86.3	97.59	78.15	96.22	68.41	94.12	84.65	96.66

pupil and corneal SR centers, and consequent accuracies of gaze detection. The reason why the accuracies in case of users wearing glasses are lower than those in case of wearing nothing is that the SR on the glasses surface can hide the pupil and corneal SR regions, which degrades the accuracy of detection of pupil and corneal SR centers, and consequent accuracies of gaze detection. The accuracies in case of wearing hat are similar to those in case of wearing nothing, which presents that our method can detect the facial feature points, pupil and corneal SR centers irrespective of hat.

Figure 9 below shows examples of images with incorrect estimation based on the SCER and LCER. As shown in Fig. 9a, incorrect estimation usually occurred when the driver was looking at point 11 in Fig. 6. This is because the MC points (39 and 42 of Fig. 3a) of the detected facial feature points were used to calculate the gaze when the corneal SR could be found, but increasing the error in detecting the MC points in turn increased the margin of error in tracking gaze.

As shown in Fig. 9b, incorrect estimation usually occurred when the driver was looking far too downward, making the eyelid

**Fig. 9** Examples of incorrect estimation based on **a** SCER and **b** LCER

occlude the upper part of the pupil, which results in inaccurate detection of the pupil center. Further, the location of the MC (points 39 and 42 of Fig. 3a), which was used when the corneal SR could not be found in the captured eye image, was detected inaccurately as well, and these two errors (errors in detecting the pupil center as well as the location of the MCs) caused another error in detecting the gaze, even when based on the LCER.

4.4 Comparisons with other methods

As the next experiment, we compared the accuracies by our method with those by previous method [10]. Previous method used deep learning method based on convolutional neural network (CNN) for estimating gaze position in car environment. For fair comparisons, same training data used in our method were used for CNN training, and accuracies were measured with same testing data used in our method. Like our method, accuracy was measured twice by exchanging training and testing data (two-fold cross-validation), and the average accuracy, which was estimated twice, became the final accuracy of gaze detection. As shown in Tables 3 and 5, our method outperforms the previous method. The reason why the accuracies by previous method are lower than those by our method is that the number of target zone (15 zones) in our research is much larger than that (8 zones) in [10]. In addition, the research [10] did not considered the cases when a driver wore hat, a glasses, a sunglasses, and took various pose including putting one hand to cheek or using mobile phone as shown in Fig. 7.

In addition, Tables 6 and 7 show the comparative accuracies by previous study [61] and a commercial gaze detection system, the Tobii EyeX System [62] with those by our method. Because the user calibration information in an indoor desktop monitor environment was not acquired beforehand for use in a vehicular environment for the Tobii system, the system was attached to the same location in the vehicle as the gaze detection system used in this study in Fig. 2, and calibration was performed on four spots where user calibration for the Tobii system was possible (points in the vicinity of 2, 3, 9, 10 in Fig. 6). Calibration was only performed on the four spots where user calibration for the Tobii system is possible inside a vehicle because the system could not acquire user calibration information when there was excessive eye or head rotation. Because the Tobii system was designed for use on a small indoor desktop monitor, it

Table 5 SCER and LCER by previous method [10]

Target Zone	Neighbors	SCER (%)	LCER (%)
1	2, 12, 15	68.6	94
2	1, 3, 12, 13, 15	65.55	98.05
3	2, 4, 12, 13, 14, 15	57.75	84.8
4	3, 5, 13, 14	68.35	96.9
5	4, 14	68.2	87.65
6	7, 9, 10	55.5	91.75
7	6, 8, 9, 10, 11	62.45	79.8
8	7, 10, 11	62.55	78.25
9	6, 7, 10, 12, 13	64.75	87
10	6, 7, 8, 9, 11, 12, 13, 14	51.5	90
11	7, 8, 10, 13, 14	68.55	94.35
12	1, 2, 3, 9, 10, 13	60.7	77.8
13	2, 3, 4, 9, 10, 11, 12, 14	68.5	81.25
14	3, 4, 5, 10, 11, 13	72.7	95
15	1, 2, 3	62	80.55
<i>Average</i>		63.84	87.81

Table 6 Comparison of SCER between proposed method and a previous method (unit: %)

	Proposed Method	Previous Method [61]
Average SCER	89.5	88.2

yielded no results in gaze detection for points 1, 4, 5, 6, 7, 8, 11, 14, and 15 in Fig. 6 when used for an experiment in a vehicle, which had an extensive area for gazing. Therefore, Table 7 shows the average SCER of the Tobii system for all points, excluding ones with no results in gaze detection. That is, we show the SCERs by our method and Tobii system in Table 7 on same gaze zones excluding the points 1, 4, 5, 6, 7, 8, 11, 14, and 15 in Fig. 6. As seen in Tables 6 and 7, the SCER of the proposed method was higher than those of the previous method [61] and the Tobii system. The reason why their accuracy [61] is lower than ours is that they did not detect pupil center but iris center for calculating gaze position. However, our method detects pupil center, which shows better accuracy. The reason why the accuracy of Tobii system [62] is lower than ours is that their system is for gaze tracking in small-sized desktop monitor environment of indoor, which does not include various illumination changes of outdoor.

Although there are various previous researches on gaze detection in car environment, most of their algorithm are not opened. Therefore, we selected these three methods [10, 61, 62] because their experimental gazing regions are comparable to ours [61], we can implement their algorithm [10], and the actual gaze position can be produced [62]. In case of the increased number of target zones, the changes in eye and head movement when driver gazes at each zone decreases in the captured image, and classification complexities of each gaze zone increase, which reduce the final accuracy of gaze detection. As the next experiment, we measured the accuracies by the methods [29, 36] as shown in Table 8. By comparing the Tables 3 and 8, we can find that our method outperforms the previous methods [29, 36]. From that, we can confirm that proposed maximum entropy criterion-based calculation of gaze position is superior to the geometric transform-based method of [29]. In addition, because the method [36] uses only head rotation for measuring driver's gaze position without considering eye movement, their accuracies are lower than ours.

For comparisons with related works, we compared the accuracies by our method and those by previous method [10, 36, 61, 62] as shown in Tables 5, 6, 7 and 8. All these previous researches do not provide the codes of their algorithms, we re-implemented their algorithms by ourselves and performed the experiments with our datasets as shown in Tables 5, 6, 7 and 8.

4.5 Processing time

Finally, Table 9 shows the processing time of the proposed method. Every algorithm was executed in the laptop computer aforementioned in Section 3.2, and the total processing time was approximately 28.6 ms as shown in Table 9. This means that the proposed method can operate at a speed of approximately 35 frames per second (1000/28.6).

Table 7 Comparisons of SCER between proposed method and the Tobii system (unit: %)

	Proposed Method	Tobii System [62]
Average SCER	86	73

Table 8 SCER and LCER by previous methods

Target Zone	Neighbors	[29]		[36]	
		SCER (%)	LCER (%)	SCER (%)	LCER (%)
1	2, 12, 15	90.1	94.3	84.5	98.1
2	1, 3, 12, 13, 15	88.7	94.2	47.3	98.3
3	2, 4, 12, 13, 14, 15	86.4	93.1	42.5	96.1
4	3, 5, 13, 14	71.1	88.5	55.8	81.2
5	4, 14	73.8	90.3	10.2	54.5
6	7, 9, 10	87.7	92.8	60.1	93.7
7	6, 8, 9, 10, 11	77.2	92.7	35.7	87.4
8	7, 10, 11	59.3	84.1	37.1	81.1
9	6, 7, 10, 12, 13	85.2	91.2	65.2	97.2
10	6, 7, 8, 9, 11, 12, 13, 14	72.7	92.4	47.3	95.4
11	7, 8, 10, 13, 14	67.4	85.7	29.1	82.5
12	1, 2, 3, 9, 10, 13	80.9	91.3	51.4	96.4
13	2, 3, 4, 9, 10, 11, 12, 14	65.9	90.2	21.5	95.1
14	3, 4, 5, 10, 11, 13	54.2	87.9	10.1	79.8
15	1, 2, 3	41.3	87.1	85.4	100
<i>Average</i>		73.5	90.4	45.5	89.1

4.6 Experiments in driving environments

We performed the additional experiments in real driving practice. To measure the accuracy, driver should gaze at 15 designated spots, but he or she is difficult to stare at the spots while actually driving. Therefore, we collected the data from the participants (sitting at the side seat of driver) gazing at 15 spots while the driver was actually driving the car as shown in Fig. 10. Each spot was indicated by attached yellow and small tape which did not distract driver's attention as shown in Figs. 10a and b. Our gaze tracking device was attached in front of dashboard (close to spot 2 as shown in Fig. 2) in previous experiments (Sections 4.1 ~ 4.5). However, because there is no dashboard in front of the participants (sitting at the side seat of driver) and it is difficult to attach our device at the spot 4 of Fig. 10a, we attached it at the position close to spot 14 as shown in Fig. 10a, from which we can show our experimental results in various positions of our gaze detection device in car. Except for this, the conditions of data acquisition were similar to those from driver in previous experiments (Sections 4.1 ~ 4.5). In order to check our method in various car environments, we used different car (model name of Daewoo Lacetti Premiere by Chevrolet [12]) from that used in the experiments of Sections 4.1 ~ 4.5. The database was collected from a total of 10 participants: 3 wearing nothing, 3 wearing only three kinds of glasses, 2 wearing only two kinds of sunglasses, and 2 wearing only hat. In addition, even the people (wearing nothing) took various pose including putting one hand to cheek or using mobile phone. Each participant stared at each spot during about three seconds and this procedure was iterated five times. Between each iteration, participant

Table 9 Processing time of proposed method (unit: ms)

	Processing Time
Detecting facial feature points (Steps (1) ~ (4) of Fig. 1)	10
Detecting pupil, corneal SR, and MC (Steps (5) ~ (7) of Fig. 1)	17
Calculating gaze position (Step (8) of Fig. 1)	1.6
Total processing time	28.6

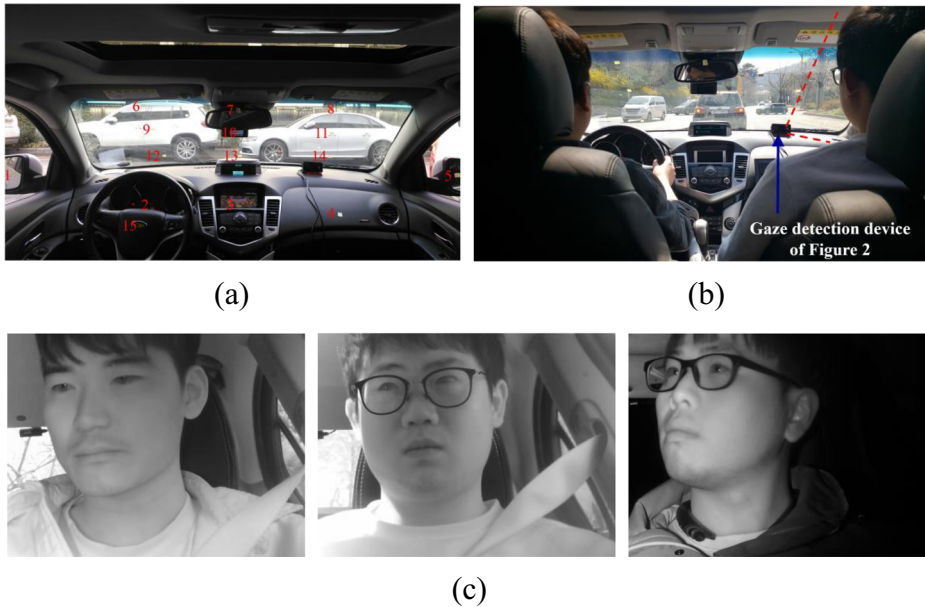


Fig. 10 Experiments in driving environments. **a** 15 gaze spots. **b** Experimental setup. **c** Examples of captured images (while gazing at spot 3 in the morning (left), spot 10 in the afternoon (middle), spot 6 at night (right) of Fig. 6)

took a rest of about 2 min. A total of about 12 or 13 min were taken for each participant. When the participants were staring at each spot, they were told to act normally, as if they were actually driving and were not restrained to one position or given any special instructions to act in an unnatural manner. To understand the influence of various kinds of external light on driver gaze detection, test data were acquired at different times of the day: in the morning, the afternoon, and at night, and they were collected while driving on various roads. We also made this collected database open through [15] in order to enable other researchers to compare the performance of their system with our database. Table 10 shows the SCER and LCER by our gaze tracking method, and they are similar to those of Table 3. From that, we can confirm the effectiveness of our gaze tracking method in real driving environments.

4.7 Discussions on the usefulness for the NIR illuminator

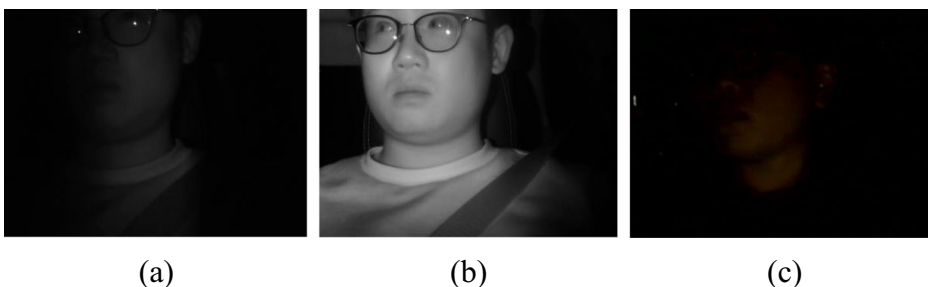
The usefulness for the NIR illuminators in our gaze detection system is as follows. First, the NIR illuminator in the gaze detection system helps capture the driver's facial image without being influenced by the changing light at night and day.

As shown in the image captured at night of Fig. 11a, the driver's face image captured without NIR illuminator using our NIR camera-based gaze detecting device is so dark that pupil center cannot be detected for gaze detection. However, as shown in Fig. 11b, pupil area is distinctive in the captured image at night with NIR illuminator using our NIR camera-based gaze detecting device, which enables the pupil center to be detected. Although it can be considered to use visible light illuminator, it causes severe dazzling to driver's eye and cannot be used in car environments. Without visible light illuminator, even with visible light camera, the pupil center cannot be detected due to too dark image at night as shown in Fig. 11c. Second, the center of corneal SR is important to calculate driver's gaze position because the

Table 10 SCER and LCER of proposed method in case of real driving environments

Target Zone	Neighbors	SCER (%)	LCER (%)
1	2, 12, 15	98.61	99.27
2	1, 3, 12, 13, 15	76.33	91.4
3	2, 4, 12, 13, 14, 15	92.85	98.18
4	3, 5, 13, 14	96.2	99.52
5	4, 14	95.94	99.07
6	7, 9, 10	70.48	89.14
7	6, 8, 9, 10, 11	79.39	98.1
8	7, 10, 11	95.6	97.56
9	6, 7, 10, 12, 13	73.72	90.81
10	6, 7, 8, 9, 11, 12, 13, 14	82.46	99.92
11	7, 8, 10, 13, 14	90.36	96.7
12	1, 2, 3, 9, 10, 13	66.68	89.99
13	2, 3, 4, 9, 10, 11, 12, 14	73.54	92.19
14	3, 4, 5, 10, 11, 13	90.06	98.53
15	1, 2, 3	57.58	90.87
<i>Average</i>		82.65	95.42

relative position of pupil center based on the center of corneal SR is used for calculating the gaze position. If only the pupil center is used for calculating gaze position, its movement is affected by driver's head movement, which degrades the gaze detection accuracy. Therefore, all the previous methods of gaze detection based on PCCR vector used the center of corneal SR to compensate the head movement for calculating gaze position [8, 9, 30, 35, 56, 62, 67, 68]. In our gaze detection system, the NIR illuminator was placed to the left of the camera as shown in Fig. 2, which can generate the corneal SR in eye image as shown in Fig. 8a, and driver's gaze position is calculated using both corneal SR and pupil center. Third, because more accurate gaze position can be detected by using pupil center instead of iris center [8, 9, 30, 35, 56, 62, 67, 68], our research utilizes the pupil center. The color of the iris (regions between pupil and sclera) is usually dependent on the amount of melanin in the anterior border layer [59]. The reflectance of the iris is also affected by the amount of melanin in the anterior border layer. The reflectance of melanin and that of the iris slightly increases as the illumination wavelength increases from visible light to NIR light [34]. That is, without NIR light, the reflectance of iris becomes small and its brightness decreases in image, which makes iris dark especially in case of (dark) brown irises (Asian people) and makes it difficult to discriminate pupil and iris region. With NIR light, the brightness of iris increases in image, which makes it easy to detect pupil region as shown in Fig. 8a. Therefore, all the previous methods using PCCR-based gaze detection used NIR illuminator [8, 9, 30, 35, 56, 62, 67, 68]. In our system,

**Fig. 11** Comparison of eye images at night **a** without NIR light, **b** with NIR light, and **c** with visible light camera

using the NIR LEDs at a wavelength of 850 nm prevented uncomfortable situations, such as the driver being blinded by the light while driving. An 850 nm band pass filter (BPF) was also mounted on the camera's lens to minimize interference due to sunlight [1].

5 Conclusions

This study proposed a new gaze detection method in a car environment. Through experiments, it was evident that the gaze detection accuracy of the proposed method was higher than that of previous methods and a commercial system at the fast processing speed. In addition, the effectiveness of our gaze tracking method was shown in real driving environments, which validated the convergence of the proposed algorithm. Because our gaze tracking camera is positioned at dashboard and captures driver's face image by upper direction (not frontal direction), more error occurs in vertical neighbors than those in horizontal neighbors. In addition, the errors increase when driver rotates his or her face severely to gaze at the position far from our gaze detection camera. These phenomena are also observed in the results of Table 5 by previous methods [10] and research [36]. These errors can be reduced by using another camera in upper or side area like previous method [61]. In future work, the method without any calibration would be researched by using deep learning-based method. In addition, emotion recognition based on the change of gaze position would be researched.

Acknowledgments This research was supported by the Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (NRF-2018R1D1A1B07041921), by the Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (NRF-2017R1D1A1B03028417), and by the National Research Foundation of Korea (NRF) grant funded by the Korea government (Ministry of Science and ICT) (NRF-2017R1C1B5074062).

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

References

1. 850nm CWL, 12.5mm Dia. Hard Coated OD 4 50nm Bandpass Filter. Available online: <https://www.edmundoptics.co.kr/optics/optical-filters/bandpass-filters/hard-coated-od4-50nm-bandpass-filters/84778/> (accessed on 7 March 2017)
2. Abtahi S, Hariri B, Shimohammadi S (2011) Driver drowsiness monitoring based on yawning detection. In Proceedings of IEEE International Instrumentation and Measurement Technology Conference, Binjiang, pp. 1–4
3. Ahlstrom C, Kircher K, Kircher A (2013) A gaze-based driver distraction warning system and its effect on visual behavior. IEEE Trans Intell Transp Syst 14:965–973
4. Batista JP (2005) A real-time driver visual attention monitoring system. In Proceedings of the 2nd Iberian Conference on Pattern Recognition and Image Analysis, Estoril, pp. 200–208
5. Bergen JR, Anandan P, Hanna KJ, Hingorani R (1992) Hierarchical model-based motion estimation. In Proceedings of European Conference on Computer Vision, Santa Margherita Ligure, pp. 237–252
6. Centre (geometry). Available online: [https://en.wikipedia.org/wiki/Centre_\(geometry\)](https://en.wikipedia.org/wiki/Centre_(geometry)) (accessed on 7 March 2017)
7. Cheng HD, Chen JR, Li J (1998) Threshold selection based on fuzzy c-partition entropy approach. Pattern Recogn 31:857–870
8. Cho D-C, Kim W-Y (2013) Long-range gaze tracking system for large movements. IEEE Trans Biomed Eng 60:3432–3440

9. Cho CW, Lee HC, Gwon SY, Lee JM, Jung D, Park KR, Kim H-C, Cha J (2014) Binocular gaze detection method using a fuzzy algorithm based on quality measurements. *Opt Eng* 53:053111-1–053111-22
10. Choi I-H, Hong SK, Kim Y-G (2016) Real-time categorization of driver's gaze zone using the deep learning techniques. In *Proceedings of IEEE International Conference on Big Data and Smart Computing*, Hong Kong, pp. 143–148
11. Cui J, Liu Y, Xu Y, Zhao H, Zha H (2013) Tracking generic human motion via fusion of low- and high-dimensional approaches. *IEEE Trans Syst Man Cybern Part A-Syst Hum* 43:996–1002
12. Daewoo Lacetti Premiere. Available online: https://en.wikipedia.org/wiki/Chevrolet_Cruze (accessed on 2 April 2018)
13. Dlib C++ Library (Real-time face pose estimation). Available online: <http://blog.dlib.net/2014/08/real-time-face-pose-estimation.html> (accessed on 7 March 2017)
14. Dong Y, Hu Z, Uchimura K, Murayama N (2011) Driver inattention monitoring system for intelligent vehicles: a review. *IEEE Trans Intell Transp Syst* 12:596–614
15. Dongguk Single Camera-based Driver Database (DSCD-DB1). Available online: <http://dm.dgu.edu/link.html> (accessed on 27 July 2017)
16. Durna Y, Ari F (2017) Design of a binocular pupil and gaze point detection system utilizing high definition images. *Appl Sci-Basel* 7:1–16
17. ELP-USB500W02M-L36. Available online: <http://www.elpcctv.com/usb20-5mp-usb-camera-module-ov5640-color-cmos-sensor-36mm-lens-p-216.html> (accessed on 7 March 2017)
18. Eye tracking. Available online: https://en.wikipedia.org/wiki/Eye_tracking (accessed on 7 March 2017)
19. Franchak JM, Kretch KS, Soska KC, Adolph KE (2011) Head-mounted eye tracking: a new method to describe infant looking. *Child Dev* 82:1738–1750
20. Fridman L, Lee J, Reimer B, Victor T (2016) "Owl" and "Lizard": Patterns of head pose and eye pose in driver gaze classification. *IET Comput Vis* 10:308–313
21. Fridman L, Lee J, Reimer B, Victor T (2016) Owl and lizard: Patterns of head pose and eye pose in driver gaze classification. *IET Comput Vis* 10:308–314
22. Fu X, Guan X, Peli E, Liu H, Luo G (2013) Automatic calibration method for driver's head orientation in natural driving environment. *IEEE Trans Intell Transp Syst* 14:303–312
23. Funke G, Greenlee E, Carter M, Dukes A, Brown R, Menke L (2016) Which eye tracker is right for your research? Performance evaluation of several cost variant eye trackers. In *Proceedings of the Human Factors and Ergonomics Society 2016 Annual Meeting*, Washington, DC, pp. 1240–1244
24. García I, Bronte S, Bergasa LM, Almazán J, Yebes J (2012) Vision-based drowsiness detector for real driving conditions. In *Proceedings of IEEE Intelligent Vehicles Symposium*, Alcalá de Henares, 618–623
25. Ghosh S, Nandy T, Manna N (2015) Real time eye detection and tracking method for driver assistance system. *Advancements of Medical Electronics*. Springer, New Delhi, pp. 13–25
26. Gonzalez RC, Woods RE (2010) *Digital Image Processing*, 3rd edn. Prentice Hall, New Jersey
27. Hansen DW, Ji Q (2010) In the eye of the beholder: a survey of models for eyes and gaze. *IEEE Trans Pattern Anal Mach Intell* 32:478–500
28. Ikonen T, Pekkanen J, Lappi O (2015) Driver gaze behavior is different in normal curve driving and when looking at the tangent point. *PLoS One* 10:1–19
29. Jang JW, Heo H, Bang JW, Hong HG, Naqvi RA, Nguyen PH, Nguyen DT, Lee MB, Park KR (2018) Fuzzy-based estimation of continuous Z-distances and discrete directions of home appliances for NIR camera-based gaze tracking system. *Multimed Tools Appl* 77:11925–11955
30. Jung D, Lee JM, Gwon SY, Pan W, Lee HC, Park KR, Kim H-C (2016) Compensation method of natural head movement for gaze tracking system using an ultrasonic sensor for distance measurement. *Sensors* 16:1–20
31. Kar A, Corcoran P (2017) A review and analysis of eye-gaze estimation systems, algorithms and performance evaluation methods in consumer platforms. *IEEE Access* 5:16495–16519
32. Kazemi V, Sullivan J (2014) One millisecond face alignment with an ensemble of regression trees. In *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, Columbus, 1867–1874
33. Khushaba RN, Kodagoda S, Lal S, Dissanayake G (2011) Driver drowsiness classification using fuzzy wavelet-packet-based feature-extraction algorithm. *IEEE Trans Biomed Eng* 58:121–131
34. Koblova EV, Bashkatov AN, Genina EA, Tuchin VV, Bakutkin VV (2005) Estimation of melanin content in iris of human eye. *Proc SPIE* 5688:302–311
35. Lee JW, Heo H, Park KR (2013) A novel gaze tracking method based on the generation of virtual calibration points. *Sensors* 13:10802–10822
36. Lee SJ, Jo J, Jung HG, Park KR, Kim J (2011) Real-time gaze estimator based on driver's head orientation for forward collision warning system. *IEEE Trans Intell Transp Syst* 12:254–267
37. Lee B-G, Lee B-L, Chung W-Y (2014) Mobile healthcare for automatic driving sleep-onset detection using wavelet-based EEG and respiration signals. *Sensors* 14:17915–17936

38. Li G, Chung W-Y (2015) A context-aware EEG headset system for early detection of driver drowsiness. *Sensors* 15:20873–20893
39. Li Z, Li SE, Li R, Cheng B, Shi J (2017) Online detection of driver fatigue using steering wheel angles for real driving conditions. *Sensors* 17:1–12
40. Li Y, Xue F, Fan X, Qu Z, Zhou G (2018) Pedestrian walking safety system based on smartphone built-in sensors. *IET Commun* 12:751–758
41. Li Y, Xue F, Feng L, Qu Z (2017) A driving behavior detection system based on a smartphone's built-in sensor. *Int J Commun Syst* 30:1–13
42. Li Y, Zhou G, Li Y, Shen D (2016) Determining driver phone use leveraging smartphone sensors. *Multimed Tools Appl* 75:16959–16981
43. Liang Y, Reyes ML, Lee JD (2007) Real-time detection of driver cognitive distraction using support vector machines. *IEEE Trans Intell Transp Syst* 8:340–350
44. Liu L, Cheng L, Liu Y, Jia Y, Rosenblum DS (2016) Recognizing complex activities by a probabilistic interval-based model. In *Proceedings of the 13th AAAI Conference on Artificial Intelligence*, Phoenix, pp. 1266–1272
45. Liu Y, Cui J, Zhao H, Zha H (2012) Fusion of low-and high-dimensional approaches by trackers sampling for generic human motion tracking. In *Proceedings of the 21st International Conference on Pattern Recognition*, Tsukuba, pp. 898–901
46. Liu Y, Nie L, Han L, Zhang L, Rosenblum DS (2015) Action2Activity: Recognizing complex activities from sensor data. In *Proceedings of the 24th International Joint Conference on Artificial Intelligence*, Buenos Aires, pp. 1617–1623
47. Liu Y, Nie L, Liu L, Rosenblum DS (2016) From action to activity: sensor-based activity recognition. *Neurocomputing* 181:108–115
48. Noris B, Keller J-B, Billard A (2011) A wearable gaze tracking system for children in unconstrained environments. *Comput Vis Image Underst* 115:476–486
49. OpenCV. Available online: <http://opencv.org/> (accessed on 7 March 2017)
50. Purkinje Images. Available online: https://en.wikipedia.org/wiki/Purkinje_images (accessed on 7 March 2017)
51. Rantanen V, Vanhala T, Tuisku O, Niemenlehto P-H, Verho J, Surakka V, Juhola M, Lekkala J (2011) A wearable, wireless gaze tracker with integrated selection command source for human-computer interaction. *IEEE Trans Inf Technol Biomed* 15:795–801
52. Ren Y-Y, Li X-S, Zheng X-L, Li Z, Zhao Q-C (2015) Analysis of drivers' eye-movement characteristics when driving around curves. *Discret Dyn Nat Soc* 2015:1–10
53. Ren Y-Y, Li X-S, Zheng X-L, Li Z, Zhao Q-C, Chen X-X (2014) Analysis and modeling of driver's gaze trajectories in curves driving. *Adv Mech Eng* 2014:1–12
54. Renault Samsung SM5. Available online: https://en.wikipedia.org/wiki/Renault_Samsung_SM5 (accessed on 7 March 2017)
55. Sahayadhas A, Sundaraj K, Murugappan M (2012) Detecting driver drowsiness based on sensors: a review. *Sensors* 12:16937–16953
56. Shih S-W, Liu J (2004) A novel approach to 3-D gaze tracking using stereo cameras. *IEEE Trans. Syst. Man Cybern. Part B-Cybern* 34:234–245
57. Smith P, Shah M, da Vitoria Lobo N (2000) Monitoring head/eye motion for driver alertness with one camera. In *Proceedings of International Conference on Pattern Recognition*, Barcelona, pp. 636–642
58. Smith P, Shah M, da Vitoria Lobo N (2003) Determining driver visual attention with one camera. *IEEE Trans Intell Transp Syst* 4:205–218
59. Sturm RA, Frudakis TN (2004) Eye colour: portals into pigmentation genes and ancestry. *Trends Genet* 20:327–332
60. Tawari A, Chen KH, Trivedi MM (2014) Where is the driver looking: Analysis of head, eye and iris for robust gaze zone estimation. In *Proceedings of IEEE International Conference on Intelligent Transportation Systems*, Qingdao, pp. 988–994
61. Tawari A, Trivedi MM (2014) Robust and continuous estimation of driver gaze zone by dynamic analysis of multiple face videos. In *Proceedings of IEEE Intelligent Vehicles Symposium*, Dearborn, pp. 344–349
62. Tobii. Available online: <http://www.tobii.com> (accessed on 7 March 2017)
63. Tsukada A, Shino M, Devyver M, Kanade T (2011) Illumination-free gaze estimation method for first-person vision wearable device. In *Proceedings of IEEE International Conference on Computer Vision Workshops*, Barcelona, pp. 2084–2091
64. van Leeuwen PM, Happee R, de Winter JCF (2015) Changes of driving performance and gaze behavior of novice drivers during a 30-min simulator-based training. *Procedia Manufacturing* 3:3325–3332
65. Vicente F, Huang Z, Xiong X, De la Torre F, Zhang W, Levi D (2015) Driver gaze tracking and eyes off the road detection system. *IEEE Trans Intell Transp Syst* 16:2014–2027

66. Vora S, Rangesh A, Trivedi MM (2017) On generalizing driver gaze zone estimation using convolutional neural networks. In Proceedings of IEEE Intelligent Vehicles Symposium, Redondo Beach, pp. 849–854
67. Wang J, Zhang G, Shi J (2016) 2D gaze estimation based on pupil-glint vector using an artificial neural network. *Appl Sci-Basel* 6:1–17
68. Yoo DH, Chung MJ (2005) A novel non-intrusive eye gaze estimation using cross-ratio under large head motion. *Comput Vis Image Underst* 98:25–51



Hyo Sik Yoon received a BS degrees in Electronics Engineering from Kang Won National University, Kangryung, South Korea in 2015. His research interests include pattern recognition, biometrics, and image processing.



Hyung Gil Hong received a BS and Master degree in electronics engineering from Dongguk University, Seoul, South Korea, in 2012 and 2014, respectively. He is pursuing PhD degree in Electronics and Electrical Engineering at Dongguk University. His research interests include biometrics and pattern recognition.



Dong Eun Lee received a BS degrees in Electronics Engineering from Chung Nam National University, Daejeon, South Korea in 2015. His research interests include pattern recognition, biometrics, and image processing.



Kang Ryoung Park received his BS and Master degrees in electronic engineering from Yonsei University, Seoul, South Korea, in 1994 and 1996, respectively. He received his PhD degree in electrical and computer engineering from Yonsei University in 2000. He has been a professor in the division of electronics and electrical engineering at Dongguk University since March 2013. His research interests include image processing and biometrics.