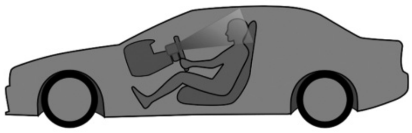
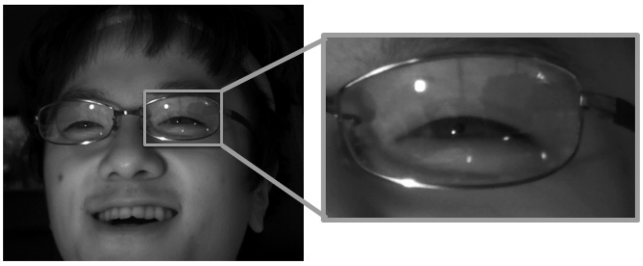
Development of Gaze Detection Technology toward Driver's State Estimation

Naoyuki OKADA
Akira SUGIE
Itsuki HAMAUE
Minoru FUJIOKA
Susumu YAMAMOTO





Abstract

In recent years, the development of advanced safety support systems and autonomous driving systems with the accelerator/brake/steering control for automobiles has been promoted. In the autonomous driving systems, in the event that the autonomous driving cannot be continued when an abnormal condition occurs, it is necessary to switch a subject of driving from the autonomous driving system to a driver. If the driver is not ready to drive an automobile when switching the subject of driving, there is a possibility of falling into a very dangerous state. Therefore, it is necessary for the system to always monitor a driver's state.

Accordingly, we considered a method of utilizing information on a driver's gaze in order to monitor the driver's state, and have developed a method of detecting a gaze direction by processing camera images of a driver's face region. In the in-vehicle environment, there was a problem in that the external scenery or light is reflected on glasses as an image, whereby exerting an adverse effect on the detection performance. With respect to this problem, we have established the gaze detection algorithm which improves the robustness during vehicle traveling, by accumulating and utilizing data on luminance information within a region in which the external scenery or light is reflected.

1 Introduction

According to the summary of information on traffic accidents in Japan in 2014 prepared by the National Police Agency, there were 4,113 fatalities, and 711,374 persons were injured. The numbers of fatalities and injuries have been decreasing recently, and the number of fatalities has been slowly decreasing in the past few years [(1) (2)] (See **Fig. 1**).

Breaking down the causes of traffic accidents, many of the fatal accidents were caused by violations of safe driving requirements, such as inattentive driving or distracted driving. These are human errors made while driving (**Fig. 2**: Number of fatal accidents in Japan classified by type of legal violation by the party principally at fault, in 2014). To eliminate the traffic accidents caused by these human errors, the development of driving support systems, such as an automatic emergency braking system and more advanced autonomous driving systems is proceeding.

In those systems, a host vehicle is controlled using the vehicle controlled status, e.g. its speed and steering wheel angle, and on the situation around the vehicle including the distance between it and other vehicles which is obtained by sensing with millimeter wave radar and/or a camera. However, we feel that it is also important to improve the technology for monitoring the driver's condition in order to realize smarter and more user-friendly driving support/ autonomous driving systems, and so have been working to develop gaze detection technology to determine the state of the driver in the vehicle traveling environment.

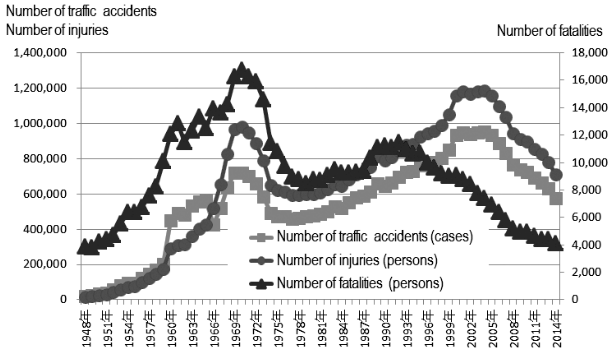


Fig.1 Changes in Number of Traffic Accident Fatalities in Japan in Each Year

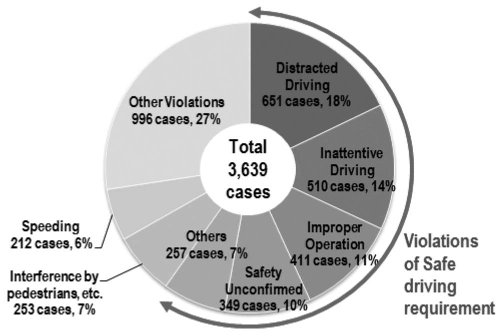


Fig.2 Number of Fatal Accidents in Japan Classified by Types of Legal Violations (First Party)(in 2014)

2 Background of Development

2.1 Problems of Driving Support Systems and Autonomous Driving Systems

In driving support systems and autonomous driving systems, a part or all of driving operations are automated

Table 1 Autonomy Levels Classified by Society of Automotive Engineers International

SAE Level	Name	Narrative Definition	Execution of Steering and Acceleration/ Deceleration	Monitoring of Driving Environment	Fallback Performance of Dynamic Driving Task	System Capability (Driving Modes)	BASt Class	NHTSA Level
0	No automation	The full-time performance by the human driver of all aspects of the dynamic driving task, even when enhanced by warning or intervention systems	Human driver	Human driver	Human driver	—	Driver Only	0
1	Driver assistance	The driving mode-specific execution by a driver assistance system of either steering or acceleration/deceleration using information about the driving environment and with the expectation that the human driver perform all remaining aspects of the dynamic driving task	Human driver and System	Human driver	Human driver	Some driving modes	Assisted	1
2	Partial autonomy	The driving mode-specific execution by one or more driver assistance systems of both steering and acceleration/ deceleration using information about the driving environment and with the expectation that the human driver perform all remaining aspects of the dynamic driving task	System	Human driver	Human driver	Some driving modes	Partially Automated	2
3	Conditional autonomy	The driving mode-specific performance by an automated driving system of all aspects of the dynamic driving task with the expectation that the human driver will respond appropriately to a request to intervene	System	System	Human driver	Some driving modes	Highly Automated	3
4	High autonomy	The driving mode-specific performance by an automated driving system of all aspects of the dynamic driving task, even if a human driver does not respond appropriately to a request to intervene	System	System	System	Some driving modes	Fully Automated	3/4
5	Full autonomy	The full-time performance by an automated driving system of all aspects of the dynamic driving task under all roadway and environmental conditions that can be managed by a human driver	System	System	System	All driving modes	—	

using vehicle control information through the accelerator, brake, and etc., and using vehicle peripheral information such as the vehicle position obtained by a sensor or a communication device. Levels are defined according to the degree of autonomy as shown in **Table 1**. In this table, there are issues in the case of switching a subject of driving between the autonomous driving system and a driver.

For example, in systems falling under the Level 1 or Level 2, there is a possibility that such the system intervenes in operations which is not intended by the driver. The system may control the vehicle so that it accelerates faster or decelerates more slowly than the driver intends, as is sometimes the case with ACC (Adaptive Cruise Control: cruise control with an inter-vehicular distance control function). Then, depending on the timing and frequency of such control, the vehicle may be put into danger or the driver may feel annoyed. Further, how the driver feels often varies depending on the driver condition, i.e. whether the driver is looking forward or not. Therefore, we consider that it is necessary to sense the driver condition and to provide the support that is the appropriate not only to the state of the vehicle and the external environment but also to the driver condition.

In addition, in autonomous driving systems falling under Level 3, although it is not normally necessary for a driver to perform driving operations, there is a possibility that autonomous driving cannot be continued due to failure of equipment in the system, system error, and external environment change, such as rapid change in the weather and road environment change. In this case, it becomes necessary for the driver to take over driving operations from the system. At this time, if the driver is not in a state to drive, the driver may make a driving operation error immediately after taking over the driving and the vehicle may be put into danger that is different from the dangers normally caused by human errors. For this reason, even in autonomous driving systems, we think that it is necessary to always monitor whether the driver is in a state to drive, in order to judge whether driving operations can be taken over by the driver.

In searching for the means to monitor the driver condition, we focused on the driver's gaze, which plays the most important role among the five senses necessary for driving.

2.2 Driver Monitoring Using Gaze

A driver is in a state to drive when all of the three driving actions, "recognition, judgment, and operation", can be performed. Driver monitoring which can judge whether the driver can perform these three driving actions is required. Among "recognition, judgment, and operation", recognition is especially important and performed first, which allows subsequent judgments and operations to be made. It is said that most of the recognition behavior during driving relies on visual input. In fact, although it is possible to engage in non-driving behavior which does not involve visual input during driving, such as listening to music and drinking a beverage, it is impossible to drive a vehicle without visual input, for example, when a blindfold is over the eyes. Visual recognition by the driver principally represents what the driver is looking at. If the driver looks at and pays attention to an object, that object is recognized.

Some driver monitoring systems utilizing images of the driver's face have already been put into practice. These systems can detect whether the driver's eyelids are open and whether the driver's face is facing forward. However, even if the driver's face is facing forward, it is possible that the driver watches the car navigation screen or the like. In such cases, the driver cannot correctly monitor the information necessary for driving, such as what is in front of the vehicle. We detected the gaze of a driver who was driving properly on an expressway, and **Fig. 3** shows the directions of the lines with the image in front of the vehicle. It is clear that the driver's gaze tend to concentrate in the vehicle's traveling direction when the driver is safely driving the vehicle. Based on these facts, we considered that it is important to include detection of the driver's gaze in driver monitoring.

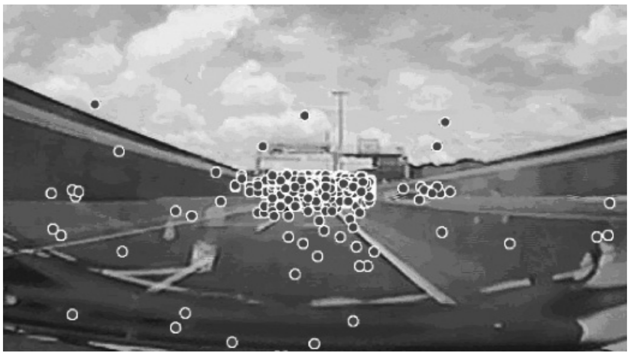


Fig.3 Gaze Distributions during Traveling on Expressway (for 30 Seconds)

3 Elemental Technologies of Gaze Detection

3.1 Mechanism of Human Eyes

First, **Fig. 4** shows the structure of a human eyeball. In general, a human viewing angle is 180 to 190 degrees. However, it is not possible to recognize the details of an object positioned near an edge of this visual field. The range in which written characters or details of an object can be recognized is only 1 to 2 degrees, a very narrow range. Therefore, as shown in **Fig. 5**, in order to clearly recognize an object, the eyeballs are rotated so that the object is placed in this range. It is a so-called "gaze" that this range of 1 to 2 degrees coincides with the object to be observed, and there is a close relationship between the orientation of the eyeball and the gaze direction. Therefore, observation of the orientation of the eyeball leads to the detection of the gaze.

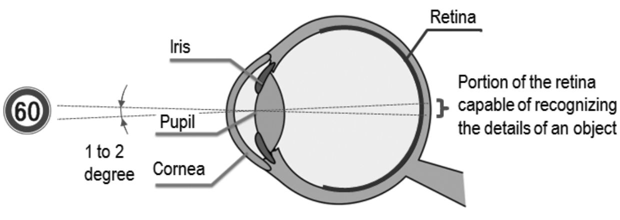


Fig.4 Structure of Human Eyeball

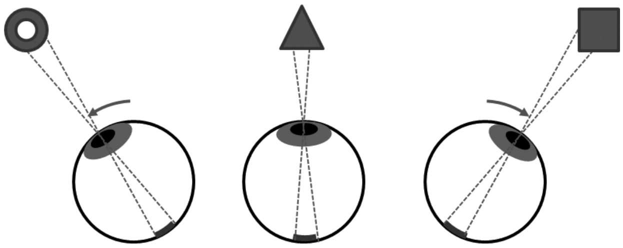


Fig.5 Movements of Eyeball when Watching Something

3.2 Gaze Detection Method

As shown in **Fig. 6**, a camera and a point light source are installed close to the gaze of the driver looking forward. In this case, both the light and the camera operate in the near infrared range; an actual reflected image thus produced is shown in **Fig. 7**. A light beam from the point light source is reflected on the cornea surface and is seen as a white dot. Since this cornea reflection dot is from an eyeball surface having a spherical shape, even if the eyeball rotates as shown in **Fig. 8**, the cornea reflection dot is always at the same position. For this reason, it is possible to obtain the three-dimensional direction of the gaze based on the position of the pupil with respect to the position of the cornea reflection dot which is used as the reference point.

In actuality, the eyeball is not a perfect sphere, and there are individual differences in the structure of the eyeball and shape of the cornea. Therefore, it is necessary to be calibrated for each individual in advance so as to make the direction of the gaze obtained from the image correspond to the direction which is actually watched by that individual. With this calibration, it becomes possible to determine the driver's gaze direction with higher accuracy.

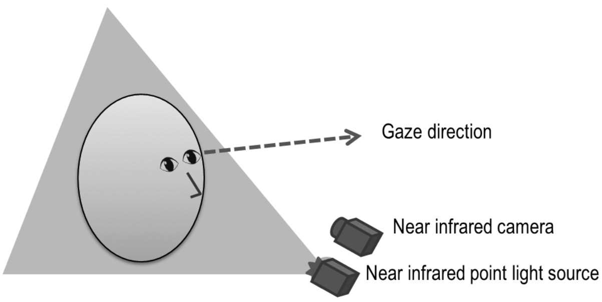


Fig.6 Gaze Detection Mechanism Using Point Light Source

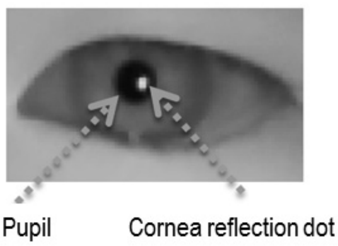


Fig.7 Eye Region Image Captured by Near Infrared Camera

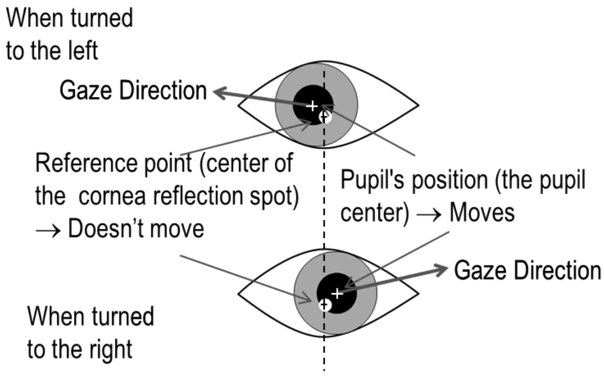


Fig.8 Positional Relationships between Pupil and Cornea Reflection

4 Efforts to In-vehicle System

4.1 Influence of Ambient Light during Vehicle Traveling

In order to detect the pupil and the cornea reflection dot from an image of the driver's face, the fact that the pupil has the lowest luminance and the cornea reflection dot has the highest luminance in the image of the eye extracted, is utilized. By finding the circular region with low luminance and identifying this as the pupil, and by finding the circular region with high luminance and identifying this as the cornea reflection dot, the position of the pupil and the position of the cornea reflection dot are determined.

However, in the vehicle driving environment, external light becomes a disturbance. In particular, when a driver wears glasses, as shown in **Fig. 9**, peripheral objects such as the scenery are reflected on the lens surface of the glasses in some cases. In this case, if the lens reflection overlaps with the vicinity of the pupil, it becomes difficult to detect the pupil. This is because the luminance of the pupil portion, which should be the lowest, becomes high due to the reflected light, and the luminance of other portions becomes relatively low, so that another portion is erroneously detected as the pupil (**Fig. 10**). Such a situation occurs quite commonly, e.g. in fine weather when external light is strong and in urban areas where there are many structures such as buildings around the vehicle. Therefore, measures to deal with this situation must be taken.

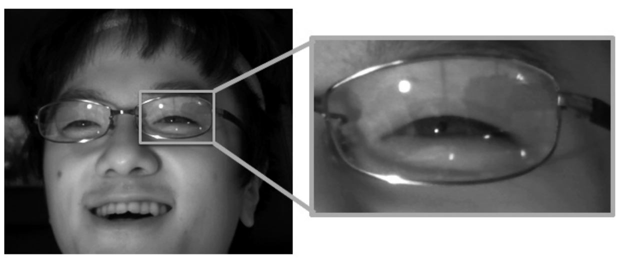


Fig.9 State in Which Peripheral Objects are Reflected on Lens Surfaces of Glasses

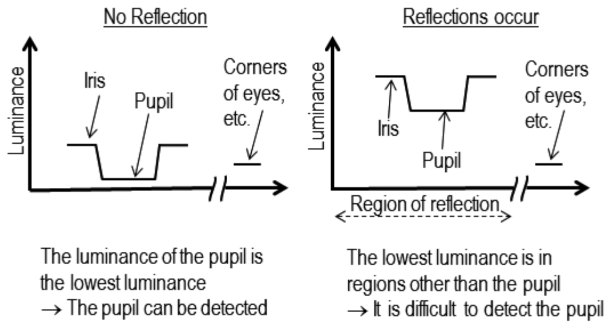


Fig.10 Luminance Changes of Eye Region Due to Reflections on Glasses

4.2 Improvement of Robustness While Driving

There are cases in which there are reflections from glasses only at one eye and at both eyes. We considered measures to deal with overlapping reflections utilizing this fact [3]. It can be judged from information on luminance throughout the entire eye region and luminance changes whether there are overlapping reflections. Hereafter, pupil detection processes where the range of overlapping reflections covers one eye only and where the range covers both eyes will be described separately.

4.2.1 Pupil Position Estimation and Data Accumulation when There are Overlapping Reflections only on One Eye

Even when there are overlapping reflections only on one eye, it is possible to detect the pupil and the cornea reflection dot of the eye with no overlapping reflection and thus obtain the gaze direction. On the other hand, since a light source with strong intensity is used and then the luminance of the cornea reflection dot appearing on an eye with overlapping reflections becomes sufficiently higher, the cornea reflection dot can be detected. Further, the relative positions of the pupil and the cornea reflection dot and their diameters in the right and left eyes are almost the same. Therefore, on the basis of the relative positions of the pupil and the cornea reflection dot of the eye with no overlapping reflection and the cornea reflection dot position on the eye with overlapping reflections, it is possible to estimate the position of the pupil in the eye with overlapping reflections. (**Fig. 11**)

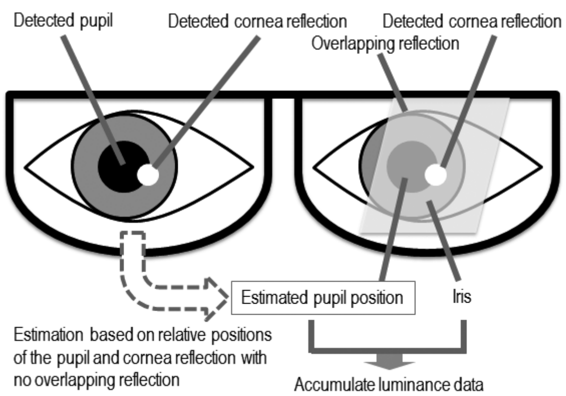


Fig.11 Pupil Estimation when Reflected on One Eye

Then, in the eye with overlapping reflections, data on the luminance at the position of the pupil thus estimated and the luminance of the iris around the pupil are accumulated. This data will be used for pupil estimation when there are reflections overlapping both eyes, which will be described in the next section.

4.2.2 When Reflections Overlap both Eyes

When reflections overlap both eyes, the pupil's position is calculated utilizing previously accumulated luminance data on the pupil and the iris, in accordance with the following procedure:

(1) The positions of the right and left cornea reflection dots are found, and the approximate iris position is presumed.
(2) In the vicinity of the pixels presumed to represent the iris, the pupil luminance corresponding to the presumed iris luminance is searched among accumulated data of correlations between the luminance of the iris and the pupil, and the pixels with the retrieved luminance are made candidate pixels of the image of the pupil.
(3) The pixels in the candidate group that form a circular region are judged to represent the pupil.

Based on the positional relationship between the pupil detected through this process and the cornea reflection dot, it is possible to obtain the eye direction. (**Fig. 12**)

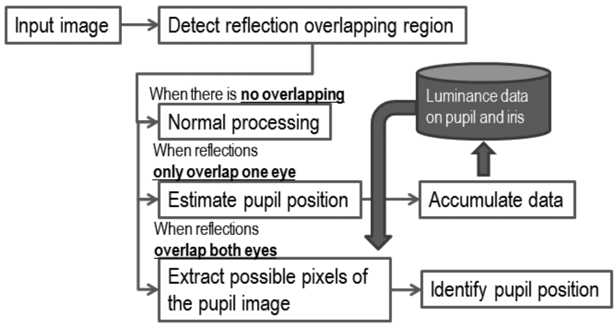


Fig.12 Pupil Detection when Reflections Occur

4.3 Evaluation

The algorithm developed in this study was introduced into the pupil detection algorithm which had been developed by us, and the effect of this introduction on pupil detection accuracy when there are overlapping reflections, was evaluated. It should be noted that the pupil detection algorithm before introduction is based on conventional technology (see Section 4.1) which performs detection using luminance characteristics of the pupil and the cornea reflection dot and the characteristic that the pupil is round, and has high robustness when glasses are worn if there is no reflection, and when the ambient brightness changes.

Images where strong reflections overlap one eye or both eyes were collected, and **Fig. 13** shows the results of the pupil detection when these images were used as inputs. The average of pupil detection rate in 5 test subjects was 10.7 % before the introduction of the above new algorithm, but the detection rate was 89.7% after the introduction of this algorithm. From this result we confirmed the effectiveness of new algorithm when there are overlapping reflections.

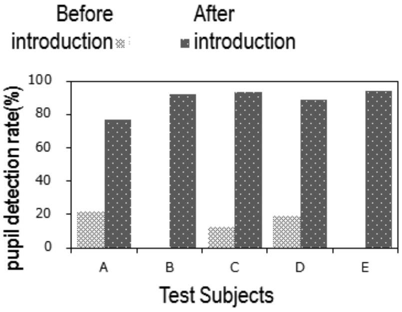


Fig.13 Comparisons of Pupil Detection Rates before and after Introducing Developed Algorithm

The conventional pupil detection algorithm exhibits robustness and performs correct detection when the images of an eye were not overlapped with reflections or when the luminance values of the eye regions uniformly increased due to partial overlapping. However, it could not perform correct detection when there were scattered reflections which made the luminance change unevenly in the eye regions.

On the other hand, after introducing the above newly developed algorithm, not only is the pupil correctly identified but also the relationship between the luminance of the pupil and the luminance of the iris at particular locations is found. Therefore, it is possible to detect the pupil even when reflections partially overlap the eye. However, it is not possible to correctly detect the pupil position when small light source reflections and the like generated on lens surfaces of glasses are erroneously detected as the cornea reflection dot, because the pupil is assumed to be close to that dot (see Section 4.2.2). It should be possible to prevent this erroneous detection by extracting cornea reflection dot and pupil candidate and then making a comprehensive evaluation of these candidates based on positional relationships of these candidate in the right and left eyes, and consistency with past detection results, etc.

5 Conclusion

In this study, we made efforts to realize in-vehicle driver monitoring, in particular gaze detection which is an important driver monitoring function. This monitoring will be essential to make possible safe and comfortable driving in the automobile society of the future. By concretely specifying problems unique to the in-vehicle environment, and by developing and evaluating measures to solve those problems, we were able to establish gaze detection technology that can deal better with interference from external light.

Through gaze detection technology, a driver's behavior can be sensed directly. Since the driver's visual recognition behavior provides most of the information necessary for driving, we feel that analysis of gaze behavior will make it possible to guess when the driver is in one of various states, such as slight sleepiness, driver distraction, and driver load while driving, which were impossible to detect in the past. These estimations of the driver condition can be utilized to provide driving support and service that is suitable for the driver condition and to make safe the take-over of driving operations by the human driver in Level 3 autonomous driving.

In the future, we will create Vehicle-ICT* using our advanced gaze detection technology and technology for detection of a driver condition utilizing that gaze detection technology, to realize a more prosperous motorized society.

6 Acknowledgements

Finally, we would like to take this opportunity to offer our sincere gratitude to Satoshi Nakajima, a principal investigator of Fujitsu Laboratories, Ltd. and other concerned parties, who cooperated in development of this technology.

References

(1) Traffic Bureau, Japan National Police Agency, "Characteristics of Fatal Traffic Accidents and State of Road Traffic Act Violation Control in 2014", 2014
(2) Japan Cabinet Office, "FY 2015 Traffic Safety White Paper", 2015
(3) Takahiro YOSHIOKA, Satoshi NAKAJIMA, Jyunichi ODAGIRI, Hideki TOMIMORI, Taku FUKUI, "Pupil Detection Robust against Environment Conditions When Wearing Glasses," Technology Report of the Institute of Image Information and Television Engineers, 2014, pp. 37-39.

*"Vehicle-ICT" is a trademark of Fujitsu.
(note) Proper nouns used herein such as products names are trademarks or registered trademarks of respective companies.

Profiles of Writers


Naoyuki OKADA
AS Engineering Group
Advanced Sensing System R&D Dept.


Akira SUGIE
AS Engineering Group
Advanced Sensing System R&D Dept.


Itsuki HAMAUE
Research & Development Planning Division
System Research & Development Dept


Minoru FUJIOKA
AS Engineering Group
Advanced Sensing System R&D Dept.


Susumu YAMAMOTO
Research & Development Planning Division
System Research & Development Dept