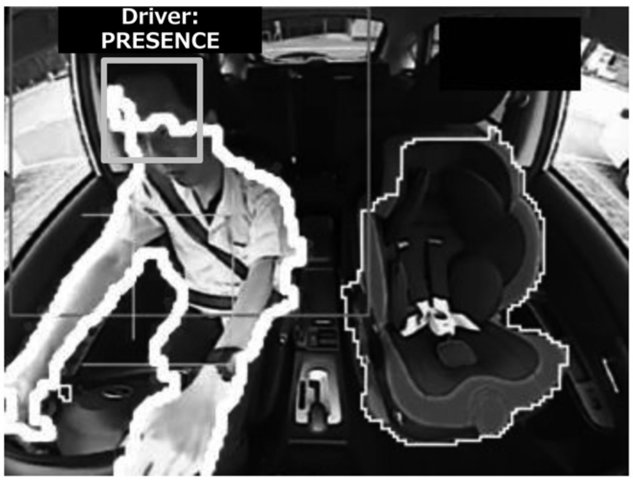
Development of a Posture-Sensing Technique Using a Wide Angle Monocular Camera (Occupant Extraction Technique)

Minoru FUJIOKA
Taiki KUWAHARA
Itsuki HAMAUE
Takahide IMAMIYA



Abstract

In recent years there has been a profuse amount of development of technology for sensing driving environments such as millimeter-wave radar and cameras in order to create even more sophisticated collision safety technology. We have been working on a posture detection technique that can extract vehicle occupants and ascertain their postures from images of a camera mounted inside a vehicle in order to support safe driving corresponding to the driver and passenger statuses by combining with 'sensing driving environments' and 'sensing driver and passenger statuses'.

This paper describes a posture-sensing technique capable of extracting vehicle occupants (driver and passengers) from monocular camera images. We have created a dynamic background updating method for use as an occupant extraction technique based on the graph-cut technique, a background subtraction method that functions well even with light fluctuation. In this paper, we explain how we created assumed actual use scenarios, and then checked the effectiveness of this occupant extraction technique in various evaluation scenarios, including those in which elements such as vehicle exterior brightness and seat conditions were varied, as a means to extract corresponding issues.

1 Introduction

The number of fatalities due to traffic accidents for all of Japan has been declining annually as the National Police Agency announced that there were 3,904 deaths in 2016, dropping it below the 4,000 figure [1]. However, there are still some 500,000 traffic accidents every year in the country [2]. To further reduce the number of traffic accidents, there is a profuse amount of research and development being conducted that focuses on creating even more sophisticated collision safety technology such as air bags, as well as improved preventive safety technology through the use of various types of sensors such as millimeter-wave radars and cameras. FUJITSU TEN has been developing technology capable of sensing the driving environment of vehicles, such as our millimeter-wave radar, with the objective of contributing to safe-driving support. We are focused on creating user-friendly safe-driving support systems by combining such drive-environment sensing with additional technology that senses driver and passenger statuses in order to induce evasive actions in accordance with the driver and passenger statuses and provide pertinent safety information. Amid our efforts regarding driver and passenger statuses, we have been working on a posture detection technique that can extract vehicle occupants and their postures from a camera mounted inside a vehicle in order to ascertain the presence and actions of the driver and passengers. This paper describes a posture-sensing technique capable of extracting vehicle occupants (driver and passengers) from monocular camera images.

2 Status Estimation from Driver and Passenger Posture

2.1 Issues with Existing Techniques

Existing techniques for sensing driver and passenger statuses within a vehicle include detecting the load distribution applied to seating sensors (load sensors) embedded in the seat surface, and measuring the charge generated between persons and sensor electrodes in a non-contact manner by using capacitive sensors also embedded in the seat surface to ascertain the position of the vehicle occupants' body.

While these sensors are capable of detecting the presence of occupants sitting in seats, they cannot detect their posture that expresses their statuses. For example, they cannot sense further details to the point of determining if the driver has a posture in which the driver's head is positioned in the middle of the headrest, the body is seated upright (not slouching), and the driver's arms are extended and holding the steering wheel.

2.2 Application of Image Recognition

For our purposes, we selected a camera capable of capturing detailed elements within the vehicle interior, and used image recognition to extract and detect the postures of vehicle occupants in order to estimate their statuses from this detected posture.

As a condition, the mounting position of this camera in the vehicle interior must be such that it has a panoramic view overlooking the entire interior in order to be able to detect the postures of multiple occupants. In other words, this means the camera must provide a wide-angle view in both horizontal and vertical directions while also being miniaturized. Additionally, in consideration of practical usage, it is crucial that it can perform detection at night also. If these requirements could be realized, we believe that it could also comply with the US collision safety standards (FMVSS 208) and the standards for seatbelt reminders of Euro NCAP.

Based on these requirements, we studied possible use of a wide-angle miniaturized camera that could be mounted to a vehicle. This study resulted in the selection of a near-infrared camera and light unit used in FUJITSU TEN's drive recorder, shown in **Table 1**, and the installation of this camera on the overhead console in the center of the vehicle's roof to provide an unobstructed view of the entire vehicle interior. This vehicle interior monocular camera can provide images such as shown in **Fig. 1**.

Table 1 Camera and Light Specifications

Near-Infrared Camera		
	Imaging element	1/4-in. color CMOS
	Pixel count	310,000 pixels
	Angle of view	135 deg. horizontal / 105 deg. vertical
Near-Infrared Light Unit		
	Emission peak wavelength	850 nm
	Light output	12 mW/sr
	Illuminating angle	Radiation angle of 140 deg. / Half-value angle of 60 deg.

Fig. 1 Image from Vehicle Interior Monocular Camera

3 Posture Detection System

3.1 System Configuration

Fig. 2 shows the configuration of a system capable of extracting vehicle occupants and their postures in order to estimate their statuses. Images captured by the vehicle interior camera are input for use in image recognition from which occupants can be extracted and their postures can be detected. This paper explains the development of a technique for extracting vehicle occupants from images of a vehicle interior camera, which is considered as Step 1 of the overall process. In the future, we will be developing techniques capable of detecting the postures of the occupants extracted from the images of the vehicle interior camera (Step 2), followed by the development of techniques capable of estimating the statuses of the occupants from their detected postures (Step 3).

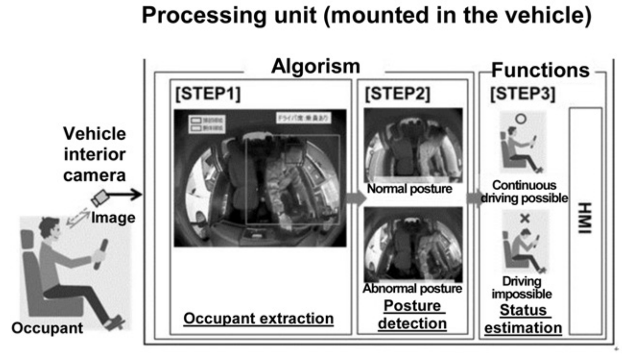


Fig. 2 System Configuration

3.2 Occupant Extraction Technique

The background subtraction method is used as a base technique for extracting occupants. Although the background subtraction method generally uses the differences in the pixels between frames to determine whether they correspond to the background or occupants, the weak point of this method is that light fluctuation makes it more difficult to detect differences in pixels. To resolve this issue, we applied the graph-cut technique proposed by Atsushi Hashimoto of Kyoto University in order to determine the difference between occupants (foreground) and the background. [3]

Graph cut is one method used for region extraction of images and is generally considered as a technique that functions well even with light fluctuation. It is a type of background subtraction method for determining the regions where the total amount of energy required for separating the foreground (image) from a previously obtained background (image) is the lowest, and separating the image accordingly.

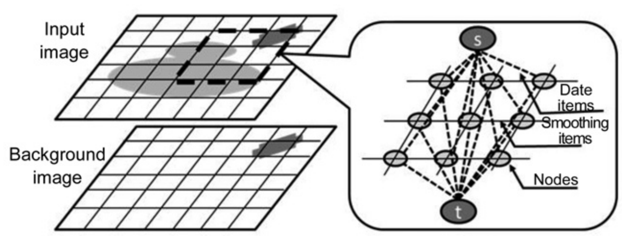


Fig. 3 Region Separation Using Graph Cut

Fig. 3 shows the foreground "s", background "t" and each adjacent node (indicated by "O" in **Fig. 3**) with these items connected by a pipe being cut to separate the foreground and background. Additionally, the flow between each node with "s" and "t" are referred to as data items (dotted lines in **Fig. 3**), and the flow between adjacent nodes themselves are referred to as smoothing items (solid lines in **Fig. 3**). Data items indicate to what degree this region appears to be in the foreground (s) or background (t), while the smoothing items indicate to what degree this region is connected with the four adjacent regions. The stronger the connection between each of these nodes, the larger the amount of energy that is required to separate them.

3.3 Issue of the Background Subtraction Method

The following issue is thought to exist if using this graph-cut technique, a background subtraction method that functions well even with light fluctuation, without modification for the interior of a vehicle.

When occupants (especially the driver) get into a vehicle, they adjust the seat position and recline it to provide a position that makes it easy to drive. However, moving the seat results in the seat acquired as part of the background data also being moved, such that the seat is then extracted not as a part of background data but rather as a part of foreground data.

This means that both the occupant sitting in the seat and the seat itself are processed as foreground data, which we can assume to negatively affect the detection accuracy of occupant extraction.

3.4 Dynamic Background Updating Method

We have focused on two points to resolve this previously assumed issue.

The first is the fact that conditions resulting in seat movement are infrequent during the period from when an occupant enters the vehicle until the occupant exits the vehicle. The second is that even after an occupant sits in the seat, the occupant will not stay completely motionless for the whole time. For these reasons, we focused on being able to detect movement at the edge (texture) of the region around the pixels, even if the pixel values are the same, as shown in **Fig. 4** to develop a technique that registers moving pixels with their surrounding texture when detecting background data so that the background data is updated when those pixel values do not change for a certain period and the texture does not change.

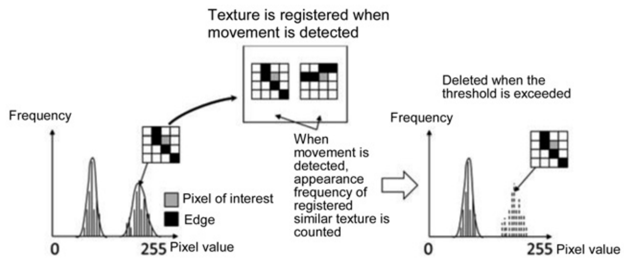


Fig. 4 Dynamic Background Updating Method

4 Evaluation of Occupant Extraction Technique Using Dynamic Background Updating

4.1 Creation of an Evaluation Environment

In order to test and evaluate the occupant extraction technique using dynamic background updating, we created an evaluation environment simulating the environment of a vehicle interior by setting up a camera, LED, seat and steering wheel in an indoor testing site as shown in **Fig. 5**.

The camera and LED used were, as indicated in Section 2, a near-infrared camera and LED light from a FUJITSU TEN infrared camera kit compatible with nighttime use (CMR-4012). A seat with forward/back sliding and reclining functions (ERGOMED-D by RECARO) was selected, and a steering wheel device with a rotating mechanism capable of reproducing a driver's driving posture and operations was installed.

It was also necessary to recreate different lighting environments such as afternoon, dusk and evening, with the assumption that a vehicle is driven at various times of the day and night. To achieve this, several of the fluorescent lamps of the testing site as well as artificial solar light irradiation equipment capable of linear light intensity adjustment (XC-500AF by SOLAX) was used to allow for adjusting the surrounding light level.

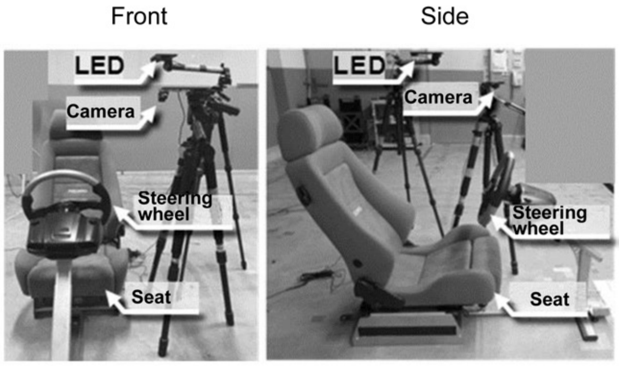


Fig. 5 Evaluation Environment

4.2 Evaluation Method

With the evaluation environment described in Section 3 obtained, camera images as shown in **Fig. 6** were read into a PC and the current occupant extraction technique of Step 1 was applied for each image frame to visually judge whether occupant extraction was properly achieved as shown in **Fig. 7**. The judgment criteria applied was as shown in **Fig. 8** such that extraction was determined to be mistaken or failed if the extracted region did not include half or more of the width of the occupant's area.

Fig. 6 Camera Image Obtained in Evaluation Environment

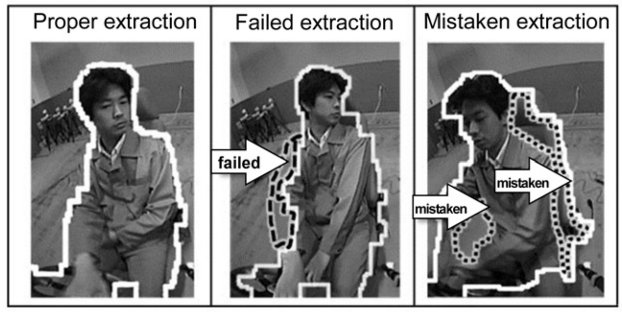


Fig. 7 Visual Judgment Results

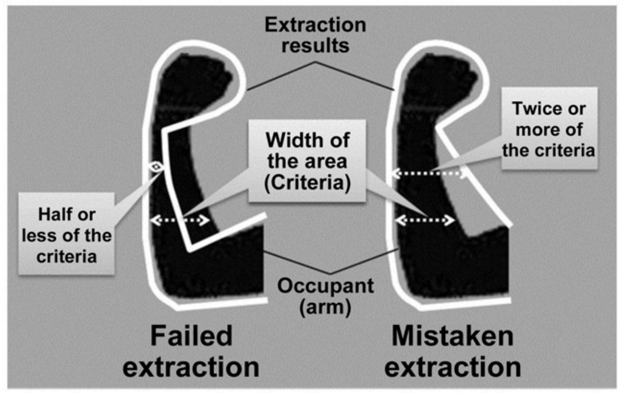


Fig. 8 Judgment Criteria for Mistaken or Failed Extraction

4.3 Evaluation

Assumed actual use scenarios were created in order to perform evaluation by varying the elements of vehicle exterior brightness, seat conditions (seat position and reclining), seat material, occupant's clothing materials and colors, and occupant's body type (adult or child). As shown in **Table 2**, the three scenarios described below were used for this evaluation in order to check the effectiveness of countermeasures against vehicle exterior brightness and seat movement, and evaluation was performed by using the obtained images. Approximately 800 frames (approx. 26 sec.) of each scenario were used.

(1) Normal driving actions were performed with the only occupant sitting in the driver's seat starting from the condition of an empty seat to the driver sitting in the seat, grasping the steering wheel, checking left and right, and operating the steering wheel with an illumination intensity similar to the level of sunlight on a cloudy afternoon (32,000 lux).

(2) Normal driving actions were performed with the fluorescent lights and artificial solar light irradiation equipment adjusted to darken the surrounding area to an illumination intensity similar to the level of that provided by street lights at night (75 lux).

(3) Normal driving actions were performed, followed by seat adjustment in which the seat position was moved forward by 20 cm and reclined forward by 33 degrees with an illumination intensity similar to the level of sunlight on a cloudy afternoon.

Table 2 Evaluation Scenarios

	Details
Scenario 1	Normal driving actions (checking left/right, steering wheel operation) with an illumination intensity similar to the level of sunlight on a cloudy afternoon (32,000 lux)
Scenario 2	Normal driving actions under nighttime street lamps (75 lux)
Scenario 3	Changes to seat position and reclining angle under sunlight on a cloudy afternoon

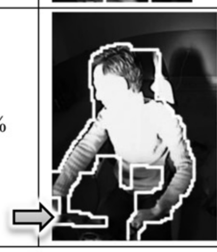
4.4 Results

This section describes the results of occupant extraction using dynamic background updating **Table 3**. The results of the evaluation showed two types of failed extraction and three types of mistaken extraction. We then analyzed the causes of these five types of failed and mistaken extraction.

Table 3 shows that extraction of the torso and arm failed because the colors of the right shoulder and seat were similar in No. 1 as seen in the reference images of Scenarios 1 and 2, and the colors of the arm and the border of surrounding objects were similar in No. 2 as seen in the reference image of Scenario 3. With use of the graph-cut technique here, the data items indicating to what degree a region appears to be in the foreground or background did not increase, while the surrounding background is propagated such that the region of the occupant is also judged to be background based on the smoothing items indicating connection with the four adjacent regions. The first issue here is that even if the color of the occupant and background are similar, parameters must be optimized by giving greater consideration to data items and similar methods when judging separation from smoothing and data items of the graph-cut technique so that more data items can be handled **Table 3**. **Table 3** shows that the seat was mistakenly extracted in No. 3 as seen in the reference image of Scenario 3, and the headrest was mistakenly extracted in No. 4 as seen in the reference images of Scenarios 1 and 2. This is due to changes in camera exposure caused by the reflection of the occupant or changes in the appearance of the background caused by the shadow of the occupant formed on the headrest, resulting in the extraction of differences regardless of the background. The second issue here is that parameters for determining changes in texture must be optimized so that dynamic background updating can be performed even under these conditions **Table 3**. **Table 3** shows that the steering wheel was mistakenly extracted in No. 5 as shown by the reference images of Scenarios 1, 2 and 3. This was due to background differences caused by changes in the appearance of the steering wheel when it was turned. The third issue here is that it must be possible to remove the regions with operating devices such as the steering wheel, seat belt retractors and shift knob from region subject to background differentiation.

Resolution of these three issues can reduce failed and mistaken occupant extraction, and we believe that they can be reduced even further by combining with time-series processing using the extraction results of multiple frames of the later development stages.

Table 3 Occupant Extraction Results

No.	Extraction Conditions	Occurrence Rate within Images			Reference Images : Occupant Extraction Results : Location of Condition Occurrence		
		Scenario 1	Scenario 2	Scenario 3	Scenario 1	Scenario 2	Scenario 3
1	Failed extraction of torso	82%	74%	0%			
2	Failed extraction of arm	56%	48%	10%			
3	Mistaken extraction of background objects around arm	54%	33%	92%			
4	Mistaken extraction of headrest	69%	55%	0%			
5	Mistaken extraction of steering wheel	27%	16%	55%			

5 Future Plans

We have developed a vehicle occupant extraction technique for detecting the presence of multiple occupants by use of a wide angle monocular camera mounted in the vehicle interior, and have also ascertained the corresponding issues. Our objective is to resolve such issues in order to first install a single camera to a vehicle interior that can detect the presence of occupants in multiple seats. Next, we plan to perform further image analysis of the detected occupants to not only determine their presence but also detect the posture of each occupant and estimate their statuses. We believe that posture sensing techniques can be applicable to various products in the future for purposes such as:

- Regulating the timing of the opening of vehicle airbags based on the body type of the seat occupant
- Dead man system that performs an emergency stop of a vehicle by determining from the driver's posture that he or she has lost consciousness while driving due to a seizure or other sudden illness
- Judgment of whether the driver is in a proper posture to assume control of the vehicle when changing from automated to manual driving

6 Acknowledgments

Finally, we would like to express our deepest gratitude to Mr. Murashita, Innovation Manager of FUJITSU LABORATORIES LTD. and all related personnel who contributed to the development of this technique.

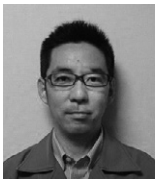
References

[1] National Police Agency Traffic Bureau, "Fatalities from traffic accidents in 2016," 2017.

[2] Institute for Traffic Accident Research and Data Analysis, "Occurrence Status of Traffic Accidents in 2016," http://www.itarda.or.jp/situation_accident.php..

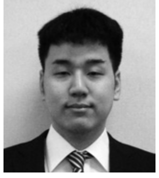
[3] Atsushi Hashimoto, Kazuaki Nakamura, Takuya Funatomi, Masayuki Mukunoki, and Michihiko Minoh, "TexCut: Background Subtraction by Texture Comparison on Graph Cut," in The IEICE Transactions on Information and Systems (Japanese Edition), 2011.

Profiles of Writers



Minoru FUJIOKA

VICT Engineering Group
Engineering R&D Dept



Taiki KUWAHARA

VICT Engineering Group
Engineering R&D Dept

Itsuki HAMAUE

VICT Engineering Group
Engineering R&D Dept

Takahide IMAMIYA

VICT Engineering Group
Engineering R&D Dept