A Study of Safety Driving Support System focusing on Driver's Head Posture Categorization

Momoyo Ito[1], Kazuhito Sato[2], and Minoru Fukumi[3]
[1]*Assistant Prof., Institute of Technology and Science, The University of Tokushima, Tokushima, Japan,*
[2]*Associate Prof., Faculty of Systems Science and Technology, Akita Prefectural University, Akita, Japan,*
[3]*Professor, Institute of Technology and Science, The University of Tokushima, Tokushima, Japan,*

Abstract

We aim to construct a driver assistance system that is able to detect such driver deviations. The system detects deviation using time-series head motion information. We analyze driver's head posture during safety verificationand propose a method for classifying head posture using two types of unsupervised neural networks: Self-Organizing Maps (SOMs) and fuzzy Adaptive Resonance Theory (ART). The proposed method has a feature based on the hybridization of two unsupervised neural networks with a seamless mapping procedure. The proposed method can generate the optimal number of cluster-generated labels for the target problem. We experimentally assess the effectiveness of the proposed method by adjusting the fuzzy ART network vigilance parameters. In addition, we indicate that driver's head posture during safety verification can be categorized according to their individual properties.

1. Introduction

Recently, driver assistance systems, which deal with a driver's state and detect if drivers are able to continue driving safely, have become increasingly popular. Many researchers have been studying the detection of driver's gaze movements and sleepiness for the estimation of unsafe driving. Cognitive errors and faulty decisions cause many traffic accidents, and the primary error factor is believed to be continual deviation from normal states. We hypothesize that head motion patterns of a driver can be used for verification of safe driver conditions by detecting deviations in such motion patterns due to inattentiveness. This study aims to construct a driver assistance system that is able to detect such deviations. Recently, active safety technology designed to prevent car accidents has been drawing significant attention. There are a number of technologies that detect drowsiness or inattentive behavior using driver's eye-gaze line or head turning motion information to alert drivers of potentially dangerous situations [1]-[3]. Active safety technology is currently available in some automobiles. However, such systems only detect a single instance of deviation from normal conditions, such as inattentiveness. Our system continually analyzes time-series data to generate a predictive driver deviation signal.

Our system quantizes driver's 3D head motions during safety verification behavior using only phase variation of 2D images taken by an in-vehicle monocular camera and modeled head motion information. In this study, we analyze head postures of a driver during safety verification at an unsignalized blind intersection and propose a head posture classification method using two types of unsupervised neural networks: Self-Organizing Maps (SOMs) [4] and fuzzy Adaptive Resonance Theory (ART) [5]. In addition, we discuss face orientation categorization using the proposed method.

The remainder of this paper is organized as follows. Section II describes related work. An analysis of safety verification is presented in Section III. The proposed method is introduced in Section IV and experimental results are presented in Section V. The conclusion and future work are presented in Section VI.

2. Related study

The highest number of reported traffic accident fatalities involves pedestrians. From automobile-focused analysis of pedestrian fatalities, the Traffic Accident Research and Analysis Center has announced that 83% of accidents occur when driving a vehicle in a straight direction. In such cases, "inattentive driving," e.g., looking at distant traffic signals or operating audio equipment, and "careless driving," e.g., idly and thinking, equally contribute (35%) to 70% of the total accidents. On the other hand, 70% of pedestrian actions are violations. Seventy-three percent are crossing violations, such as

crossing immediately before or after a vehicle, crossing outside crosswalks, and ignoring traffic signals. There is a very high risk of fatal accidents when drivers succumb to inattentive or careless driving, which easily occurs when driving in a straight direction and when pedestrians violate crossing regulations. Here, "careless driving" refers to operating an automobile in a distracted state because of psychological and physiological factors. Drivers operating a vehicle in a distracted state are at risk of not perceiving potential dangers or may demonstrate delayed responses to such dangers. In a distracted state, the driver's mental and physical resources are distributed without focusing on driving behaviors and requirements, which is a common contributor to serious accidents [6]. Several studies have examined such distracted states. Homma et al. [7] focused on what they referred to as the vague state, and Abe et al. [8] examined what they referred to the thinking state. These studies promoted the experimental verification of such states and confirmed the existence of characteristic scenes in which discovery delay or oversight of changes in an ambient environment occurs.

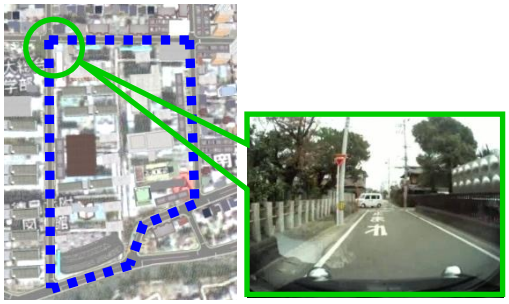
The prediction of driver's behavior is effective for the overall prevention of traffic accidents. However, individual driver characteristics are greatly influenced by states of mind, such as emotional stress and mental burden, which are not always constant. Emotional stress is an important factor in traffic accidents; irritability and impatience increase risky driving behaviors, such as short and narrow inter-vehicle distance, rapid acceleration and deceleration, and driving at high speeds. In addition, excessive anxiety generates carelessness or oversight because it interferes with cognitive processes. Matthews [9] stated that the various emotional stresses during driving must be broadly classified into the following three factors. The "degree of involvement in the driving task (Task Engagement)" involves the level of driver engagement and fatigue. "Puzzle-distress (Distress)" includes feelings of tension, pleasure–displeasure, and anger. "Anxiety (Worry)" involves cognitive interference. Individual thresholds for anger or frustration caused by conflict with another car, e.g., the other car does not run as expected, vary for each driver. However, regardless of the different thresholds and tolerances, these emotions exist as stress and affect driving behaviors. In addition, acting within time constraints or delays can easily lead to impatience, e.g., you cannot proceed as intended because of a traffic jam. Such situations can evoke emotional stress. However, driving behavior prediction that considers operating characteristics on the basis of the mental and emotion state of drivers has not yet been realized.

Many researchers have been working on the estimation of drivers' face orientation, head postures, and gaze. S. J. Lee et al. proposed a vision-based real-time gaze zone estimator that works in both day and night conditions and is sufficiently robust to recognize facial image variation caused by eyeglasses [10]. In another system, driver vigilance has been estimated by the percentage of eye closure and a fuzzy classifier using infrared images [11]. Another study focusing on natural driving environments presented an automatic calibration method and categorized the head position using 12 gaze zones with a particle filter [12]. In addition, drivers' head positions have been estimated using localized gradient orientation histograms of the facial region as input to support vector regressors [13]. A robust driver's head and facial feature tracking system, which is capable of detecting occluded eye and mouth features, has also been proposed [14]. These methods use facial orientation, driver's gaze, and the degree of eye openness to estimate the degree of driver's concentration and fatigue; these factors are realized by detecting and tracking the corresponding facial feature points. However, these approaches have some inherent technical issues, such as the failure of tracking and mismatch in the relationship between corresponding facial feature points, because changes in the driving environment lead to different degrees of light. The proposed method does not require the detection and tracking of facial feature points. We focus on time-series information of geometric phase changes in a two-dimensional space captured by a single video camera, with respect to the neutral driving position seated to the fixed position of an individual driver.

3.Analysis of safety verification behavior

Most car accidents occur near or in intersections. If safety verifications are insufficient and the driver is operating a vehicle in an abnormal state, the probability of a car accident is significantly higher. In our study, we constructed a safety verification behavior model according to an individual's behavior during safety verification motions at an unsignalized blind intersection. To construct the safety verification behavior model, we quantize head posture changes made during safety verification motions. In this section, we analyze a driver's upper body posture motions during safety verification behavior and discuss the granularity of quantization.

Figure 1. Monocular in-vehicle camera



(a) Course map (b) Unsignalized intersection
Figure 2. Driving course.

3.1. Dataset

To analyze a driver's safety verification posture, actual driving environment data were collected. The subject, a man in his twenties, gave informed consent prior to engaging in the driving exercise. The subject practiced driving the designated course before the actual data collection. The course proceeded around the University of Tokushima, and it took approximately 15 min to complete one lap. A monocular in-vehicle camera (Anshin-mini, Anshin Management Co., Ltd., see Figure 1) was set on the windshield in front of the driver's seat. We recorded the subject's upper body data (driver's images). Another camera was set behind the rearview mirror to record images in front of the car (driving scene images). The image resolution was 640 × 480 pixels, and the frame rate was 30 fps. There were some unsignalized blind intersections in the course. The course is shown in Figure 2 (a), and an example of an unsignalized blind intersection is shown in Figure 2 (b). Data were collected on three different days, resulting in three unique datasets: A, B, and C. The subject drove the course once per day.

3.2. Dataset Analysis

We analyzed right turn behavior at the intersections, as shown in Figure 2 (b). Figure 3 illustrates the subject's head posture changes during safety verification. From dataset A, the subject verified the safety of the intersection with deeply bent head postures. On the other hand, from datasets B and C, the subject made smaller bending movements. From all the dataset results, we were able to recognize a common posture while checking the convex mirrors when entering the intersection. For sufficient quantization of head postures, the classification of frontal facial features, head postures when checking a convex mirror, small head bending postures, deep head bending postures, and right and left checking postures is the minimum requirement.

4. Proposed method

The proposed method has a feature based on the hybridization of two unsupervised neural networks with a seamless mapping procedure comprising the following steps. First, on the basis of the similarity of the spatial phase structure of images, we identify a local neighborhood region containing the order of phase changes. The region is mapped to one-dimensional space equivalent to more than the optimal number of clusters. Labels that match the optimal number of clusters are generated by additional learning that is in accordance with the order of the one-dimensional maps formed in the neighborhood region.

Specifically, SOMs have excellent topological properties for spatial mapping. Using one-dimensional SOMs with a non-circular mapping layer, the neighborhood region can form an independent one-dimensional space that locally maintains order on the basis of the similarity of the structural phase. In addition, after mapping, sparse input data are reflected in the weight vectors of the mapping units. The weight vectors for each unit of the one-dimensional mapping layer maintaining this order are positioned to enter the ART network by sequentially placing them into the ART input space from the start or end of the terminal units. In reality, ART encourages the formation of adaptive category combinations with stability and plasticity to achieve the optimal number of cluster-generated labels for the target problem. Figure 4 depicts an overview of the hybridization of the two unsupervised neural networks. The proposed method has a seamless mapping procedure comprising the following two steps, as shown in Figure 4.

The first step is to classify the input patterns in feature space using one-dimensional SOMs with a non-circular mapping layer. The second step is to integrate the weight vectors of the one-dimensional SOMs into proper categories using fuzzy ART networks.

The algorithm of SOMs comprises the following steps.

1). Let wi,j(t) be the weight from the input unit i to the mapping unit j at time t. The weights are initialized with random numbers.

2). Let xi(t) be the input data to the input unit at time t.

3). The Euclidean distance dj between xi(t) and wi,j(t) is calculated as follows.

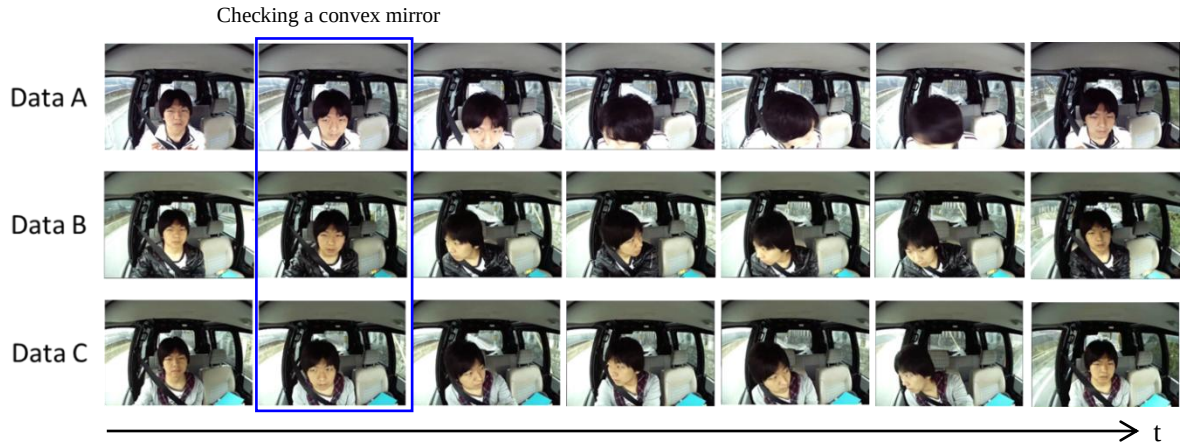


Figure 3. Head posture change during safety verification.

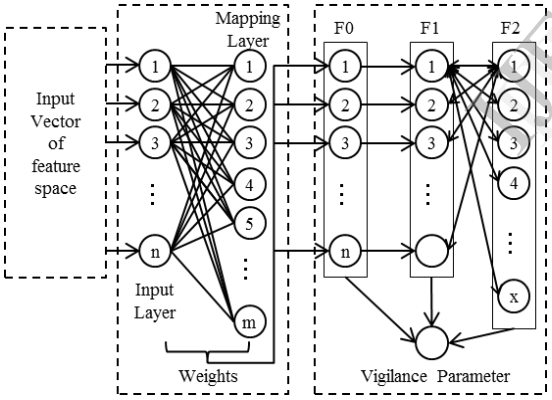


Figure 4.Network structure of proposed method.

(a) Original image (320 × 240 pixels)

(b) Region of interest

(c) Gabor wavelet filter image

(d) Coarse-grained image (24 × 24 pixels)

Figure5. Input data of SOMs.

$$d_j = \sqrt{\sum_{i=1}^{I}\left(x_i(t) - w_{i,j}(t)\right)^2} \quad (1)$$

4). The winning unit dj becomes a minimum. Let Nc(t) be the units of the winning unit neighborhood. The weight wi,j(t) inside Nc(t) is updated.

$$w_{i,j}(t+1) = w_{i,j}(t) + \alpha(t)\left(x_i(t) - w_{i,j}(t)\right) \quad (2)$$

Here, $\alpha(t)$ is the training coefficient, which decreases with time. Training ends when the iterations reach the maximum number.

In this study, the initial value of α(t) is set as 0.5, and that of Nc(t) is set as 2/3 of the number of mapping layer units, such that the values decrease linearly with time. The number of learning operations is empirically set as 1,000.

The fuzzy ART algorithm is presented below. Fuzzy ART network architecture consists of the following three fields: Field 0 (F0) receives input data, Field 1 (F1) is for feature representation, and Field 2 (F2) is for category representation.

1). wi denotes the weights between each F2 i and each corresponding F1. All wi values are initialized as one.
2). Input data x is the SOM weight wi,j for F0.
3). For each unit i in F2, the choice function Ti is defined as follows.

$$T_i = \frac{|x \wedge w_i|}{a + |w_i|} \quad (3)$$

4). T_c is a winning category, where$c$ gives the maximum value of T_i. The category with the smallest index is chosen if more than T_i is maximal. When T_c is selected for a category, the T_i unit on the c of F2 is set to one and other units are set to zero.
5). Resonance or resetting is assessed. The match function that $x \wedge w_c$ to F1 of the signal from the unit in the c of F2.

$$\frac{|x \wedge w_c|}{|i|} \geqq p \quad (4)$$

Here, x and c are resonant. Resonance occurs if the match function of the selected category meets the vigilance criterion. Then, the weight vector wi0 is updated as follows.

$$w_{i0} = r(x \wedge w_{i0}) + (1 - r)w_c \quad (5)$$

If x has no resonance with Tc, then Tc is reset. The network seeks the next category where Ti is maximal and reselects it. The network determines resonance or resets. If all categories are reset, then a unit is created on F2 and a new category is registered. Here, r denotes the learning speed parameter.

The units of F2 and SOM's weight correspond in fuzzy ART. And similar feature is integrated of category.

5. Experimental results

We examined the effectiveness of the proposed method on the basis of the safety verification analysis.

Datasets A, B, and C were used in this experiment. Figure 5 shows details of the experimental preprocessing. Figure 5 (a) is the original image (320 × 240 pixels). Figure 5 (b) is the region of interest (240 × 240 pixels) extracted from the center of the original image. Figure 5 (c) is Gabor wavelet filter image of Figure 5 (b). Figure 5 (d) is a coarse-grained image (24 × 24 pixels) of Figure 5 (c). The input features of the SOMs are the coarse-grained image pixel values (576 dimensions). We empirically set the mapping layer to 25 units.

5.1. Effectiveness of proposed method

The advantage of the proposed method is the ability to adaptively integrate categories of one-dimensional SOM mapping results to maintain neighborhood units by adjusting the fuzzy ART vigilance parameter. Tables 1–3 show the results for vigilance parameters from 0.93 to 0.97 for datasets A, B, and C, respectively. The SOM's unit number (1–25) corresponds to the category number of the integrated result from fuzzy ART. In these tables, N indicates that no data are mapped to the unit. In addition, category integration is achieved with N as the boundary and integration is conducted within the boundary areas. Specifically, focusing on mapping units from 1 to 9 in Table 2, similar units are integrated to make a neighborhood as the vigilance parameter decreases from 0.97 to 0.93. When the vigilance parameter is 0.97, units 1 and 2 are integrated into category 1. Similarly, units 3, 4, and 5, units 6 and 7, and units 8 and 9 are integrated into categories 2, 3, and 4, respectively. In contrast, when the vigilance parameter is 0.93, units 1–5 and units 6–9 are integrated into categories 1 and 2, respectively. As shown in Table 2, the boundary unit changes adaptively with the vigilance parameter.

Table 1. Category integration results (dataset

N: no data mapped to unit

Dataset A		SOMs mapping unit number																								
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25
Vigilance parameter	0.93	1	1	1	1	1	2	2	N	3	3	3	N	4	4	4	4	5	5	N	6	6	6	6	7	7
	0.94	1	1	1	2	2	N	3	3	N	4	4	4	4	5	5	5	5	6	N	7	7	7	7	8	8
	0.95	1	1	1	2	2	N	3	3	N	4	4	4	5	5	6	6	6	6	N	7	7	7	8	8	8
	0.96	1	1	2	2	2	3	3	4	4	5	5	5	6	6	7	7	7	8	8	N	9	9	10	10	11
	0.97	1	1	2	2	3	N	4	4	N	5	5	6	6	7	8	8	9	9	N	10	10	11	11	12	12

Table 2. Category integration results (dataset

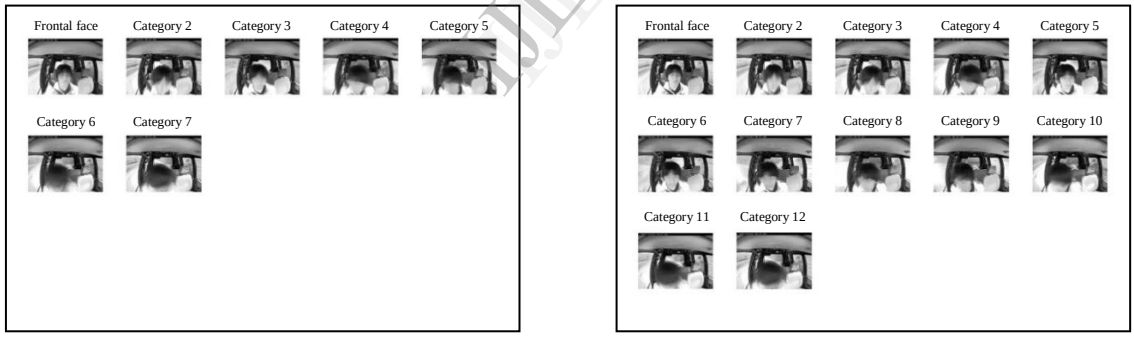
N: no data mapped to unit

Dataset B		SOMs mapping unit number																								
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25
Vigilance parameter	0.93	1	1	1	1	1	2	2	2	2	3	3	3	3	3	4	4	N	4	N	5	5	5	5	6	6
	0.94	1	1	1	1	2	2	2	3	3	3	4	4	4	5	6	N	5	N	7	7	6	N	6	7	7
	0.95	1	1	1	2	2	2	3	3	3	4	4	4	5	5	5	6	N	7	N	8	8	8	7	9	9
	0.96	1	1	1	2	2	2	3	3	4	4	5	5	6	6	7	7	7	N	8	8	9	9	10	10	11
	0.97	1	1	2	2	2	3	3	4	4	N	5	5	6	6	7	8	8	9	9	10	10	11	11	12	12

Table 3. Category integration results (dataset

N: no data mapped to unit

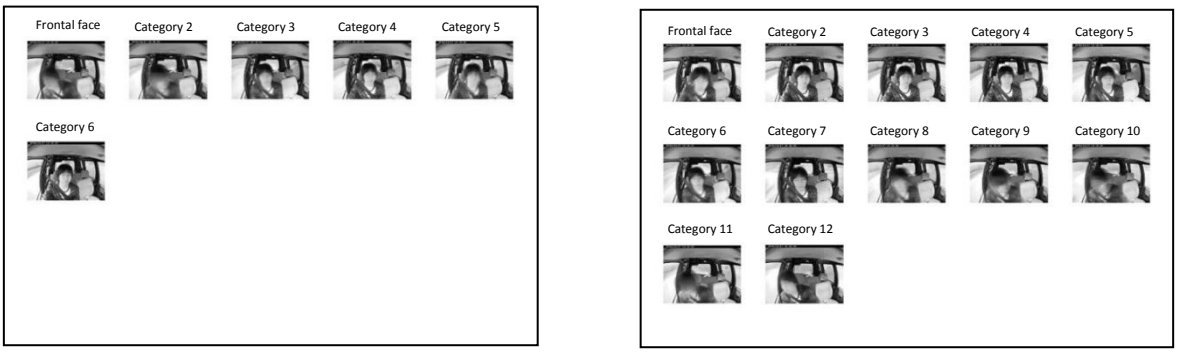
Dataset C		SOMs mapping unit number																								
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25
Vigilance parameter	0.93	1	1	1	1	1	1	2	N	2	3	3	N	3	4	4	4	5	5	5	5	6	6	6	7	7
	0.94	1	1	1	2	2	2	2	3	3	3	4	4	4	5	5	5	6	6	6	N	7	7	7	7	7
	0.95	1	1	1	2	2	2	3	3	4	4	4	5	5	N	6	6	6	7	7	7	8	8	9	9	9
	0.96	1	1	2	2	2	3	3	N	4	N	5	5	6	6	7	7	8	8	N	9	9	10	10	11	11
	0.97	1	1	2	N	3	3	4	5	5	6	7	7	N	8	N	9	N	10	10	11	11	12	12	13	13



(a) Vigilance parameter: 0.93 (b) Vigilance parameter: 0.97

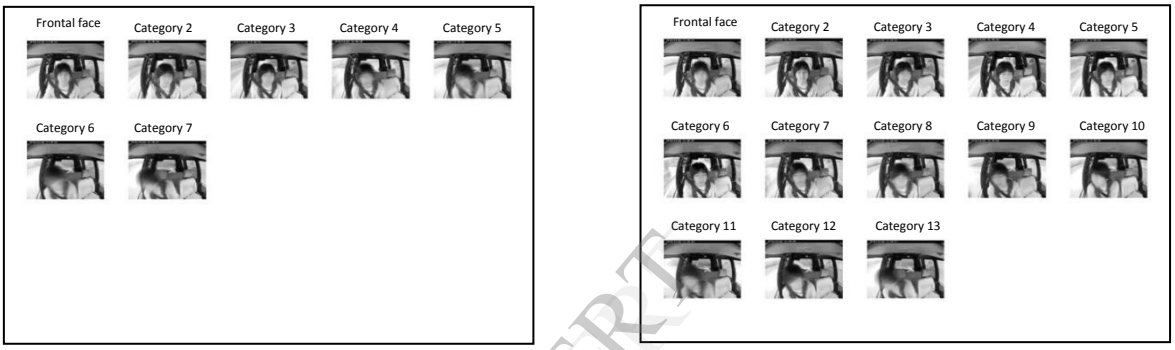
Figure 6. Average image for each integrated category (dataset A)

Figures 6, 7 and 8 show the average images from datasets A, B and C, respectively. In these figures, each safety verification posture is ordered according to the phase variation of the images on the basis of the frontal face category. Figures 6, 7 and 8 illustrate that the safety verification motion mapped by the one-dimensional SOMs is integrated with a similar head posture category and that face orientation phases adaptively with the vigilance parameter control. On the basis of the front-facing head posture, Figures 6, 7 and 8 illustrate the results of sorting and quantifying the extent of head posture changes (i.e., geometric phase changes) associated with safety



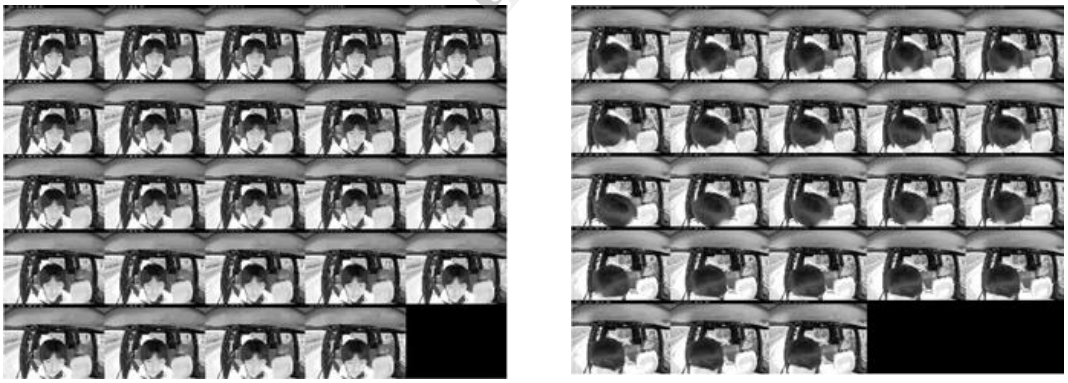
(a) Vigilance parameter: 0.93 (b) Vigilance parameter: 0.97

Figure 7. Average image for each integrated category (dataset B)



(a) Vigilance parameter: 0.93 (b) Vigilance parameter: 0.97

Figure 8. Average image for each integrated category (dataset C)



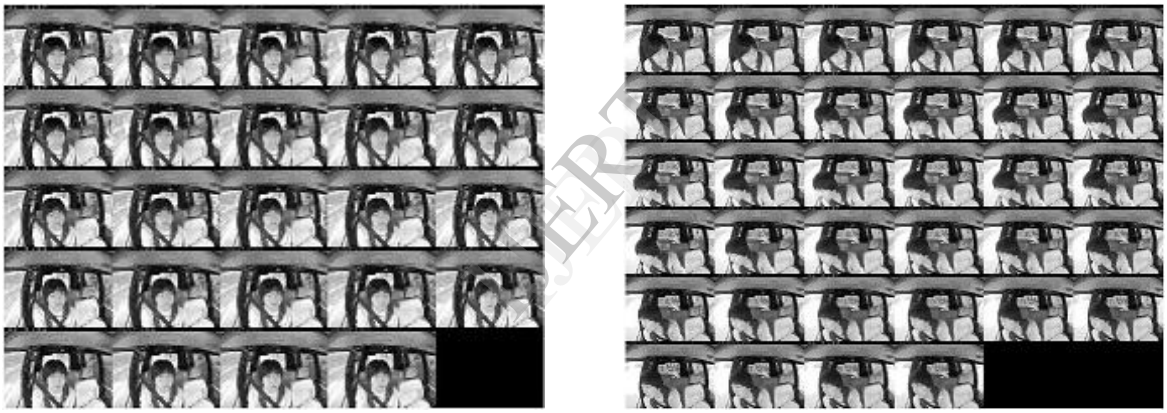
(a) Frontal face category (b) Deep head bending category

Figure 9. Head posture classification results (dataset

verification behavior. A front-facing head posture represents the neutral driving position seated to the fixed position of an individual driver. The average image in each category is calculated. All captured images corresponding to the safety verification behavior period are classified using our proposed method. Momentary head postures vary according to driving conditions. In particular, left or right head postures due to safety verification in a non-signalized intersection significantly depend on the degree of visibility. In addition, head postures associated with looking into the intersection have been confirmed. By adaptively controlling vigilance parameters, we were effectively able to analyze changes in the head posture during safety verification behavior in time-series. When vigilance parameters were set to relatively low values, e.g., 0.93, excessive and abrupt head posture changes during safety verification were

(a) Frontal face category (b) Small head bending category

Figure 10. Head posture classification results (dataset B).



(a) Frontal face category (b) Small head bending category

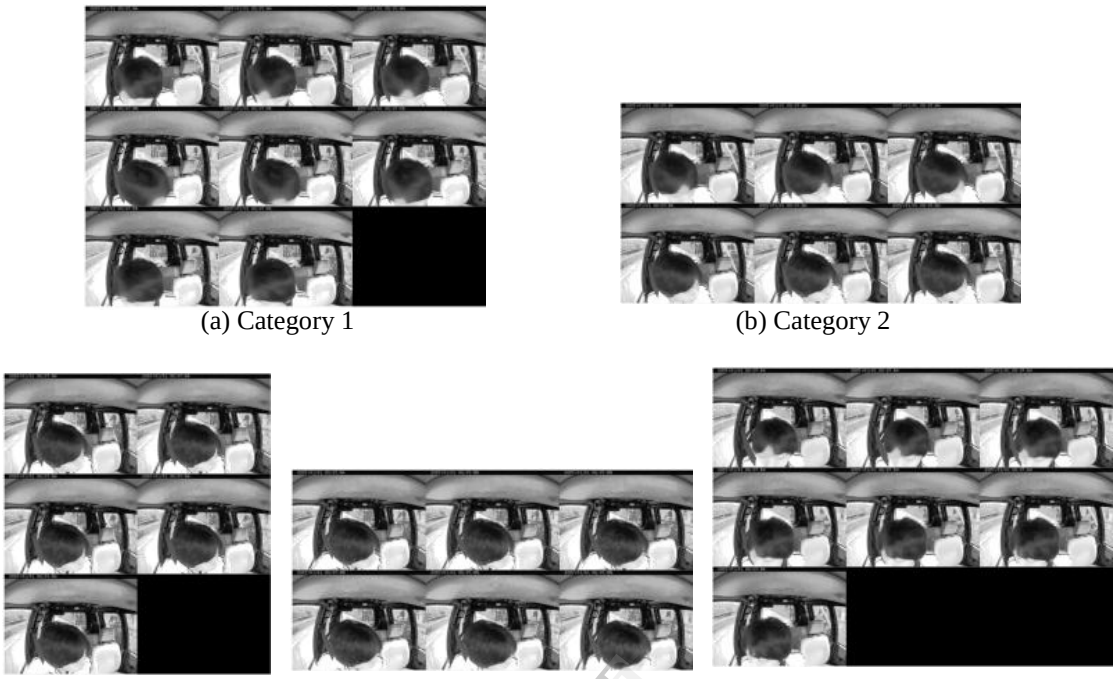
Figure 11. Head posture classification results (dataset C).

captured; however, the proposed method could not guarantee the required classification accuracy for analysis of time-series changes. On the other hand, when vigilance parameters were set to a more appropriate value, e.g., 0.97, it was possible to generate categories that appropriately represent the time-series variation of the safety verification behavior. Furthermore, by considering the safety verification behavior of a single subject at the same intersection and the analysis of datasets on different driving days, the proposed method can adaptively categorize subject-specific safety verification behaviors. For safety verification behavior with the head bending posture, as can be seen in Figures 9 (b), 10 (b) and 11 (b), right and left postures are in the same category because the proposed method quantizes the phase variations of the full region of interest and uses this information as a foundation for categorization and integration.

5.2. Recursive categorization

To sensitively separate right and left head postures and recognize the degree of head bending during safety verification, recursive categorization using the proposed method with a mixed category is expected to be effective.We examined the possibility of recursive categorization to separate right and left face posture on previous clause results.

Figure 12 depicts recursive categorization results. The proposed method could separate right and left head posture, respectively.



(a) Category 1 (b) Category 2

(c) Category 3 (d) Category 4 (e) Category 5

Figure 12. Recursive categorization results (dataset A).

6. Conclusion and future plan

In this study, we analyzed driver's head posture during safety verification at an unsignalized blind intersection and proposed a method of classifying head postures using two types of unsupervised neural networks: SOMs and fuzzy ART to quantize driver's head motion for the detection continual deviation signals. Moreover, we recursively applied the mixture head posture results to proposed method. It was demonstrated that the proposed method was able to appropriately approximate head posture categories for a driver assistance system.

In future, we will conduct proposed method on more data and define a head posture space to analyse time-si.

References

[1] T. Yonekawa, "Vision and Problem of Active Safety Technology," Symposium of JSAE, 2006.

[2] T. Inagaki, "Smart Collaboration between Humans and Machines based on Mutual Understanding," Annual Reviews in Control, vol. 32, Dec. 2008, pp. 253–261.

[3] K. Ishida, S. Hachisuka, T. Kimura, and M. Kamijo, "Comparing Trends of Sleepiness Expressions Appearance with Performance and Physiological Change Caused by Arousal Level Declining," JSAE, vol. 40, June 2009, pp. 885–890.

[4] T. Kohonen,"Self - organized formation of topologicallycorrect feature maps," Biological Cybernetics, vol. 43, Jan. 1982, pp.59–69.

[5] T. Kamio, S. Soga, H. Fujisaka, and K. Itsubori, "An adaptive state space segmentation for reinforcement learning using fuzzy-ART neural network," Proc. IEEE MWSCAS 2004, vol. 3, July 2004, pp. 117–120.

[6] J. D. Lee, M. A. Regan, and K. L. Young, "What Drivers Distraction? Distraction as a Breakdown of Multilevel Control," in Driver Distraction: Theory, Effects, and Mitigation. CRC, 2008, ch.4, pp. 41–54.

[7] R. Homma, G. Abe, K. Kikuchi, R. Iwaki, T. Fujii, and M. Fukushima, "Characteristics of visual attention while driving under the state of drowsiness," JSAE, vol. 42, Sept. 2011, pp. 1217–1222.

[8] G. Abe, K Kikuchi, R. Iwaki, and T. Fujii, "Effects of Cognitive Distraction on Driver's Visual Attention," Mechanical Systems (C), vol. 76, July 2010, pp. 14–20.

[9] G. Matthews, A. K. Emo, and G. J. Funke, "The Transactional Model of Driver Stress and Fatigue and its Implications for Driver Training," Driver Behavior and Training Volume II, Hampshire, Ashgate, 2005, pp. 273–286.

[10] S. J. Lee, J. Jo, H. G. Jung, K. R. Park, and J. Kim, "Real-Time Gaze Estimator Based on Driver's Head Orientation for Forward Collision Warning System,"IEEE Trans. Intell. Transp. Syst., vol. 12, Sept. 2010, pp. 539–553.

[11] L. M. Bergasa,J. Nuevo, M. A. Sotelo, R. Barea, and M. E. Lopez, "Real-Time System for Monitoring Driver Vigilance," IEEE Trans. Intell. Transp. Syst., vol. 7, March 2006, pp. 63–77.
[12] X. Fu, X. Guan, E. Peli, H. Liu, and G. Luo, "Automatic Calibration Method for Driver's Head Orientation in Natural Driving Environment," IEEE Trans. Intell. Transp. Syst., vol. 14, March 2013, pp. 303–312.
[13] E. Murphy-Chutorian and M. M. Trivedi, "Head Pose Estimation and Augmented Reality Tracking: An Integrated System and Evaluation for Monitoring Driver Awareness," IEEE Trans. Intell. Transp. Syst., vol. 12, June 2010, pp. 300–311.
[14] P. Smith, M. Shah, and N. da Vitoria Lobo, "Determining Driver Visual Attention With One Camera," IEEE Trans. Intell. Transp. Syst., vol. 4, Dec. 2003, pp. 205–218