

A Study on Feature Extraction Methods Used to Estimate a Driver's Level of Drowsiness

Kyong Hee Lee, Whui Kim, Hyun Kyun Choi, Byung Tae Jang

Intelligent Robotics Research Division, Electronics and Telecommunications Research Institute, 218 Gajeong-ro, Daejeon, Republic of Korea

{leekh, khsunkh, choihk, jbt}@etri.re.kr

Abstract— Recently, in addition to autonomous vehicle technology research and development, machine learning methods have been used to predict a driver's condition and emotions in order to provide information that will improve road safety. A driver's condition can be estimated not only by basic characteristics such as gender, age, and driving experience, but also by a driver's facial expressions, bio-signals, and driving behaviours. Recent developments in video processing using machine learning have enabled images obtained from cameras to be analysed with high accuracy. Therefore, based on the relationship between facial features and a driver's drowsy state, variables that reflect facial features have been established. In this paper, we proposed a method for extracting detailed features of the eyes, the mouth, and positions of the head using OpenCV and Dlib library in order to estimate a driver's level of drowsiness.

Keywords— drowsiness, blinking, yawning, head position

I. INTRODUCTION

Analyses of a driver's intrinsic characteristics and status are useful to provide interactive services for safe driving. For example, an elderly driver who feels tired after driving a long distance can be warned to rest by a warning signal on a vehicle's instrument cluster or by an audio alarm. Also, automotive air conditioning can be adjusted depending on driver's condition, as shown in Figure. 1. An automatic recognition of driver's fatigue can be used to automatically operate the air conditioner or to provide fresh air by opening the window in a vehicle.

It is necessary to develop a driver perception psychological behaviour model which reflects driver behaviour, perception and cognitive characteristics, judgment ability, and psychological characteristics. To determine the condition of a driver, we can monitor the motion of a vehicle, driving patterns, a driver's vital signs, and a driver's face using a camera.

As the latest machine-learning capable data is released, a variety of professional algorithms are constantly being introduced. Also, as training speed is improving due to the use of graphics processing units (GPU), image analysis is receiving attention in many fields. Therefore, in this paper, possible driver condition estimation techniques will be explained using image feature extraction analyses.

II. IDENTIFYING A DRIVER'S CONDITION

In general, drowsiness can be defined as a state in which the head is tilted up, tilted to one side, tilted back, or remains motionless for a few seconds. Also, the level of drowsiness can be identified by eye movement. Prior to drowsiness, there is no movement of the pupil, the frequency of blinking is reduced, or a state of yawning may be observed.

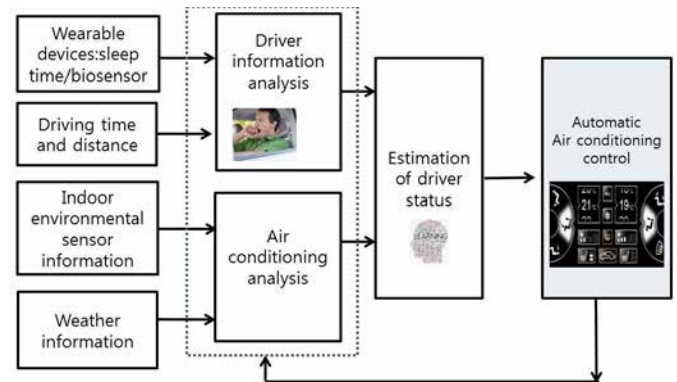


Figure 1. Smart air-conditioning using driver information analyses

To establish various spatial landmark features on the eyes, the nose, and the mouth, the facial region has to be extracted. In previous studies, these spatial features were entered into a convolutional neural network (CNN). Then, for temporal analysis, the extracted features of each frame were analysed using a recurrent neural network (RNN) [1]. However, for an accurate estimation, much training data is necessary when an RNN is used. Thus, due to a lack of data, an RNN was not used in our study.

A. Facial Characteristics During a State of Drowsiness

When people start driving, they tend to gaze at the left or right sides of the road after setting a destination. However, as time passes by, the frequency of blinking decreases as a driver becomes fatigued. A reduced rate of blinking may result in drivers closing their eyes unconsciously, and their heads may flop. These temporal behaviours can be observed continuously. Therefore, temporal changes of facial features may be analyzed to identify changes in a driver's condition. According to the flow of the time, continuous frames with spatial face landmark features have been analyzed using

variable length frames. Short video clips comprising five frames were entered into a deep convolutional neural network in [2]. In a 3D convolutional neural network, 20 continuous frames were used [3].

While yawning, blinking is not considered because some people yawn with their eyes open. It is assumed that the elements necessary for predicting a driver's state are the position of a driver's head, eye state, and yawning. A driver's state is assumed to develop from a normal state to a fatigued state and to a drowsy state. The following rule-based modelling was used to distinguish each state of a driver. The direction of the head, blinking, and yawning information were used, as shown in TABLE 1. The position of the head can be classified as either upright or tilting downward by considering the head's yaw angle or pitch angle. An interactive warning service can be employed depending on a driver's condition

TABLE 1. STATUS FEATURES AND DESCRIPTION

	Head Pose		Eye Blink	Mouth Status
Facial behavior	Upright {0,1}	Down {0,1}	Closure {0,1}	Yawn {0,1}
Status	Normal	Drowsy	Fatigue, Drowsy	Fatigue
Interactive service	None	Warning	Warning	Warning

B. Head Yaw and Pitch Angle Analysis for Estimation of Drowsiness

An analysis based on a driver's head movement examines yaw direction and pitch direction. Yaw movement refers to the left and right movements of a driver's head and pitch movement refers to up and down movements. In order to measure pitch movement, the horizontal line of one point on the face is used as a reference line. Yaw and pitch angles can also be used to determine the position of a driver's gaze. The standard for a frontal gaze was within ± 15 degrees of the yaw angle [4].

C. Monitoring Yawning for Indications of Mild Fatigue

Yawning is an unconscious behaviour done to emit carbon dioxide, to maintain body temperature, or a phenomenon expressing sympathy for others. One method employed to detect mild fatigue is to study the yawning behaviour of a driver [5].

As one indicator of fatigue, yawning corresponds to stretching of the mouth in the action unit (AU) of the Facial Action Coding System (FACS) [6]. Because yawning is characterized by a slow and wide opening of the mouth unlike talking, singing, or expressing surprise, detecting the sides of the mouth and measuring the size and shape of the mouth is necessary to identify a yawn.

D. Monitoring Eyes to Identify the Degree of Fatigue and the Level of Drowsiness

A driver's degree of fatigue or drowsiness has also been determined by prolonged eyelid closure, eye-rolling, or staring without blinking for over three seconds [5]. A person is judged to be in a state of drowsiness if his or her eyes remain closed for more than the threshold time. Even though the head is positioned upright with little variation in the yaw angle of head rotation, a driver gazing blankly indicates that the driver is in a state of drowsiness.

Fatigue is estimated by analysing blinking and yawning. Blinking is used to calculate the percentage of eyelid closure (PERCLOS) over the whole video frame or per a certain time. PERCLOS can be one of the fundamental indicators of fatigue or drowsiness. The threshold value of PERCLOS used to determine a state of drowsiness was 0.2 [7] and 0.3 ~ 0.4 [8]. PERCLOS was computed for 15, 30, and 60 seconds in [9] and for 10 minutes in [8]. In order to calculate PERCLOS, eye closure has to be defined. Eye aspect ratio (EAR) is the ratio of the height to the width of an eye, and this is used to distinguish whether eyes are open or closed. When EAR is more than the threshold value of 0.2, eyes were considered as closed in [10].

III. DATA ANALYSIS

In the previous studies, the NTHU-DDD dataset has been used to identify driver status [11]. For this dataset, subjects' facial expressions—slow blinking, prolonged eye closure, yawning, and nodding—were recorded while they were behind the wheel of a driving simulator. We used part of this dataset to examine how facial features change by utilizing three basic data: angle of head movement, rate of blinking, and yawning.

At first, in order to detect the facial landmarks on standard definition (640 x 480) video obtained from the NTHU-DDD data, which was recorded at 30 frames per second and consists of 2,670 frames over one minute and 29 seconds, we used Python 3.5 and imported libraries of OpenCV 3.4.2 and Dlib 19.4.0. These libraries can also be applied to face swapping or face morphing functions. The OpenCV library provides a pre-trained classifier for the face or the eyes as well as a detector. Dlib, which provides a face detector based on a histogram of oriented gradients (HOG) and a support vector machine (SVM), uses coordinates of 68 facial landmarks using iBUG 300-W face landmark dataset [9].

As shown in Figure. 2, the contour outlines of the two eyes consists of 12 points which are numbered from 37 to 48. The outline of the mouth consists of 20 points, numbered from 49 to 68, which were detected using OpenCV and Dlib.

Accordingly, some features were extracted from the video using OpenCV and Dlib, as shown in Figure. 3. All of the features are shown in TABLE 2. There are spatial factors such as the angle of movement of the head (AMH), eye vertical length (EVL), eye aspect ratio (EAR), mouth aspect ratio (MAR), and mouth height length (MHL) per frame. There are also temporal factors such as percentage of eye closure (PERCLOS), blinking frequency (BF), eye close count (ECC), and eye close duration (ECD), which accumulated over a certain time.

As shown in Figure. 4, the correlation heat map drawn using Pandas library presents the relationships between each factor. These are expressed in numerical values and colours. The closer to one or the closer to white, the more significant a relationship is. When we analysed the relationships between yawning and other variables to identify the level of fatigue, the correlation between yawning and MHL, YAR, EAR, and EVL was 0.77, 0.72, 0.14, and 0.13, respectively. When we analysed the relationships between slow eye movement and other variables used to determine the level of fatigue or drowsiness, PERCLOS, ECD, BF, eye close count, head angle, and yawning, the correlations measured 0.51, 0.46, 0.46, 0.41, 0.13, and 0.11, respectively. When the relationships between nodding and angle of head position, PERCLOS, ECD, the rate of eye closure, BF, and slow eye movement were analysed, the correlations measured 0.82, 0.35, 0.33, 0.27, 0.24, and 0.13, respectively.

Figure. 5 shows MAR with an open mouth, which demonstrates yawning being analysed to estimate a driver's state of fatigue. Figure. 6 shows PERCLOS, blink frequency, and ECD used to identify levels of fatigue and drowsiness. Figure. 7 shows angle of the head, PERCLOS, ECD, and head position from an upright to a downward direction per frame.

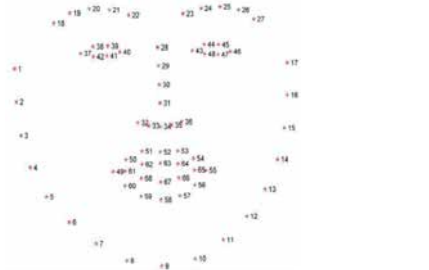


Figure 2. The 68 facial landmark points from the iBUG 300-W dataset [8]

TABLE 2. DETAILED FEATURE EXTRACTION AND DESCRIPTION

Features	Detailed Description		
	Detailed Feature Parameters	Spatial Features	Temporal Features
Head upright	Angle of movement of the head (AMH)	V	
Head down	Angle of movement of the head (AMH)	V	
Eye blinking	Percentage of eye closure (PERCLOS)		V
	Blinking frequency (BF)		V
	Eye close count (ECC)		V
	Eye close duration (ECD)		V
	Eye aspect ratio (EAR)	V	
Mouth	Eye vertical length (EVL)	V	
	Mouth aspect ratio (MAR)	V	
	Mouth height length (MHL)	V	

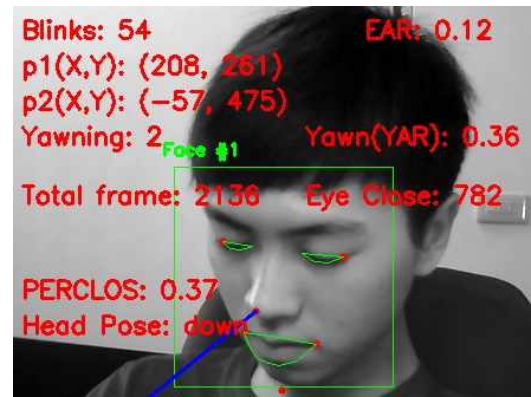


Figure 3. Features extracted from video using OpenCV and Dlib facial landmark points

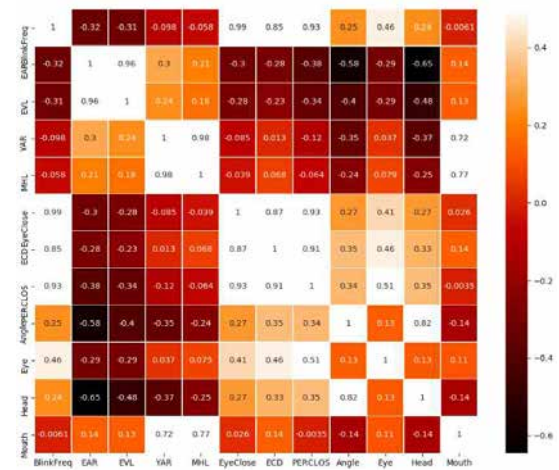


Figure 4. A heat map showing the correlation between features

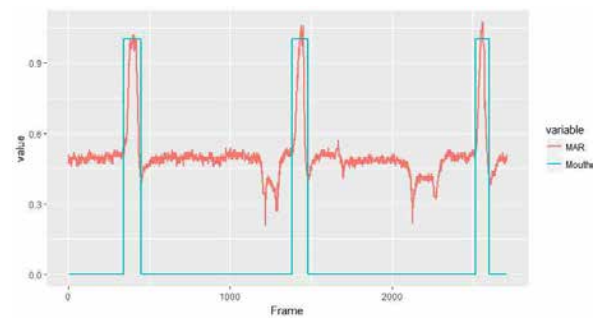


Figure 5. MAR, mouth open

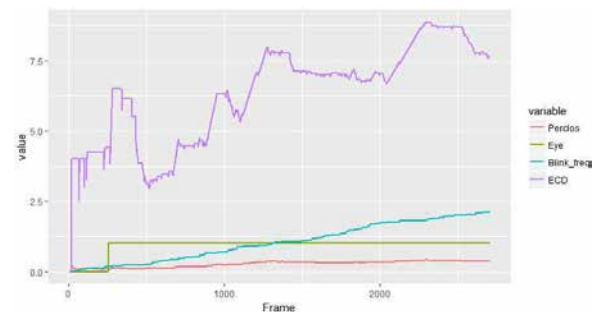


Figure 6. PERCLOS, blink frequency, ECD, drowsiness

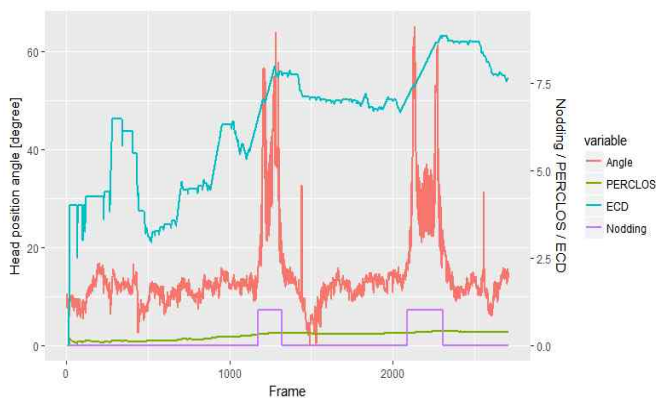


Figure 7. Head position, PERCLOS, ECD, Changes in head position

IV. CONCLUSIONS

In order to estimate a driver's condition, certain facial features were identified on the video of a male individual driving using a driving simulator. Using OpenCV and Dlib in Python, the features examined were frequency of eye closure, PERCLOS, ECD, frequency of yawning, and head positions.

There will undoubtedly be limitations concerning the detection of drivers' behaviours and facial expressions due to factors such as light reflection, the wearing of sunglasses, darkness, and obstructions by drivers' hands. Therefore, a driver's condition can be also measured by bio sensors such as EEG and ECG sensors. Furthermore, information obtained by in-vehicle environmental sensors during actual driving would be valuable. Temperature, humidity, and CO2 sensors can be used to predict a driver's condition. Monitoring travel routes and a driver's behavioural characteristics, such as driving speed, steering, acceleration, and deceleration, is also necessary to estimate a driver's fitness for driving.

ACKNOWLEDGMENT

This work was supported by an Electronics and Telecommunications Research Institute (ETRI) grant funded by the Korean government and Ulsan Metropolitan City. [18ZS1300, Development of smart context-awareness foundation techniques for major industry enhancement] [18AS1300, Development of smart HSE system and digital cockpit system based on ICT convergence for enhanced major industry]

REFERENCES

- [1] J. Jo, H.G. Jung, K. R. Park, J. Kim, "Vision-based method for detecting driver drowsiness and distraction in driver monitoring system", *Optical Engineering*, 50(12), 127202, Dec. 2011.
- [2] P. Ekman, W.V. Friesen, J.C. Hager, *Facial Action Coding System (FACS): Manual*, CD Rom, San Francisco, CA, 2002.
- [3] G. Li, and W.Y. Chung, "Detection of driver drowsiness using wavelet analysis of heart rate variability and a support vector machine classifier," *Sensors*, 13(12), pp. 16494-16511, 2013
- [4] T. Hayami, K. Matsunaga, K. Shidoji, Y. Matsuki, "Detecting Drowsiness while Driving by Measuring Eye Movement-A Pilot Study," *The IEEE 5th International Conference on Intelligent Transportation Systems*, Singapore, Sep. 2002
- [5] T. Vogelpohla, M. Kühn, T. Hummel, M. Vollrath, "Asleep at the automated wheel—Sleepiness and fatigue during highly automated

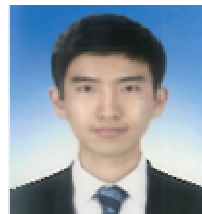
driving", <https://doi.org/10.1016/j.aap.2018.03.013>, *Accident Analysis and Prevention*, 2018

- [6] J. Lyu, "Long-term Multi-granularity Deep Framework for Driver Drowsiness Detection", *MultiarXiv:1801.02325v1*, 2018
- [7] J. M. Yu, Representation Learning, "Scene Understanding and Feature Fusion for Drowsiness Detection", in *Asian Conference on Computer Vision Workshop*, 2016
- [8] Xuan-Phung Huynh, "Detection of driver drowsiness using 3D deep neural network and semi-supervised gradient boosting machine", in *Asian Conference on Computer Vision Workshop*, 2016
- [9] C. Sagonas, G. Tzimiropoulos, S. Zafeiriou, M. Pantic. "A semi-automatic methodology for facial landmark annotation", *Proceedings of IEEE Int'l Conf. Computer Vision and Pattern Recognition (CVPR-W)*, 5th Workshop on Analysis and Modeling of Faces and Gestures (AMFG 2013), June 2013
- [10] T. Soukupova and J. Cech, "Real-time eye blink detection using facial landmarks", *21st Computer Vision Winter Workshop*, 2016
- [11] C. H. Weng, Y. H. Lai, S. H. Lai, "Driver drowsiness detection via a hierarchical temporal deep belief network", In *Asian Conference on Computer Vision Workshop on Driver Drowsiness Detection from Video*, Taipei, Taiwan, Nov. 2016



Kyong Hee Lee is a senior researcher in the Intelligent Robotics Research Division, Electronics and Telecommunications Research Institute (ETRI), Daejeon, Republic of Korea. She received her B.S. and M.S. degrees in information and communication engineering from Chonbuk National University in 1999 and 2001, respectively. She joined ETRI in 2001. Her research interests include microwave amplifiers, antennas, RFID, IoT, and machine

learning.



Whui Kim is a researcher in the Intelligent Robotics Research Division, ETRI. He received his B.S. degree in electronics and information engineering from Korea University in 2014 and his M.S. degree in computer software in 2016 from the University of Science and Technology (UST). He joined ETRI in 2016. His research interests include convolutional neural networks (CNNs) and machine learning.



Hyun kyun Choi is a director of the ETRI-Ulsan Joint Research Section, ETRI. He received his B.S. and M.S. degrees in electronic engineering from Kyungpook National University in 1995 and 1997, and his Ph.D degree in electronic engineering from Chungnam National University in 2015. He joined ETRI in 2000. His interests include intelligent transport systems, human-machine interaction, and vehicle-to-everything communications.



Byung Tae Jang is a director of the Industry IT Convergence Research Group, ETRI. He received his B.S. degree in atmospheric science from Seoul National University in 1989 and his M.S. and Ph.D degrees in computer science from Chungnam National University in 1994 and 2001. From 1989 to 1996, he was a senior researcher at the Software Engineering Research Institute, Daejeon. He joined ETRI in 1997. His current research topics include human computer interaction (HCI), and smart health, safety, and environments in shipyards.