Drowsiness Detection Method during Driving by using Infrared and Depth Pictures

Gang-chon You[1], Do-hyun Park[1], Soon-kak Kwon[1,*]

Abstract

In this paper, we propose the drowsiness detection method for car driver. This paper determines whether or not the driver`s eyes are closed using the depth and infrared videos. The proposed method has the advantage to detect drowsiness without being affected by illumination. The proposed method detects a face in the depth picture by using the fact that the nose is closest to the camera. The driver's eyes are detected by using the extraction of harr-like feature within the detected face region. This method considers to be drowsiness if eyes are closed for a certain period of time. Simulation results show the drowsiness detection performance for the proposed method.

Key Words: Drowsiness Detection, Eyes Detection, Haar-like Feature, Infrared Video, Depth Video.

I. INTRODUCTION

According to traffic accident statistics of highway traffic in the last 10 years, the traffic accidents caused by drowsiness have been higher than the traffic accidents caused by over speeding. The number of traffic accidents may be reduced significantly if actions to try to wake up driver`s sleeping is taken through detecting drowsiness during driving. Several methods to detect drowsiness during driving have been studied. Among the methods for detecting drowsiness, there is a method of continuously observing the driver`s condition by utilizing biological signals such as EDA, HRV, ECG, and EOG [1]. The method of eye detection through corneal reflex of pupil using two lights, an on-axis light and an off-axis light, to detect drowsiness is studied [2]. These methods detect drowsiness accurately, but they are needed to install the additional devices. Methods of detecting drowsiness through the color image analysis without the installation of these additional devices have been studied as follows: the method of extracting facial feature points using regression tree ensemble algorithm to detect yawning, eyes, and a direction of head[3]; the method of determining the drowsiness by the frequency of eye blinking and the duration of eye closure[4]; the method for detecting the region that the sum of absolute difference(SAD) of the reference image is more than a threshold[5]; the method for determination of the eye closing when the number of black pixels is below the threshold[6]; the method of determining the drowsiness by measuring the time when the black area of the eye area extracted using the Haar-like Feature[7]. However, the conventional methods depended on the color picture has disadvantages that the accuracy of the face detection is reduced when the face is on the side. In addition, these methods little detect the face in a dark environment such as a nighttime. This is the fatal disadvantage because drowsiness during driving occurs mainly at night.

The infrared or depth picture can be used instead of color picture in detection of face and eyes. The infrared and depth pictures have the advantage that the change of pixel is less in change of illumination compared to color image.

In this paper, the infrared and depth pictures are used to detect the face and eyes. The face of a driver is detected by using the depth picture. After that, eyes in the detected face is found by using the infrared picture. In the eyes detection, the eyes are found by Haar-like

features.

II. Detection Method of Drowsiness during driving

In this paper, we detect the drowsiness during driving by using both the infrared and depth pictures. First, we detect the face by using the depth picture. After that, we detect the drowsiness by extracting Haar-like feature. The flowchart of the proposed method is shown in Fig. 1.
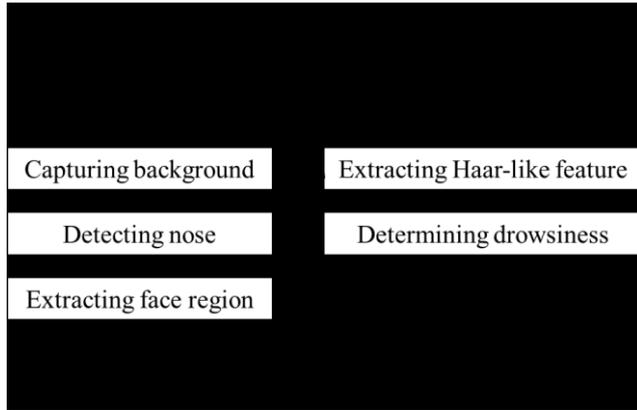

Fig. 1. Flowchart of proposed method.

2.1. Haar-like Feature

Regions used in the Haar-like features are defined by rectangles that have various patterns consisting of in the light and dark areas as shown in Fig. 2. Haar-like method finds the meaningful features based on differences in pixel values in each region. To extract the features, the Haar cascade method is used. Haar cascade method subtracts the brightness value corresponding to the black part and the white part in the image, and finds out the threshold value. Since the addition and subtraction of brightness values are inefficient, the integral image method is used.
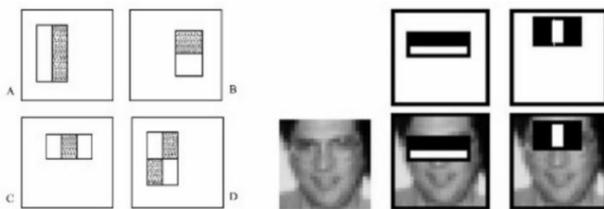

Fig. 2. Haar-like features.

The integral image method starts from the (0, 0) point of each pixel in the original image in Fig. 3, and continuously accumulates the pixel value. The pixel is moved to the next pixel to generate an integral image. If the Haar cascade method is only used, the pixel value in D must be added to obtain the sum of the brightness of D. However, if the sum of brightness of D is obtained by using the integral image, the sum of brightness of D is calculated by the sum of brightness in 1, 2, 3, and 4 in

Fig. 3. Therefore, if the sum of the brightness of the rectangle from the origin to each pixel is stored as the integral image, the sum of pixel brightness values of a specific region can be obtained by adding and subtracting the brightness value of each region.
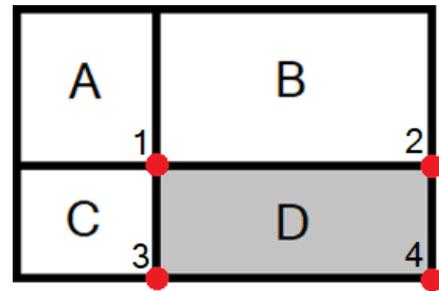

Fig. 3. Integral method of Haar cascade method.

The next step is AdaBoost, which is the combination of 'Adaptive' and 'Boosting'. AdaBoost is a way to amplify the performance of the final strong classifier by learning step by step supplementing simple weak classifiers complemented as shown in Fig. 4. In the boosting process, we combine weak classifiers with low prediction performance to generate one strong classifier with better performance. In the adaptive procedure, when the weak classifiers are sequentially learned one at a time, the information obtained by misclassifying the learned classifiers is reflected in the learning of the next classifier to compensate for the disadvantages of the previous classifiers. The user can focus on the data that is misclassified and learn and classify the data. The final strong classifier can be obtained by applying weights to each weak classifier and combining them.
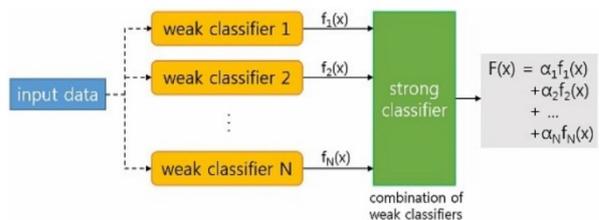

Fig. 4. Weak and strong classifiers in Adaboost.

After the Adaboost process, we use the cascade process to detect the object by sliding the boosted strong

190

classifier on the input image. This maintains the characteristics of the object, slides a simple strong classifier first, and slides a stronger classifier with a slightly stronger probability of not being an object when passing through. In this case, if we fail to pass, we will skip the non-object part in such a way that the part will be skipped during the next sliding, finding the final result. As the cascade method progresses, the area used for detection becomes smaller, so the amount of computation decreases and the speed increases.
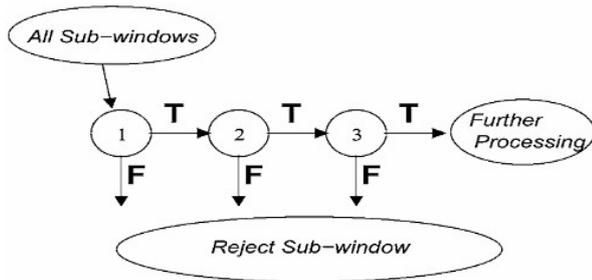

Fig. 5. Cascade process in Haar cascade method.

2.2 Detection of Drowsiness by using Infrared and Depth Pictures

First, the background image is taken to separate the driver object from the depth image. In the background image which is photographed later, the reference background is subjected to the arithmetic operation, labeling, and morphological operation to extract the object and remove the noise. If the drowsiness is detected through the depth image. Even if the passengers other than the driver who is in the car are photographed at the same time, since the infrared ray and the depth sensor are located closest to the driver, the driver is generally the largest object. It is possible to label the driver as the largest object without disturbing the user.
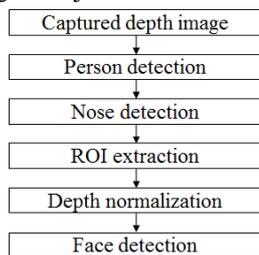

Fig. 6. Face detection flow chart.

To detect the driver's face, the driver's body is captured by the depth camera. In order to extract the driver from the captured picture, we obtain the background by the depth camera. Background and foreground are separated by obtained background picture.

The nose is generally located at the closest distance to the depth camera, and the position of the nose becomes the pixel with the minimum depth value. The neighboring pixel feature and the facial structural feature

are utilized in order to prevent the jaw, hand, chest, etc. other than the nose end point from being mistaken.
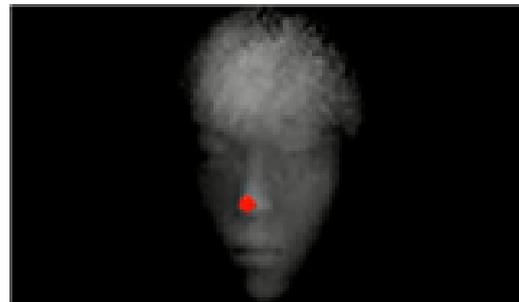

Fig. 7. Nose end point recognition in depth image.

The feature of the adjacent pixel is that the depth value of the upper and lower right and left pixels of the nose end point, the depth value of the lower jaw and the depth value of the jaw are larger. The face structural feature is that the nose end point is located at the center of the face, the background area and the width of a common human face. In order to detect the position of the eye in the extracted driver object, a pixel corresponding to the nose end point is searched, and the eye region is detected by cropping the pixel corresponding to the nose end point by a predetermined magnitude. In order to find the end point of the nose, the binarization is performed and the search is performed in the horizontal direction. Then, the N consecutive pixels from the left to the right with respect to the P_i pixel are searched for pixels having a depth value larger than that of the immediately preceding pixel by equation (1).

$$p_k > p_{k+1} \ (k = i - N, \dots, i) \qquad (1)$$
$$p_k < p_{k+1} \ (k = i, \dots, i + N),$$

where, i is the position of current pixel and N is the range of finding pixels.

This is applied to both vertical and horizontal pixels, and a pixel satisfying both of them is used as a candidate pixel. Thereafter, eight neighboring points separated from P_i by a predetermined number of pixels are compared with respect to the candidate pixels, and it is checked whether the compared pixels have a greater depth value than the candidate pixel P_i. This is based on the fact that the nose around the face is more protruded than the other areas. Then, the coordinates of the fixed region based on the coordinates of the nose of the depth image obtained by depth acquisition are assigned to the infrared image and the region is designated as the region of interest as shown in Fig. 8[9].

Fig. 8. Extraction of eye region from infrared image.

In order to detect the eyes of the driver in the extracted eye region of Fig. 9, Haar-like feature is applied and two different cascades are applied to improve the accuracy.
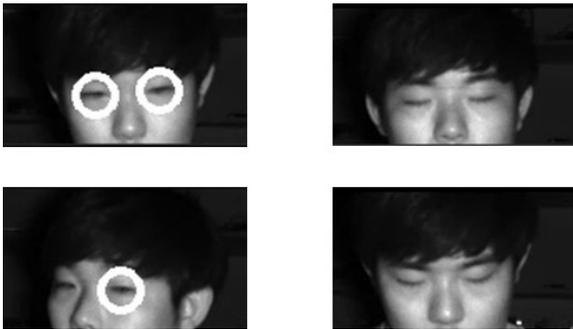

Fig. 9. Eye detection through Haar-like Feature.

From the features of the eyes, the drowsiness state is found as follows [8]. In normal state, the fast and sharp flicker can be found. At the beginning of drowsiness, long flicker is repeatedly found. In drowsiness, closed eyes is found continuously. The each eyes are detected through the Haar-like feature, the eye can be detected even if the driver turns the direction of the head to the side. Drivers at early and drowsy states can often determine whether or not they are drowsy in the eye. Therefore, if the eyes are not detected for about 2 seconds when the face of the driver is normally detected, it is determined that the driver is in a drowsy state and the driver is warned of drowsy driving.

III. Simulation Result

In order to measure the accuracy of the proposed drowsiness detection, we use Kinect v2 as capturing the depth picture and the infrared picture. The specification of device for simulation is shown in Table 1.

Table 1. Specification of Kinect V2.

Color	Resolution	1920×1080
	FPS	30
Depth	Resolution	512×424
	FPS	30
Depth acquisition range		0.5 ~ 8.0m
Person detection range		0.5 ~ 4.5m
Degree	Horizontal	70 degrees
	Vertical	60 degrees

In the bright environment and the dark environment, depth images of the front and side of the person are taken for 30 seconds for each measurement, and the total frames and the frames in which the eyes were detected were measured. Table 2 shows the accuracy measurement results. In this result, the face is detected with the accuracy of 88.3% in the bright environment and is detected with the accuracy of 90.8% in the dark environment.

Table 2. Accuracy measurements of face detection.

Brightness of illumination	Face direction	Total frames	Eye detection frames	Detection rate
Bright	Front	2089	1804	86.4%
	Side	2088	1887	90.4%
Dark	Front	2098	1864	88.8%
	Side	1971	1834	93.0%
Total		8246	7389	89.6%

Table 3 shows the accuracy of the drowsiness detection. We determine that it is the drowsiness state when eyes are kept closed for more than 1500ms in consideration of temporary blinking of eyes. In this simulation, the drowsiness detection accuracy is more than 99% irrespective of the illumination environment.

Table 3. Measurement of accuracy of the drowsiness detection.

Brightness of illumination	Total frames	Detection Frames	Accuracy of drowsiness detection
Bright	1356	1354	99.86%
Dark	1454	1444	99.32%
Total	2810	2798	99.57%

IV. Conclusion

In this paper, we propose the method of the drowsiness detection by using the depth and infrared pictures. The proposed method uses the depth picture to detect the face and obtain the region of the detected face. After that, the obtained region is assigned to the infrared picture. The assigned region in the infrared picture becomes an input of the extracting Haar-like features. The eyes are detected by Haar-like features of the face. If the eye is not detected for a certain time while the face is detected,

the method determines the face as a drowsy state. This method solves the problem that eyes are not detected in the dark environment. By detecting the drowsiness in dark environment through the proposed method, it is expected that the number of traffic accident caused by drowsiness is reduced significantly.

Acknowledgement

REFERENCES

[1] W. S. Park, J. W. Choi, T. M. Kim, and Y. K. Yang, "Drowsy-driving Prevention Techniques by BP Algorithm Using Electro Oculomoor Graphy and HRV," *Proceeding of the Conference of Korean Society for Geospatial Information System*, pp. 194-199, 2007.

[2] X. G. Zhang, J. G. Kim, and J. I. Park, "Eye Blink Detection Method for Drowsy Driving Detection System," *Proceeding of the Conference of Institute of Electronics Engineers of Korea*, pp. 482-483, 2016.

[3] M. Y. Oh, Y. S. Jeong, and K. H. Park, "Driver Drowsiness Detection Algorithm based on Facial Features," *Journal of Korea Multimedia Society*, vol. 19, no. 11, pp. 1852-1861, 2016.

[4] S. M. Kang and K. M. Huh, "Development of a Drowsiness Detection System using Machine Vision," *Journal of Institute of Control, Robotics and Systems*, vol. 22, no. 4, pp. 266-270, 2016.

[5] Y. K. Lee, S. K. Yeom, "Drowsy driver warning with eye recognition," *Proceeding of the Conference of Institute of Electronics Engineers of Korea*, pp. 329-330, 2010.

[6] D. M. Kim, H. J. Wi, J. H. Kim, H. C. Shin, "Drowsy Driving Detection Using Facial Recognition System", *Proceeding of the Conference of Korea Information Science Society*, pp. 2007-2009, 2015.

[7] H. S. Noh, P. S. Shin, "Drowsiness Warning System using image processing," *Proceeding of the Conference of the Korean Institute of Electrical Engineers*, pp. 369-371, 2012.

[8] B. J. Kim, S. S. Park, S. G. Oh, I. Y. Kim, N. G. Kim, "A Study on the Driver's Drowsiness Detection and Monitoring System," *Journal of Institute of Control, Robotics and Systems*, vol. 1, no. 8, pp. 887-890, 1997.

[9] S. K. Kwon, H. J. Kim, and D. S. Lee, "Face Recognition Method Based on Local Binary Pattern using Depth Images," *Journal of the Korea Industrial Information Systems Research*, vol. 22, no. 6, pp. 39-45, 2017.

Authors

Gang-chon You(enemflsmsto@naver.com) is currently an undergraduate student in the Department of Computer Software Engineering at Dongeui University. His research interest is in the areas of image recognition.

Do-hyun Park (waxc@naver.com) is currently an undergraduate student in the Department of Computer Software Engineering at Dongeui University. His research interest is in the areas of image recognition.

Soon-kak Kwon(skkwon@deu.ac.kr) received the B.S. degree in Electronic Engineering from Kyungpook National University, in 1990, the M.S. and Ph.D. degrees in Electrical Engineering from Korea Advanced Institute of Science and Technology (KAIST), in 1992 and 1998, respectively. From 1998 to 2000, he was a team manager at Technology Appraisal Center of Korea Technology Guarantee Fund. Since 2001, he has been a faculty member of Dongeui University, where he is now a professor in the Department of Computer Software Engineering. From 2003 to 2004, he was a visiting professor of the Department of Electrical Engineering in the University of Texas at Arlington. From 2010 to 2011, he was an international visiting research associate in the School of Engineering and Advanced Technology in Massey University. Prof. Kwon received the awards, Leading Engineers of the World 2008 and Foremost Engineers of the World 2008, from IBC, and best papers from Korea Multimedia Society, respectively. His biographical profile has been included in the 2008~2014, 2017 Editions of Marquis Who's Who in the World and the 2009/2010 Edition of IBC Outstanding 2000 Intellectuals of the 21st Century. His research interests are in the areas of image processing, video processing, and video transmission.