Driver Monitoring System Based on Gaze Classification by Using Eye Center Location Recognition

Juyong Lee[1], Seungdo Jeong[2,*] and Jihoon Lee[3]

*[1,2,3]Department of Smart Information & Telecommunication Engineering,
Sangmyung University, Cheonan, South Korea*
Juyonglee0208@gmail.com, sdjeong@smu.ac.kr, vincent@smu.ac.kr

Abstract

Due to the rapid development of IT technology, a variety of technologies are being developed to improve the safety of drivers by using various sensors and cameras. Recently, various studies have been conducted to improve the safety of the drivers by using a relatively low-cost camera instead of the expensive sensors. However, the problem is that the existing camera system cannot accurately predict the driver's sight due to various constraints such as camera angle and illuminance. Therefore, in this paper, we propose an algorithm that can improve the safety of the driver by using predicted gaze and drowsiness prediction using eye center information. The proposed algorithm consists of two steps that predict the gazing region by recognizing the center of the eye and a stage that recognize the drowsiness based on the closing time of the eyes. Through performance analysis, we confirmed that the proposed algorithm accurately predicts the sight of the driver from various angles. In addition, we set up scenarios for driving and conducted additional performance analysis, which confirmed that it could help improving driver safety.

1. Introduction

As IT technology has developed rapidly, ADAS (Advanced Driver Assistance System) that assures driver safety has been rapidly developed. Especially, the intelligent vehicle safety technology focusing on the stability enhancement is most actively developed and studied because it is directly related to the safety of the driver and the pedestrian. The intelligent vehicle safety systems are equipped with cameras or sensors inside and outside the vehicle to detect the danger which is difficult to detect for the drivers. Therefore, cameras or sensors to determine the risks are the most important factor in the intelligent vehicle safety technology. Sensors used in the intelligent vehicle safety systems include active sensors such as Radar and passive sensors that utilize visible light band images. Although active sensor-based technology has advantages of high processing speed the sensor is bulky and expensive. On the other hand, passive sensor-based technology is a method of detecting lanes, vehicles, pedestrians, and obstacles by analyzing images acquired with a camera. Image-based technology has less accuracy and slow processing speed than active sensor-based technology. However, the acquired images can perform various roles and distinguish the shape of the object. Also, it is advantageous because of the small size and low cost image sensor and it can be applied without additional cost. However, the number of serious car crashes is still increasing regardless of the intelligent vehicle safety technologies and new sensor technologies. There are studies that drivers'

---

drowsiness and carelessness cause many accidents [1-3]. In summary, driver drowsiness and carelessness are critical factors behind road accidents. To reduce the number of road accidents, it is necessary to monitor driver and driving behavior. If it could predict unsafe situations in advance, this would contribute to safe driving. It is an important factor for a driver monitoring system to recognize the drowsiness and carelessness of the drivers. Thus, eye movement and blinking information are very important factor [4- 6]. Also, to detect driver distraction, it is necessary to extract head pose or gaze information efficiently. We propose an isophote based eye gaze recognition and drowsiness detection system based on driver fatigue to design a monitoring system ensuring driver safety. The organization of this paper is as follows. Section2 describes a classifier for object recognition and an isophote for finding the center of the eye. Section 3 describes our proposed monitoring system for driver safety. The performance analysis for the proposed monitoring system is described in Section 4. Section 5 summarizes our paper in briefly.

2. Related Work

2.1. Harr Feature-based Cascade Classifier

Among diverse object detection methods, Haar feature-based cascade classifiers is known as an effective and efficient method. Since it is machine learning based approach, it needs a lot of positive and negative images to extract the feature of object. The positive image means an image of the targeting object, and the negative image means an image containing others than the targeting object. After training using these images, we need to extract features from it. Each feature is a single value obtained by subtracting the total numbers of the pixels under white and black rectangle. But, it needs so many computations. For example, 24x24 window generate over 160,000 features. To solve this computation overhead, the Adaboost algorithm appeared to recognize faces in videos or images [7]. The AdaBoost algorithm performs a linear combination to construct a strong classifier. Thus, the Adaboost algorithm combines weak classifiers to create a strong classifier that performs better. This makes it possible to detect objects efficiently because it performs computations faster than existing classifiers.
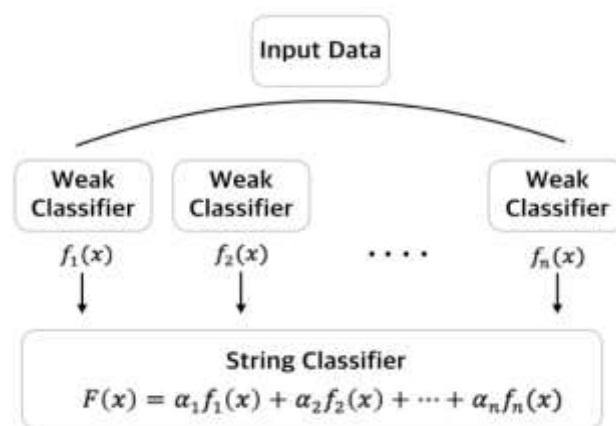

Figure 1. The Adaboost Algorithm Operation Process

2.2. Isophote Curvature

The isophote of images means curves connecting points of equal intensity. That is, the isophote is considered as contour lines obtained by slicing the intensity landscape with a horizontal flat. Above all, isophote has been used for object detection sphere because it is not affected by rotation and linear lighting changes [8]. In every point of the image, a

local coordinate frame is fixed. Also, the local coordinate frame points in the direction of the maximal change of the intensity, which corresponds to the direction of the gradient. Gauss filtering method is used to compute image derivatives [9].

$$\hat{w} = \frac{\{L_x, L_y\}}{\sqrt{L_x^2 + L_y^2}} ; \quad \hat{v} = \perp \hat{w} \tag{1}$$

The equation (1) represents frame vectors ($\hat{w}$) and $\hat{v}$ of the reference frame $\{v, w\}$ which also referred to as that *gauge coordinates*. . L_x and L_y of x and y dimension function are the first-order derivatives of the luminance function. In the case of the *gauge coordinate* system, $L(v, w(v)) = constant$ is an isophote and w'' of the tangent vector is an isophote curvature. The equation (2) is derived by differentiating the isophote definition to v. In the gauge state, $w' = 0$ because $L_v = 0$.

$$L_v + L_w w' = 0; \quad w' = -\frac{L_v}{L_w} \tag{2}$$

The equation (3) is derived by differentiating the equation (2) with *v* again.

$$L_{vv} + 2L_{vw}w' + L_{vw}w'^2 + L_w w'' = 0 \tag{3}$$

Considering $\mathcal{K}$ as isophote curvature, which is the reciprocal of the subtended radius r, can be computed as the equation (4). (Because $L_v = 0$ and $\mathcal{K} = 0$, as mentioned above.)

$$\mathcal{K} = \frac{1}{r} = w' = -\frac{L_v}{L_w} = -\frac{L_y^2 L_{xx} - 2L_x L_{xy} L_y + L_x^2 L_{yy}}{(L_x^2 + L_y^2)^{\frac{3}{2}}} \tag{4}$$

In the isophote curvature, both direction and orientation are important factors for finding the center of the eye. The orientation can be estimated from the gradient. Also, the direction of the orientation points towards the highest change in the luminance. Because the sign of the isophote curvature depends on the intensity of the outer side of the curve, they multiply the gradients and the inverses of the isophote curvature for better estimation [10].

$$D(x, y) = \frac{\{L_x, L_y\}}{L_w}\left(-\frac{L_w}{L_{vv}}\right) = -\frac{\{L_x, L_y\}}{L_{vv}} = -\frac{\{L_x, L_y\}(L_x^2 + L_y^2)}{L_y^2 L_{xx} - 2L_x L_{xy} L_y + L_x^2 L_{yy}} \tag{5}$$

Also, they define the displacement vectors $D(x, y)$ to predict the center of the eye. It also can be mapped into an accumulator and it is defined as *center map*. Because they generate center map using the equation (5), they convolve the accumulator so that each cluster of votes will generate a single center estimate. In real world, howerver, we have no guarantee that the boundaries of an object have the same intensity. In such case, every vote for the isophote center will produce inaccurate results since the shape of the isophotes differs from the shape of the object. To improve accuracy and cope with this problem, they use the curvedness that indicating how much curved a shape is.

$$curvedness = \sqrt{L_{xx}^2 + 2L_{xy}^2 + L_{yy}^2} \tag{6}$$

The curvedness has low response on flat surfaces and edges, whereas it yields high response in places where the isophote density is maximal. That is, the isophote density is maximal around the edges of an object and they select the parts of the isophotes where the

curvedness is maximal. Therefore, we can discriminate between dark and bright centers by utilizing the sign of the curvature.

3. The Proposed Monitoring Algorithm for Driver Safety

The main purpose of the proposed monitoring algorithm is to recognize driver's gaze and drowsiness to guarantee the safety of the driver. Figure 2 shows the proposed monitoring algorithm for the safety of the driver. The proposed monitoring algorithm consists of two steps: one is to predict the eye gaze and the other is to recognize drowsiness. By predicting the exact gaze of the driver and drowsiness, it can improve the safety of the driver. Therefore, the proposed monitoring algorithm can help improve the safety of the driver.
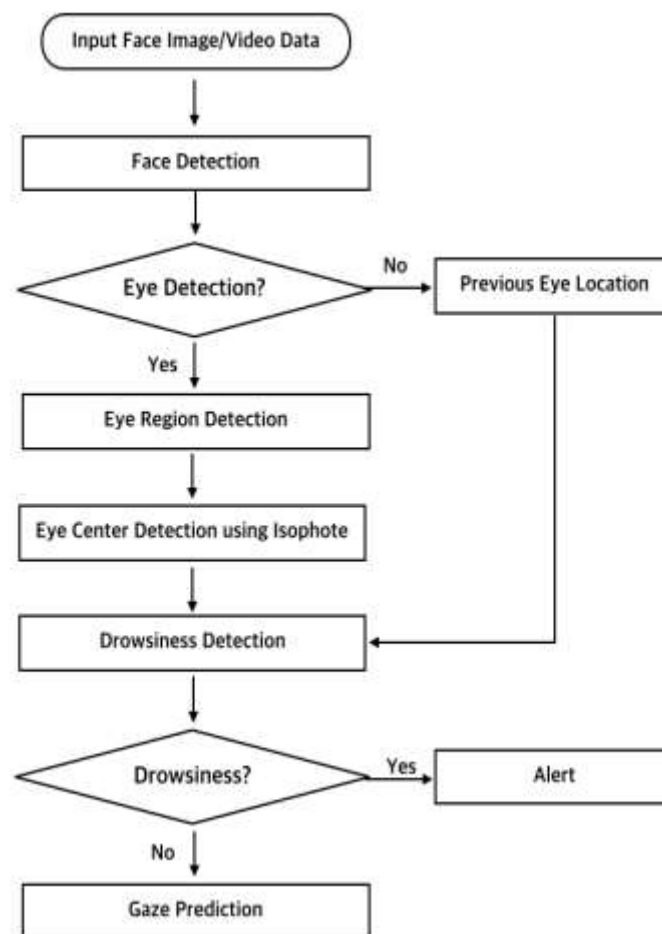

Figure 2. The Proposed Algorithm Flow Chart for the Safety of the Driver

3.1. Gaze Prediction

The most common way to detect the position of the eye is to utilize a specific classifier [11, 12]. However, the problem is that the position of the eye cannot be accurately detected when the eye flickers. So, the safety of the driver cannot be guaranteed. To increase the accuracy of eye detection, the proposed algorithm uses a method of detecting an area of the eye. This is because there have been problems that the position of the eye cannot be accurately detected when the eye flickers or closed. Figure 3 shows a case where the eyes are not recognized when the eyes are closed. To solve this problem, the

proposed algorithm additionally operates a method of detecting an area of the eye based on the previous range of the eyes.
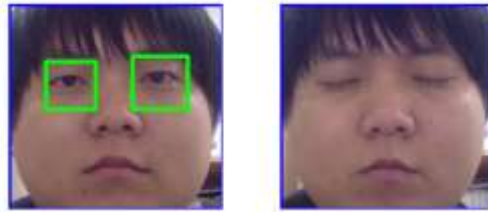

Figure 3. Eye Detection and Cognitive Problem in Eyes Closed State

When the eyes are detected, it detects the range including the two eyes. If the eye has not been recognized due to the eye flickers or eye close, it has a range including two eyes in the database and predicts the position based on the database. Figure 4 shows the situation in which the position of the eyes are predicted based on the range data of the eyes. In other words, the proposed algorithm utilizes the database to recognize the position of the eyes, so that it is possible to flexibly cope with closing or blinking eyes.

(a) Eye region detection from the face   (b) Eye region detection based on previous state
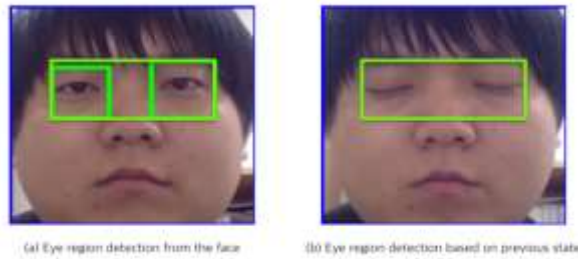
Figure 4. The Examples of the Eyes Region Detection
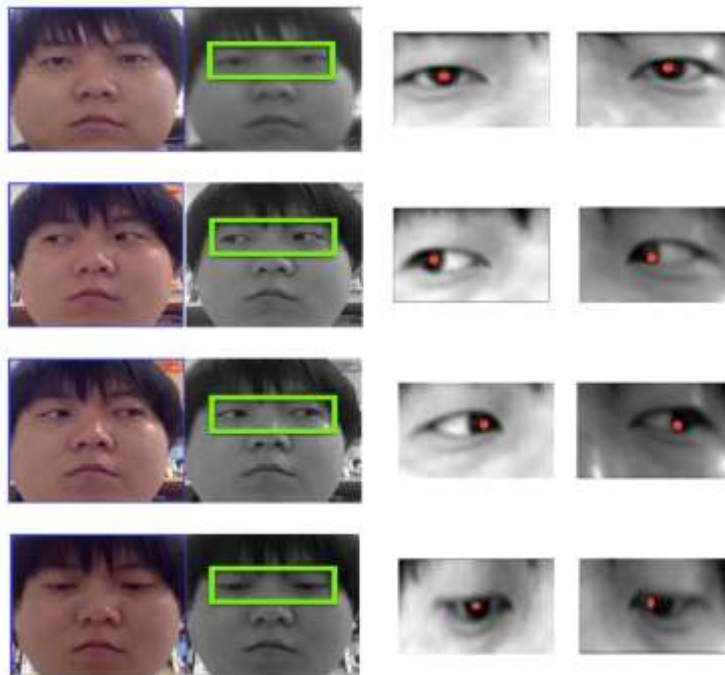
**Figure 5. Eye Center Image Examples with the Isophote Applied**

When the position of the eyes is detected, a step of calculating the isophote curvature is performed to find the center of the eye. To calculate the isophote curvedness, we also apply Gaussian blur effect to eye region. Details of the isophote curvature and the isophote curvedness operations are described in Section 2. Figure 5 shows the eye center images with isophote applied.

If the head position is located at the center as shown in Figure 5, the driver's gaze can be predicted by using the center position of the eye. However, if the position of the head is located to the left or right as shown in Figure 6, it is difficult to predict the gaze accurately with only the center position of the eye.

(a) Head Position – Center   (a) Head Position – Right

Figure 6. Inaccuracy of Eye Gaze Prediction by Head Position

To predict the field of the view efficiently regardless of the face position, the proposed algorithm predicts the direction based on the difference of the center coordinates of each eye. In other words, if the position of the head is inclined toward the right, the position of the left eye is relatively higher than the position of the right eye, so that additional information for visual prediction can be obtained.

(a) Head Position – Left   (b) Head Position – Right
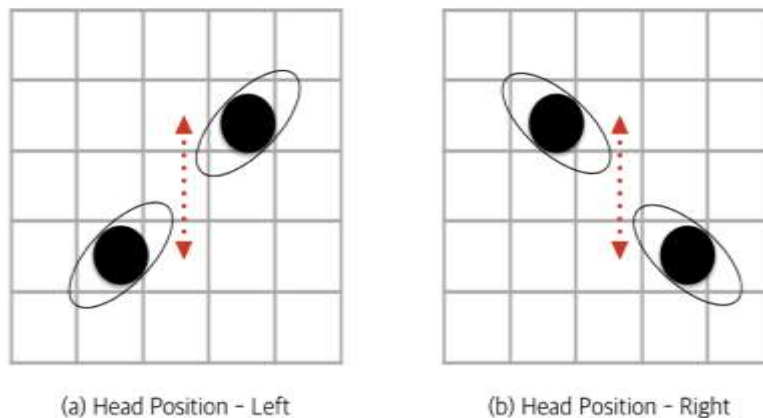
Figure 7. The Examples of Eye Prediction Considering Head Angle

If the head position is known, we should apply a different kind of sight prediction methods to that position. For this, the proposed algorithm improves the accuracy of gaze prediction by using the method of dividing the detected eye into four regions as shown in the Figure 8.
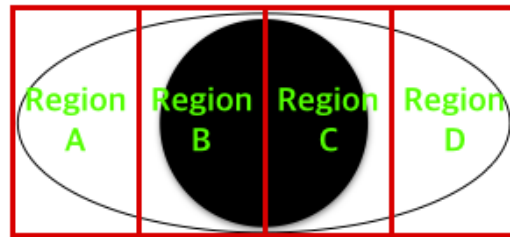
Figure 8. Divided Eye Regions

For example, when driver turns his/her gaze to the left, we can predict gaze because the center of both eyes moves similarly to the left. That is, the proposed method utilizes the characteristic that the position of the eye moves similarly when the driver recognizes a specific object. The proposed monitoring algorithm improves driver's safety by using head position prediction based on eye position and gaze recognition function using eye center position.

3.2. Drowsiness Detection

Fatigue reduces driver perception and decision-making capability to control the car. Research studies that usually after 2-3 hours of continuous driving, driver is fatigued and the driving performance deteriorated [13, 14]. To detect driver's drowsiness, the proposed monitoring algorithm uses the driver's fatigue-based approach. To detect drowsiness, the proposed algorithm uses a method based on the eye closeness. To perform the proposed method, the process of setting a threshold value must be preceded. Since the degree of blinking of eyes is different for each driver, the proposed algorithm conducts a training process to set a threshold value. When the threshold is set and the eyes are closed, the timer is started. After the threshold is set, the timer is started when the eyes are closed and alerts the driver when the time the eyes are closed exceeds the threshold. That is, when the time of the closeness of the eyes exceeds the threshold value, it is recognized as drowsiness and the warning is notified to the driver.

4. Performance Analysis

To evaluate the performance of the proposed algorithm, experiments were conducted on face and eye recognition, and eye gaze prediction in various situations. In addition, we conducted a performance experiment on whether the proposed algorithm works efficiently through scenarios setting that the driver moves while moving the field of view. All the images needed to carry out the experiment were used in the image database of PubFig (Public Figures Face Database) [15].

4.1. The Position of the Head Facing Forward

To In the proposed algorithm, face and eye recognition should be preceded to predict gaze of driver. Figure 9 shows the result of detecting the center of the face and eyes using the image of the database. In the proposed method, when the face is facing to the front, it shows good results in detecting the center of the eyes.

Figure 9. The results of Detecting the Center of the Face and Eyes

On the other hand, Figure 10 shows the image when the detection fails. Figure 10 also shows the image when eye and center of eye detection fails. When the face is detected but the eye is not found in the image, the center of both eyes of the driver cannot be found, so the result of the driver's gaze cannot be predicted.

**Figure 10. The Examples That Failed Eye Gaze Detection**

4.2. The Various Position of the Head

Figure 11 shows the driver's gaze predictions for various face positions. In the proposed method, if the face is facing the front, the center of the eye can be searched with a high probability. However, as shown in Figure 11, when the angle or direction of the face is varied, the driver's gaze may not be detected. To solve this problem, existing methods such as making a 3D model based on the feature of face exist. However, there are many computation overheads to create a 3D model, which can cause problems that cannot be immediately addressed. On the other hand, because the proposed method uses heuristics to make a quick decision, it is possible to cope with the situation quickly, though it is less accurate than the method of generating a 3D model.

(a) Head Position – Center   (a) Head Position – Left

Figure 11. The Example of a Variety of Face Angles
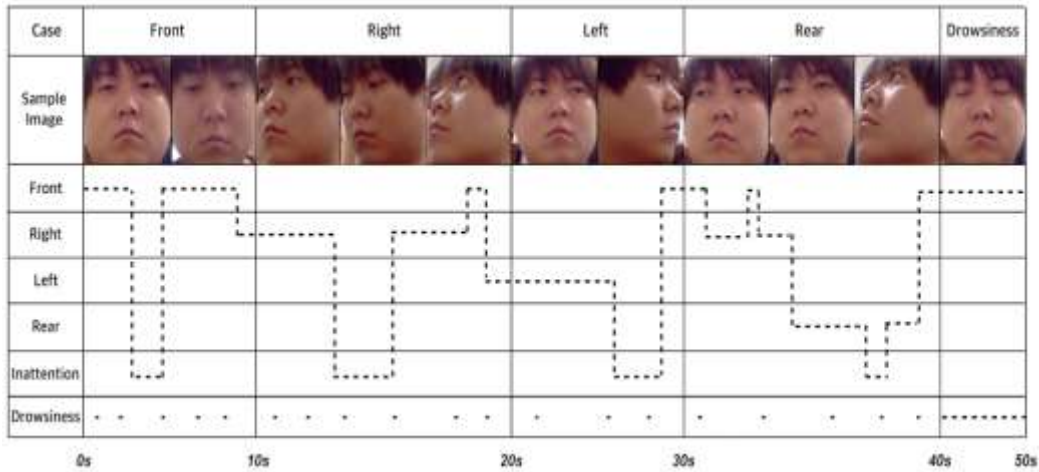
Figure 12. Graph of the Eye Gaze Prediction in Scenario Situation

4.3. Scenario-based Performance Analysis

To analyze the performance of the proposed algorithm in the driving situation, we set the scenarios gazing at the room mirror, side mirror, the front, inattention and drowsiness. We used 30 frame images per second for performance analysis and experimented with the front, right side mirrors, left side mirrors, room mirrors and drowsiness in an order.

Figure 12 shows the flow of the scenario images. First, the eye gaze prediction experiments were performed in scenarios with each image. When a driver is gazing at the front, the center of the eye is located close to the eye center regardless of the angle of the face, which shows a high prediction result. When a driver is viewing both side mirrors in a state where the face is facing the front, the centers of both eyes move to the same direction, so that prediction results are good. However, if the angle of the face changes greatly, the accuracy of the vision cannot be predicted by the center of the eye. When the face is facing the front and gazes at the room mirror, the eye is accurately predicted by eye-center information. However, even if the driver is looking at the room mirror, there is a case where the center of the eye is located at the center depending on the face angle. In the case of drowsiness, the system operates by recording the time when the eyes are closed. Therefore, when the driver flickers the eyes occurs in each scenario, the proposed method may misunderstand it as a case of drowsiness. In addition, the proposed method performs the eye gaze prediction operation when the eyes are closed by using the previous gaze prediction results. Therefore, we show the results of previous gaze prediction with rear-mirror as well as drowsiness predictions as shown in Figure 12. Also, since the information of the eyes' center changes at each scenario boundary, it shows incorrect prediction results.

Table 1. Accuracy of Gaze Prediction in Scenario Situation

View	Accuracy
Front	86.33%
Left Side Mirror	78.66%
Right Side Mirror	80.66%
Room Mirror	47.66%

That is, when the eye gaze is predicted only by the center position of the eye, only the instantaneous state is determined, which results in the relatively low accuracy. To solve this problem, we used a method which estimates the gaze by estimating the interval with higher hit rate for a certain period of time. We set the interval time to 400ms to prevent the accident by recognizing the driver's gaze. [16]. For example, as shown is Figure 12 in Section 1, it is possible to predict slightly more precisely because there is a greater probability of gazing at the front, even if there are some wrong results of predicting the right in the situation of gazing at the front. That is, since the gaze prediction is performed by collecting information of a specific period rather than each situation, a higher prediction rate can be confirmed.

Table 2. Accuracy of Gaze Prediction using Higher Hit Estimation in Scenario Situation

View	Accuracy
Front	90.66%
Left Side Mirror	84.33%
Right Side Mirror	85.67%
Room Mirror	57.66%

Table 2 shows improved results than Table 1. In other words, since the proposed system detects the non-observing situation with a high probability, so it is possible to cope with the risk situation effectively. Figure 12 also shows the inattention results. If the driver's gaze is pointing downward, the problem may arise because it is perceived as looking at the inattention region. However, since the proposed method detects the inattention region, the driver can receive a warning to cope with the danger. Also, when the driver is looking at the side mirror, there is a case where the head is judged according by the angle and the position of the head, but it is possible to cope with a dangerous situation because the eye is predicted through the section. In other words, the proposed system detects the non-observing situation with a high probability, so it is possible to cope with the risk situation effectively.

5. Conclusion

In this paper, we presented a driver monitoring algorithm to improve driver's safety by recognizing gaze and drowiness. Our approach begins by detecting the driver's face and eyes. Based on the detective eyes, we generate the area of the eye and then use the isophote to find the center of the eye and predict the driver's gaze. Also, the sleeping prediction function is performed by recording the time for how long the eyes are closed. To evaluate the performance of the proposed monitoring algorithm, the driver's eye gaze prediction were performed by varying the head and eye positions. In addition, since the

gaze of the eye is estimated by using the entire gaze information of each section, it has relatively high certainty. Experimental results show that our proposed method is able to predict driver's gaze and drowsiness with high accuracy in various environments. Although the accuracy of the prediction does not seem high in each situation, it is not a big problem when is is vieved systematically. The proposed algorithm guarantees the driver's safety because it can quickly recognize the driver's non - residential and dangerous situations then it can alert the driver. In other words, the method proposed in this paper has a feature of not requiring expensive equipment and fast execution speed because it uses eye center and interval information to predict eye line.

Acknowledgments

References

[1] Q. Wang, J .Yang, M. Ren and Y. Zheng, "Driver fatigue detection: a survey", The Sixth World Congress on Intelligent Control and Automation, **(2006)**, pp. 8587-8591.

[2] J. May and C. Baldwin, "Driver fatigue: The importance of identifying causal factors of fatigue when considering detection and countermeasure technologies", Transportation Research Part F: Traffic Psychology and Behaviour, vol. 12, no.3, **(2009)**, pp. 218-224.

[3] L. Hartley, T. Horberry, N. Mabbott and G, Krueger, "Review of fatigue detection and prediction technologies", National Road Transport Commission, **(2000)**.

[4] E. Ohn-Bar and M. Trivedi, "Hand gesture recognition in real time for automotive interfaces: A multimodal vision-based approach and evaluations", IEEE Transaction on Intelligent Transportation System, vol. 15, no. 6, **(2014)**, pp. 2368-2377.

[5] R. Mbouna, S. Kong and M.-G. Chun, "Visual analysis of eye state and head pose for driver alertness monitoring", IEEE Transaction on Intelligent Transportation System, vol. 14, no. 3, **(2013)**, pp. 1462-1469.

[6] Q. Ji, Z. Zhu and P. Lan, "Real-time nonintrusive monitoring and pre- diction of driver fatigue", IEEE Transaction on Vehicle Technology, vol. 54, no. 4, **(2004)**, pp. 1052-1068.

[7] M. Hamouz, J. Kittlerand, J. K. Kamarainen, P. Paalanen, H. Kalviainen and J. Matas, "Feature-based affine-invariant localization of faces", IEEE Transaction on Pattern Analysis and Machine Intelligence, vol. 27, no. 9, **(2005)**, pp. 1490-1495.

[8] A. Georghiades, P. Belhumeur and D. Kriegman, "From few to many: Illumination cone models for face recognition under variable lighting and pose", IEEE Transaction on Pattern Analysis and Machine Intelligence, vol. 23, no. 6, **(2001)**, pp. 643-660.

[9] J. Geusebroek, A. Smeulders and J. van de Weijer, "Fast anisotropic gauss filtering", IEEE Transaction on Image Processing, vol. 12, no. 8, **(2002)**. pp. 938-943.

[10] P. Campadelli, R. Lanzarotti, and G. Lipori, "Precise eye localization through a general-to-specific model definition", Proceedings of the British Machine Vision Conference, Edinburgh, UK, **(2006)**.

[11] S. Asteriadis, N. Nikolaidis, A. Hajdu and I. Pitas, "An eye detection algorithm using pixel to edge information", In International Symposium on Control, Communication and Signal Processing, **(2006)**.

[12] M. Asadifard and J. Shanbezadeh, "Automatic adaptive center of pupil detection using face detection and cdf analysis", Proc. of the international multiconference of engineers and computer scientists, **(2010)**.

[13] G. Matthews and D. Paula A, "Task-induced fatigue states and simulated driving performance", The Quarterly Journal of Experimental Psychology: Section A, vol. 55, no. 2, **(2002)**, pp. 659-686.

[14] J. F. May, E. P. Bryan and J. W. Catesby, "The deterioration of driving performance over time in drivers with untreated sleep apnea", Accident Analysis & Prevention, vol. 89, **(2016)**, pp. 95-102.

[15] N. Kumar, A. C. Berg, P. N. Belhumeur and S. K. Nayar, "Attribute and Simile Classifiers for Face Verification", IEEE International Conference on Computer Vision, **(2009)**.

[16] M. S. Horswill, S. Helman, P. Ardiles and J. P. Wann, "Motorcycle accident risk could be inflated by a time to arrival illusion", Optometry & Vision Science, vol. 82, no.8, **(2015)**, pp. 740-746.

Authors

Juyong Lee, he received his B.S., M.S degrees in information telecommunication engineering from Sangmyung University, Cheonan, Korea in 2014, 2016, respectively. He is currently Ph.D. candidate in electronic information system engineering at Sangmyung University, Cheonan, Korea. His research interests include mobile edge computing, content centric networking, content-based networking/routing, and software defined networking.

Seungdo Jeong, he received the M.S. and Ph.D. degrees in Electrical and Computer Engineering from Hanyang University, Seoul, Korea at 2001 and 2007, respectively. From 2009 to 2014, he served as an assistant professor of the Department of Information and Communication Engineering at Hanyang Cyber University, Seoul, Korea. He is currently an assistant professor at the Department of Smart Information and Telecommunication Engineering, Sangmyung University. His research interests include multimedia information retrieval, computer vision, multimedia contents processing, and tensor analysis.

Jihoon Lee, he received his B.S., M.S, and Ph.D. degrees in Electronics engineering from Korea University, Seoul, Korea in 1996, 1998, and 2001, respectively. From 2002 to 2011, he had worked at Samsung Electronics as a senior research member. He is currently an assistant professor at the Department of Smart Information and Telecommunication engineering, Sangmyung University. His research interests include information centric networking, context-aware networking, mobile cloud, and network security.