A New System for Driver Drowsiness and Distraction Detection

Mehrdad Sabet*, Reza A. Zoroofi*, Khosro Sadeghniiat-Haghighi†, Maryam Sabbaghian*
*School of Electrical and Computer Engineering, University of Tehran, Tehran, Iran
m.sabet@ece.ut.ac.ir, zoroofi@ut.ac.ir, mariam@seas.harvard.edu
†Tehran University of Medical Sciences, Center For Research on Occupational Medicine
sadeghniiat@sina.tums.ac.ir

***Abstract*—Drowsiness especially in long distance journeys is a key factor in traffic accidents. In this paper a new module for automatic driver drowsiness detection based on visual information and Artificial Intelligence is presented. The aim of this system is to locate, track and analyze both the driver's face and eyes to compute a drowsiness index to prevent accidents. Both face and eye detection is performed by Haar-like features and AdaBoost classifiers. In order to achieve better accuracy in face tracking, we propose a new method which is combination of detection and object tracking. Proposed face tracking method, also has capability to self correction. After eye region is found, Local Binary Pattern (LBP) is employed to extract eye characteristics. Using these features, an SVM classifier was trained to perform eye state analysis. To evaluate the effectiveness of proposed method, a drowsy person was pictured, while his EEG signals were taken. In this video we were able to track face by an accuracy of 100% and detecting eye blink by accuracy of 98.4%. Also we can calculate face orientation and tilt using eye position which is valuable knowledge about driver concentration. Finally, we can make a decision about drowsiness and distraction of the driver. Experimental results show high accuracy in each section which makes this system reliable for driver drowsiness detection.**

***Keywords*-face detection, face tracking, eye state analysis, AdaBoost and LBP.**

I. INTRODUCTION

Sleep Related Vehicle Accidents has the biggest part in traffic accidents. 30% of all injuries and fatalities have been caused by drowsiness globally [1]. Annually 20% of all accidents are caused by fatigue and distractions [2]. In the USA 100,000 vehicle accidents had been occurred every year which lead to $12 billion [3]. 20% of road accidents in the UK are because of this problem [4].

According to this statistic there is huge need to build a system for monitoring the drivers and measuring their level of attention. These kinds of systems are categorized in Advance Driver Assistance System (ADAS). The goal of such system is to increase vehicle safety with employing new technologies and also decreasing dangerous situation that might happen for passengers and vehicle.

There are three ways to deal with this problem. First way to do this is making vehicle more intelligent. This work can be done by using multiple sensors and subsystems such as wheel angle measuring sensors and laser scanners for measuring distance between vehicle and obstacle and pedestrian detector which is based on image processing technique [5], [6], [7]. However this method is not intrusive but they depend on driver skills and road or vehicle type which cause to degrade the performance of this kind of ADAS.

Another way is based on vital signals such as EEG and ECG [8], [9]. This method is very accurate and also is the best reference for making decision about driver drowsiness and fatigue. However multiple sensors must attach to driver which make the driver inconvenient.

In the third method, detecting driver drowsiness and distraction are performed using facial image processing [10], [11], [12]. Fortunately human face produces distinctive characteristics in different states. In drowsiness situation many visual cues can be detected in human face, such as eye blinking, yawning and head movement. According to this characteristics, analysis of facial image is the most feasible and appropriate way to do this. The goal of proposed system is to monitor and alarm the driver to prevent any accident that might be caused by drowsiness and distraction.

The rest of this paper is organized as follows. A general framework of the proposed system is presented in Section 2. In Section 3 we introduce our face tracking method and other subsystems which have been used in it. Sections 4 and 5 discuss eye detection and state analysis. Section 6 explains the way we measure distraction. Section 7 shows experimental results and discussion are given about proposed method. Section 8 presents conclusion.

II. WARNING SYSTEM DESIGN

Every drowsiness detection system has several main modules such as face and eye detection, tracking, etc. In [13] a new algorithm for detecting iris, pupil and lips based on color information is proposed. This algorithm could analyze eye state and make decision about level of fatigue. D'Orazio et al. [14] searches eye region in whole image with this assumption that sclera is always brighter than iris. Then employed Hough transform to gathering some candidate region. Finally using Neural Network, they classify eyes and non-eyes regions.

The main problems of this system are that it can only work when eyes are visible and it is not robust to illumination changes. Horng et al. [15] used eye color model in HSI space

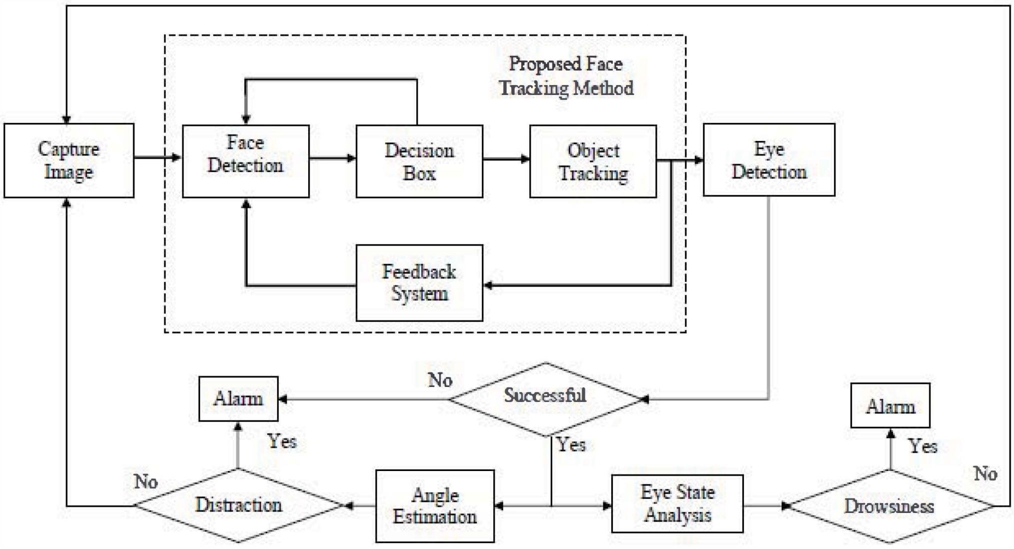


Fig. 1. Block diagram of proposed system for drowsiness and distraction detection

for detecting eyes. Driver is detected as drowsiness if the eyes are closed for five consecutive frames.

According to what has been presented, we propose a new structure for driver drowsiness detection system. Block diagram of proposed system illustrated in Fig. 1. It would be interesting to note that proposed system operates using gray level images. We will describe each block of proposed system in following sections.

III. Face Tracking

Face tracking system must be robust to head movement, rotation, pose variation and illumination changes. To achieve this goal we propose a method to use face detection and object tracking systems simultaneously. This combination gives us the opportunity to utilize advantages of two programs together.

A. Face Detection

For the face detection process, proposed system uses Viola and Jones(VJ) method [16] which based on machine learning approach for visual object detection. This method takes advantage of three different features which are integral image, AdaBoost technique and the cascade classifier. With these features it can achieve near real-time speed and correct detection rate as high as 90%. One of the important advantages of this method is robustness to light condition changes [10].

It cannot work properly when face turns and exits from the frontal view to the camera. This method can detect faces that are tilted up to ± 15 degrees in plane and ± 45 degrees out of plane [16]. Fig. 2 shows true face position and VJ estimated position for 100 frames of Hide [17] video. True position was determined manually by hand. As we can see VJ method is very accurate but it can just detect the face in a few frames.

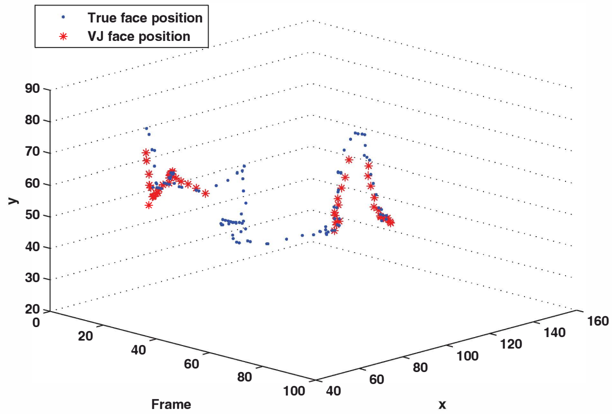


Fig. 2. 100 frame sequence where VJ method dose not find the face in all frames.

B. Object Tracking

For object tracking we use Ning et al. method [18]. This method is based on mean shift. In addition to use color histogram, this method utilized LBP [19] pattern and propose joint color-texture histogram method which represent target very distinctive and effective.

The LBP operator labeling pixels in image [19]. This process was done by thresholding center pixel with its neighborhoods. LBP operator can be defining as follows:

$$LBP_{P,R}(x_c, y_c) = \sum_{p=0}^{P-1} s(g_p - g_c)2^p \tag{1}$$

where g_c is gray level of center pixel and g_p is gray level of P pixels around center with radius R and the function s(x) is defined as follows:

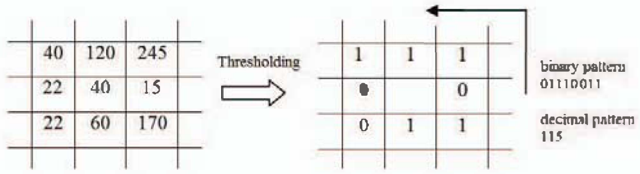


Fig. 3. An example of $LBP_{8,1}$ texture model.

$$s(x) = \begin{cases} 1, & x \geq 0 \\ 0, & x < 0 \end{cases} \tag{2}$$

Fig. 3 shows an example of $LBP_{8,1}$. One drawback to this object tracking method in face tracking process is the size of surrounding rectangle is not adaptive. Therefore if the distance between head and camera reduces, face appears much bigger than the rectangle size, which is determined at first frame. So this size of rectangle could not cover whole face. Another problem of this method is that it could not work properly in gray level images whereas mean shift method depends on color histogram [20]. This causes mean shift algorithm achieves poor representation of the target and leads to divergence and inaccurate target tracking. This problem is more obvious when face came back to frontal view from side view, because in this situation, face posture changes very fast.

C. Feedback System

Consider the object tracking module is the main part of proposed face tracking subsystem, as realized it could diverge or loose target. Thus to overcome this problem, we proposed a feedback system to detect and measure any divergence from true position at the output of object tracking system and correct this error. Therefore if true representation of face was available, this image could be stored as a reference and compared to output of object tracking system to measure accuracy of rectangle position that contains face image.

As illustrated in Fig. 2, if face detection system could find face position, this position is very accurate and very close to the true one. Because cascade structure that is used in VJ method makes this face detection system to have very low false positive rate [16]. So, this advantage gives us confidence that the rectangle which determined by this method, contains the face and is not a false object. For making a reference image, we use LBP operator. As mentioned earlier one of the advantages of LBP textures is that they don't depend on color or illumination of picture.

Once LBP image of reference and output have been calculated, two dimensional correlation could be used to measure the differences between these two images. The bigger the number, the smaller the divergence. Therefore when divergence is bigger than the threshold, feedback system pass control to detection system to find true face position and re-initialize object tracking system. Smaller threshold leads to better face localization in each frame but also causes more switches between detection and tracking system and more computational time. So, determining optimum threshold is necessary with constraint on computational power. We use 0.08 for this value which came from experiment and lead to good performance.

D. Decision Rules

Because the detection and tracking systems are imperfect, there is strong need to enforce several simple rules. At the starting point of proposed system and when the first frame loaded to system, decision box always pass control to detection system until the face is found. After passing this level, if detection system couldn't find the face, decision box passes control to the next section and then initializes tracking system with the last face position. If the assumption of soft head movement is valid, in each frame that face detection system, outputs the position whereas is far from the last position, we assume face detection system make a mistake and that output is false alarm. So decision box passes control to the next section and uses last position of face.

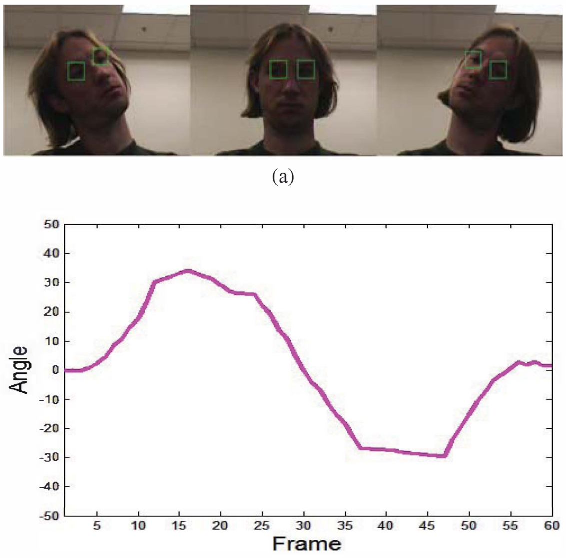
IV. Eye Detection

Locating the position of eye is difficult task due to many factors such as lighting condition, expression, facial shadowing, etc. After the position of face has been obtained, locating the eye can be done with better accuracy. As mentioned about the benefit and reliability of Adaboost [16] we utilize this method to detecting eyes. As what was said before this method is not robust to pose variation. However we use this disability as an advantage in distraction detection system. If eye could not be detected we can assume that the driver don't look at forward. So this situation can be categorized in distraction state and must alarms the driver.

V. Eye State Analysis

To detecting drowsiness it is necessary to know eye state that is open or close. Eye state classification is difficult due to some parameters. According to the efficiency and low computational time of Support Vector Machine (SVM), proposed system use this method to analysis eye state. After eye has been detected, LBP operator has been used to extract eye characteristics. Independency of LBP from illumination lead to more precision in eye state analysis. In the next stage, feature vector obtained from sub-window that introduced by Chen and Kubo [21]. Feature vector d compute from LBP image G(x,y) with the sub-window size 5 by 6 as follow:

$$d_i = 1/40 \sum_{y=1:5} \sum_{x=1:6} G(x, y) \quad i = 1, ..., 40. \tag{3}$$



(a)

(b)

Fig. 4. An example of face orientation measuring: (a) frame 15, 30, 45 of James video [17] are displayed. (b) corresponding angle for 60 frame.

VI. Distraction

It is estimated that 20% of traffic accidents are caused by driver distraction [2]. To detect this characteristic the drivers face should be studied because the pose of the face contains information about ones attention, gaze, and level of fatigue [10]. To verify driver distraction the following procedure has been implemented. Face orientation is estimated using the eye position, with the following equation:

$$\theta = \tan^{-1}(\Delta x / \Delta y) \tag{4}$$

where $\Delta x = x2 - x1, \Delta y = y2 - y1$, (x1, y1), and (x2, y2) correspond to the left and right eye center positions, respectively. Fig. 4 shows an example of face orientation and its corresponding angle for 60 consecutive frames. Decision boundary could be define as follows:

$$FaceOrientation = \begin{cases} Right, & \theta > 8^\circ \\ Front, & |\theta| \leq 8^\circ \\ Left, & \theta < -8^\circ \end{cases} \tag{5}$$

VII. Experimental Result

In this section we demonstrate the robustness and efficiency of proposed system using multiple video and quantitative results. Proposed system implemented in MATLAB version 7.8.0.347 and run on a PC with Intel Core 2 Duo 2.8 GHz CPU and 4 GB RAM. All videos frames are scaled to 350×240 pixels for faster execution. Four video sequences were employed to assess the performance of proposed method. Hide and James video obtained from Honda/UCSD Video Database [17], third video is well known Dudek sequence [22] and the fourth video is Smiley [23].

TABLE I
Quantitative Comparisons in Terms of RMSE and Speed

Method	Sequence	Total Frames	RMSE	Speed(fps)
Proposed Method	Smiley	456	5.15	99.7
Ning et al. [18]			10.92	110
IVT [24]			41.77	1.87
Proposed Method	Dudek	573	6.78	88.23
Ning et al. [18]			24.6	110
IVT [24]			5.32	1.1
Proposed Method	Hide	459	6.95	92.3
Ning et al. [18]			8.73	110
Proposed Method	James	220	6.42	84.5
Ning et al. [18]			7.4	110

TABLE II
Results of Eye Blink Detection

Sequence	Eye Blink	Correct Rate
ZJU1	24	100
ZJU2	20	100
ZJU3	24	96
ZJU4	23	95.65
Talking Face	13	100
Average	—	98.33

For measuring accuracy of proposed face tracking method we directly compare our tracking results to IVT [24] and Ning et al. [18] methods by employing the average root mean square error (rmse) between the ground-truth and estimated positions to assess tracking performance. Table I presents these values and also shows computational speed for each method in each video. Proposed face tracking method outperforms Ning et al. method [18] in all sequences and IVT [24] in Smiley video and also much faster than IVT method.

In eye state analysis the main objective of the SVM training is to obtain the best parameters and the best kernel. After several SVM training experiments, it was decided to use the RBF kernel with the Sigma value equal to 30. Two datasets were used to test this method. The first one is the ZJU Eyeblink [25], which comprises 80 short videos having 11,800 frames (350×240) captured from 20 Chinese people at a capturing speed of 30 fps. They include 255 blinks totaly. The second database is the Talking Face Video database, which contains one video having 5000 frames captured from one male engaged in conversation. It includes 61 blinks totaly. This corresponds to approximately 200 seconds of recording; the consequent capturing speed can be estimated to be approximately 25 fps. Table II shows the results of eye blink detection (every peroson has 4 video in ZJU dataset, we showing the results of these videos in one row).

To evaluate the effectiveness of proposed system, a drowsy person was pictured. The video length is approximately 45

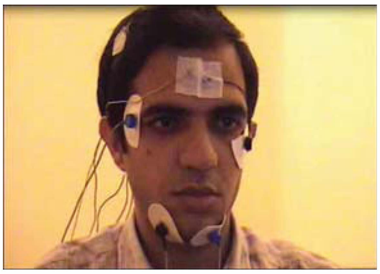


Fig. 5. An example of drowsy video.

minutes and was taken around the midday. Also the EEG signals were recorded along with imaging. As we mentioned before EEG signals are very good reference to detecting drowsiness. Fig. 5 shows one frame of drowsy video. One of these signals is taken from the eye which shows eye closure state and we use this one as a reference for eye blink detection. Finally by employing proposed system we could track face by an accuracy of 100% and detecting eye blink by accuracy of 98.4%.

VIII. Conclusion

In this paper a new system has been presented to monitor and detect driver drowsiness and distraction. This system uses advance technologies based on computer vision and artificial intelligence. For face tracking we employed feedback system using LBP, to prevent divergence and losing the target. Because of feedback system, detection and tracking modules can cooperate with each other. Eye state analysis performed by SVM with features which extracted using LBP operator.

The proposed algorithm for face tracking and eye state analysis is shown to be robust and accurate for varying light, background changes and facial orientation. The system is also observed to provide agreeable results.

References

[1] L. Fletcher, L. Petersson, and A. Zelinsky, "Driver assistance systems based on vision in and out of vehicles," *in Proceedings of The IEEE Symposium on Intelligent Vehicles*, p. 322327, 2003.

[2] T. Brand, R. Stemmer, B. Mertsching, and A. Rakotonirainy, "Affordable visual driver monitoring system for fatigue and monotony," *in Proceedings of the IEEE International Conference on Systems, Man and Cybernetics*, vol. 7, pp. 6451–6456, October 2004.

[3] NHTSA, "Evaluation of techniques for ocular measurement as an index of fatigue and the basis for alertness management," National Highway Traffic Safety Administration, Washington DC, USA, Technical Report DOT HS 808762, 1998.

[4] L. Hagenmeyer, "Development of a multimodal, universal human-machine interface for hypovigilance-management systems," PhD Thesis, University of Stuttgart, Stuttgart, Germany, 2007.

[5] M. Sarshar and M. Rezaei, "A novel system for advanced driver assistance systems," *In Proceedings of the IEEE International Conference on Systems*, pp. 529–534, 2010.

[6] J. M. Armingola, A. de la Escalera, C. Hilario, J. M. Collado, J. P. Carrasco, M. J. Flores, J. M. Pastor, and F. J. Rodriguez, "Ivvi: Intelligent vehicle based on visual information," *Robotics and Autonomous Systems*, vol. 55, p. 904916, September 2007.

[7] A. Sathyanarayana, P. Boyraz, and J. H. Hansen, "Information fusion for robust 'context and driver aware' active vehicle safety systems," *Information Fusion*, vol. 10, no. 16, June 2010.

[8] G. Yang, Y. Lin, and P. Bhattacharya, "A driver fatigue recognition model based on information fusion and dynamic bayesian network," *Information Sciences*, vol. 180, pp. 1942–1954, January 2010.

[9] C. T. Lin, C. J. Chang, B. S. Lin, S. H. Hung, C. F. Chao, and I. J. Wang, "A real-time wireless braincomputer interface system for drowsiness detection," *IEEE Transactions on Biomedical Circuits and Systems*, vol. 4, no. 4, pp. 214–222, August 2010.

[10] M. J. Flores, J. M. Armingol, and A. de la Escalera, "Driver drowsiness warning system using visual information for both diurnal and nocturnal illumination conditions," *EURASIP Journal on Advances in Signal Processing*, June 2010.

[11] Q. Ji and X. Yang, "Real-time eye, gaze, and face pose tracking for monitoring driver vigilance," *Real-Time Imaging*, vol. 8, no. 5, pp. 357–377, February 2002.

[12] X. Fan, Y. Sun, B. Yin, and X. Guo, "Gabor-based dynamic representation for human fatigue monitoring in facial image sequences," *Pattern Recognition Letters*, vol. 31, p. 234243, 2010.

[13] P. R. Tabrizi and R. A. Zoroofi, "Drowsiness detection based on brightness and numeral features of eye image," *Fifth International Conference on Intelligent Information Hiding and Multimedia Signal Processing*, pp. 1310–1313, 2009.

[14] T. D'Orazio, M. Leo, and A. Distante, "Eye detection in faces images for a driver vigilante system," *in Proceedings of the IEEE Intelligent Vehicles Symposium*, pp. 14–17, June 2004.

[15] W. Horng, C. Chen, Y. Chang, and C. H. Fan, "Driver fatigue detection based on eye tracking and dynamic template matching," *in Proceedings of the IEEE International Conference on Networking, Sensing and Control*, pp. 7–12, March 2004.

[16] P. Viola and M. J. Jones, "Robust real-time face detection," *International of Computer Vision*, vol. 57, no. 2, pp. 137–154, 2004.

[17] Honda/UCSD Video Database. [Online]. Available: http://vision.ucsd.edu/ leekc/HondaUCSDVideoDatabase/HondaUCSD.html

[18] J. Ning, L. Zhang, D. Zhang, and C. Wu, "Robust object tracking using joint color-texture histogram," *International Journal of Pattern Recognition and Artificial Intelligence*, vol. 23, no. 7, pp. 1245–1263, 2009.

[19] T. Ojala, M. Pietikainen, and T. Maenpaa, "Multiresolution gray-scale and rotation invariant texture classification with local binary patterns," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, no. 7, pp. 971–986, July 2002.

[20] S. Avidan, "Ensemble tracking," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 29, no. 2, pp. 261–271, February 2007.

[21] Y. W. Chen and K. Kubo, "A robust eye detection and tracking technique using gabor filters," *in Proceedings of the 3rd International Conference on Intelligent Information Hiding and Multimedia Signal Processing*, pp. 109–112, 2007.

[22] D. Ross. [Online]. Available: http://www.cs.toronto.edu/ dross/ivt/

[23] Colorado State University. [Online]. Available: http://www.cs.colostate.edu/ vision/SMC-FaceTracking/

[24] D. A. Ross, J. Lim, R.-S. Lin, and M.-H. Yang, "Incremental learning for robust visual tracking," *Int. J. Comput. Vis.*, vol. 77, no. 1-3, pp. 125–141, May 2008.

[25] Z. Eyeblink. [Online]. Available: http://www.cs.zju.edu.cn/?gpan/