Facial Features Monitoring for Real Time Drowsiness Detection

Conference Paper · November 2016

DOI: 10.1109/INNOVATIONS.2016.7880030

Manu Bn
Independent Researcher/Computer Vision

Facial Features Monitoring for Real Time Drowsiness Detection

Manu B.N
Independent Researcher
Bangalore, India
manubn88@gmail.com

***Abstract*—This paper describes an efficient method for drowsiness detection by three well defined phases. These three phases are facial features detection using Viola Jones, the eye tracking and yawning detection. Once the face is detected, the system is made illumination invariant by segmenting the skin part alone and considering only the chromatic components to reject most of the non face image backgrounds based on skin color. The tracking of eyes and yawning detection are done by correlation coefficient template matching. The feature vectors from each of the above phases are concatenated and a binary linear support vector machine classifier is used to classify the consecutive frames into fatigue and nonfatigue states and sound an alarm for the former, if it is above the threshold time. Extensive real time experiments prove that the proposed method is highly efficient in finding the drowsiness and alerting the driver.**

***Keywords*-drowsiness detection, Viola Jones, correlation coefficient template matching, k means, Sobel edge, SVM.**

I. Introduction

Every year thousands of people in India lose their lives due to traffic accidents. The role of human factor plays a key role in the accidents. In general, the driver fatigue alone accounts for around 25 percent of the road accidents and up to 60 percent of road accidents result in death or serious injury [1]. A main cause of fatigue is sleeplessness or insomnia. So a drivers' drowsiness state is a major factor in severe road accidents that claims thousands of lives every year.

In the recent years, use of intelligent algorithms in cars has developed considerably. These systems use WSNs to monitor and transmit the condition of the car and the driver. Smart cars that use software techniques to control engine speed, steering, transmission, brake etc. has improved the quality of driving significantly. Ad hoc networks were the first to develop the automatic navigation systems in cars. A noticeable weakness of these systems is that they don't respond in real time to the environmental changes. It is especially important in driving where time is a critical factor in driver's decision. On the other hand, another method to check the driver fatigue is monitoring the physical condition and facial expressions of the drivers, which wireless sensor networks are unable to process and transmit these information with adequate precision and a good recall. Hence it is very essential to develop an efficient drowsiness detection system.

A. Survey of Existing systems

In the past decade drowsiness has been investigated by various researchers. In [2] the authors have used the HOG SVM approach for drowsiness detection and comparisons are made with that of a human rater. The popularity of HOG SVM lies with object detection and not with the eye blink detection. Further, comparisons with a human rater in highly unreliable as human decisions are prone to more errors than an automated one. Li, G., Lee, B. L., & Chung [3] have proposed a smart watch and a headband containing sensors to identify drowsiness. But this is limited by the choice of the driver who might not like to wear a headband and might wear a different watch. In [4] the authors used a supervised learning method which needed a highly reliable ground truth. The authors in [5] used PERCLOS features for eye status detection. All the human-defined eye features used in their research were calculated from eyelid movements. This means the features are a subset of information provided by eyelid movements. Hence, extracting and utilizing only those artificial features can lead to loss of some meaningful information. Researches in [6] have used speaking and smile detection as emotion detection parameters. In the absence of which a possible drowsiness condition can occur. Use of such a system can only make the algorithm more complex to identify drowsiness. Kim et al [7] have used deep learning and facial expression recognition. But, this approach has the shortcoming of requirement of a huge amount of data to train a neural network to work with a high level of accuracy. Trivedi et al [8] have considered only the iris status for drowsiness detection. They don't consider yawning detection or head lowering which could lead to a system with a much better accuracy. Singh, R. K., et al. have integrated an electrode within a steering wheel to monitor the heart rate and sound alarm in case of fatigue. The approach is still at a very nascent phase as stated in their paper [9].

This paper presents a new method of drowsiness detection which monitors the status of eyes and mouth and does not require any sensors or wearable devices and works well under various illumination conditions.

The rest of the paper is organized as follows: Section II contains the algorithm for the proposed method. Section III gives details about real time implementation results and finally section IV presents the conclusions and future scope.

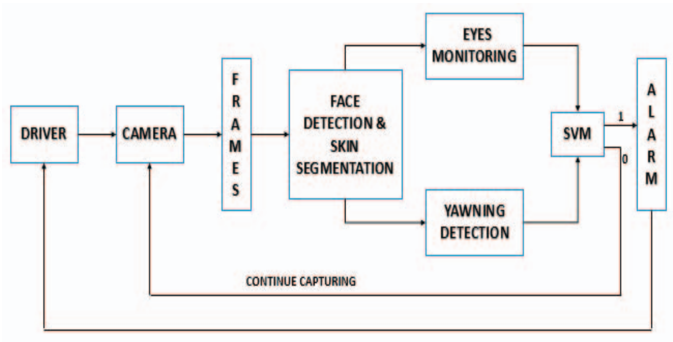


Fig.1 The proposed framework

II. DROWSINESS DETECTION FRAMEWORK

The proposed framework is illustrated in Fig.1. Initially a camera acquires the video of the driver. Then this video is broken down into frames. The below sections give a description of the methodology followed once the frames are acquired.

A. *Face Detection and Skin Segmentation*

The face detection is done by Viola Jones [10]. The main aim of face detection is to minimize the false detections in identifying facial expressions. The importance of this part is to accurately locate the position of the eyes and the mouth. Once the face is detected, skin segmentation [11] is performed by converting the image to YCbCr domain. The biggest advantage of converting the image to the YCbCr domain is that, the influence of luminosity can be eliminated by considering only the chromatic components. In the RGB domain, each component of the image i.e red, green and blue has a different brightness. But, in the YCbCr domain all the brightness information is given by the Y-component, since the Cb (blue) and Cr (red) components are completely independent of the luminosity. The domain conversions are used to segment the RGB image into Y, Cb, Cr components. Despite the fact that skin colors differ from person to person, and race to race, it was found that [11] the color remains distributed over a very tiny region in the chrominance plane. This method detects skin regions over the entire face image and rejects most of the non face image. Fig. 2 shows the detected face and the corresponding skin regions using the YCbCr segmentation mentioned in [11].

B. *Eye Tracking*

The most important factor which helps detect driver fatigue is the state of eyes, i.e. open or closed. In the state of drowsiness, the eyelid muscles subconsciously gravitate to accelerate the process of going to sleep.

Fig.2 Detected face and the corresponding skin regions

The position of the driver's eyes are determined by using Viola Jones [10]. Then the two eyes are separated using edge detection and in accordance with the symmetrical properties [6] of the eye, the center of the eye is determined [6]. Finally the pupil is identified. If eyes are open then it is treated as the normal state during which the alarm is not set off. If eyes are closed then it is treated as the fatigue state during which the alarm is set on.

Edge detection can be viewed as the process of localizing pixel intensity transitions. There exist several edge detection methods such as Sobel. Such methods have been proposed for detecting transitions in images. In the proposed work, Sobel which is an edge detection method is preferred over others methods like Canny. The Canny edge and other such algorithm solved these problems by first blurring the image slightly, followed by an algorithm that can effectively thin the edges to one-pixel. This constitutes a slower process, hence, Sobel operator is highly recommended over Canny in massive data communication especially in image transfer. The Sobel detector convolves the image with a filter that has small, separable, and integer valued function in both horizontal and vertical directions. Hence it is a relatively inexpensive method of computation. Secondly the gradient approximation produced by it is relatively crude, which is best suited for high frequency variations such as eye blinking during fatigue.

The status of the eyes is determined in every frame using the correlation coefficient template matching method. By considering the variations to the connected pixels and the similarity ratio to the eye pixels sufficiently, the specific region of the eyes is obtained. The Sobel edge detection method is also used to detect the eyes' precise and exact boundaries. The technique starts from left and right side, to find eyes, therefore can detect the eyes separately. The detected eyes are segmented from the image and are used to generate an eye template, by this means one can obtain a rather stable eye template for the status analyzing and also reduce the influence of light reflections. Fig.3 shows the eye template generation process. To distinguish the fatigue, the eyes' states should be recognized accurately. The two factors which can affect the size of the eyes are described here.

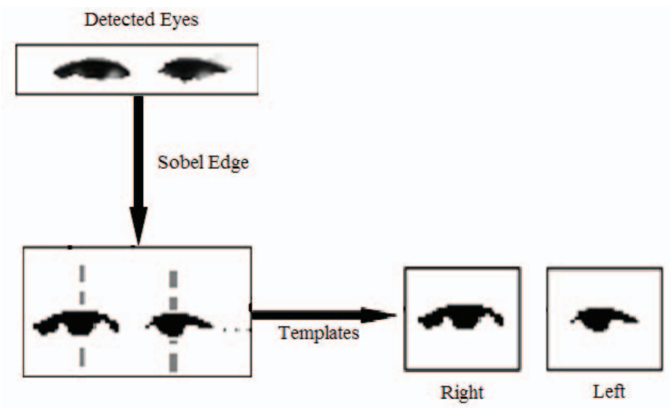


Fig.3 Eye template generation process [6]

TABLE I. Eye Template Information

	Eye Region	Area (no. of pixels)	Template	Avg. Height	Ratio
Full Open		201		7.62	2.87
Half Open		156		6.79	3.04
Closed		115		6.02	3.17

The human eyes vary in size. Secondly, the distance between driver and the camera varies in every frame. So it was decided to normalize the eye template to a pre fixed size of 12×30 and then extract the features. For each eye template, area, average height and width to height ratio are the crucial features to determine the eye's status which is shown in Table. I. The eye states can be divided into three categories: full open, half open and completely closed. From the Table I. one can see that eyes of different states have different characteristics. The different eye states of full open and half open are not well distinguished most of the time and can cause more false alarms and the variable movement of the drivers head can result in the driver's eyes tracking failure.

C. Yawning Detection

Another distinct sign of fatigue during driving, which is manifested in a person's face, is yawning that occurs due to body reflexes when a person is tired and is about to fall asleep. Once the regions of mouth area are found using Viola Jones, the mouth region alone is segmented by K means [5] clustering and tracked using correlation coefficient template matching [12]. K means partitions the objects into K no. of mutually exclusive clusters, so that objects in each cluster are closest to each other, and farthest from objects in other clusters. Each of the K clusters is characterized by its centroid, or center point. The function K-means performs K-Means clustering, by using an iterative algorithm that assigns objects to each clusters so that the sum of distances from each object to its corresponding cluster centroid, over all K clusters, is a minimum. The objective function aims to obtain the minimum distance between the classes, or basically between the image pixels. The minimum function is defined in (1) and (2).

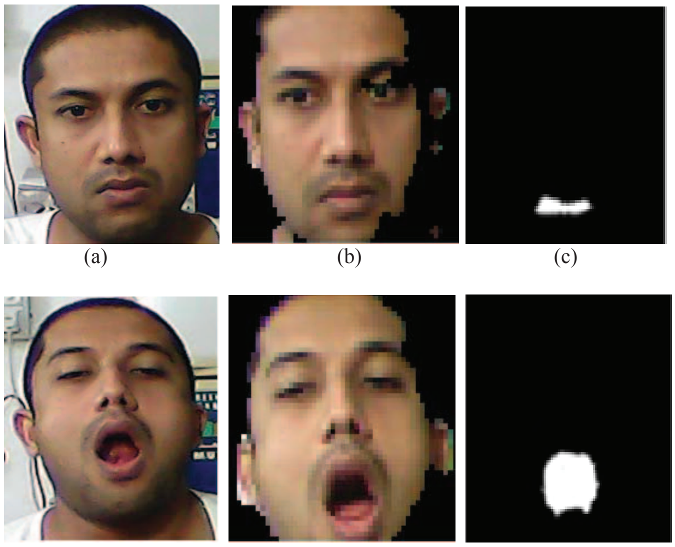
$$c_j = (x_i \,|\min(\,\|x_i - x_j\|\,)) \quad (1)$$

$$\operatorname{argmin} \Sigma\|c_j - x_j\| \quad (2)$$

In equations (1) and (2) x_i is i^{th} pixel, x_j is the center of class j and c_j are pixels belonging to class j. Classification of image pixels is based on the brightness intensity. Finally, a large part of the segmented area in the image demonstrates the exact position of the mouth and detects the yawning using the templates generated with K = 2 as illustrated in Fig.4. All the open and close templates are of fixed size 38x62.

D. Training

For training purpose 100 different templates are created for each type i.e. open and close for each of the eyes and the mouth and their correlation is measured w.r.t to the



(a) (b) (c)

(d) (e) (f)

Fig.4 Mouth template generation process

completely closed ones. The feature vector is given in (3) which is normalized over the total number of frames.

$$\text{Feature} = [LC\ RC\ MC\ TC]/TF \quad (3)$$

In (3) LC, RC, MC and TC stands for results of correlation coefficient template matching for left eye closure, right eye closure, mouth closure and the total closure for both the eyes and mouth. TF stands for total number of frames. For better real time efficiency partially closed templates i.e. 30% - 50% closed are also trained and assigned the label of close.

III. RESULT ANALYSIS

A binary SVM classifier with a linear kernel is used for classification. A 15fps 40M pixels camera is used for acquiring the video frames using Matlab 2013. If the frames continue to exhibit the drowsiness state for more than 3 seconds, an alarm is turned on.

The average accuracy (AAC), the detection rate (DR) and false alarm rate (FAR) in accordance with equations (4), (5) and (6) has been calculated. These three measures, which have been proposed for evaluating the accuracy, indicate the acceptable performance of the proposed algorithm in detecting the signs of fatigue in driver's face at the time of driving. Where N_{TP}, N_{TN}, N_{FP} and N_{FN} are number of frames in which fatigue has been detected, failed to detect fatigue, no sign of fatigue and the algorithm has not detected any by mistake and no sign of fatigue but the algorithm has detected some by mistake respectively. ROI is the objective area in the image which includes all the signs of fatigue in the face. To further validate the performance of the method various trials are conducted and tabulated in Table II. Trials 1 and 3 are conducted by dark complexion person in dim and bright light conditions. Trials 2 and 4 are conducted by fair complexion person under dim and bright light conditions. Table III gives comparisons with recent methods. Fig. 6 shows two frames with normal and drowsiness detected. The processing speed is 0.0395 seconds per frame on i3 processor with a 4GB RAM.

$$AAC = (N_{TP}+N_{TN}) / (N_{TP}+N_{TN}+N_{FP}+N_{FN}) \quad (4)$$

$$DR = \text{True Positives} / \text{Drowsy Frames} \quad (5)$$

$$FAR = \sum \text{Detected ROI} / \text{No. of Frames} \quad (6)$$

TABLE II. RESULT ANALYSIS

Trial No.	No. of frames	No. of frames with fatigue		No. of frames without fatigue		(AAC)	(DR)	(FAR)
		NTP	NFN	NTN	NFP			
1	90	31	2	46	6	90.58%	93.93%	11.53%
2	90	5	0	16	1	95.45%	100%	5.88%
3	90	11	2	39	2	92.59%	84.61%	4.87%
4	90	58	1	22	3	97.38%	98.57%	5.22%
Total	360	105	5	123	12	**94.58%**	**94.27%**	**6.87%**

TABLE III. RESULTS COMPARISON

Method	HOG+SVM [2]	Eyelid Movement [5]	Wearable EEG [3]	Proposed
Accuracy	91.6%	90.45%	92%	**94.58%**

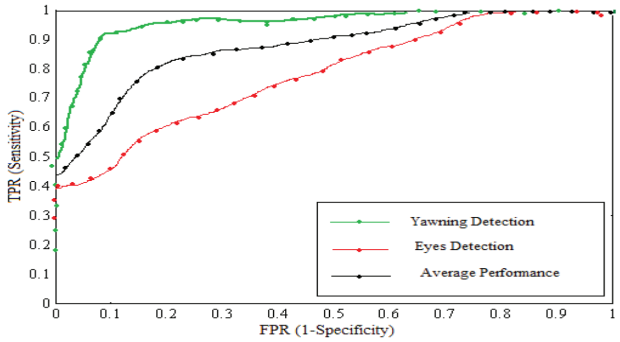


Fig. 5 ROC diagram for the total no. of frames used

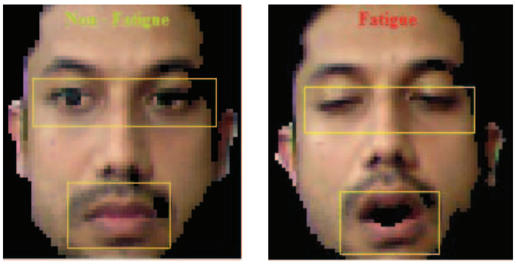


Fig. 6 A non-fatigue and a fatigue frame

IV. CONCLUSIONS AND FUTURE WORK

This paper presents the real time implementation of drowsiness detection which is invariant to illumination and performs well under various lighting conditions. Correlation coefficient template matching provides a super-fast way to track the eyes and mouth. The proposed system achieves an overall accuracy of 94.58% in four test cases, which is highest in comparison to the recent methods. A high detection rate and reduced false alarms makes sure that this system can efficiently reduce the number of fatalities every year.

Despite the highly satisfactory performance, the system was unable to predict drowsiness when the head was tilted towards right or left. Head lowering prediction is also to be included with a threshold. The accuracy drops by 8% with spectacles. In the future, efforts will be made to make the system rotation invariant. The code is available in [13].

REFERENCES

[1] https://data.gov.in/keywords/indian-road-accident-data

[2] Pauly, Leo, and Deepa Sankar. "Detection of drowsiness based on HOG features and SVM classifiers." 2015 IEEE International Conference on Research in Computational Intelligence and Communication Networks (ICRCICN). IEEE, 2015.

[3] Li, G., Lee, B. L., & Chung, W. Y. (2015). Smartwatch-Based Wearable EEG System for Driver Drowsiness Detection. Sensors Journal, IEEE, 15(12), 7169-7180.

[4] R. Nopsuwanchai, Y. Noguchi, M. Ohsuga, Y. Kamakura, and Y. Inoue,"Driver-independent assessment of arousal states from video sequencesbased on the classification of eyeblink patterns," in Intelligent Transportation.

[5] Rezaee, Khosro, et al. "Real-time intelligent alarm system of driver fatigue based on video sequences." Robotics and Mechatronics (ICRoM), 2013 First RSI/ISM International Conference on. IEEE, 2013.

[6] Du, Yong, et al. "Driver fatigue detection based on eye state analysis." Proceedings of the 11th Joint Conference on Information Sciences. 2008.

[7] Choi, In-Ho, Sung Kyung Hong, and Yong-Guk Kim. "Real-time categorization of driver's gaze zone using the deep learning techniques." 2016 International Conference on Big Data and Smart Computing (BigComp). IEEE, 2016.

[8] Tawari, Ashish, Kuo Hao Chen, and Mohan Manubhai Trivedi. "Where is the driver looking: Analysis of head, eye and iris for robust gaze zone estimation" Intelligent Transportation Systems (ITSC), 2014 IEEE 17th International Conference on. IEEE, 2014.

[9] Singh, R. K., et al. "A real-time heart-rate monitor using non-contact electrocardiogram for automotive drivers." 2016 IEEE First International Conference on Control, Measurement and Instrumentation (CMI). IEEE, 2016.

[10] Viola, Paul, and Michael J. Jones. "Robust real-time face detection." International journal of computer vision 57.2 (2004): 137-154.

[11] Basilio, Jorge Alberto Marcial, et al. "Explicit image detection using YCbCr space color model as skin detection." Proceedings of the 2011 American conference on applied mathematics and the 5th WSEAS international conference on Computer engineering and applications, IEEE ICCE. 2011.

[12] Mahmood, Arif, and Sohaib Khan. "Correlationcoefficient-based fast template matching through partial elimination." Image Processing, IEEE Transactions on 21.4 (2012): 2099-2108.

[13] https://in.mathworks.com/matlabcentral/fileexchange/55152-drowsiness-detection-using-a-binary-svm-classifier