Automatic Fatigue Detection of Drivers through Pupil Detection and Yawning Analysis

Conference Paper · January 2010

DOI: 10.1109/ICICIC.2009.119 · Source: IEEE Xplore

3 authors:

Tayyaba Azim
Institute of Management Sciences, Peshawar, P…

Arfan Jaffar
Imam Muhammad bin Saud Islamic University

Anwar Mirza
King Saud University

Automatic Fatigue Detection of Drivers through Pupil Detection and Yawning Analysis

Tayyaba Azim
Department of Computer Science, National University of Computer and Emerging Sciences, Islamabad,Pakistan.
Email: tayyaba.azim@nu.edu.pk

M.Arfan Jaffar
Department of Computer Science, National University of Computer and Emerging Sciences Islamabad,Pakistan.
Email: arfan.jaffar@nu.edu.pk

Anwar Majid Mirza
Department of Computer Science, National University of Computer and Emerging Sciences Islamabad,Pakistan.
Email: anwar.m.mirza@nu.edu.pk

***Abstract*—This paper presents a non-intrusive fatigue detection system based on the video analysis of drivers. The system relies on multiple visual cues to characterize the level of alertness of the driver. The parameters used for detecting fatigue are: eye closure duration measured through eye state information and yawning analyzed through mouth state information. Initially, the face is located through Viola-Jones face detection method to ensure the presence of driver in video frame. Then, a mouth window is extracted from the face region, in which lips are searched through spatial fuzzy c-means (s-FCM) clustering. Simultaneously, the pupils are also detected in the upper part of the face window on the basis of radii, inter-pupil distance and angle. The monitored information of eyes and mouth are further passed to SVM (Support Vector Machines) that classify the true state of the driver. The system has been tested using real data, with different sequences recorded in day and night driving conditions, and with users belonging to different race and gender.**

***Keywords*-SVM;yawning detection;s-FCM**

I. Introduction

The increasing number of road accidents due to fatigue and low vigilance level has become a serious problem for the society. Statistics reveal that between 10% to 20% of all the traffic accidents are due to drivers with a diminished vigilance level[1].This problem has increased the need of developing active safety systems, that can prevent road accidents by warning the drivers of their poor driving condition. Drivers with a diminished vigilance level and fatigue suffer from a marked decline in their abilities of perception, recognition, and vehicle control, and therefore pose serious danger to their own life and to the lives of other people. Fatigue of the driver is defined as a mental or physical exhaustion due to which he/she either feels drowsy or cannot respond appropriately to the situation. The prevention of such accidents is a major focus of effort in the field of active safety research.

Most of the previous research focuses on the collection of ocular parameters to detect the fatigue level of driver[1],[2],[3]. In most of them, the system initializes with the detection of eyes to gather the ocular parameters like blinking frequency, PERCLOS, eye state etc. These systems may fail to predict the driver's true state if the eyes are not detected due to varying lighting conditions or vibrations experienced in real driving scenario. It happens because of their reliance on merely ocular parameters. Hence, it cannot be denied that an approach depending on different visual facial features will work more effectively than the one depending on just one feature. For this reason, our proposed system incorporates yawning feature along with the eye state analysis to determine the fatigue level of driver in real time. The computational overhead is reduced by decreasing the search area for the mouth and pupils. The mouth region is detected by an improved Fuzzy C-Means clustering technique that deploys spatial as well as spectral information along with an optimal number of cluster statistics to segment the lips accurately. Other check of parameters further increase the reliability of mouth region detection for yawning analysis.

The paper is arranged as follows: In Section 2, a review of the previous studies in this line is presented. Section 3 describes the overall system architecture, explaining its main modules. The experimental results are shown in Section 4 and finally we present the conclusion and future studies in Section 5.

II. Literature Review

Different authors have tried different algorithms and combination of parameters to solve the vigilance and fatigue detection problem. Following is the review of some of the notable computer vision techniques used by different authors:

Bergasa, Nuevo, Sotelo, Barea and Lopez[1] proposed a non-intrusive computer vision system for monitoring driver's vigilance in real time. They chose six parameters: Percent Eye closure (PERCLOS), eye closure duration, blinking frequency, nodding frequency, face position, and fixed gaze. These parameters were combined using a fuzzy classifier to infer the level of inattentiveness of the driver.

Qiang Ji and Yang[2] discussed a real time prototype for monitoring driver vigilance. There were two CCD cameras embedded on the dashboard of the vehicle; the first camera was a narrow angle camera, focusing on the driver's eyes to monitor eye-lid movement while the second camera was a wide angle camera that focused on the driver's head to track and monitor head movement. The system used visual bio-behaviors like eye-lid movement, face orientation and gaze movement (pupil movement) that typically characterize a driver's level of vigilance.

Smith, Shah and Lobo[3] proposed a system that relied on estimation of global motion and color statistics to robustly track a person's head and facial features. The system classified driver's rotation in all viewing directions and detected eye/mouth occlusion due to head rotation.It also detected eye blinking and eye closure, and recovered the 3D gaze of the eyes.

Saradadevi and Bajaj[6] proposed a method to locate and track driver's mouth using cascade of classifiers proposed by Viola-Jones for faces. SVM is used to train the mouth and yawning images. During the fatigue detection, mouth is detected from face images using cascade of classifiers. Then, SVM is used to classify the mouth state and detect yawning for fatigue monitoring.

III. Proposed Approach

We present an algorithm that detects driver fatigue by analyzing changes in the mouth and eye state. The method proposed can be divided into the following phases:

- Image Acquisition
- Face Detection and Tracking
- Pupil Detection
- Mouth Detection
- Yawning Analysis and Eye State Analysis
- Driver State Analysis

The outline of the proposed approach is shown in Figure1. In the following sections, each of the phase will be discussed in detail:

A. *Image Acquisition*

The purpose of this stage is to acquire video images of driver's face so that directly observable visual cues could be gathered for fatigue determination. The image acquisition system consists of a low-cost charged coupled device (CCD) micro camera sensitive to near-IR. The acquired images should be relatively invariant to light conditions and should facilitate pupil and mouth detection. This purpose is served by near-IR illuminator which highlights our facial features of interest and reduces the effect of ambient light[1]. The video data is collected in real driving scenarios and is obtained from[1].

B. *Face Detection and Tracking*

In order to collect the yawning information of driver, the mouth should be localized first. But, since, it is a cumbersome task to look for the mouth in whole frame with changing light and background conditions, hence the area of search is reduced by looking first for the face in video frame. The face is detected using Viola and Jones[11] real time face detection technique based on AdaBoost algorithm. The face detection technique in AdaBoost is comprised of three aspects: the integral image, a strong classifier comprising of weak classifiers based on the AdaBoost learning algorithm, and an architecture comprising of a cascade of a number of strong classifiers. A 25-layer cascade of boosted classifiers is trained to detect multi-view faces. A set of sample face and non-face (termed as background) images are used for training. AdaBoost face detection algorithm detects faces in a rapid and robust manner with a high detection rate at 15 frames per second.The face detector is trained such that it is able to detect faces that are tilted up to about ± 15 degrees in plane and about ± 45 degrees out of plane (towards a profile view). However, it becomes unreliable with rotation more than this.

Face tracking is carried out using Kalman filter. The Kalman filter is an efficient recursive filter that estimates the state of a dynamic system from a series of incomplete and noisy measurements. It estimates the position and uncertainty of the moving objects in next time frame i.e. where to look for face, and how large a region should be searched in next frame, around the predicted position to be sure to find face with certain confidence. It works reasonably well with frontal face orientations but fails if the face is not bright enough from the background due to external illumination interferences. It also fails when a sudden head movement occurs because the assumption of a smooth head motion is violated. To overcome this problem, an adaptive search window could be used whose size is determined automatically based on the face position, face velocity and location error[1].

C. *Pupil Detection*

The goal of pupil detection task is to constantly monitor the eye state of the driver for determining his/her vigilance and fatigue level. Since, eyes lie in the upper part of the face region, hence the face detected is cropped, and the eyes are searched in the upper half of the face image. This further helps us in reducing the computing cost of the fatigue detection system and makes it a good candidate for real time use.

The pupils are detected by searching the cropped image region for two bright blobs that satisfy certain geometric constraints. The cropped image is first binarized using adaptive threshold for detecting the bright blobs in image. This image further goes under morphological operations so that spurious bright pixels in upper face region are eliminated.

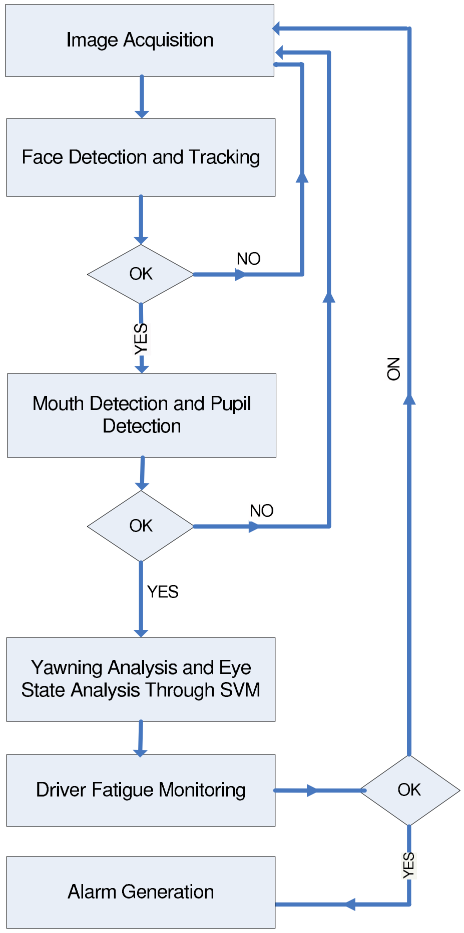

Figure 1. Block diagram of fatigue detection system

Figure 2. (a) Image obtained after performing dilation and opening (morphological operators) on adaptively thresholded image (b) Image obtained after eight connected component analysis (c) Possible circular regions initially detected and marked through red asterisks (d) Image obtained after applying geometric constraints, the pupils finally detected are represented by red plus symbol .

After that, a standard eight connected component analysis is applied to identify the binary blobs that satisfy certain size and shape constraints. The size and shape parameters chosen are: radius, inter-pupil distance, and angle between the pupils. The best candidates are selected depending on their size, intensity, position and distance, whereas the circular regions that do not satisfy these geometric constraints are removed. All the possible pairs are evaluated and the pair with the highest aptitude is selected as pupils. See Figure2 for illustration of pupil detection phase.

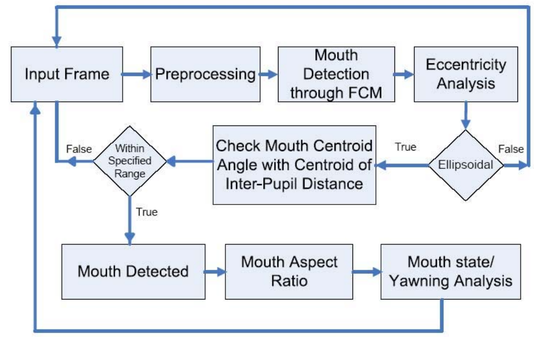

Figure 3. Block diagram of mouth detection and yawning analysis phase

D. Mouth Detection

The mouth region is extracted from the face window for yawning analysis of the driver. This further boosts up the detection speed and accuracy of system as research shows that only a single visual cue takes more time to raise an alarm in fatigue monitoring applications[2]. The mouth is searched in the lower most part of the face window. The block diagram of the mouth detection phase is shown in Figure3.

After the mouth region is extracted from face image, it is passed to the FCM module. Fuzzy c-means (FCM) is a method of clustering which allows one piece of data to belong to two or more clusters. The FCM algorithm creates a fuzzy partition that properly describes the image, using only the pixel intensity feature. This phase involves the calculation of cluster centers and membership functions in the spectral domain only. In order to keep the system totally autonomous, an automatic way to determine the right number of clusters, is needed. This was done by iterating the FCM algorithm for a range of hypothesized number of clusters and choosing the best option, based on a cluster validity measure. After finding out the optimal number of clusters, we pass this information to the spatial FCM. Spatial FCM also considers neighbouring information of each pixel along with the spectral information and returns fuzzy membership matrix. Spatial information incorporated by FCM helps us to remove noise in the image. After the convergence of s-FCM, defuzzification is applied to assign a specific cluster to each pixel for which the membership is maximal. Thus, the output of FCM module is a segmented lip image in binary form[12].

The image further goes under two tests to verify that mouth has been detected successfully. The first test is based on eccentricity analysis. Since, the lips region is more like an ellipsoid; hence its eccentricity is closer to 1.Thus a check is made on the eccentricity of detected regions; if it is closer to 1, it is checked further for the second test else the next consecutive video frame is grabbed to detect the

Figure 4. Illustration of condition 2 for mouth detection.The detected mouth centroid should be in a specific angle range from the inter-pupil centroid

correct mouth pixels. The second test checks the centroid of the detected lip region to be in a specific angle range with the inter-pupil centroid. In a frontal face pose, the centroid of the lip region should be perpendicular to the point which is the centre of inter-pupil distance. But if the whole lip region is not detected successfully, then the centroid might be misplaced, in that case the flexibility in angle range ensures the true presence of lip region. The formula for calculating the angle between centroids is given as:

$$\theta = \arctan\left(|y_2 - y_1|/|x_2 - x_1|\right) \tag{1}$$

where (x_1, x_2)represents the centroid of the detected lip region and (y_1, y_2) denotes the inter pupil centroid. The detected centroid point passing the two tests is a true representative of lips.This concept is best illustrated in Figure4

We then calculate the height and width of the lip region to determine the degree of mouth openness in different frames. To avoid scaling of mouth in different frames and the variations between subjects, we use the aspect ratio of mouth bounding rectangle to represent degree of mouth openness. The degree of mouth openness is defined as follows:

$$DoO = w/h = w/(h * \cos\theta) \tag{2}$$

where w is the width of mouth bounding rectangle, which is the distance between the two lip corners, and h is the height of mouth bounding rectangle, which is measured by the distance between two lines running through the upper and lower lip boundaries.

E. Yawning Analysis and Eye State Analysis

The decision of drivers fatigue depends on the eye and mouth state of the driver. Thus, the width to height ratios of eyes and mouth are passed as features to SVM (Support Vector Machines) for classification. SVM is a useful technique for data classification whose goal is to produce a model which predicts target value of data instances in the testing set which are given only the attributes. The kernel function used in SVM is Radial Basis Function (RBF). The RBF kernel has several advantages over linear kernel[6].

There are several advantages of SVM. The most important advantage is that during the training process, only a few vectors out of the training set are selected to become support vectors. This reduces the computational cost and provides a better generalization. Another advantage is that there are no local minima in the quadratic program, so the found solution is always the optimum of the given training set. Finally, the main advantage is that the solution is not dependent on initial conditions unlike neural networks[6].

F. Driver State Analysis

The state of the driver is determined by the result produced by SVM. If the SVM observes that driver's eyes are closed or half open for several consecutive frames or finds that the person is yawning along with the closed or half open eyes, it immediately concludes that the person is feeling fatigue and an alarm is generated. This saves the system from generating false alarm when he/she is talking while driving. Similarly, if just closed eyes are observed in several consecutive frames, then the system does not wait for the true yawning response by SVM, it immediately rings the alarm assuming the severity of situation. In short, yawning alone cannot correctly predict the state of the driver.

IV. Experiments and Results

The proposed system was implemented on Pentium 4, 1.7 GHz, using the Matlab environment. The video data is collected in real driving scenarios and is recorded in day and night lighting conditions. The input video format is RGB with 400x320 size images. The frame rate of the video acquired from camera is 15frames per second. The experimental results shown by fatigue detection system on real driving video sequences are 92% on the average.

We tested the method on six different videos acquired in real driving conditions. The system fails to detect the fatigue when the driver rotates his head to an extent that the mouth or pupil detection is missed. The system reinitializes when the driver moves his head to front. The detection results are not affected by the exposure of teeth due to the optimal number of cluster check in spatial fuzzy c-means clustering. Moreover, the selection of dynamic threshold segments the lip region successfully and is not effected by the lighting conditions unless the brightness is so high that no contrast in image is present. In that case, the frame under consideration is skipped and next frame is checked for yawning analysis. But if the bad lighting condition persists, the system fails to detect the driver's fatigue.

V. Conclusion and Future Work

This paper discusses a real time fatigue detection system that makes use of yawning and eye state analysis to measure the level of inattentiveness of driver. The system was tested using different sequences recorded in real driving scenarios, with different users. The system works robustly at night time because of the IR illuminator being used. The performance of the system decreases during daytime, especially in bright days. A possible solution to this problem could be the incorporation of IR filters in the car's glasses[1].

There are many future directions for driver's fatigue detection system. At the moment, the system does not detect accurately the yawning gesture if the driver puts hand on his mouth while yawning (a typical yawning gesture).The mouth cannot be segmented in that frame and when the driver puts off his hand, the yawning gesture is over too. Hence, some technique must be incorporated to solve this problem. Moreover, at the moment, the system also fails to detect the pupils accurately, if the driver is wearing glasses. Hence, this improvement must also be incorporated in the system to make it work efficiently.

ACKNOWLEDGMENT

The authors Ms. Tayyaba Azim and Mr. M. Arfan Jaffar would like to acknowledge the Higher Education Commission (HEC), Govt. of Pakistan and NU-FAST for providing funds and required resources to complete this work. We are also thankful to Jesus Nuevo for sharing the videos of drivers with us. It would have been impossible to complete this effort without their continuous support.

REFERENCES

[1] Luis M. Bergasa,Jesus Nuevo,Miguel A. Soteli, Rafeal Barea, Maria Elena Lopez, "Real time system for monitoring driver Vigilance", *IEEE Transactions on Intelligent Transportation Systems*, Vol.7, No.1, March 2006.

[2] Qiong Wang, Jingyu Yang, Mingwu Ren, Yujie Zheng,"Driver Fatigue Detection: A Survey", *Proc. 6th World Congress on Intelligent Control and Automation*, Dalian, China, June 21-23, 2006.

[3] Qiang Ji, Xiaojie Yang, "Real-Time Eye, Gaze, and Face Pose Tracking for Monitoring Driver vigilance", *Real Time Imaging*, Vol.8, No.5, pp. 357-377, Oct. 2002.

[4] Ilkwon Park, Jung-Ho Ahn and Hyeran Byun, "Efficient Measurement of the Eye Blinking by Using Decision Function for Intelligent Vehicles", in *ICCS 2007, LNCS 4490*, Part IV, 2007, pp. 546-549.

[5] P. Smith, M. Shah, and N. D. V. Lobo, "Determining driver visual attention with one camera", *IEEE Trans. Intell. Transp. Syst.*, Vol. 4, No. 4, pp. 205218, Dec. 2003.

[6] Mandalapu Saradadevi, Preeti Bajaj, "Driver Fatigue Detection Using Mouth and Yawning Analysis", *IJCSNS International Journal of Computer Science and Network Security*, Vol.8, No.6, pp.183-188, June 2008.

[7] Yoshio Matsumotoy, Alexander Zelinsky, "An algorithm for Real time Stereo Vision Implementation of Head Pose and Gaze Direction Measurement", *Proc Int. Conf. Series Vision Vehicle (VIV9)*, Brisbane, Australia, Aug. 2001.

[8] Qiang Ji, Zhiwei Zhu, Peilin Lan, "Real Time Non-intrusive Monitoring and Prediction of Driver Fatigue", *IEEE Transactions on Vehicular Technology*, Vol. 53, No. 4, July 2004.

[9] M. Imran Khan and A. Bin Mansoor, "Real Time Eyes Tracking and Classification for Driver Fatigue Detection", *Springer Verlag*, 2008, pp. 729-738.

[10] Jorge P. Batista, "A Real-Time Driver Visual Attention Monitoring System", in *LNCS*, Vol. 3522, May 2005, pp. 200-208.

[11] Paul Viola, Michael J. Jones, "Robust Real-Time Face Detection", *IEEE ICCV Workshop on Statistical and Computational Theories of Vision*, pp. 137-154, 2001.

[12] M. Arfan Jaffar, Ayyaz Hussain, Anwar M. Mirza, Asmat ullah Ch, "Fuzzy Entropy and Morphology based fully automated Segmentation of Lungs from CT Scan Images", *International Journal of Innovative Computing, Information and Control (IJICIC)*, Vol.5, No.12, December 2009.

[13] J. Healyand R. Picard, "Smart-Car: Detecting Driver Stress", *Proc. 15th Int. Conf. Pattern Recognition*, Barcelona, Spain, Vol.4, 2000, pp. 218-221.