Riyadisi – Automated Driver Attention Assistance System

K U G S Darshana, M D Y Fernando, S S Jayawardena, S K K Wickramanayake, C R de Silva

Abstract

In recent years, road crashes have produced a large number of deaths and injuries the world over. It has been calculated that 10% to 20% of fatal traffic accidents are caused by driver inattention. Two major causes for driver inattention are driver drowsiness and driver distraction. In order to help in reducing these fatalities this paper introduces a new Driver Attention Assistance System. The proposed system is a real time automated noninvasive system which uses computer vision and machine learning techniques to detect driver drowsiness and distraction and alerts the driver appropriately. A video stream is analyzed in detecting driver inattention. In this system PERCLOS, yawning analysis and nodding off are used to detect drowsiness whereas gaze detection and head movements are used to detect distraction

1. Introduction

Around the world nearly 1.3 million people die each year due to road crashes. In addition, 25-50 million are injured or disabled. It has been calculated that 10% to 20% of traffic accidents with driver fatalities are caused by driver inattention [1]. Drowsiness and distraction of drivers can be identified as main causes for driver inattention. Drowsiness is unexpected, unavoidable and beyond the control of the driver. Stress, illnesses, certain type of medications, monotonous driving and liquor can make the driver drowsy. Sleepiness increases reaction time which is critical in driving. Reduced attention and increased reaction time increase the probability of road accidents.[2] Sleep related accidents appear to be more severe, possibly because of the higher speeds involved and the driver being unable to take any avoiding action, or even brake, prior to the collision. [1] Distractions such as looking away, using mobile phones or navigation systems also make a driver lose his focus.

The aim of Riyadisi is to contribute to road accident reduction by assisting drivers to maintain their attention throughout the journey. The proposed system uses typical visual features of a drowsy person; decreased eye blinking rate, yawning and nodding off, to detect drowsiness, whereas it uses eye-gaze movement and head movement in distraction detection. This paper presents how the introduced system carries out attention assistance.

1.1 Background

Driver assistance technologies are evolving rapidly. Studies done on driver attention assistance systems can be broadly categorized into three sections.

- Bio-physiological methods
- Vehicle based performance methods
- Visual cue based methods

[1] K.U.G.S. Darshana, M.D.Y. Fernando, S.S. Jayawardena, S.K.K. Wickramanayake, C.R. De Silva are with the Department of Computer Science and Eng., Univ. Moratuwa. (e-mail: sankadarshana@gmail.com; fdileepa@gmail.com; sadari.jayawardena@gmail.com; sw0900@gmail.com; chathura@uom.lk)

Physiological methods provide the most accurate results as they depend their processing on bio-physiological data of the driver such as brain wave patterns, respiratory dynamics and heart rate variations [3]. However due to intrusive nature the usability of those systems, in real time environments is not acceptable. Further, accuracy of this technique decreases due to perspiration effect. Vehicle - based performance methods use speed variations, steering wheel movements (SWM), standard deviation of lane position (SDLP), etc for monitoring driver attention. Automobile companies have already developed driver assistance systems on their own, using sensors and cameras. Volvo, Mercedes Benz, Ford and Toyota are among them. According to Sahayadha et al., [4], SWM's accuracy depends on the geometrical characteristics of the road and the kinetic characteristics of the vehicle and SDLP depends on the external factors such as road marking, climatic and lighting conditions. On the other hand these methods are specific to a particular brand and they may require advanced mechanical components. Further systems that use vehicle-based parameters to detect drowsiness alerts the driver only after sleepiness of the driver adversely affects driving.

The issues of these two methodologies can be eliminated if visual cues based methods are used to monitor driver attention. Visual cues based methods are totally non-intrusive, but they have to cope with different challenges. The main challenge is different illumination conditions. To overcome that challenge Riyadisi uses an IR camera which works both in the day and at night. On the other hand these systems should be able to monitor human subjects of different ethnic backgrounds, gender and ages; also with/without glasses.

This vision based system performs a series of classification tasks in face detection, eye detection and eye state detection. Some of these classifications are image based and some are feature based which will be discussed later. In either case there is a significant processing overhead involved. In order to reduce processing overhead in classification and extracting geometric detail in the image, preprocessing techniques such as edge detection, histogram equalization and filtering are used.

In addition, there are some other preprocessing mechanisms which can be used generally in any classification task that are not restricted to image classification. These techniques deal with the major concern, number of attributes/dimension. Attributes we consider in an image are pixels. There are thousands of pixels in the images portion of interest, and so there are thousands of dimensions that we have to deal with. Hence, we need a mechanism to reduce dimension. Principal component analysis is a statistical technique which projects the data along non correlated set of axes (Eigen vectors). This method has been used by similar experiments in vision based drowsiness detection.(Rami et al.,2011)[5]. First principal component gives the maximum variance for the data set along a direction and the second component contains the maximum variance among the directions which are orthogonal to the first principal component and so on. Dimensionality reduction is achieved by omitting the axes with lower variances. Due to the independence between the axes PCA should provide the attributes which would help to easily discriminate between classes (in our case classes of images). Unfortunately, in real world situations this is not straightforward since the first few principal components contain the details about noise which is common to all the classes. Therefore, it is evident that we cannot use the first few principal components for discrimination. One alternative is omitting the very first principal components from the calculation which should be done with care, because there is a possibility of losing important information as well. In order to overcome this difficulty another statistical method called "Linear Discriminant Analysis" can be used.[6]

Visual cue based drowsiness detection mainly consists of three steps.

- Face detection
- Facial feature extraction
- Facial feature analysis for fatigue detection

Face detection techniques can be categorized into two broad categories; feature-based approaches and image-based approaches.

Using pixel properties such as edges, pixel intensity, skin color etc. to detect visual features comes under feature-based approaches [7]. Skin-like pixel extraction process makes multi-view face detection straightforward but it is not robust in varying light condition. Hence, it is not appropriate for night time operating systems. Further skin color depends on the ethnic background of the person. Using information from face geometry to identify various facial features is another aspect of the feature-based approach [7]. This approach works very well for frontal face detection. However it is orientation and scale invariant.

Image-based approaches make use of pattern recognition techniques for face detection [7]. This approach is followed by a training procedure, in which example images are classified as faces and non-faces. A comparison between these two classes allows deciding the existence of a face in a given image. Template matching is the simplest image-based approach. The disadvantage associated with this method is the difficulty in handling variations in size, orientation, and expressions etc due to the static nature of the templates [8]. Most of the other image-based methods apply window scanning technique to detect faces. The size of the scanning window, the subsampling rate, the step size, and the number of iterations depends on the window scanning algorithm [7].

'Viola Jones' is a feature based face detection algorithm that takes the advantage of cascade classification and AdaBoost. Using the cascade classifiers reduces the overhead of calculating feature attributes for all the input images. The first few classifiers reject a large number of negative images; hence computation task of last few classifiers is significantly reduced. In our context, minimizing false negative ratio is more important. Here, AdaBoost algorithm is used to create an optimal integration of weak classifiers which results in a guaranteed upper bound for the error rate. In order to minimize the false negative ratio, we have to give more weight to false negative examples compared to false positive examples, in each iteration of AdaBoost. [9] Main techniques for drowsiness detection are eyelid movement analysis, yawning analysis and nodding-off detection. Out of these three, eyelid movement analysis has become the most popular technique. Eye localization can be done under either normal illumination (passive eye detection) or IR illumination (active eye detection). Detecting eye location in different light conditions and when the driver is wearing sun glasses are challenging with passive eye detection. However, eye detection with IR illumination eliminates these challenges [10].

Yawning analysis consists of both yawning detection and drowsiness prediction. Classification task will be complicated since visual data is not sufficient to distinguish between yawning mouth and normal mouth.

Among all the candidates, SVM and neural network based algorithms are more appropriate for eye state classification. In [11](Marco Javier Flores et al.) researchers have achieved 94% accuracy with SVM with RBF kernel for eye state classification. In [12] 91.8% accuracy was achieved with back propagation neural networks for the above task. Since SVM can be described as a linear model which automatically minimizes over fitting by design, we chose it over neural networks. When the data points are linearly separable, SVM gives the maximum margin separation.[13] Otherwise, we will have to look for a suitable kernel function as a discriminant between the two classes. This will project the input image vector to a higher dimensional space which will make training and classification computationally intractable. The kernel trick can be used in calculating the training model and doing the classification to make the computational complexity almost similar to that in linear separation. However, SVM's classification complexity is significantly higher than neural networks. Since we are doing real time classification we have to worry about the classification complexity. In this scenario, neural network is the better candidate. If the designed SVM's classification time is tolerable we can always move to SVM option due to its high accuracy. An elegant choice of a kernel function will make the SVM highly accurate.

Final output from the classification process will produce PERCLOS and yawning rate measures. Choice of the eye state and yawning classifier will depend on the level of accuracy asked by the above measures for a successful classification of drowsiness levels and the minimum time needed to trigger a drowsy situation.

In detecting distraction, Computer Vision based methods such as gaze detection, head movement analyses and hand movement analysis can be used. Ahlstrom and Kircher use a predefined field relevant for driving (FRD); the area where the driver is looking when driving, for looking away detection [*9*]. The measurement of looking away detection is then defined as duration of glances residing outside the field relevant for driving. Instead of same weighted gazes, using a weighted gaze score increases the accuracy because, glances in the rearview mirror and looking in the side mirrors for a turn should not be considered as driver distractions. Furthermore, scenarios in which the driving has been paused should be skipped, since looking away at that time does not constitute a significant distraction to driving.

2. Methodology

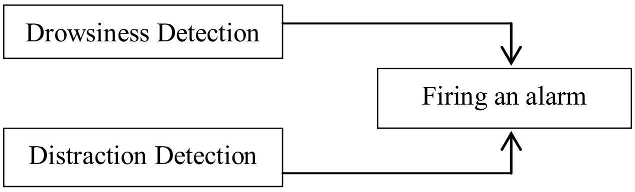

Figure 1: Overview of the system

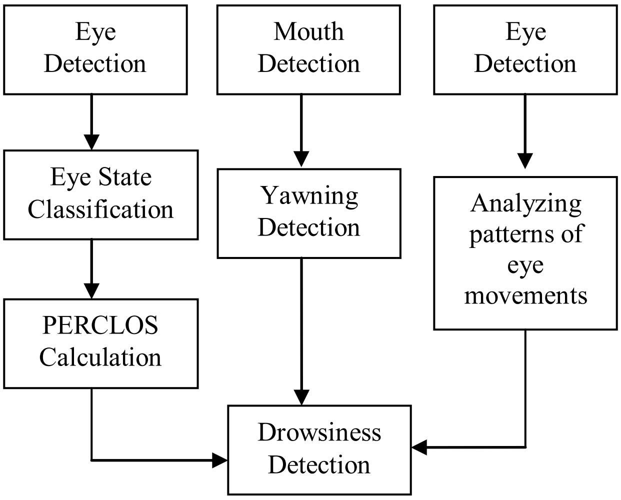

Figure 2: Drowsiness detection

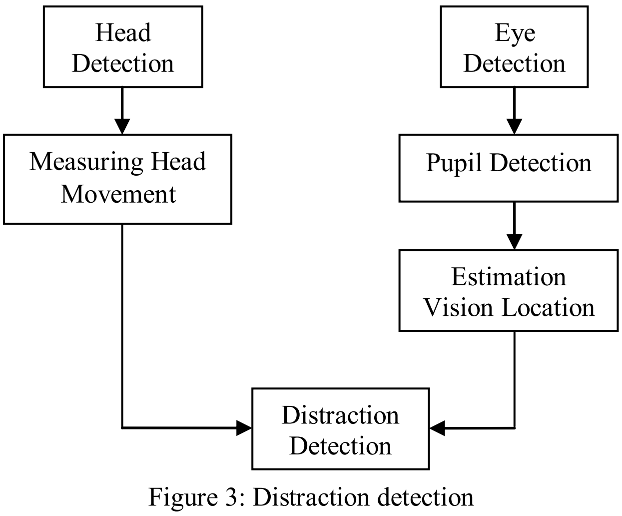

Figure 3: Distraction detection

2.1 Visual Cues Detection

First phase of the system proposed in this paper is face detection. In order to identify all the possible actions of the driver, the face should be detected regardless of face orientation. Frontal face detection and rotations in all three directions should be simultaneously processed here. However, in our context we can assume that driver's head will not be rotated beyond the 180 degree front space in any direction.

A video stream is fragmented into a set of frames and each frame is processed to decide whether a face can be located. The face detection methodology used in the proposed system is Viola-Jones algorithm presented in [10]. This system considers 11 different face poses – frontal, left-profile, right-profile, looking-up, looking-down, three clockwise positions and three anti-clockwise positions-. Clockwise and anti-clockwise positions are -90°,-60°, -30° +30°, +60°, +90° rotations from the Y axis. Detecting with an angle difference larger than 30 degrees has relatively low accuracy. On the other hand, using low angle difference would reduce the performance. It is not necessary to train classifiers for each of these 11 positions. Apart from frontal face detection, frontal face classifier can be used to detect rotations around the Y axis as well. Both left and right profile faces can be detected using classifier for left-profile.

Detecting the face at the first stage has an advantage because it reduces the search area for other facial features. As an example, this system is not searching the lower half of the face during eye location detection. The next state of our system is eye and mouth localization. Eye detection will also make use of the Viola-Jones algorithm. Mouth detection becomes difficult for the people with a beard. When using Viola-Jones algorithm, more than one classifier is needed for accurate mouth localization. Hence, in our system, the largest blob located in the middle of the face is identified as the mouth. When an eye or a mouth is detected it is saved as a grayscale image. Once saved, those images are normalized to obtain a standard scale.

A state detection component uses identified eye and mouth locations and identifies their binary states –closed & open eyes states and normal & yawning mouth states-. For state detection, we must have sufficient training data corresponding to different eye states and mouth states. Eye states are namely closed and open, while mouth states are normal and yawning. An SVM will be trained to detect different eye states and mouth states. To make the eye state detection training data more representative, we gather images of people wearing spectacles, people with different skin colors, people from both genders and people from different cultures. Additionally this set should also include images taken in different illumination conditions. Similarly, mouth state detection training set would include images of people with a beard, people with different skin colors etc. A principal component analysis may be performed to reduce dimensionality of the images depending on the required efficiency level.

Gaze detection is one of the most important parameters used in detecting driver distraction. Integrated with the drivers facial features, driver's pupils are used for gaze detection. After detecting the pupils of both eyes, the position of the pupils with reference to the face and facial features such as nose and eyes are used to decide where the driver's gaze is at. Although there are methods to gain higher accuracy using features of the pupils, due to practical reasons such as the processing power requirements, our system uses the above method, which although a less accurate method, will give a higher performance when coupled with other distraction detection features of the system.

Most of the existing systems depend only on one parameter such as PERCLOS, yawning frequency etc. to detect either drowsiness or distraction. However predicting fatigue situations based on one parameter is not very reliable. This system considers three parameters - PERCLOS, yawning frequency, nodding-off frequency- to detect drowsiness and two parameters - facial pose, gaze point, - to detect distraction. The results of these parameters will be used by a classification model. If the system identifies a fatigue situation, it will start an alarm and advise the driver to stop the vehicle and take a break.

2.2 Decision Making

In the decision making component, we are researching on the most appropriate way in which we can implement the required functionality with minimal resources. The basic model of the decision making component will be processing all the measurements and signals coming from the sensory components and making the decision whether the driver is in drowsiness or in distraction, thereby deciding whether he or she should be alerted or not.

The method we are currently researching on is the feasibility of using neural networks to analyze the time series generated from the sensory components, and to generate minimal level profiling data through sensory data. Profiling is required to generate accurate results for a particular user, yet at a very minimal level so the system would not depend on a specific user, but a more general user model. This neural network component is expected to weigh each of the data streams acquired from sensory components due to the fact that different drivers tend to show different levels of drowsiness and distraction signs. For an example, one driver might have drowsy-like eyes even when he is not drowsy and the variations between signs tend to differ greatly from one driver to the other. Therefore the system is developed generically so that it would be able to manage the drivers specifically.

3. Results and Discussion

3.1 Results of Face Detection

Face detection process was tested using MIT CBCL Face Database [14]. It contains 52 face images including frontal and profile faces of 10 subjects. Table 1 shows the results of the experiment.

Table 1: Results of multi-view face detection

	Count	Percentage
True positives	49	94.23%
False positives	2	3.85%
False negatives	1	1.92%

Up to now the system has a considerable accuracy. The necessity is to improve the true positive rate while reducing the false negative rate. False positive rate does not affect the accuracy of the system because it does not lose any data related to the driver.

3.1 Results of Eye State Detection

Training set of eye images containing both open and closed eyes were input to SVM training model. Principal component analysis was performed to reduce the dimensionality from 60*40(image resolution) to a few hundreds. Model was validated by testing against test images of different human subjects captured in different illumination conditions.

Table 2: Results of eye state detection

Number of open eye images in training set	Number of closed eye images in training set	Number of Principal Components in training set	Test-open	Test-close	Accuracy	Classification time	Train time
143	69	204	Training set		100%	0.06s	2.7144
143	69	204		Training set	100%	0.03s	2.7456
143	69	204	test-open		100%	0.03s	2.9640
143	69	204		test-close	50%	0.03s	2.8080
143	69	204	eye		58%	0.03s	2.5896
143	69	204		eye-closed	86%	0.03s	2.5428

References

[1] The royal society for the prevention of accidents, “Driver fatigue and road accidents a literature review and position paper”, Edgbaston, UK, 2001.
[2] ChristerAhlström and KatjaKircher Review of Real-time Visual Driver Distraction Detection Algorithms in *Proceedings of Measuring Behavior 2010*, Eindhoven, The Netherlands, 2010, pp 310 - 313.
[3] David Sommer, Martin Golz, Thomas Schnupp, JarekKrajewski,UdoTrutschel, Dave Edwards. “A measure of strong driver fatigue” *International Driving Symposium on Human Factors in Driver Assessment, Training and Vehicle Design* 2011.
[4] ArunSahayadhas, Kenneth Sundaraj, MurugappanMurugappan., “Detecting Driver Drowsiness Based on Sensors: A Review”. *Sensors*, vol. 12, pp. 16937-16953, 2011.
[5] Rami N. Khushaba, SarathKodagoda, Sara Lal, GaminiDissanayake, Uncorrelated fuzzy neighborhood preserving analysis based feature projection for driver drowsiness recognition, Fuzzy Sets and Systems, Vol 221, 16 June 2013, pp 90-111, ISSN 0165-0114, http://dx.doi.org/10.1016/j.fss.2012.12.003.
[6] Theodoridis and K. Koutroumbas, ”Pattern Recognition: Fourth Edition”, Academic Press, Elsevier, USA, 2009
[7] EsraVural, Mujdat Cetin, AytulErcil, Gwen Littlewort, Marian Bartlett, Javier Movellan. “Automated Drowsiness Detection for Improved Driving Safety” in Proc. ICAT, 2008.
[8] Meynet, Julien. "Fast face detection using adaboost." University of Trier. July 2003
[9] Jaeik Jo, Sung Joo Lee, Ho Gi Jung, Kang Ryoung Park, Jaihie Kim, “Vision-based method for detecting driver drowsiness and distraction in driver monitoring system” Optical Engineering 50(12), 127202 December 2011.
[10] Marco Javier Flores, Jose Maria Armingol and Arturo de la Escalera (2012, May 25) Real-Time Drowsiness Detection System for an Intelligent Vehicle [online] Available: http://www.uc3m.es/portal/page/portal/dpto_ing_sistemas_automatica/investigacion/lab_sist_inteligentes/publications/iv08.pdf
[11] Marco Javier Flores, José MaríaArmingol and Arturo de la Escalera (2012, May 25) Real-Time Drowsiness Detection System for an Intelligent Vehicle [online] Available : http://www.uc3m.es/portal/page/portal/dpto_ing_sistemas_automatica/investigacion/lab_sist_inteligentes/publications/iv08.pdf
[12] Y. Ying , S. Jing, Z. Wei, “The Monitoring Method of Driver’s Fatigue Based on Neural Network”, In: Proc. International Conf. on Mechatronics and Automation, China, 2007 , pp. 3555-3559
[13] CHRISTOPHER J.C. BURGES(2013, Apr 28) A Tutorial on Support Vector Machines for Pattern Recognition [Online]. Available:http://research.microsoft.com/pubs/67119/svmtutorial.pdf
[14] Dr. Bernd Heisele, Mary Pat Fitzgerald (1996). MIT CBCL Face Database