ISSN 2348 – 7968

Head Pose and Eye State Monitoring (HEM) for Driver Drowsiness Detection: Overview

Vinay K Diddi[1], Prof S.B.Jamge[2]

[1] Department of M.E Electronics, Walchand Institute of Technology, Solapur University, Solapur, Maharashtra, India

[2] Department of M.E Electronics, Walchand Institute of Technology, Solapur University, Solapur, Maharashtra, India

Abstract

This proposed method presents visual analysis of eye state and head pose (HP) for continuous monitoring of alertness of a vehicle driver. Most of the previous techniques of visual detection to non-alert driving patterns rely either on eye closure or head nodding angles to determine the driver drowsiness or distraction level. Compared with other approaches that only use a discrete number of gaze fixation areas [8], the proposed approach considers all directional head and eye movements of the driver. The proposed method uses considers visual features such as eye index (EI), pupil activity (PA), and head pose (HP) to extract critical information of a vehicle driver to detect whether driver is alert or not. (EI) determines if the eye is open, half closed, or closed from the ratio of pupil height and eye height. Pupil activity (PA) measures the rate of deviation of the pupil center from the eye center over a time period.

Head pose (HP) finds the amount of the driver's head movements by counting the number of video segments that involve a large deviation of three Euler angles of HP, i.e., nodding, shaking, and tilting, from its normal driving position. HP provides useful information on the lack of attention, particularly when the driver's eyes are not visible due to occlusion caused by large head movements.

.

Keywords: *Driver drowsiness, Eye state, head pose, pupil activity*

1. Introduction

Road traffic accidents are a major cause of death with over one million people losing their lives and a further fifty million seriously injured each year worldwide. According to a 2012 poll conducted by the National Sleep Foundation, one in five pilots admit that they have made a serious error, and one in six train operators and truck drivers say that they have had a "near miss" due to sleepiness . In 2008, the National Highway Traffic Safety Administration estimates that 100 000 police reports on vehicle crashes were direct results of driver drowsiness resulting in 1550 deaths, 71 000 injuries, and $12.5 billion in monetary losses. However, recent research into the causation of road accidents have found that momentary lack of attention and drowsiness are the main cause of road accidents. Some researchers claim that lack of attention is the main cause of accidents as factors such as fatigue, alcohol or drug use, distraction and speeding all impair the driver's capacity to pay attention to the vehicle and road conditions. These factors have motivated research efforts that aim to improve driver performance and thus help to reduce accidents.

Driver drowsiness and distraction represents diminished attention to activities that are critical for safe driving in the absence of a competing activity. Driver distraction is a diversion of attention away from activities critical for safe driving toward a competing activity. The term *drowsiness* refers to a combination of symptoms such as impaired performance and a strong desire to sleep. Unlike driver distraction, driver drowsiness involves no triggering event but instead, is characterized by a progressive withdrawal of attention from the road and traffic demands. Both driver drowsiness and distraction, however, might have the same effects, i.e. decreased driving performance, longer reaction time, and an increased risk of crash involvement. Three main approaches have been developed to detect driver inattention, i.e., physiological, driving-behavior-based, and visual-feature-based approaches.

Physiological approach-These approaches focus on using sensors to detect the driver's behavior while driving. Physiological approaches involve analysis of vital signals such as brain activity, heart rate, and pulse rate. As an example, Khushaba et al [1]., developed a fuzzy mutual-information-based wavelet packet transform model to estimate the drowsiness level from a set of electroencephalogram, electrooculogram, and electrocardiogram signals. However, physiological approaches often require electrodes that are attached to the driver's body, which are intrusive in nature and, therefore, may cause annoyance to the driver.

Driving behavior based approaches-Driving-behavior-information-based approaches evaluate the driver's performance over time. Based on the variations in the lateral position, speed, steering wheel angle, acceleration, and breaking, the system determines if the driver is alert or not. Liang *et al.* [2] developed a real-time approach to detecting distraction using the driver's eye movements and

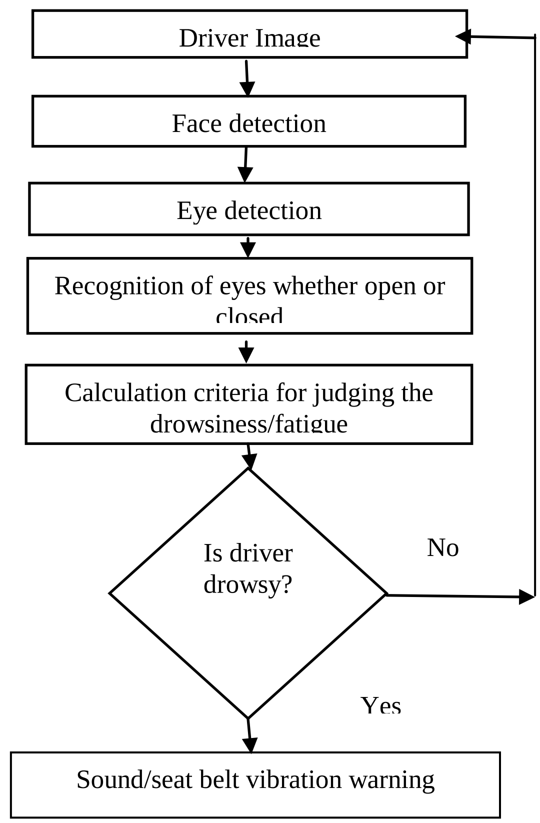
ISSN 2348 – 7968

driving performance data collected in a simulator environment called the in-vehicle information system. Then, the data were used to train and test both support vector machine (SVM) and logistic regression models to detect driver distraction. The analysis by Liang *et al.* suggested that the SVM outperformed the traditional approach of logistic regression in detecting driver distraction. An advantage of this approach is its convenient signal acquisition. However, they highly depend on the vehicle type, driver experience, and the road condition. If a driver falls asleep on a straight road, such systems may fail because the car would not provide any significant information

Visual-feature-based approaches-The feature based approach analyzes visual features from the driver's facial images. Drowsy people often produce unique visual features on the face such as eye blinking, yawning, and eye and head movements. Hammoud *et al.* [6] proposed a driver drowsiness detection system that estimates the status of the eyes in the near-infrared spectrum. Moriyama *et al.* [3] estimated the eye state by creating detailed templates of the shape and texture of the eyelid. As a widely accepted visual measure for drowsiness detection, the percentage of eyelid closure (PC) counts the number of eye blinks of the driver [7]. More recently, Jimenez *et al.* [8] have proposed a gaze fixation system based on a stereo camera system to detect the driver's distraction level in a driving simulator. From the viewpoint of practical applications, visual-feature-based approaches are preferred since they are natural and inherently nonintrusive to the driver.

2. Methodology

Drowsiness is the transition from the awake state to the sleep state, were the ability to concentrate greatly reduces, posing the greater risk of the crash involvement. The objective of this research is to develop a system as specific countermeasures to reduce collisions associated with driver drowsiness. This proposed method presents visual analysis of eye state and head pose (HP) for continuous monitoring of alertness of a vehicle driver. Most existing approaches to visual detection of non-alert driving patterns rely either on eye closure or head nodding angles to determine the driver drowsiness or distraction level. The proposed method is based on using a single camera for continuous monitoring of alertness of a vehicle driver without the use of additional source of light. The proposed scheme finds in real time the eye and pupil centers and HP angles from a face object in a live video stream captured by a camera. The proposed method brings eye state and HP together to make a decision if a driver is not alert. Compared with other approaches that only use a discrete number of gaze fixation areas, the proposed approach considers all directional head and eye movements of the driver. SVMs are based on the statistical learning technique and can be used for pattern classification and inference of nonlinear relationships between variables. This method has been successfully applied to the detection, verification, and recognition of faces, objects, handwritten characters and digits, text, speech, and speakers, and the retrieval of information and images. An SVM classifier, which is trained with the three visual features of EI, PA, and HP, is used to learn the driving patterns of the driver to classify if the subject is either alert or non-alert. The non-alert state represents that the driver is either drowsy or distracted. The flowchart shows the basic steps involved in the drowsiness detection system-:



3. Feature Extraction

3.1 Pupil center and eye detection

The pupil center detection can be done by using a ADABOOST technique, which is also known as adaptive boosting. It is one of the most widely method for visual feature extraction. It is also known as viola jones [9] algorithm. Unlike neural networks and SVMs, the

AdaBoost training process selects only those features known to improve the predictive power of the model, reducing dimensionality and potentially improving execution time as irrelevant features do not need to be computed. A boosted classifier is a classifier in the form

$$F_T(x) = \sum_{t=1}^{T} f_t(x)$$

where each is a weak learner that takes an object as input and returns a real valued result indicating the class of the object. In adaptive template matching, the previously detected face region is used as template *T*.In the next processing cycle, we match template *T* against each pixel within a search window in the next image frame *I*. Then, the normalized sum-of-squares difference *S* is used as a metric of match between the face template and the search region of the face from the previous frame, i.e.

$$S(x,y) = \frac{\sum_{x',y'}[T(x',y') - I(x+x',y+y')]^2}{\sqrt{\sum_{x',y'} T^2(x',y') \sum_{x',y'} I^2(x+x',y+y')}}$$

Where, $T(x, y)$ and $I(x, y)$ denote the brightness intensity of template *T* and source image *I* at (x, y). The matched image becomes a new template to be used for template matching in the following frame.

To detect the center of the pupil only the region of interest is taken into an account, which also reduces the computational load while image processing. Considering the problem the image of the face is divided into four quadrants and only upper two parts are used for the further feature extraction. To improve and detect the pupil objects clearly the eye images are up sampled by two and binarized images are eroded, normalized using morphological operators. The pixels of intensity below a threshold are labeled as the pupil.

Fig 1 shows pupil detection when eye is open,closed or halfly closed

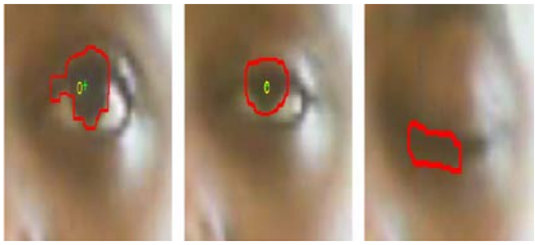


Fig 1

The EI and PA that describe the state of the eye and the pupil. EI is defined as the ratio of the pupil height and the height of the eye in pixels, i.e.

$$EI = \frac{\text{Pupil Height}}{\text{Eye Height}}$$

The pupil tends to show an isotropic shape when the eye is open. When the eye is half closed, the eye region becomes similar to a more rectangular shape. When the eye is closed, on the other hand, the detected eye region becomes to a flat and long shape that goes beyond the geometric constraint to be a pupil. The thresholds for the EI are chosen to determine the three states of the eye, i.e., open, half closed, and closed. We determine that an eye is closed if EI < thresh1, open if EI >thresh2, and half closed if thresh1 < EI < thresh2.

PA measures the temporal activity of eye movements, which gives useful information to determine drowsiness. The coordinates (px, py) in pixels represent the pupil center with the eye center as a reference point. The relative displacement of the pupil between the two consecutive video frames is $(\Delta px, \Delta py)$, where $\Delta px = |px(t + 1) - px(t)|$ and $\Delta py = |py(t + 1) - py(t)|$. Then, the PA index is defined as the sum of the average displacements of the pupil movements in horizontal and vertical directions. When a subject is drowsy or distracted, the PA value tends to be greater. Thus

$$PA = \Delta px + \Delta py.$$

Various detection algoritm and there detection in pecentage table

Algorithm	Detection Rate
Adaboost	91.8%
Cumulative distribution function	96.0%
Generalized projection function	94.8%
Adaboost with adaptive thresholding	97.2%

3.2 Headpose Estimation

Head pose estimation is a key element of human behavior analysis. Many applications would benefit from automatic and robust head pose estimation systems such as driver drowsiness detection,it gives the vital informatiom. While 2D video presents ambiguities hard to resolve in real time, systems relying on 3D data have shown very good results. The approach is to estimate the head pose from monocular images. We can determine the head pose in the three-dimensional space by three Euler angles of rotation around three axis orthogonal to each other.The fig 2 shows the movement of head.

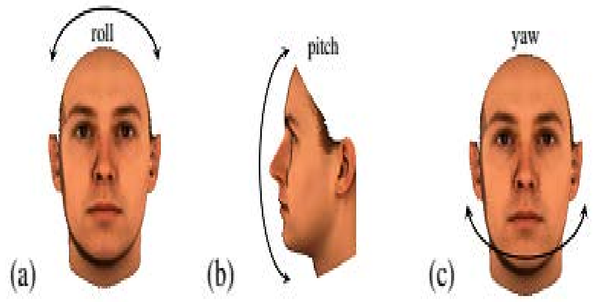


Fig 2

While most of the existing approaches for head pose estimation deal only with poses that vary around the vertical axis.It is possible to estimate the 3D rotation and translation of a 3D object from a single 2D photo, if an approximate 3D model of the object is known and the corresponding points in the 2D image are known. A common technique known as "POSIT" algorithm has been recently developed, where the 3D pose is estimated directly from the 3D model points and the 2D image points, and corrects the errors iteratively until a good estimate is found from a single image. Most implementations of POSIT only work on non-coplanar points. The relationship between a point on the 3-D head model and a point in the 2-D image is expressed as

$$s\begin{bmatrix}u\\v\\1\end{bmatrix}=\begin{bmatrix}f_x & 0 & c_x\\0 & f_y & c_y\\0 & 0 & 1\end{bmatrix}\begin{bmatrix}r_{11} & r_{12} & r_{13} & t_1\\r_{21} & r_{23} & r_{23} & t_2\\r_{31} & r_{32} & r_{33} & t_3\end{bmatrix}\begin{bmatrix}X\\Y\\Z\\1\end{bmatrix}$$

Where, (X, Y, Z) denote the coordinates of a 3-D point in the world coordinate space, and (u, v) denote the coordinates of the projection point in pixels. The main idea is to determine the correspondences between 2D image features and points on the 3D model curve. The POSIT algorithm requires image coordinates of at least four non coplanar object's points.

The algorithm for determining pose estimation is based on the Iterative Closest Point algorithm.

a) Reconstruct the projection rays from the image
b) for each projection ray R;
c) For each 3D contour
  c1) Estimate the nearest point p1 of ray R to a point on the contour
  c2) if (n==1) choose p1 as actual p for the point-line correspondence
  c3) else compare p1 with p
    if dist (P1,R) is smaller than dist(P,R)
      then choose P1 as new P
d) Use (P,R) as correspondence set.
e) estimate pose with this correspondence set
  f)Transform countours , goto (b)

Changes in HP angles provide good information to determine the drowsiness and distraction of the driver. When a person is drowsy, the nodding angle is expected to be high compared with a person that is fully alert and is keeping his head straight.

$$\mathrm{HP} = w1ND + w2SH + w3T\,L$$

where ND, SH, and $T\,L$ denote the number of consecutive video segments of nodding, shaking, and tilting, respectively, that exceed a threshold of 15 degree. When a person is distracted and looking away from the road, the HP difference is expected to be high compared with a person that is driving alert and staring at the road up front.
To track the changes in the head position, we apply the Lucas–Kanade optical flow method [12]. Optical flow finds an estimate of the feature points between two video frames extracted using the good features to track method

$$R = Rz(\phi)Ry(\theta)Rx(\gamma)$$

Given rotation matrix R, HP angles γ, θ, and ϕ are then computed by equating each element in R with its corresponding element from the rotation matrix.

4. Conclusions

This paper represents the visual analysis of eye state and head pose for the driver drowsiness detection, by using a web camera for continuous monitoring of the driver state. This proposed method extracts the visual features from the eyes and head movements of a driver. EI measures eye closures, PA finds dynamic motion of the eye, and HP calculates all directional head movements. This method is a non-contact type drowsiness detection system as compared to that of other system. It does not require any illumination. The proposed scheme offers high classification accuracy with acceptably low errors and false alarms for people of various ethnicity and gender in real road driving conditions

Acknowledgments

We would like to take this opportunity to thank one and all who have provided their valuable advice, without their guidance this work would not have been a success, we have to thank who have helped us directly or indirectly since they have given us more than just guidance.
Our profound thanks to **Dr. S R Gengaje**, Head of the Department, Department of Electronics Engineering, Walchand Institute of Technology, Solapur, for his invaluable advice and constant encouragement to complete this work in a successful manner.
I would like to convey our sincere thanks to management and supportive staff of Walchand Institute of Technology, Solapur, for encouraging us to come up with this paper work

References

[1] R. N. Khushaba, S. Kodagoda, S. Lal, and G. Dissanayake, "Driver drowsiness classification using fuzzy wavelet-packet-based featureextraction algorithm," *IEEE Trans. Biomed. Eng.*, vol. 58, no. 1, pp. 121–131, Jan. 2011.

[2] Y. Liang, M. L. Reyes, and J. D. Lee, "Real-time detection of driver cognitive distraction using support vector machines," *IEEE Trans. Transp.Syst.*, vol. 8, no. 2, pp. 340–350, Jun. 2007.

[3] T. Moriyama, T. Kanade, J. Xiao, and J. F. Cohn, "Meticulously detailed eye region model and its application to analysis of facial features," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 28, no. 5, pp. 738–752, May 2006.

[4] J. C. Williams, "2012 Sleep in America poll: Transportation workers'sleep," Nat. Sleep Foundation, Arlington, VA, USA, Tech. Rep., 2012.

[5] T. A. Ranney, E. Mazzae, R. Garrot, and M. J. Goodman, "NHTSA driver distraction research: Past, present, and future," Nat. Highway Traffic Safety Admin., East Liberty, OH, USA, Tech. Rep., 2000.

[6] R. I. Hammoud, G. Witt, R. Dufour, A. Wilhelm, and T. Newman, "On driver eye closure recognition for commercial vehicles," *SAE Int. J. Commercial Veh.*, vol. 1, no. 1, pp. 454–463, Apr. 2009.

[7] Y. Dong, Z. Hu, K. Uchimura, and N. Murayama, "Driver inattention monitoring system for intelligent vehicles: A review," *IEEE Trans.Transp. Syst.*, vol. 12, no. 2, pp. 596–614, Jun. 2011.

[8] P. Jimenez, L. M. Bergasa, J. Nuevo, N. Hernandez, and I. G. Daza,"Gaze distraction system for the evaluation of driver distractions induced by IVIS," *IEEE Trans. Transp. Syst.*, vol. 13, no. 3, pp. 1167–1178,Sep. 2012.

[9] P. Viola and M. J. Jones, "Robust real-time face detection," *Int. J. Comput.Vis.*, vol. 57, no. 2, pp. 137–154, May 2004.

[10] S. Choi and D. Kim, "Robust head tracking using 3D ellipsoidal head model in particle filter," *Pattern Recognit.*, vol. 41, no. 9, pp. 2901–2915,Sep. 2008.

[11] Z. Zhou and X. Geng, "Projection functions for eye detection," *Pattern Recognit.*, vol. 37, no. 5, pp. 1049–1056, May 2004.

[12] B. D. Lucas and T. Kanade, "An iterative image registration technique with an application to stereo vision," in *Proc. Int. Joint Conf. Artif. Intell.*,1981, pp. 674–679.