Computer Vision Based Gaze Tracking for Accident Prevention

Mrs.P.Sheela Rani*, P.Subhashree**, N.Sankari Devi**
*Assistant Professor, Dept of Information Technology, Panimalar Institute of Technology
**Student, Dept of Information Technology, Panimalar Institute of Technology
rpsheelarani2014@gmail.com, subashree95@gmail.com, devishankari2@gmail.com

I. Abstract

Distracted driving is one of the main causes of vehicle collisions in India. Passively monitoring a driver's activities constitutes the basis of an automobile safety system that can potentially reduce the number of accidents by estimating the driver's focus of attention. Automotive vehicles are increasingly being equipped with accident avoidance and warning systems for avoiding the potential collision with an external object, such as another vehicle or a pedestrian. Upon detecting a potential factor, such systems typically initiate an action to avoid the collision and/or provide a warning to the vehicle operator. In this paper a complete accident avoidance system is proposed by determining the driver's behavior. As the main causes of vehicle accident were related to human factors, they could be labeled in one of the two main driver's distraction categories (Alcohol Consumption, Drowsiness and distracted vision). The aim of the proposed system is to help in analyzing the factors associated with driver's behavior for the development of accident avoidance systems. The main causes of the traffic accidents, discovered in the analysis of the driver behavior with the help of our system, will be used for the development of assistant devices and alarm systems that could help the driver to avoid risky situations. In this project we are implementing two image processing tool to get the facial geometry based eye region detection for eye closure identification, combined tracking and detection of vehicles. Frequencies of eye blinking and eye closure are used as the indication of sleepy and warning sign is then generated for recommendation; (b) outside an ego vehicle, road traffic is also analyzed.

Keywords --- Driver assistance system, accident avoidance, gaze tracking, head pose estimation.

II. Introduction

One of mankind's most major senses is its eyesight. The eye is different from the other body parts that make up the human's sensor array. It is different, as through the eyes, a lot can be read with regards to a human's expressions. For example, it can be assumed that a person's attention is generally focussed on where they are looking and because of this, tracking the eyes movement can be useful.

Fig.1. Camera fitted in the steering to monitor the driver distraction.

Our motivation for developing a technique of eye detection and gaze estimation is primarily derived from interest in a novel interface for laptop computers. This interface would allow a user to control the computer through simply looking at features of interest, and perhaps blinking to express action. Applications such as driver fatigue assessment, surveillance, advertising, sociological studies, etc. would benefit from this technology. For example: a car could safely come to a stop should the driver experience a microsleep. Law enforcement personnel could direct their attention to subjects prolongedly looking at an unexpected object. An advertising company could target ads based on what catches the user's attention. The crux of a novel gaze-driven interface is, not surprisingly, eye localization and gaze estimation. The technique explained in this paper estimates a subject's gaze direction after calibrating with an image of the subject looking straight ahead. The subject need not be a member of the set of images used to train the eye detection module.

A large amount of research has been conducted in the field of eye detection. It is agreed that face recognition is the easiest way to detect gaze or attention. To detect a person's state of awareness there are other easier methods than pupil tracking. For example, monitoring the heads outline is more convenient. This is because it is harder to make the eye gaze system non- intrusive. New techniques are also constantly being designed to compensate for the head movement in non-steady conditions. Edge Detection is used in conjunction with error and blink checking to derive an approximate corresponding x, y co- ordinate within the frame boundaries.

There is a potential in the automobile industry to have gaze detection for control loops within cars. As cars are becoming increasingly computer controlled, the feasibility of using gaze detection in cars is growing. Gaze detection is useful in letting the computer know where the driver's attention is focused. Using gaze detection is also beneficial in being able to tell if the driver is checking the mirrors or even falling asleep. Gaze detection in virtual reality allows the display image to become much more lifelike, as eye movement can compensate the image. Virtual Reality benefits with the integration of a vision tracking system. In accommodating for pupil motion, the experience is enhanced.

III. Related Work

Driver monitoring has been a long standing research problem in computer vision. Driver distraction is mainly taken as a big issue in current situation. Previously, eye blink sensors are utilized for the drowsiness detection. But efficiency wise it difficult to wear and also affects the eye. Also automatic braking systems have not been employed in the existing automobile systems. Fiction has had people operating things using sight rather than hands a long time ago, yet the technology hasn't developed into the mainstream as of yet.

FaceLAB is a stereo-based eye tracker that detects eye movement, head position and rotation, eyelid aperture and pupil size. FaceLAB uses a passive pair of stereo cameras mounted on the car dashboard. The system has been used in several driver assistance and inattention systems, such as However, stereo-based systems are too expensive to be installed in mass-produced cars and they require periodic re-calibration because vibrations cause the system calibration to drift over time.

Though several accident avoidance systems have already been deployed in the vehicles they were seem to be less efficient and costlier compared to our proposed system.

IV. System Description

There are four main modules: Image acquisition, facial feature detection, head pose estimation, gaze tracking.

i. Image Acquisition

The image acquisition module is based on a low-cost CCD camera placed on top of the steering wheel column. The CCD camera was placed over the steering wheel column for two reasons:

(1) It facilitates the estimation of gaze angles, such as pitch, which is relevant for detecting when the driver is texting on a phone (a major threat to safety).

(2) From a production point of view, it is convenient to integrate a CCD camera into the dashboard.

Fig.2. Capturing the image of the driver

For night time operation, the system requires an illumination source to provide a clear image of the driver's face. Moreover, the illumination system cannot impact the driver's vision. To this end, an IR illuminator was installed on the car dashboard.

ii. Facial Feature Detection

Facial features, such as eyes, eyebrows, nose and mouth, and their spatial arrangement, are important for the facial interpretation tasks based on face images, such as face recognition, facial expression analysis and face animation. Therefore, locating these facial features in a face image accurately is a crucial step for these tasks to perform well. However, in reality, the appearance of the facial features in the images varies significantly among different individuals. Even, for a specific person, the appearance of the facial features is easily affected by the lighting conditions, face orientations and facial expressions, etc. Therefore, accurate facial feature detection and tracking still remains a very challenging task, especially under different illuminations, face orientations and facial expressions, etc.

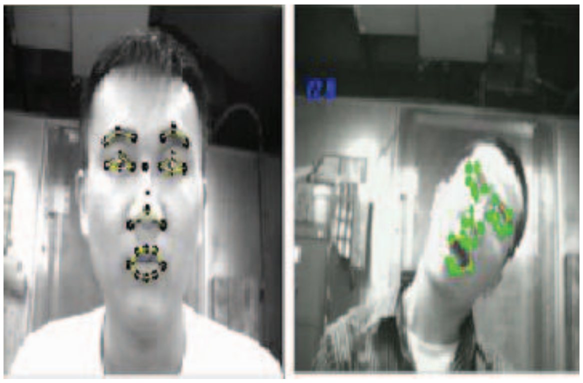


Fig.3. Real-time facial Feature tracking

In our research, we proposed an effective approach to detect and track facial features from the face images with different facial expressions under various face orientations in real time. The improvements in facial feature detection and tracking accuracy are resulted from:

(1) the use of pyramidal Gabor wavelets for efficient facial feature representation;

(2) dynamic and accurate model updating for each facial feature to eliminate any error accumulation;

(3) imposing the global geometry constraints to eliminate any geometrical violations.

By these combinations, the accuracy of the facial feature tracking reaches a practical acceptable level. Subsequently, the extracted spatio-temporal relationships among the facial features can be used to conduct the facial expression classification successfully.

iii. Head Pose Estimation

The appearance of the human face varies significantly as the facial expression changes, hence, the human face is not a rigid object. In addition, the human can move the head freely in front of the camera. Therefore, the motion of the face is the sum of two independent (or different) motions: rigid motion and nonrigid motion. The rigid motion consists of the global motion of the face describing the rotation and translation of the head, or face pose; while the nonrigid motion consists of the local motion of the face describing the contraction of facial muscles, or facial expression. When captured by the camera, both motions are mixed together to form a 2D face motion in the image plane.

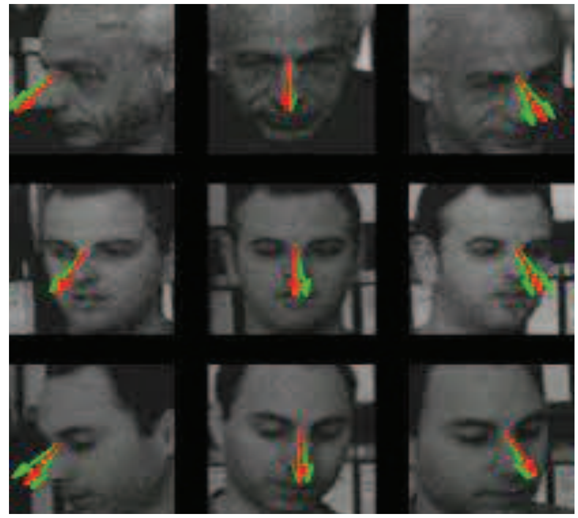


Fig.4. Estimating head pose of the driver

In our research, a novel technique is proposed to recover 3D face pose and facial expression simultaneously from a monocular video sequence in real time. First, facial features are detected and tracked robustly under various face orientations and significant facial expressions. Second, after modelling the coupling between face pose and facial expression in the 2D image as a nonlinear function, a normalized SVD (N-SVD) decomposition technique is proposed to recover the pose and expression parameters analytically. Subsequently, the solution obtained from the N-SVD technique is further refined via a nonlinear technique by imposing the orthonormality constraints on the pose parameters. Therefore, our proposed method can recover the face pose and facial expression from the face images robustly and accurately.

iv. Gaze Tracking

As one of the most salient features of human face, eyes play an important role in interpreting and understanding a person's desires, needs, and emotional states. Eye gaze is defined as the line of sight of a person. In particular, the eye-gazes, indicating where a person is looking, can reveal the person's focus of attention. Eye gaze tracking has been an active research topic for many decades because of its potential usages in various applications such as Human Computer Interaction (HCI), Virtual Reality, Eye Disease Diagnosis and Human Behavior Study, etc.

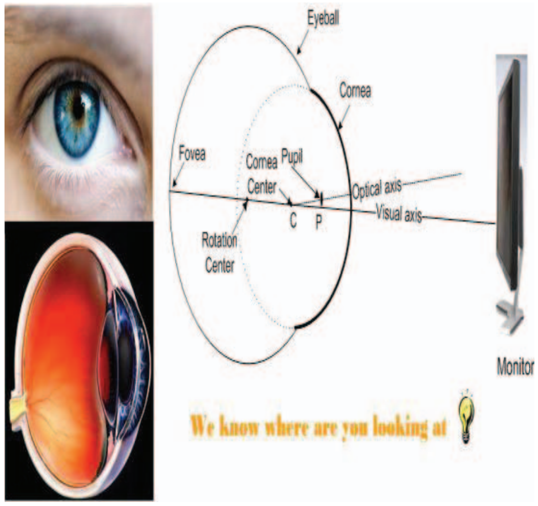


Fig.5. Tracking the line of sight

A good eye tracker is a prerequisite of eye gaze monitoring. Robust techniques for eye detection are of particular importance to eye-gaze tracking systems. Many eye tracking methods rely on intrusive techniques such as measuring the electric potential of the skin around the eyes or applying special contact lenses that facilitate eye tracking. This causes serious problems of user acceptance. We have developed a non-intrusive eye tracker that can detect and track a user's eyes in real-time as soon as the face appears in the view of the camera. Unlike most of the existing gaze tracking techniques, which often require a static head to work well and require a cumbersome calibration process for each person, our gaze tracker can perform robust and accurate gaze estimation under natural head movements with only one-time calibration.

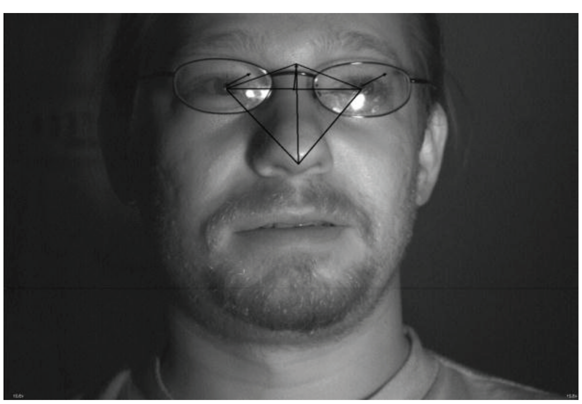


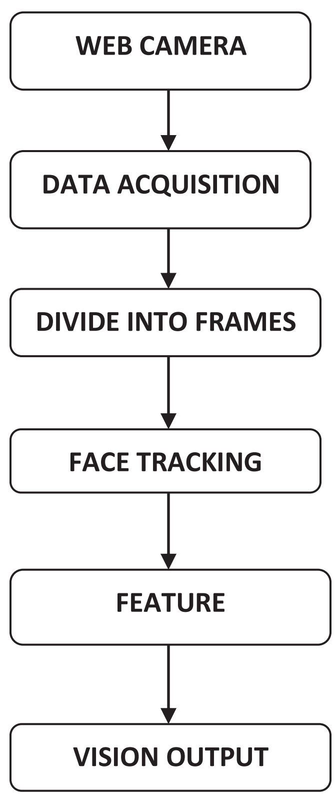
Fig.6. The detected triangle, eyes' centers and the nose bottom, together with the gaze line.

We propose an eye gaze tracking system under natural head movements. The system consists of one CCD camera and two mirrors. Based on geometric and linear algebra calculations, the mirrors rotate to follow head movements in order to keep the eyes within the view of the camera. We also introduce a hierarchical generalized regression neural networks (H-GRNN) scheme to map eye and mirror parameters to gaze, achieving a gaze estimation accuracy of 92% under head movements. The use of H-GRNN also eliminates the need for personal calibration for new subjects since H-GRNN can generalize. Preliminary experiments show our system is accurate and robust in gaze tracking under large head movements.

V. Proposed Methodology

The Algorithm is developed to detect the driver drowsiness and to alert the driver and also to intimate to hardware to stop the car. Frequencies of eye blinking and eye closure, distracted vision are used as the indication of sleepy and warning sign is then generated for recommendation. Outside an ego vehicle, road traffic is also analyzed. Ultrasonic sensor also employed to detect the distance between the front and the rear vehicles. If the distance seems to be very less, then an alert will be given to the driver to slow down the vehicle. If the speed is not reduced by the driver then an automatic braking will be activated to stop the vehicle.

The following flow chart gives the architectural block diagram of the system:



We utilize three algorithms presented to do this. The first is called Longest Line Scanning (LLS).

i. Longest Line Scanning (LLS)

Longest line scanning algorithm is used for iris centre detection. Eye position is not sufficient enough for tracking the eye accurately. The direction of visual attention of the eyes must be measured more precisely from the eye image. The reason for choosing iris is because sclera is light and the iris boundary is dark, therefore iris boundary can easily be optically detected and tracked.

LLS is a algorithm based on the following useful property: the center of an ellipse lies on the center of the longest horizontal line inside the boundary of the ellipse. Though we do fit a circle to the iris, it will appear as an ellipse in the image if the subject is not looking directly into the camera. So we simply scan the iris from top to bottom, each time plotting a horizontal line between the edges. The longest line found will give us a candidate for the center of the iris. See Figure 7. below.

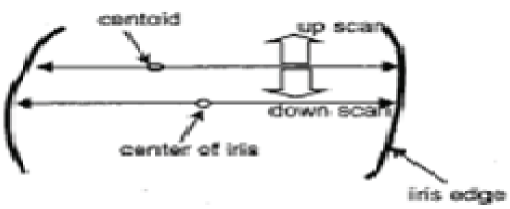


Fig.7. Principle of LLS

We take the candidate center from the LLS algorithm, and use it as an input to our next algorithm, Occluded Circular Edge Matching.

ii. Occluded Circular Edge Matching (OCEM)

Unfortunately, the LLS algorithm may not be ideal due to the presence of eyelids. The longest line may actually be hidden by the subject's eyelids. Although the LLS method detects the center of the iris, the following problems arise: intra-iris noise and rough iris edge. Obviously, if the edge of the iris is noisy, the horizontal line drawn in LLS will not be easily defined.

OCEM takes both the candidate center of the iris and the edge image as inputs, and approximates the shape of the iris with a circle. The center of that circle is our chosen center for the iris.

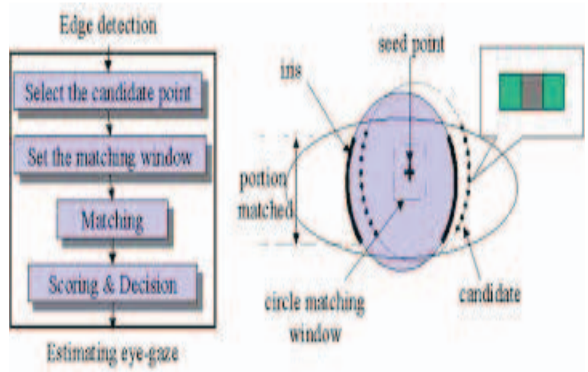


Fig.8. Principle of OCEM

iii. Blink Detection

Blink detection algorithm separates the voluntary and the involuntary blinks and detects single voluntary blinks or sequence of blinks. (Applications, like fatigue monitoring, human–computer interfacing and lie detection.) There are two types of blinks:

1. Short blinks which are spontaneous and are involuntary blinks and are ignored by the blink detection algorithm. (shorter than 200 ms)
2. Long blinks which are user specified blinks or voluntary blinks. This blinks are considered for the blink detection algorithm.(longer than 200 ms)

Eye-Blink detection algorithm consists of 3 major steps:

1. Eye region extraction
2. Eye-blink detection
3. Eye-blink classification

Fig.9. Eye blink detection

Correlation coefficient is the similarity between the current eye images with respect to a saved eye image. Correlation Coefficient can be used to measure the openness of the eye. If the coefficient for any two consecutive frames is lower than the predefined threshold value TL or greater than predefined threshold value TH, then eye blink is detected. If the blink is detected between 250 ms and 2s then such blink is considered as voluntary control blink otherwise involuntary eye-blink detection.

There are various types of blinking explained as follows:

Spontaneous blink – It can be called as the natural blinking that occurs without the person realizing about any such eye movements. Spontaneous blinking helps in clearing the tears away that are constantly produced by the human eye. It is the most common type of blink and it is so short that your brain is unable to notice them, if you are not focusing on it.

Reflexive blink – The kind of blink that occurs in response to an event or a situation is known as the reflexive blinking. Like a loud sound or an object approaching your eye. It helps in protecting your eye from external threats as the eye closes when a stimulus occurs. These are automatic reflexes that are intractable.

Voluntary blink – Voluntary blinks are caused intentionally by the person. They last for a slightly longer time than the spontaneous blink and can occur for different reasons. For example, if someone‘s eyes start to feel dry or they feel uncomfortable, they might blink to wet and comfort their eyes.

VI. Conclusion

This paper describes a real-time gaze tracking system using the video from a monocular camera installed on steering wheel column. The proposed system is able to detect at day and night, and under a wide range of driver's characteristics. The system does not require specific calibration or manual initialization. More importantly, no major re-calibration is necessary if the camera position is changed. This is due to the explicit use of 3D geometric reasoning. Hence, the installation of the system in different car models does not require any additional theoretical development. Thus, when positioning the camera, it is beneficial to the detection accuracy to maximize the degree of variation between the open and closed eye images of the user. Finally, with respect to the clinical environment, this system provides an unobtrusive alternative. Our experiments showed that our head pose estimation algorithm is robust to extreme facial deformations. The system achieved an accuracy above 90 % for all of the scenarios evaluated.

VII. References

[1] H. Yin, P. Fu, and S. Meng, "An efficient face detection method in color images," in Knowledge-Based Intelligent Information and Engineering Systems, ser. Lecture Notes in Computer Science, R. Khosla, R. Howlett, and L. Jain, Eds. Springer Berlin Heidelberg, 2005, vol. 3684,pp.875–880. [Online].Available:http://dx.doi.org/10.1007/11554028 122

[2] T. Ahonen, A. Hadid, and M. Pietikainen, "Face description with local binary patterns: Application to face recognition," Pattern Analysis and Machine Intelligence, IEEE Transactions on, vol. 28, no. 12, pp. 2037–2041, Dec 2006.

[3] RAINER STIEFELHAGEN, Gaze Tracking for Multimodal Human-Computer Interaction, Diplomarbeit, Universitiit Karlsruhe, September, 1996.

[4] C. Ware and H.T. Mikaelian, __An Evaluation of an Eye Tracker as a Device for Computer Input,‘‘ Proc. ACM CHI+GI‘87 Human Factors in Computing Systems Conference pp. 183-188 (1987).

[5] Krolak, A., Strumillo, P.: Eye-blink controlled human-computer interface for the disabled. Advances Intell. Soft Comput. 60, 133-144(2009).

[6] Stephen Bialkowski, Zeb Johnson, Kathryn Morgan, Xiaojun Qi, and Donald H. Cooley, "A NonIntrusive Approach to Gaze Estimation," Proc. of the 8th International Conference on Computer Vision, Pattern Recognition, and Image Processing (CVPRIP'07), July 18-24, Salt Lake City, UT, USA, 2007

[7] David Bäck, "Neural Network Gaze Tracking using Web Camera". [Online].Available:www.imt.liu.se/mi/Publications/pdfs/back05.pdf