Eye-Gaze and Augmented Reality Framework for Driver Assistance

Muhammad Shahid[1,3], Tabassam Nawaz[2], Hafiz Adnan Habib[1]

[1]Computer Engineering Department, University of Engineering and Technology, Taxila, Pakistan
[2]Software Engineering Department, University of Engineering and Technology, Taxila, Pakistan
[3]Pakistan Council for Science and Technology, Ministry of Science and Technology, Islamabad, Pakistan
shahid_pcst@yahoo.com, tabassam.nawaz@uettaxila.edu.pk, adnan.habib@uettaxila.edu.pk

Abstract: Driver inattention or cognitive overload is among the leading contributor to road accidents. Driver behavioral cues could be employed in advanced driver assistance systems (ADASs) to alleviate such accidents. Facial and eye gaze are among the most import behavioral cues to reflect driver active state within the vehicle. With the intent to improve the human error related accidental controls in mind, a simple and smart ADAS is proposed that could assist the driver based on continues monitoring of facial and eye gaze information. The system measures driver eye gaze within near frontal facial positions and project future position of vehicle on windscreen based on the vehicle parameters and driver's active eye gaze estimates, assuming the movement of vehicle in a straight line. The projections adapt to shift in driver's perspective. The projection over the wind screen give visualization as if the lines are physically drawn over road according to the width of the vehicle. The system also warns the driver when there is constant shift in head or eye gaze from normal forward facing positions, beyond some threshold period. The system is based on live video input from low-end webcam. The system will reinforce driver ability in effective estimation of future positions of vehicle and leverage better control while driving. In addition it can also be employed to assist novice drivers to keep track of the width of vehicle during the training sessions.
[Muhammad Shahid, Tabassam Nawaz, Hafiz Adnan Habib. **Eye-Gaze and Augmented Reality Framework for Driver Assistance**. *Life Sci J* 2013;10(3):1571-1578] (ISSN:1097-8135). http://www.lifesciencesite.com. 236

Key words: Advanced Driver Assistance System, Vehicle Future Position Projection, Road Centre Inattention Detection , Eye Tracking, Gaze Estimation, Augmented Reality.

1.INTRODUCTION

According to recent statistical estimates, road accident is among the top ten leading causes of death and attributed to approximately 1.3 million deaths annually (WHO Report, 2012). Human errors including driver inattention or cognitive overload lead to misjudgments and delays in recognition of environment or future course of driving and is considered among the leading contributing factor to road accidents. Even though developments in passive safety technologies such as seatbelts, airbags and crumple zones etc. have partially reduced damages and improved safety during accidents, nevertheless further progress in these technologies is limited due to their inherited limitations. Active safety technologies based on advanced driver assistance systems (ADASs) offer new opportunities to improve driving experience by actively guiding driver during the course of driving. ADASs continually monitor vehicle environment and assist driver in realizing and responding to potential challenges in a timely manner. Unlike traditional passive safety technologies which are primarily effective in post-accidental condition, ADASs help avoid accidents. ADASs is an active research domain having applications ranging from general purpose driver information and warning systems to more advanced intervening and fully automated systems.

Driver inattention or cognitive overload primarily affects the driver's perception of future events of the vehicle. In addition, change in drivers' perspective vis-à-vis windscreen within the vehicle also affects driver's perception of vehicle future position. Figure 1(a)–(g) illustrate how driver's perception about the appearance of future course of vehicle change with the change in driver's position within the vehicle from extreme right to left. Since vehicle future path is not actually drawn on the road, such sort of head displacements from normal forward facing position lead to misjudgment in future course of the vehicle and add vulnerabilities to accidents.

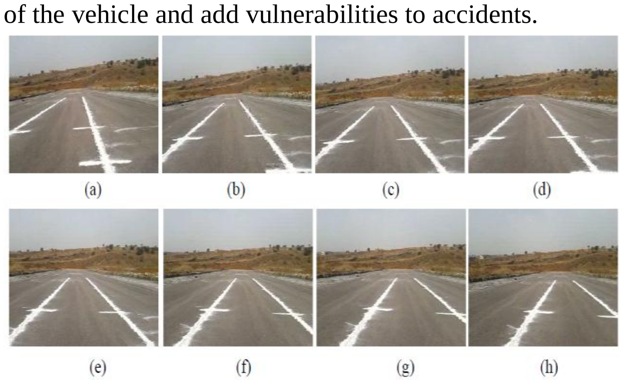

Figure 1. Change in appearance of future path of vehicle based on driver's in-vehicle head shifts from extreme right to left.

Here we present an advanced driver information and assistance system that augments driver ability to maintain effective longitudinal control. In addition, the system also supervises driver level of attentiveness to mitigate the prospect of accidents by timely warning during moments of inattention or drowsiness. Facial and eye gaze information are used as cues to determine driver's level of attentiveness. The system alerts the driver in situations where the driver is observed inattentive beyond some threshold period. The system also projects future position of the vehicle over windscreen based on the driver active eye gaze state and vehicle parameters. Vehicle parameter remains constant while drivers' eye gaze estimates are updated from real-time video input from a video camera.

The rest of the paper is organized as under. Section 2 presents a brief review of the related research and section 3 describes the proposed system. Experimental setup details are discussed in section 4. In section 5 we report some results of our preliminary study on application of eye gaze in driver monitoring and assistance system. Section 6 concludes the paper.

2.EXISING METHOD

2.1 ADAS and Computer Vision

The use of computer vision as an instrument in the ADASs is not particularly a new concept. Computer vision based techniques have been explored in the recent past in a number of ADASs that are geared toward improving driving experiences and to mitigate the causes of accidents such as lane departure warning systems, automatic cruise control systems, pedestrian protection systems, automatic parking, traffic signboard recognition and driver's drowsiness detection systems etc. These systems are based on algorithmic designs that learn the vehicle inside and outside environments based on some sensors input and communicate or assist the driver about approaching events or of potential hazards. Camera is the most widely employed sensor in computer vision based systems though some applications employed use of other sensors including radar, lidar, ultrasonic and laser scanner etc.

2.2 Augmented Reality

Augmented reality (AR) enrich real world scene with synthetic graphical, textual and audio information to improve its meaningfulness to the observer (Krevelen et. al, 2010). AR is largely being the focus of attention of the mobile industry where the technology is embedded in new communication devices giving rise to use of the platform for smart and innovative applications. Like many other fields, the technology is also finding innovative applications in automotive industry where new model are being equipped with AR-enabled head-up displays of information earlier found on the vehicle dashboard such as speedometer, fuel gauge and navigational information etc. A great deal of research is underway on development of enhanced AR-enabled ADASs using array of sensors and cameras. These systems will detects all sort of vehicle inside and outside environment such as driver head posture and eye gaze movement, road congestion, on-road obstacles, pedestrian, road edge especially under fog conditions etc. and assist driver in any associated hazards.

2.3 Face Detection

Face detection is considered enabling technology having immense application domains such as surveillance, smart card applications, general identity verification and access controls, human-computer interface enabled adaptive multimedia environments etc. Earlier research in face detection was only restricted to analysis of simple facial characteristics such as shape, color, texture and motion. Most recently reported techniques include AdaBoost classifier (Schneiderman et. al., 1998)(Viola et. al., 2004), Neural Networks (Rowley et. al, 1998), Support Vector Machine (Shih et. al., 2004), and the Bayes classifier (Pham et. al., 2001). Efficiency of eye gaze estimation application is highly influenced by the performance of face detection algorithm, typically employed to narrow down region of interest for performance optimization and pose estimation. Adaboost proposed by voila and jones (Viola et. al., 2004) is among the most successful and widely used face detection algorithm having real world applications in digital cameras and image organization software etc.

2.4 Eye Socket Detection

Eye gaze estimation and tracking applications are well reliant on effectiveness of eye sockets detection. Geometric and photometric appearances of eye are significant cues for eye sockets detection (Hansen et. al, 2010). Several vision based techniques have been explored found on such characteristics and are largely categorized as shape-based, appearance-based and feature-based. Shape-based techniques (Xie et. al., 1994)(Lam et. al., 1996) involve defining eye template based on the structural appearance of eye and then searching the template within the target image for maximum correlation. These techniques are although accurate but computationally quite expensive. Appearance-based techniques (Huang et. al., 2000) adopt holistic approach based on photometric appearance as characterized by intensity distribution of eyes and their surrounding region. These techniques use classifiers such as support vector machine (Shih et. al., 2004) and neural networks (Rowley et. al., 1998) for eye detection. The classifiers are trained over a large dataset of eyes of different subjects under varied scale, orientation and lighting conditions. Feature-based techniques (Feng et. al., 2001) take into consideration a set of features that are unique and mostly related to eye region such as dark/bright pupil, iris, eye corners and corneal reflections etc.

2.5 Gaze Estimation

Technically, gaze estimation is determination of human line of sight and its translation on a two dimensional surface. The point of intersection of line of sight on the surface is regarded as point of regard. Earlier techniques of gaze estimation were quite intrusive in nature using different type of body attachments. These techniques were only suited for experimental or laboratory arrangements with no or limited acceptability in daily usage. Rapid advancements in computing and imaging capabilities have given rise to new research direction involving nonintrusive vision based techniques for gaze estimation. Pupil center corneal reflection (Zhu et. al., 2004) is among the most widely used nonintrusive technique that employs infrared light source to produce glint on the corneal surface. Distance vector between the pupil center and the glint is employed to estimate the gaze direction. Cross-ratio method (Yoo et. al., 2005) used multiple infrared light sources fixed to four corners of the screen. The method employed position of the pupil center and the four glints produced on the corneal surface to estimate the gaze direction. One-circle method (Hansen et. al., 2005) uses elliptical iris directions under normal eye movement and natural lighting conditions to estimate gaze direction.

A typical process of eye-gaze detection involving face detection, eye sockets detection and gaze estimation is depicted in Figure 2.

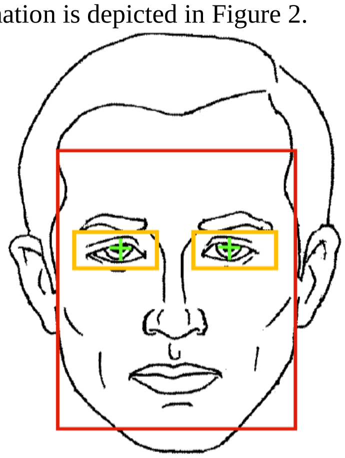

Figure 2. Face detection, eye socket detection and gaze estimation.

3.MATERIALS AND METHOD

3.1 System Overview

The system proposed is based on the consideration that driver cannot maintain constant level of attentiveness during the course of driving. Driver drowsiness, sleepiness or fatigue all affects the driving safety. In addition, constant head-shift due to several on– and off–windscreen distractions such as outside people, objects and events and in-vehicle adjustment to devices, talking to partners and listening to phone calls etc. also adds to these vulnerabilities. During such states, drivers not only lose their prime focus of attention but also change in driver’s perspective vis-à-vis windscreen, as a result of these shifts, may also lead to misperception or misjudgment of future path of the vehicle.

The main goal of the proposed system is to:

1. improve driving safety by supervising driver level of attentiveness and alerting them where observed possible sign of distraction or drowsiness, and to
2. augment driver ability in judgment of future occupancy of vehicle during change in driver’s perspective vis-à-vis windscreen.

The key idea is very trivial. Driver’s perception of future path of the vehicle is built on the projections of the road reaching drivers visual senses through the windscreen. A system that could constantly measure those visual projections build in the drivers mind and translate or augment them over the windscreen based on the prevalent physical state of driver (facial and eye-gaze estimate) within the vehicle, will give the driver perception that the projections are actually marked on the road according to width of the vehicle. This will assist driver in effective estimate of future position of the vehicle and maintain greater longitudinal control.

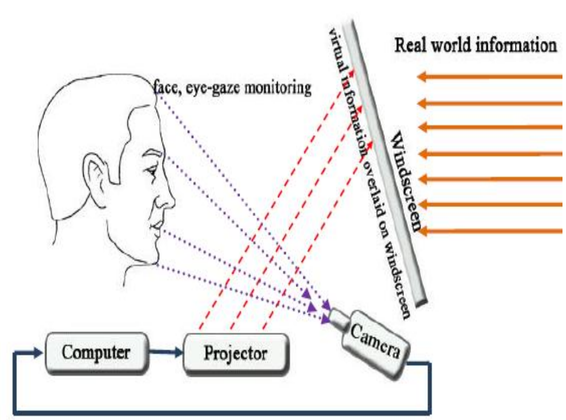

Figure 3. Eye-gaze driven virtual information projection over windscreen.

Figure 3 shows schematic diagram of the proposed system. The proposed system is based on the following assumptions.

1. The vehicle is equipment with a personnel computer, an infrared illumination source and camera connected via USB port.
2. The camera is mounted on the dashboard pointing straight to the driver under neutral forward facing posture.
3. The motion of the vehicle is on a straight path.

The system, to be used effectively during both day and nighttime, an additional infrared illumination source is used. The infrared illumination source is turned on during nighttime or instances of low ambient lighting conditions to effectively detect facial features while ambient source is used during the broad day light driving.

Figure 4 outlines the design flowchart. C1 and C2 represent counters; C1 counts number of successive frames driver face is off the camera field of view; C2 counts successive frames driver eye-gaze is off the neutral forward facing position. The system initializes with the driver moving in and starts the vehicle.

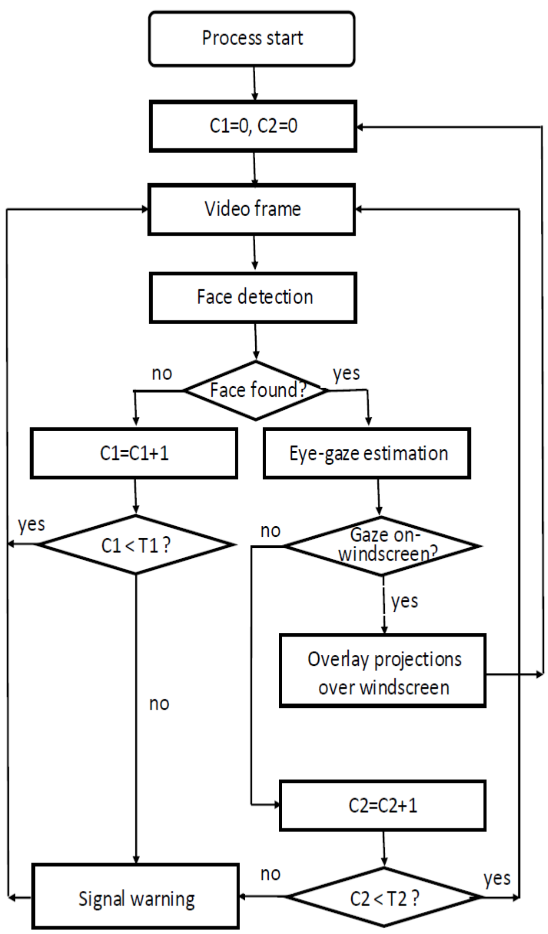

Figure 4. Flowchart of the proposed eye-gaze and augmented reality framework for driver assistance.

The system begins with driver face detection. Successful face detection is followed by eye socket detection within the cropped facial region and gaze estimation. The estimated gaze position is translated by overlaying projection over the windscreen, representing zoom in version of vehicle future path projections in the driver visual senses. The measurement of projections i.e. in-between widths and respective angles are based on actual width of the vehicle and the driver's eye gaze position as well as distance from the windscreen. The width of the vehicle is constant while change in eye gaze position affects the projections. If the driver face is not detected for T1 number of successive frame, the system signals warning to the driver of possible instance of distraction or drowsiness. If eye gaze is off the normal forward looking position for more than T2 successive frames, the driver is intimated of possible distraction as well.

3.2 Face Detection

The objective of face detection is to determine whether or not driver is in neutral forward looking position. In events driver face is not detected within the camera field of view, means significant shift from normal driving position. The system initiates counter C1 and non-detection of face for successive frame until threshold T1 will invoke warning signal to the driver of possible vulnerability due to continues inattention. Once the driver face is detected again, the counter C1 is reset. Face detection also empower to narrow down region of interest for subsequent processing.

Face detection is carried out using the methodology proposed by voila and jones (Voila et. al, 2004) with some minor adjustment. The implementation involves a variant of Adaboost learning algorithm for selection of small number of critically important features. The rectangular shaped Haar-like features are computed over special pre-computed image representation, known as integral image. The rectangular features represent basic image structures like edges and bars etc., act as classifier and are pooled in form of cascade to yield extremely potent classifier. While considering the distance factor of the driver to the camera and to ensure optimal performance and quality, a minimum face size of 64x64 is used for detection.

3.3 Eye Sockets Detection

The face region is cropped using the principal face coordinates. To boost the efficiency of the algorithm, eye detection is applied to only reduce space determined based on geometric position of the eye sockets within the face region and eye coordinates from the preceding frame. Eye detection is performed using the approach mentioned in (Voila et. al., 2004).

3.4 Eye Features Detection

Inner eye corners, pupil/iris centers are the principal features employed for accurate eye gaze determination. For pupil/iris detection, method proposed in (Shahid et. al., 2010)(Zhu et. al., 2002) is used. Pupil has the lowest intensity values is the darkest region within the eye sockets. Eye brows and eye lashes are other features having similar intensity values. Pupil prospect point is first determined using lowest average neighborhood intensity values. The most probable location of pupil prospect point is within the pupil boundary. Outward horizontal projections from pupil prospect point are drawn individually to labeled edge map of each eye sockets. The projections intersecting the first edge on either side having identical label is the pupil. Pupil center coordinates are obtained by measuring the center of longest horizontal and vertical lines of the pupil edge. The horizontal projections are further extended from the edge of the pupil until finding edges having maximum gradient magnitude on either side which is the iris-sclera boundary. The gradient magnitudes of the two projections are correlated as sanity check during the course of projections. Lines between the detected iris-sclera edges are analyzed to find the longest line. The center of the longest line is inferred as the iris center.

Inner eye corners are the points where the lower and upper eyelids intersect. We adapted Harris corner detector (Harris et. al., 1988) to detect eye corners. Harris algorithm provides an efficient mechanism to extract feature points. Applying Harris corner detector fetch many corner points when applied to complete facial space. Since we are interested in inner eye corner, we apply the algorithm separately to each eye socket space. From experimental findings, it is observed that in case of Harris corner detection we need to select the specific points we are interested in the after detection. Generally, the detected points distribute intricately. We pick the farthest left and right points as the inner eye corner points for right and left eyes respectfully. The vertical position or height of the eye corners must satisfy the following criteria: top

$$P_{highest} < C_{vertical} < P_{lowest}$$

Where $P_{highest}$ and P_{lowest} denote the highest and lowest pixel positions of the detected pupil edge. $C_{vertical}$ denotes the vertical pixel position of the eye corner. Geometric filters (Corcoran et. al., 2012) are also applied to verify the validity of features determined. Eye sockets coordinates are saved. Subspace around eye sockets are used for eye detection in successive frame.

3.5 Gaze Estimation

Pupil center corneal reflection (PCCR) vector is among the most widely used technique in eye gaze application. The technique uses infrared light emitting diodes (IR-LEDs) to illuminate eye for corneal reflection. However, interference due to ambient infrared illumination during daytime outdoor environment makes it very difficult to detect correct and distinct corneal reflection.

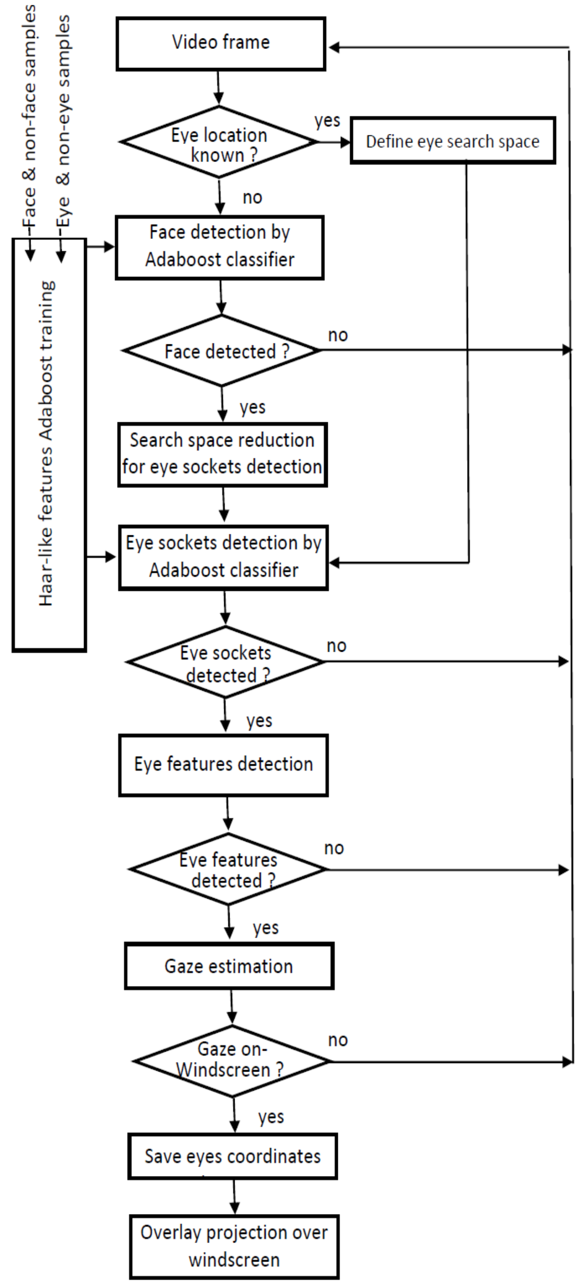

Figure 5. Eye Gaze Tracking Algorithm.

Eye corners also maintain stable state during change in eye gaze postures and are considered good substitute for corneal reflection, as an anchor/reference point for eye gaze vector. For such reasons, iris center eye corner (IC-EC) vector as an alternate to PCCR is employed for eye gaze estimation. The technique is particularly suitable during outdoor broad day time environment. In order to make the application equally effective during nighttime when there is no ambient illumination, an additional infrared illumination source is used.

Figure 5 shows the proposed eye gaze tracking algorithm. For each input frame where eye gaze coordinates from preceding frame are known, eye sockets are searched only within a small sub space i.e. +10% of the detected eye sockets space in the preceding frame, otherwise complete frame is searched for face detection (Moshnyaga, 2012). Face detection is followed by eye socket detection within facial sub space. Adaboost classifiers are used for both face and eye socket detection. Gaze is estimated using eye feature vector. Eye socket coordinates are updated for fast eye socket detections in subsequent frames.

Real-time output of the proposed system is shown in Figure 6. The blue rectangle is the detected face region. The red rectangle represents the reduced search space for eye detection estimated based on eye socket geometric location within facial region as well as eye sockets coordinate from earlier frame. Detected eye sockets are marked with green rectangles. Green dots denote iris centers.

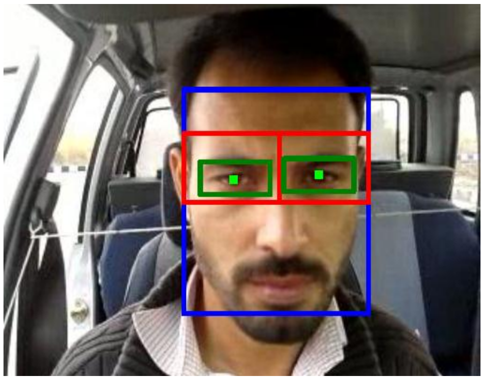

Figure 6. Face, reduced region of interest, with eyes sockets detected by AdaBoost classifier.

4. EXPERIMENTAL SETUP

Driver head movement along horizontal and vertical axis from stable forward facing driving position is quite normal during the course of driving. Horizontal movement are mainly invoked by the vehicle outdoor environments e.g. viewing side mirrors, obstacle on road etc. while vertical movement are primarily result from in-vehicle activities like interacting with devices. These head shifts, change the driver's perspective vis-a-vis windshield which result in change in appearance or projection of future path of the vehicle. Figure 7 shows the different head position that the driver take during these movements. To translate such projection from driver's visual senses to windscreen, the dimensions of the projections were determined through experimental setup.

Figure 7. Typical head movement within vehicle. The driver (e) looks on the windscreen, normal position, (a)–(d) distracted due to outside event, not looking on the windscreen, (f)–(h) talks to person on left.

Two parallel line, each 40 feet in length and 4.6 feet part, were drawn on the surface facing the vehicle as shown in Figure 8 (a)–(b). The lines represent the straight path vehicle will follow and width between the lines represents the actual width of the vehicle. The windscreen of the vehicle was divided up into 24 cell grid consisting of 4 rows and 6 columns. The cells were uniformly spaced. While on neutral forward facing driving position, visual appearance of lines as perceived by the driver were marked on the windscreen. The dimension of the lines including their respective lengths, internal angles and distance between them were recorded. Associated images of the driver were also captured and eye gaze vector calculated. Measurements were recorded while driver gaze at the center of each windscreen cell. The measurements serve for eye gaze adaptive projection over the windscreen.

Figure 8. Parallel lines drawn on the surface facing vehicle.

Suzuki Mehran having dimension of 3300mm x 1405mm x 1410mm (L x W x H) was used as platform for the measurement. The system was primarily developed and tested on a Microsoft Windows 7 PC with an Intel i7 Core™ CPU (Q740) with 1.73 GHz processor and 4.0 GB RAM. Video input was obtained using Microsoft LifeCam VX-5000 for broad daytime operations. For nighttime operation, a different Microsoft LifeCam VX-5000 webcam which was specially modified for infrared application was used. Image processing was performed using Emgucv and Intel OpenCV library.

5. EXPERIMENTAL RESULTS

The proposed system was tested on two users both under day and nighttime conditions. Certain constraints were considered during real-time evaluation of the system such as placement of the camera vis-à-vis driver (in between distance, height and angle), non-occlusion of eyes and illumination conditions etc. The performance of the system was evaluated on two conditions 1) distraction detection and 2) eye gaze tracking for visual projection. Distraction detection does not involve any complex analysis and only parameter to decide distraction was shift from neutral forward facing position. Initially, the users were asked to remain in forward facing position, enabling the system to detect eye gaze feature vector under neutral forward facing position. The detected feature vector was used as benchmark for distraction detection during subsequent frames. For purpose of evaluation of detection accuracy, every instance of deviation from neutral position was considered distraction without considering the threshold interval. The mean accuracy of distraction detection was found to be 96.6%.

For evaluation of eye gaze tracking for visual projections, user calibration was initially performed. The users were asked to gaze at the centers of each grid cell for predefined interval in specified sequence. Eye gaze feature vectors were recorded. During the second phase, the users were allowed to freely gaze through the windscreen. The detected eye feature vector was correlated to prerecorded data to determine the gaze direction and draw associated projection. Mean accuracy of 93.2% was recorded during this phase.

Figure 9 shows some sample output of the proposed system in tracking driver's eye gaze and drawing associated projection. (a)–(c) shows successful estimation of driver's eye gaze from different face positions. (d)–(e) shows the respective future position projection over windscreen based on respective eye gaze estimates.

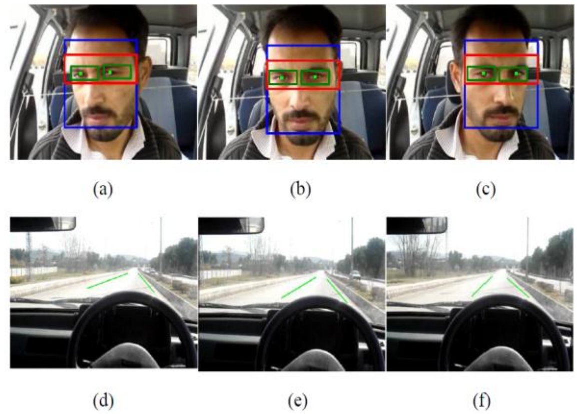

Figure 9. Real-time projections based on driver eye gaze estimate. (a)–(c) shows driver different face position with successful eye gaze estimation. (d)–(f) respective projections on windscreen.

6. CONCLUSION

A real-time eye gaze tracking technology as elaborated in this paper has a wide potential to be employed in automobile industry as real-time driver assistance system that could not only predict future path of the vehicle for effective judgment and longitudinal control but also alerts driver in critical situation like extended inattentions or drowsiness. In addition, the system can also serve as an effective tool to assist novice drivers to keep track of the width of vehicle during training sessions. The unique aspect of the system is that it is based on non-invasive technique. The system employs general purposed camera. The camera passively records the data without any hind of user involvement, thus giving a comfortable feeling to the driver.

While our initial studies have focused only on in-vehicle horizontal and vertical head shifts within narrow head movement, it is believed that by using other techniques such as depth perception we can adapt the projections for user movement on forward-backward lines as well.

REFERENCES

[1] Road Traffic Injuries. Technical Report: World Health Organization 2012.
[2] Van Krevelen D W F, Poelman R A survey of augmented reality technologies, application and limitations. The International Journal of Virtual Reality. 2010; 9:1-20.
[3] Schneiderman H, Kanade T. Probabilistic modeling of local appearance and spatial relationships for object recognition. Proc. Computer Vision and Pattern Recognition, 1998; pp. 45-51.

[4] Viola P, Jones M J. Robust real-time face detection. International Journal of Computer Vision. 2004; 57:137-154.
[5] Rowley H A, Baluja S, and Kanade T. Neural network-based face detection. IEEE Transaction on Pattern Analysis and Machine Intelligence. 1998; 20:23-38.
[6] Shih P, Liu C. Face detection using discriminating feature analysis and support vector machine in video. Proc. International Conference on Pattern Recognition, 2004; 2: 407-410.
[7] Pham T V, Worring M, Smeulders A W M. Face Detection by Aggregated Bayesian Network Classifiers. Proc. International Workshop on Machine Learning and Data Mining in Pattern Recognition. 2001; pp. 249-262.
[8] Hansen D W, Ji Q. In the eye of beholder: a survey of models for eyes and gaze. IEEE Transaction on Pattern Analysis and Machine Intelligence. 2010; 32:478-500.
[9] Xie X, Sudhakar R, Zhuang H. On improving eye feature extraction using deformable templates. pattern recognition. 1994;27:791–799.
[10] Lam K M, Yan H. Locating and extracting the eye in human face images. Pattern Recognition. 1996;29: 771–779.
[11] Huang W, Mariani R Face Detection and Precise Eyes Location. Proc. International Conference on Pattern Recognition. 2000.
[12] Feng G C, Yuen P C. Multi-cues eye detection on gray intensity image. Pattern Recognition. 2001; 34:1033–1046.
[13] Zhu Z, Ji Q. Eye and gaze tracking for interactive graphic display. Machine Vision and Applications. 2004; 15:139–148.
[14] Yoo D H, Chung M J. A novel non-intrusive eye gaze estimation using cross-ratio under large head motion. Computer Vision and Image Understanding. 2005; 98:25–51.
[15] Hansen D W, and Pece A E C. Eye tracking in the wild. Computer Vision and Image Understanding. 2005; 98:182–210.
[16] Shahid M, Nawaz T. Iris detection based on pupil prospect point and horizontal projections. Proc. SPIE International Conference on Digital Image Processing, Singapore, 2010; 7546.
[17] Harris C and Stephens M J. A combined corner and edge detection. Proc. 4th Alvey Vision Conference. 1988; pp. 147-151.
[18] Corcoran P M, and Bigioi P. Real-time eye gaze tracking for gaming design and consumer electronics systems. IEEE Transaction on Consumer Electronics. 2012; 58:347-355.
[19] Moshnyaga V. G. A camera-driven power management of computer display. IEEE Transaction on Consumer Electronics 2012. 22:1542-1553.
[20] Zhu J and Yang J. Subpixel eye gaze tracking. Proc. 5th IEEE International Conference on Automatic Face and Gesture Recognition 2002; pp. 131-137.
[21] Yang C. Eye feature extraction based on gray information and harris corner detection. Proc. International E-Conference on Computer Science 2006; pp. 285-288.

8/21/2013