DEVELOPMENT OF AN EMPIRICAL MODEL TO ASSESS ATTENTION LEVEL AND CONTROL DRIVER'S ATTENTION

Munira Akter Lata, Md Mahmudur Rahman, Mohammad Abu Yousuf,
Md. Ibrahim Al Ananya, Devzani Roy and Alvi Mahadi

Institute of Information Technology, Jahangirnagar University, Dhaka, Bangladesh

ABSTRACT

Any kind of vehicle driving is one of the most challenging tasks in this world requiring simultaneous accomplishment of numerous sensory, cognitive, physical and psychomotor skills. There are various number of factors are involved in automobile crash such as driver skill, behaviour and impairment due to drugs, road design, vehicle design, speed of operation, road environment, notably speeding and street racing. This study focuses a vision based framework to monitor driver's attention level in real time by using Microsoft Kinect for Windows sensor V2. Additionally, the framework generates an awareness signal to the driver in case of low attention. The effectiveness of the system demonstrates through board experiments in case of hostile light conditions also. Experimental result illustrates the quite well functionality of the framework with 11 participants and measures the attention level of participants with equitable precision.

KEYWORDS

Kinect, Software Development Kit (SDK), Eye State (ES), Face Angle (FA), Mouth Motion (MM), Attention with Eye State (AE), Attention with Face Angle (AF) and Attention with Mouth Motion (AM).

1. INTRODUCTION

Day by day the increasing number of automobile crashes in every year becomes an alarming issue for each nation. Traffic collision, which is caused by human or mechanical failure, carelessness or a combination of many other factors, should be dealt with the principles of anticipation, attention and compensation.

Hence, traffic safety enhancement turn into a high priority task for different government agencies over the world such as National Transportation Safety Administration (NTSA) in USA and Observatoire National Interministeriel de la SecuriteRoutiere (ONISR) in France. Researchers recommended that 57% of crashes were due solely to driver factors, 27% to combined roadway and driver factors, 6% to combined vehicle and driver factors, 3% solely to roadway factors, 3% to combined roadway, driver and vehicle factors, 2% solely to vehicle factors and 1% to combined roadway and vehicle factors [1]. In fact, inattentive driving is responsible for 20-30% of road deaths and this statistic reaches 40-50% in particular crash types, such as fatal single vehicle semitrailer crashes [2].

Since, there are no standard rules to measure driver's attention level; the unique solution is to observe driver's attention level continuously. Unfortunately driver's attention level goes down due to several reasons such as sleep deprivation, fatigue, talking over mobile phone, texting over cell phone, talking with passengers, looking away out of the direction, eyes off the road, hypnotic

DOI : 10.5121/ijcseit.2018.8102

drugs, driving more than two hours without break and driving in a monotone road. Most of the time drivers are not alert when their attention level goes down and as a result, catastrophic road accident may occur within a second. So, driving with proper level of attention plays an important role in reducing the accident rate as well as promises safe journey. At the same time continuous monitoring of driver's attention level is also a significant issue for the improvement of the driving enactment.

There is numerous research activities have been conducted on designing an intelligent system related to safe driving. Although there are plenty of issues remained unsolved related to the driving. A. Tawari et al., [3] have conducted an Urban Intelligent Assist (UIA) project which aim is to make urban day-to-day driving safer and stress free. A 3year collaboration between Audi AG, Volkswagen Group of America Electronics Research Laboratory and UC San Diego, the driver assistance portion of the UIA project focuses on two main use cases of vital importance in urban driving. The first, Driver Attention Guard, applies novel computer vision and machine learning research for accurately tracking the driver's head position and rotation using an array of cameras. The system then infers the driver's focus of attention, alerting the driver and engaging safety systems in case of extended driver inattention. The second application, Merge and Lane Change Assist, applies a novel probabilistic compact representation of the on road environment, fusing data from a variety of sensor modalities.

F. S. C. Clement et al., [4] have developed driver fatigue detection system based on Electrocardiogram (ECG), Electromyogram (EMG), Electrooculogram(EoG) and Electroencephalogram (EEG) approaches. But these approaches suffer from various limitations [5]: the first limitation is cost. Both facial recognition and EEG technologies require expensive sensors that are not affordable to the general public. This limits both the exposure to the technology and its applications. The second problem is that of invasiveness. Sensors that need to be worn on the body or limit the mobility of the learner are more difficult to introduce and gain acceptance.

This study represents a vision based framework to monitor the driver's attention level in real time by using Microsoft Kinect for Windows sensor V2. The Kinect sensor provides color images, IR images, 3D information and also adapted when driving at hostile light conditions. Different parameters information such as eye state, face angle and mouth motion is collected with the help of face tracking SDK. Later the parameters information is used to assess the level of attention in percentage. Furthermore, the framework generates a warning signal as a safety measure when driver's attention level goes down.

2. PROPOSED METHOD

As it is mentioned in the introduction of this study, the main purpose of this work is to design a sensible framework to measure the attention level of vehicle driver in real time. In spite of considering lots of factors that are involved in measuring the attention level of vehicle driver, this study considers three key parameters such as: Eye State (ES), Face Angle (FA) & Mouth Motion (MM). The steps of designing a sensible framework to measure the attention level of vehicle driver are illustrated as:

1. Vehicle driver should sit within the viable range of the Kinect sensor to collect the attention level of him. The RGB color camera of Kinect captures the vehicle driver and IR emitter emits dotted light patterns through the human.

2. The Kinect depth sensor read dotted light pattern that emits by IR emitter.

3. Then a depth image of the vehicle driver is produced by Kinect depth sensor.
4. The Kinect depth sensor sends continuous depth information to the Kinect device driver for further processing. Primary processing is done within the Kinect sensor.
5. The Kinect device driver receives continuous depth information from Kinect depth sensor.
6. Then depth information sends to personal computer for far ahead processing.
7. After that, face of the vehicle driver is tracked with the help of Microsoft Face Tracking Software Development Kit for Kinect for Windows (Face Tracking SDK), together with the Kinect for Windows Software Development Kit (Kinect for Windows SDK).
8. Then assess level of attention by using previously detected face parameters.
9. The feedback about attention level provided to the driver to inform him whether he drives attentively or inattentively.

The schematic framework for assessing the attention level of vehicle driver is illustrated in Figure 1.

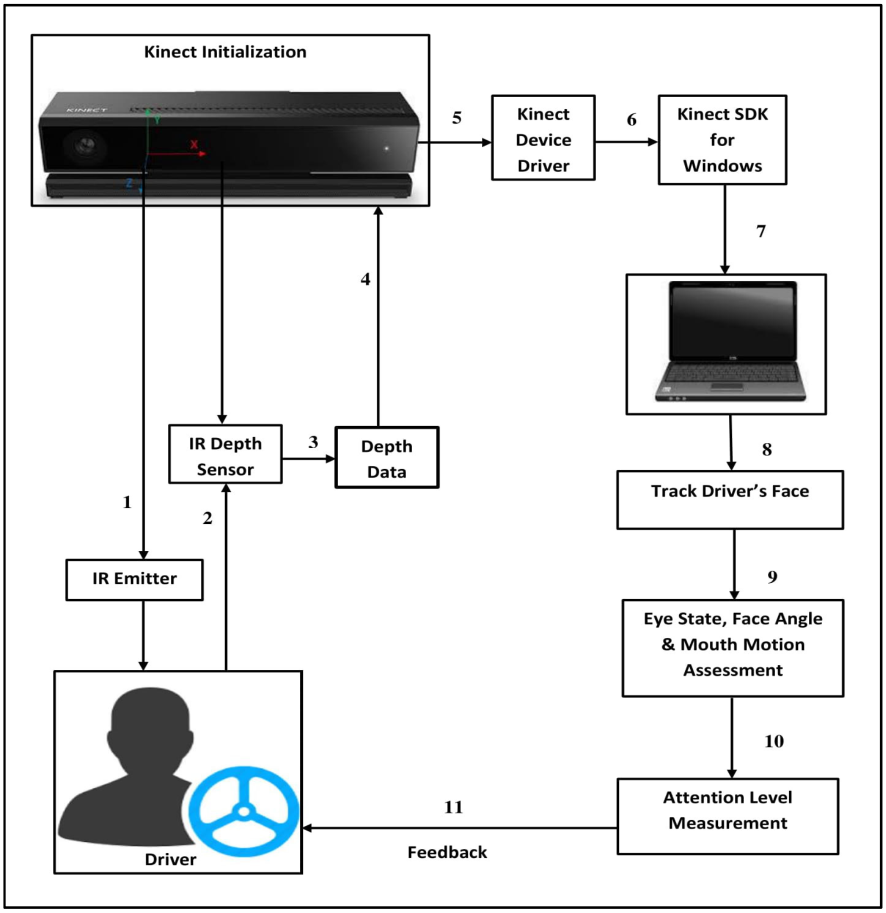


Figure 1.Schematic framework for assessing the attention level of vehicle driver

2.1. Depth Data Encapsulation

Microsoft Kinect is composed of an infrared (IR) depth sensor, a built in color camera, an infrared (IR) emitter and a microphone array, enabling it to sense the location and movements of human. The color camera is responsible for capturing and streaming the color video data and to detect there is an element (i.e. human). Depth camera also captured the human and produces a depth image of that human. To obtain the X, Y and Z coordinate values on specific point of human, IR emitter and the IR depth sensor perform combined function. And with up to 3 times higher depth, fidelity, the Kinect V2 sensor provides significant improvements in visualizing small objects and all other objects more clearly.

2.2. Depth Information Processing

After capturing and detecting that there is an element (i.e. human) by color camera, the IR emitter of Kinect persistently emits infrared light in a random dot pattern over everything in front of it. The dotted light reflects off various objects and normally these dotted light unable to be seen by us. But by using an IR depth sensor it is possible to capture their depth information. Hence, the dots in the scene read by IR depth sensor and processed by it. IR depth sensor sends the primary depth information from which they were reflected. Based on the distance of the individual light points of reflection, IR depth sensor start reading the secondary data and pass it to the prime sense chip within the device. After getting the secondary data, the prime sense chip analyses the received data and creates a depth image. The depth stream data as a succession of the depth image frame returned by Kinect sensor with 16 bit gray scale format.

2.3. Face Tracking with Kinect

After getting the depth stream data as a succession of the depth image frame by Kinect sensor, face tracking SDK installed in laptop computer help to track the human face. The face tracking SDK is performed on the basis of Active Appearance Model [5]. An Active Appearance Model (AAM) is a computer vision algorithm for toning a statistical model of object shape and entrance to a new image. An image set, combined with coordinates of marks that appear in all of the images, is provided to the training supervisor. Principal Component Analysis (PCA) is used to find the mean shape and main variations of the training data to the mean shape. After finding the Shape Model, all training data objects are deformed to the main shape and the pixels converted to vectors. Then, PCA is used to find the mean appearance (intensities) and variances of the appearance in the training set. Both the Shape and Appearance Model are combined with PCA to one AAM model. Any example [6] can then be approximated by using Eq. 1.

$$\mathrm{a} = \bar{\mathrm{a}} + \mathrm{U_s}\,\mathrm{x_s} \qquad (1)$$

where, a is a vector that represent the aligned points of all image sets into a common coordinate frame, $\bar{\mathrm{a}}$is the mean shape, $\mathrm{U_s}$ is a set of orthogonal modes of shape variation and $\mathrm{x_s}$is a set of shape parameters.

By displacing the parameters in the training set with a known amount, a model can be created which gives the optimal parameter update for a certain difference in model intensities and normal image intensities. This model is used in the search stage. To drive an optimization process, AAM uses the transformation between the current assessment of appearance and the target image. Formally [7] this can be written as Eq. 2.

$$\mathrm{dI} = \mathrm{I_{image}} - \mathrm{I_{model}} \qquad (2)$$

where, dI is the difference vector.

In this way the fit can be enhanced by adjusting the model and pose parameters to fit the image in the best possible way. The searching process of AAM can call be a prototype search [7] – the search path and the optimal model parameters will be unique in each search where the initial and final model configuration matches this configuration. These prototype searches can then be made at model building time; thus saving computationally expensive high dimensional optimization. AAM can match to new images very swiftly by taking advantage of the least squares techniques.

The face tracking SDK utilizes Kinect's depth data, so it can track faces/heads in 3D. The person can be detected either while standing or sitting and facing the Kinect. The head pose tracking of the current frame depends on the tracking result of previous frame. In the failure case of head tracking in previous frame, the Kinect sensor starts processing of head detection for current frame.

If head detection is also failed, the sensor skip current frame from being considering. If head detection is successful, the sensor immediately tracks the head to found the head pose angle. Yaw, pitch and roll are three head pose angles can assess by the sensor, respectively. If the sensor cannot found any head to track, it skips the current frame. In the success case of head tracking in previous frame, the sensor assesses head movement and based on the movement, the head is tracked. If head tracking is successful, three head pose angles namely, yaw, pitch and roll can assess by the sensor, respectively. Otherwise, the sensor skips the current frame. The work flow of the human head tracking system is presented in a flowchart [8], as shown in Figure 2.

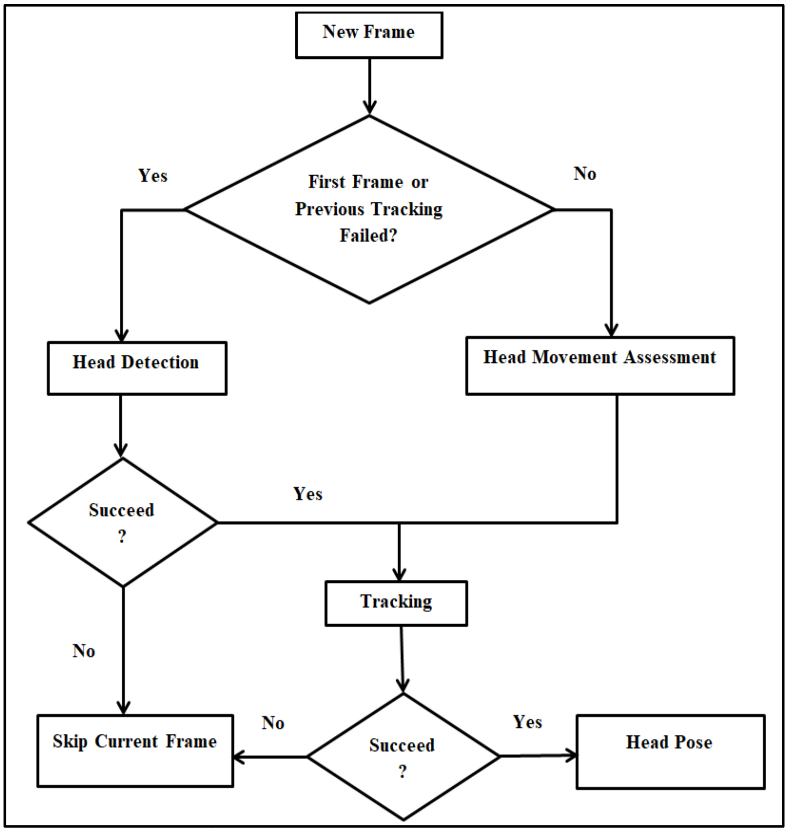


Figure 2.Workflow of human head pose tracking system

2.4. Procedure to Measure Attention Level

In this study, attention can be considered as a function of three parameters such as, Eye State (ES), Face Angle (FA) and Mouth Motion (MM), as shown in Eq. 3.

$$M(A) = f\ (ES, FA, MM) \qquad (3)$$

where, $M(A)$is the Measurement of attention, ES is the Eye State with respect to Kinect, Face Angle and Mouth Motion of the vehicle driver are represented by FA and MM respectively. ES has three sub-parameters: left eye, right eye and looking away. FA has three sub-parameters: face yaw, face pitch and face roll. MM has one sub-parameter: mouth open.

Procedure to measure level of attention of vehicle driver represented in Algorithm 1.

Algorithm 1:Algorithm for measuring driver's attention level

```
 1. Input face image through kinect sensor in real time.
 2. Find parameters as well as sub-parameters values with the help of face tracking SDK
    interface (using C++). These are:
            Left Eye Closed (Yes/No),
            Right Eye Closed (Yes/No),
            Looking Away (Yes/No),
            Mouth Open (Yes/No),
            Face Yaw,
            Face Pitch and
            Face Roll
 3. If Left and Right Eye == Closed > = threshold value (2 seconds) then
   |  set attention level with Eye State (AE=0) in percentage
end
 4. Else if Left Eye and Right Eye == Open and Mouth == Close then {
   |  find highest degree from face yaw, face pitch and face roll;
   |  use cosine function to find attention level;
   |  set attention level as Attention with Face Angle (AF) in percentage
   |                                                   }
end
 5. Else Left Eye, Right Eye and Mouth == Open then {
   |  find highest degree from face yaw, face pitch and face roll;
   |  calculate Attention with Mouth Motion (AM) as,
   |
   |                        AM = 19 (AF)/20          (4)
   |
   |  set attention level as Attention with Mouth Motion (AM) in percentage
   |                                                   }
end
 6. Display assessed attention level A (in percentage) as AE, AF or AM according to
    parameters as well as sub-parameters values.
```

2.5. Classification of Attention Level

Range of attention level required to specified to find whether driver drive the vehicle attentively or inattentively. The range of the attention level classified from full attention to no attention of driving based on attention distraction.If driver's attention distraction is maximum or drive the vehicle with no attention, the framework will warn the driver to provide required attention of driving. Table 1 shows the classification of attention level.

Table 1.Classification of attention level.

Attention Level (%)	**Classification**
A = 0	No Attention
A < 10	Maximum Attention Distraction
A < 20	Very High Attention Distraction
A < 40	High Attention Distraction
A < 60	Medium Attention Distraction
A < 70	Low Attention Distraction
A < 90	Very Low Attention Distraction
A <= 100	Full Attention

2.6. Procedure to Design Alarm Generating Segment

Proposed framework generates an alarming sound to alert the driver when driver is inattentive for a threshold time period such as 2 seconds or more than 2 seconds. Figure 3 shows the work flow of alarm generating segment. At first the proposed framework determines whether the driver's eyes are closed or not. If the eyes are closed for a threshold time period, it generates an awareness signal. If eyes are not closed but driver looking away out of the direction for a threshold time period, it also generates an awareness signal.

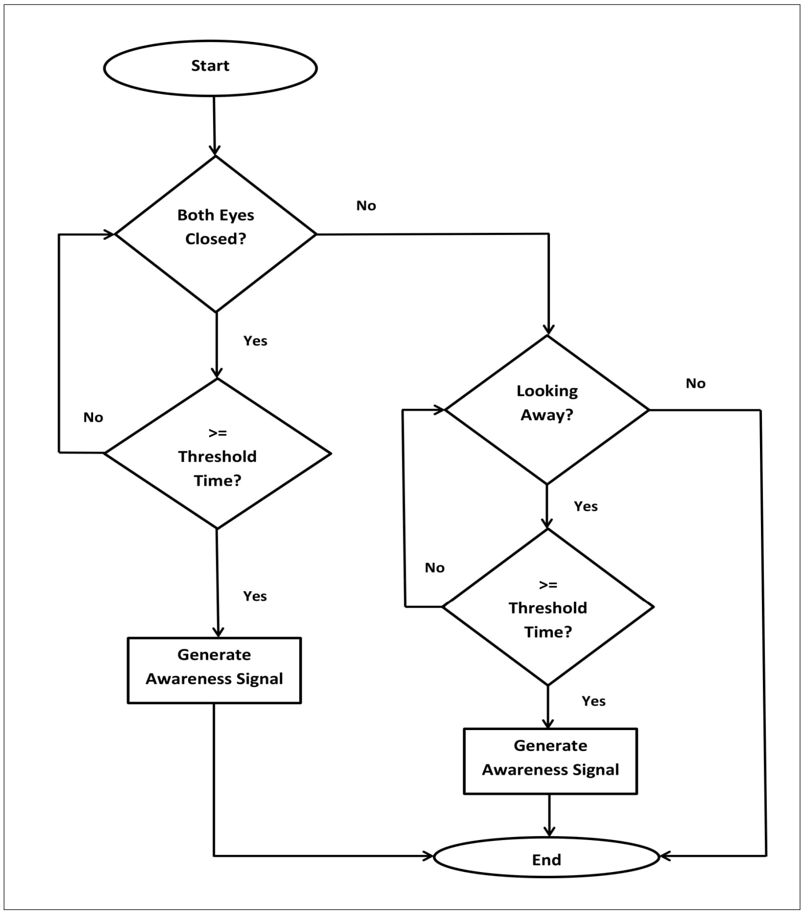


Figure 3.Workflow to design alarm generating system

While driving, it is very common that the driver may feel sleepy or tired and the attention of driving might distracted. As a result serious accident can happen within a second. To prevent the situation it is very important to inform the driver that he is not attentive. There are several ways to warn the driver, this study choose the acoustic alarm as a warning signal when attention level of driver gradually goes down.

3. EXPERIMENTAL RESULTS

3.1. Experimental Setup

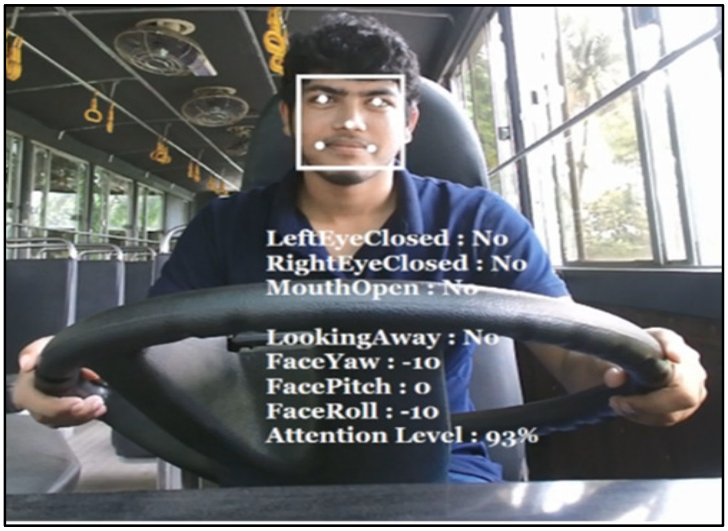
To create experimental environment, Microsoft Kinect is connected with the computer via power adapter and the sensor device set in a proper position in front of the driver.While testing, the driver has to be seated directly in front of the system environment which is established with Microsoft Kinect and laptop. To manipulate the Application Interface, laptop's built in keyboard and mouse are used. The Kinect sensor is attached at a distance of 15 inches infront of the steering of the vehicle and facing directly at the subject.

3.2. Participants

11 participants (10 graduate students and 1 bus driver) are recruited in the real vehicle environment to run the experiment. The average age of participants are 24 years (SD= 8.04). In this experiment, different people are positioned in the driving seat within the device range of area to get the appropriate attention level of driver. One by one all participants take part in the interacting experiment system.

3.3. Graphical User Interface

To monitor the measured attention level of the driver as well as values of three parameters, eye state, face angle and mouth motion, a Graphical User Interface is developed. The total Graphical User Interface (GUI) of the system with participant shown in Figure 4.The assessed values of these three parameters are very important to recognize the deviation of the attention level over time and also important to alert the driver if the attention level riskily falls down through warning signal.



LeftEyeClosed : No
RightEyeClosed : No
MouthOpen : No

LookingAway : No
FaceYaw : -10
FacePitch : 0
FaceRoll : -10
Attention Level : 93%

Figure 4.Graphical User Interface (GUI) of the system with participant

After founding a human instantaneously, the system sensor start detecting the human and after detection, the system records depth data of that part in real time.The face of the human is tracked by a mask of mesh. To evaluate the functionality of detection and tracking process, participants are asked to close their eyes, move their face left, right, up and down randomly and speak to detect the mouth motion as well as move the vehicle slowly. It correspondingly tests whether the system is resulting real time variation of eye state, face angle and mouth motion along with level of attention. Each trial is started with a static positioning of a participant and it takes approximately 3-4 hours to perform the experiment.

3.4. Evaluation of ES, FA and MM estimation module

The Kinect sensor is positioned within the viable range from vehicle steering to capture each participant in driving seat. The trial duration is more than 120 seconds and in this study first 30 seconds of trial duration are considered for performance evaluation of ES, FA and MM module for each participant. Statistical mode is used to find the most frequently appeared frame within one second and this frame taken into consider for evaluation performance by using Eq. 5.

$$M(A) = mode\ (An_{n=1}^{n=fps}) \qquad (5)$$

According to evaluation results, the proposed framework provides similar result for various participants. Such as, firstly, the evaluated values for 1st, 5th, 7th and 9th participant are same. Secondly, the evaluated values for 2nd, 3rd, 4th and 11th participantare same and thirdly, the evaluated values are same for 6th and 10th participant.

The evaluated result for 1st, 5th, 7th and 9th participant is shown in Table 2, where the proposed framework detect eye state, face angle and mouth motion with 100% accuracy. Table 3 shows the evaluated result for 2nd, 3rd, 4th and 11th participant, where the proposed framework detect eye state and face angle with 100% accuracy and mouth motion detected with 93.33% accuracy.

Table 2.Evaluation of ES, FA and MM estimation module for 1st, 5th, 7thand 9th participant.

Parameters	**Duration (sec)**	**Number of Times of Correct Decision (sec)**	**Number of Times of Wrong Decision (sec)**	**Accuracy (%)**	**Error Rate (%)**
Eyes State (ES)	30	30	0	100%	0%
Face Angle (FA)	30	30	0	100%	0%
Mouth Motion (MM)	30	30	0	100%	0%

Table 4 shows the evaluated result for 6th and 10th participant, where the proposed framework detect eye state and face angle with 100% accuracy and mouth motion detected with 96.67% accuracy.Table 5 shows the evaluated result for 8th participant in terms of ES, FA and MM estimation module, where the proposed framework detect eye state and face angle with 100% accuracy and mouth motion detected with 83.33% accuracy.

Table 3.Evaluation of ES, FA and MM estimation module for 2nd,3rd, 4thand 11th participant.

Parameters	**Duration (sec)**	**Number of Times of Correct Decision (sec)**	**Number of Times of Wrong Decision (sec)**	**Accuracy (%)**	**Error Rate (%)**
Eyes State (ES)	30	30	0	100%	0%
Face Angle (FA)	30	30	0	100%	0%
Mouth Motion (MM)	30	28	2	93.33%	6.67%

Table 4.Evaluation of ES, FA and MM estimation module for 6th and 10thparticipant.

Parameters	**Duration (sec)**	**Number of Times of Correct Decision (sec)**	**Number of Times of Wrong Decision (sec)**	**Accuracy (%)**	**Error Rate (%)**
Eyes State (ES)	30	30	0	100%	0%
Face Angle (FA)	30	30	0	100%	0%
Mouth Motion (MM)	30	29	1	96.67%	3.33%

Table 5.Evaluation of ES, FA and MM estimation module for 8thparticipant.

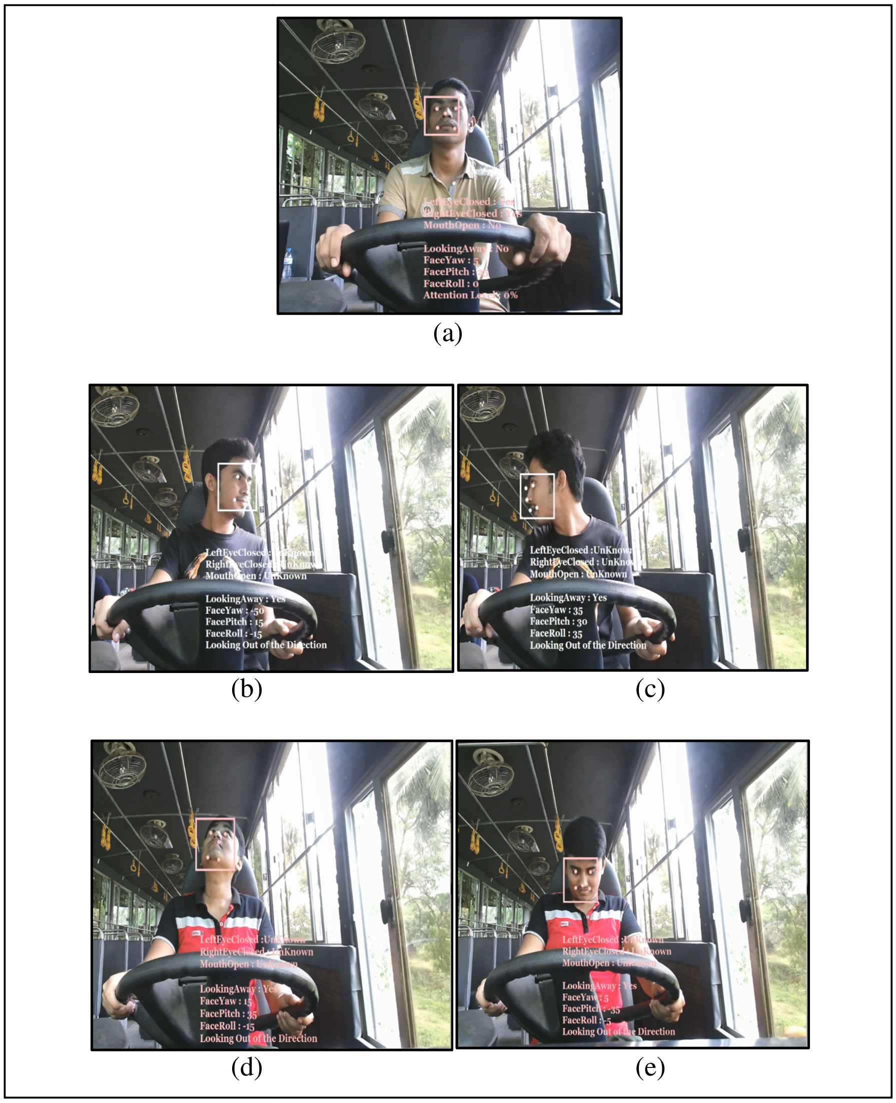
Parameters	**Duration (sec)**	**Number of Times of Correct Decision (sec)**	**Number of Times of Wrong Decision (sec)**	**Accuracy (%)**	**Error Rate (%)**
Eyes State (ES)	30	30	0	100%	0%
Face Angle (FA)	30	30	0	100%	0%
Mouth Motion (MM)	30	25	5	83.33%	16.67%

With the help of evaluation results of ES, FA and MM estimation module as represented in Table 2, 3, 4 and 5 respectively, it can possible to find overall performance accuracy of theproposed framework. According to performance evaluation results in Table 2, 3, 4 and 5 respectively, it can be stated that, the proposed framework detect eye state and face angle with average accuracy 100% and detect mouth motion with average accuracy 95.45% by considering total participants.

So, the proposed framework performed with overall accuracy rate 98.48% and error rate 1.52%. Due to higher accuracy rate and lower error rate the proposed framework functioning quite well as far expected.

3.5. Alarm Generating Segment

There may a possibility of vehicle crashes when driver drive the vehicle inattentively. For this reason it is important to aware the driver when the attention level is very low for a while.



(a)

(b) (c)

(d) (e)

Figure. 5. Examples of alarm generating cases:(a) when both eyes are closed for 2 seconds or more than 2 seconds,(b) looking out of the direction (left side),(c) looking out of the direction (right side),(d) looking out of the direction (up) and(e) looking out of the direction (down)

Proposed framework generates an alarming sound to alert the driver when he/she becomes inattentive for 2 seconds or more than 2 seconds. If the driver is not being attentive after listening the alarm, then the system will generate continuous sound to draw the driver's attention until the driver's high attention level is certain.

Figure 5 shows the examples of alarm generating cases. In first case, the proposed framework generates a warning signal when driver's both eyes are closed for threshold time period such as, 2 seconds or more than 2 seconds. For rest of the cases the proposed framework generates a warning signal when driver looking out of the directions such as left side, right side, up and down for threshold time period.

4. CONCLUSION

This study presents a framework to monitor driver's attention level by considering three important parameters such as eye state, face angle and mouth motion. As the depth data has been considered rather than that of RGB image processing, performance is quite improved overcoming the environmental effects. Image based processing requires complex algorithm and computational costs, since simpler algorithm for dealing with depth data considered in this study along with the main application which is Visual Studio 2013 (C++). The face of the vehicle driver was tracked and the recorded features coordinate values provided to measure driver's attention level.The proposed framework is evaluated in different environments with different participants and results reveal that it is able to measure the level of attention of the driver with a reasonable accuracy.
The experimental results have shown that:

I. The evaluated average accuracy of the proposed framework is almost 98.48% and error rate 1.52%.

II. The measured attention level for each participant is quite satisfactory as far subjective and objective evaluations.

III. The proposed framework generate awareness alarm in the form of acoustic signal to alert the participant when his eyes are closed or looking out of the directions for two or more than two seconds, as illustrated in Figure 5.

This study suffered from certain problems with the Kinect such as:

I. This kind of Kinect needs a specific requirements:

   - Physical dual core 2.6 GHz - 2 logical cores per physical or faster processor
   - USB 3.0 controller dedicated to the Kinect for Windows V2 sensor
   - 4 GB of RAM
   - Graphics card that supports DirectX 11
   - Windows 8 or 8.1

II. Sometimes the frame per second is down in (18-24 FPS) and the recognition became little bit slow. This also happened with the older version of Kinect V1.

III. It is required to place the Kinect camera directly facing to the test subject.

REFERENCES

[1] https://en.wikipedia.org/wiki/Traffic_collision Date accessed: December 2016.

[2] State of the Road, A fact sheet of CARRS-Q, Centre of Accident Research and Road Safety Queensland, Queensland, Australia. Date accessed: March 2013.

[3] A. Tawari, S. Sivaraman, M. M. Trivedi, T. Shannon and M. Tippelhofer, “Looking-in and Looking-out Vision for Urban Intelligent Assistance: Estimation of Driver Attentive State and Dynamic Surround for Safe Merging and Braking,” Proc. of IEEE Intelligent Vehicles Symposium (IV), doi: 978-1-4799-3637-3/14, 2014.

[4] F. S. C. Clement, A. Vashistha and M. E. Rane, “Driver Fatigue Detection System,” Proc. of the IEEE International Conference on Information Processing, doi: 978-1-4673-7758-4/15, 2015.

[5] D. Stanley, “Measuring Attention using Microsoft Kinect,” Master's Thesis, Dept. of Computer Science, College of Computing & Information Sciences, Rochester Institute of Technology, May 10, 2013.

[6] G. J. Edwards, T. F. Cootes and C.J. Taylor, “Face Recognition Using Active Appearance Models,” Proc. of European Conference on Computer Vision (ECCV), 1998, pp. 581-595.

[7] M. B. Stegmann, “Active appearance model: Theory, Extensions and Cases,” 2nd edition, September 2000.

[8] A. Yargic and M. Dogan, “A Lip Reading Application on MS Kinect Camera,” Proc. of IEEE International Symposium on Innovations in Intelligent Systems and Applications (INISTA), doi: 10.1109/INISTA.2013.6577656, 2013.

[9] G. L. Masala and E. Grosso, “Real time detection of driver attention: Emerging solutions based on robust iconic classifiers and dictionary of poses,” Article in Transportation Research Part C, 2014, pp. 32–42.

[10] J. Li, G. Ngai, H. V. Leong and S. Chan, “Multimodal Human Attention Detection for Reading,” Proc. of the 2016 ACM Symposium on Applied Computing (SAC' 16), 2016.

[11] T. McWilliams, B. Reimer, B. Mehler, J. Dobres and H. McAnulty, “A Secondary Assessment Of The Impact Of Voice Interface Turn Delays On Driver Attention And Field Conditions,” Proc. of the Eighth International Driving Symposium on Human Factors in Driver Assessment, Training and Vehicle Design, 2015.

[12] P. Chowdhury, L. Alam and M. M. Hoque, “Designing an Empirical Framework to Estimate the Driver’s Attention,” Proc. of the 5th IEEE International Conference on Informatics, Electronics and Vision (ICIEV), 2016, pp. 513 – 518.

AUTHORS

Munira Akter Lata received her B.Sc(Hons.) degree in Information Technology from Jahangirnagar University, Savar, Dhaka, Bangladesh in 2016. Currently, she is serving as Lecturer in the Department of Computer Science and Engineering of City University, Dhaka, Bangladesh. Her current research interests include computer vision, image processing, machine learning, artificial intelligence and network security.

Md Mahmudur Rahman received his B. Sc (Hons.) degree in Information Technology from Jahangirnagar University, Savar, Dhaka, Bangladesh in 2017. Currently, he is serving as Lecturer in the Department of Computer Science and Engineering of University of South Asia, Dhaka, Bangladesh. His current research interests include computer vision, image processing and network security.

Dr. Mohammad Abu Yousuf is currently serving as Associate Professor in Institute of Information Technology at Jahangirnagar University, Savar, Dhaka, Bangladesh. His current research interests include computer vision, human-robot interaction and medical image.

Md. Ibrahim Al Ananya received his B.Sc (Hons.) degree in Information Technology from Jahangirnagar University, Savar, Dhaka, Bangladesh in 2017. Currently, he is pursuing his M.Sc degree in Information Technology from Jahangirnagar University, Savar, Dhaka. His current research interests include computer vision, image processing and network security.

Devzani Roy received her B.Sc (Hons.) degree in Information Technology from Jahangirnagar University, Savar, Dhaka, Bangladesh in 2017.Currently, she is pursuing her M.Sc degree in Information Technology from Jahangirnagar University, Savar, Dhaka. Her current research interests include computer vision, image processing and network security.

Alvi Mahadi received his B.Sc (Hons.) degree in Information Technology from Jahangirnagar University, Savar, Dhaka, Bangladesh in 2016. Currently, he is serving as Junior Software Engineer at Nascenia Ltd. Bangladesh. His current research interests include machine learning, computer vision and artificial intelligence.