Non Intrusive Physiological Measurement for Driver Cognitive Distraction Detection: Eye and Mouth Movements

Afizan Azman
Research School of Informatics,
Loughborough University,
LE11 3TU, UK
a.azman@lboro.ac.uk

Qinggang Meng
Research School of Informatics,
Loughborough University,
LE11 3TU, UK
q.meng@lboro.ac.uk

Eran Edirisinghe
Research School of Informatics,
Loughborough University,
LE11 3TU, UK
e.a.edirisinghe@lboro.ac.uk

***Abstract*- Driver distractions can be categorized into 3 major parts:-visual, cognitive and manual. Visual and manual distraction on a driver can be physically detected. However, assessing cognitive distraction is difficult since it is more of an "internal" distraction rather than any easily measured "external" distraction. There are several methods available that can be used to detect cognitive driver distraction. Physiological measurements, performance measures (primary and secondary tasks) and rating scales are some of the well-known measures to detect cognitive distraction. This study focused on physiological measurements, specifically on a driver's eye and mouth movements. Six different participants were involved in our experiment. The duration of the experiment was 8 minutes and 49 seconds for each participant. Eye and mouth movements were obtained using the FaceLab Seeing Machine cameras and their magnitude of the r-values were found more than 60% thus proving that they are strongly correlated to each other.**

Keywords:- Driving Safety, Cognitive distraction, Pearson-r Correlation, Bayesian Network

I. INTRODUCTION

Driving safety issues include inattention, fatigue, lack of concentration, behavior, alcohol, drug, using mobile phone, listening to a radio and music and many more. From all those issues, the most important issue to be considered is the driver's concentration as well driver's inattention or distraction.

Driver distraction can be defined as a process that draws away drivers' attention from the road. It disturbs the driving and the vehicle control process. Generally concentration and attention are two similar words. Concentration is when someone is doing something to focus on, to give one's attention. Concentration can also be defined as an ability to re-focus attention if and when one is distracted, or immediately after any rest or one or more breaks, on the same task or other tasks. The strength or the ability to concentrate is actually achieved by mental conditioning. Attention at the other hand is part of focus and a component of intelligence. Attention usually only lasts for a few seconds. Some tests suggested that a person can only put his attention up to 90 seconds, however, 30 seconds is generally regarded as maximum period of putting an attention. Attention is the act or process of focusing on one or more particulars in a content of one's consciousness so as to give an essential or priority to restricting one's input and ignoring unwanted aspects [1]. Thus, when it comes to a driver's attention and concentration definition, it is related to where a driver's looking, doing and thinking that are related to driving.

In this paper, the authors will explain about driving safety issues that are related to driver's distraction and how the distraction can be detected by using non-intrusive equipment. The detection is basically covered the physiological measurement where eye and mouth movements are captured and studied. Experiment was conducted with six participants were involved and it has been found that, the movements from eye and mouth were highly correlated when the participants were distracted. The experimental setup will be explained in details and the results are also discussed. Data collected were analyzed initially by Pearson-r correlation. Towards the end of this project, the data will be analyzed with a data mining technique called Bayesian Network.

A. Driving and Perceptual Factors

With so many In Vehicle Information System (IVIS) in a vehicle, such as radio, cassette player, navigator system, cellular phone and others, the hands might get away from the steering wheel, eyes off the road and the mind may stop processing information related to the road and vehicle safety. Ideally, while driving, a driver is not only required to look and handle the car properly, but also required to think about driving safety information only. This seems and sounds impossible because in a real-time environment, a driver can be easily distracted thinking about something else. In psychology this is referred to as a subconscious mind phenomenon and in a driving environment, this situation is in line with Baumann and Krems (2007) claimed as an inattention blindness or looked-but-did-not-see phenomenon [2].

When people are looking around, the three-dimensional world is perceived immediately and effortlessly. By likening the eye with a camera, it produces a leading impression that both are entailed in perception. Both have a common procedure in the physics of image formation. However, cameras merely record an image and it will be interpreted by the visual system. Seeing is actually not synonymous with image formation. Eyes can see, and it can be a little difficult for the interpretation to be grasped. One attempt to understand the perceptual aspects of driving is that, the purpose of visual perception is not just to transmit an image from the eye to the brain, or to clean up or make sense of the imperfect image on the back of the retina, but the perceptions actually consist of a set of processes that have been developed by evolution to help the people to survive, finding food and mates and avoiding predators and dangers [3]. Thus, the perceptual process was actually designed to supply useful information about what is happening in the environment. But it is not necessary for the information to contain all details. In a driving environment, this perceptual process gives a high priority for a driver to find edges and signs and detecting movement objects. Image perceived by a driver involves many tasks and each can be quite tricky. With a limited period of time to process certain images received and limited resources to process the information, a driver really needs to focus while driving.

B. Non-Intrusive Apparatus

Generally, many problems can be solved if the assessment can help to understand that a person is actually thinking about something else when he is driving. It will not only solve active and passive safety issues but it can also give a new idea about vehicle intelligence system development. There are a few methods available to assess a driver's cognitive distraction. Physiological measurements can be seen as one of the most popular method to use because it represents human body gesture, facial expression and body or organ movements. Physiological measurement measures and monitors a range of physiological parameters, usually the face, head, mouth, eyes, skin, heart etc. Compared to other types of measurement, physiological measurements are available in real time. Primary tasks, secondary tasks and behavior performance cannot be detected in a real time situation. Real time cognitive assessment would be a useful tool for building and testing more detailed cognitive models and the measures obtained could be used in real time to modify a task relative to the changing cognitive measures. Physiological measurements can be intrusive or unobtrusive. The unobtrusive physiological measures are measured and objectively analyzed in a real time without impacting user performance and user ratings.

However, to use this physiological measurement, many researchers are inclined to use intrusive equipment where the devices are mounted onto the human body. Some existing equipments are either head mounted devices like Head Mounted Display (HMD), wearable glasses like Eye-R Sensor System, Eye Touch System (ETS), ASL Mobile Eye and EyeLink II or attachable wire like using Galvanic Skin Response (GSR) System are not only inconvenient to the participants, but they are also impractical to be used in real daily life. On top of those weaknesses these types of devices have a drawback that they can induce drowsiness or a headache.

In this study, a non-intrusive device was used. Thus, the participants in the experiment were not required to wear or hold anything during the experiment. The images of the participant's facial expression were captured using faceLab Seeing Machine cameras (version 4.1).

C. Eye and Mouth Movement

Many studies in driver cognitive distraction detection, have investigated the patterns of eye movements. Eye movements can give a variety of information based on fixation, saccade, eyelid movement (height and width), blinking duration, blinking frequency, pupil diameter, gaze direction and rotation etc. Liang and Miyaji used the eye movement as their main feature to detect cognitive distraction on a driver [4][5][6]. Miyaji used blinking, saccade, eyelid movement and pupil diameter [4] whereas Liang used the characteristics of fixations, saccades and smooth pursuits to recognize the patterns of eye movements [5][6]. Two experiments in out study were conducted to collect data regarding these movements of the eye with a 60 Hz using FaceLab Seeing Machine eye tracking system.

In human science and psychology studies, it has been proved that mouth movement is a good indicator of a human's state of mind. Mouth movement can also be thought of as a form of body language. Body language can be used to obtain information about whether a person is distracted or not. Rongben[7], has monitored the relationship between mouth movement and driver fatigue or distraction using a camera. Normally the mouth is hardly open when the driver is alert. The maximum width (W_{max}) and maximum height (H_{max}) can indicate different levels of distraction. The height (H_m) between top lip and the bottom lip varies greatly when one is talking, yawning or even thinking. Thus, the mouth movement can be represented in the feature vector as $Z = (W_{max}, H_{max}, H_m)$ [7].

Data from eye and mouth movements were captured from the FeatureSet output from faceLab cameras. Each FeatureSet contains two FeaturePoint sub-objects [8]:

a) The position of the Left pupil, in normalized image coordinates

b) The position of the Right pupil, in normalized image coordinates

II. EXPERIMENTAL SETUP

The experiment was conducted in a laboratory setup. Six participants were chosen for the experiment. All the participants were experienced drivers. On an average, they have been driving for five years. Therefore, all the participants are familiar with the driving conditions and environment. The participants were aged between 28 to 38 years old and all of them were familiar with the route shown and ever had used the route before.

A video sequence of a real road environment around Loughborough town, Leicestershire, UK was shown to all participants. The video was 8 minutes and 49 seconds long. Two sets of experiment were conducted for each participant. The first set was a controlled experiment, where the participants were only required to watch the video and there wasn't any kind of distraction. The second set of the experiment was run with several types of distractions. As pointed out by the American Automobile Association (AAA) in their report about passenger vehicles and road safety, there are mainly 13 different types of distractions [9]. However in this experiment the participants were distracted with only some specific types of distractions such as trying to recognize a pedestrian or talking with other passengers in the vehicle, listening to music or memorizing a seen object or a an event from the video. Mostly, the distractions occurred from verbal distraction where they were asked a few questions. Questions have been categorized into a few groups:-Open questions, Radio-related questions, Safety-related questions and Simple Arithmetic questions. While watching the video, the participants were also listening to a recorded streaming radio. This was done to create a real-driving experience.

To distinguish both sets of experiments, the authors used the *annotation_id* button from FaceLab Seeing Machine System to annotate the time or frame number when a distraction started and ended. et

III. RESULTS

Initially the data collected was analyzed by a statistical modeling tool, SPSS. A correlation between eye movements and mouth movements are studied. Distractions were annotated every time during the course of the experiment, and they were fixed variables in this study. Distraction is annotated with annotation button 1 and No Distraction is with default value 0. Eye and mouth movements varied in values. From the analysis it has been proved that both eye and mouth movements have a very high correlation to each other. The correlation between both features was analyzed using Pearson-r Correlation which also known as Bivariate Correlations Algorithms [10]. Pearson correlation is defined as the covariance of the two variables divided by the product of their standard deviations.

Pearson-r correlation is typically used to describe the strength of the linear relationship between two quantitative variables. r-values were ranged from -1.00 to +1.00. The sign of r-value provides information about the direction of the relationship between those two variables.

The magnitude of the r-values showed the strength of the relationship between those two variables [11].

a) 0.0 to 0.3 = negligible correlation

b) 0.3 to 0.5 = low correlation

c) 0.5 to 0.7 = reasonable correlation

d) 0.7 or more = good or strong correlation

It was surprisingly good, that whole data collected showed a magnitude either in a reasonable correlation range or strong correlation range. However, some of the data collected especially towards the end of the experiment process, some data were unacceptable to be analyzed because the participants were moving their head far from the boxes frame that capture their face.

Mouth and Eye Movements data are contained in FeatureSetsbyCamera class. It contains data related to the position of facial features in each camera image. Thus, the position will show the movement on both eye and mouth whenever a distraction occurs. Data from the columns below are the observed nodes in the Bayesian Network Model that will be used later.

a) +leftEyeRect() : FeatureRect

b) +rightEyeRect() : FeatureRect

c) +mouthRect() : FeatureRect

Specifically, the movements between eye and mouth were studied on the basis of their width and height information: +_width:float and +_height:float.

TABLE 1: EYE'S HEIGHT AND MOUTH'S HEIGHT

Correlations

		REYERECT CA H	LEYERECT CA H	REYERECT CB H	LEYERECT CB H	MOUTRECT CA H	MOUTRECT CB H
REYERECT _CA_H	Pearson Correlation	1.000	.704**	.955**	.873**	.923**	.937**
	Sig. (2-tailed)	.000	.000	.000	.000	.000	.000
	N	183292	183292	183292	183292	183292	183292
LEYERECT_ CA_H	Pearson Correlation	.704**	1.000	.518**	.924**	.776**	.628**
	Sig. (2-tailed)	.000	.000	.000	.000	.000	.000
	N	183292	183292	183292	183292	183292	183292
REYERECT _CB_H	Pearson Correlation	.955**	.518**	1.000	.749**	.868**	.940**
	Sig. (2-tailed)	.000	.000	.000	.000	.000	.000
	N	183292	183292	183292	183292	183292	183292
LEYERECT_ CB_H	Pearson Correlation	.873**	.924**	.749**	1.000	.899**	.829**
	Sig. (2-tailed)	.000	.000	.000	.000	.000	.000
	N	183292	183292	183292	183292	183292	183292
MOUTRECT _CA_H	Pearson Correlation	.923**	.776**	.868**	.899**	1.000	.955**
	Sig. (2-tailed)	.000	.000	.000	.000	.000	.000
	N	183292	183292	183292	183292	183292	183292
MOUTRECT _CB_H	Pearson Correlation	.937**	.628**	.940**	.829**	.955**	1.000
	Sig. (2-tailed)	.000	.000	.000	.000	.000	.000
	N	183292	183292	183292	183292	183292	183292

**. Correlation is significant at the 0.01 level (2-tailed).

The table above shows a relationship between the variables right eye's height and left eye's height with the mouth's height. It clearly showed that both the variables have a strong or good relationship. Right eye's height magnitude ranged from 0.868 (the lowest) to 0.940 (the highest). However, the left eye's height magnitude ranged from 0.628 (the lowest) to 0.899 (the highest). These r-values definitely mean that whenever the eye's height changes or eye moves, the mouth's height also relatively changes and mouth moves. Another interesting conclusion that can be made from the result is that the right eye's height is more related (53.94%) to the mouth's height as compared to the left eye's height (46.05%). Based on a research in the area of 'Eyes Direction Language' it has been shown that many psychologists have defined that when a person eye's direction is at left or right, it indicates that they are thinking about something. The thing to look out for is the direction someone's eyes are looking in when they are thinking. Looking to the left indicates that a person is reminiscing or trying to remember something. On the other hand, looking to the right indicates that a person is doing more creative thoughts and this is often interpreted as a potential sign that someone may be being deceitful in some situation for an instance, creating a version of events or trying to find solution for a problem [12]. Often, a person is looking to right more than looking to left because a person usually involved with visually and auditory remembered images (Vr and Ar), and internal dialog (Ai) when he is "talking to himself".

TABLE 2: EYE'S WIDTH AND MOUTH'S WIDTH

Correlations

		REYERECT_ CA W	LEYERECT CA W	REYERECT CB W	LEYERECT CB W	MOUTRECT CA W	MOUTRECT CB W
REYERECT_ CA_W	Pearson Correlation	1.000	.710**	.962**	.880**	.946**	.958**
	Sig. (2-tailed)	.000	.000	.000	.000	.000	.000
	N	183292	183292	183292	183292	183292	183292
LEYERECT_ CA_W	Pearson Correlation	.710**	1.000	.521**	.931**	.788**	.635**
	Sig. (2-tailed)	.000	.000	.000	.000	.000	.000
	N	183292	183292	183292	183292	183292	183292
REYERECT_ CB_W	Pearson Correlation	.962**	.521**	1.000	.754**	.887**	.957**
	Sig. (2-tailed)	.000	.000	.000	.000	.000	.000
	N	183292	183292	183292	183292	183292	183292
LEYERECT_ CB_W	Pearson Correlation	.880**	.931**	.754**	1.000	.916**	.843**
	Sig. (2-tailed)	.000	.000	.000	.000	.000	.000
	N	183292	183292	183292	183292	183292	183292
MOUTRECT_ CA_W	Pearson Correlation	.946**	.788**	.887**	.916**	1.000	.963**
	Sig. (2-tailed)	.000	.000	.000	.000	.000	.000
	N	183292	183292	183292	183292	183292	183292
MOUTRECT_ CB_W	Pearson Correlation	.958**	.635**	.957**	.843**	.963**	1.000
	Sig. (2-tailed)	.000	.000	.000	.000	.000	.000
	N	183292	183292	183292	183292	183292	183292

**. Correlation is significant at the 0.01 level (2-tailed).

Table 2 shows a correlation between variables right eye's width and left eye's width with respect to the mouth's width. Interestingly, the result for eye's width from both sides were also significantly similar to the eye's height result as per discussed before. All magnitudes for r-value from the eye's width are also in a group of 0.7 or more which indicates a strong relationship. Right eye's widths ranged from 0.887 to 0.958 Left eye's widths ranged from 0.635 and 0.916. Similar to the right eye's height, right eye's width is also highly correlated to mouth's width. Right eye is 54.08% correlated to the mouth's width and left eye is only 45.92%. The right eye's width is more than the left eye's width is also due to the similar reasons explained before.

Those results justify that eye and mouth are relatively correlated to each other whenever a driver is distracted while driving. These results were initially used for verification and justification purposes. The data collected from faceLab will be later analyzed with a Bayesian Network (BNT) using toolbox [13].

This study used a Dynamic Bayesian Network (DBN). Hence, a transition of data for every parameter from the hidden nodes and observed nodes are checked at every single time slice, t. Bayesian approach learning starts with some prior knowledge about the model structure, the set of arcs in the BNs and model parameters. This initial knowledge is represented in a form of prior probability distribution and updated using data to obtain a posterior probability distribution over models and parameters.

Assume that a prior distribution over models structure is P(M) and a prior distribution over parameters for each model is P(p | M), a data set D is used to form a posterior distribution over models using Bayes Rule [14]:

$$P(M \mid D) = \frac{\int P(D \mid p, M) P(p \mid M) dp P(M)}{P(D)}$$

which integrates out the uncertainty in the parameters. For a given model structure, the posterior distribution over the parameters is computed by:

$$P(p \mid M, D) = \frac{P(D \mid p, M) P(p \mid M)}{P(D \mid M)}$$

The authors DBNs model used information of distractions to define driver cognitive state either distracted or distracted respectively as the hypothesis node. The evidences used in this BN models will sometimes called as instances. The authors BN model contained 24 nodes of instances measured either with discrete or continuous values. Prior knowledge for this study is, whenever a driver is distracted, some instances will produce a measurement to indicate a movement. In DBNs, it is not necessary for every instance to produce similar measurement at every time slice, t, but they are conditionally dependent to each other. The prior distribution is possibly effect the posterior distribution. Thus, all instances are checked on every time slice to determine whether the driver is cognitively distracted or not.

Initial Bayesian Network Model for this study is shown below as in Static Bayesian Network.

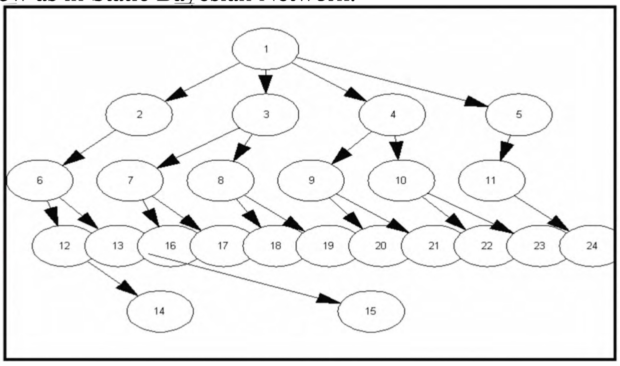

Figure 1: Static Bayesian Network (SBN)

As shown, there are 24 nodes in the model. The 24 nodes represented as:-

TABLE 3: BAYESIAN NETWORK NODES

1	Cognitive distraction	2	Visual physical
3	Face	4	Eyelid
5	Head	6	Pupil
7	Lips	8	Mouth open
9	Eyebrows	10	Eye open
11	Track state	12	Gaze quality
13	Pupil diameter	14	Gaze rotation
15	Pupilary response	16	Inner contour
17	Outer contour	18	Mouth width
19	Mouth height	20	Left eye
21	Right eye	22	Eye width
23	Eye height	24	Head rotation

SBNs consider instance and beliefs at a single time point. But DBNs can model the instances in a time series event. DBNs connect two SBNs that are identical in structure from two successive time slices. The state of a node at time t is conditionally dependent on the nodes in both the previous time slice *(t-1)* and current time slice *(t)*. This explanation defined the essential issue of building a DBN model.

Figure 2 is a DBN model used in this study.

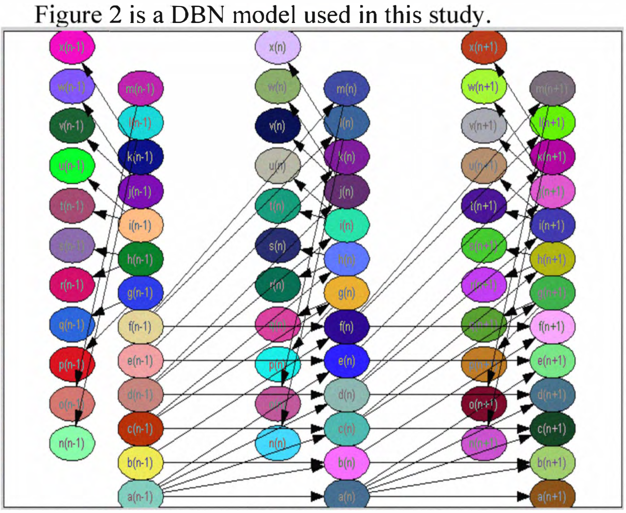

Figure 2: Dynamic Bayesian Network (DBN)

IV. CONCLUSION

It is shown from this study that mouth movements and eye directions/movements are new features that are probably useful to indicate a driver is distracted cognitively or not. Psychology study has proved that when a person is thinking, their mouth and eyes are moving, thus, in this paper it has been justified with the experimental results that mouth and eyes are correlated to each other when a person is thinking or cognitively distracted. Two important conclusions from this study are: (1) mouth and eye movements are highly correlated to each other; and (2) Right eye is more correlated to mouth movement either from eye's height or width compared to the left eye. Data collected from the faceLab Seeing Machine is only a simulated study and attempts are being made to install a real time face tracking system in a car for a road environment. The effect of eyebrow and lips in detecting a driver's cognitive distraction will also be investigated in the future.

REFERENCES

[1] *Concentration.* Retrieved February, 2009, from http://www.webster-dictionary.org/definition/concentration

[2] M.Baumann, J. F. K. (2007). Situation awareness and driving: A cognitive model. In Pietro C. Cacciabue (Ed.), *Modelling driver behaviour in automotive environments: Critical issues in driver interactions with intelligent transport systems* (illustrated ed., pp. 253) Springer.

[3] Graham Hole. (2007). Perceptual factors in driving. *The psychology of driving* (pp. 20) Lawrence Earlbaum Associates.

[4] Masahiro MIYAJI, Mikio DANNO, Haruki KAWANAKA, Koji OGURI. (2008). Driver's cognitive distraction detection using AdaBoost on pattern recognition basis. *Proceedings of the 2008 IEEE International Conference Vehicular Electronics and Safety,* Columbus, OH, USA. 51.

[5] Yulan Liang, John D. Lee, Michelle L. Reyes. (2007). Nonintrusive detection of driver cognitive distraction in real time using bayesian networks. *Transportation Research Record: Journal of the Transportation Research Board.*

[6] Yulan Liang, John D. Lee, Michelle L. Reyes. (2007). Real time detection of driver cognitive distraction using support vector machine.*8, No.2*(IEEE Transactions on intelligent transportation system), 340.

[7] Wang Rongben, Guo Lie, Tong Bingliang, Jin Lisheng. (2004). Monitoring mouth movement for driver fatigue or distraction with one camera. IEEE Intelligent Transportation Systems, Washington, D.C, USA. (IEEE Intelligent Transportation Systems) 314.

[8] *faceLab and faceAPI from seeing machines.* (2009). Retrieved August, 2009, from http://www.seeingmachines.com/product/facelab/

[9] Jane C. Stutts,Donald W. Reinfurt,Loren Staplin,Eric A. Rodgman. (2001). THE ROLE OF DRIVER DISTRACTION IN TRAFFIC CRASHES No. 202/638-5944) AAA Foundation for Traffic Safety.

[10] Blalock, H. M. (1972). Social statistics. New York: McGraw-Hill.

[11] Nancy Leech, Karen Barrett, George A. Morgan. (2007). *SPSS for intermediate statistics* Psychology Press.

[12] Richard Bandler, J. G.,. (1979). Frogs into princes: Neuro linguistic programming Real People Press.

[13] Kevin Murphy. (31 October 2005.). Software packages for graphical models / bayesian networks. http://www.cs.ubc.ca/~murphyk/Bayes/bnsoft.html

[14] Zoubin Ghahramani. (1997). *Learning dynamic bayesian networks*