ANALYSIS OF DRIVER BEHAVIORS DURING COMMON TASKS USING FRONTAL VIDEO CAMERA AND CAN-BUS INFORMATION

Jinesh J Jain and Carlos Busso

Multimodal Signal Processing Laboratory (MSPLab)
The University of Texas at Dallas, Department of Electrical Engineering
Richardson, TX - 75080, USA
jinesh.jain@utdallas.edu, busso@utdallas.edu

ABSTRACT

Even a small distraction in drivers can lead to life-threatening accidents that affect the life of many. Monitoring distraction is a key aspect of any feedback system intended to keep the driver attention. Toward this goal, this paper studies the behaviors observed when the driver is performing in-vehicle common tasks such as operating a cellphone, radio or navigation system. The study employs the UTDrive platform – a car equipped with multiple sensors, including cameras, microphones, and Controller Area Network-Bus (CAN-Bus) information. The purpose of the analysis is to identify relevant features extracted from a frontal video camera and the car CAN-Bus data that can be used to distinguish between normal and task driving conditions. Statistical hypothesis tests are used to assess whether the differences observed in the selected features are significant. Then, these features are used in binary classification tasks (normal versus task). For most of the considered tasks, features extracted from the frontal video camera are found to be the most prominent indicators to distinguish between normal and task driving conditions (e.g., head pitch and yaw). The features from the car CAN-Bus data slightly improve the classification accuracy, from 76.7% (using features only from the frontal video) to 78.9% (using all features).

Index Terms— Driver behavior, In-vehicle environment, Visual Distraction, CAN-Bus car information, Real world driving database.

1. INTRODUCTION

Distraction and fatigue in drivers is one of the main causes of road accidents. According to The National Highway Traffic Safety Administration (NHTSA), over 25% of police-reported crashes involved inattentive drivers [1]. The 100-car Naturalistic Study concluded that over 78% of crashes and 65% of near crashes involved some form of inattention [2]. Driver attentiveness is important at all times as any small mistake can lead to everlasting consequences, not just to the passengers in the car but also to other people. Monitoring driver distraction and fatigue can help to reduce the number of accidents, improving the overall driving experience. This paper analyzes the driver behaviors observed during common activities such as operating a radio, Global Positioning System (GPS) and cellphone that can affect the attention level of the driver.

An "on-road driver in a car" platform provides different sources of information that can be used to estimate the driver behaviors. The information can be clustered in two main groups: information about the driver, and information about the vehicle. Driver information includes head pose and eye gaze information to estimate the field of concentration, and eye closure information to indicate attentiveness [3]. Recent studies have considered the use of monocular, infrared (IR) and stereo cameras to track these behaviors. Invasive sensors such as EEG, ECG, EOG have also been proposed to estimate biometric signals [4, 5, 6]. Likewise, getting information about the car is another part of the puzzle. Important driving behaviors can be inferred by analyzing the steering wheel movements, vehicle speed and braking dynamics. In previous work, this information has been obtained from CAN-Bus data obtained either in simulated environments or in real world driving platforms [7, 8, 9, 10]. Most of these cues are directly affected by changes in the driver behaviors, especially in situations when they get distracted. Therefore, studying the patterns observed in the CAN-Bus data can lead to new metrics to detect the driver's attention level. While the driver cues and CAN-Bus data have being separately analyzed, we hypothesize the potential benefits of combining these complementary sources of information to infer the attention level of the driver.

Our long term goal is to detect relevant patterns observed from the driver and the car that can be employed to quantify the attention level of the driver. Towards this goal, this paper studies the driver behaviors observed during common tasks such as programming a GPS, changing the radio, using a cellphone, and engaging in spontaneous conversation with another passenger. The proposed work derives its novelty from the fact that the analysis is being done on both driver video data and car CAN-Bus information. The database is collected with actual drivers in a real world scenario using the UTDrive platform. Multiple sensors are used in the car to obtain the required car information. Head pose and eye closure information is extracted from a video from a frontal camera, which is processed using the AFECT software [11]. This information is then analyzed to identify prominent features that can be employed to classify between different tasks and normal driving. The analysis includes testing whether the selected features are

978-1-61284-350-6/11/$26.00 ©2011 IEEE

significantly different, using matched pairs t-test. Binary K-Nearest Neighbor classifiers are trained to detect between task and normal conditions using the selected features.

The result of the analysis indicate that Head - Yaw mean, Head - Pitch Mean, Eye Closure and Vehicle Speed are the most prominent features. The results of the classification task show that features derived from the frontal video camera and the car CAN-Bus can be employed to distinguish task and normal conditions with high accuracy. Visually intensive distraction tasks such as observing pictures, operating a radio, a GPS, and a phone are detected with accuracies over 74% using features derived from frontal video camera. When the classifiers are also trained with CAN-Bus data (multimodal approach), the accuracy improves 2.2% (absolute).

The paper is organized as follows. Section 2 gives an overview of the experiment setup. It describes the protocol and the tasks performed during the collection of the database. Section 3 describes the extraction and preprocessing of the data. It gives details about the different features derived from both modalities (frontal video and CAN-Bus data). Section 4 presents the analysis and statistical tests on the selected features. Section 5 describes classification experiments that aim to assess the discriminant power of the features. Section 6 concludes the paper with the discussion, final remarks, and future directions of our research.

2. EXPERIMENT

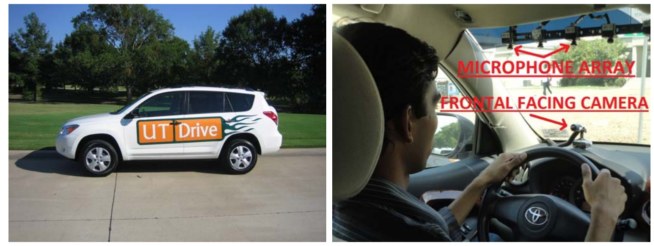
The study involves 8 subjects consisting of university students and employees. The database was collected in dry days with good light conditions to reduce the impact of environment in the study. The UTDrive platform and the protocol used during the recording are described in this section. The readers are referred to [12] for further details about the UTDrive project.

2.1. UTDrive

A 2006 Toyota RAV4 is used as platform for the UTDrive project (Fig. 1-a). The UTDrive car provides a perfect platform towards understanding the driver actions in relation to his/hers attention and fatigue levels. It is custom fit with data acquisition systems with multiple sensors including camera and microphone arrays. The system also records CAN-Bus data which provides the car driving information such as brake, gas, acceleration, vehicle speed and steering wheel information. Key to this work, is the video being recorded from a frontal camera (PBC-700H) facing the driver (set up right on the dashboard behind the steering wheel - Fig. 1-b). The video is recorded at 320x240 resolution at 30fps. There is also a camera facing the road ahead with potential applications in the future. This camera also records at 320x240 resolution but at 15fps. The GPS is fixed at the center of the front windshield. The information is simultaneously stored in a Dewetron computer.

2.2. Protocol

The protocol for the experiment is as follows. A route, shown in Figure 2, lasting about 5.6 miles is decided upon, starting and ending at the university premises. The route includes



(a) UTDrive (b) Setup

Fig. 1. UTDrive Car and sensors Setup

Fig. 2. Route used for the recording (5.6 miles long). Subjects drove this route two times. In the first lap, the subjects performed the tasks in order, starting with the *Radio* task and ending with the *Conversation* task. In the second lap, the subjects drove the same route without performing any task

many traffic lights, stop signs, heavy and low traffic zones, residential areas and also a school zone. The data collection per subject involves two runs of driving around the same route. The overall run for each lap takes about 12 - 16 minutes. In the first lap, the subjects were asked to perform the following tasks:

- Operating the in-built car radio (Fig. 2, red route): The *Radio* task involves changing the radio channels to target stations.
- Operating and following instruction from the GPS (Fig. 2, green route): This requires the driver to input a specific predecided address into the GPS and then follow the instructions to the desired destination. This task is subdivided into *GPS - Operating* and *GPS - Following* (preliminary results suggested that driver behaviors were different for these two activities).
- Operating and talking on the phone (Fig. 2, navy blue route): Here, the driver is asked to make a phone call from his/her cell phone to an airline automatic flight information system to obtain flight information between two US cities. This task is also subdivided into *Phone - Operating* and *Phone - Talking* (similar reason from GPS task).
- Describing pictures (Fig. 2, orange route): The driver is asked to look at randomly selected pictures and describe them. This gives exaggerated data from distractions such as billboards, sign boards and shops.
- Conversation with a passenger (Fig. 2, black route): The

last task is a spontaneous conversation between the driver and a second passenger in the car (the first author of the paper). The driver is asked few general questions in an attempt to get the driver involved in a conversation.

After subdividing the phone and GPS tasks in 2, there are 7 tasks considered in the analysis.

The second lap involves normal driving without any of the aforementioned tasks. The data collected from this lap is used as normal reference. Since the same route is used for both normal and task conditions, the analysis is less dependent on the selected road.

3. DATA EXTRACTION AND PREPROCESSING

The data collected has two parts: the files from the Dewesoft software, which provides the car CAN-Bus information, and the synchronized video files. The data is extracted in two steps: The Dewesoft software provides easy means to select and export required data into MATLAB files. The video of the driver is processed with the software AFECT [11].

3.1. CAN-Bus Information

The exported MATLAB files include all the CAN-Bus information such as steering wheel angle, brake value and brake pedal pressure values, vehicle speed, acceleration in RPM and gas pedal pressure. Among these variables, the steering wheel angle, brake and gas pedal pressures and vehicle speed are imported into MATLAB to be used as features for the analysis.

The study considers the jitter in the steering wheel, which is obtained as the sequence of variance over a period of 5 sec frames on the steering wheel angle signals. This along with brake pedal and gas pedal information are directly affected by driver behavior. We hypothesize that a driver involved in a task will produce more "jittery" behaviors (steering wheel movements, brake and gas pedal pressures). Therefore, the derivatives of the brake and gas pedal pressure features are computed (difference equation). Then the mean and standard deviation is estimated. These statistics are considered as features for the analysis. Vehicle speed is also included as a possible feature, since it is hypothesized that drivers tend to drive at different speed when involved in a task.

3.2. Information from frontal video

The AFECT software uses the video from the frontal camera as input. It provides various information from the driver such as head pose information, eye closure and emotion information (action units) [11]. It uses the algorithm described in [13] for head pose estimation. The algorithm is shown to be robust for large datasets and different illumination conditions. The results are imported into MATLAB.

Among all the features provided by AFECT, we consider head pose information, represented by the yaw and pitch angles, and the eye closure count. Since most of the tasks that are being studied require visual involvement, it is expected that head yaw and pitch will be more effected by the tasks. Thus, head roll movements are not included in the analysis.

The analysis also considers the eye closure information. Percentage of eye closure is defined as the percentage of

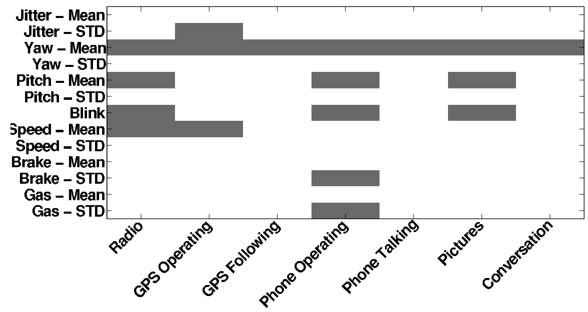


Fig. 3. Results of the Matched Pairs t-test: Features vs Tasks. For a particular task, gray regions indicate the features that are found to have significant differences (p-value = 0.05)

frames in which the eyelids are lowered below a given threshold. This threshold is set at the point where the eyes are looking straight at the frontal camera.

The AFECT software has its limitations and advantages. Some information is lost when the head is rotated beyond a certain degree or when the face is occluded by the driver' hands. The algorithm produces empty data in those cases. However, one of the primary advantages of AFECT is that the estimation is done frame by frame. This is important as the errors do not propagate across frames.

3.3. Preprocessing

During stops, the drivers may choose to perform various activities such as look around, which are not related to driving activities. These behaviors can adversely affect the analysis. Since the goal of the study is to analyze the driver behaviors when the car is moving, the data when the car is not moving or closing to stop is not included (if speed is less than 5kph). Likewise, for simplicity the analysis includes only data when the vehicle is moving straight (avoiding most of the segments in which AFECT does not provide information). Therefore, segments are discarded when the car is making turns greater than 20 degrees. The CAN-Bus information about the vehicle speed and the steering wheel are used as indicators, respectively. Any remaining gap in the data due to face rotation or hand obstruction while driving is interpolated. This procedure is performed for the data collected during the normal and task conditions.

4. ANALYSIS OF FEATURES

This section analyzes prominent features that can be used to detect the different common tasks considered in Section 2. The proposed approach consists in analyzing the differences in each feature between task and normal conditions. For the normal condition, only the data collected in the corresponding route for the target task is considered. This eliminates the route conditions factor introduced in the analysis. For example, route segments have different speed limits.

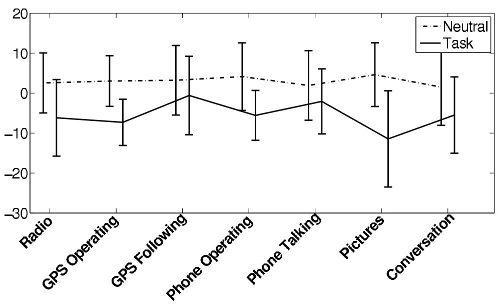
A Matched Pairs hypothesis test is estimated to assess whether the differences in the features between each task and the corresponding normal condition are significant. Since the current version of the database contains only 8 subjects, a t-test for small sample is used. Figure 3 shows the features

Table 1. Percentage of eye closure in task and normal conditions. The values for normal conditions are estimated over the features observed in the corresponding route of the tasks

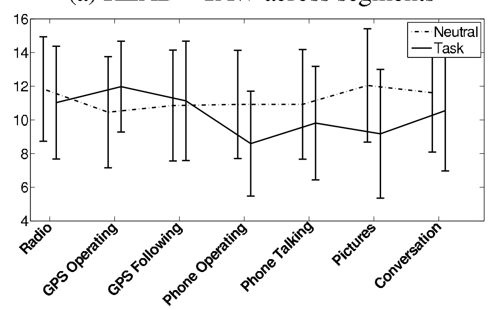
	Radio	GPS Operating	GPS Following	Phone Operating	Phone Talking	Pictures	Conversation	Mean across tasks
Neutral	78.52	67.06	73.57	74.69	84.94	80.01	75.58	76.34
Task	65.90	70.47	70.65	44.05	84.47	65.46	73.37	67.77

Table 2. Number of samples for normal and task conditions after segmenting the data in 5 sec samples

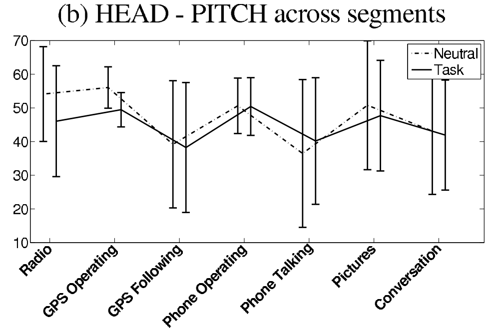
	Radio	GPS Operating	GPS Following	Phone Operating	Phone Talking	Pictures	Conversation
Neutral	47	28	91	11	87	48	50
Task	57	32	92	11	76	53	52



(a) HEAD - YAW across segments



(b) HEAD - PITCH across segments



(c) VEHICLE SPEED across segments

Fig. 4. Error-bar plots displaying the mean and standard deviation for: a) Head - Yaw b) Head - Pitch c) Vehicle Speed

that are found significant (p-value = 5%). The gray regions indicate the features that are significantly different. The figure shows that Head Yaw - Mean presents significant differences for each of the 7 tasks considered in the experiments. The features Head Pitch - Mean, Blink and Vehicle Speed - Mean are also significantly different for some of the tasks. Some features such as Brake - STD (standard deviation) and Gas - STD are prominent for tasks such as *Phone - Operating*. The figure also shows that there are tasks such as *GPS - Following*, *Phone - Talking*, and *Conversation* in which few of the selected features presents significant differences. This result suggests that the behavior of the driver may not be significantly affected by these tasks. The discriminant analysis presented in Section 5 confirms this result.

Figure 4 gives the error-bar plots for Head - Yaw, Head - Pitch and Vehicle Speed. These plots are calculated across all subjects. Figure 4-a shows that head yaw patterns change when the driver is performing each of the tasks. This is clearly observed for tasks that require to operate a device (phone, radio or GPS). These results suggest that these tasks deviate the focus of the driver from the road. Similarly, Figure 4-b and Figure 4-c show the error-bar plots for head pitch patterns and the vehicle speed, respectively. The figures show that for some of the tasks the features show differences. The Pitch - Mean increases when the driver is operating the GPS. This result is directly related to the placement of the GPS device in the car. Figure 4 also shows differences in the features during normal conditions across tasks. These differences are inherently dependent on the road. This result suggests that the characteristic of the route is an important variable that should be considered in the design of automatic feedback systems.

Table 1 shows the percentage of eye closure for the normal and task conditions. It can be seen that the closure rates differ from the patterns observed during normal condition for tasks such as *Radio*, *Phone - Operating*, *GPS - Operating* and *Pictures*. This result suggests that this feature is relevant and should be included in the classification task presented in the next section.

5. BINARY CLASSIFICATION

This section analyzes the discriminative power of the features by conducting classification experiments. For each of the tasks, a binary classifier is built to distinguish between normal and task conditions. Notice that a multi-class classification problem is not considered in this preliminary analysis, since the goal of the paper is to identify prominent features that can be used as metric to assess the driver attention. Although the classification task is dependent on the route, the experiments will inform valuable information to design a system to track the behavior of the driver.

The data is segmented into 5 sec windows, which are labeled according to the task performed by the drivers. Each of these segments is considered as an independent sample. For the normal condition, only the samples collected in the corresponding route of the target task are considered (i.e., the set of normal samples for different routes is different). The goal of the binary classifier is to assign a label – normal or task – to each of these 5 sec segments. Table 2 gives the total number of samples that are available in the database. Segmenting the data in 5 sec windows yields a reasonable number of samples per task. For *Phone - Operating* and *GPS - Operating*, the database contains less samples, since the duration of these tasks is shorter.

Forward feature selection (FFS) is used to determine the best features for each of the tasks. Features were selected one-by-one until the binary classification performance stopped increasing. The procedure is separately applied for each set of features (frontal video and CAN-Bus data), and for both sets of features. The binary classifiers are trained with the K-Nearest Neighbor algorithm, k = 5 (empirically chosen based on performance). Since the size of the database is relatively small, the experiments are implemented using "leave-one-out" cross validation. This approach ensures that the results are driver independent (testing data contains only the features from the driver that is not included in the training set).

The average results across the drivers are reported in Table 3. The accuracy (*Acc*), precision (*Pre*), recall (*Rec*) and F - score (*2*Pre*Rec/(Pre + Rec)*) of the classifiers are reported when the following sets of features are used: features from the frontal video camera, features from the CAN-Bus data, and features from both sets of features. This table is estimated with the best features selected by the FFS algorithm for each condition (task and feature set). For precision and recall, the "task" condition is considered as the relevant class. The baseline used in the experiment corresponds to the case in which the label of the class with higher prior is assigned to all the samples. The Δ improvement is calculated as the difference between the accuracy of the classifiers using all the features, and the best accuracy of the system trained with a single feature set (frontal video or CAN-Bus data).

Table 3 shows that the average accuracy across tasks is 76.7% when the features derived from the frontal video camera are used. The average accuracy of the classifier with CAN-Bus features is similar (76.5%). When all the features are combined, the average accuracy increases to 78.9%, which represents a 2.2% improvement (absolute) over the systems separately trained with either modalities.

Table 3 shows that accuracies higher than 81% can be achieved for the tasks *Radio*, *GPS - Operating*, *Phone - Operating* and *Pictures*. These results are explained by the highly visually intensive nature of these tasks. For other tasks such as *GPS - Following*, *Phone - Talking* and *Conversation*, the accuracies of the classifiers are lower. These results suggest that the behaviors observed from these tasks do not significantly differ from the behaviors observed during normal driving conditions (at least in the analyzed features).

For the tasks *GPS - Operating* and *Phone - Talking*, the use of CAN-Bus data slightly degrades the performance of the multi-modal classifiers. This result shows that the driving style was preserved during these tasks (speed, braking and gas pedal dynamics). However, these features are found to be useful for other tasks such as *Radio*, *GPS - Following Phone - Operating* and *Conversation*.

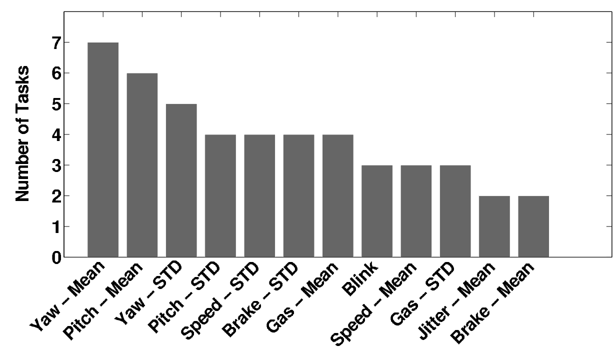


Fig. 5. Frequency in which the features are selected by the *Forward Feature Selection* (FFS) algorithm across tasks (e.g., Brake - Mean is selected in four out of seven tasks)

Figure 5 shows the number of time that each feature was selected by the FFS algorithm across the 7 tasks. This figure is presented when both sets of data are used in the binary classifiers. Features are included only when they increase the overall performance of the system. Therefore, this figure indicates how useful the features are in the binary classifiers. The figure shows that the mean of Head Yaw is a prominent feature, since it is selected in each of the tasks. Although the top 4 features are from the frontal video camera, the figure shows that CAN-Bus features are important. The fact that they are frequently selected indicates that they provide complementary information. Therefore, both feature sets should be included in the analysis of driver behaviors.

6. CONCLUSIONS

Any fall in the attention level of a driver can lead to disastrous consequences. An attempt has been made to collect a real world database consisting of real driving data. The corpus is a comprehensive multi-modal database consisting of videos, audio and car CAN-Bus information. The subjects were asked to complete different tasks that are commonly performed by drivers such as operating a GPS and a phone. The driver behaviors extracted from a frontal video camera and the CAN-Bus data were compared with the patterns observed during normal conditions. Hypothesis test, error-bar plots and classification tasks were analyzed to determine the relevant features that are informative of the considered tasks.

The results indicate that the data obtained from the frontal video camera is useful to distinguish between normal and task conditions (especially the mean of the head yaw motion). For the tasks considered in the experiments (which are mainly visually intensive), the features from the CAN-Bus data provide a small but significant improvement. This finding shows that looking at both the frontal video camera and the car CAN-Bus data gives complementary information that can be potentially used to track the attention level of the driver.

The work presented in this paper is our first step towards obtaining a metric to determine the attention level of the

Table 3. Accuracy, Precision, Recall, F-Score for features from frontal video camera, CAN-Bus data and both modalities. For the baseline, the labels are assigned to the class with higher prior. The table also gives the Δ improvement achieved when the multimodal classifiers are used over unimodal classifiers

	Features from frontal video camera				Features from CAN-Bus				All Features				Baseline	Δ Impro.
	Acc	Pre	Rec	F - Score	Acc	Pre	Rec	F - Score	Acc	Pre	Rec	F - Score		
Radio	0.886	0.905	0.845	0.868	0.896	0.894	0.863	0.874	0.910	0.903	0.918	0.904	0.556	0.013
GPS - Operating	0.929	0.975	0.885	0.906	0.898	0.975	0.813	0.857	0.916	0.975	0.854	0.894	0.534	-0.013
GPS - Following	0.628	0.680	0.608	0.605	0.629	0.691	0.589	0.609	0.635	0.649	0.647	0.626	0.540	0.007
Phone - Operating	0.740	0.714	0.750	0.889	0.740	0.714	0.750	0.889	0.813	0.857	0.875	0.905	0.542	0.073
Phone - Talking	0.636	0.702	0.779	0.697	0.570	0.594	0.793	0.656	0.591	0.612	0.774	0.663	0.579	-0.045
Pictures	0.918	0.935	0.918	0.921	0.906	0.906	0.913	0.901	0.918	0.935	0.918	0.921	0.530	0.000
Conversation	0.632	0.652	0.628	0.629	0.719	0.755	0.669	0.696	0.742	0.758	0.710	0.719	0.539	0.023
Mean across tasks	0.767	0.795	0.773	0.788	0.765	0.790	0.770	0.783	0.789	0.813	0.814	0.805	0.546	0.022

drivers. We are currently collecting data from more subjects to increase the size of the database, which will be used to validate the results presented in this paper. We are interested in identifying relevant features from different modalities. We believe that other car CAN-Bus features can help to increase the performance of the classifiers. Likewise, we are planning to analyze the acoustic signal collected from the microphone array. For example, the energy of the audio can be used to distinguish between task and normal driving condition for activities that are not visually intensive (i.e., *Phone - Talking* and *Conversation*). Also, we are looking at other sensing technologies that are invariant to illumination changes (IR).

Another direction that we are considering is to estimate a single multi-class classifier to recognize relevant non-driving tasks that can affect the attention level of the drivers. For example, detecting the proposed tasks can be used as a midlevel representation to detect driver distractions. Ideally, this classifier should be independent of the road. A real-time algorithm with such capabilities will have an impact on designing a feedback system to alert the driver when attention level falls, preventing potential accidents, and therefore, improving the overall driver experience.

Acknowledgment

The authors would like to thank Dr. John Hansen for his support with the UTDrive Platform. We want to thank the Machine Perception Lab (MPLab) at The University of California, San Diego for providing the AFECT software. The authors are also thankful to Ms. Rosarita Khadij M Lubag for her support and efforts with the data collection.

7. REFERENCES

[1] T.A. Ranney, W.R. Garrott, and M.J. Goodman, "NHTSA driver distraction research: Past, present, and future," 2000, Driver Distraction Internet Forum.

[2] V.L. Neale, T.A. Dingus, S.G. Klauer, J. Sudweeks, and M. Goodman, "An overview of the 100-car naturalistic study and findings," 2005, available at "http://www-nrd.nhtsa.dot.gov/pdf/nrd-01/esv/esv19/05-0400-W.pdf".

[3] Y. Dong, Z. Hu, K. Uchimura, and N. Murayama, "Driver inattention monitoring system for intelligent vehicles: A review," in *IEEE Intelligent Vehicles Symposium*, Xi'an, China, June 2009, pp. 875 –880.

[4] C.T. Lin, R.C. Wu, S.F. Liang, W.H. Chao, Y.J. Chen, and T.P. Jung, "EEG-based drowsiness estimation for safety driving using independent component analysis," *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 52, no. 12, pp. 2726 – 2738, Dec 2005.

[5] C. Berka, D.J. Levendowski, M.N. Lumicao, A. Yau, G. Davis, V.T. Zivkovic, R.E. Olmstead, P.D. Tremoulet, and P.L. Craven, "EEG correlates of task engagement and mental workload in vigilance, learning, and memory tasks," *Aviation, Space, and Environmental Medicine*, vol. 78, pp. B231–B244(14), May 2007.

[6] I.G. Damousis and D. Tzovaras, "Fuzzy fusion of eyelid activity indicators for hypovigilance-related accident prediction," *IEEE Transactions on Intelligent Transportation Systems*, vol. 9, no. 3, pp. 491 –500, Sept 2008.

[7] P. Angkititrakul, M. Petracca, A. Sathyanarayana, and J.H.L. Hansen, "UTDrive: Driver behavior and speech interactive systems for in-vehicle environments," in *IEEE Intelligent Vehicles Symposium*, Istanbul, Turkey, June 2007, pp. 566 –569.

[8] A. Perez, M.I. Garcia, M. Nieto, J.L. Pedraza, S. Rodriguez, and J. Zamorano, "Argos: An advanced in-vehicle data recorder on a massively sensorized vehicle for car driver behavior experimentation," *IEEE Transactions on Intelligent Transportation Systems*, vol. 11, pp. 463 –473, June 2010.

[9] E. Murphy-Chutorian and M.M. Trivedi, "Head pose estimation and augmented reality tracking: An integrated system and evaluation for monitoring driver awareness," *IEEE Transactions on Intelligent Transportation Systems*, vol. 11, no. 2, pp. 300 –311, June 2010.

[10] A. Giusti, C. Zocchi, and A. Rovetta, "A noninvasive system for evaluating driver vigilance level examining both physiological and mechanical data," *IEEE Transactions on Intelligent Transportation Systems*, vol. 10, pp. 127–134, March 2009.

[11] M. Bartlett, G. Littlewort, T. Wu, and J. Movellan, "Computer expression recognition toolbox," in *8th IEEE International Conference on Automatic Face Gesture Recognition (FG 2008)*, Amsterdam, The Netherlands, Sept 2008, pp. 1 – 2.

[12] P. Angkititrakul, D. Kwak, S. Choi, J. Kim, A. Phucphan, A. Sathyanarayana, and J.H.L. Hansen, "Getting start with UTDrive: Driver-behavior modeling and assessment of distraction for in-vehicle speech systems," 2007.

[13] J. Whitehill and J.R. Movellan, "A discriminative approach to frame-by-frame head pose tracking," in *8th IEEE International Conference on Automatic Face Gesture Recognition, FG 2008.*, Amsterdam, The Netherlands, Sept 2008, pp. 1 –7.