DETECTING AND CLASSIFYING DRIVER DISTRACTION

Amardeep Sathyaranayana, Pongtep Angkititrakul, John H.L. Hansen

Center for Robust Speech Systems (CRSS)
Erik Jonsson School of Engineering and Computer Science
The University of Texas at Dallas, Texas, USA
{axs063000, angkitit, john.hansen}@utdallas.edu

ABSTRACT

With rapid advancements in technology, new devices such as voice interactive systems, navigation systems, hands-free mobile communications, and entertainment systems (e.g., IPOD etc.) have introduced a significant range of user control/demands within vehicles. Even though these applications assist the multi-tasking ability for the driver, they introduce a variety of distraction levels which could divert the driver's attention from the primary driving task. In this study, we describe and model driver's normal driving using vehicle's CAN-Bus signals and sensory data from the UTDrive corpus. Possible distracted driving behaviors of a given driver are identified and categorized as different distraction levels (i.e., low, medium, and high). As further advancements are made for in-vehicle systems, distraction classification should act as a recommendation for evaluating risk factors and recommendations to reduce accidents caused by these distractions.

1. INTRODUCTION

Today, society depends more on the mobility of people, with further advances for vehicle technology in terms of comfort, features and safety. These technological advancements have revolutionized the automotive industry, from a purely mechanical system to a microcontroller managed system. In general people spend more time in their cars today than ever before. For this reason, there is a need for further advancement of sub-systems for driver comfort and also to enable him to be multi-tasking while driving the car. (e.g., internet in the car, mobile communications, car navigation, etc.) There is no doubt that such technology provides improved support, comfort and flexibility to the driver but it can also be viewed as a source of distraction. Such technology could divert the attention of the driver from the road, hence acting as a source of distraction. Distraction could be defined as anything which diverts the attention of driver causing any deviation from the normal driving pattern. Causes of distraction can be broadly classified into visual, cognitive, biomechanical and auditory [1]. There is an assumption that the wide range of devices accessible to the driver may lead to an increase in the number of accidents due to driver distractions. Surveys suggest that mobile phones are a major source of driver distraction. Active research work is underway in a number of labs to find the sources of distraction and ways to mitigate their impact on driving performance. Studies have also shown that drivers can achieve better and safer driving performance while using speech interactive systems to operate in-vehicle systems compared to hand operated interfaces [2]. Though it provides a better interface, operating a speech interactive system will still divert a driver's attention away from his or her primary driving task with varying degrees of distraction. These distractions can have varied impact on normal driving patterns and could result in slight, severe or fatal accidents. With advancing technology, there is a need for intelligent vehicles which could help drivers operate vehicles safer. Such intelligent vehicles should be able to identify a driver, compare his current driving pattern with his normal driving pattern, and take necessary actions when there is a noticeable deviation from the normal driving pattern. This paper proposes detecting possible distractions and categorizing them as different distraction levels (i.e., low, medium and high).

A subset of the UTDrive Project [7] database is used to detect and classify driving under distractions. UTDrive project is part of an on-going international collaboration to collect and research rich multi-modal data recorded for modeling driver behavior for in-vehicle environments. The corpus consists of audio, video, gas/brake pedal pressure, forward distance, GPS information, and CAN-bus information collected across a large number of drivers [7].

The reminder of this paper is organized as follows. Section 2 describes the CAN-Bus signals and the signals which help define driving characteristics. Section 3 gives a generic overview of the different models which could be used in building driver behavior models. Section 4 gives the classification framework used in detecting and classifying driver distractions. Section 5 focuses on the nature of data used for the experiment and also the experiments performed. Section 6 gives the experimental results. It consists of an elaborate explanation of how the results where obtained and the interpretation of those results. Section 7 summarizes the paper with conclusion and future work.

2. DATA ANALYSIS

With the rapid advancement in the automotive industry, the need for advanced high speed operations in cars becomes necessary. This leads to the introduction of microcontrollers in cars performing specific operations. Data transfer between these microcontrollers is mainly accomplished using the Controller Area Network (CAN) protocol [6]. With a very low data error rate and high speed communication, data is transferred on twisted wire pair which is referred to as the CAN-Bus. This CAN-Bus data contains all the information about the current status of the vehicle, for example, engine speed, vehicle speed, vehicle acceleration, brake pedal pressure and position, engine temperature, etc. All the data on the CAN-Bus is transmitted as frames in binary format and are uniquely identified by the header. Without knowing the header, the CAN-Bus information cannot be used. Using the OBD 2 port, CAN-Bus signals can be read by the outside world.

Since detecting and classifying driver distraction is the main motive of this paper, the focus is on using CAN-Bus signals which are driver dependent. Signals like engine temperature, air bag status, oil level monitor, etc. give the current vehicle conditions but have nothing to do with the current driver condition. A driver's primary interaction with the car would be pressing the gas pedal, brake pedal and handling the steering wheel to direct the car. These signals could be used to define a driver's behavior or mood as these are directly in his control. For example say the driver is tensed, he could hit the gas pedal harder, or could move the steering wheel erratically.

Individually, these signals could change independently. Changes could be due to road traffic demand, but collectively it could give a better description of the driver's general behavior and general trends leading to distraction. Therefore, the signal obtained from the CAN-Bus contains the logistics of the driving route for the vehicle, as well as the subtle driver specifics on how the vehicle is operated. In a manner similar to speech recognition, the route information in the CAN-Bus signal is similar to the text of what someone speaks, the particular variations on how the driver achieves that route is similar to speech signal characteristics such as who is speaking and how they speak (emotion, accent, stress, etc.). This paper could therefore verify that a combination of such driver dependent data would be useful in driver behavior modeling and distraction detection.

3. DRIVER BEHAVIOR MODELING

3.1. CAN Raw Feature Vectors

The raw CAN signals of vehicle speed, acceleration, brake pedal pressure and steering wheel angle could define a driver's characteristic driving behavior. These raw signals are clubbed together on a sample by sample basis to form a 4 dimensional feature vector. This feature vector is then sliced according to requirements and used in building models.

3.2. Gaussian Mixture Model

A Gaussian Mixture Model (GMM) is a statistical model which models the distribution of feature vectors. These feature vectors could be extracted from any non deterministic signal [3]. A GMM estimates a probability density function using the expectation-maximization algorithm. The GMM built for neutral driving comprises the entire space which a normal driving pattern occupies. Hence a raw signal can be evaluated to match either a neutral or a distraction model.

3.3. Hidden Markov Model

A Hidden Markov model (HMM) is a statistical model which determines the hidden parameters from the observable set of parameters. The system being modeled is assumed to be a Markov process with unknown parameters. Any transitions in the parameters can also be traced using the HMM [4]. The extracted model parameters can then be used to perform further analysis such as matching driver patterns with neutral and distracted models.

3.4. Kullback–Leibler divergence

Kullback–Leibler (KL) divergence or KL distance is a measure of the differences between reference probability distribution P to an arbitrary probability distribution Q [5]. For discrete probability distributions, p={p_1, ..., p_n} and q={q_1, ..., q_n}, the KL-distance is defined to be

$$KL(p,q) = \sum_i p_i . \log_2 (p_i / q_i) \qquad (1)$$

4. CLASSIFICATION FRAMEWORK

4.1. Distraction Detection

Anything which diverts the attention of a driver that causes deviation from the normal driving pattern can be termed as distraction. The cause for distraction could be visual, cognitive, biomechanical and auditory. It could be as simple as seeing something interesting happening in the pedestrian area or it could be a temporary black out. Any form of a distraction has an impact on the driver resulting in some change in his driving pattern. This impact is reflected in small variations in the driver's normal driving pattern which can be detected using Gaussian Mixture Models (GMM) or Hidden Markov Models (HMM) models.

4.2. Distraction Classification

The impact of distraction could vary. The impact in terms of accidents could be classified as slight, severe or fatal accidents [1]. Based on how closely a given raw data can be detected as distraction or neutral driving, classification can be made as low, medium, and high distraction levels. For example if a given raw data can be detected as distraction data with a very low Equal Error Rate (EER), then the raw data is classified under high distraction level. Also the KL distance between the neutral driving model and distraction driving model gives a good indication of the level of distraction since it tells how far apart the two models are spaced from each other. If there is little difference between neutral and distraction models, then both the models are very close to each other and the KL distance is small. In such a scenario, there is a higher probability of getting a high EER value suggesting that there is not too much difference between the neutral driving and driving with distraction tasks.

5. EXPERIMENT

5.1. Data/ Route description

A subset of the UTDrive Project database is used to detect and classify driver distractions. Each driver is required to drive through the route twice which is shown in the Figure 1. The average time to complete the route is around 10 minutes and it includes a residential area and a commercial/school zone. The first run is neutral driving, where no distractions are introduced. In the second run, there are secondary tasks which the driver must perform while driving on the specific road. The tasks are labeled in the figure within red block beside the part of the route where that task is performed. There are 4 different tasks performed in the second run: conversation, lane changing, calling up an automated dialog system (commercial voice portal), and performing some common tasks like tuning the radio, reading license plates of other cars, adjusting the AC/heater levels, etc. Though lane changing and conversation are not major distractions, they have been included so that a validation check could be performed while detecting distractions. One point of consideration is that only a high level transcription is done on the entire route, splitting it into data specific on each major leg of the route. This transcription is done with the help of a visual aide. The driver might be distracted all the time on a specific leg, or it is possible he is distracted in only short bursts. Therefore, it may not be possible to successfully classify the entire leg where a secondary task is performed if the distraction only occurs for a small portion of the time.

Figure 1: Data collected on the route with labeled tasks

Further more in this paper the drivers are numbered sequentially from 1 to 8, and the tasks performed are labeled as LC (Lane Changing), CO (Passenger Conversation), MP (Using mobile phone for voice dialog/portal) and CT (Common Tasks).

5.2. Validation Method

Only a subset of the UTDrive project database has been used consisting of 8 different drivers completing the same route twice, once without any distraction and once with distraction. The training procedure used is N leave-one-out type (round-robin), where 7 drivers' neutral driving data and driving with distraction data was used for training and the remaining one driver's data is used for validation.

6. EXPERIMENTAL RESULTS

6.1. Route Dependent Models

The original data comprising of vehicle speed, vehicle acceleration, brake pedal pressure and steering wheel variations in degrees obtained from CAN bus is split into regions where the four different tasks were performed. Even the neutral driving data is split into these regions. Next, a 4 dimensional feature vector is built comprising of the 4 CAN-bus raw signals. Models are built based on the region of the route, and the data used for testing is also from the same region of the route but from the one driver set aside while training the model.

6.1.1 GMM and HMM comparison

The 4 dimensional feature vectors are used to train a 64 mixture Gaussian Mixture Models (GMM) for neutral and distraction driving using region specific data. The round-robin (N leave-one-out) training procedure is followed. The same data is used to train the neutral and distraction Hidden Markov Models (HMM) comprising 4 states and 16 mixtures in each state. From the one driver's data which was not used for training, 5 seconds long test data sets are created and tested against both GMM and HMM. Scores are obtained and tabulated in Table 1 and an Equal Error Rate is computed for both scenarios.

DISTRACTIONS	EER	
PARTS OF ROUTE (5sec)	GMM	HMM
LC	37.0474	44.8881
CO	52.8945	48.4345
MP	39.0074	35.5109
CT	37.4045	40.5619

Table 1: Comparison of EER from GMM and HMM

As mentioned earlier, lane change (LC) and conversations (CO) are not major distractions and using mobile phone (MP) and common tasks (CT) are distractions. From the table, it can be seen that the EER for MP and CT are smaller and that GMM and HMM show similar performance. Hence GMM is selected to be used for all further computations throughout the paper.

EER in this case can go above 50% since the entire length of the region need not be under distraction. Distraction could happen in small bursts. To show this, the KL distance is computed between the neutral GMM and distraction GMM and the results are tabulated in Table 2. The results show greater separation.

6.1.2 Distraction Detection

The optimum time needed for detecting a distraction is also computed. For this the same procedure as followed in computing EER for GMM is followed. The only change is that the feature vectors are split into 1 second, 2 seconds and up to 10 seconds of separate data, and all individual data segments are used to train and test the models. EER for all the cases are computed and an average EER over all drivers is plotted in Figure 2, with the optimum time for detecting a distraction obtained.

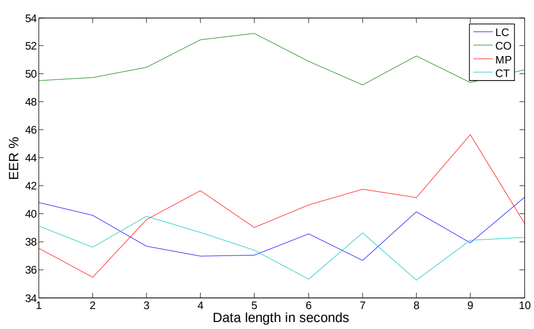

Figure 2: Average EER over all drivers from 1 second to 10 seconds long data.

Since lane change (LC) and conversation (CO) are not considered major causes of distraction, an evaluation of the results obtained on using mobile phones (MP) and common tasks (CT) show that the discrimination between various distraction and detecting distraction itself can be best made with 1 to 6 seconds worth of data and the average value residing at around 5 seconds data. Hence the optimum time duration segment for detecting distraction is considered to be 5 seconds.

6.1.3 Distraction Classification

KL distance is calculated between the neutral driving GMM and distraction GMM for a particular region of the route for every driver on a 5 second per frame data. The results are tabulated in Table 2.

KL	LC	CO	MP	CT
1	10.6436	16.1577	18.5362	19.0907
2	14.2011	14.6433	19.6111	15.5726
3	19.4781	17.4117	32.9928	16.8599
4	14.2380	14.9699	18.5042	17.9632
5	15.0899	13.0063	20.2903	18.7232
6	10.7808	14.8051	22.0906	20.7059
7	15.8742	30.2861	14.4747	25.8047
8	12.9468	12.4495	14.5173	12.7812
AVG	14.1566	16.7162	20.1272	18.4377
Result	NO	LOW	HIGH	MEDIUM

Table 2: KL distance between neutral and distraction driving GMM

From Table 2, it can be seen that the KL distance for lane changing and conversation tasks against their respective neutral driving GMMs is small. This indicates that these tasks generally do not distract the driver. Therefore lane change (LC) is classified as “NOT A DISTRACTION”. The KL distance for conversation suggests that even this does not cause much distraction. Hence conversation can be classified with a distraction level “LOW”. Common tasks like tuning the radio, adjusting the AC/heater level, checking if all doors are locked and windows are up, etc. actually cause distraction. It can also be seen that the common tasks are detected with reduced EER compared to others. These common tasks cause the driver to focus his attention on other secondary things and distract him from focusing on the road. Hence common tasks could be classified with distraction level “MEDIUM”. Mobile phone conversation is classified with distraction level "HIGH" and it is justified as the driver generally tends to focus more attention on the conversation than on the road ahead. The average KL distance is highest for neutral MP distraction model assessment, and also for 6 of the 8 drivers, confirming the consistency across subjects.

6.2. Route Independent Models

Similar to the route dependent model, route independent models are built by training all the data available for the entire route for all the drivers leaving one for test (N leave-one-out type). Two such models, neutral driving and distraction driving models are built each time. The raw data from a driver outside this model is tested against both the models and a decision is taken on whether it is distraction or neutral driving data. The average EER over all the drivers is tabulated in Table 3 below. Hence, we see that the average EERs are quite large for the four tasks.

EER	LC	CO	MP	CT
AVG	42.0015	47.3484	48.8894	41.20762

Table 3: Average EER over all drivers for whole route.

Instead of using EER for driver distraction, the sum of all the scores could be obtained, with testing on each leg of the route using route independent models. An average of these scores over all the drivers is tabulated in Table 4 below.

EER	LC	CO	MP	CT
Distraction	-9.998	-11.459	-11.946	-10.310
Neutral	-2.641	-9.2403	-6.7837	1.9773

Table 4: Average of sum of scores in particular region.

It can be observed that by setting appropriate thresholds the above tasks can be detected as different distractions.

7. CONCLUSIONS AND FUTURE WORK

Safety being the prime concern, recent research in the automotive industry has shown that much can be done to improve driving performance. Though the number of accidents per mile has been declining, distractions due to in-vehicle technology has raised concerns for road safety. There is a need for intelligent systems which could detect driver distractions and aide the driver in maneuvering the vehicle in a safe and comfortable manner. In this paper, an effort has been made to understand and detect distraction against neutral driving. Also, an effort has been made to classify distraction into low, medium and high risk categories. This is a useful step in building intelligent systems which can assist drivers under abnormal conditions. Knowing the level of risk, the system can either decide to give the driver some amount of freedom or take total control over the car by reducing the speed/ assisting brake control and calling for assistance. By evaluating the optimum time required to detect any distraction properly, better information can be given to the driver. As a consequence of early detection, a warning could be given to the driver before he moves into medium or high risk category of distraction, hence preventing any possible accidents. Further work will be done on building behavior models for each driver which not only detects distractions and make early predictions, but also learns and adapts with time to any variations in neutral driving patterns. Also, cognitive load on the driver under various tasks could be studied to better understand and model driver behavior and make better predictions.

REFERENCES

[1] Michael Pettitt, Gary Burnett and Alan Stevens, "Defining Driver Distraction," World Congress on Intelligent Transport Systems, San Francisco, November 2005.

[2] C. Carter and R. Graham, "Experimental Comparison of Manual and Voice Controls for the Operation of In-Vehicle Systems," in Proceedings of the IEA2000/HFES2000 Congress, Santa Monica, CA.

[3] D.A. Reynolds, "Speaker identification and verification using Gaussian Mixture Mixture speaker Models", Speech Communication, vol. 17, pp 91-108, 1995.

[4] Lawrence R. Rabiner, "A Tutorial on Hidden Markov Models and Selected Applications in Speech Recognition", Proceedings of the IEEE, 77 (2), p. 257–286, February 1989.

[5] Kullback, S., and Leibler, R. A., 1951, On information and sufficiency, Annals of Mathematical Statistics 22: 79-86.

[6] www.semiconductors.bosch.de/pdf/can2spec.pdf

[7] Pongtep Angkititrakul, Matteo Petracca, Amardeep Sathyanarayana, John H.L. Hansen, "UTDrive: Driver Behavior and Speech Interactive Systems for In-Vehicle Environments", IV 2007, Istanbul, Turkey, June 2007 (To be presented).