2013 IEEE Intelligent Vehicles Symposium (IV)
June 23-26, 2013, Gold Coast, Australia

Driver Classification and Driving Style Recognition using Inertial Sensors

Minh Van Ly[†], Sujitha Martin[†] and Mohan M. Trivedi[†]

Abstract— **Currently there are many research focused on using smartphone as a data collection device. Many have shown its sensors ability to replace a lab test bed. These inertial sensors can be used to segment and classify driving events fairly accurately. In this research we explore the possibility of using the vehicle's inertial sensors from the CAN bus to build a profile of the driver to ultimately provide proper feedback to reduce the number of dangerous car maneuver. Braking and turning events are better at characterizing an individual compared to acceleration events. Histogramming the time-series values of the sensor data does not help performance. Furthermore, combining turning and braking events helps better differentiate between two similar drivers when using supervised learning techniques compared to seperate events alone, albeit with anemic performance.**

I. INTRODUCTION

Recently there is interest in utilizing smartphones as data collectors because they are ubiquitous, small, and cheap. In addition, they have a myriad of embedded sensors and are, most importantly, an extension of the users it serves. This makes them ideal candidates for unobtrusive data collection for various research. In particular their sensors are ideal to monitor weather[14], traffic events[10], or user behaviors[17].

Though smartphones have a myriad of sensors, only a subset are useful when it comes to processing and interpreting certain events, such as car maneuvers. These sensors include: cellular signal strength, battery capacity, internal phone and external temperature, barometer, GPS, accelerometer, gyroscope, magnetometer, ambient light sensor, cameras, microphone as well as possible other sensors in future phones. Currently live video processing on the phone is taxing on the battery. Therefore, limiting or eliminating camera sensor processing serves best in these environments. Hence this paper will limit the features to using only inertial data.

Using inertial sensor data to understand driving style can help reduce dangerous driving. The authors in [5] show drivers react differently to different types of feedback stimuli when driving. For instance, an aggressive driver tends to follow instructions more often than a passive driver would. Therefore, identifying the person as well as their driving style is invaluable information in any effective vehicle feedback system.

Before the system can provide proper feedback, it needs to build a profile of the current driver. It needs to learn individual driver's driving habits. Not only does it need to learn this, but it also needs to differentiate between different drivers. In a typical family, there might be one or two additional driver for a single car. Therefore, driver differentiation, in this case, is a 2-class or a 3-class problem.

Though a simple camera system might be able to distinguish between different people quickly and efficiently, our system strides to distinguish the driver using only the inertial sensors and possibly other sensors widely available on a phone because sometime images, video, or voice data may not be available, too difficult, or taxing to use. Such situation where voice data is not available can arise during a phone call. Good video data can also be absent or noninformative if the phone's camera is incorrectly mounted. However, an inertial base system will enable a discrete, unobtrusive, and seamless identification of the driver.

Previous research have shown the potential in using inertial and orientation sensors to detect driving events. However, few have explored possible uses of these data such as uniquely identifying someone from their accelerating, braking and turning events in a vehicle. This paper explores how correlated the driver's actions are with the inertial sensors and shows whether those events are sufficient to differentiate between different drivers. This opens up the possiblity of using Smartphones and their inertial sensors to do similar tasks.

II. RELATED STUDIES

There have been studies on using inertial sensors as event detectors. Furthermore some have used them in identifying the person of interest.

A. Driving Event Detection and Classification using Inertial Sensors

Johnson et al [8] detected and classified driving maneuvers using a smartphone's accelerometer and gyro sensors mounted in the car. Subsequent research has focus on implementing these systems on a smartphone. Sathyanarayana et al [15] used SVM and other classifying techniques to show detecting driving events from smartphones using inertial sensors works equally well, if not better, compared to using the CAN signals from the vehicles. However these research only focus on driving event detection and classification. They stop short of using these data to classify different drivers.

B. Human identification using Inertial Sensors

In [11] and later in [16], they attempted to identify a person using inertial sensors. They detect and classify behavior and events and showed it's possible to separate the data into classes. However, these events are highly unstructured.

[†]Laboratory of Intelligent and Safe Automobiles, UCSD, CA 92092, USA
{m6ly, scmartin, mtrivedi} at ucsd.edu

A person's movement and actions can range wildly. Using a vehicle platform structures and constrains the data. A vehicle is limited to accelerating, braking, and turning events. Complex maneuvers can be a composition of these three basics events. For instance, a U-turn can be classified as two turning events.

This paper is structured in the following way. The first section talks about the data capture test bed and data used in the experiement. The second section gives a description of the dataset and how classification performance was done. The third section focuses on the results of these experiments. The fourth section concludes this paper with a few observations and future research ideas.

III. Experimental Setup

A. Experimental Testbed

A 2008 Volkswagen Passat called LISA-X collects all the data used in this research. This experimental vehicle is outfitted with many sensors and vision system. In particular, the sensors suite include: Front side radar, CAN signals from the car, such as engine speed, brake pressure, acceleration, pedal pressure, vehicle speed, angular rotation, headpose analysis, GPS, and Front/rear camera view. Fig. 1 shows the data capture vehicle. However the inertial sensors data is the focus of this research because these sensors are widely available on smartphone. In particular, LISA-X is equipped with Lateral (Left/Right), Longitudinal (Forward/Backward) Acceleration, and Yaw Angular Velocity sensors.

Fig. 1. LISA-X Testbed in Experimental Research

B. Path Characteristics

Since we are looking for distinct characteristics within data from different drivers, we need to look at generalize driving data. Eventually, this system needs to be able classify these drivers regardless of the path the driver takes. For instance, one person in a family might use the car for one purpose; hence the data explains a certain route characteristics. However, a different person in the household might use it for another purpose and take a totally different route resulting in different route characteristics. The system still needs to work in this scenario and needs to be route independent. Therefore, GPS data is not considered for this initial research. However, one can easily extend to use GPS data as another feature. Using both inner street driving and highway driving data generalizes driving data. Moreover, these captures are taken at different times of the day to get a variation of congested and non-congested driving. Two drivers are used in this research and we generated data using the test vehicle on multiple runs on both similar and different routes.

IV. Data Processing

Only inertial sensors, as well as other key events, are extracted from the time synchronized vehicle's sensors data during each capture. All types of vehicle maneuvers can be represented using three types of events: Braking, Acceleration, and Turning, (either right or left turn).

A. Ground Truth Data Association

These three types of events are simple to extract from the CAN data. Below are criteria defining each event. The start of a braking events is defined when the brakelight indicator turns from off (0) to on (1). The end of a braking event is defined as when BrakeLight indicator is 0 or when the vehicle speed is 0. This takes into account braking events when the end vehicle speed is not zero, i.e not just braking events at stop signs or red lights. Thus Braking events are defined as

$$\begin{array}{ll} \text{StartBrakeEvent:} & BrakeLight = 1 \\ \text{EndBrakeEvent:} & BrakeLight = 0 \quad \text{OR} \\ & |VehicleSpeed| = 0 \end{array} \tag{1}$$

To eliminate false braking events into our data, braking events shorter than 1 second are removed. Furthermore, braking events with zero velocity sections are shortened to include only the relevant part. These events happen when the driver is waiting at a red light or in heavy traffic.

The criteria for acceleration events is defined as:

$$\begin{array}{ll} \text{StartAccEvent:} & AccPedal > \epsilon_a \\ \text{EndAccEvent:} & AccPedal \leq \epsilon_a \end{array} \tag{2}$$

Similar to Braking events, micro-acceleration events less than 1 second in duration are removed. Here ϵ_a is 1.

Steering angle and vehicle speed are used to segment and extract turning events. The vehicle speed eliminates premature classification of turning event. For instance, if the driver only turns the steering wheel, but has not yet engaged the throttle to complete the turning event then this is not considered a turning event. Turning events are defined as:

$$\begin{array}{ll} \text{StartTurnEvent:} & |SteeringAngle| > \epsilon_t \quad \text{AND} \\ & VehicleSpeed > 0 \\ \text{EndTurnEvent:} & |SteeringAngle| < \epsilon_t \end{array} \tag{3}$$

Here ϵ_t is 30 degrees. Events less than 0.25 seconds in duration are excluded because those can be events in which the driver have not performed a valid turning maneuver but just turned the steering wheel beyond the threshold. This definition also allows for swerving maneuvers. Due to unknown reasons, some turning events do not have valid gyro data in them. These events are removed by looking at the gyro data themselves to verify they are nonzero. Therefore we have implemented automatic extraction and classifications of driving events based on these three definitions. Fig. 2 shows typical acceleration, braking and turning events, and their associated sensor outputs.

B. Event cues from Inertial Data

As stated in the previous section, the signals of interest are taken from the inertial sensors. The definition in (1), (2), and (3) only segments the inertial data. Whereas the data defining acceleration and braking events only use longitudinal accelerometer, data defining the turning events use both gyro and longitudinal accelerometer data. For turning events, longitudinal data provides insight on how hard the driver accelerated into or out of the turn. Lateral Accelerometer data can also be used, however, that information was corrupted in the CAN data for this particular vehicle. Table I includes the details of each event and the number of events per driver.

TABLE I
EVENTS AND CUES

Event Type	Event Cues	Number of Events for Driver A	Number of Events for Driver B
Acceleration	Acceleration	171	207
Braking	Acceleration	254	159
Turning	Yaw Angular Velocity & Longitudinal Acceleration	80	80

Fig. 3 shows how correlated the chosen inertial signals are with underlining truth signals (i.e. steering wheel position in the instance of turning events). This is shows the signals are correlated and the inertial sensor data are indeed another way to reflect the driver's acceleration, braking, and turning habits.

C. Feature Vectors

Histogramming the extracted time series vector [9] into 5 bins helps to reduce computation complexity because no resampling and interpolation are needed. Moreover, preliminary results using 10 bins did not yield any significant performance improvements. Using 5 bins histogram reduces the feature vectors size and helps alleviate over fitting when learning with limited data size. Typical signal statistics such as the min, max, mean, and variance are included in the feature vector. Additionally, the duration of the event is also included. The feature vectors for turning events include both the histogram of the both the Angular Velocity and the Longitudinal Acceleration. Though there are a myriad of other features we can append to these feature vectors, this paper explores this subset. See (4) and (5) for how the feature vectors are constructed.

$$AccFeatureVect/BrakeFeatureVect:$$

$$\begin{array}{l} \{ \\ \quad Histogram(AccYSignal, 5) \\ \quad Min(AccYSignal), Max(AccYSignal) \\ \quad Mean(AccYSignal), Variance(AccYSignal) \\ \quad Duration(AccYSignal) \\ \} \end{array} \tag{4}$$

$$TurningFeatureVect:$$

$$\begin{array}{l} \{ \\ \quad Histogram(GyroSignal, 5) \\ \quad Min(GyroSignal), Max(GyroSignal) \\ \quad Mean(GyroSignal), Variance(GyroSignal) \\ \quad Histogram(AccYSignal, 5) \\ \quad Min(AccYSignal), Max(AccYSignal) \\ \quad Mean(AccYSignal), Variance(AccYSignal) \\ \quad Duration(AccYSignal) \\ \} \end{array} \tag{5}$$

The duration of the GyroSignal in the turning events captures the duration for both the Gyro and Accelerometer data and thus is not repeated in the feature vector.

D. Datasets

The datasets are split into two types of test datasets. One is with the full feature vector including the histogram of signals of interest. This is described in (4) and (5). The other dataset includes selected statistics including min, max, mean, and variance of the feature vector. See Table II for details about all the test data. The '+' denotes concatenation of feature vectors.

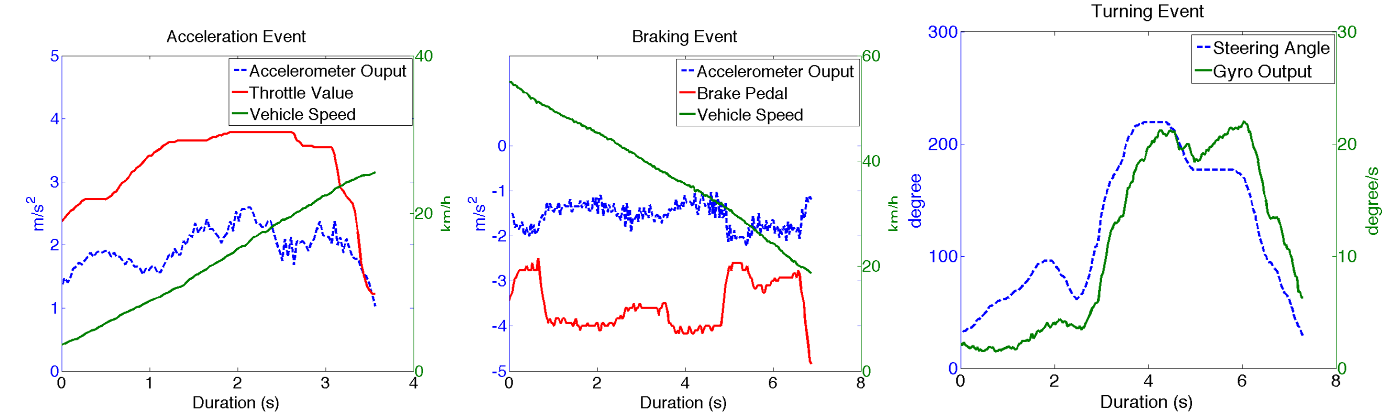
TABLE II
VARIANTS IN THE DATASET

Full Dataset	Selected Features Dataset
Acceleration Events	Acceleration w/o histogram
Braking Events	Braking w/o histogram
Turning Events	Turning w/o histogram
Acceleration + Braking Events	Acceleration + Braking w/o histogram
Braking + Turning Events	Braking + Turning w/o histogram

To generate the concatenated features dataset, one event is added to the end of another event. However, it is important to randomize the larger number of events feature vectors to help create an unbiased dataset. Running a few performance evaluation iterations using this method resulted in comparable performance in both supervised learning and the unsupervised learning methods.

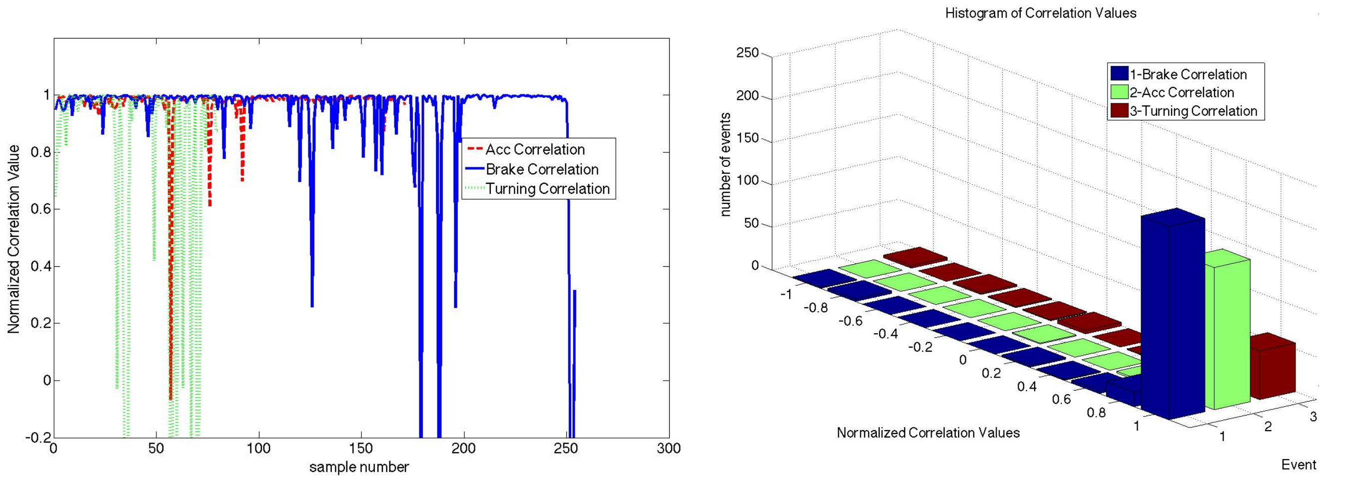
E. Performance Evaluation

Separating them into test and training subsets enables performance evaluation when dataset size is limited. We cross-validate all these dataset and take the average of the performances. Essentially we want to see the effect of decreasing cross-fold validations, with 10% of the data as 10-fold cross-validation to 50% of the data as the 2-folds cross-validation. The numbers you see afterwards are the averages of each subset cross-validation testing. The SVM [1] method



(a) (b) (c)

Fig. 2. (a) Acceleration Event. The Accelerometer output roughly follows the throttle value. (b) Typical Braking Event. The Accelerometer output closely reflects pedal value. (c) Typical Turning Event. The Gyro ouput closely follows the Steering wheel values



(a) (b)

Fig. 3. (a) 0-delay cross-correlation results. For Acceleration events, the 2 signals are accelerometer data and the throttle data (dotted-red). For Brake events, the 2 signals are accelerometer data and brake pedal data (solid-blue). Finally, for the Turning data, the 2 signals are Gyro data and the steering wheel data. The negative correlations for a few samples are probably caused by noise in the data and occur predominately during low vehicle speed and light throttle/brake pedal values. For the most part, the data are correlated. (b) Histogram of the correlation values. Most events' normalized correlation value are 1, therefore most events are correlated.

uses the RBF kernel, and cost and parameters grid listed with power intervals of 2 below.

$$C = [2^{-15}, 2^{15}]$$
$$\gamma = [2^{-15}, 2^{15}]$$

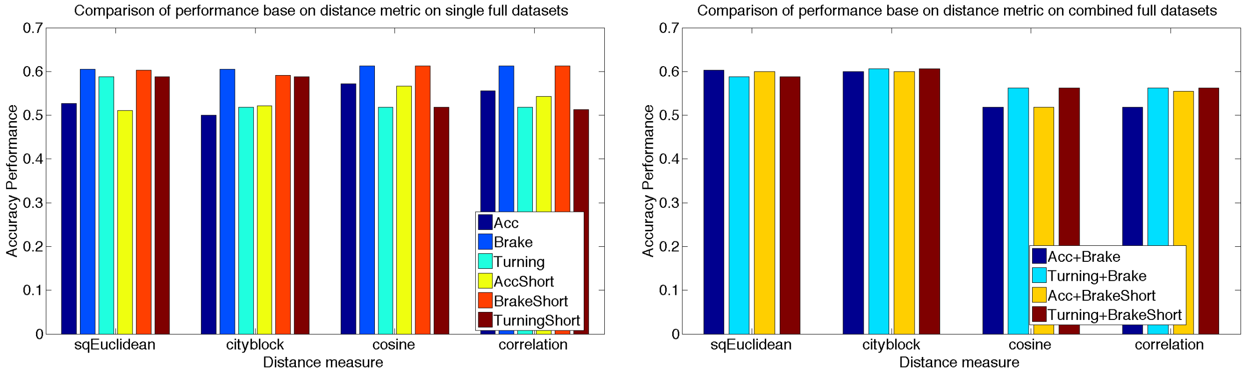
The same process is also applied for evaluating the K-mean Clustering method.

V. Experimental Analysis

SVM [1] and k-mean clustering are used as our training algorithm for a two person classification problem to reduce complexity. The following is broken up into unsupervised and supervised learning method.

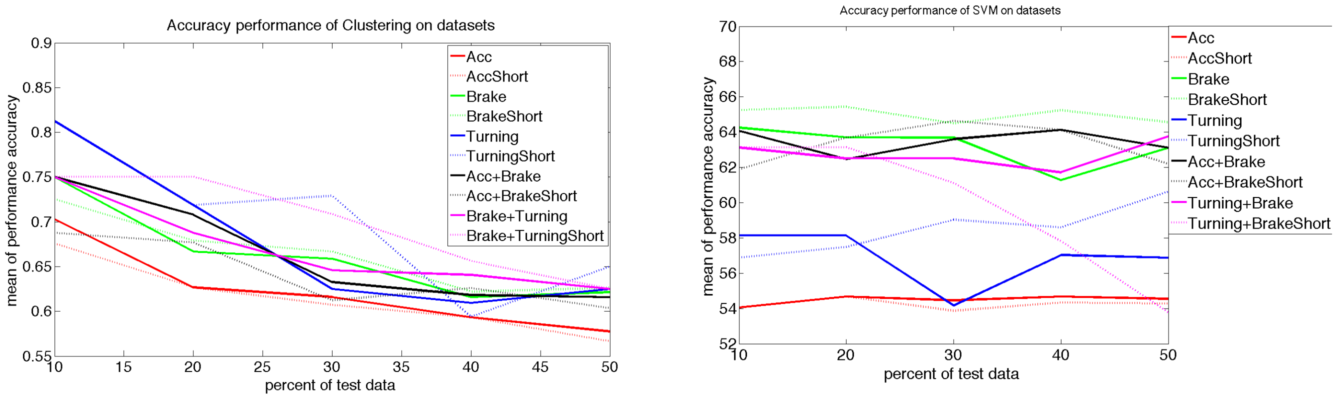
A. *Unsupervised Learning (K-means Clustering)*

We used k-mean clustering with k being two and evaluate on simple distance metric. Fig. 4a shows the performance of some datasets. The first thing to note is most distance metric performs similarly; therefore we can use L2 or even L1 norm to reduce complexity. The second thing to note is the performance is similar for both the shortened feature vector dataset compared to the full feature vector case; therefore we can use the statistical importance of the signal as opposed to keeping the full data. Finally, Braking is the most distinguishing feature vector out of the three types of vehicle maneuver events. From the cross-correlation results in the experiment setup section, we can see the data are correlated for both the acceleration with the throttle signal and the acceleration and the brake pedal. However, from this result, we can see braking is much more distinctive for the two drivers in this test case. Typically people will have their foot on the acceleration pedal, hence requiring the driver to switch pedals. This switching might cause people to press the brake pedal distinctively. Future research using foot cam is needed to verify this.



(a) (b)

Fig. 4. (a) Clustering performance of the dataset using different similarity metrics. The best we can achieve is roughly 60% when using all the data samples. Using Acceleration events only is not informative and including the histogram of the data does not increase performance, therefore, we should only use the shortened data and leave out acceleration events. (b) Cluster performance of the combined dataset using different similarity metrics. Observe the similar trend as in Fig. 4a



(a) (b)

Fig. 5. (a) Performance when we do cross validation where we limit ourselves a subset of training and testing data. The thing to note here is as we have more data to train, the performance is higher. Hence our downward trend as we increase our test dataset percentage. (b) Performance of SVM method. We can see a similar downward trend as we increase the number of test sample compared to training sample. However this method seems to be more resilient to the training samples. The other thing to note is Acceleration data itself is not too informative just as in the k-mean clustering method. Almost in all, the shortened feature vector equally performs or outperforms the full feature vector.

Fig. 4b shows the performance of the combined dataset. Similar to Fig. 4a, the first thing to note is L1 and L2 norm distance metric performs the best and equally well. Moreover, it does not seem like performance increases compared to the individual events when combining the feature vectors. This might indicate we only need to look at the single feature vector vs the combined. Finally we note shortening the dataset has little effect on the performance.

Fig. 5a shows the performance vs testing percentage for clustering. The dotted lines represent the shortened feature vector. The figure correctlys satisfies our intuition of increased performance accuracy with more training data. However if we increase the test samples and reduce the training samples, then performance decreases.

B. Supervised Learning (SVM)

We see a similar trend here using a Supervised learning method shown in Fig. 5b. Accuracy performance increases overall as expected, but is still relatively low. Again Acceleration events do not add much to our performance. Similarly we see using the full dataset does not increase our performance compared to the shortened dataset. Though we still see a similar downward trend in performance as we increase the number of test samples, it is more flat indicating the supervised learning is more constant in performance. Finally in almost all cases the shortened dataset outperforms the full data set. This might be attributed to the RBF kernel, which works well in low dimensionality vectors while the euclidean or polynomial works well in higher dimensionality vectors. Fig. 6 shows the ROC curve of each dataset using the SVM method. We can verify the shortened dataset works well or better compared to full dataset and the combination of braking and turning events gives the highest performance in differentiating between drivers.

VI. Conclusion

In this research we showed there is a potential in using inertial sensors to differentiate between different drivers. While features associated with acceleration events did not play a significant role in differentiating between drivers, features

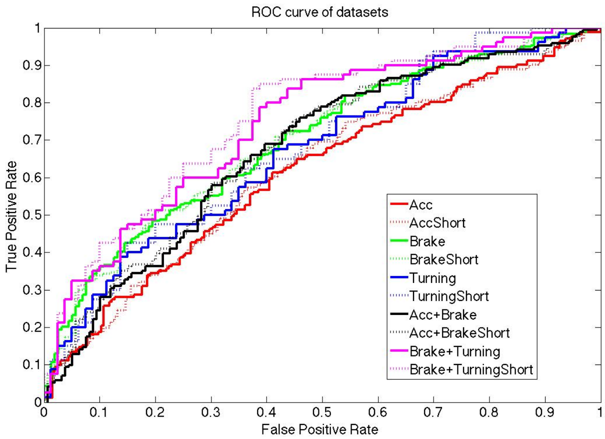


Fig. 6. ROC Curve of all the dataset explored. We can see the concatenated Braking and Turning events feature vectors performs the best out of all our datasets. In general the shortened datasets outperforming the full length datasets.

associated with braking and turning events showed significant potential in differentiating between drivers. Moreover, the histograms of statistical features were not informative enough to affect performance results when compared with statistical features alone. The results presented in this paper are promising enough to determine who is driving because normally the driver is the same within a driving session. Therefore, we can implement a majority count system where if the events are classified 65% of the time as driver A, we will count all driving events in that session as driver A at the end of the drive.

While the scope of the features explored in this work is limited, we would like to explore other combination of features as well as other learning algorithms and distance metric to further improve our classification accuracy. Some additional features can include the driver body posture, head dynamics and hand location. [12], [13], [7] By looking at the driver inside the vehicle, we can learn where a driver positions his/her hands, how fast and how wide he/she moves his/her head to examine the surrounding situation, and any changes in over all body posture[4], [6]. Using these as additional features will go a long way towards understanding driving style and intent so robust and reliable active safety and driver assistance systems can be designed[2], [3].

Other future works include differentiating between more people and implementing such system on a smart phone to verify such classification accuracy still holds. This work can be extended to differentiate how individual drivers vary in their style of driving from day-to-day. This will generate profiles for drivers, which can help to provide the best possible feedback mechanism to reduce dangerous driving.

VII. ACKNOWLEDGMENT

We acknowledge support of the UC Discovery Program and associated industry partners. We also thank our UCSD LISA colleagues who helped in a variety of important ways in our research studies, especially Eshed Ohn-Bar for helping to generate the dataset. Finally, we thank the reviewers for their constructive comments.

REFERENCES

[1] C.-C. Chang and C.-J. Lin. Libsvm: a library for support vector machines. *ACM Transactions on Intelligent Systems and Technology (TIST)*, 2(3):27, 2011.

[2] A. Doshi, S. Y. Cheng, and M. M. Trivedi. A novel active heads-up display for driver assistance. In *Systems, Man, and Cybernetics, Part B: Cybernetics, IEEE Transactions on*, pages 85–93. IEEE, 2009.

[3] A. Doshi, B. T. Morris, and M. M. Trivedi. On-road prediction of driver's intent with multimodal sensory cues. In *IEEE Pervasive Computing*, pages 22–34. IEEE, 2011.

[4] A. Doshi and M. M. Trivedi. On the roles of eye gaze and head dynamics in predicting driver's intent to change lanes. In *Transactions on Intelligent Transportation Systems*, pages 453–462. IEEE, 2009.

[5] A. Doshi and M. M. Trivedi. Examining the impact of driving style on the predictability and responsiveness of the driver: Real-world and simulator analysis. In *Intelligent Vehicles Symposium (IV), 2010 IEEE*, pages 232–237. IEEE, 2010.

[6] T. Gandhi and M. M. Trivedi. Parametric ego-motion estimation for vehicle surround analysis using an omnidirectional camera. In *Machine Vision and Applications*, pages 85–95. IEEE, 2005.

[7] K. S. Huang and M. M. Trivedi. Robust real-time detection, tracking, and pose estimation of faces in video streams. In *IEEE Int Conf on Pattern Recognition*, pages 965 – 968 Vol.3. IEEE, 2004.

[8] D. A. Johnson and M. M. Trivedi. Driving style recognition using a smartphone as a sensor platform. In *Intelligent Transportation Systems (ITSC), 2011 14th International IEEE Conference on*, pages 1609–1615. IEEE, 2011.

[9] N. C. Krishnan, G. N. Pradhan, and S. Panchanathan. Recognizing short duration hand movements from accelerometer data. In *Multimedia and Expo, 2009. ICME 2009. IEEE International Conference on*, pages 1700–1703. IEEE, 2009.

[10] M. Litzenberger, H. Glasl, B. Kohn, B. Schalko, and G. Fernández. Sensor fusion on an embedded system for traffic data analysis-etrada-v system. In *Intelligent Transportation Systems, 2008. ITSC 2008. 11th International IEEE Conference on*, pages 894–899. IEEE, 2008.

[11] J. Mantyjarvi, M. Lindholm, E. Vildjiounaite, S.-M. Makela, and H. Ailisto. Identifying users of portable devices from gait pattern with accelerometers. In *Acoustics, Speech, and Signal Processing, 2005. Proceedings.(ICASSP'05). IEEE International Conference on*, volume 2, pages ii973 – ii976 Vol. 2. IEEE, 2005.

[12] S. Martin, C. Tran, A. Tawari, J. Kwan, and M. M. Trivedi. Optical flow based head movement and gesture analysis in automotive environment. In *Intelligent Transportation Systems (ITSC), 2012 15th International IEEE Conference on*, pages 882 – 887. IEEE, 2012.

[13] E. Ohn-Bar and M. M. Trivedi. In-vehicle hand localization and abnormal event detection. In *IEEE Intelligent Vehicles Symposium*. IEEE, 2013.

[14] K. Rangan and T. Vigneswaran. An embedded systems approach to monitor green house. In *Recent Advances in Space Technology Services and Climate Change (RSTSCC), 2010*, pages 61–65. IEEE, 2010.

[15] A. Sathyanarayana, S. O. Sadjadi, and J. H. Hansen. Leveraging sensor information from portable devices towards automatic driving maneuver recognition. In *Intelligent Transportation Systems (ITSC), 2012 15th International IEEE Conference on*, pages 660–665. IEEE, 2012.

[16] W. Shi, J. Yang, Y. Jiang, F. Yang, and Y. Xiong. Senguard: Passive user identification on smartphones using multiple sensors. In *Wireless and Mobile Computing, Networking and Communications (WiMob), 2011 IEEE 7th International Conference on*, pages 141–148. IEEE, 2011.

[17] H. Thiruvengada, S. Srinivasan, and A. Gacic. Design and implementation of an automated human activity monitoring application for wearable devices. In *Systems, Man and Cybernetics, 2008. SMC 2008. IEEE International Conference on*, pages 2252–2258. IEEE, 2008.