Driving Signature Extraction

Ekim Yurtsever*, Chiyomi Miyajima*, Selpi and Kazuya Takeda***

* *Graduate School of Information Science, Nagoya University, Furo-cho, Chikusa-ku, Nagoya, 464-8603, JAPAN (e-mail: {yurtsever.ekim, miyajima, takeda}@g.sp.m.is.nagoya-u.ac.jp)*

** *Dept. of Applied Mechanics, Chalmers University of Technology, 412 96 Goteborg, SWEDEN (e-mail: selpi@chalmers.se)*

Abstract: This study proposes a method to extract the unique driving signatures of individual drivers. We assume that each driver has a unique driving signature that can be represented in a *k* dimensional principal driving component (PDC) space. We propose a method to extract this signature from sensor data. Furthermore, we suggest that drivers with similar driving signatures can be categorized into driving style classes such as aggressive or careful driving. In our experiments, 122 different drivers have driven the same path on Nagoya city express highway with the same instrumented car. GPS, speed, acceleration, steering wheel position and pedal operations have been recorded. Clustering methods have been used to identify driving signatures.

Keywords: Driving style recognition, driver identification, driver modeling

1. INTRODUCTION

Understanding the distribution of driver behaviour is very important for traffic analysis. Especially identifying aggressive or dangerous drivers is crucial as many researchers such as [1], [2], [3] and [4] showed the relationship between aggressiveness of drivers and car accidents. Even though recent developments in road safety and driver assisting systems improved vehicle safety, they did not eliminated human factors in traffic incidents.

Previous researches introduced various approaches on driver classification and driving style recognition. [5] utilized inertial sensors and used braking, acceleration and turning as three main driving events for driver classification with support vector machine (SVM) and k-mean clustering algorithms. [6] divided driving style into two categories as typical (non-aggressive) and aggressive. Dynamic time warping (DTW) has been used for recognizing aggressive driving and a smartphone as a sensor device. [7] developed Gaussian mixture models (GMM) for car following and pedal operation modelling. Driver identification with these models has been done for evaluation. [8] suggested and categorized drivers into five driving categories: flow conformist, extremist, hunter/tailgater, planner, and ultraconservative. Drivers has been categorized according to the distance they keep with the proceeding vehicle and their velocity. [9] investigated driving behaviours at a roundabout. Speed and acceleration models have been developed using Bayesian inference methods. A detailed survey for driving style recognition has been made by [10].

Driving style recognition and observing the changes of the driving style classes' populations could lead to better road design in terms of safety. For example, if population of aggressive driving style increases for a given amount of drivers on a particular highway turn, it will indicate that turn stimulates aggressive driving. Finding dangerous driving patterns and tracking the history of these patterns could manifest reasons and show patterns for risky behaviour; evaluating of these information could lead to the design of better safety systems. However before achieving such demanding goals, data mining and feature extraction processes should be done. In this study we focus on this aspect. The driving features will be extracted from sensor data first, then these features will be evaluated and meaningful relationships between drivers and features will be established.

This study aims to develop a method to recognize driving styles and extract driving signatures. We believe that driving behaviour of a human driver is a complex structure which has many sub features. We call these sub features principal driving components (PDC). Each driver behaviour is a unique combination of these PDC. Instead of predetermining driving behaviour classes as aggressive, dangerous or safe and categorizing drivers into these classes, we propose to group drivers which shows similar PDC distributions together. This paper explains the PDC extraction out of driving data.

2. EXPERIMENTAL SETUP

122 drivers' driving data on Nagoya city highway in Japan has been obtained in a one year period. Each driver followed the same path during the experiment and each driving session lasted approximately ten minutes. All of the data has been collected with the same vehicle, a Toyota Estima. The intra city highway area where the experiment was held is shown on Google Earth in Fig. 1. 122 different drivers followed this route. 52 of the drivers were female and 70 of them were male. Seven driving signals obtained from this experiment have been used in this study. These signals are: GPS, speed, longitudinal and lateral acceleration, steering wheel position, force on the brake pedal and the gas pedal. Units of these signals are shown in Table 1. Sampling rate of the GPS is 1 Hz while the original sampling rates of the other signals were 16 kHz. However, in our study those signals are downsampled at 10 Hz. Histograms

Fig. 1. Google Earth image of the highway area

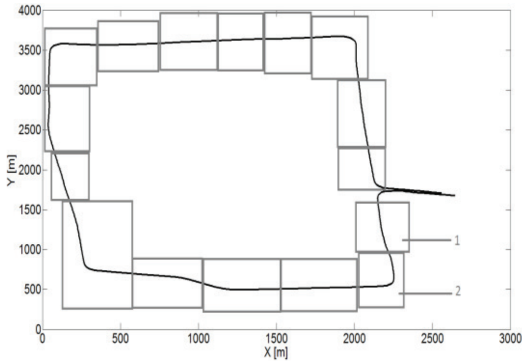

Fig. 2. Sections of the highway. Section 2 and Section 1 are shown. (Traffic is clockwise)

of driving signals are shown in Fig. 3. and driving signal examples are shown in Fig. 4. In Fig 3. x axis represents normalised (z score) values and y axis represents data count. Detailed information about the corpus can be found in [11].

3. EXTRACTION OF DRIVING SIGNATURE

We assume that each driver has a unique driving signature. Categorizing drivers without a complete driving analysis into predetermined driving style classes such as aggressive or normal may lead to wrong categorization. Further, we assume that each driver's unique driving signature ($\boldsymbol{v}$) is a vector in a k dimensional `principal driving component` (PDC) space. We propose a method to determine the dimension of this space and construct the driving signature vector.

$$\boldsymbol{v} = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_k \end{pmatrix} \tag{1}$$

After each drivers $\boldsymbol{v}$ have been extracted, these vectors can be clustered. The final clusters can represent the driving styles.

3.1 PDC space

It is assumed that driving on a straight section and on a curved section of a highway is different for the driver. Because of this, the road network, the highway loop in our case, has been divided into sections. In this paper, the Nagoya highway loop is divided into 16 sections, as shown in Fig. 2.

PDC extraction has been done separately section by section. It should be noted that if PDC extraction would have been done without sectioning, the clusters (PDC's) would have been influenced by the road shape. For example, all of the drivers increase their lateral acceleration and rotate their steering wheel on a turn while most of them would drive straight on a straight section of the road. Therefore, without sectioning, the drive vectors of all drivers on a turn would have been clustered together. Then, differences between particular drivers' driving styles could not be recognized.

The collected driving signals have been also divided into corresponding sections for each driver. GPS data has been used to divide the signals with respect to location after time stamps of GPS have been synchronized with other signals. There are 16 sections and 122 drivers in total. s is used for representing the sections and i for the drivers.

$$\begin{gathered} s: 1,2, \ldots, S \\ i: 1,2, \ldots, D \\ n: 1,2, \ldots, N \end{gathered} \tag{2}$$

n is the time stamp. N is the sample count of a particular driver at a particular section. A faster driver will have a lower N value than a slower driver at the same section.

Derivative values of each signal were calculated. Then, the driving signals of each driver at each section was vectorised to create drive vector. In our study, there are 12 signals used (see Table 1).

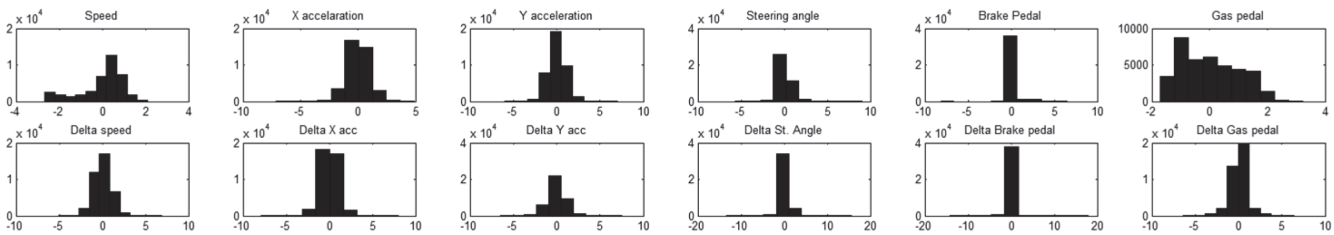

Fig. 3. Histograms of the driving signals at Section 2

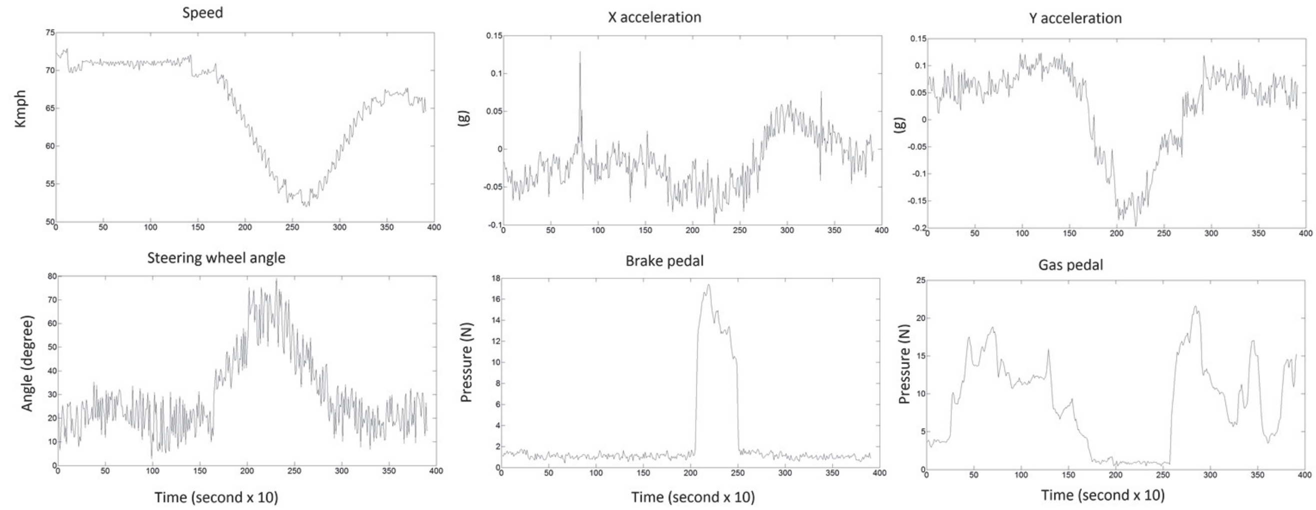

Fig. 4. Drive signal examples

$$\Delta x[m] = \frac{\sum_{k=-K}^{K} kx[m+k]}{\sum_{k=-K}^{K} k^2} \tag{3}$$

The 12 dimensional drive vector ($\boldsymbol{x}$) at a particular section (s) for a particular driver (i) with a particular time stamp (n) is as follows.

$$\boldsymbol{x}(i,s,n) = \begin{pmatrix} x_1(i,s,n) \\ x_2(i,s,n) \\ \cdot \\ \cdot \\ x_{12}(i,s,n) \end{pmatrix} \tag{4}$$

Table 1. Variables

Variable	Physical quantities	unit
x_1	speed	Km/h
x_2	longitudinal acceleration	G
x_3	lateral acceleration	G
x_4	steering wheel angular position	degree
x_5	Force on brake pedal	Newton
x_6	Force on gas pedal	Newton
x_7	Δx_1	
x_8	Δx_2	
x_9	Δx_3	
x_{10}	Δx_4	
x_{11}	Δx_5	
x_{12}	Δx_6	

All drive vectors on the same section were used to create various number of drive clusters. K means clustering method was used for this purpose. With observations $(O_1, O_2, \ldots, On)$ and k $(\leq n)$ sets Q = $\{Q_1, Q_2, \ldots, Q_k\}$:

$$\text{argmin}_Q \sum_{i=1}^{k} \sum_{O \in Q_i} \|O - \mu_i\|^2 \tag{5}$$

Dimension of the PDC space is equal to the meaningful number of clusters, k. An empirical approach has been followed to determine k.

k was selected as 2 at first, then increased incrementally to 15. Average of sums of point to cluster centroid, distortion has been calculated for each number of clusters. The relationship between number of clusters and average distortion in our experiment is shown in Fig.5. Different curves represent different sections.

The `elbow` point, where the curve bends and starts to show linear behaviour is selected as the meaningful number of clusters, dimension of the PDC space. It is selected as 6 in our experiments.

Unit vectors ($\boldsymbol{e_1}, \ldots, \boldsymbol{e_k}$) of the PDC space represents the clusters.

$$\boldsymbol{e_1} = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}_{kx1}, \boldsymbol{e_2} = \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}_{kx1} \ldots \ \boldsymbol{e_k} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}_{kx1} \tag{6}$$

3.2 Driver signature

The drive vectors of driver i at section s is as follows

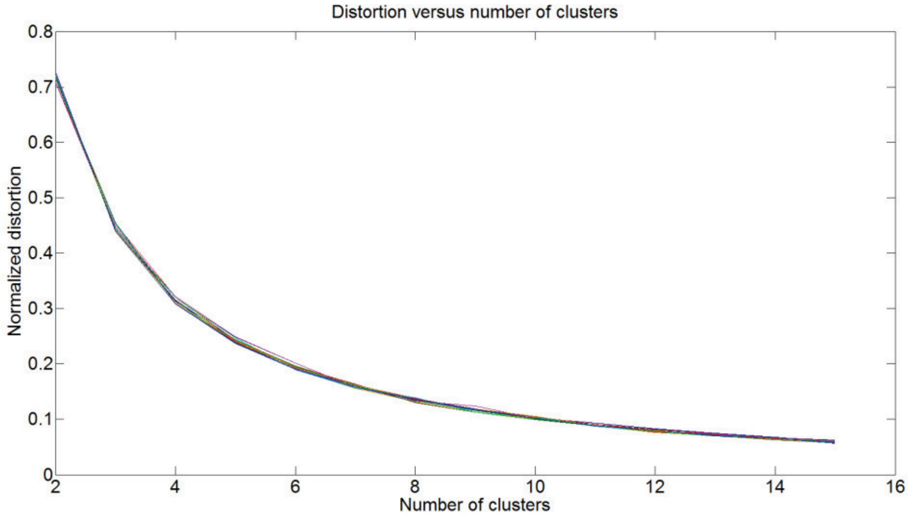

Fig. 5. Distortion versus number of clusters.

$$\boldsymbol{x}(i,s,1),\dots,\boldsymbol{x}(i,s,N) = \underbrace{\overbrace{\begin{pmatrix} x_1(i,s,1) \\ x_2(i,s,1) \\ \cdot \\ \cdot \\ x_{12}(i,s,1) \end{pmatrix}}^{\mathrm{IDX}_1}, \dots, \overbrace{\begin{pmatrix} x_1(i,s,N) \\ x_2(i,s,N) \\ \cdot \\ \cdot \\ x_{12}(i,s,N) \end{pmatrix}}^{\mathrm{IDX}_N}}_{N} \quad (7)$$

where N is sample count of the driver at a particular section and IDX is the cluster index for the corresponding vector.

After k has been selected and all of the drive vectors have been clustered accordingly, the indexes (IDX) of each drive vector (index shows which cluster the particular drive vector belongs to) of one driver at a particular section were summed up and divided to N. N depends on the driver, a faster driver will have a lower N value than a slower driver at the same section. Then, a driver index function (f) was constructed for each driver. For driver i at section s, f is as follows:

$$f(i,s) = p_1\boldsymbol{e_1} + p_2\boldsymbol{e_2} + \dots + p_k\boldsymbol{e_k} \quad (8)$$

$$p_m = \frac{\sum \mathrm{IDX}(m)}{N} \quad (9)$$

Driver signature vector ($\boldsymbol{v}$) is constructed after this with driver index functions. $\boldsymbol{v}$ for driver i is as follows:

$$\boldsymbol{v}(i) = \begin{pmatrix} a_1(i) \\ a_2(i) \\ \vdots \\ a_k(i) \end{pmatrix} \quad (10)$$

Where a_m is:

$$a_m = \frac{\sum_{s=1}^{S} p_m(i,s)}{S} \quad (11)$$

Coefficients of f, p_m, of three random drivers from our experiments are shown in Fig. 6. Each driver has a unique coefficient distribution as can be seen in this figure.

Each driver has a unique $\boldsymbol{v}$. We propose that K means can be used again to cluster these vectors. These final clusters will represent the driving styles. In this study we focused on driving signature extraction. Categorization of drivers based on driving styles will be investigated in our future works.

4. CONCLUSIONS

The behaviour of drivers, their driving signatures are mathematically represented within a PDC space in our proposed method. However, the meaning of PDC's have not been identified and driving styles (for example aggressive driving or careful) have not been categorized yet.

We would like to utilize this method for driver identification also. In our future works we will work on these issues.

ACKNOWLEDGMENT

We thank the Swedish Foundation for International Cooperation in Research and Higher Education (STINT) for the Initiation Grant that enables us to have meetings and discussions that led to this work.

REFERENCES

[1] González-Iglesias, B., Gómez-Fraguela, J. A., & Luengo-Martín, M. Á. (2012). Driving anger and traffic violations: Gender differences. Transportation research part F: traffic psychology and behaviour, 15(4), 404-412.

[2] Chliaoutakis, J. E., Demakakos, P., Tzamalouka, G., Bakou, V., Koumaki, M., & Darviri, C. (2002). Aggressive behavior while driving as predictor of self-reported car crashes. Journal of safety Research, 33(4), 431-443

[3] Shinar, D. (1998). Aggressive driving: the contribution of the drivers and the situation. Transportation Research Part F: Traffic Psychology and Behaviour, 1(2), 137-160.

[4] Deffenbacher, J. L., Deffenbacher, D. M., Lynch, R. S., & Richards, T. L. (2003). Anger, aggression, and risky behavior: a comparison of high and low anger drivers. Behaviour research and therapy, 41(6), 701-718.

[5] Van Ly, M., Martin, S., & Trivedi, M. M. (2013, June). Driver classification and driving style recognition using inertial sensors. In Intelligent Vehicles Symposium (IV), 2013 IEEE (pp. 1040-1045). IEEE.

[6] Johnson, D., & Trivedi, M. M. (2011, October). Driving style recognition using a smartphone as a sensor platform. In Intelligent Transportation Systems (ITSC), 2011 14th International IEEE Conference on (pp. 1609-1615). IEEE.

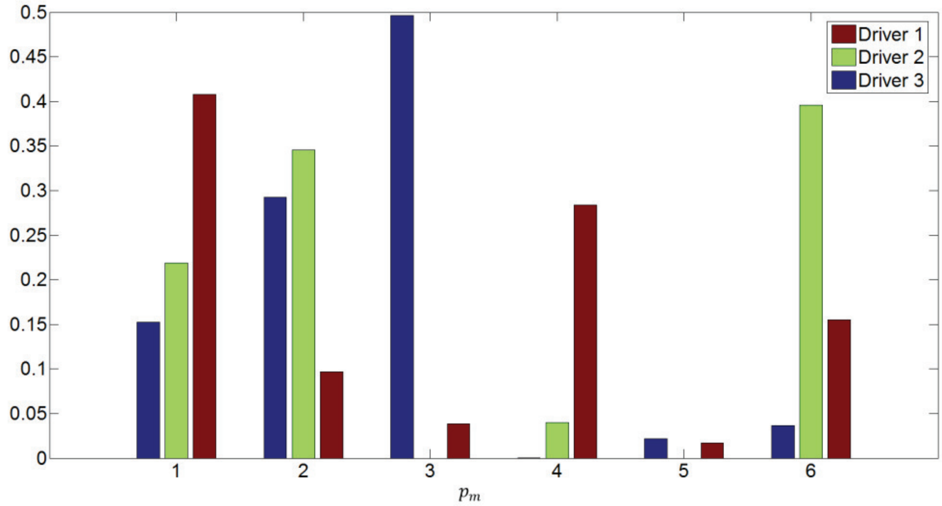

Fig. 6. Coefficients of driver index function for three random drivers from our data

[7] Miyajima, C., Nishiwaki, Y., Ozawa, K., Wakita, T., Itou, K., Takeda, K., & Itakura, F. (2007). Driver modeling based on driving behavior and its evaluation in driver identification. Proceedings of the IEEE, 95(2), 427-437.

[8] Fancher, P. (1998). Intelligent cruise control field operational test (No. UMTRI-98-17).

[9] Mudgal, A., Hallmark, S., Carriquiry, A., & Gkritza, K. (2014). Driving behavior at a roundabout: A hierarchical Bayesian regression analysis. Transportation research part D: transport and environment, 26, 20-26.

[10] Bolovinou, A., Amditis, A., Bellotti, F., & Tarkiainen, M. (2014). driving style recognition for co-operative driving: A survey. In ADAPTIVE 2014, the sixth international conference on adaptive and self-adaptive systems and applications (pp. 73-78).

[11] Takeda, K., Hansen, J. H., Boyraz, P., Malta, L., Miyajima, C., & Abut, H. (2011). International large-scale vehicle corpora for research on driver behavior on the road. Intelligent Transportation Systems, IEEE Transactions on, 12(4), 1609-1623.