

## Driving Style Analysis Using Data Mining Techniques

Z. Constantinescu, C. Marinoiu, M. Vladoiu

**Zoran Constantinescu**

ZealSoft Ltd., Bucharest, Romania

E-mail: zoran@zealsoft.ro

**Cristian Marinoiu, Monica Vladoiu**

PG University of Ploiesti, Romania

E-mail: cmarinoiu@upg-ploiesti.ro, monica@unde.ro

**Abstract:** This paper investigates the modeling of the personal driving style of various vehicle drivers based on several driving parameters. The purpose of such an endeavor is to classify the drivers according to their risk-proneness within the larger context of increasing traffic safety, which is a major concern worldwide. This information is valuable especially for those involved in fleet management and it can be used to improve and to make safer the driving style of various individuals who serve within that fleet. Equally important, such information could help any driver to see the danger within his or her driving style. Cluster and principal component analyses from exploratory statistics have been used to identify and explain drivers grouping according to their driving behavior. The driving parameters (behavioral indices) are collected from urban traffic by an in-house developed GPS-based device and sent to a data server for analysis.

**Keywords:** driving style, driving parameters, real time vehicle tracking, data mining.

## 1 Introduction and Related Work

Our society is changing at an amazing rate and there is no domain of human activity that is not affected by this process. Nowadays, people are overwhelmed with information and are under continuous pressure of being "in time" with some processing of that information. Consequently, everything happens or goes faster and faster each and every day. If we imagine having a look from above at the life pace in the 19th Century and if we compare this image with the current view, the difference is striking. This is also true for people trying to reach their destinations by using motorized vehicles that rush in various directions showing a throng-like view. In the last decades the number of such means of transportation has increased dramatically, their performances have improved at a fantastic rate and therefore, traffic conditions have worsened, especially in major cities. A direct effect of these changes can be seen on the driving behavior of city drivers, which becomes increasingly aggressive and incident-prone, reducing therefore the traffic safety. In this larger context, there is a major interest in categorizing the driving style of city drivers based on their driving behavior, which can be abstracted by means of various driving parameters.

In spite of this interest, the undertaking of extracting "standard behaviors from raw data of real human drivers has not yet been tackled and will be an area of future research" [1]. Even in cases where there have been made field tests that involved real drivers in real world experiences, "the driver's performance in terms of driving style was defined in each test through the subjective judgment of experts present at test" corroborated with fuel consumption [3]. The need for an objective method to understand daily driving behavior which derives from the driving style is emphasized in many works [1, 2, 4, 5]. There are some works that tries to determine the

driving style, seen as "the attitude, orientation and way of thinking for daily driving", based on questionnaires' surveys [4,5]. More recent works use a virtual driving simulator to collect realistic driving data from human drivers and to model human driving behavior [18], or classify driving style by combining objective rank method with recurrent learning based on Elman's type neural network [20]. There are also related works on modelling traffic flow, driving course decisions, or drivers behavior in emergency situations [19]. However, very few studies explore the modeling of personal driving style of various vehicle drivers based on several driving parameters, especially in urban traffic [7–9].

This work is about modeling the personal driving style of various vehicle drivers based on several driving parameters (behavioral indices). The purpose of such an endeavor is to classify the drivers according to their risk-proneness within the larger context of increasing traffic safety, which is a major concern worldwide. This information is valuable especially for those involved in fleet management and it can be used to improve and to make safer the driving style of various individuals who serve within that fleet. Equally important, such information could help any driver to see the danger within his or her driving style.

The paper is organized as follows: section 2 describes shortly the Gipix system and the GPS-based tracking device used for collecting the raw data, section 3 presents the driving parameters extracted from the raw data from the device, section 4 details the used methods, the analysis of the data and the interpretation of the results, and section 5 gives some concluding remarks and possible future work.

## 2 The Gipix System for Vehicle Tracking

Gipix is a system for real-time vehicle tracking, which offers very accurate positioning that is based both on state-of-the-art GPS technology and GSM/GPRS data transmission. The system can be used both for individual vehicles and fleets. The core of the system consists of a data server that processes the maps for the main Romanian cities and roads, the monitored vehicles, the tracks for those vehicles, the drivers, critical events, predefined tracks, specific reports etc. Gipix collects the data of interest by using a GPS-based device, called Gipix-102B that must be installed on each monitored vehicle. This device automatically transmits the vehicle's positions and signals various critical events, both to the system and to the interested users. The Gipix System has been developed in-house to overcome some of the limitations of the commercially available GPS-based solutions [10].

The main advantages of our solution are one-second acquisition interval, high sensibility (the antenna works in difficult conditions), integrability with other applications, adaptability to users' needs, local storage for approximately 100 hours of data for areas which are not GSM covered and positioning without GPS signal if the antenna fails, based on the GSM cells. From all these benefits, the most important for the current work has been the one-second acquisition interval, which is crucial for the success of the statistical methods that we have used.

A screenshot of the Gipix System is depicted in Figure 1, in which the cars of some of the drivers who have been involved in the analysis can be seen. In this screen is displayed also a sample track (the purple curved line). An upside down drop-like cursor can be moved along the track path and the given information will be tailored accordingly. A small info box, which contains data about instantaneous speed ( $v$ ), altitude( $h$ ), acceleration ( $a$ ), GPS error ( $e$ ) can be seen on the right side of the screen. At the lower part of this screen the speed-acceleration plot is available for this track (red for speed and blue for acceleration). Several options can be selected from the left upper main menu: positioning, tracks, routes, personal or general points of interest, settings, configuration of users, car devices, vehicles and so on [10].



Figure 1: Gipix tracking system - driver's track analysis

### 3 Raw Data and the Driving Parameters

The raw data that has been considered for analysis is obtained from the tracking device in real time, by using GPRS and Internet as communication media. All the data is collected in the central server, where it is analyzed. It consists of GPS positions, time and speed values. Data is sampled at 1 second interval. From the speed values we calculate longitudinal acceleration at each time step (using numerical derivation), as well as the mechanical work (as the energy required to increase the speed over the time). A sample speed and acceleration diagram for one random track and driver, and for a small time interval is shown in Figure 2. Values for speed and acceleration are measured in [km/h] and [km/hs], respectively.

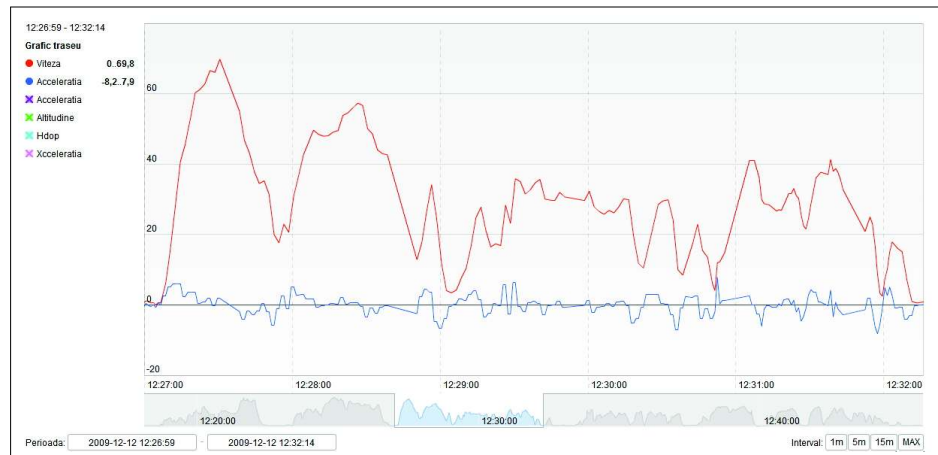


Figure 2: Speed and acceleration plot over time

For the statistical analysis we have been using the following driving parameters (extracted from the raw data):

- *speed over 60 km/h*: percent of time ( $V_{60}$ ) - represents the percentage of time, from the

total time the vehicle is moving (thus excluding all stops of the vehicle), in which the speed is larger than 60 km/h - considered as the speed limit to be fined;

- *speed*: mean value ( $V_{mn}$ ) and standard deviation ( $V_{sd}$ ) - statistical values for the driving speed, the vehicle is also considered as moving;
- *acceleration*: standard deviation ( $A_{sd}$ ) - statistical value for the acceleration, the vehicle is considered as moving as well;
- *positive acceleration*: mean value ( $A_{+mn}$ ) and standard deviation ( $A_{+sd}$ ) - statistical values for all positive accelerations, when the vehicle is considered to increase its speed; this is an indirect measure of the acceleration pedal position;
- *braking*: mean value ( $Br_{mn}$ ) and standard deviation ( $Br_{sd}$ ) - statistical values for negative accelerations, when the vehicle is considered to decrease its speed, and ignoring free decelerations (decrease in speed without breaking); because we don't have any indication of the actual brake pedal position, this is done by setting a threshold for the negative acceleration, above which all values are considered to be free decelerations;
- *mechanical work*: ( $W$ ) - this is calculated as the sum, over the time, of all positive kinetic energy values required to increase the vehicle speed.

Data is collected for a total of 23 different drivers, with two additional controlled test drives. For each driver we collected several tracks (an average of 9 tracks per driver), over a short period of time (2-5 working days), and in similar conditions (same city: Bucharest). The total number of tracks analyzed is 200. The additional two test drives are done with extreme driving styles: a very aggressive one (D91) and a slow, non-aggressive, more economical one (D94). The resulting data is presented in Table 1.

sample	driver	$V_{60}$	$V_{mn}$	$V_{sd}$	$A_{sd}$	$A_{+mn}$	$A_{+sd}$	$Br_{mn}$	$Br_{sd}$	$W$
		%	km/h	km/hs	km/h	km/hs	km/hs	km/hs	km/hs	xJ
1	D1	5.5	25.6	17.5	2.45	2.23	1.46	3.07	1.57	58.42
2	D2	6.2	31.5	17.5	3.24	2.67	1.81	3.93	2.07	81.08
3	D3	14.4	34.7	19.5	2.78	2.30	1.59	3.89	2.23	64.43
4	D4	10.0	32.8	19.4	3.08	2.65	1.72	3.72	2.06	76.67
5	D5	6.4	29.1	19.1	2.81	2.48	1.68	3.34	1.90	70.55
6	D6	0.6	24.2	14.9	3.20	2.64	1.72	4.02	2.30	78.99
7	D7	8.0	27.9	18.3	3.32	2.72	1.91	3.89	2.15	82.17
8	D8	7.4	27.5	18.6	2.89	2.61	1.70	3.31	1.71	69.58
9	D9	2.8	24.7	17.6	2.99	2.68	1.75	3.36	2.03	74.64
10	D10	6.1	26.9	14.1	2.78	2.49	1.70	3.33	1.61	63.82
11	D11	4.6	27.9	17.5	3.27	2.67	1.71	4.18	2.33	76.33
12	D12	3.6	26.0	17.9	2.88	2.48	1.76	3.45	1.82	74.65
13	D13	6.3	30.0	18.5	2.74	2.39	1.61	3.66	1.87	64.96
14	D14	10.2	35.2	18.0	2.79	2.39	1.98	3.16	1.78	61.55
15	D15	6.6	27.0	18.9	2.88	2.50	1.75	3.52	2.15	69.77
16	D16	5.2	33.4	17.6	1.98	1.66	1.08	3.13	1.54	42.42
17	D17	11.3	29.7	21.5	2.45	2.21	1.67	3.01	1.69	53.39
18	D18	3.6	26.3	17.0	3.13	2.74	2.02	3.57	1.84	74.26
19	D19	2.3	25.3	16.6	2.97	2.39	1.99	3.47	2.35	67.71
20	D20	8.0	28.6	20.5	2.66	2.22	1.57	3.32	1.92	65.34
21	D21	18.4	37.8	21.7	3.80	3.02	1.97	4.25	2.76	95.35
22	D22	1.2	23.0	15.1	2.36	2.19	1.48	2.93	1.45	60.92
23	D23	7.2	27.1	17.6	3.37	2.71	1.79	4.15	2.49	87.87
24	D91	14.9	36.8	21.7	3.81	3.27	2.06	4.49	2.53	94.25
25	D94	2.1	27.1	13.9	2.09	1.81	1.14	3.30	1.74	45.16

Table 1: Driving Parameters

## 4 Data Analysis

### 4.1 Analysis Methods

The matrix that represents the driving parameters has been used to conduct the multivariate analysis (Table 1). We have approached two widely used methods: Hierarchical Cluster Analysis (HCA) and Principal Component Analysis (PCA). They allow the evaluation of sample similarities according to determined variables. HCA classifies the drivers according to some variables so that homogeneity within and heterogeneity among groups are obtained. PCA linearly transforms the original variables, and thus composes a new set of independent variables (components), which can be used to identify the important variables that explain sample grouping.

Cluster Analysis (CA) is a method of unsupervised learning, and a statistical methodology used to categorize individual objects into groups with similar meanings (homogeneous). CA is typically used when the researcher does not know the number of groups in advance and wishes to establish groups and to analyze group membership. It seeks to identify groups that both minimize inter-group variation and maximize outer-group variation. We have performed a hierarchical cluster analysis, starting with individual points as clusters, then successively merging two clusters until only one cluster remains. We have used Ward's method and Euclidean distance [11]. Compared to other hierarchical methods, it uses an analysis of variance approach to evaluate the distances between clusters, and it is regarded as very efficient in creating clusters [12]. The method is described more formally in Algorithm 1.

---

**Algorithm 1 - HCA** - basic Hierarchical Clustering Algorithm

---

```

Compute the proximity matrix.
repeat
  Merge the closest two clusters.
  Update proximity matrix.
until only one cluster remains.
```

---

Principal Component Analysis is a statistical method for arranging large arrays of data into interpretable patterning match. It transforms a number of (possibly) correlated variables into a (smaller) number of uncorrelated variables called *principal components*. The first principal component accounts for as much of the variability in the data as possible, and each succeeding component accounts for as much of the remaining variability as possible. The principal components are computed from the matrix of correlations between the variables, outputting their eigenvalues (the amount of variance accounted for by each component) and the component loadings (how the variables correlate with the principal component) [13–15]. This analysis attempts to plot and arrange these variables in a lower-dimensional space, where more closely related items are plotted closer to each other than the less closely related items. A simple formal description of PCA is shown in Algorithm 2.

---

**Algorithm 2 - PCA** - Principal Component Analysis

---

- 1: Organize the data as a  $m * n$  matrix, where  $m$  is the number of measurement types and  $n$  is the number of trials.
  - 2: Subtract off the mean for each measurement type or row in the matrix.
  - 3: Calculate the covariance matrix.
  - 4: Calculate the eigenvectors and eigenvalues of the covariance matrix.
  - 5: Rearrange the eigenvectors and eigenvalues in order of decreasing eigenvalue.
-

## 4.2 Data Analysis

Using our input data, the result of cluster analysis is visualized as a tree-diagram or dendrogram in Figure 3. The dendrogram [12–14] represents the pattern of clustering among the drivers; the longer the connecting lines further to the right the more distance is between the clusters and/or the drivers. Based on the distance, we can distinguish between 6 major clusters (labeled 1 to 6 and colored differently). We can clearly see the two "extreme" clusters (no. 2 and no. 6), associated with the two controlled test drives: D91 (sample 24, cluster no. 6) and D94 (sample 25, cluster no. 2). This is also clear in some of the two-variable plots of the data in Figure 4. However, the other clusters need to be further investigated for a better interpretation.

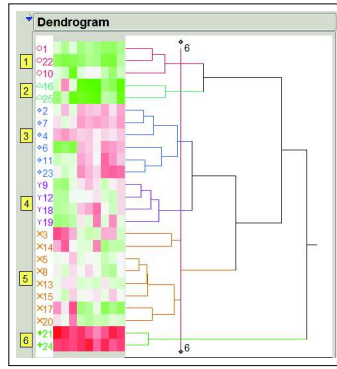


Figure 3: Hierarchical Cluster Analysis

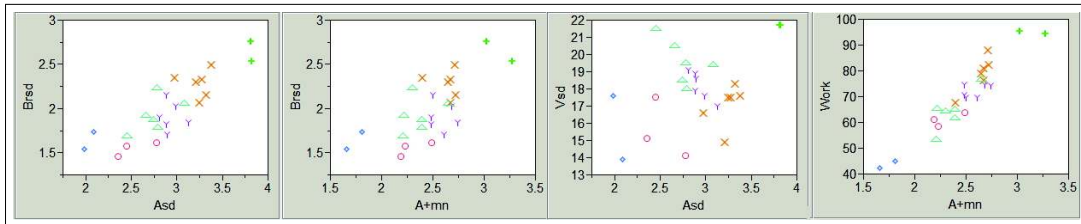


Figure 4: Different plots for Cluster Analysis

Using PCA on our data, we have obtained the principal components by computing the eigenvalues and eigenvectors of data correlation matrix as shown in Figure 5. There are a few common criteria for deciding how many components to keep: a) visual interpretation of the scree plot for the "elbow", b) the number of eigenvalues larger than 1.0, c) required meaningful percentage of variance (80-90%), and d) how many components are interpretable. Using the criteria a) and c), we can account for 2-3 components, with 2 components accounting for 84.5% of the data, and 3 components for more than 92% of the data. Based on the second criteria (b), a number of 2 components would seem more appropriate.

We will decide how many components are significant by further analyzing the correlation between the variables and the principal components, the 2- and 3-component biplots and the clustering of the data.

The correlations between the original variables and the principal components (PC1, PC2, PC3) are presented in Table 2. We can see that the first principal component have a good correlation to all acceleration- ( $A_{sd}$ ,  $A_{+mn}$ ,  $A_{+sd}$ ) and braking- ( $Br$ ,  $Br_{mn}$ ,  $Br_{sd}$ ) related variables, but also to the mechanical work ( $W$ ), whereas the second component has a strong positive correlation to the speed-related variables ( $V_{60}$ ,  $V_{mn}$ ,  $V_{sd}$ ). The third principal component has

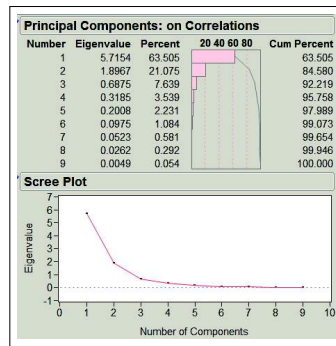


Figure 5: Principal Component Analysis

	PC1	PC2	PC3	RC1	RC2	RC3
$V_{60}$	0.6388762	<b>0.7239720</b>	0.0500097	0.1770218	0.1951046	<b>0.9302670</b>
$V_{mn}$	0.5083209	<b>0.7715614</b>	-0.1699350	-0.0693270	0.2716954	<b>0.8966329</b>
$V_{sd}$	0.5800321	<b>0.6531136</b>	0.2707885	0.3075579	0.0069074	<b>0.8612093</b>
$A_{sd}$	<b>0.9620060</b>	-0.2584930	0.0190901	<b>0.7478074</b>	<b>0.6315966</b>	0.1857643
$A_{+mn}$	<b>0.9105883</b>	-0.2841430	0.1992828	<b>0.8397556</b>	0.4688754	0.1568073
$A_{+sd}$	<b>0.7888608</b>	-0.2625120	<b>0.4604541</b>	<b>0.9192378</b>	0.1912318	0.1471884
$Br_{mn}$	<b>0.8581642</b>	-0.1517880	<b>-0.4571810</b>	0.3312685	<b>0.9062566</b>	0.1935466
$Br_{sd}$	<b>0.8658366</b>	-0.1163580	-0.3475490	0.3933960	<b>0.8199581</b>	0.2385597
$W$	<b>0.9216296</b>	-0.3125500	0.0299027	<b>0.7488098</b>	<b>0.6104728</b>	0.1207866

Table 2: Correlation between variables and components (factor loadings)

correlations (positive and negative) only to the acceleration and braking. Apparently, the first two principal components would give a good correlation with all the variables, however they cannot provide any indication on differentiating the use of the acceleration or of the braking pedal. By calculating the rotation of the principal components using the varimax method, we have obtained the rotated components RC1, RC2, RC3, presented also in Table 2 [15, 16]. We can now see a clear distinction between RC1 and RC2 in explaining the acceleration and braking. The third rotated component, RC3 will have a very good correlation with the speed variables, similarly to PC2.

We further look in Figure 5 at the 2D and 3D scatterplots of the two principal components and of the first two and three rotated components. In the first, 2D plot, we can clearly see the drivers grouped in the six main groups by the cluster analysis. However, there is small overlap between clusters 3 and 4, which is difficult to separate using only the first two components. The second and third plots are made by using the rotated components: the 2D plot of the first 2 components does not separate easily the clusters; however, the 3D plot of all 3 components does indeed separate the clusters well.

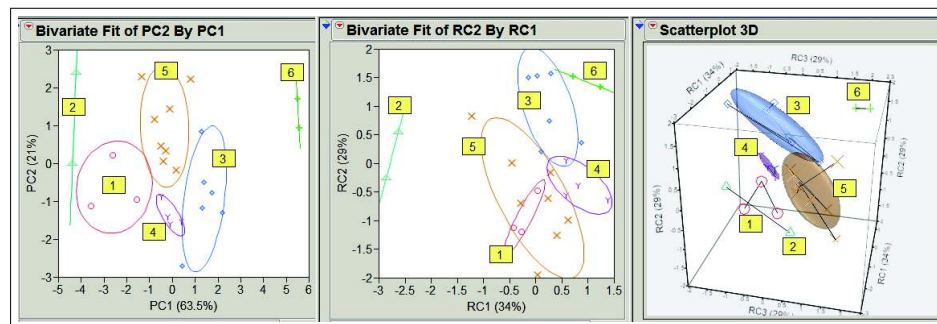


Figure 6: Scatterplots for data: a) PC1-PC2, b) RC1-RC2, c) RC1-RC2-RC3

### 4.3 Data Interpretation

Based on the analysis of the first two principal components (PC1, PC2) and of their values and correlations to the clusters, we can derive an interpretation (Table 3). Based on first principal component (PC1), we can suggest 5 categories of "aggressiveness": from non-aggressive (drive test D94) to a very aggressive (drive test D91). Using the second principal component (PC2), we can further obtain two categories: a tendency to drive with high speed in the city, and a more moderate speed driving.

By analyzing the three rotated components (RC1, RC2, RC3) and their values and correlation to the clusters, we can suggest three intervals for each of the factors: (very) small (below 1.0 or 1.5), near zero-one (between -1.0 and 1.0), and large (above 1.0). Their interpretation is given in Table 3. We combine the interpretations of the clusters, based on principal components and rotated components into Table 4.

Component	Values	Interpretation (driving style)	Clusters
PC1 (63.5%)	Very small ( $< -5$ )	Non-aggressive	2
	Small ( $-5 < -1$ )	Somewhat non-aggressive	1
	Between ( $-1 < 1$ )	Neutral	4 5
	Large ( $1 < 5$ )	Moderately aggressive	3
	Very large ( $> 5$ )	Very aggressive	6
PC2 (21.0%)	Negative	Low-moderate speed	1 2 4 5 3
	Positive	Tendency to high speed	2 5 6
RC1 (34%)	Very small ( $< -2$ )	Lower acceleration usage	2 5
	Between ( $-1 < 1$ )	Moderate acceleration usage	1 2 3 5
	Large ( $> 1$ )	Higher acceleration usage	4 6
RC2 (29%)	Very small ( $< -1.5$ )	Smooth braking	1 5
	Between ( $-1.5 < 1$ )	Moderate braking	1 2 3 4 5
	Large ( $> 1$ )	Sudden braking	3 6
RC3 (29%)	Small ( $< -1$ )	Tendency to lower speed	1 2
	Between ( $-1 < 1$ )	Moderate speed	1 2 3 4 5
	Large ( $> 1$ )	Tendency to high speed	5 6

Table 3: Interpretation by Principal Components and by Rotated Components

Cluster	Test	Aggressivity (PC1)	Speed (PC2,RC3)	Accelerating (RC1)	Braking (RC2)
1		Moderately low	Low-Moderate	Moderate	Smooth-Moderate
2	D94	Very low	Low-Moderate	Low-Moderate	Smooth-Moderate
3		Moderately high	Moderate	Moderate	Sudden
4		Neutral	Moderate	High	Moderate
5		Neutral	Moderate-High	Low-Moderate	Moderate-Sudden
6	D91	High	High	High	Sudden

Table 4: Interpretation of the Clusters

## 5 Conclusions and Future Work

Driving quality can be thought as the "combination of an energy saving driving style with a behavior that is respectful of the environment (noise, pollution and safety) and of the vehicle and is comfortable for the passengers as well" [3].

We are aware of the fact that many things affect drivers' behavior, such as driving environment (weather, road condition, day/night etc.), traffic context, driver's particular condition (upset, ill, tired, sleepless, distracted etc.), and, of course, driver's individual characteristics (gender, age, driving experience and frequency, annual mileage, familiarity and confidence with respect to driving, personal attitude, etc.). Unfortunately, for the time being we have not been able to access all that information for our set of drivers. The results of this work can be significantly improved by taking into consideration, at least, some of these factors, which can be easily provided (individual characteristics). To tackle other factors, such as the drivers state of mind is more difficult, but can be done with the right technical solution (real-time computer vision analysis of video streams from the driver and other sensors attached to the driver). Finally, there are some issues, as



anxiety related to possible traffic accidents, which are very hard to assess objectively, from an observer's viewpoint, without having the drivers interviewed or questioned explicitly about those matters. Though, the questionnaires-based studies may fall into the pitfall of answers from the perspective of socially desirable behavior. Anyway, when the external factors such as traffic situation or road environment have a smaller effect, "the driving behavior will be regulated more by the driving style, which is one of the internal factors" [4].

The data analysis for this paper was generated using Version 9 of the SAS System. Copyright 2008 SAS Institute Inc., Cary, NC, USA [17].

## Bibliography

- [1] Rigolli, M., Brady, M., Towards a Behavioural Traffic Monitoring System, International Conference on Autonomous Agents, *Proceedings of the 4th International Joint Conference on Autonomous Agents and Multiagent Systems*, pp. 449-454, 2005.
- [2] Rygula, A., Driving Style Identification Method Based on Speed Graph Analysis, *International Conference on Biometrics and Kansei Engineering*, pp. 76-79, 2009.
- [3] Vangi, D., Virga, A., Evaluation of Energy-Saving Driving Styles for Bus Drivers, Proc. Instn Mech. Engrs, Vol. 217 Part D: *J. Automobile Engineering*, pp. 299-305, 2003.
- [4] Ishibashi, M., Okuwa, M., Doi, S., Akamatsu, M., Indices for Characterizing Driving Style and their Relevance to Car Following Behavior, *SICE Annual Conf.*, pp. 1132-1137, 2007.
- [5] O. Taubman-Ben-Ari, M. Mikulincer and O. Gillath, The multidimensional driving style inventory-scale construct and validation, *Accident Analysis and Prevention*, Vol. 36, pp. 323-332, 2004.
- [6] D. J. French, R. J. West, J. Elander and J. M. Wilding, Decision-making style, driving style, and self-reported involvement in road traffic accidents, *Ergonomics*, Vol. 36, No. 6, pp. 627-664, 1993.
- [7] Bonsall, P., Liu, R., Young, W., Modelling Safety-related Driving Behaviour-Impact of Parameter Values, *Transportation Research Part A* 39, pp. 425-444, 2005.
- [8] Mierlo, J., Maggetto, G., Burgwal, E., Gense, E., Driving Style and Traffic Measures-Influence on Vehicle Emissions and Fuel Consumption, Proc. Instn Mech. Engrs Vol. 218 Part D: *J. Automobile Engineering*, pp. 43-50, 2004.
- [9] Cherrett, T., Pitfield, D., Extracting Driving Characteristics from Heavy Goods Vehicle Tachograph Charts, *Transportation Planning and Technology*, Vol. 24, No. 4, pp. 349-363, 2001.
- [10] M. Vladoiu, Z. Constantinescu, Toward Location-based Services using GPS-based Devices, *Proceedings of ICWN 2008 - the 2008 International Conference of Wireless Networks*, a Conference of World Congress on Engineering 2008 (WCE 2008), July, London, UK, Vol. I, pp. 799-804, 2008.
- [11] Ward, J.H., Hierarchical grouping to optimize an objective function. *Journal of the American Statistical Association*, Vol. 58, pp. 236-244, 1963.
- [12] Anderberg, M. R., *Cluster Analysis for Applications*. Academic Press, New York, 1973.

- 
- [13] Sokal, P.H.A., Sneath, R.R., *NUMERICAL TAXONOMY: The Principles and Practice of Numerical Classification*. Freeman, San Francisco, 1973
  - [14] MacQueen, J.B, Some Methods for classification and Analysis of Multivariate Observations, *Proceedings of 5th Berkeley Symposium on Mathematical Statistics and Probability*, Berkeley, University of California Press, Vol. 1, pp. 281-297, 1967.
  - [15] Jolliffe, I.T., *Principal Component Analysis*, Springer, 2002.
  - [16] Dauxois, J., Pousse, A. and Romain, Y., Asymptotic theory for the principal component analysis of a vector random function: some applications to statistical inference. *Journal of Multivariate Analysis*, Vol. 12, pp. 136-154, 1982.
  - [17] SAS Institute Inc., *SAS 9.1.3 Help and Documentation*. Cary, NC: SAS Institute Inc., 2000-2004.
  - [18] Hattori, Hiromitsu, Nakajima, Yuu and Ishida, Toru, Agent Modeling with Individual Human Behaviors, *Proc. of 8th Int'l. Conf. on Autonomous Agents and Multiagent Systems (AAMAS 2009)*, pp. 1369-1470, 2009.
  - [19] Plochl, Manfred and Edelmann, Johannes, Driver models in automobile dynamics application, *Vehicle System Dynamics*, Vol. 45, No. 7, pp. 699-741, 2007.
  - [20] Augustynowicz, A., Preliminary Classification of Driving Style with Objective Rank method, *International Journal of Automotive Technology*, Vol. 10, No. 5, pp. 607-610, 2009.