A Support Vector Clustering Based Approach for Driving Style Classification

Yuxiang Feng, Simon Pickering, Edward Chappell, Pejman Iravani, and Chris Brace

Abstract—**All drivers have their own habitual choice of driving behavior, causing variations in fuel consumption. It would be beneficial to classify these driving styles and extract the most economical and ecological driving patterns. However, driving style of each driver is not consistent and may vary within a single trip. Therefore, this paper proposes a novel technique to robustly classify driving style using the Support Vector Clustering approach, which attempts to differentiate the variations in individual's driving pattern and provides an objective driver classification. It is part of a research program aiming to replicate some humans' driving behaviors on chassis dynamometer using a robot driver. Moreover, it can potentially be used in developing more economical and personalized advanced driver assistance systems (ADAS) and humanized autonomous driving strategies. With the easily accessible on-board diagnostics (OBD) data on modern vehicles, both vehicle state and traffic information of three drivers were collected using an instrumented vehicle, which had external forward-looking radar and a monocular dashcam. For data processing, each trip data was first segmented into separate event groups. Prominent factors were then extracted by applying Principal Component Analysis (PCA) on both statistical and spectral features of all signals. Afterwards, Support Vector Clustering (SVC) was performed to classify driving style during the trip. The trained classifier was used to indicate the driving pattern variations in percentage. The validity of the proposed method was evaluated using the jerk profile, where a high correlation was found between the classification results and jerk distributions. Moreover, a positive relation between fuel consumption and driving aggressivity was also confirmed. Furthermore, it was found that weather condition, time of the day and ultimately, the driver's eagerness, can cause significant variations in driving style.**

Index Terms—**Driving style analysis, fuel consumption, real-world driving data, support vector clustering.**

I. Introduction

With the rapid growth of eco-driving training programs [1]-[3], the potential benefit of improving driving style on fuel consumption has been widely acknowledged in the recent decade. While the exact fuel reduction remains debatable, and may vary among different individuals, most previous studies indicated a fuel saving of 10% - 15% can potentially be achieved through the optimization of driving style [4], [5]. Alongside with the promotion of eco-driving, the variances of driving styles have also been recognized.

With the earliest driving style research occurred in 1949 [6], most of these studies were primarily focusing on improving driving safety, and various self-report instrument questionnaires had been adopted for data collection [7]. However, owing to the overwhelming energy crisis and the advent of vehicular sensing technologies, another trend emerged in the last decade, which utilized recorded vehicle information to analyze driving styles and their relations to fuel consumption. Many pattern recognition methods have hence been adopted for classification. For example, Aljaafreh *et al.* [8], Al-Din *et al.* [9], and Dörr *et al.* [10] developed fuzzy logic based classifiers. Meanwhile, Macadam *et al.* [11] and Meseguer *et al.* [12] used Neural Network to differentiate driving styles. Moreover, other methods, such as K-means and hierarchical clustering [13], K-nearest neighbors [14], and self-organizing map [15] have also been applied for driving style classification.

Although these studies have successfully classified drivers into three or four labelled groups (aggressive, normal, defensive, etc.), they tended to assume each participant's driving style remained consistent within each trip, and neglect the potential driving style variations. While this assumption may be suitable for these studies, a more plausible approach is to divide each entire trip into several segments, as even an extreme aggressive driver may not maintain driving aggressively during the entire trip. Therefore, this paper proposes an event based classification approach which aims to differentiate the variations in each individual's driving pattern, and provides a more objective driving style classification.

Meanwhile, most of existing studies only use vehicle state information for classification, as shown in Table I. While these vehicle-related parameters can certainly reveal different driving styles, it should be noted that the interaction with traffic flow can also be a major cause to the variances. For example, in some car following scenarios, the headway distance to the leading vehicle can be a critical reveal of different driving styles, and also the underlying trigger of the change in vehicle state. Therefore, some influential factors on driving styles can be lost if excluding traffic information, as drivers of different driving styles have diverse traffic anticipation preferences [7]. Thus, driver's interaction with traffic flow should be included to improve the performance of driving style classification. The headway distance to the leading vehicle is hence identified as a prominent factor to represent this interaction. Unlike vehicle state information, which can be easily retrieved from the Engine Control Unit (ECU), the headway distance in this study was obtained from external Continental radar and a monocular camera, using the data fusion approach described in previous work [16].

Alongside with the data collection, Support Vector

Clustering (SVC) was selected for driving style classification. Inspired by Support Vector Machine (SVM), SVC is an unsupervised clustering algorithm initially proposed by Ben-Hur *et al.* [17] in 2001. Owing to its performance in detecting arbitrary shape clusters with a hierarchical structure in high dimensional data [18], SVC has previously been used for pattern recognition [17] and image segmentation [19]. As SVC is a relatively new classification algorithm, it hence hasn't been adopted for driving style research before. However, its supervised version (SVM) has already been used in some related research. For instance, Wang and Xi used vehicle speed and throttle opening as feature parameters, and developed an algorithm that combined K-means clustering and SVM to classify drivers into aggressive and moderate [20]. Meanwhile, both driving performance and physiological measurements were used in a SVM classifier to distinguish drunk and normal driving [21]. As the performance of this set of support vector based algorithms has been validated in many studies [17]-[21], SVC was hence adopted as the classification algorithm in this research.

TABLE I: Feature Parameters for Driving Style Classification

Author	Feature Parameters
Aljaafreh *et al.* [8]	Longitudinal acceleration Lateral acceleration Vehicle speed
Al-Din *et al.* [9]	Longitudinal acceleration Vehicle speed Following distance
Dörr *et al.* [10]	Acceleration/deceleration Vehicle speed Time gap ACC activation
Macadam *et al.* [11]	Range Range rate
Meseguer *et al.* [12]	Acceleration Vehicle speed Engine speed
Constantinescu *et al.* [13]	Acceleration Vehicle speed Mechanical work
Vaitkus *et al.* [14]	Acceleration
Albers and Albrecht [15]	Acceleration Engine speed Transmission and wheels Path of gap and clutch pedals

This paper proposes to differentiate driving style variations during each trip and provide a percentage based driving style classification using OBD data. The performance of this method is evaluated using real driving data of three human drivers. The major contributions of this paper are the investigation of driving style variations for each driver, the inclusion of traffic information, and the consideration of spectral features. Meanwhile, the potential influence of weather condition on driving styles was also investigated. The ultimate aim of this study is to replicate some humans' driving behaviors using a robot driver, and hence introduce the variance of driving styles in drive cycle research. Moreover, as driving styles can influence Real Driving Emissions (RDE) tests, this classification can help to increase the validity of RDE results, and facilitates the replication of these tests on chassis dynamometer. Furthermore, it can also contribute in developing more economical and personalized advanced driver assistance systems (ADAS) and autonomous driving strategies.

II. METHODOLOGY

The adopted data collection method and general procedures are introduced in this section. Related algorithms, such as event detection, feature parameters selection, PCA, and SVC are included in this section.

A. Data Collection

1) Test equipment

The real driving data used in this study was collected from an instrumented 2014 VW Sharan, which was equipped with an Influx Rebel data logger, a 77 GHz long range Continental radar ARS 308, and a Nextbase dashcam. The vehicle state information was directly retrieved from ECU with the data logger. Meanwhile, in order to record headway distance information, a sensor fusion approach was previously proposed, which performs Kalman filter to fuse radar and dashcam measurements [16].

2) Trip data

Three drivers participated in the driving data collection phase. Real driving data of 12 separate trips were collected, with a total covered distance of approximately 1106 km. Meanwhile, all three drivers were requested to drive the vehicle in similar weather condition and time of the day, which is to minimize the potential disturbances caused by these external factors. Moreover, in order to evaluate the influence of weather conditions on driving style variations, extensive driving data of one particular driver was collected in different weather conditions, such as sunny, foggy, rainy and dark. The speed distributions of obtained driving data are illustrated in Fig. 1, with different trip data of the same driver combined.
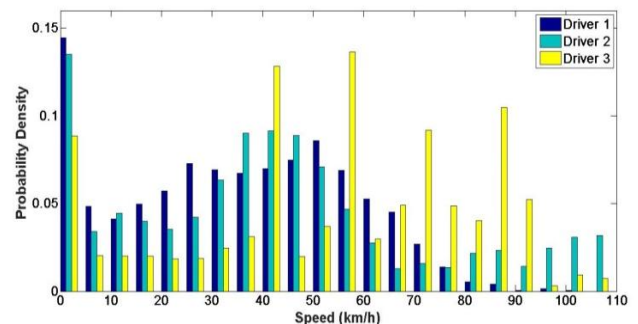

Fig. 1. Vehicle speed distribution of each driver.

This histogram has a bin size of 5 km, and ranges from 0 to 110 km/h, as the maximum speeds of each driver are 100.8 km/h, 112 km/h and 106 km/h respectively. It can be noted that the speed preferences of these three drivers are distinct. The first driver shows a larger proportion (41.4%) in low speed range (0 – 30 km/h). Meanwhile, the third driver has the largest proportions at 40 – 45 km/h and 55 – 60 km/h, which are 12.8% and 13.6% respectively. Moreover, the second driver has a relatively smoother speed distribution in speed range (0 -70 km/h), and a larger proportion (17.8%) in high speed range (70 -110 km/h).

Alongside with the vehicle speed, several other signals, such as engine speed, throttle pedal position, and headway distance, were also extracted as feature parameters for classification. This is because these parameters have direct linkages with the driver's choice, and are hence assumed capable of revealing driving style variances [9], [12].

Meanwhile, the corresponding instantaneous fuel consumption was also recorded to investigate their correlation.

B. Event Detection

With all trip data collected, synchronizations among different signals were implemented within each individual trip. This is because radar and OBD measurements were logged at different frequencies. After the synchronization, a dynamic sliding window approach was developed to segment each trip into separate event groups. Four classes of driving events were derived as accelerating, braking, maintaining, and stop. Each class of event was defined by fixed conditions, and can occur multiple times during each trip. The transitions between different events are illustrated in Fig. 2.
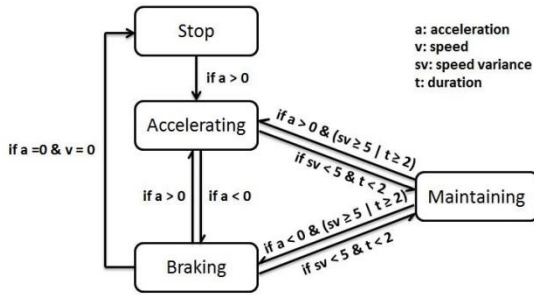

Fig. 2. Transitions between events.

As shown in Fig. 2, short accelerating or braking events with small speed changes are regarded as maintaining. Therefore, each trip data can be segmented into these four event groups using the above transition conditions. One segmented trip is illustrated in Fig. 3 to demonstrate this proposed method. It can be noted that there are 170 accelerating events, 131 braking events, 258 maintaining events and 19 stop events detected during this trip.
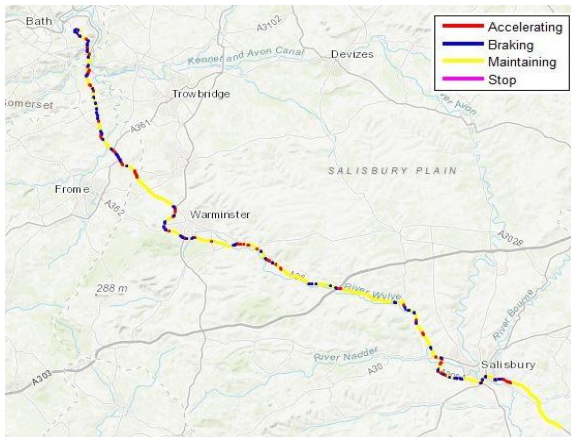

Fig. 3. Event segmentation result of one selected trip.

C. Feature Parameters Selection

1) Statistical features

With all driving data segmented into separate groups, the statistical features of each signal were first extracted. Four typical statistical features were identified as mean, standard deviation, maximum and minimum values. With four input signals, a 16-dimension statistical feature parameter was hence computed for each detected event.

2) Spectral features

Alongside with statistical features, spectral features of

each signal were also captured, as ignoring the temporal dependencies can lead to skewed results [22]. This is because the measurement at one state is highly correlated with measurements at adjacent states. While most existing driving style studies only use statistical features for classification, spectral features have been adopted in two previous studies. One was implemented by Žylius *et al.* in 2014 [23]. They used short-time Fourier transform to analyse accelerometer signals and classify driving styles as aggressive and safe. Meanwhile, Discrete Wavelet Transform (DWT) was implemented by Hallac *et al.* to achieve driver identification using driving data from a single turn [22]. While both methods can be used to extract spectral features, DWT was selected in this study owing to its flexible time-frequency window.

3) Discrete wavelet transform

While there are several different types of wavelet, Haar wavelet was selected for the transformation as it is one of the most frequently used wavelets for non-smooth functions [24]. Originally proposed by Haar in 1910 [25], its mother wavelet function is defined as,

$$\psi(t) = \begin{cases} 1 & 0 \le t < 0.5 \\ -1 & 0.5 \le t < 1 \\ 0 & \text{otherwise} \end{cases} \quad (1)$$

The core of DWT is to compute approximation and detail coefficients by passing the original signal through a series of filters. As the feature of the original signal is preserved in these coefficients, they can hence be treated as the spectral feature parameters [22].

Moreover, in order to ensure the computed spectral feature parameters have the same dimensions, measurements of events within a same group were resampled to a unified length. This length was dynamically defined by the longest event within the group. Using this resampling process, the dimensions of DWT vectors within each event groups were hence aligned.

D. Principal Component Analysis

After both statistical and spectral features were extracted, they were combined as feature parameters for classification. However, it should be noted that the dimensions of these parameters were quite large, and hence unsuitable for further processing. In order to reduce the dimensions of feature parameters and improve the clustering speed, Principal Component Analysis (PCA) was hence implemented. PCA is a statistical method that uses orthogonal transformation to convert correlated variables into linearly uncorrelated variables. The converted variables are referred to as principal components, with each principle component accounts for as much of the variability in the original data as possible. Therefore, deciding the number of principle components is crucial to dimension reduction. While there are four commonly used criteria for this selection, which are, a) visual interpretation of the scree plot for the "elbow", b) eigenvalues larger than 1.0, c) meaningful percentage of variance, and d) interpretable components, the third criterion was selected in this study, as it can efficiently reduce the dataset to 2-3 components [13]. Therefore, selected principal components were required to represent at least 95% variance of original data. The number of principal components and the corresponding percentage of variance of each trip are listed in

Table II.

TABLE II: PCA RESULTS OF EACH TRIP

Trip Number	Accelerating		Braking		Maintaining	
	no	pct. (%)	no	pct. (%)	no	pct. (%)
1	3	97.4	2	96.9	2	97.7
2	3	96.8	2	96.7	2	96.7
3	4	97.6	2	97.2	2	97.5
4	3	97.3	2	95.6	2	97.7
5	3	95.2	3	95.3	2	97.0
6	3	95.0	3	95.9	2	97.1
7	3	97.2	3	96.1	2	96.3
8	3	96.0	4	96.7	2	97.1
9	3	96.7	4	96.2	2	96.4
10	3	97.2	3	96.1	2	96.3
11	3	97.1	3	96.4	2	96.2
12	3	95.5	3	95.4	2	97.0

E. Support Vector Clustering

After the PCA process, Support Vector Clustering (SVC) was hence performed on the selected principal components to classify each event into different driving style groups. According to Ben-Hur *et al.* [17], SVC has two main steps, which are SVM Training and Clustering Labelling. The first step aims to construct cluster boundaries. The original data is mapped into a high dimensional feature space using a Gaussian kernel function, which can be represented as,

$$K\left(x_i, x_j\right) = e^{-q||x_i - x_j||^2} \tag{2}$$

where q is the width parameter.

Afterwards, the smallest sphere that encloses the image of feature points is searched, which can be described as,

$$||\phi\left(x_j\right) - a||^2 \leq R^2 + \xi_j \tag{3}$$

where a is the sphere center; R is the radius; ξ_j is the slack variable.

Lagrangian with penalty term is hence introduced to solve this problem,

$$L = R^2 - \sum_j \left(R^2 + \xi_j - \left||\phi(x_j) - a\right||^2\right)\beta_j \\ - \sum_j \xi_j \mu_j + C \sum_j \xi_j \tag{4}$$

where β_j and μ_j are Lagrange multipliers; $C \sum_j \xi_j$ is the penalty term.

Therefore, the distance of point x to sphere center in feature space can be derived as,

$$R^2(x) = K(x,x) - 2\sum_j \beta_j K\left(x_j, x\right) + \sum_{i,j} \beta_i \beta_j K\left(x_i, x_j\right) \tag{5}$$

The cluster boundaries can hence be determined by contours that enclose the points in data space given by,

$$\{x | R(x) = R\} \tag{6}$$

During the second step, cluster labels are assigned to each data point. As this cluster labelling process can be time consuming, several different approaches have been proposed to improve the efficiency and accuracy of this procedure, such as Complete Graph (CG), Delaunay Diagram (DD), Minimum Spanning Tree (MST), K-Nearest Neighbours (K-NN), and Reduced Complete Graph (R-CG) [26]. While these approaches were proposed in different studies, their performances have already been compared on the same data set [26]. As the dimensions of feature parameters in this research were reduced to 2-3 using PCA, DD was hence selected as the cluster labelling algorithm considering the trade-offs between labelling accuracy and time complexity.

III. RESULTS

The results are presented in three sections, discussing the classification results of the proposed method, their correlation with fuel consumption, and the influence of weather conditions respectively.

A. Driving Style Classificaiton

To demonstrate the proposed driving style classification method, two trip data sets for each driver were classified separately. It can help to investigate the variations of each individual's driving style within one trip, and also the consistency between different trips. Moreover, these drivers were also compared to further assess their driving style.

1) Driver 1

The classification results of the first driver were shown in Table III.

TABLE III: CLASSIFICATION RESULT OF DRIVER 1

Trip	Driving Style	Driving Event (%)		
		Accelerating	Braking	Maintaining
1	Aggressive	15.2	12.8	10.0
	Normal	73.4	45.8	57.9
	Defensive	11.4	41.4	32.1
2	Aggressive	21.4	3.8	14.7
	Normal	69.3	44.5	61.2
	Defensive	9.3	51.7	24.1

It can be noted that the classification results were rather consistent between two trips, which indicates that this driver possesses a relatively stable driving style. Meanwhile, this driver tends to be a normal driver, as most driving events were classified as normal driving. Moreover, the similar proportions of normal and defensive during braking events also demonstrate that driving style can vary during each trip. Therefore, this driver can be classified as a mixed normal and defensive driver, with a higher tendency towards normal driving. Based on the average of both trips' data, this driver can be defined as 13.0% aggressive, 58.7% normal and 28.3% defensive.

2) Driver 2

TABLE IV: CLASSIFICATION RESULT OF DRIVER 2

Trip	Driving Style	Driving Event (%)		
		Accelerating	Braking	Maintaining
1	Aggressive	88.7	5.4	41.7
	Normal	0.8	86.5	56.3
	Defensive	10.5	8.1	2
2	Aggressive	93.6	3.6	63.8
	Normal	5.6	2.7	23.3
	Defensive	0.8	93.7	12.9

As shown in Table IV, the second driver has a mixed driving style. While aggressive was the dominant driving style during accelerating events in both trips, this driver's driving style in braking and maintaining events varied dramatically, which could be caused by some external factors, such as traffic condition and tight schedule. It can be noted in the second trip that while most braking events were classified

as defensive, a higher tendency of aggressive driving was detected during both accelerating and maintaining events. It demonstrates that this driver might have a tight schedule during this trip, and he still concentrated on driving safety (93.7 defensive braking). Nevertheless, this driver tends to be an aggressive driver, especially in accelerating events. From percentage perspective, this second driver can be classified as 49.5% aggressive, 29.2% normal and 21.3% defensive.

3) Driver 3

It can be noted from Table VI that this driver tends to possess a mixed normal and defensive driving style. The performance of this driver was rather consistent between two trips. While defensive driving occupied a dominant proportion (88.0% and 94.7%) in braking events, both normal and defensive driving styles were detected during accelerating and maintaining events. Therefore, this driver can be classified as a mixed normal and defensive driver, with a higher tendency of defensive driving. This driver can be defined as 5.6% aggressive, 26.8% normal and 67.6% defensive.

TABLE V: CLASSIFICATION RESULT OF DRIVER 3

Trip	Driving Style	Driving Event (%)		
		Accelerating	Braking	Maintaining
1	Aggressive	11.9	9.8	4.8
	Normal	30.1	2.2	43.1
	Defensive	58.0	88.0	52.1
2	Aggressive	3.2	1.4	2.3
	Normal	36.1	3.9	45.4
	Defensive	60.7	94.7	52.3

4) Drivers Comparison

With the driving style of each driver separately classified, a cross comparison among these three drivers was performed. As illustrated in Fig. 4, while all three driving styles were detected in each driver's data, these three drivers tend to be more normal, more aggressive, and more defensive separately.
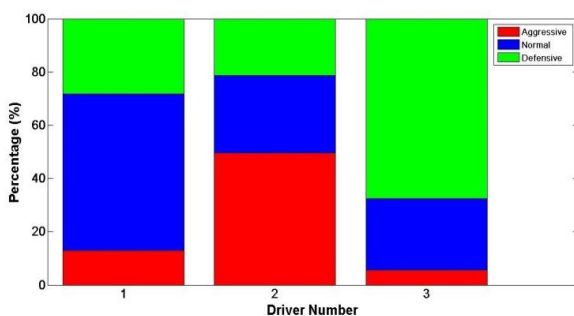

Fig. 4. Driving style classification of three drivers.

Moreover, in order to evaluate the performance of the proposed classification approach, the jerk profile of each driver was computed for validation. Defined as the rate of change in acceleration or deceleration, jerk is widely recognized as a crucial factor in determining the driver's aggressiveness [27]. While classification using jerk only focuses on the driver's reaction on speed control, and neglects the influence of traffic condition, its validity in revealing driving style variance has been confirmed [27], [28]. Therefore, the jerk profiles of each driver were computed, and the distributions of absolute jerk were shown in Fig. 5.
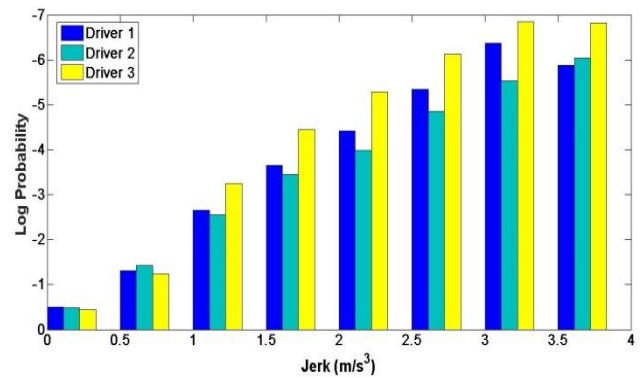

Fig. 5. Log probability of each driver's jerk profile.

It can be noted that the jerk distributions had a high correlation with the classified driving styles, especially in range [0.5, 4], where the aggressive second driver had a larger proportion in large jerk range (14.24% in [1, 4]), and the defensive third driver had a larger proportion in small jerk range (29.20% in [0.5, 1]). Meanwhile, in range [0, 0.5], while the third driver still occupied the largest proportion (64.75%), the proportion of the second driver (61.67%) was slightly larger than the first driver (60.88%), which is contradictory to common expectations. However, it should be noted that this phenomenon might be caused by the second driver's extreme defensive behavior during braking events. Moreover, the mean absolute jerks of these three drivers were also computed as 0.4182, 0.4561, and 0.3131 m/s^3. Therefore, based on the jerk analysis, these three drivers can be classified as normal, aggressive and defensive drivers respectively, which correspond to the proposed classification results. Thus, the performance of the proposed SVC based driving style classification approach is validated.

B. Correlation with Fuel Consumption

With the driving styles of each driver classified, their correlation with fuel consumption was also investigated. The results obtained were illustrated in Fig. 6.
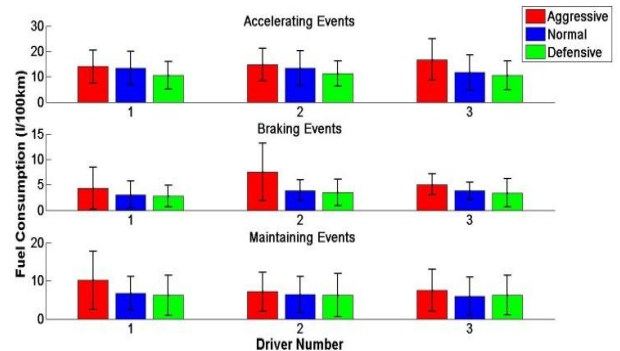

Fig. 6. Average fuel consumption of each driver.

As shown in Fig. 6, the average fuel consumptions of each driver were computed for different driving styles and events. Meanwhile, the error bars were used to represent the standard deviation, which can indicate the variations of fuel consumption. It can be noted that for each separate driver, their average fuel consumptions generally satisfied the common expectation that aggressive driving consumed most fuel, and defensive consumed least. Moreover, an interesting finding is that the average fuel consumption of the classified aggressive driver (driver 2) was not the largest in some driving events. For instance, in the aggressive driving part of

accelerating events, the average fuel consumption of driver 2 (aggressive driver) was less than driver 3 (defensive driver). This indicates that a defensive driver can consume more fuel than an aggressive driver when driving aggressively. Nevertheless, driver 2 had a higher frequency of aggressive driving, which still led to most fuel consumed.

C. Weather Influence

While the trip data of three drivers were recorded in the same weather condition to minimize external disturbances, six more trips of the third driver were recorded to investigate the potential influence of weather condition. The eight trips of this driver can be categorized into four pairs, each representing sunny, foggy, rainy and dark.

The classification results of these eight trips were listed in Table VI. It can be noted that while this driver's driving style was rather consistent between trips in the same weather, the variances caused by different weather conditions were large.

TABLE VI: CLASSIFICATION RESULTS OF DRIVER 3 IN DIFFERENT WEATHERS

Trip	Driving Style	Driving Event (%)		
		Accelerating	Braking	Maintaining
1	Aggressive	4.2	8.4	10.1
	Normal	75.2	40.8	82.3
	Defensive	20.6	50.8	7.6
2	Aggressive	6.9	11.4	8.2
	Normal	90.4	74.3	84.9
	Defensive	2.7	14.3	6.9
3	Aggressive	11.9	9.8	4.8
	Normal	30.1	2.2	43.1
	Defensive	58.0	88.0	52.1
4	Aggressive	3.2	1.4	2.3
	Normal	36.1	3.9	45.4
	Defensive	60.7	94.7	52.3
5	Aggressive	0.4	5.5	1.4
	Normal	96.9	3.9	3.9
	Defensive	2.7	90.6	94.7
6	Aggressive	7.3	10.4	6.0
	Normal	92.1	12.0	17.3
	Defensive	0.6	77.6	76.7
7	Aggressive	8.6	26.1	2.1
	Normal	44.6	71.1	91.4
	Defensive	46.8	2.8	6.5
8	Aggressive	10.4	28.0	5.4
	Normal	35.8	63.7	93.9
	Defensive	53.8	8.3	0.7

As shown in Fig. 7, the driving style of this driver tended to be more normal in sunny and dark, and more defensive in foggy and rainy conditions. While the variances between sunny, foggy and rainy can be direct evidence of how weather can affect driving style, the percentage distribution of dark indicates that time of the day may also have considerable influence. This is because a major difference introduced by these three weathers (sunny, foggy and rainy) is the visibility. Meanwhile, this driver also performed more defensively with reduced visibility in foggy and rainy conditions. However, the driver showed a higher aggressiveness in dark condition, when visibility was also restricted. This indicates that the time of the day and the eager to get home may have a larger influence on driving style than weather conditions. Thus, incorporating these factors in future research may further improve the performance of driving style classification. Nevertheless, the potential influence of weather on driver's driving style is validated.
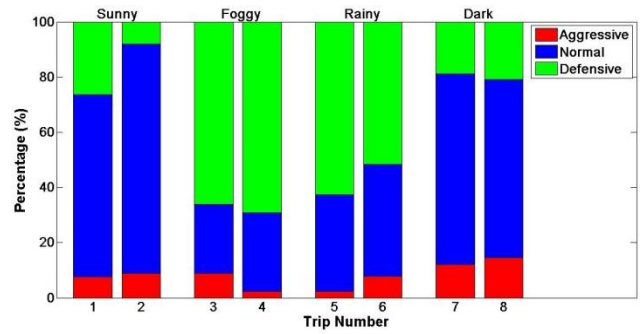

Fig. 7. Driving style classification of each trip.

IV. CONCLUSION

The primary aim of classifying driving style using Support Vector Clustering was achieved. During the data collection phase, 12 trip data from three drivers were collected. Both headway distance and vehicle state information were recorded to reveal the driving style variations. Afterwards, a dynamic sliding window approach was proposed to segment each trip into different event groups. During the data analysis phase, Discrete Wavelet Transform was first performed to extract spectral features of collected data. Afterwards, Principal Component Analysis was used to reduce the dimension of feature parameters, and identify prominent factors from the combined statistical and spectral features. Support Vector Clustering was then performed on these prominent factors to classify driving styles and indicate the driving pattern variations of each driver.

The performance of this proposed approach was evaluated using the collected data of three human drivers. Alongside with differentiating driving style variations of each driver, the validity of the classification results was also examined using the jerk profile. Moreover, the correlation between the classified driving styles and fuel consumption was also investigated. Furthermore, the influence of weather condition on driving style was evaluated using extensive trip data of the third driver.

It was found that the proposed classification approach can efficiently differentiate the driving style variations during each trip, and the classified driving styles have a high correlation with fuel consumption. Meanwhile, weather condition, time of the day and the driver's eagerness can also cause variations in driving style.

Therefore, it can be noted that the proposed Support Vector Clustering based driving style classification approach can effectively differentiate the variations in individual's driving pattern. More importantly, it validates the hypothesis that each driver's driving style is not consistent and can be affected by many factors. Thus, a more complete driving style classification method that incorporates these external factors is recommended in future research.

Moreover, it should be noted that owing to the limited participants, this study mainly focuses on validating the proposed approach. More data samples will be collected to reach a generalized classification in following studies.

This work is an extension to its earlier version, which was presented at the 2018 3rd International Conference on Intelligent Transportation Engineering, held at Singapore.

REFERENCES

[1] S. H. Ho, Y. D. Wong, and V. W. C. Chang, "What can eco-driving do for sustainable road transport? Perspectives from a city (Singapore) eco-driving programme," *Sustainable Cities and Society*, vol. 14, pp. 82-88, February 2015.

[2] EcoDriven Project, "European campaign on improving driving behaviour, Energy-efficiency and traffic safety," Final Progress Report, Belgium: Intelligent Energy Europe, 2009.

[3] R. Luther and P. Baas, "Eco-driving scoping study," *Energy Efficiency and Conservation Authority*, 2011.

[4] A. Alessandrini, A. Cattivera, F. Filippi, and F. Ortenzi, "Driving style influence on car CO_2 emissions," in *Proc. 20th International Emission Inventory Conference – Emission Inventories – Meeting the Challenges Posed by Emerging Global, Nat., and Regional and Local Air Quality Issues*, 2012.

[5] R. J. Vermeulen, "The effects of a range of measures to reduce the tail pipe emissions and/or the fuel consumption of modern passenger cars on petrol and diesel," TNO Report, 2006.

[6] W. A. Tillmann and G. E. Hobbs, "The accident-prone automobile driver, A study of the psychiatric and social background," *The American Journal of Psychiatry*, vol. 106, no. 5, pp. 321-331, 1949.

[7] F. Sagberg, S. Selpi, G. F. Piccinini, and J. Engström, "A review of research on driving styles and road safety," *Human Factors*, vol. 57, no. 7, pp. 1248-1275, November 2015.

[8] A. Aljaafreh, N. Alshabatat, and M. S. N. Al-Din, "Driving style recognition using fuzzy logic," *Vehicle Electronics and Safety (ICVES)*, pp. 460-463, 2012.

[9] M. S. N. Al-Din, A. Aljaafreh, N. Albdour, and M. Saleh, "Driving styles recognition using decomposed fuzzy logic system," *International Journal of Electrical, Electronics & Computer Systems*, vol. 16, no. 1, pp. 1-5, September 2013.

[10] D. Dörr, D. Grabengiesser, and F. Gauterin, "Online driving style recognition using fuzzy logic," in *Proc. 2014 IEEE 17th International Conference on Intelligent Transportation Systems (ITSC)*, 2014, pp. 1021-1026.

[11] C. Macadam, Z. Bareket, P. Fancher, and R. Ervin, "Using neural network to identify driving style and headway control behaviour of drivers," *Vehicle System Dynamics Supplement*, vol. 28, pp. 143-160, 1998.

[12] J. E. Meseguer, C. T. Calafate, J. C. Cano, and P. Manzoni, "Characterizing the driving style behaviour using artificial intelligence techniques," in *Proc. IEEE 38th Conference*, 2013.

[13] Z. Constantinescu, C. Marinoiu, and M. Vladoiu, "Driving style analysis using data mining techniques," *International Journal of Computers Communications & Control*, vol. 5, no. 5, pp. 654-663, December 2010.

[14] V. Vaitkus, P. Lengvenis, and G. Zylius, "Driving style classification using long-term accelerometer information," in *Proc. 19th International Conference on Methods and Models in Automation and Robotics*, 2014, pp. 641-644.

[15] A. Albers and M. Albrecht, "Method using a self-organising map for driver classification as a precondition for customer oriented design," in *Proc. International Conference on Engineering Design*, 2005.

[16] Y. X. Feng, S. Pickering, E. Chappell, P. Iravani, and C. Brace, "Distance estimation by fusing radar and monocular camera with Kalman filter," in *Proc. SAE 2017 Intelligent and Connected Vehicles Symposium*, 2017.

[17] A. Ben-Hur, D. Horn, H. T. Siegelmann, and V. Vapni, "Support vector clustering," *Journal of Machine Learning Research*, 2001, vol. 2, pp. 125-137.

[18] J. Yang, V. Estivill-Castro, and S. K. Chalup, "Support vector clustering through proximity graph modeling," in *Proc. the 9th International Conference on Neural Information Processing (ICONIP '02)*, 2002, vol. 2, pp. 898-903.

[19] Y. Q. Wang and X. Liu, "Improved support vector clustering algorithm for color image segmentation," *Engineering Review*, vol. 35, no. 2, pp. 121-129, 2015.

[20] W. S. Wang and J. Q. Xi, "A rapid pattern-recognition method for driving styles using clustering-based support vector machines," in *Proc. 2016 American Control Conference (ACC)*, 2016.

[21] H. Q. Chen and L. Chen, "Support vector machine classification of drunk driving behaviour," *International Journal of Environmental Research and Public Health*, 2017.

[22] D. Hallac, A. Sharang, R. Stahlmann, A. Lamprecht, M. Huber, M. Roehder, R. Sosič, and J. Leskovec, "Driver identification using automobile sensor data from a single turn," in *Proc. 2016 IEEE 19th International Conference on Intelligent Transportation Systems (ITSC)*, 2016, pp. 953-958.

[23] G. Žylius, V. Vaitkus, and P. Lengvenis, "Driving style analysis using spectral features of accelerometer signals," in *Proc. 9th International Conference on Intelligent Techologies in Logistics and Mechatronics Systems (ITELMS'2014)*, 2014.

[24] S. Mallat, *A Wavelet Tour of Signal Processing: The Sparse Way*, Academic Press, 2008.

[25] A. Haar, "Zur Theorie der orthogonalen funktionensysteme," *Mathematicsche Annalen*, vol. 69, pp. 331-371, 1910.

[26] J. Lee and D. Lee, "Dynamic characterization of cluster structures for robust and inductive support vector clustering," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 28, no. 11, pp. 1869-1874, September 2006.

[27] Y. L. Murphey, R. Milton, and L. Kiliaris, "Driver's style classification using jerk analysis," in *Proc. IEEE Workshop on Computational Intelligence in Vehicles and Vehicular Systems*, pp. 23-28, 2009.

[28] C. F. Lee and P. Öberg, "Classification of road type and driving style using OBD data," SAE Technical Paper 2015-01-0979, 2015.

Yuxiang Feng is currently pursuing his Ph.D. degree at the Department of Mechanical Engineering, University of Bath, UK. He received his MSc degree in Mechatronics with distinction at Univeristy of Bath in 2014, and obtained his BSc degree in Mechanical Engineering in 2012 from Tongji University, China.

His research interests mainly include autonomous vehicle, humanized driver modelling, sensor fusion, and drive cycle and RDE test.

Simon Pickering completed his Ph.D. degree in 2005 at the Department of Mechanical Engineering, University of Bath, UK. He received his MEng degree at University of Bath in 2000.

He has been working as a research officer at the Powertrain and Vehicle Research Centre, University of Bath, between 2005 and 2014, and been promoted as a research fellow since 2014.

Edward Chappell received his Ph.D. degree in 2015 at the Department of Mechanical Engineering, University of Bath, UK. He received his MEng degree at University of Bath in 2009.

He was a knowledge transfer fellow at the Powertrain and Vehicle Research Centre, University of Bath, between 2010 and 2013. He has temporarily worked as an experimental officer in 2014, and been promoted as a research associate between 2014 and 2017. He is currently working as a research fellow at the university.

Pejman Iravani is a senior lecturer at the Department of Mechanical Engineering, and the coordinator of the Robotics Lab at the University of Bath, UK.

He worked as a lecturer at the University of Bath bewteen 2006 and was promoted as a senior lecturer since 2017.

His research interest mainly involves computer vision, machine learning, dynamic modelling and control systems.

Chris Brace is a professor of automotive propulsion, and the deputy director of the Powertrain and Vehicle Research Centre at the University of Bath, UK.

He worked as a lecturer at the University of Bath between 2000 and 2006, and was promoted as a senior lecturer bewteen 2006 and 2012. After that, he worked temporarily as a reader between 2012 and 2014, and was assigned the professor of automotive propulsion since 2014.

His research interest mainly includes low-carbon combustion engines, turbocharging, and emission test.