Driving Style Estimation by Fusing Multiple Driving Behaviors: A Case Study of Freeway in China

Yulin Ma[1], Zhixiong Li[2, 3], Yicheng Li[4], Hongtai Li[1], R. Malekian[5]

[1.] National Center of ITS Engineering and Technology, Research Institute of Highway, Ministry of Transport, Beijing, 100088 China
[2.] School of Mechatronic Engineering, China University of Mining and Technology, Xuzhou 221000, China
[3.] School of Mechanical, Materials, Mechatronic and Biomedical Engineering, University of Wollongong, Wollongong, NSW 2522, Australia
[4.] ITS Research Center, Wuhan University of Technology, Wuhan 430063, China
[5.] Department of Electrical, Electronic & Computer Engineering, University of Pretoria, Pretoria 0002, South Africa

Abstract

Traffic accident is one of the most serious issues in traffic problems. In China, more than 50 thousand people die in each year from traffic accidents. To alleviate the incidence of traffic accidents, this paper proposes a driving style estimation method by fusing multiple driving behaviors for Chinese drivers. Firstly, we invite Chinese volunteers to operate a driving simulator. Massive driving data are collected by the simulator. Then, a driving dataset is set up by the collected data. Furthermore, we adopt the collected driving data to represent behaviors by using SVM. Last but not least, a novel classification method is proposed to estimate driving styles, which is called multiple decision tree. The method can fuse multiple behaviors and explore the relationship between driving styles and behaviors. As a result, 20 volunteers and a freeway in China is selected for case study. After test, the proposed method has a 95% accuracy for style estimation. However, about 25% volunteers have a Risk style and these volunteers should change their driving habits. It also reveals the high incidence of accidents in China. Hence, the proposed method can alert the driver with bad styles and is helpful to ease traffic accidents.

Key Words: traffic accident; driving style estimation; driving behavior; multiple decision tree; SVM

1. INTRODUCTION

Traffic accident is one of the most serious issues in traffic problems. It is the key cause of death of the population in recent years. To alleviate this issue, many auxiliary driving systems have been designed [1], such as Forward Collision Warning (FCW), Adaptive Cruise Control (ACC), Lane Departure Warning System (LDWS), etc. The above systems can alleviate incidence of traffic accidents. However, there are still many people dead in vehicle crashes. Table 1 shows the traffic accident statistics from 2009 to 2013 in China [2]. From this table, we can find that although the death toll decreased, it was still more than 50,000 in 2013. The statistics indicate that traffic accident remission cannot only depend on high-tech systems. So it is critical to implement suitable prevention strategies to improve the safety [3].

TABLE 1 Traffic accident statistics from 2009 to 2013 in China

Years	Traffic accidents (events)	Death toll (number of people)
2009	238351	67759
2010	219521	65225
2011	210812	62387
2012	204196	59997
2013	198394	58539

Chinese government has conducted further analysis and investigation of traffic accidents. The research shows that hazardous driving is one of the main reasons to cause traffic accidents. Hence, driving behavior has become a hot topic all over the world [4] . Various researches are focus on hazardous driving [5], angry driving [6], etc. These researches mainly explore driving behaviors. When drivers operate hazardous drivings, these methods will send a warning signal to drivers. However, in some dangerous situations, the drivers usually have no time to change their operations before the warning signals are sent to them. Hence, another method should be proposed. The method should have an ability to make long term observation for drivers and classify the styles of drivers. Therefore, if drivers have known which driving style the drivers belong to, they will avoid hazardous driving behaviors when in dangerous situations. To this end, this paper proposes a driving style estimation method to decide what kind of driving style a driver belongs to. The proposed method is aimed at reducing the incidence of traffic accidents.

2. RELATED WORKS

2.1 Review of driving behavior

Driving behavior is different from driving style, but it has an influence on judgment of driving styles. Driving style is decided by multiple driving behaviors. A driver maybe drive with a lot of behaviors including risk, moderation and safety. However, he has only one style. For instance, if a driver has a lot of risk behaviors, his style will be risk. Hence, the work presented here is strongly related to driving behavior analysis, from which we draw many inspirations, particularly in clustering analysis, driving behaviors classification [7, 8].

From literature, various methods have been proposed to classify driving behaviors. In these methods, vehicle speed, acceleration, position, braking are the main data to analyze driving behaviors. For instance, Aarts et al. [9] analyze the relationship between speed and road crashes. They conclude that vehicle speed not only affects the severity of collisions, but also increases the risk of collisions. Based on this, Wu et al. [10] set up a driving behavior clustering model by using GPS data mining. In this model, they adopt vehicle speed and acceleration to classify driving behaviors into four kinds: driving behavior of acceleration-deceleration, pro-speeding, acceleration and deceleration. In addition, Zito et al. [11] investigate the use of GPS data for traffic monitoring. In their method, GPS data can compute the vehicle position and travelling time. Vehicle position and travelling time combined with hazardous situations can help drivers recognize their hazardous driving behaviors. Moreover, Yan et al. [12] set up a model to detect hazardous situations by using Markov Blanket and sequential minimal optimization. From the above literature, the goal of driving behavior classification is to detect driving safety. Vehicle speed, position and travelling time are closely related with driving safety. However, the researches discussed above are mainly focus on some specific traffic events. To evaluate the whole journey, driving style should be researched.

2.2 Review of driving style

Driving style is based on driver's habits. Previous researches reveal that drivers with risk style are usually driving faster, braking harder than drivers with safety style [13, 14]. Hence, to study driving style, we can start from driving data.

Guo et al. collect naturalistic driving data on various roads and then adopt these data to detect behaviors [15]. Furthermore, they divide drivers styles into three risk levels by using K-means cluster method. However, they only adopt braking data as the features to evaluate driving styles. Similarly, Wang et al. [16] collect naturalistic driving data on real Chinese roads. The driving database includes vehicle status, potential crash objects, driver information, actions, etc. They fuse these kinds of data and cluster different driving-risk levels involved in near-crashes by using nine rules. Moreover, in [17], Bonsal et al. set up a model for personal driving style classification based on several safety-related parameters. From this model, they confirm the parameters which effect traffic safety most as key parameters. Li et al.[18] propose a method to identify driving styles. Driving behavior is represented by 12 kinds of behaviors. Then they employ a conditional likelihood maximization method to extract driving features. Finally, driving styles are classified by using these features.

From the mentioned literature, we can find that multi-dimensional naturalistic driving data are important for driving style analysis. How to establish the relationship among the different kinds of data is a key point to classify driving styles. Hence, in this paper, we first set up a driving database for multi-dimensional driving data. Furthermore, we adopt the driving data to represent multiple driving behaviors. Last but not leaset, we propose a method to classify driving styles by fusing the multiple driving behaviors.

3. MULTI-DIMENSIONAL DRIVING DATABASE CREATION

Various factors affect the classification of driving styles, which is shown in Figure 1. According to the figure, driving style contains three levels, driving decisions, driving behaviors and driving data. Driving behaviors are based on driving data, such as vehicle speed, acceleration, braking, etc. In other words, we can adopt sequence of driving data to represent behaviors.

We utilize a driving simulator to create the database. The driving simulator is shown in Figure 2. This simulator contains the main freeways in China. We randomly select more than 200 volunteers from China whose driving age are more than 1 year. These volunteers can represent the styles of Chinese drivers. In the step of database creation, volunteers first select a destination and then select a route for driving. The route may be a time-saving route, short-distance route, etc. When volunteers operate the simulator, various data are collected at the same time. These data contain GPS data, vehicle speed, acceleration, braking, yaw angle. Hence, there are more than 7 kinds of behaviors are selected for each volunteer. Each volunteer drives more than one hour. More than 10,000 data are collected for each volunteer.There are more than 2 million data in the dataset. All of these data have timestamps. Table 2 shows part of data in multi-dimensional driving database.

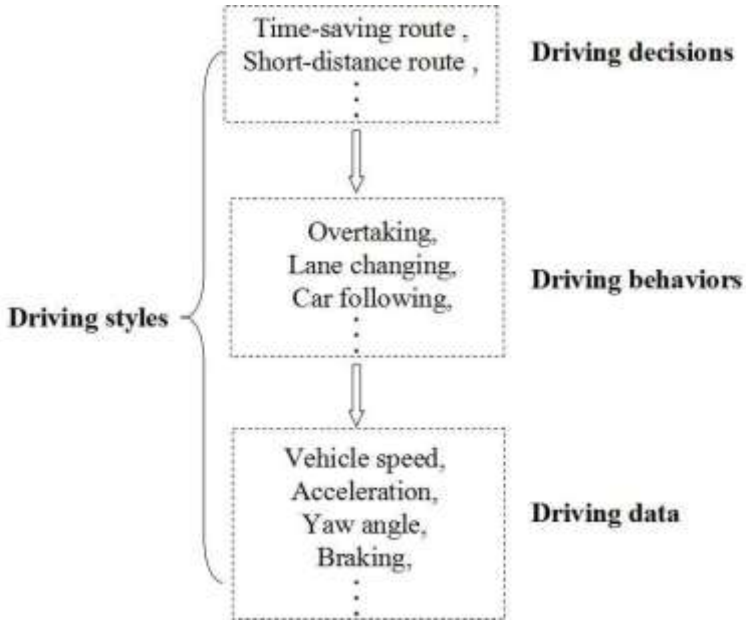


FIGURE 1. Representation for driving styles

There are two methods to determine that which driving style does a driver belong to. The first one is to collect crash rate of driver in many years. The other one is evaluated by experts. Obviously, the first one is not practical as it will spend a lot of times and resources. The second method is adopted for styles classification [19, 20]. Hence, in the process of data collection by volunteers, five experts vote driving style for types, Safety, Moderation and Risk. After data collecting, the type with highest vote is the style of the driver.

FIGURE 2. Driving database creation system

TABLE 2 Part of data in multi-dimensional driving database

Driver#	Longitude	Latitude	Time	Yaw angle (°)	Speed(m/s)	Acceleration(m/s^2)	Braking
1	117.531543	38.966128	19:35:30	15.34	30.3	1.7	0
1	117.531894	38.966192	19:35:31	16.75	31.2	0.9	0
1	117.531295	38.966291	19:35:32	14.39	28.3	-1.9	0
…	…	…	…	…	…	…	…
2	117.286169	39.113681	20:07:12	8.32	13.3	0.1	0
2	117.286164	39.113573	20:07:13	7.19	12.9	-0.4	0
2	117.286155	39.113440	20:07:14	5.87	14.8	1.9	0
…	…	…	…	…	…	…	…
50	117.357180	39.088108	14:48:58	10.8	7.6	-0.9	0.175
50	117.357238	39.088083	14:48:59	11.3	4.7	-2.9	0.183
50	117.357283	39.088073	14:49:00	8.9	2.3	-2.4	0.183

4. METHODOLOGY

Driving styles are estimated based on the created driving database. In this step, our goal is to classify driving styles by driving data. The difficulty is that there are multi-dimensional data which are not easy to find the relationship. As we know, driving style is determined by series of driving behaviors. Hence, we adopt the driving data to represent driving behaviors. Thereafter, we set up the relationship between driving behaviors and driving style. Before we represent driving behaviors, the first thing we should do is data preprocessing.

4.1 Data preprocessing

Driving data from database have huge quantity. We need to reduce data quantity. Considered that the data collection frequency in simulator follows the frequency of GPS data in real word. These types of data have a high frequency with about 1 Hz. Hence, some data keep constant in some scenarios. This problem results in data redundancy.

These data are reduced based on distance. For each type of data, we compute their averages in a constant distance, which is 200 m. The computation is shown as follow:

$$\overline{d_j} = \sum_{i=1}^{n} d_j^i / n \tag{1}$$

where $d_j, (j = 1,2,3)$ is the j th type of data, which d_1 denotes vehicle speed, d_2 denotes vehicle acceleration and d_3 denotes yaw angle. Meanwhile, we also select GPS data in every 200 m. Thus, it can be ensured that each data collection node includes vehicle speed, acceleration, GPS and yaw angle. In this way, the data quantity can be reduced at least 6 times. Note that we also set up two thresholds for acceleration and yaw angle, respectively. If either of $d_j \left(j = 2,3 \right)$ exceeds the threshold, we will stop to remain the original data in this 200-meter route.

4.2 Driving behavior representation

In this step, the driving data are adopted to represent behaviors. We first divide driving behaviors into characteristic and non characteristic. Characteristic behavior includes the lane changing, overtaking, turning, etc. These behaviors are represented by vehicle speed, acceleration, braking, position, yaw angle, etc. To the contrary, non characteristic behavior indicates that vehicle drives in uniform linear motion. Non characteristic behavior is represented by vehicle speed, driving time for uniform linear motion. It is easy to represent non characteristic behaviors. The characteristics behaviors are distinguished from driving behaviors in two ways, GPS based selection and acceleration-yaw-angle based selection.

In GPS based selection, we first match GPS data with some special scenarios, such as crossroads, traffic lights, etc. When volunteers drive to these scenarios, GPS data are matched with these scenarios. In this time, we start to collect data for behaviors representation. Similarly, in acceleration-yaw-angle based selection, we adopt the two thresholds mentioned above for selection. It is because if acceleration or yaw angle exceeds the threshold, the vehicle is faced for lane changing, overtaking, deceleration or other behaviors. These behaviors can usually reflect driving style for driver. Hence, in this time, we also start to collect data for behavior representation.

Furthermore, we represent characteristic behaviors by collected data. First of all, the GPS data are mapped in the global coordinate. The heading angle in each position is computed as follow:

$$\delta^i = d_3^i - d_3^{i-1} \tag{2}$$

where i denotes the number of position. The vehicle speed and acceleration create trajectories. The driving behaviors are shown in Figure 3. The figure shows four behaviors, turning, lane changing, overtaking and car following. The length of arrows denote vehicle speed while the direction of arrow denotes acceleration or deceleration. Black dots denote braking in this position.

(a) (b) (c) (d)

FIGURE 3 Representation for driving behaviors

4.3 Driving behavior classification

We have divided driving behaviors into characteristic and non characteristic. As we know, driving style is decided by multiple driving behaviors. Hence, we divide driving behaviors into three types which are corresponding to driving styles, Safety, Moderation and Risk.

As characteristic behavior only includes two types of data, vehicle speed and driving time, it is easy to evaluate whether the behavior is risk or non risk. Furthermore, we select two thresholds $\sigma_i^j, (i=1,2; j=1,2)$ for each type of data, where σ_1^j denote the thresholds of vehicle speed. σ_2^j denote the thresholds of driving time. The thresholds selection are based on the traffic rules in the route. The behavior type classification for non characteristic is shown in Table 3, where *v* denotes speed; *t* denotes time; R denotes Risk; S denotes Safety; M denotes Moderation.

TABLE 3 Driving behavior type classification for non characteristic

	$t > \sigma_2^{(1)}$	$\sigma_2^{(1)} \geq t > \sigma_2^{(2)}$	$\sigma_2^{(2)} \geq t$
$v > \sigma_1^{(1)}$	R	R	M
$\sigma_1^{(1)} \geq v > \sigma_1^{(2)}$	M	M	M
$\sigma_1^{(2)} \geq v$	M	S	S

There are many types of driving data to represent characteristic behaviors. Hence, we should set up a model to classify these behaviors. In this step, we adopt Support Vector Machine (SVM) to classify characteristic behaviors (Chang et al., 2011). SVM can efficiently perform a non-linear classification and implicitly map the driving data into high-dimensional feature space. In this model, the training data are described as $\{x_i, y_i\}$ $(i=1,2,...q. y_i \in \{-1,1\})$, where x_i denotes the driving data, y_i denote their corresponding labels. The goal for using SVM model is to find a hyperplane which is the best to distinguish the three types of driving behaviors. A hyperplane can be formulated as follow:

$$\omega \cdot x + b = 0 \tag{3}$$

where, ω is the vector perpendicular to the hyperplane. *x* is a point on the hyperplane. *b* is a deviation constant. The illustration for SVM is shown in Figure 4. Therefore, the decision function is formulated as follow:

$$f(x) = \text{sgn}(\omega \cdot x + b) \tag{4}$$

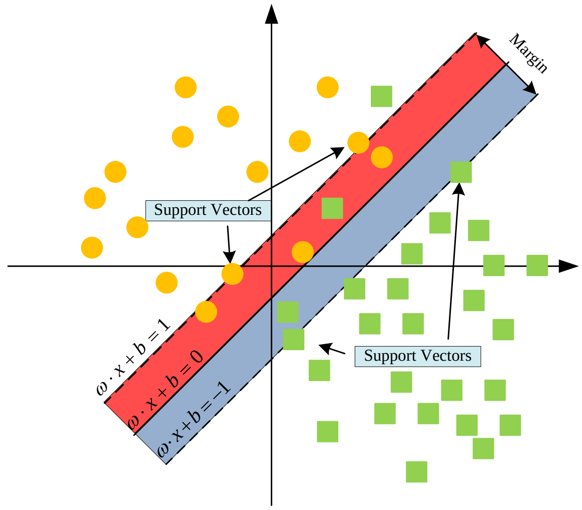


FIGURE 4 Illustration for data clustering by SVM

The margin of hyperplane is denoted as $\frac{2}{\|\omega\|}$. Then the hyperplane can be computed as follow:

$$\min_{\omega,b} \frac{\|\omega\|^2}{2} \tag{5}$$

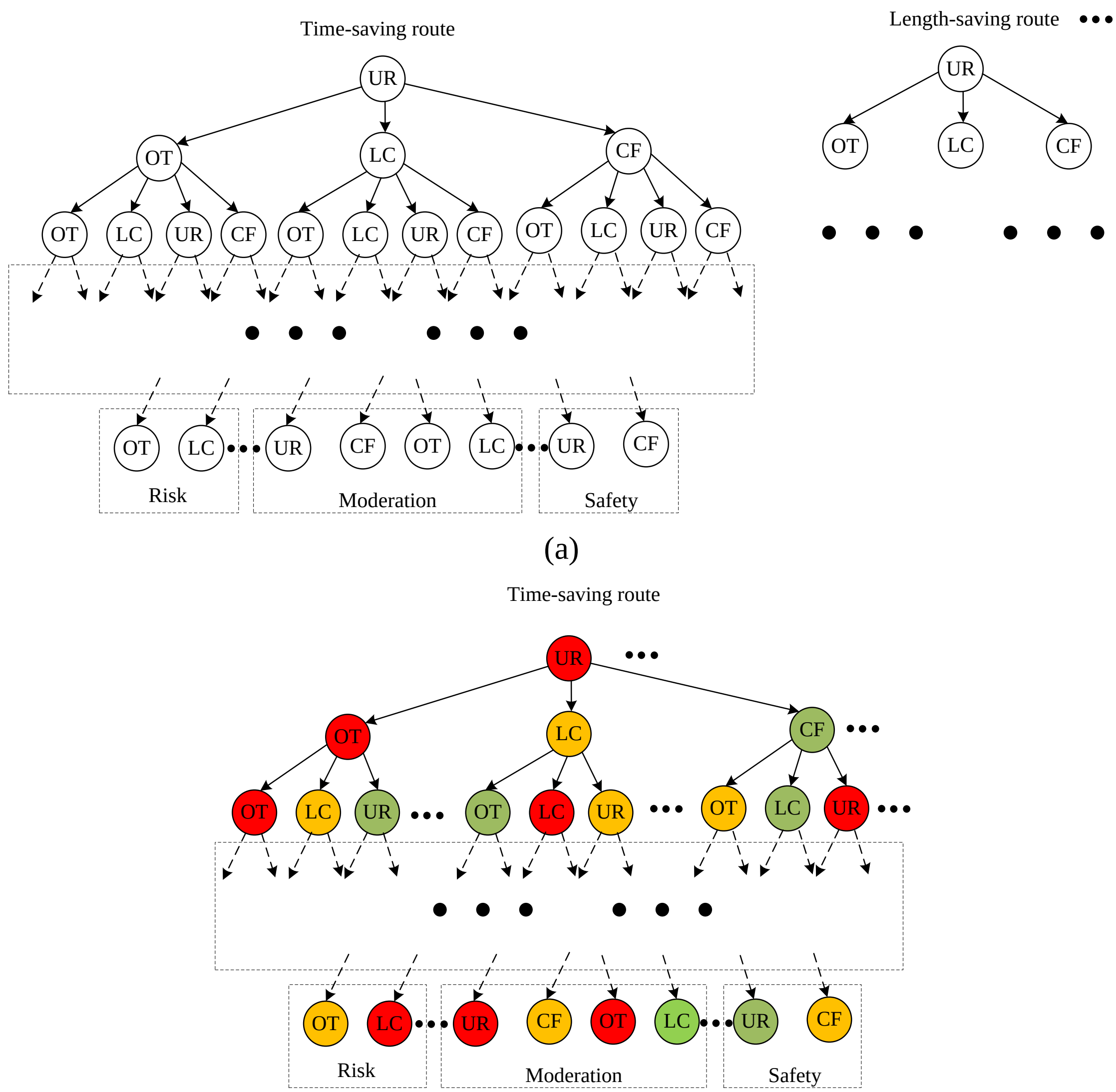
Hence, the driving data can be classified in two types, risk and non risk. Furthermore, we adopt SVM again to classify the non-risk data. As a result, the non-risk data are divided into two types, moderation and safety.

4.4 Driving style classification by multiple decision tree

We have classified the behaviors types. To classify the types of driving styles, we should set up the relationships between multiple behaviors. In this step, we set up a multiple decision tree to make driving style classification.

A driver does a lot of driving behaviors when driving. These behaviors can be represented by decision trees, which is shown in Figure 5 (a), where OT is short for overtaking; UR is short for uniform rectilinear motion; LC is short for lane changing; CF is short for car following. However, only represent these behavior cannot evaluate driving behaviors clearly. Hence, we adopt the classification results of behaviors to evaluate driving styles. Therefore, we set up multiple decision trees for behaviors representation, which is shown in Figure 5 (b). In this sub-figure, the three types of behaviors are represented in three different colors. Red circle denotes Risk; orange circle denotes Moderation and green circle denotes Safety. In the database, experts have evaluated the drivering style for each volunteer. Thus, their corresponding behaviors are also labeled by one type of style. Hence, we can encode each type of driving style by multiple decision trees. As a result, a style description database is set up.

Therefore, when new collection data are input, we first transfer the data into behaviors. Then, behaviors are classified in three types. Last but not least, multiple decision trees are set up and the behaviors are encoded. As a result, we match the created codes of behaviors with style description database. A series of codes which is most similar with the input codes is selected. We decide that the style for the most similar codes is the style of input codes.



(b)

FIGURE 5 Decision trees and multiple decision trees for driving behaviors

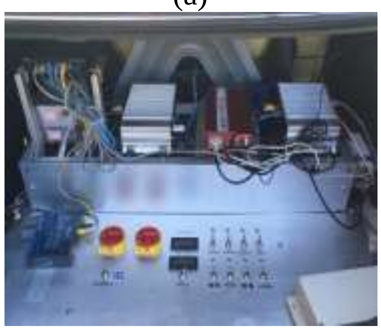
5. CASE STUDY

In case study, we adopt actual roadway to evaluate the driving styles. The test route is a freeway selected in Tianjin, China, which is shown in Figure 6. The total length of test route is about 17 km. 20 drivers are invited as volunteers to collect driving data. The test vehicle is equipped with laser scanner, cameras, odometry, GPS receiver and many other sensors to collect the real-time driving data, which is shown in Figure 7. All the behaviors are in strict compliance with local regulations and they are not deliberately maintained through the test. The collection data include GPS data, speed, acceleration, yaw angle, braking. Then, we adopt these collection data to match with dataset and evaluate driving style for each volunteer. To ensure the accuracy of evaluation, we also invite the same experts who have evaluated the styles in database generation. Three of them are in the test vehicle while the others are in the following vehicle.

Before we evaluated the driving style by using the collection data, we first eliminate outliers by Random Sample Consensus (RANSAC) method. Unlike the data collected by simulator, the data collected in roadway would have some outliers, such as drift error for GPS data. The outliers should have effect on behaviors representation. RANSAC method is good at outlier detection and it is very efficiency to remove the outliers [21].

FIGURE 6 Test route selection

(a) (b)



(c) (d)

FIGURE 7 Data collection system

When the outliers are removed, the next steps are data preprocessing and behavior representation. Furthermore, driving behaviors are classified in three types by the method presented in previous section. As we know, driving behaviors classification is a key task for driving style estimation. To ensure the accuracy, we eliminate some uncertain types of behaviors. We use precision-recall curve to evaluate the behavior classification, which were shown in Figure 8. “Recall” here is the proportion of the number of corrected classification to the number of behaviors detections. We adopted this curve to estimate each type of behaviors, where (a) shows the Risk type, (b) shows the Moderation type, (c) shows the Safety type. In each type, we select the main characteristic behaviors of OT, UR, LC and CF to evaluate the classification. From this figure, when the precision rates achieve 100%, all the behaviors have recalls more than 70% in Risk, more than 50% in moderation, more than 70% in Saftey. And all the recalls of behaviors are more than 80% in each type when the precision rates achieve 80%. More than half of behaviors have recalls more than 90% when the precision rates achieve 80%. These results show that the behaviors classification method has a high accuracy and the results can be efficiently used for further estimation.

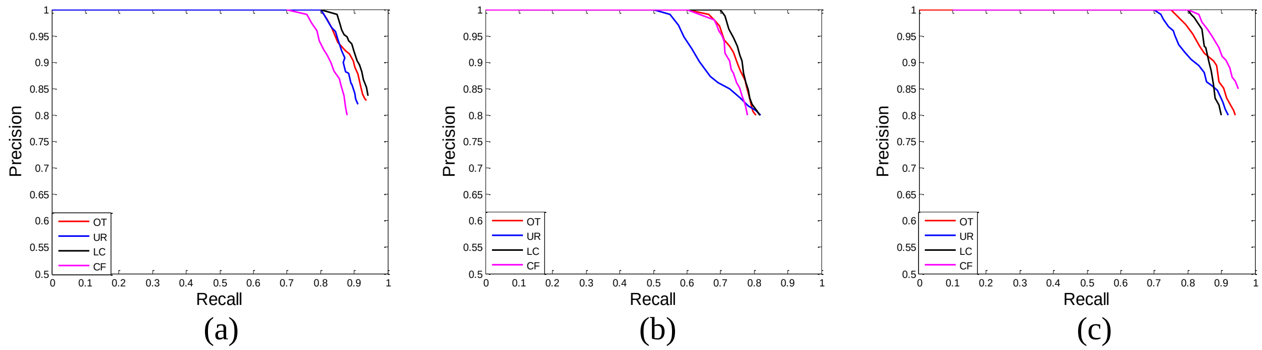


FIGURE 8 Precision-recall curves for types of behaviors: (a) Risk; (b) Moderation; (c) Safety

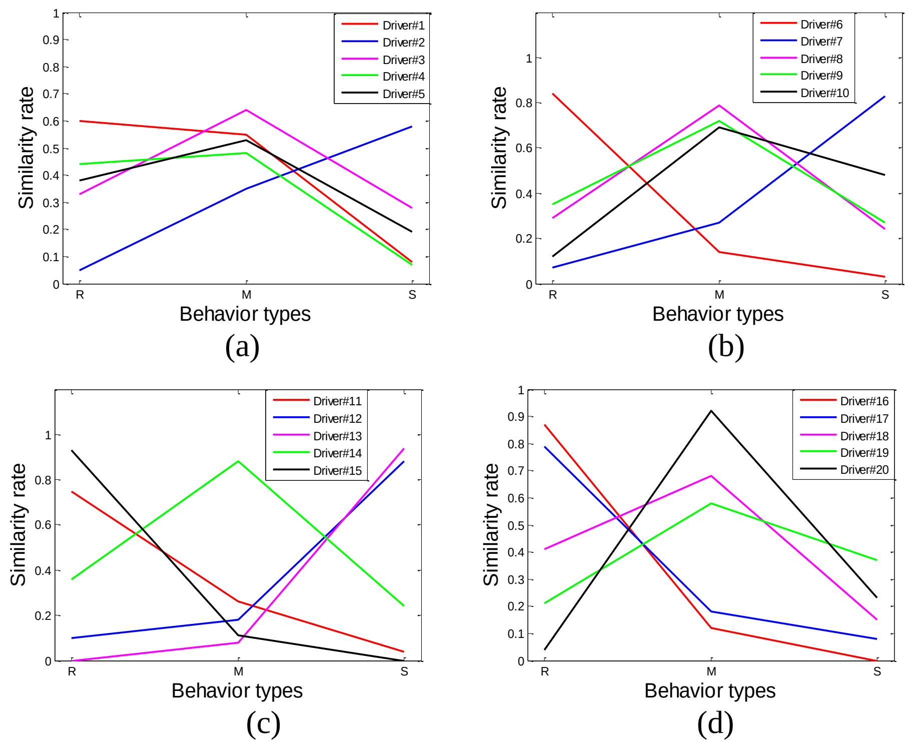


FIGURE 9 Style classification for volunteers: (a) driver 1 to driver 5; (b) driver 6 to driver 10; (c) driver 11 to driver 15; (d) driver 16 to driver 20.

The behavior classification results are used to estimate the style of each driver. Each set of driving data is matched with database by using multiple decision trees. There would be a computed similarity rate for each type of style. As a result, we obtain three similarity rates for all the styles. The style type with the highest rate is selected for volunteers. The matched results for volunteers are shown in Figure 9. From this figure, we can find that there are 10 drivers who have the highest rate for Moderation type. The number of drivers with a Risk style is 6. The other 4 drivers have a Safety style. However, what is accuracy of styles estimation? The results are compared with the ground truth which are collected by five experts. The comparison results are shown in Table 4. From this table, we can find that all the estimation is correct except driver 1. The ground truth is Moderation but the estimation is Risk. It is because the two similarity rates are close which can be found in Figure 9 (a). This driver has a tendency to become a Risk style due to the similarity rate. Hence, it is helpful to estimate driver 1 to a Risk style. Overall, the proposed method has a 95% accuracy for style estimation, which proves that this method is reliable.

According to case study, we can find that about a quarter of volunteers have a Risk style. It is very dangerous not ony for drivers themselves, but for their passengers and pedestrians. They need to enhance security awareness and change their driving styles. Moreover, there are also some drivers with a Safety style. Although they are more safe in driving, they may cause a waste of time and energy. The target of our study is to appeal to the drivers with bad styles to change their styles. We will get a safer trip when travelling by vehicle.

TABLE 4 Results comparison with ground truth

Driver #	Estimated by the proposed method	Ground truth	Driver #	Estimated by the proposed method	Ground truth
1	Risk	Moderation	11	Risk	Risk
2	Safety	Safety	12	Safety	Safety
3	Moderation	Moderation	13	Safety	Safety
4	Moderation	Moderation	14	Moderation	Moderation
5	Moderation	Moderation	15	Risk	Risk
6	Risk	Risk	16	Risk	Risk
7	Safety	Safety	17	Risk	Risk
8	Moderation	Moderation	18	Moderation	Moderation
9	Moderation	Moderation	19	Moderation	Moderation
10	Moderation	Moderation	20	Moderation	Moderation

6. CONCLUSIONS

This paper presents a driving style estimation method to enhance drivers' awareness of traffic safety. The method is realized by fusing multiple driving behaviors which are based on actual driving data. Firstly, the method adopts actual driving data to represent driving behaviors. Then, driving behaviors are classified by SVM model. Last but not least, driving styles are estimated by exploring the relationship between driving behaviors. As a result, we select 20 volunteers in China to estimate their styles as casy study. The results show that the proposed method has a 95% accuracy for style estimation. This paper also reveals that about a quarter of drivers have a style of Risk and many drivers have a style of Safety in China. The study presented in this paper can alert the driver with bad styles to change their habits. It is helpful to ease traffic accidents and save travel time. In the future, we will consider the factor of gender difference. It is also a very important factor to affect the driving styles. Furthermore, we will also conduct some experiments in ordinary urban roads to evaluate the proposed method.

ACKNOWLEDGEMENTS

The work presented in this paper was funded bythe National Key Research and Development Program (2016YFB0100903) and the National Natural Science Foundation of China (61503284 and 51505475), Yingcai Project of CUMT (YC2017001), Priority Academic Program Development of Jiangsu Higher Education Institutions, and the UOW Vice-Chancellor's Postdoctoral Research Fellowship.

REFERENCES

1. Milanés, V., S. E. Shladover. Handling cut-in vehicles in strings of cooperative adaptive cruise control vehicles. *Journal of Intelligent Transportation Systems*, Vol. 20, No. 2, 2016, pp. 178-191.
2. Li, Z, Y. Jiang, Q. Guo, C. Hu, Z. Peng. Multi-dimensional variational mode decomposition for bearing-crack detection in wind turbines with large driving-speed variations. *Renewable Energy*, Vol. 116, 2016, pp. 55-73.
3. Traffic Management Bureau of the Public Security Ministry. Statistics of China's Road Traffic Accidents. Wuxi: Traffic Management Bureau of the Public Security Ministry, 2013.
4. Wang, J., C. Xiong, M. Lu, and K. Li. Longitudinal Driving Behavior on Different Roadway

Categories: an Instrumented-vehicle Experiment, Data Collection and Case Study in China. *IET Intelligent Transport Systems*, Vol. 9, No. 501 5, 2015a, pp. 555-563.

5. Staubach, M. Factors Correlated with Traffic Accidents as a Basis for Evaluating Advanced Driver Assistance Systems. *Accident Analysis & Prevention*, Vol. 41, No. 5, 2009, pp. 1025-1033.
6. Lei, H., X. Yan, C. Wu et al. A Study of Chinese Motorist's Physiological Characteristics under Angry Driving. *In Transportation Research Record: Journal of the Transportation Research Board*, Washington, D.C., 2013.
7. Kulkarni, P and V. G. Sangam. Smart Helmet for Hazardous event Detection and Evaluation in Mining Industries using Wireless Communication. *Journal of Communication Engineering and Its Innovations*, Vol. 3, No. 1, 2017, pp. 11-16.
8. Behr, C. J., A. Kumar, G. P. Hancke. A Smart Helmet for Air Quality and Hazardous Event Detection for the Mining Industry. *Industrial Technology (ICIT)*, 2016 IEEE, IEEE, 2016, pp. 2026-2031.
9. Aarts, L. and I. V. Schagen. Driving Speed and the Risk of Road Crashes: A Review. *Accident Analysis & Prevention*, Vol. 38, No. 2, 2006, pp. 215-224.
10. Wu, C., C. Sun and D. Chu. Several Typical Driving Behavior Characteristics Clustering of Commercial Vehicle Based on GPS Data Mining: A Case Study of Highways in China. *In Transportation Research Record: Journal of the Transportation Research Board*. 2016, Vol. 2581, pp. 154-163.
11. Zito, R., G. d'Este and M. A. P. Taylor. Global Positioning Systems in the Time Domain: How Useful a Tool for Intelligent Vehicle-highway Systems? *Transportation Research Part C: Emerging Technologies*, Vol. 3, No. 4, 1995, pp. 533 193-209.
12. Yan, L., Y. Zhang, Y. He, et al. Hazardous Traffic Event Detection Using Markov Blanket and Sequential Minimal Optimization (MB-SMO). *Sensors*, Vol. 16, No. 7, 2016, pp. 10847.
13. Xiong, H., L.N. Boyle, J. Moeckli, et al. Use Patterns Among Early Adopters of Adaptive Cruise Control. *Human Factors*, Vol. 54, No. 5, 2012, pp. 722-733

14 Sagberg, F., G. Piccinini, J. Engström, et al. A Review of Sesearch on Driving Styles and Road Safety. *Human Factors*, Vol. 57, No. 7, 2015, 1248-1275.

15. Guo, F., Y. Fang. Individual Driver Risk Assessment Using Naturalistic Driving Data. *Accident Analysis & Prevention.* Vol. 61, 2013, pp. 3-9.
16. Wang, J., Y. Zheng, X. Li, et al. 2015b. Driving Risk Assessment Using Near-crash Database through Data Mining of Tree-based Model. *Accident Analysis & Prevention.* Vol. 84, 2015b, pp: 54-64.
17. Bonsall, P., R. Liu, and W. Young. Modelling Safety-related Driving Behaviour-impact of Parameter Values. *Transportation Research Part A: Policy and Practice*, Vol. 39, No. 5, 2005, pp. 425-444.
18. Li G., S. E. Li, B. Cheng, et al. Estimation of Driving Style in Naturalistic Highway Traffic Using Maneuver Transition Probabilities. *Transportation Research Part C: Emerging Technologies*, Vol. 74, 2017, pp. 113-125.
19. Holman, A. C. and C. E. Havârneanu. The Romanian Version of the Multidimensional Driving Style Inventory: Psychometric Properties and Cultural Specificities. Transportation research part F: Traffic Psychology and Behavior, Vol. 35, 2015, pp. 45-59.
20. Chang, C. and C. Lin. LIBSVM: A Library for Support Vector Machines. *ACM Transactions on Intelligent Systems and Technology*, Vol. 2, No. 3, 2011, pp. 289-396.
21. Naredo, E. and S. Silva. RANSAC-GP: Dealing with Outliers in Symbolic Regression with Genetic Programming. *Genetic Programming: 20th European Conference*, (EuroGP 2017), 2017, Springer, 2017, pp. 114.