Sensors & Transducers, Vol. 167, Issue 3, March 2014, pp. 194-202

Sensors & Transducers

http://www.sensorsportal.com

A Study of Individual Characteristics of Driving Behavior Based on Hidden Markov Model

Xingjian Zhang, Xiaohua Zhao, Jian Rong
Beijing University of Technology, 100124, Beijing, China
Tel: +86-10-67396075, fax: +86-10-67391509
E-mail: jzhyoung@163.com

Received: 15 March 20014 /Accepted: 30 March 2014 /Published: 31 March 2014

Abstract: Drivers' individual difference is one of the key factors to influence the accuracy of driving behavior model. The accuracy of model should include the effect characteristics of individual difference on driving behavior. The overtaking process was the research object to study the individual characteristics of driving behavior. The operation data of accelerator and steering wheel of each driver was analyzed with the character of time series. Based on both of the operation data, hidden Markov model (HMM) was employed to model the individual characteristics of driving behavior. Two individual models were built for each driver, one trained from accelerator data and one learned from steering wheel angel data. The models can be used to identify different drivers and the accuracy can reach to 85 %. It proved that individual difference is one factor which cannot be ignored in driving behavior model, and HMM has effectiveness in modeling it. *Copyright © 2014 IFSA Publishing, S. L.*

Keywords: Individual characteristics, Driving behavior, Hidden Markov model, Driver identification.

1. Introduction

As the increasing number of vehicle and road mileage, traffic safety is facing great pressure. Nearly 80 % of the traffic accidents are caused by drivers themselves in road transport system [1-3]. These accidents may be just because of that drivers have no timely or proper operations when needed. In order to improve traffic safety level, it is one of the key points to deeply analyze the characteristics of driving behavior and study the effects of driving behavior on vehicles and traffic system. There is an emerging consensus that driving behavior is the direct response in driving process.

Driving behavior has attracted attention of many researchers. They attempt to analyze and describe driving behavior characteristics with a variety of mathematical models. Amata studied the feature of driving operation behavior with linear regression and Bayesian network model, and proved that the characteristics of driving behavior were different among different drivers [4]. Based on different driving behavior features of different drivers, Wakita as well as Miyajima modeled it with Mixture of Gaussians to identify different drivers [5, 6]. Suzuki proved that stochastic switched autoregressive exogenous model held remarkable potential to function as a behavior recognizer [7]. Some researchers paid their attention to the prediction of vehicle running state. They forecasted the running state or stop probability with HMM to model driving behavior in different way and proved the effectiveness of Hidden Markov Model (HMM) [8-11]. D. Mitrovic recognized driving events based HMMs with a high recognition rate of 93.8 % using only vehicle speed and acceleration as raw data [12].

Article number P_RP_0097

And there are also an amount of researches and applications on driver assistance technology with driving behavior model [13, 14].

The above researches indicate that drivers' personality have significant impact on driving behavior. But it can be seen that most of the researchers do not consider drivers' individual difference as the key influencing factor. However, different drivers have different driving characteristics. Individual differences are due to the different driving characteristics and driving habits of drivers. The difference makes driving behavior models even more complicated. Many studies mainly focus on the modeling and description with some mathematical model. The most important thing of building model is accuracy. But it is very difficult to find a very accurate model to represent the features of driving behavior. To be more accuracy, the model should not be one macroscopic driving behavior model, but individual model with mechanisms of learning and training.

An accurate model should be the combination of driving behavior analysis and individual difference. Driving behavior generally includes drivers' operation behaviors, such as steering wheel, accelerator, brake, and running state of vehicle, which consists of speed, acceleration and lane position. Generally, drivers adopt to simple control laws instead of complex nonlinear control law when driving. Sometimes much complex model is not only too complicated, but makes it difficult to understand the physical meaning of the drivers' maneuver [7]. Considering the time series feature of driving behavior, we employ HMM, which has the mechanisms of learning and training, to model each driver's driving behavior. The aim of this study is to validate that individual difference have significant influence on driving behavior and it can be modeled with mathematical model. Then we can identify each driver according to the models.

2. Method

The main purpose of this experiment is to collect data of driving behavior, including various kinds of behavior data under common vehicle operating status, such as acceleration, deceleration and turning. In order to collect data more conveniently and ensure that the vehicle was operated in an ideal driving status, the experiment virtual road scenario and environment were designed based on the AutoSim driving simulator. Subjects were required to accomplish the experimental tasks according to the designed process.

2.1. Apparatus

Driving simulator can provide virtual three-dimensional driving scenarios for subjects with vehicle dynamic model based on computer simulation. The system provides a screen with 130 angle of view in front, two rearview mirrors on each side and one screen back. The system can record all data of the drivers' operation behaviors and vehicle running state with the frequency of 30 Hz to support related study.

2.2. Scenario

The scenario design included road facilities design and event design. The road facilities included route, traffic sign and relative driving environments. We designed 3 different scenarios to avoid the impact of same driving scenario on driver. The experimental road had two lanes in each direction. Basic parameters of the road as well as facilities and markings were designed according to the national standard of China. The road alignment and facilities were designed based on the research needs. The goal of experiment was to obtain driving behavior data including operational data and vehicle running data. Thus, the design should provide different road conditions so that subjects could perform various driving behavior. The experimental road is shown as Fig. 1.

The road was 14 kilometers long, including 12 straight lines, 5 horizontal curves, 4 obstacle zones and 3 parking bays. They were combined to introduce different driving behavior. The expected driving behaviors for each component are shown in Table 1.

Except the road facilities design, the events were designed at the same time. The purpose of event design was getting enough driving behavior data. The limiting speed was designed respectively 80 km/h, 60 km/h, 30 km/h, corresponding to straight lines, horizontal curves and obstacle zones. Other vehicles were added in scenarios to impel following and overtaking process. We designed 9 different events with the combination between road facilities and driving tasks. Subjects could accomplish various expected operations. The driving events and expected operational status are shown in Table 2. Some of the events were controlled by the audio instruction, such as overtaking and stop. The instructions were given when the vehicle reached the design locations.

2.3. Participants

Male and female drivers perform few differences at common driving circumstances, however, there are differences in manner and psychology level when emergency occurs. Drivers aging 21~40 are more likely to encounter traffic accidents. Besides, because of fewer driving experiences, there are more obvious potential safety hazards related to drivers who have less than 3 driving years, otherwise, drivers who have more than 3 driving years can drive more smoothly and steadily [15].

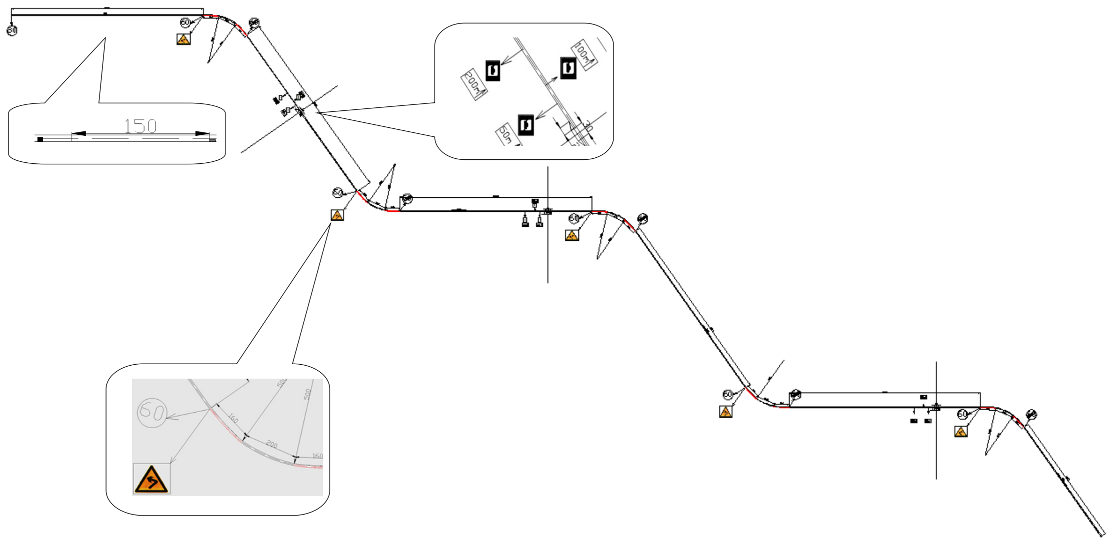

Fig. 1. Route in experiment.

Table 1. Expected driving behavior for each component.

Road elements	Expected driving status	Expected driving behavior
Straight line	Starting, acceleration, deceleration, driving in a constant speed	Accelerator, brake, clutch, controlling speed
Horizontal curve	Turn left, turn right	Steering wheel
Obstacle zone	Acceleration, deceleration, changing lanes	Brake, accelerator, steering wheel
Parking bay	Parking and starting	Gears, accelerator, brake, clutch

Table 2. Expected driving behavior for different event.

Event	Number	Expected driving behavior
Start	E1	gear, accelerator, acceleration
Uniform velocity	E2	speed, steering wheel
Deceleration	E3	brake, gear, steering wheel
Change lane	E4	accelerator, brake, steering wheel
Acceleration	E5	accelerator, acceleration
Turning	E6	steering wheel
Following	E7	accelerator, acceleration, brake, steering wheel
Overtaking	E8	accelerator, acceleration, steering wheel
Stop	E9	acceleration, brake, steering wheel

This research was a kind of model analysis of the regular pattern of driver's individual driving behavior, which was aiming at serving to the development of traffic safety technology and reducing the number of traffic accidents. Meanwhile, this research also focused on testifying the individual variation. We assumed that there were different characteristics of driving behavior even among the same kind group. Therefore, 20 male subjects were recruited, and all of them had almost the same attribute. According to the request, they were all in healthy conditions, completely sober while driving, and not fatigue, drunken or taking medicine. The ages and driving years of 20 subjects are shown in Fig. 2.

2.4. Procedure

Some researchers suggest that the fatigue will increase rapidly when a driver drives more than 3 hours. Moreover, drivers are less likely to be fatigued at 9:00-11:00 am, 4:00-5:30 pm, and 7:00-9:00 pm [16-19]. To avoid the interference from fatigue, the experiment was conducted at the time above.

To reduce the contingency, each subject was required to drive at forenoon, afternoon, and evening on different day. The scenario time of all experiments were set 8:00 pm by sending time command of simulator. Each experiment took place for almost 30 minutes. They were required to obey the limiting speed and instructions, and keep driving on the right lane except essential condition like overtaking or obstacle zone. The experimental procedure for each subject is shown in Fig. 3.

2.5. Data Collection

The experiment was conducted in strict accordance with the designed request and process so as to guarantee the effectiveness of data objectively. Through the experiment, basic data of 20 subjects' driving behavior were collected, which included expected driving operations and fitted the needs of this study. Subjects were all in normal conditions and out of such special circumstances as fatigued or drunken.

The similarity of simulator was evaluated and graded by each subject. They evaluated the simulator according to their own feel, and the evaluation items include overall feel, vehicle manipulation, experimental scene, and so on. The best refers to 10 and the worst refers to 0, the average scores are shown in Fig. 4. The results suggested that the simulator had not a perfect biofidelity. Especially the brake imitated real operation poorly. It may affect the findings' value of practical application. But because some experiments cannot be taken place in real driving, and the simulator provides the foundations and conditions for the research. At the same time, we hoped to model every driver's driving behavior and prove the characteristics and personal habit in this research. All the drivers drove in the same simulator. The experimental driving operations had the same basic environment, and then the data collected could support the research.

3. HMM Introduction

Time series is one of the most important features of driving behavior, especially in some driving processes. HMM is good at modeling the feature. So it was employed to model driving behavior in this paper. The mathematic theory of it was introduced from the following aspects.

3.1. HMM

HMM, which is expressed with parameters, is a model used to describe the statistical properties of a random process. It is also a double stochastic process: Markov chains and general random process. Markov chains are used to describe the transfer of states, and are expressed by transition probability. General random process is used to describe the relationship between state and observation sequence, and is expressed by observation probability. Because the states of HMM are unobservable, the model is called "Hidden" Markov Model. The structure of HMM is shown as Fig. 5.

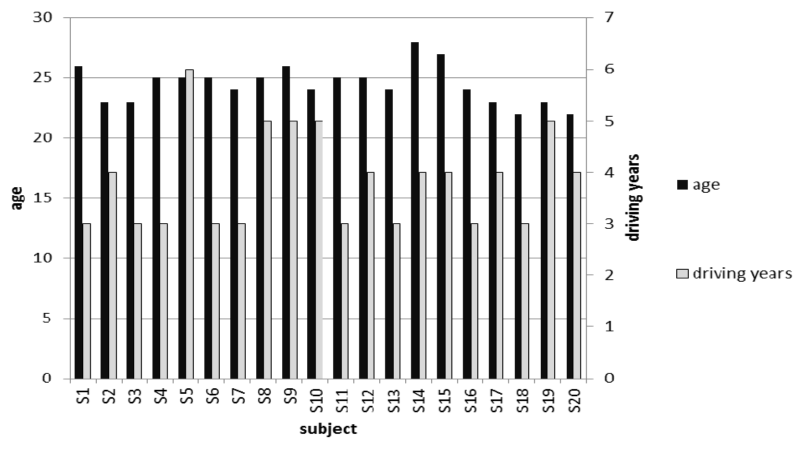

Fig. 2. Age and driving years of each subject.

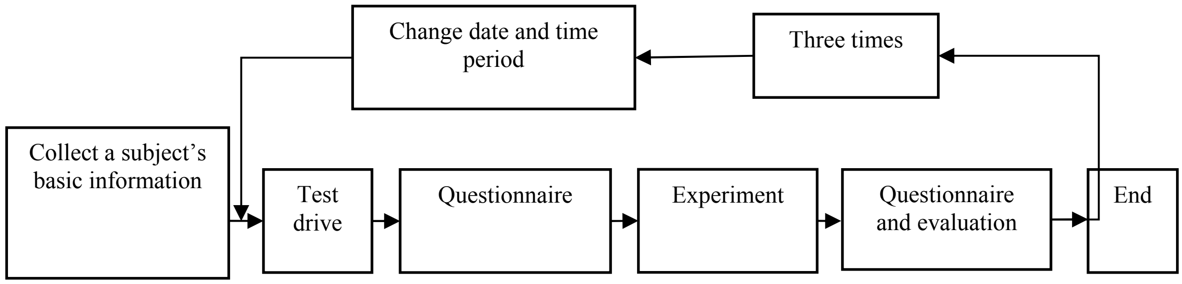

Fig. 3. Experimental procedure for each driver.

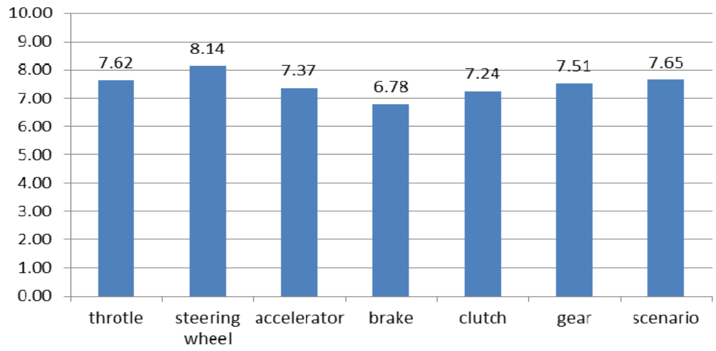

Fig. 4. Average score of subjective evaluation.

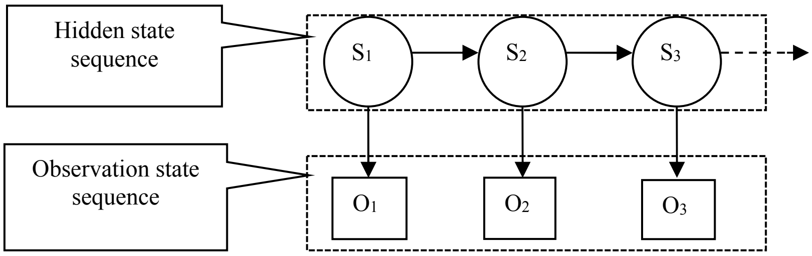

Fig. 5. Structure of HMM.

HMM is defined as

$$\delta(t+1) = \sum_i a_{i,j}\delta_t(t)$$
$$O(t) = N(\mu_i, \sum_i) \quad (1)$$

where t is the discrete time, $\delta_t(t)$ is the probability of state i at time t. $a_{i,j}$ is the state transition probability from state i to j. $O(t)$ is the observable value vector at time t. $N(\mu, \sum)$ is normal distribution and its mean is μ, covariance matrix is Σ.

3.2. HMM Algorithm

The three basis questions of using HMM are as following. They are the key points of HMM.

Assessment: Given observation sequence and model how to calculate the probability of observation sequences' emergence.

Decoding: Given observation sequence and model, seek the optimal hidden state sequence in a meaningful case.

Learning: Given observation sequence, how to regulate the model parameters to get the optimal model.

Assessment and learning problems are used in this paper. The mainly algorithm of them are as following.

3.3. HMM Application

HMM is built based on probability theory. Its biggest superiority is the capability to model the transformation character of time series. At the same time, time series is one of the most important characters of driving behavior. HMM is a naturally suitable tool to model driver behavior for two reasons [20, 21]:

1) HMMs can model the stochastic nature of the driving behavior, providing sufficient statistical smoothing while offering effective temporal modeling. It works very well in practice for several important applications.

2) The variations in the driving signals across the drivers can be modeled.

Generally, the use process of HMM is mainly three key points:

3) Define the hidden state and observation sequence. Get the number of hidden state and make pretreatment to observation data for training model.

4) Learn and train model.

5) Validation model.

4. Result

4.1. Driving Status

The mainly goal of this section was to build individual driving behavior model for each driver and each model could include the driving behavior features of corresponding driver. HMM was adopted to account for dynamic behavior model. It modeled each dynamic human behavior by optimizing its parameters, which were learned from training driving behavior data. Each driver has his own a group of HMM parameters. When the parameters of every model are trained and optimized thoroughly, the model can be used to recognize the unknown dynamic behavior data. In the process of recognition, the test data sequence will be graded by every model and the one whose model gives the highest score will be selected to be the identification result [22]. The score is log-likelihood. It indicates the similar level between model and driving behavior or between model and test data. The higher of the log-likelihood, the more similar they are. The maximum value of it is 0.

The study and training of model need different diving behavior data in driving process, such as starting, turning and accelerating process. The most superiority of HMM is its description capability for time series characteristics of signal. In this paper, the studied target we select was overtaking process, which was common driving process and contained accelerating, changing lane and overtaking processes. The mainly process was shown as Fig. 6. Accelerator and steering wheel were the mainly operation in overtaking process. The two operations affected each other and the overtaking process needed they match well. At the same time, Fig. 4 showed that the simulator had a higher biofidelity for drivers on steering and accelerator. So we selected the two driving behavior data sequence and model each of them with HMM.

In the experiment, each driver had 6 times to overtake totally. The average speed of overtaking was about 80 km/h and the distance was around 300 meters. It was needed 15 seconds to complete the overtaking process, including 5-6 seconds for accelerating, 4-5 seconds for turning and 4-6 seconds for overtaking. We cut out 15 seconds data for each overtaking process, which included 450 data point. It meant that each driver had 2700 data sequences for each of accelerator and steering wheel angle.

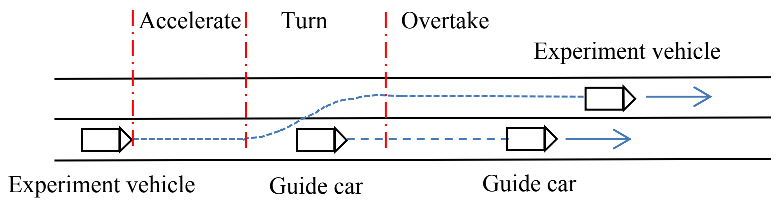

Fig. 6. Process of overtaking.

4.2. Model Training

We trained HMMs with each of accelerator and steering wheel angel data for every driver. All the data was preprocessed and standardized and we got a data matrix of size 6×450 from each of the two operations for every driver. The Baum-Welch algorithm was employed to learn 5×450 of them for the HMMs and the Forward-Backward algorithm was used to evaluate the other 1×450 data segments. On the process of overtaking, drivers mainly have 5 mental states, including following, overtaking, accelerating, decelerating and turning. The current vehicle state include 6 states, which are uniform linear motion, linear acceleration, linear deceleration, turning uniform motion, turning acceleration and turning deceleration. So based on drivers' mental state and current vehicle state, we divided the hidden driving state to 30 states. The standardized data were divided 100 levels, so we defined the observation-state as 100 states. The 30-state HMM model was trained for 20 drivers. We iterated 20 times with the data for every model based on the Baum-Welch algorithm. Each driver has two individual HMM models, one model trained by accelerator data and one trained by steering wheel data. The initial matrix of each model was initialized with uniform segmentation of each data sequence.

The log-likelihood values derived from the 20 iterations of HMM training process were shown on Fig. 7 and 8. Fig. 7 showed the log-likelihood variation tendency of 20 HMMs learn from accelerator data sequence and Fig. 8 showed it from steering wheel data. According to Baum-Welch algorithm, HMM will be more accuracy as the increasing iteration times. In Fig. 7, the HMMs parameters would be stable after about 12 iterations and they needed at least 14 times in Fig. 8.

4.3. Identification Result Verification

The Forward-Backward algorithm was used to evaluate the test data segment. Each sequence data were evaluated by 20 drivers' individual HMMs to get 20 log-likelihood values. We will select the maximum value and the corresponding HMM will be as the identification object. So the identification methodology was to confirm that which HMM had the maximum log-likelihood.

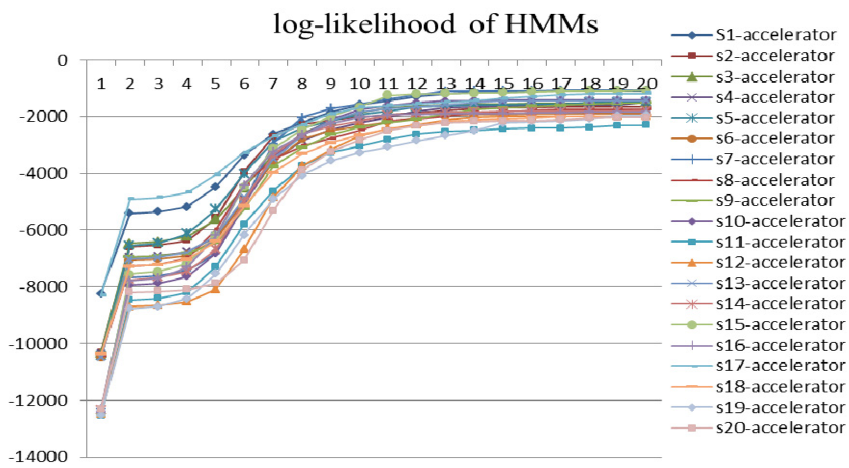

Fig. 7. Log-likelihood of HMMs trained from accelerator.

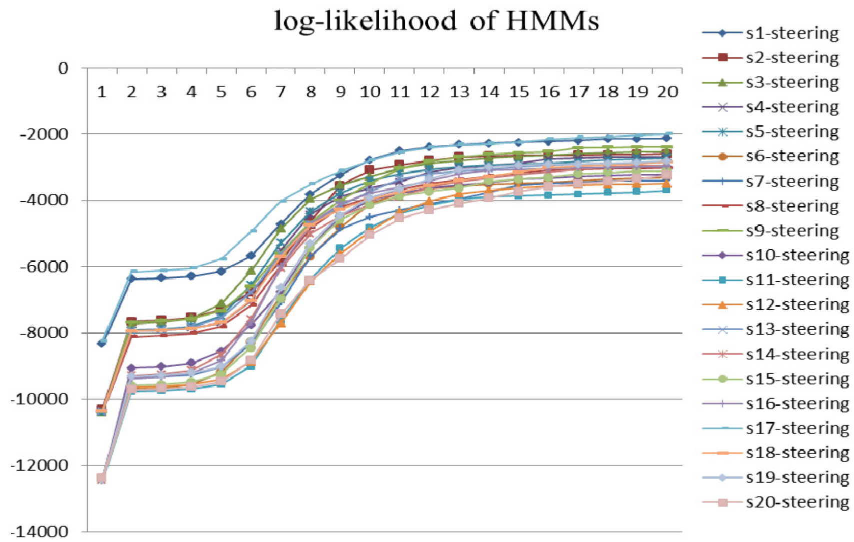

Fig. 8. Log-likelihood of HMMs trained from steering wheel.

According to the method of identification, we could identify which driver the new data sequences belong to. The test data included 20 accelerator data segments and 20 steering wheel angel data segments, and they were all from the 20 subjects. The accuracy rate of recognition used only accelerator HMMs was 70 %. And it was 75 % used only steering wheel HMMs. We combined both of them and calculated new identification score for every driver. First, we standardized all the log-likelihood values to get two new scores for each driver. And the new identification score for every subject was the sum of the two values. Then, we recognized the drivers with the same method. The accuracy rate reached to 85 %. Table 3 shows the identification details.

Based on the personality differences, we built individual models for everyone. Most of the models can represent the characteristics of corresponding driver and we can recognize unknown data with the models. It indicated that personality differences were important factors to affect driving behavior which could not be ignored. We can model it through HMM, considering the time series feature of driving behavior.

Table 3. Identification details with each model.

Identification with models	Correct/total number	Percent
Accelerator	14/20	70 %
Steering wheel	15/20	75 %
Accelerator & steering wheel	17/20	85 %

5. Conclusion and Discussion

In the analysis of traffic safety, the driver's driving behavior has been studied in many aspects, and driving behavior's modeling has also become the focus of driving behavior research. However, no model was studied to model individual differences characteristics of driving behavior. This paper focused on proving the individual characteristic in the driving behavior. We tried to build individual models to improve the driving behavior model's accuracy. Based on the time series feature of driving behavior data, HMM was employed to be the model framework. We can get the following conclusions:

1) Individual difference is one key factor which cannot be ignored when building driving behavior model.

2) HMM hold the effectiveness of modeling individual difference of driving behavior.

3) HMMs trained from each driver's driving behavior data can be used in driver identification. The identification accuracy will be higher with combination of more driving behavior data.

4) The analysis of driving behavior model had a far-reaching significance in many aspects, which had a great role in raising the traffic safety level. The establishment of driving behavior feature model and its application has proved the individual feature model's feasibility and effectiveness. It promotes the establishment of a more accurate and comprehensive model to a certain degree. However, studying from only one state, this model is still at the preliminary stage of the research. A more accurate and comprehensive driving behavior model requires an all-round analysis of various driving states and driving process, especially individual difference of drivers' characteristics. We believe that the establishment of driving behavior model based on the drivers' individual difference, with training and learning method, will have better results, and it will get further applications in the related research.

In the following research, by expanding the sample size, increasing drivers' driving states and collecting more driving behavior data in various conditions, we will build driving behavior database for each driver. It will provide a rich data base for the establishment of feature models. We will try to build a more accurate and applicable individual feature model of driving behavior for each driver.

Acknowledgements

This study is supported by the National Nature Science Foundation of China project: The Study of the Mechanism for Traffic Signs Influence on Driving Behavior and Its Cognitive Model, No. 51108011; and Beijing Nature Science Foundation project: The Study of Identification Method of Drunk Driving Based on Driving Behavior of Driving Personal Character, No. 8112004.

References

[1]. Ravenholt, R. T., Addiction mortality in the United States, 1980: Tobacco, alcohol, and other substances, *Population and Development Review*, 1984, pp. 697-724.

[2]. Hayakawa, H., Fischbeck, P. S., Fischhoff, B., Traffic accident statistics and risk perceptions in Japan and the United States, *Accident Analysis and Prevention*, 32, 6, 2000, pp. 827-835.

[3]. Norris, F. H., Matthews, B. A., Riad, J. K., Characterological, situational, and behavioral risk factors for motor vehicle accidents: a prospective examination, *Accident Analysis and Prevention*, 32, 4, 2000, pp. 505-515.

[4]. Amata, H., Miyajima, C., Nishino, T., Kitaoka, N., Takeda, K., Prediction model of driving behavior based on traffic conditions and driver types, *in Proceedings of the 12th International IEEE Conference on Intelligent Transportation Systems, (ITSC'09),* 2009, pp. 747-752.

[5]. Wakita, T., Ozawa, K., Miyajima, C., Igarashi, K., Katunobu, I. T. O. U., Takeda, K., Itakura, F., Driver identification using driving behavior signals, *IEICE Transactions on Information and Systems*, 89, 3, 2006, pp. 1188-1194.

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

[7]. Sekizawa, S., Inagaki, S., Suzuki, T., Hayakawa, S., Tsuchida, N., Tsuda, T., Fujinami, H., Modeling and recognition of driving behavior based on stochastic switched ARX model, *Intelligent Transportation Systems, IEEE Transactions,* 8, 4, 2007, pp. 593-606.

[8]. Oliver, N., Pentland, A. P., Graphical models for driver behavior recognition in a smart car, in *Proceedings of the IEEE Intelligent Vehicles Symposium,* 2000, pp. 7-12.

[9]. Kumagai, T., Sakaguchi, Y., Okuwa, M., Akamatsu, M., Prediction of driving behavior through probabilistic inference, in *Proceedings of the 8th International Conference on Engineering Applications of Neural Networks,* 2003, pp. 8-10.

[10]. Abe, K., Miyatake, H., Oguri, K., A study on switching AR-HMM driving behavior model depending on driver's states, in *Proceedings of the Intelligent Transportation Systems Conference, (ITSC'07),* 2007, pp. 806-811.

[11]. Kishimoto, Y., Oguri, K., A Modeling Method for Predicting Driving Behavior Concerning with Driver's Past Movements, in *Proceedings of the IEEE International Conference on Vehicular Electronics and Safety*, 2008, pp. 132-136.

[12]. Mitrovic, D., Reliable method for driving events recognition, *IEEE Transactions on Intelligent Transportation Systems,* 6, 2, 2005, pp. 198-205.

[13]. Akamatsu, M., Sakaguchi, Y., Okuwa, M., Modeling of driving behavior when approaching an intersection based on measured behavioral data on an actual road, in *Proceedings of the Human Factors and Ergonomics Society Annual Meeting,* 47, 16, 2003, pp. 1895-1899.

[14]. Torkkola, K., Venkatesan, S., Liu, H., Sensor Sequence Modeling for Driving, *in Proceedings of the FLAIRS Conference,* 2005, pp. 721-727.

[15]. Ma, Y., Study on Characteristics of Driving and Its Countermeasures to Road Safety, Doctoral Degree Dissertation, *Harbin Institute of Technology*, 2007.

[16]. Barr, L. C., David Yang, C. Y., Hanowski, R. J., Olson, R., Assessment of driver fatigue, distraction, and performance in a naturalistic setting, *Transportation Research Record: Journal of the*

Transportation Research Board, 1937, 1, 2005, pp. 51-60.

[17]. Kim, K. E., Yamashita, E. Y., Asleep at the wheel: spatial and temporal patterns of fatigue-related crashes in Honolulu, *Transportation Research Record: Journal of the Transportation Research Board,* 1779, 1, 2001, pp. 46-53.

[18]. Pack, A. I., Pack, A. M., Rodgman, E., Cucchiara, A., Dinges, D. F., Schwab, C. W., Characteristics of crashes attributed to the driver having fallen asleep, *Accident Analysis and Prevention,* 27, 6, 1995, pp. 769-775.

[19]. Lenné, M. G., Triggs, T. J., Redman, J. R., Time of day variations in driving performance, *Accident Analysis and Prevention,* 29, 4, 1997, pp. 431-437.

[20]. Rabiner, L. R., A tutorial on hidden Markov models and selected applications in speech recognition, in *Proceedings of the IEEE,* 77, 2, 1989, pp. 257-286.

[21]. Sathyanarayana, A., Boyraz, P., Hansen, J. H., Driver behavior analysis and route recognition by hidden Markov models, in *Proceedings of the IEEE International Conference on Vehicular Electronics and Safety (ICVES'08),* 2008, pp. 276-281.

[22]. Meng, X., Lee, K. K., Xu, Y., Human driving behavior recognition based on hidden Markov models, in *Proceedings of the IEEE International Conference on Robotics and Biomimetics (ROBIO'06),* 2006, pp. 274-279.