Analysis of drivers’ characteristic driving operations based on combined features

Min Wang, Shuguang Li, Lei Zhu and Jin Yao
School of Manufacturing Science and Engineering, Sichuan University, Chengdu, China

Abstract

Purpose – Analysis of characteristic driving operations can help develop supports for drivers with different driving skills. However, the existing knowledge on analysis of driving skills only focuses on single driving operation and cannot reflect the differences on proficiency of coordination of driving operations. Thus, the purpose of this paper is to analyze driving skills from driving coordinating operations. There are two main contributions: the first involves a method for feature extraction based on AdaBoost, which selects features critical for coordinating operations of experienced drivers and inexperienced drivers, and the second involves a generating method for candidate features, called the combined features method, through which two or more different driving operations at the same location are combined into a candidate combined feature. A series of experiments based on driving simulator and specific course with several different curves were carried out, and the result indicated the feasibility of analyzing driving behavior through AdaBoost and the combined features method.

Design/methodology/approach – AdaBoost was used to extract features and the combined features method was used to combine two or more different driving operations at the same location.

Findings – A series of experiments based on driving simulator and specific course with several different curves were carried out, and the result indicated the feasibility of analyzing driving behavior through AdaBoost and the combined features method.

Originality/value – There are two main contributions: the first involves a method for feature extraction based on AdaBoost, which selects features critical for coordinating operations of experienced drivers and inexperienced drivers, and the second involves a generating method for candidate features, called the combined features method, through which two or more different driving operations at the same location are combined into a candidate combined feature.

Keywords Machine learning, Advanced driver assistant systems, Driver behaviors and assistance

Paper type Research paper

1. Introduction

With an increasing volume of automobiles, a number of traffic problems, including frequent traffic accidents and severe shortage of energy efficiency, are also on the rise (Sagberg et al., 2015). The World Health Organization (2015) report on the status of global road safety stated that road traffic accidents were a major cause of death in the world and the leading cause of death among people of 15-29 years of age, with about 1.25 million people having died in 2013. To reduce traffic accidents and improve energy efficiency, many studies have been conducted with different results. For instance, Kato and Kobayashi (2008) found that fuel consumption could be reduced by 10-30 per cent while driving in eco-mode, which underscored the significance of driving behavior. Bingham et al. (2012) also found that calm drivers tend to have a lower fuel rate than aggressive drivers in similar situations. For the purpose of honing the skills of inexperienced drivers, research studies focused on driving skills by establishing a driver classification model. Wahab et al. (2009) applied the driving style questionnaire (DSQ) method to define individual driving styles and then collected driving data from drivers to train a classifier. Generally, the DSQ method needs a lot of time, efforts and resources to investigate driver behaviors. Aoude et al. (2012) divided the driving data into two driving styles (compliant and violating) and trained a classifier using a combination of the SVM-Bayesian filter (SVM-BF) and the hidden Markov model (HMM). Sundbom et al. (2013) collected the labeled data from drivers who drove normally or aggressively to train a classifier, based on a probabilistic autoregressive eXogenous model. Naivala et al. used feature extraction and classifier modeling to establish a classification model of driver’s driving skill when passing corners. They
adopted principal component analysis (PCA) to extract critical characteristics. And then, the discriminant model of driver's driving skill was established by using SVM, K-nearest neighbor (KNN) and probabilistic neural networks (PNN) (Chandrasiri et al., 2010, 2012, 2016). Ly et al. (2013) also used a support vector machine (SVM) to recognize driving styles based on the labeled information of the vehicle’s inertial sensors. To model and analyze driving styles semantically, Wang et al. (2017) gave a new framework for driving style analysis using primitive driving patterns with Bayesian nonparametric methods, a hierarchical structure (HDP-HSMM) was developed by combining hierarchical Dirichlet process (HDP) and hidden semi-Markov model (HSMM), which could learn a set of expected primitive driving patterns in car-following behaviors. Wang et al. (2017) used a k-means clustering method for drivers' labeling and applied a semi-supervised approach, namely, a semi-supervised support machine (S3VM), to classify various driving styles, the data labeling required a prior is greatly reduced and S3VM improved classification accuracy by about 10 per cent. Li et al. (2013, 2014) studied drivers’ driving skills under a specific curve by using wavelet analysis to extract critical features and established the algorithm of experienced driver's behavior extraction based on AdaBoost. The above three studies were based on curved roads, using indirect features that reflected the potential specifics of practiced drivers and unpracticed drivers as candidate features. The studies analyzed drivers’ lateral driving traits and longitudinal driving characteristics at the same time. Although drivers’ driving skills can be better reflected in lateral and vertical operations under the cornering condition, the method of generating candidate feature results in a driving skill analysis only based on several single features, which cannot reflect driving skill on drivers’ co-occurrence of driving operation; although signals of different frequency components can be found in the same feature, it is still limited to a single feature.

This paper took advantage of candidate combined features reflecting the consistency of driving operations; critical features were extracted using AdaBoost at the same time. Section 1 of this paper introduces the main achievements in terms of drivers’ driving level. Section 2 involves a battery of experiments designed for driving data collection based on driving simulator. Section 3 describes data processing method and data analyzing approaches. Section 4 discusses relevant data analyzing result. Section 5 states the conclusions. The main research process is shown in Figure 1.

2. Experiment

This experiment was carried out with a driving simulator (DS) (Figure 2), which consisted of a visual system with a field of view of 140° around, a sound system and a dynamic model. The driving environment for the experiment (Figure 3) was a city road with six curves with left turn, and these curves, with different radii and lengths, were numbered 1-6 according to the travel direction (Figure 4). The speed limit of 60 km/h at 50 and 100 m before the start of each curve required drivers to maintain a speed of about 60 km/h before entering the curve. The collected data contained the position of accelerator and brake, front wheel angle, vehicle speed, lateral acceleration, longitudinal acceleration and yaw rate, with a sampling frequency of 60 Hz. To obtain sufficient experimental data, a total of 16 drivers of different driving levels participated in the experiment. Each driver completed 12 laps, the first two of which were test drives. Basic information of drivers is shown in Table I.
2.1 Data normalization

All data collected on the basis of time were normalized with a certain distance according to the travel direction utilizing linear interpolation so that the same curve at different laps had comparability. The normalized data with the same data length are shown in Figure 5.

2.2 Driving skill labeled

Murphey et al. (2009) suggested that a smaller jerk or steady driving process would result in less fuel consumption and higher safety. This means that the smaller the jerk, the higher the driving skill. Complex jerk on behalf of a changed rate of acceleration at a distance was used for showing driver’s driving skill in a curve. \( J \), representing the complex jerk, is given in equation (1). In the condition of the same average speed as described above, the bigger the variate \( J \), the lower the driving skill:

\[
J = \sum_{i=1}^{N} \sqrt{J_{\text{lateral}}^{2} - i + J_{\text{longitudinal}}^{2} - i}
\]

where \( J_{\text{lateral}}^{2} \) and \( J_{\text{longitudinal}}^{2} \) stand for lateral and longitudinal accelerations, respectively, at the i-th point in one curve and the variate \( N \) is the total number of standard points in the same curve.

3. Method for data processing

3.1 Generation method for candidate combined features

Candidate combined features were decided by driving data, including steering wheel angle, accelerator petal position, brake petal position and corresponding operation and vehicle speeds. We chose the average in a distance of 9 m, which contained 30 standard points as candidate features, to decrease the error caused by operating occasionality, and the averages were extracted every other point:

\[
y(P) = \begin{cases} 
2 & \text{if } P_{i+1} - P_i > \Delta \\
1 & \text{if } |P_{i+1} - P_i| \leq \Delta \\
0 & \text{if } P_{i+1} - P_i < -\Delta 
\end{cases}
\]

where variable \( y \) represents the change of single feature \( P \).

This paper referred to the feature co-occurrence for face detection (Mita T et al., 2005), which combined two or more different features into one feature, called the combined feature. The following gave the combined principle of two features at the same point: for a single feature \( P_i \), the current feature \( P_{i+1} \) was compared with the previous adjacent feature \( P_i \), and a threshold value \( \Delta \) was set empirically for each kind of the feature \( P \). Then, ternary numbers 2, 1 and 0 were used to indicate that the difference of \( P_{i+1} \) and \( P_i \) was greater than \( \Delta \), equal to \( \Delta \) and less than \( \Delta \), respectively. The variable \( y \) for a sample \( P \) is figured in equation (2).

With the above processing for two features at a certain point of a certain curve, we could obtain a two-dimensional \( N \times 2 \) array, according to the rule of converting a ternary number into a decimal number [equation (3)], the ternary array of \( n \times 2 \) was converted to the decimal array of \( n \times 1 \) and the decimal array only contained the elements of 0-8, representing the nine kinds of candidate combined features, as shown in Table II:

\[
G = 3\Delta_1 + \Delta_2
\]

where \( \Delta_1 \) and \( \Delta_2 \) are both ternary numbers.

3.2 Method for feature extraction

The feature extraction processing using AdaBoost is shown in Feature extraction processing using AdaBoost:

1. Given example of labeled data \( (x_1, y_1), (x_2, y_2), \ldots, (x_m, y_m) \), where \( x_i \in X, y_i \in \{-1, +1\} \)
2. Initialize weight \( w_0 = \frac{1}{2}, y_i = 0, 1 \), where 0 and 1 are on behalf of experienced driver and
Min Wang, Shuguang Li, Lei Zhu and Jin Yao

Characteristic driving operations

Table II Combined feature method

<table>
<thead>
<tr>
<th>Combined feature</th>
<th>G (Decimal number)</th>
<th>Δ1</th>
<th>Δ2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Feature difference of a single feature P (Ternary number)</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>6</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>7</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>8</td>
<td>2</td>
<td>2</td>
</tr>
</tbody>
</table>

Notes: *2 means \( P_{i+1} - P_i > \Delta \), feature \( P \) increased; 1 means \( |P_{i+1} - P_i| \leq \Delta \), feature \( P \) unchanged; 0 means \( P_{i+1} - P_i < -\Delta \), feature \( P \) decreased.

4.1 Feature extraction

The relationship between number of weak classifiers and error rate of strong classifier in Figure 6 was a critical step for deciding the number of weak classifiers using AdaBoost. The number of weak classifiers was the number of features extracted. We found that the error rate of strong classifiers was less than three per cent as the number of weak classifiers was increased. This paper stipulated that when the accuracy of classifiers satisfied 97 per cent, the process for feature extraction was completed. Figure 7 shows the concrete locations of a part of the 15 features extracted. A major difference between skilled and unskilled drivers was obvious at the entrance. The details of those features are provided in Table III. For example, the first combined feature consisting of velocity and steering angle appeared at the site that was 62.7 m away from the origin of 1-th curve, and the feature value of this combination was bigger than 1.5 as to inexperienced drivers.

4.2 Features distribution characteristics

Curves were divided into five parts, including 50 m before curve, 50 m after curve and trisection of the remaining curve in Figure 8. They were named sections AB, BC, CD, DE and EF along the travel direction. Figure 9 shows all features’ distribution on the five sections of curved proposed above. Most features occurred at the entrance and exit, which were in line with actual driving as drivers got used to adjusting driving operations at those parts. In contrast, there were a few operations in the middle of the curves, seen in section CD. Combined features of “steering wheel operation speed and accelerator operation speed” and “accelerator pedal position and steering wheel operation speed” were the most frequently extracted, which meant that the difference between the two groups of drivers was mainly in these two combined features.

In section AB, it was found that the combined feature of steering wheel angle and accelerator operation speed was more frequently extracted. In fact, drivers changed the steering wheel angle and velocity constantly at the entrance to adapt to the
changing course. The cause for the difference of the two groups in speed had first appeared via accelerator operation instead of velocity itself; thus, the adjustment of steering wheel angle and accelerator operation frequency most exhibited drivers’ driving skills in section AB. With that, the operations on accelerator directly resulted in the imparity of velocity in the next section BC, and experienced and less experienced drivers differed greatly in section BC during the combined feature of velocity and steering operation speed. Accelerator and steering operation speeds became the predominant driving behavior in sections CD and DE. Drivers had accomplished the flexural road in section EF, skilled drivers mainly sped up and almost did not adjust the steering wheel angle, which reflected to driving operations were steering operation speed almost kept unchanged and accelerator operation speed changed, and this was the combined feature of steering operation speed and accelerator operation speed.

5. Conclusion

This paper proposed a method for driving operations characteristics analysis, using AdaBoost and feature co-occurrence. When the driving operations went through the curves at a special course, they were studied based on DS. In the end, all features corresponding to relevant curves were selected and extracted using the proposed method. The result illustrated that most features came out at the entrance and exit of all curves, which conformed to actual behavior when drivers entered or left curves.

We just studied driving feature extraction, which was a part of fundamental research in the field of driving operations characteristics. In the future, we plan to enrich the driving environment and not keep it restricted to courses consisting of curves alone. We are also keen to develop a driving assistant system that will help improve inexperienced drivers’ driving skills through driving behavior analysis, so as to decrease traffic accidents.

References


Corresponding author
Shuguang Li can be contacted at: lsg042@163.com

For instructions on how to order reprints of this article, please visit our website: www.emeraldgrouppublishing.com/licensing/reprints.htm
Or contact us for further details: permissions@emeraldinsight.com