Vol.2010-ITS-41 No.1
2010/6/18

Developing of Objective Similarity Measures for Real-World Driving Behaviors

Pongtep Angkititrakul†, Ryuta Terashima†, and Toshihiro Wakita†

Recent advances in ITS allow us to collect large amount of real-world multi-modal driving data for research study and analysis. In order to effectively utilize database, ability to automatically mine driving situations of interest is one of the essential steps. In this paper, we propose a probabilistic technique to measure similarity of driving behaviors based on posterior probability of driving modes in a driving space. The similarity distance is then obtained from correlation coefficient between two feature matrices. In addition, the framework allows adaptation scheme of driver model to better fit individual driving characteristics. Experimental results on car-following situations have showed that the proposed framework achieved 40% accuracy, while human achieved 49%, in the similarity-ranking task

1. Introduction

On account of advances in vehicle-based ITS technology, instrumented vehicle has become possible for extensive data acquisition of non-intrusive measurement of driver behavior in realistic driving environment. For instance, driving behavior (e.g., steering maneuver, brake/gas pedal operation), vehicle status (e.g., velocity), and driving environment (e.g., following distance) can be observed synchronously with the video signals capturing forward driving scenes and driver's facial region. Such realistic, multi-modal driving data is very valuable for studying driver behavior and developing advanced driver assistance technologies. Therefore, ability to manage and organize a large amount of data is essential to utilize the driving database. For example, Figure 1 shows an image of a driving-data browser developed at Nagoya University [4]. This system allows us to browse and retrieve driving data/scene of interest (i.e., query), and can be accessed via networks from PCs or smart phones. The developing system can be utilized in driver-risk assessment as a feedback of driving performance to reduce risky driving behaviors. Similarly, in driving education or driving monitoring system, one would like to analyze decision and behavior of a driver under the similar situations or compare driving behaviors before and after receiving the training, and then assess the driving performance. Moreover, in developing advanced driving recording system, one would like to predict a hazardous situation before hand in order to appropriately activate the recording process. If a particular situation or behavior commonly leads to a hazardous situation, the system should be alerted and prepared when encounters the similar situation.

Objectively measuring similarity of driving behaviors is one of the key components to achieve these abovementioned goals. However, to the best of our knowledge, at present there is no unified protocol or standard definition to measure similarity between driving behaviors. Besides, there are several factors contributing to driving behaviors ranging from driver itself to vehicle status and driving environment. Measures of driving similarity can be subjective. One can assess similarity between two driving events using different criteria. For example, in car-following situations, one could decide similarity criteria by how driver takes action corresponding to the driving environment such as the distance to a lead vehicle (e.g., tailgating, keeping distance), smooth following behavior, frequent decelerating and accelerating, and sudden and hard braking. Intuitively, one might judge driving similarity by safety or risk levels—two driving behaviors are similar if they both are either safe or unsafe.

In this paper, we took an initiative step to develop an objective measure of similarity between two driving behaviors. In particular, we focused on the car-following situation which characterizes longitudinal behavior of a driver when he or she follows another vehicle in front. We proposed a probabilistic framework to measure similarity of driving behaviors based on posteriori probability of driving modes in a driving space. The hypothetical driving mode is a set of particular manners of driving operations under a given environmental condition. Such latent patterns of driving modes can be captured and modeled by observing driving behaviors from large amount of driving data. In this study, we employed Gaussian mixture model (GMM) to represent driving modes due to its stochastic and unsupervised-training properties. Given a set of driving-mode models, a feature matrix representing a driving event is the projection of its observed driving data in a driving space—obtained by computing posteriori probability of observations generated by each driving-mode model. Consequently, a similarity distance, as well as a similarity score, between two driving events can be determined by correlation between two representative feature matrices. Furthermore, the proposed framework is capable of model adaptation (e.g., Maximum A Posterior or Bayesian adaptation) to better fit individual driving characteristics, and being adjusted its parameters continuously to suit the expanded database as new data are obtained.

A series of experimental evaluations was conducted to validate similarity scores generated by the proposed algorithm in three different aspects: (1) Can similarity score differentiate driving similarity and dissimilarity? (2) Correlation of similarity scores and degrees of similarity judged by human annotator, and (3) Validation of similarity between

†豊田中央研究所
TOYOTA Central R&D Laboratories, Inc.

query and search results obtained by the algorithms. All validation processes were performed with subjective assessment by human annotator. In addition, for comparison, two other methods based on raw observations and histogram counts were introduced to generate similarity scores.

This paper is organized as follows. In the next section, we discuss our basic framework of similarity measure, together with two baseline methods. In Section 3, we introduce our proposed method based on driving modes and adaptation scheme. The real-world driving corpus is described in Section 4. Section 5 discusses the experimental evaluation, followed by conclusion in Section 6.

Figure 1 Driving-data browser developed at Nagoya University.

2. Driving Similarity

To measure a similarity between two driving events, a feature matrix is first extracted from each sequence of observed driving signals. The similarity measure is then based on the distance between two representative feature matrices, as shown in Figure 2. Here, correlation coefficient between two feature matrices is employed as an inverse relation of similarity distance—the higher correlation between two feature matrices, the closer distance between them, and therefore the more similar two driving events are. The correlation coefficient between matrices A and B can be computed as,

$$r = \frac{\sum_m \sum_n (A_{mn} - \bar{A})(B_{mn} - \bar{B})}{\sqrt{(\sum_m \sum_n (A_{mn} - \bar{A})^2)(\sum_m \sum_n (B_{mn} - \bar{B})^2)}},$$

where $\bar{A} = mean2(A)$ and $\bar{B} = mean2(B)$ are the means or averages of matrix elements. The correlation coefficient has a value between -1 and +1, which respectively correspond to negative and positive perfect relationship. In the paper, we also refer to this correlation as a similarity score. As it approaches zero there is less of a similarity, and closer the similarity score is to +1, the stronger the similarity between two driving events.

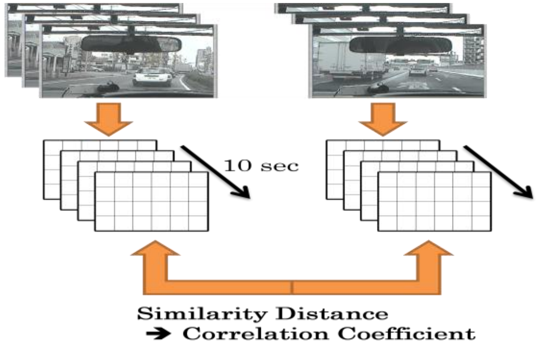

Figure 2 Similarity measure of two driving events.

The important part of this scheme is that a feature matrix representing driving event should encode and preserve meaningful characteristics of driving behavior. Before we proceed to describe our proposed algorithm, it is worthwhile to discuss some other conventional methods as for the baseline comparison.

2.1 Raw Driving Signals

A time series of multi-dimensional driving signals is employed directly as a feature matrix. The utilized driving signals are following distance, vehicle velocity, gas/brake pedal pressure, and their delta parameters (i.e., the first derivatives). The feature matrix is then a two-dimensional matrix with each row contains a vector of one driving signal.

2.2 Multi-dimensional Histogram Count

The histogram method counts the number of values of multi-dimensional driving signals that fall between the elements in the edge vectors (i.e., bins) obtained from driving signals of interest. In this work, two-dimensional histogram count between following distance and pedal operation[a] is computed every time sample, as shown in Figure 3. Consequently, the feature

[a] The pedal operation signal is obtained by combining both gas-pedal and brake-pedal signals into one signal by subtracting brake-pedal signal from gas-pedal signal. That is, gas-pedal signal represents positive value, and brake-pedal signal represents negative value.

matrix is obtained by adding all counts from the driving event, and has a size equals to numbers of bins used by both signals (i.e., *ct* x *bt* in Figure 2).

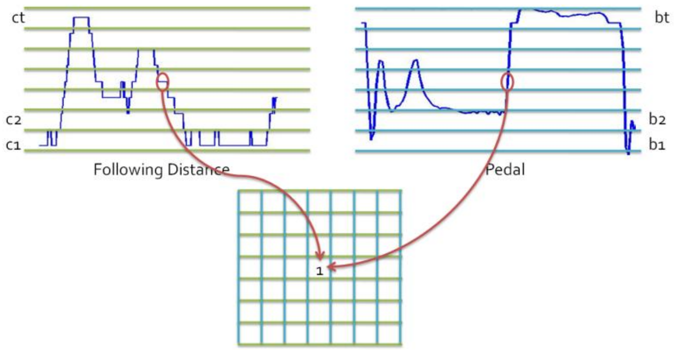

Figure 3 Illustration of 2-D histogram count between following distance and pedal operation

3. Similarity Measure based on Driving Mode Representative

The motivation of our proposed algorithm is that, in car following, drivers normally perform a few different driving modes given a range of following distance, i.e., how they accelerate and decelerate vehicle velocity at a particular distance between vehicles. For instance, when a following vehicle approaches a lead vehicle with a very close distance, it is more likely that the following vehicle will be in a driving mode with hard-braking operation. On the other hand, when a distance between vehicles is far enough, the following vehicle tends to be in the driving modes with accelerating or maintaining velocity more than the ones with decelerating. Such latent driving modes might not be easy to clearly defined or formulated, but the relationship between driving observations representing driving modes can be captured and modeled with stochastic methods using a wide range of training data. Figure 4 illustrates hypothetical driving modes in a driving space of pedal operation and following distance. In this figure, it is assumed that there exist four driving modes within each bin of following distance (e.g., two gas-pedal operating modes and two brake-pedal operating modes can possibly be performed given a range of following distance.) Each driving mode is represented by a Gaussian distribution. Therefore, given a driving space with assigned driving modes, two driving events are considered 'similar' if they traverse through the same driving modes—their observations touch the same joint distributions between driving operation and driving environment. Based on this assumption, a feature matrix (i.e., driving mode representative) can be obtained by projecting driving observations onto a driving space using posterior probability of driving observations generated by driving-mode models, and the similarity distance between two driving events is measured in this space. Next, we will describe how to create driving-mode models.

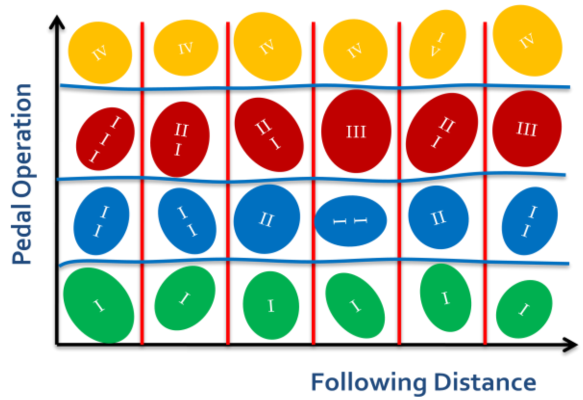

Figure 4 Illustration of four driving modes in each following-distance bin.

Let a driving event be represented by a time series *{X(t)}, t=1,…,T.* Here, *X(t)* is a multi-dimensional driving-signal observations, i.e., $X(t) = [p(t)\ d(t)]^T$, where *p(t)* represents pedal operation, and *d(t)* represents following distance between host vehicle and lead vehicle. Driving mode representative of a driving event can be obtained as follows. First, the range of possible following-distance values is partitioned into *N* bins, i.e., *{0, edge(1), …, edge(n), …, edge(N)}*. The boundaries between bins [*edge(n)-edge(n-1)*] can be linear scale or log-like scale, subsequently *X(t)* belongs to the *i-th* bin if *edge(i-1) <= d(t) < edge(i)*, as

$$X_i(t) = \{X(t): edge(i-1) \le d(t) \le edge(i), \forall t\}.$$

For each following-distance bin, *M* driving modes can be obtained by training an *M*-mixture GMM[b] [5] from a pool of *X(t)* belonging to that bin. That is, each Gaussian

[b] Each mixture θ_m of Gaussian Mixture Model (GMM) is defined by a mixture weight ($weight_m$)、a mean vector (μ_m), and a covariance matrix (Σ_m). The weighted Gaussian probability density function (pdf) $p(x|\theta_m)$ is defined as $p(x|\theta_m) = \frac{weight_m}{\sqrt{(2\pi)^D|\Sigma_m|}} \cdot \exp\left\{-\frac{1}{2}(x-\mu_m)^T\Sigma_m^{-1}(x-\mu_m)\right\}$

component of GMM represents one driving mode within that following-distance bin.

$$Mode_{i,j} = Gaussian(i,j), \qquad 1 \le i \le N, 1 \le j \le M.$$

where $\{Gaussian(i,j), weigth_{i,j}\}$, $\forall 1 \le j \le M$, constitute the GMM_i. Thus, there are total $N \times M$ driving modes characterizing the whole driving space. The driving-mode models are trained from the pool of a large number of driving observations of several drivers from the database.

Given a driving space, at time *t*, the driving observations of each driving event *{X(t), t=1,…,T}* can be projected onto the driving space by computing posterior probabilities of the driving signals given driving-mode models, as

$$F_{i,j}(t) = p(X(t)|Mode_{i,j}), t = 1, \ldots T$$

By projecting all samples in this manner, we obtain a series of posterior probabilities in the driving space. Since consecutive observations are strongly correlated, their projections in driving space are close to each other. Furthermore, the posterior probability will be high if the observations of a driving trajectory are close to that driving mode, and the probability will decrease as neighboring driving modes are further from the observed trajectory. Finally, the two-dimensional mode representative of a driving event is calculated by

$$MR_{i,j} = weight_{i,j} \cdot \sum_{t=1}^{T} F_{i,j}(t), t = 1, \ldots, T$$

The similarity score between two driving events $E1$ and $E2$ is then defined by the correlation coefficient of two normalized mode representatives $MR^{(E1)}$ and $MR^{(E2)}$ as

$$S(E1, E2) = corrcoef\left(norm(MR^{(E1)}), norm(MR^{(E2)})\right),$$

where $norm(MR)$ is the normalization of *MR* by its total summation of all elements, as

$$norm(MR) = \frac{MR_{i,j}}{\sum_i \sum_j MR_{i,j}}.$$

3.1 Model Adaptation

It is widely accepted that driving behaviors vary among individual drivers, depending on driver's experience, background, age, etc. Even for the same driver, driving behavior may be altered from situation to situation. In addition, each driver may perform diverse driving operation under the same driving environment or may have the preferred driving modes, as a result of individual driving style. Therefore, the joint distributions between driving operation and driving environment, or the driving-mode models, should be adapted to fit individual driving characteristics. The advanced algorithms capable of adapting their model parameters are desired to develop a more reliable system.

In this section, we discuss model adaption of driving modes. The aim of adaptation scheme is to adjust the relationship between pedal operation and following distance by relocating the position of driving modes (i.e., shifting the means of distributions), as shown in Figure 5. The driver-independent model, as described in the previous section, captures the average driving characteristics, while the adapted model, namely driver-dependent model, captures individual driving characteristics. The adaptation scheme can be applied as follows.

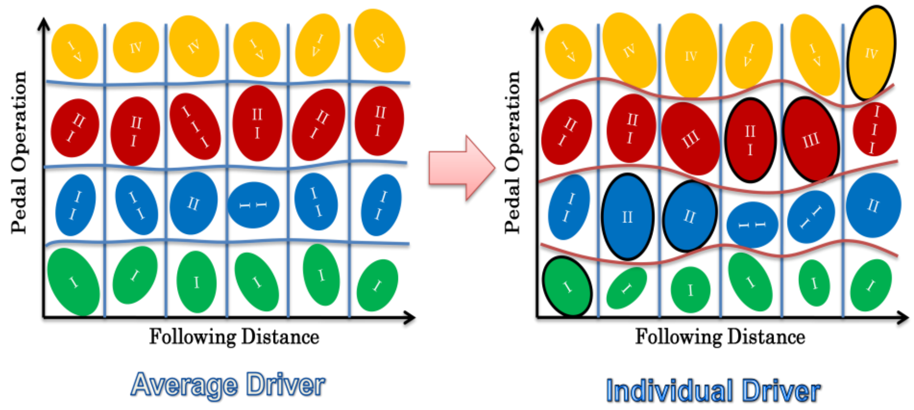

Figure 5 Illustration of driving-mode adaptation

Let $Mode_{i,j}^{(I)}$ represent a set of the driver-independent, driving-mode models trained from a large number of drivers as discussed previously, $\{X^{(A)}(\tau)\}, \tau = 1, \ldots, \mathrm{T_A}$ represent a set of driving observations belonging to driver A, and $\{X^{(B)}(\tau)\}, \tau = 1, \ldots, \mathrm{T_B}$ represent a set of driving observations belonging to driver B. The driver-dependent mode models representing driver A and driver B can be obtained by adapting the driver-independent models $Mode_{i,j}^{(I)}$ with the observations corresponding to each driver. That is,

$$Mode_{i,j}^{(A)} = Adapt\left(Mode_{i,j}^{(I)}, X^{(A)}\right),$$

$$Mode_{i,j}^{(B)} = Adapt\left(Mode_{i,j}^{(I)}, X^{(B)}\right).$$

Similar to the training process of the driver-independent modes, only $X^{(A)}(\tau)$ belonging to the *i-th* bin will be used to adapt the *M* driving modes representing the *i-th* bin. In this work, we employed Maximum-A-Posteriori (MAP) or Bayesian adaptation framework to adapt driving-mode models. MAP adaptation is an effective, well-defined scheme, widely used with GMM. More details about this framework can be found in [6]. Note that the adaption is simultaneously performed on all driving modes belonging to the same following-distance bin, as

$$GMM_i^{(A)} = MAP(GMM_i^{(I)}, \{X_i^{(A)}(t)\})$$

Given the driver-dependent modes of driver A and B, and the corresponding driving events $\{X^{(A)}(t)\}$ and $\{X^{(B)}(t)\}$, $t = 1, \ldots,$ T, the driver-dependent mode features can be obtained by

$$F_{i,j}^{(A)}(t) = p\left(X^{(A)}(t) \mid Mode_{i,j}^{(A)}\right), t = 1, \ldots T$$

$$F_{i,j}^{(B)}(t) = p\left(X^{(B)}(t) \mid Mode_{i,j}^{(B)}\right), t = 1, \ldots T$$

Finally, the two-dimensional mode representatives of two driving events are represented by

$$MR_{(i,j)}^{(A)} = weight_{i,j}^{(A)} \cdot \sum_{t=1}^{T} F_{i,j}^{(A)}(t), t = 1, \ldots, T$$

$$MR_{(i,j)}^{(B)} = weight_{i,j}^{(B)} \cdot \sum_{t=1}^{T} F_{i,j}^{(B)}(t), t = 1, \ldots, T$$

In a similar manner, the similarity score between two driving events $MR^{(A)}$ and $MR^{(B)}$ is the correlation coefficient of two mode representatives as

$$S(A,B) = corrcoef\ \left(norm\left(MR^{(A)}\right), norm\left(MR^{(B)}\right)\right).$$

Constrained Adaptation

Furthermore, to avoid the adapted driving modes to shift across the reasonable pedal-operation boundary (e.g., from gas-pedal pressure to brake-pedal pressure, and vice versa) due to the sparse data of some drivers, a simple constraint can be applied to avoid this issue. Hence, we modified the adaptation scheme to apply gas-pedal data to adapt the acceleration modes, and brake-pedal data to adapt the deceleration modes; otherwise, remain the original models when there is no proper adapt data, as follows

$$Gaussian_{i,j}^{(A)} = \begin{cases} MAP(Gaussian_{i,j}^{(A)}, \left\{X_i^{(A)}(t): p_i^{(A)}(t) \geq 0\right\}, & Mode_{i,j}^{(A)}: pedal\ mean\ \geq 0 \\ MAP(Gaussian_{i,j}^{(A)}, \{X_i^{(A)}(t): p_i^{(A)}(t) < 0\}, & Mode_{i,j}^{(A)}: pedal\ mean < 0 \end{cases}$$

4. Real-World Driving Corpus

All the observable driving signals are collected synchronously employing an instrumented vehicle, a TOYOTA Hybrid Estima, developed by Takeda Lab, Graduate school of Information Science, Nagoya University, Japan. The vehicle is equipped with a wide range of sensors and data recording systems. The rich multi-modal data contains twelve channel speech, three-channel video, driving behavior including gas and brake pedal pressures, steering angle, and vehicle velocity, physiological signals including driver's heart rate, skin conductance, and emotion-based sweating on the palms and soles, etc. In particular, the steering angle is obtained by a potentiometer. The brake and accelerator pedal pressure are obtained by the pressure sensors. The vehicle velocity is measured from the output of the JIS5601 pulse generator. The following distance from a lead vehicle is acquired by two types of distance sensors mounted in front of the vehicle in order to locate the lead vehicle in both short and long ranges. More details about the vehicle setup and related work can be found in [1] [2] [3] [7]. The utilized driving data in the study are from 77 drivers, with balance in both genders. Each driver drove the instrumented vehicle around the Nagoya area under a variety of driving environments and traffic conditions.

5. Experimental Evaluation

To validate the proposed objective similarity measures, we conducted a series of experiments and compared the performances between algorithms and human annotator. An annotator watched all the forward-scene videos obtained by the front-view camera, and subjectively assessed the similarity of driving behaviors. In addition, the annotator clarified the criteria used to decide the similarity. There is no single rule the annotator used to judge the similarity; it depends on situation case by case. However, the common criteria used by the annotator in judging similarity of driver behaviors in car following are smooth following, keeping comfortable distance, following-too-closely, tailgating, sudden braking, safe/unsafe following, etc. Three evaluation schemes were used to demonstrate the performance of similarity score in measuring driving similarity, as shown in Figure 6.

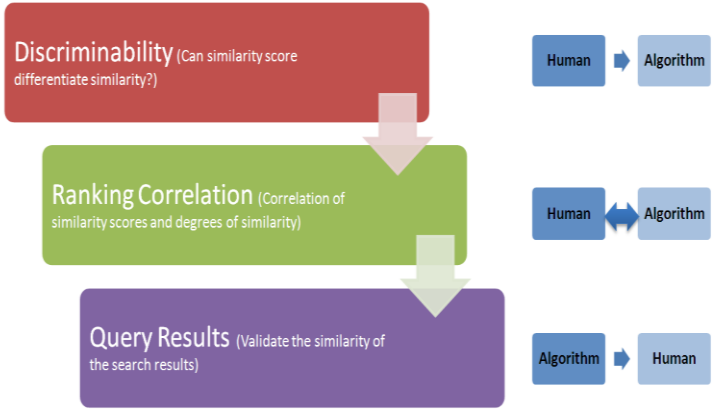

Figure 6 Evaluation cycle of similarity measurement

1) Discrimability of similarity level: Can estimated similarity score differentiate similarity and dissimilarity of driving behaviors?
2) Ranking Correlation: What is correlation between similarity score and degree of similarity judged by annotator?
3) Top-10 Similarity: How well can algorithms select similar behaviors from the database, given a reference query?

To simplify the problem, in this study all the driving events were equally partitioned into a fixed length 10 seconds, and all driving signals were re-sampled to 10 Hz. The necessary pre-processing was performed to reduce noises from data acquisition. For comparison, the similarity scores are computed from different algorithms as described previously:

1) RCORR (raw driving signals): Each feature matrix has dimension 7x100 (i.e., seven features and 100 samples)
2) HISTC (histogram count): Each feature matrix has dimension 4x6 (i.e., four bin for pedal operation and six bins for following distance)
3) MODE (driver-independent driving mode): Each feature matrix has dimension 4x6 (i.e., four driving modes and six following-distance bins)
4) MAP (driver-dependent modes based on MAP adaptation): Each feature matrix has dimension 4x6
5) ADAPT (driver-dependent modes based on constrained adaptation): Each feature matrix has dimension 4x6

Note that the numbers of parameters (i.e., bins, modes) were obtained empirically to achieve the optimal performance of each algorithm.

5.1 Discriminability of Similarity Level

In this first experiment, we want to verify that similarity score estimated by algorithm is meaningful. That is, how well can the estimated similarity score discriminate similar driving behaviors from dissimilar driving behaviors?

Experimental Setup

Several pairs of driving events were randomly selected from the corpus. We first selected a reference driving event, and then selected a set of other driving behaviors to pair up with the reference one. Human annotator assessed the pair-wise similarity and assigned similarity levels of each pair—compared with the reference behavior. Here, the similarity level was classified into 4 scales where level 1 represents 'not similar' and level 4 represents 'very similar'. This subjective scale is judged relatively across all selected pairs.

Next, those pairs with assigned similarity level 1 and level 4 (total 229 pairs) were used in a binary classification task, where level-1 pairs belong to one class and level-4 pairs belong to another class. The estimated similarity scores generated by different algorithms were used to discriminate these two classes.

Experimental Results

Figure 7 illustrates the Detection Error Tradeoff (DET)[c] curves of all algorithms, and Table 1 shows the Equal Error Rate (EER)[d] of all algorithms. All the algorithms performs better

---

[c] DET curve is one type of the Receiver Operating Characteristic (ROC) curves. It is a plot between two types of detection errors. The closer the plot to the lower left corner (i.e., zero), the better detection performance
[d] EER is the point where both false alarm and false rejection errors are equal. Again, lower EER generally implies better classification performance.

than chance (i.e., EER = 50%), and the proposed driving modes outperforms the other two baseline algorithms with almost 9% absolute EER. The MAP adaptation can further reduce EER of driver-independent models and shows overall better performance. This experiment demonstrates effectiveness of similarity score in discriminating similar/dissimilar behaviors.

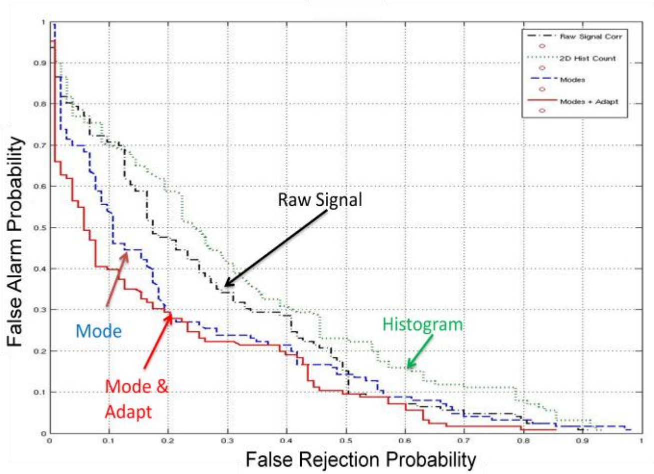

Figure 7　DET curves of similarity classification (Level 1 vs Level 4)

Table 1 Equal Error Rate (EER) of different methods

Algorithm	Equal Error Rate (%)
Baseline I (RCORR)	31.89%
Baseline II (HISTC)	34.94%
MODE	25.80%
MAP	24.43%

5.2 Ranking Correlation

Having shown that the estimated similarity score is meaningful, we now want to investigate the correlation of similarity score in ranking similarity with that of human assessment. We believe that the higher correlation between algorithm and human annotator implies higher performance of the algorithm.

Experimental Setup

For each driving event as a reference driving event, we randomly selected the other three candidates of driving events from the corpus excluding the reference one. Human annotator subjectively ranked the similarity of all three candidates compared with the reference driving events, where 1st is the most similar, 2nd is the second, and 3rd is the least similar driving events. Since subjective evaluation is subject to inconsistency errors, the annotator was asked to perform assessment twice without knowing that the list was the same but being shuffled the order. Table 2 shows the consistency performance of human annotator in ranking similarity. From the total number of 599 reference behaviors, human annotator was able to rank order perfectly matched from both rounds only 48.75%. However, human annotator was able to consistently rank the candidate with the most similar behavior to the reference behavior up to 67.11%. On the average, the correlation coefficient of two ranking assessments was approximately 0.45.

Table 2 Ranking correlation performance of human annotator

Human consistency test	Performance
Avg. Correlation Coefficient (ρ)	0.4454
Perfectly Matched (%)	48.75
First-rank Matched (%)	67.11

Experimental Results

Next, we picked up those reference behaviors and corresponding three candidates which human annotator consistently ranked similarity in both rounds (i.e., 292 cases). For each reference event, the pair-wise similarity scores between the reference event and each of candidate events were estimated. Consequently, the similarity ranking was based on the estimated similarity scores with the highest similarity score represents the most similar behavior. Table 3 shows the ranking performance of similarity scores generated by each algorithm and compared with ranking by human annotator. The proposed algorithm with constrained adaptation demonstrates the best performance at 39.73% in perfect matching with human ranking, compared with the histogram-based algorithm which is only 28.42%. In overall, the proposed algorithm outperforms the other two baseline algorithms and shows higher correlation with the human assessment. Again, slight but consistent improvement can further obtained by adaptation scheme.

Table 3 Ranking correlation performance of algorithms compared with human annotator

	Baseline 1 (RCORR)	Baseline 2 (HISTC)	MODE	MAP	ADAPT
Avg. Corr. Coeff. (ρ)	0.2312	0.2072	0.2945	0.3271	0.3271
Perfectly Matched (%)	33.22	28.42	35.62	36.64	39.73
First-rank Matched (%)	54.79	55.82	55.82	58.22	56.85

5.3 Top-10 Similarity

In the last experiment, we want to validate the effectiveness of the algorithms on selecting a collection of similar driving behaviors from a database to a reference one.

Experimental Setup

For each driving event as a reference or query, the algorithms computed its pair-wise similarity scores with all the other driving events in the corpus (i.e., 779 events). All the similarity scores were ranked from the highest to the lowest. Only the reference behavior and its associated Top-10 driving behaviors with highest similarity scores were selected for human validation.

Experimental Results

Human annotator assessed similarity between the reference behavior and its corresponding Top-10 similar events selected by different algorithms. Table 4 illustrates average percentage of dissimilar events among Top-10 events judged by human annotator. As we can see, although the driver-independent modes shows the lowest performance (i.e., on the average, 2.73 selected events are not similar to the query), with the adaptation scheme the proposed algorithm shows the lowest errors.

Table 4 Human judge on top-10 similarity

Comparison to Human Annotation	Baseline I (RCORR)	Baseline II (HISTC)	MODE	ADAPT
'Not-Similar' Percentage	21.1	25.8	27.3	20.5

6. Conclusions

In this paper, we investigated objective similarity measure of driving events based on posterior probability of driving modes. We have showed the advantages of the proposed algorithm over the other two conventional methods (i.e., raw signals, and histogram count), and have demonstrated its promising performance toward human annotation. Future work will consider variable durations of driving events and comparison with human annotators who are expert drivers (e.g., driving instructors). We believe that ability to measure similarity of driving behaviors will pave the way to better understanding and modeling driver behavior.

Acknowledgement

This work was support by Strategic Information and Communications R&D Promotion Programme (SCOPE) of Ministry of Internal Affairs and Communications (MIC) Japan under grant No. 082006002.

The authors would like to thank Profs. Kazuya Takeda, Chiyomi Miyajima, and Tatsuya Suzuki at Nagoya University for their advice and supports.

References

(1) P. Angkititrakul, R. Terashima, T. Wakita, K. Takeda, C. Miyajima, and T. Suzuki, "Evaluation of Driver-Behavior Models in Real-World Car-Following Task," IEEE International Conference on Vehicular Electronics and Safety (ICVES), India (2009)

(2) L.Malta, C. Miyajima, N. Kitaoka, and K. Takeda, "Multimodal real-world driving data collection and analysis," 4th Biennial Workshop on DSP for In-Vehicle Systems and Safety, Dallas, Texas (2009).

(3) C. Miyajima, T. Kusakawa, T. Nishino, N. Kitaoka, K. Itou, and K. Takeda, "On-going data collection for driving behavior signal," 2007 Biennial on DSP for In-Vehicle and Mobile Systems (DSPINCARS 2007), Istanbul, Turkey, (2007).

(4) M. Naito, C. Miyajima, T. Nishino, N. Kitaoka, and K. Takeda, "A Browsing and Retrieval System for Driving Data," IEEE-Intelligence Transportation, San Diego (2010)

(5) D.A. Reynolds and R.C. Rose, "Robust text-independent speaker identification using Gaussian-mixture speaker models," IEEE transaction on Speech and Audio Processing, (1995)

(6) D.A. Reynolds, T.F. Quatieri, and R.B. Dunn, "Speaker Verification using Adapted Gaussian Mixture Models," Digital Signal Processing, Vol. 10(1), pp. 19-41 (2000)

(7) A. Ozaki, S. Hara, T. Kusakawa, C. Miyajima, T. Nishino, N. Kitaoka, K. Itou, and K. Takeda, "In-car speech data collection along with carious multimodal signals," Language Resources and Evaluation Conference (LREC2008), pp. 22-19 (2008)