Behavior Signal Processing for Vehicle Applications

Chiyomi Miyajima, Pongtep Angkititrakul, Kazuya Takeda
Nagoya University, Nagoya, Japan
E-mail: {miyajima, pongtep, takeda}@g.sp.m.is.nagoya-u.ac.jp Tel: +81-52-789-4432

***Abstract*—Within the past decade, analyzing and modeling human behavior by processing large amounts of collected data has become an active research topic in advanced human-machine interaction systems. The research community strives to find improved ways to explain and represent meaningful behavioral characteristics of humans in order to develop efficient and effective cooperative interactions between humans, machines, and the environment. This paper provides a summary of progress achieved to date of our research on behavior signal processing, with a focus on the driver-vehicle-environment interaction. First, we describe the method of data collection used to develop our real-world driving corpus, which contains multimodal driving signals capturing relevant information regarding driver, vehicle, and environment. Then, the paper provides an overview of our signal processing and data-driven approaches used to analyze and model driver behavior for a wide range of practical vehicle applications. We then perform experimental validation using the realistic driving behavior of several drivers. In particular, the vehicle applications include driver identification, behavior prediction (i.e., car following and lane change), driver frustration (emotion) detection, and driver education. We hope that this paper will provide some insight to researchers who have interest in this field, and help identify areas and applications where further research is needed.**

I. Introduction

Human behavior plays an important role in any system involving human-machine interaction. In regards to driving, when analyzing driver-vehicle-environment interactions (Fig. 1), human errors contribute to more than 90% of fatal traffic accidents. Understanding human/driver behavior can be useful in preventing traffic collisions [21], as well as enhancing effectiveness of the interaction between driver, vehicle, and environment. The study of driver behavior is a very challenging task due to its stochastic nature with high degree of inter- and intra-driver variability. To cope with these issues, over the past decades data-centric approaches have gained much attention in the research community [8], [14], [15], [18]. In this research, we focused on the understanding of human behavior from a signal processing perspective, and on developing a methodology to analyze and model the extracted meaningful behavioral information.

To analyze and model driver behavior, the first step is to collect a reasonable amount of realistic multi-modal observations. Here, observations or driving signals represent behavioral variables as a time series which possesses particular dimensions of behavioral characteristics. We took extra care in designing and developing our instrumented vehicle to collect a broad range of driving signals which could represent relevant information regarding driver, vehicle, and environment. This paper describes our data collection and driving corpus, which is one of the largest realistic driving corpora with more than 550 participants at the time this article was written. Subsequently, we discuss the behavior processing and modeling methods with practical vehicle applications.

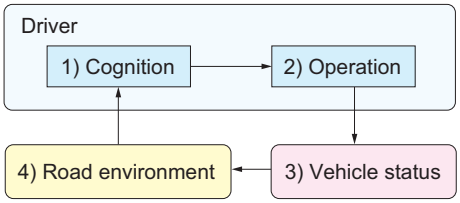

Fig. 1. Cyclic Interaction between driver, vehicle, and driving environment.

Driver behavior conveys multiple layers of information regarding a driver such as a driver's mental, physical, and cognitive states, driver intention, and driver identity. Therefore, modeling driver behavior can be used to detect or predict behaviors of interest during vehicle operation, or can be used to investigate and assess driver behavior after vehicle operation. In particular, we showed that driver models obtained from driving signals could be used to recognize driver identity, predict driver operation, detect driver frustration, and assess recorded driving behavior.

We have developed a driver-behavior model based on a probabilistic Gaussian mixture model (GMM) framework. The GMM model was applied to capture relationships among the related parameters of car-following behavior. We showed that the GMM models representing the patterns of pedal operation in the spectral domain could achieve an accuracy rate of 89.6% in recognizing the identities of 276 drivers. Furthermore, the GMM-based behavior modeling framework was extended to predict driver behavior in terms of pedal operation given observed driving signals such as following distance and vehicle velocity. The modeling framework is also capable of model adaptation which allows the adapted model to better represent particular driving characteristics such as individual driving style. The experimental results showed that the framework could achieve a prediction performance of 19 dB (signal-to-deviation ratio).

In addition, employing a Bayesian network (BN), driver frustration could be detected at a true positive rate of 80% with a 9% false positive rate. Finally, utilizing the automatically detected hazardous situations, we developed a web-based system for drivers to navigate and review each hazardous

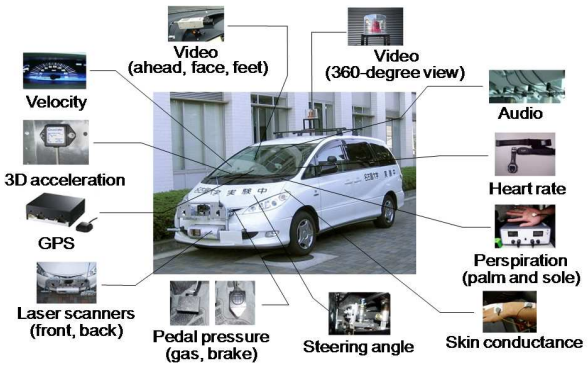

Fig. 2. Instrumented vehicle.

situation of their own recorded driving activity. The system provides feedback on each risky driving behavior and suggests how the users can appropriately respond to such situations in a safe manner. The experimental evaluation showed that safe-driving behavior improved significantly after using the proposed system. In conclusion, in this paper we describe signal processing approaches for collecting, analyzing, modeling, and assessing human behavior, and demonstrate the advantages for vehicle applications.

This paper is organized as follows. We first introduce our driving corpus and the data collection in the next section. In Sec. III, we describe the first application of driver modeling in evaluating driver identification. Then, in Sec. IV, driver behavior prediction is discussed for both car-following and lane-change tasks. Sec. V describes our analysis of real-world driver's frustration using combination of speech signal and pedal actuation signals. The last application regarding driver education is discussed in Sec. VI. Finally, we summarize and discuss the future work in Sec. VII.

II. Driving Corpus

A data collection vehicle was designed for synchronously recording audio with other multimedia data [10]. Various sensors were mounted on a Toyota Hybrid Estima with 2,360 cc displacement and automatic transmission, as shown in Fig. 2. Table II summaries all driving data recorded by the system. Participants drove the instrumented vehicle on city streets and expressways in the city of Nagoya, Japan. During the experiment, drivers performed secondary tasks carefully designed to provide activities that were most likely to occur during everyday driving. Data collection vehicle, route, equipments, and treatment conditions were the same for all drivers. Drivers performed secondary tasks in the same order at very similar locations, so data from different drivers can be readily compared. An experimenter monitored the experiments from the rear seat.

A. Collection Protocol

To develop a technique for quantifying the stress level of drivers, driving data are recorded under various conditions with four different tasks. The details of the tasks are described as follows with examples of spoken sentences.

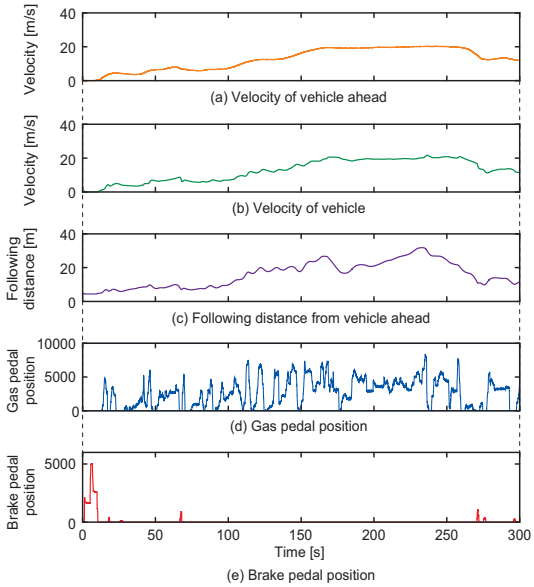

Fig. 3. Examples of driving behavior signals.

- Signboard reading task Drivers read aloud words on signboards such as names of shops and restaurants seen from the driver seat while driving, e.g., 7-11 and Dennys.
- Navigation dialogue task Drivers are guided to an unfamiliar place by a navigator on a cell phone with a hands-free headset. Drivers do not have maps, and only the navigator knows the route to the destination.
- Alphanumeric reading task Drivers repeat random four-letter strings consisting of alphabet a-z and digits 0-9, e.g., UKZC, IHD3, and BJB8. The instruction of the four-letter strings is heard by earphone.
- Music retrieval task Drivers retrieve and play music by a spoken dialogue interface. Music can be retrieved by artist name or song title, e.g., Beatles or Yesterday.

Each driver starts from Nagoya University and returns after about 70 minutes of driving. Driving data are recorded under the above four task conditions on city roads and two conditions on an expressway. Driving data without any tasks are recorded as references before, between, and after the tasks. Fig. 3 illustrates some samples of pre-processed driving signals.

After completing the route, the participant was also asked to assess his/her subjective level of frustration by referring to the front-view and facial videos as well as the corresponding audio. A user interface for such assessment was designed so that drivers used an intensity scale and slid a bar from zero to thirty, i.e., from none to extremely frustrated. The interface output was a continuous signal: the level of frustration was recorded every 0.1 seconds.

B. Data Annotation

An effective annotation of multimedia information is crucial for providing a more meaningful description of the situations drivers experience. In this study, we proposed a data annotation protocol that covers most of the factors that might affect

TABLE I
SUMMARIZATION OF DRIVING DATA ACQUISITION

Recording System	Specification	Channel	Rate	Unit	Data Captured
Microphones	Omni-directional condenser microphones (SONY ECM-77B) and a closed-talk mic.	12	16 kHz	Voltage	Audio and speech
Video Cameras	CCD cameras (SONY DXC-200A)	4	29.4118 fps	images (pixels)	Driver's face (left and right) and feet, front-view road
Pedal Sensors	Pressure sensors (LPR-A-03KNS1 and LPR-R-05KNS1)	2	16 kHz	N	Gas and brake pedal pressure
Steering sensor	Potentiometer (COPAL M-22E10-050-50K)	1	16 kHz	degree	Steering angle
Speed Sensor	Pulse generator JIS5601	1	16 kHz	m/s	Vehicle velocity
Distance Sensors	SICK DMT-51111 and MITSUBISHI MR3685	2	16 kHz	m	Distance to a lead vehicle
Physiological Sensors	Chest belt POLAR S810i	1	100 Hz	bpm	Heart rate
	Perspiration meter (SKINOS SKD-2000)	1	16 kHz	mV	Driver's sweat
	Electrodermal meter (SKIN SKSPA)	1	16 kHz	mV	Driver's sweat
Accelerometers	Crossbow CXL04LP3	3	16 kHz	G	Acceleration in x-y-z
Laser Scanners	Front: RIEGL LMS-140i-80	1	20 Hz	m	Objects in front
	Back: RIEGL LMS-Q120i	1	50 Hz	m	Objects behind
Omni-directional Camera	Point Grey Ladybug	1	6x15 fps	images	Driving environment
GPS	Navicom GPS-M1zz	1	1 Hz	standard	GPS information

drivers and the drivers responses. The annotation labels are comprised of six major groups: drivers affective state (level of irritation), driver actions (e.g. facial expression), drivers secondary task, driving environment (e.g. type of road, traffic density), vehicle status (e.g. turning, stopped), and speech / background noise. The annotation protocol designed in this research is comprehensive, and can be used in a wide range of research elds. We are currently annotating data from all drivers in our database [6].

III. DRIVER IDENTIFICATION

Pedal operation patterns also differed among drivers. Fig. 4 shows examples of gas pedal operation signals of 150 s collected in the driving simulator for two drivers. They were all recorded with the same moving pattern of a lead vehicle. Pedal operation patterns are similar in the same driver but different between the two drivers.

We modeled the differences in the gas and brake pedal operation patterns with GMMs using the following two kinds of features [9].

A. Modeling Based on Raw Pedal Operation Signals

Pedal operation patterns can be visualized with histograms of raw pedal operation signals. They significantly differ from each other. Driver 1 has a symmetric distribution around 4,000, whereas driver 2 has a peak around 2,000 and a wider spread to the right than to the left. Driver 2 also has a sharp peak near zero because he usually puts his foot on the gas pedal evenwhen not accelerating. These distributions of raw pedal operation signals were modeled with Gaussian mixture model (GMMs).

GMM is a statistical model widely used in pattern recognition including speech and speaker recognition [17]. It is defined as a mixture of multivariate Gaussian, and the probability of D-dimensional observation vector $\boldsymbol{o}$ for GMM λ is obtained as follows:

$$b(\boldsymbol{o} \mid \lambda) = \sum_{i=1}^{M} w_i \mathcal{N}_i(\boldsymbol{o}), \tag{1}$$

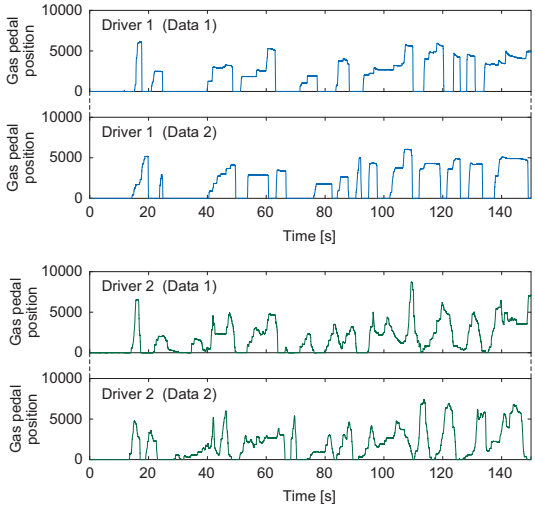

Fig. 4. Examples of gas pedal operation patterns for two drivers (Top: Driver1, Bottom:Driver2).

where M is the number of the Gaussians of the GMM and $\mathcal{N}_i(\boldsymbol{o})$ is the D-variate Gaussian distribution of the i-th component defined with mean vector $\boldsymbol{\mu}_i$ and covariance matrix $\boldsymbol{\Sigma}_i$:

$$\mathcal{N}_i(\boldsymbol{o}) = \frac{1}{\sqrt{(2\pi)^D \, |\boldsymbol{\Sigma}_i|}} \exp\left\{-\frac{1}{2}(\boldsymbol{o}-\boldsymbol{\mu}_i)'\boldsymbol{\Sigma}_i^{-1}(\boldsymbol{o}-\boldsymbol{\mu}_i)\right\}, \tag{2}$$

where $(\cdot)'$ and $(\cdot)^{-1}$ denote transpose and inverse matrices, respectively. w_i is a mixture weight for the i-th component and satisfies $\sum_{i=1}^{M} w_i = 1$.

B. Modeling Based on Spectral Features of Pedal Operation Signals

Cepstrum is the widely used spectral feature for speech and speaker recognition [16], defined as the inverse Fourier transform of the log power spectrum of the signal. As shown in Fig. 5, cepstral analysis allows us to smooth the structure

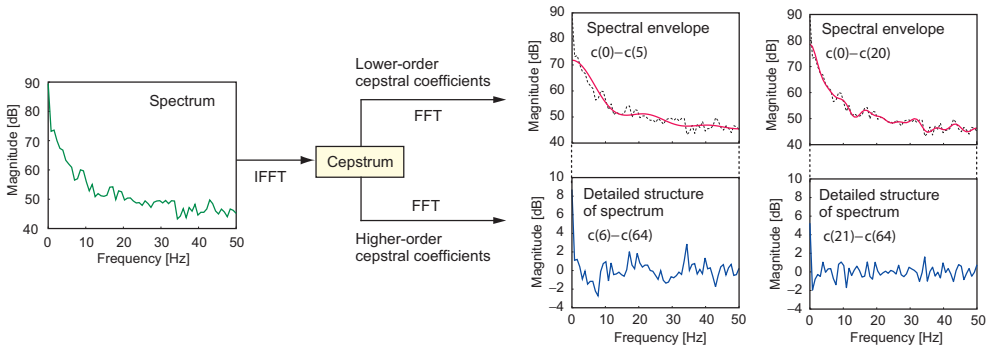

Fig. 5. Examples of spectral envelopes extract through spectral analysis using lower order cepstral coefficients c(0)-c(5) or c(0)-c(20)

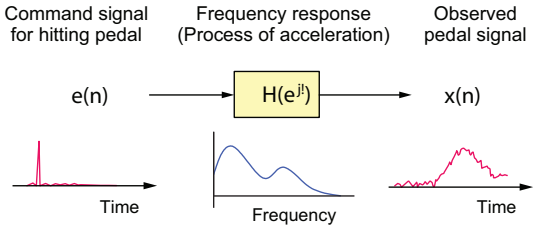

Fig. 6. General modeling of driving signal.

of the spectrum by keeping only the first several lower-order cepstral coefficients and setting the remaining coefficients to zero. Assuming that individual differences in pedal operation patterns can be represented by the smoothed spectral envelope of pedal operation signals, we modeled the pedal operation patterns of each driver with lower-order cepstral coefficients.

As shown in Fig. 6, in driver modeling, we assume that command signal for hitting a pedal $e(n)$ is filtered with driver model $H(e^{j\omega})$ represented as the spectral envelope, and the output of the system is observed as pedal signal $x(n)$, e.g., in gas pedal operation, a command signal is generated when a driver decides to hit the gas pedal, and $H(e^{j\omega})$ represents the process of acceleration. This can be described in frequency domain as follows:

$$X(e^{j\omega}) \quad = \quad E(e^{j\omega})H(e^{j\omega}), \tag{3}$$

where $X(e^{j\omega})$ and $E(e^{j\omega})$ are the Fourier transforms of $x(n)$ and $e(n)$, respectively. We focus on driver characteristics represented as frequency response $H(e^{j\omega})$.

Assuming that the spectral envelope can capture the differences between the characteristics of different drivers, we focused on the differences in spectral envelopes represented by cepstral coefficients (cepstrum), which were also modeled with GMMs.

1) Dynamic Features of Driving Signals: Similar to speech and speaker recognition, we found that the dynamic features of the driving signals offer much information on driving behaviors. Dynamic features are defined as the following linear regression coefficients:

$$\Delta \boldsymbol{o}(t) = \frac{\sum_{k=-K}^{K} k\boldsymbol{o}(t+k)}{\sum_{k=-K}^{K} k^2}, \tag{4}$$

where $\boldsymbol{o}(t)$ is a static feature of raw signals or cepstral coefficients at time t and K is the half window size for calculating the Δ coefficients. We determined the regression window to be $2K = 800$ ms from preliminary experiments for both raw pedal signals and cepstral coefficients. If $\boldsymbol{o}(t)$ is a D-dimensional vector, D dynamic coefficients are obtained from the static coefficients, combined into a $2D$ dimensional vector, and modeled with GMMs.

C. Driver Identification Experiments

1) Experimental Conditions: The driving data of 276 drivers collected on city roads in the data collection vehicle were used, excluding data collected while not moving. Driving signals of three minutes were used for GMM training and another three minutes for testing. We used both brake and gas pedal signals in the real vehicle experiments because drivers use the brake pedal more often during city than expressway driving.

Cepstral coefficients obtained from the gas and brake pedal signals are modeled with two separated GMMs, and their log-likelihood scores were linearly combined. In driver identification, the unknown driver was identified as driver $\hat{k}$ who gave the maximum weighted GMM log-likelihood over the gas and brake pedals:

$$\begin{aligned}\hat{k} = \arg\max_{k} \{&\gamma \log P(\boldsymbol{G} \mid \lambda_{G,k}) \\ &+(1-\gamma)\log P(\boldsymbol{B} \mid \lambda_{B,k})\}, \quad 0 \le \gamma \le 1,\end{aligned} \tag{5}$$

where $\boldsymbol{G}$ and $\boldsymbol{B}$ are the cepstral sequences of the gas and brake pedals and $\lambda_{G,k}$ and $\lambda_{B,k}$ are the k-th driver models of the gas and brake pedals, respectively. γ $(0 \le \gamma \le 1)$ is a linear combination weight for the log-likelihood of gas pedal signals.

2) Experimental Results: The results for the 16-component GMMs are summarized in Fig. 7. The identification performance was rather low when using raw driving signals: the best identification rate for raw signals was 47.5% with $\alpha = 0.80$.

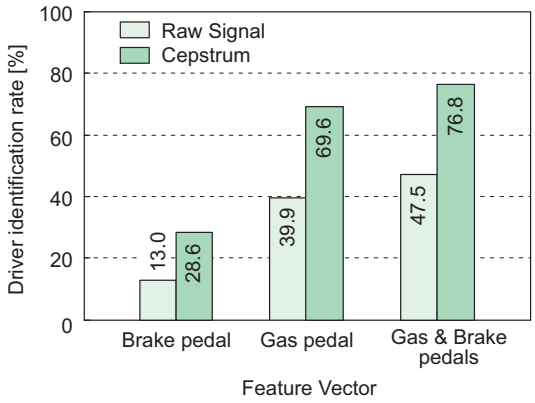

Fig. 7. Comparison of identification rates for raw pedal signals and cepstral coefficients

By applying cepstral analysis, however, the identification rate increased to 76.8% with $\alpha = 0.76$. Although the pedal sensors were different from those of the driving simulator, similar results were obtained. We thus conclude that cepstral features captured the individualities of driving behavior better than raw driving signals and achieved better performance in driver identification.

IV. Driver Behavior Prediction

Driver behavior models can be employed to predict driver operation given the available observations at the prediction time. Here, two driving tasks are considered: car following [12], [1] and lane change [13].

A. Car Following

Car following characterizes longitudinal behavior of a driver while following another vehicle in front [2]. In this study, we focus on car following in the sense of the way driver behavior of the following vehicle is affected by the driving environment (i.e., leading vehicle) and the own vehicle status. There are several contributory factors in car-following behavior such as relative position and velocity of following vehicle with respect to lead vehicle, acceleration and deceleration of both vehicles, and perception and reaction time of the followers. Fig. 8 shows a basic diagram of car following and corresponding parameters, where $v_t^f, a_t^f, f_t, \mathbf{x}_t^f$ represent vehicle velocity, acceleration/deceleration, distance between vehicles, and observed feature vector at time t, respectively.

The GMM-based driver-behavior modeling representing the patterns of pedal operation corresponding to the observed vehicle velocity and following distance. The underlying motivation of this modeling framework is that as a driver determines gas and brake pedal operation in response to the stimulus of vehicle velocity and following distance, accordingly, such patterns can be modeled by the joint distribution of all correlated parameters, as shown in Fig. 9.

1) Feature Extraction and Model Representation: To model the pedal pattern, an observed feature vector at time t, $\mathbf{x}_t$, consists of vehicle velocity, following distance, and pedal

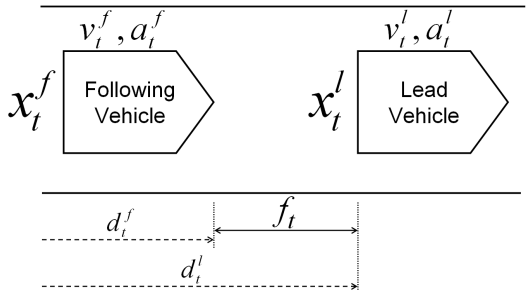

Fig. 8. Car following with corresponding parameters.

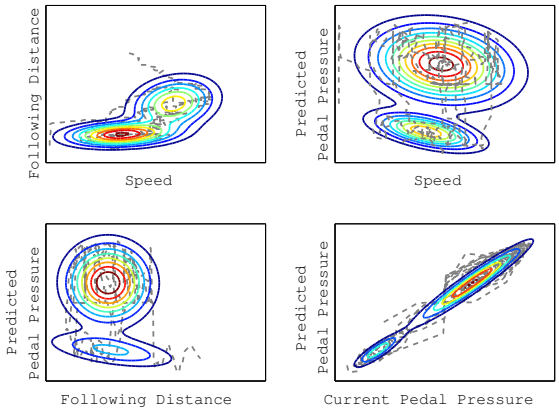

Fig. 9. Car-following trajectory in parameter space (gray dashed line) overlaid with 2-mixture GMM distribution.

pattern (G_t) with their first- (Δ) and second-order (Δ^2) derivatives as

$$\mathbf{x}_t = [v_t^f, \Delta v_t^f, \Delta^2 v_t^f, f_t, \Delta f_t, \Delta^2 f_t, G_t, \Delta G_t, \Delta^2 G_t]^T, \quad (6)$$

where the $\Delta(\cdot)$ operator of a parameter is defined as

$$\Delta x_t = x_t - \frac{\sum_{\tau=1}^{\mathbb{T}} \tau x_{t-\tau}}{\sum_{\tau=1}^{\mathbb{T}} \tau}, \quad (7)$$

where $\mathbb{T}$ is a window length (e.g., 0.8 s). Next, let us define a set of augmented feature vectors $\mathbf{y}_t$ as

$$\mathbf{y}_t = [\mathbf{x}_t^T \; G_{t+1}]^T. \quad (8)$$

Consequently, the joint density between the observed driving signals $\mathbf{x}_t$ and the next pedal operation G_{t+1} can be modeled by a GMM Φ, with a mean vector μ_k^y and a covariance matrix Σ_k^{yy} of the $k-th$ mixture expressed as

$$\mu_k^y = \begin{bmatrix} \mu_k^{\mathbf{x}} \\ \mu_k^G \end{bmatrix} \text{ and } \Sigma_k^{yy} = \begin{bmatrix} \Sigma_k^{\mathbf{xx}} & \Sigma_k^{\mathbf{x}G} \\ \Sigma_k^{G\mathbf{x}} & \Sigma_k^{GG}. \end{bmatrix} \quad (9)$$

2) Pedal Prediction: In this work, the predicted gas pedal pattern $\hat{G}_{t+1}$ is computed by the weighted predictions resulting from all mixture components of GMM, as

$$\hat{G}_{t+1} = \sum_{k=1}^{K} h_k(\mathbf{x}_t) \cdot \hat{G}_{t+1}^{(k)}(\mathbf{x}_t) \quad (10)$$

where $\hat{G}_{t+1}^{(k)}(\mathbf{x}_t)$ is a maximum a posteriori (MAP) prediction of the observed parameters $\mathbf{x}_t$ given the k-th mixture component which is given by

$$\begin{aligned}\hat{G}_{t+1}^{(k)}(\mathbf{x}_t) &= arg\max_{G_{t+1}}\{p(G_{t+1}|\mathbf{x}_t,\phi_k)\} \\ &= \mu_k^G + \Sigma_k^{Gx}(\Sigma_k^{xx})^{-1}(\mathbf{x}_t - \mu_k^x) \end{aligned} \quad (11)$$

The term $h_k(\mathbf{x}_t)$ is the posterior probability of the observed parameter $\mathbf{x}_t$ belonging to the k-th mixture component, given by

$$h_k(\mathbf{x}_t) = \frac{\alpha_k p(\mathbf{x}_t|\phi_k^x)}{\sum_{i=1}^K \alpha_i p(\mathbf{x}_t|\phi_i^x)}, 1 \le k \le K \quad (12)$$

where $p(\mathbf{x}_t|\phi_i^x)$ is the marginal probability of the observed parameter $\mathbf{x}_t$ generated by the i-th Gaussian component, and α_k is the prior probability of the k-th mixture component.

3) Model Adaptation: We applied Bayesian or Maximum-A-Posterior (MAP) adaptation to reestimate the model parameters individually by shifting the original statistic (i.e., mean vectors) toward the new adaptation data [17]. The *universal* driver-behavior models were first obtained from a pool of driving data of several drivers from the training set. The universal driver models represent average or common driving characteristics shared by several drivers. In this study, to enhance the model capability, we took a further step to adapt the parameters of the universal driver models in the following two scenarios.

- Driver Adaptation: The aspect of driver adaptation is to adapt the model parameters to better represent individual driving characteristics. In this scenario, the driving data belonging to each particular driver are used to adapt the universal model to obtain the adapted driver models, namely *driver-dependent* or *personalized* driver models. That is, each driver will be associated with an individualized and unique driver model.
- On-line Adaptation: The driving data at the beginning of each car-following event were used to adapt the universal model, and subsequently, the on-line adapted driver model was used to represent driving behavior for the rest of that car-following event. The objective of the on-line adaptation is to capture the overall unique car-following characteristics of such event (e.g., driver and environment) that deviates from the average characteristics of the universal models.

4) Experimental Evaluation: The evaluation is performed using approximately 300 minutes worth of clean and realistic car-following data from 68 drivers. Manual annotation is exploited to verify that only concrete car-following events with legitimate driving signals that last more than 10 seconds are considered. Fig. 10 compares the prediction performance of the universal, driver-adapted, and on-line-adapted (using 30 sec of driving data) driver models with 4, 8, 16, 32 Gaussian components in terms of Signal-to Deviation Ratio (SDR).

The Signal-to-Deviation Ratio (SDR) is defined as follows

$$SDR = 10\log_{10}\frac{\sum_{t=1}^T G^2(t)}{\sum_{t=1}^T (G(t)-\hat{G}(t))^2}[dB], \quad (13)$$

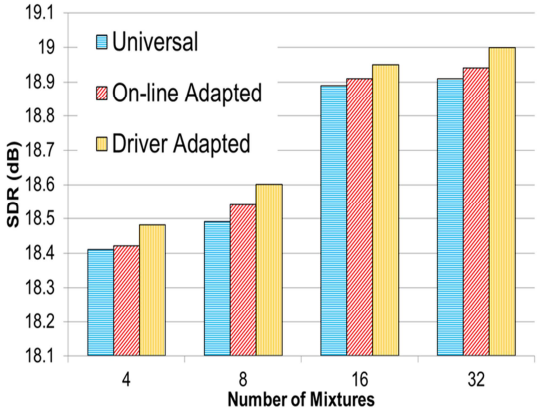

Fig. 10. Comparison of prediction performance

where T is the length of signal, $G(t)$ is the actually observed signal, and $\hat{G}(t)$ is the predicted signal. The driver-adapted models showed the best performance.

B. Lane Change

Since lane change activity consists of multiple states (i.e., examining the safety of traffic environments, assessing the positions of other vehicles, moving into the next lane, and adjusting driving speed to traffic flow) [3], a single dynamic system cannot model vehicle trajectory. In addition, the boundaries between states cannot be observed from its trajectory.

To study lane-change behavior, a set of vehicle movement observations was measured using a driving simulator. Relative longitudinal and lateral distances from the vehicle's position when starting the lane change, $x_i[n], y_i[n]$, and the velocity of the vehicles, $\dot{x}_i[n], \dot{y}_i[n]$, were recorded every 160 ms. Here $i = 1, 2, 3$ is an index for the location of surrounding vehicles (Fig. 11), and $(x_0[n], y_0[n])$ represents the position of the drivers own vehicle. The duration of lane-change activity, $n = 1, 2, ..., N$, starts when $V0$ (drivers own vehicle) and $V2$ are at the same longitudinal position and ends when $V0$'s lateral position reaches the local minimum as shown in Fig. 11.

1) Modeling Trajectory using HMM: We used a three-state HMM to describe the three different stages of a lane change: preparation, shifting, and adjusting. In the proposed model, each state is characterized by a joint distribution of eight variables:

$$v = [\dot{x}_0, y_0, \Delta\dot{x}_0, \Delta\dot{y}_0, \Delta^2\dot{x}_0, \Delta^2\dot{y}_0, \dot{x}_1, \dot{x}_2]^T. \quad (14)$$

In general, longitudinal distance, x_0, monotonically increases in time and cannot be modeled by an i.i.d. process. Therefore, we use longitudinal speed $\dot{x}_0$, as a variable to characterize the trajectory. Finally, after training the HMM using a set of recorded trajectories, the mean vector μ_j and covariance matrix Σ_j of the trajectory variable v are estimated for each state $j = 1, 2, 3$. The distribution of duration N is modeled using a Gaussian distribution.

The shape of a trajectory is controlled by the HMM and the duration of the lane change activity. When the driver performs

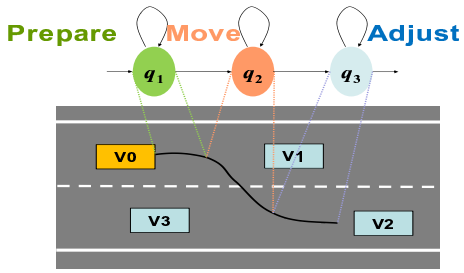

Fig. 11. Lane change trajectory and geometric positions of surrounding vehicles.

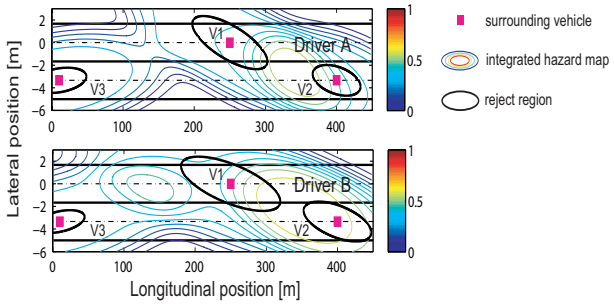

Fig. 12. Hazard maps for two drivers when the same positions of surrounding vehicles were given.

a lane change in a shorter time, this results in a sharper trajectory. We generate a set of probable lane-change trajectories by determining state durations d_j using uniform resampling. Once a set of state durations is determined, we applied the maximum likelihood HMM signal synthesis algorithm (ML method) [20] or the sampling algorithm to generate the most probable trajectory. Simply repeating this process will produce a set of probable vehicle trajectories which characterize a trained drivers typical lane-change behavior

2) Trajectory Selection: Although various natural driving trajectories may exist, the number of lane-change trajectories that can be realized under given traffic circumstances is limited. Furthermore, the selection criteria of the trajectory, based on the traffic context, differs among drivers, e.g., some drivers are more sensitive to the position of the front vehicle than that of the side vehicle. Therefore, we model the selection criterion of each driver with a scoring function for lane-change trajectories based on vehicular contexts, i.e., relative distances to the surrounding vehicles.

In the proposed method, a hazard map function M is defined in a stochastic domain based on the histograms of the relative positions of the surrounding vehicles $r_i = [x_i - x_0, y_i - y_0]^t$. To model sensitivity to surrounding vehicles, we calculated covariance matrix R_i for each of three distances, $r_i, i = 1, 2, 3$, using training data. Since the distance varies more widely at less sensitive distances, we use the quadratic form of inverse covariance matrices R_i as a metric of the cognitive distance. Then we calculate hazard map function M_i for surrounding vehicle V_i as follows:

$$M_i = \frac{1}{1 + exp\{\alpha_i(r_i^t R_i^{-1} r_i - \beta_i)\}} \tag{15}$$

where α_i is a parameter of the minimum safe distance defined so that the minimum value of cognitive distance $r_i^t R_i^{-1} r_i$ of the training data corresponds to the lower 5% distribution values, and β_i is the mean value of $r_i^t R_i^{-1} r_i$.

Hazard map M_i can be regarded as an a posteriori probability of being in the safe driving condition under range distances $Pr(safe|r_i)$, when the likelihood is given as a exponential quadratic form. Therefore, integrating the hazard maps for all surrounding vehicles can be done simply by interpolating three probabilities with weights λ_i into an integrated map M. Once the positions of the surrounding vehicles at time n, $r_i[n]$, are determined, M_i can be calculated for each point in time, and by averaging the value over the lane-change duration we can compare the possible trajectories. Then the optimal trajectory that has the lowest value is selected from the possible trajectories.

3) Experimental Evaluation: Thirty lane-change trials were recorded for two drivers using a driving simulator which simulated a two-lane urban expressway where the traffic was moderately dense. The drivers were instructed to pass the lead vehicle once during each trial, when they were able to. The trained hazard maps M for the two drivers shown in Fig. 12 depicts differences in sensitivity to surrounding vehicles.

We generated possible lane-change vehicle trajectories over a 20-second period using two methods as mentioned, and then selected the optimal trajectory. For quantitative evaluation, we calculated the difference between the predicted and actual trajectories based on dynamic time warping (DTW), using the normalized square difference as a local distance, and measured it in terms of SDR. Average SDRs of the best trajectory hypothesis (best) and all trajectory hypotheses (mean) using the maximum likelihood method (left) and the sampling method (right) are shown in Fig. 13. The sampling method was better at generating vehicle trajectories similar to actual driver trajectories than the ML method. Fig. 13 also shows the SDRs when driver A's model was used for predicting driver B's trajectory and vice versa. The SDR decreased by 2.2 dB when the other driver's model was used to make the prediction. This result confirmed the effectiveness of the proposed model for capturing individual characteristics of lane-change behavior. We also tested our method using the actual lane-change duration, i.e., $I = J$. When the actual lane-change duration N is given, the root mean square error (RMSE) between the predicted and actual trajectories can be calculated. The average RMSE for 60 tests was 17.6 m, which was a good result for predicting vehicle trajectories over a distance of about 600 m (i.e., for a 20-second time period).

V. Driver Frustration

In this section, we propose a method for estimating a driver's frustration that integrates features of a different nature. The designed model is based on the assumption that emotions are the result of an interaction with the environment and are usually accompanied by physiological changes, facial

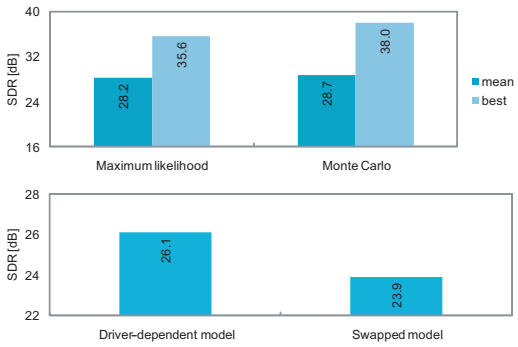

Fig. 13. Average SDRs of lane trajectory. Top: the best and mean trajectories using ML method (left) versus sampling method (right). Bottom: using a driver's own model (left) versus using the other driver's models (right).

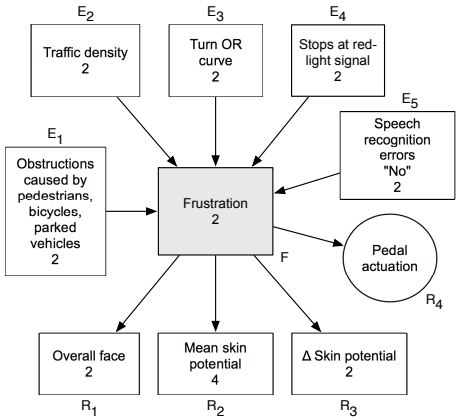

Fig. 14. Proposed BN structure. Squares represent discrete (tabular) nodes, and the circle represents a continuous (Gaussian) mode. The number inside each node represents the number of mutually exclusive states that the node can assume. Labels outside nodes identify random variable.

expressions, or actions [7].

A. Analysis

A method for combining all of the different features and annotation results in an efficient language was needed, and a Bayesian network (BN) [11] was the natural choice to deal with such a task. One of the important characteristics of a BN is the capability to infer the state of an unobserved variable, given the state of observed ones. In our case, we wanted to infer a participant's frustration given the driving environment, speech recognition errors (communication environment), and the participant's responses measured through his/her physiological state, overall face, and pedal actuation.

The graph structure proposed to integrate all of the available information is shown in Fig. 14. This model was based on the following assumptions: (1) environmental factors that may have an impact on goal-directed behavior (traffic density, stops at red-light signals, obstructions, turn or curve, and speech recognition errors) may also have a direct effect on frustration; (2) a frustrated driver is likely to present changes in his/her facial expression, physiological state, and gas- and brake-pedal actuation. In Fig. 14, squares represent discrete (tabular) nodes and the circle represents a continuous (Gaussian) node. The number inside each node represents the number of mutually exclusive states that the node can assume. Random variables were identified by a label outside each node: "F" for frustration, "E" for environment, and "R" for responses.

In addition to the graph structure, it is necessary to specify the parameters of the model, obtained here using a training set. During parameterization, we calculate the Conditional Probability Distribution (CPD) at each node. If the variables are discrete, this can be represented as a table (CPT), which lists the probability that the child node takes on each of its different values for each combination of values of its parents. On the other hand, if the variable is continuous, the CPD is assumed as a Gaussian distribution. For example, the continuous node *Pedal actuation*, which has only one binary parent, was represented by two different multivariate Gaussians, one for each emotional state: frustrated and not frustrated. For each observed environment (driving and communication) and the corresponding driver responses, we can use Bayes' rule to compute the posterior probability of frustration, as:

$$P(F|E_1,E_2,E_3,E_4,E_5,R_1,R_2,R_3,R_4) = P(F|E_1,E_2,E_3,E_4,E_5) \cdot \frac{P(R_1|F)\cdot P(R_2|F)\cdot P(R_3|F)\cdot P(R_4|F)}{P(E_1,E_2,E_3,E_4,E_5,R_1,R_2,R_3,R_4)} . \quad (16)$$

The denominator was calculated by summing (marginalizing) out F. In addition, in this study we set a uniform Dirichlet prior to every discrete node in the network. This was done in order to avoid over-fitted results due to the Maximum Likelihood approach used for calculating the conditional probability tables. Without a prior, patterns that were not observed in the training set would be assigned zero probability, compromising the estimation.

The network input data are all of the available data—pedal actuation, skin potential and other binary signals. At a given time step t, frames of sizes L and M were used to extract features from the skin potential and pedal actuation signals, respectively. Results served as network inputs. The value of each binary label at the current time step was directly entered in the network without further processing. Frame shift was kept fixed at 0.5 seconds. For two consecutive frames, the value of, for example, current traffic density has an effect on future skin potential and pedal actuation signals in order to account for delayed physiological and behavioral reactions. In addition, frustration was estimated continuously, i.e., we did not pre-select segments where we were certain about frustration or neutrality and then ignore ambiguous regions.

B. Experimental Evaluation

Within the data used in experiments, 129 scenes of frustration (segments with original scale above 0) were found. On average, participants got frustrated 6.5 times while driving. The mean strength of frustration scenes was 10.5, and the mean duration was 11.8 seconds.

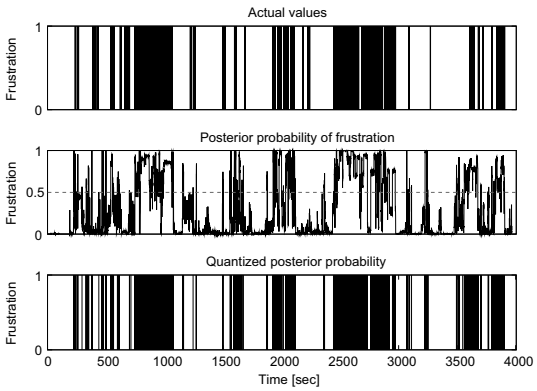

Fig. 15. Results for individual drivers (arranged side by side) calculated using the Full network. Comparison between (top) actual frustration from all drivers, (center) posterior probability of the frustration node, and (bottom) its quantized version using a threshold of 0.5.

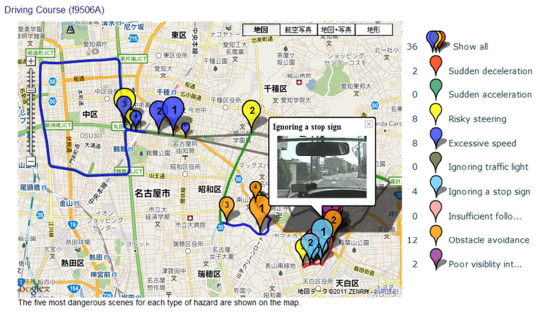

Fig. 16. Interface summarizes hazardous situations on a driving map.

Fig. 15 shows estimation results for individual drivers arranged side by side: actual frustration from all 20 participants (top); the posterior probability of the frustration node calculated using the *Full* network (center); and quantized posterior probability using a threshold of 0.5 (bottom). The quantized probability of each driver was further median-filtered to remove spikes. The overall result was achieved with a true positive (TP) rate of 80% and a false positive (FP) rate of 9% (i.e., the system correctly estimate 80% of the frustrated and, when drivers were not frustrated, made mistakes 9% of the time). Furthermore, the results suggest that information on the driving and communication environment, as well as pedal actuation, was effective in improving the model accuracy.

VI. Driver Education

Employing driver-behavior modeling, we developed a next-generation Event Data Recorder (EDR), which is capable of detecting a wide range of potentially hazardous situations that would not be captured by conventional EDR [4], for enhancing safe driving behavior. Our automated diagnosis and self-review system was developed on a server computer as a web application using CGI for easy access via networks from PCs or smart phones such as the iPhone [19]. The system automatically detects nine types of potentially hazardous situations from the driver's own recorded historical driving data as:

1) Sudden deceleration
2) Sudden acceleration
3) Risky steering
4) Excessive speed
5) Ignoring traffic light
6) Ignoring stop sign
7) Insufficient following distance
8) Potentially risky obstacle avoidance
9) Potentially risky behavior at poor-visibility intersection

The current version will display up to five of the most hazardous scenes of each hazard type by automatically gauging the hazard level using the magnitude of the difference from the pre-defined thresholds (for hazard types 1-7), or from the magnitude of the likelihood ratio between the risky and safe driving models (for hazard types 8-9)[1]. The system allows users to browse through each detected hazardous situation, represented by a balloon icon, on an actual driving map. Each balloon represents one hazardous situation with different colors corresponding to different hazard types, as shown in Fig. 16. The system also provides statistic of all the hazardous situations the driver encountered from all recorded data using a pie chart (e.g., the number of occurrences of each hazard type). Therefore, the system could identify a tendency toward risky driving behavior, or other personality traits possessed by an individual driver.

After clicking on the balloons on the driving map, the corresponding video and driving signals are displayed, along with instructions on how the user can improve their driving safety. The user can also examine different kinds of driving signals related to that particular driving scene. The safety instructions were prepared in advance for each type of detected hazardous situations, based on the above-mentioned manual. In general, the system will inform the user the reason why a particular driving behavior in that situation is considered unsafe, and then tell the user how they can improve their driving behavior. Fig. 17 shows an example of the interface diagnosing a hazardous situation at an intersection. The system notifies the user that the user did not stop at the stop sign and crossed the intersection at a speed of 17 km/h. Then, the system suggests that in this driving situation the driver should completely stop at the stop sign, and confirm that it is safe to cross the intersection before taking action.

A. Experimental Evaluation

In order to validate the effectiveness of our developed system in reducing the number of detected hazardous situations, we recruited 35 drivers, including 6 expert drivers, to participate in our experiment. The subjects were asked to drive the instrumented vehicle three times on three different days, following the same route, which takes approximately one and a half hour to complete. We used data from the second and the third sessions for our analysis, because we wanted

[1]Here, two GMM-based driver-behavior models were used to represent safe driving behavior and risky driving behavior. Risky driving behavior could be determined by performing a hypothesis test of the observations against a pre-defined threshold [5]

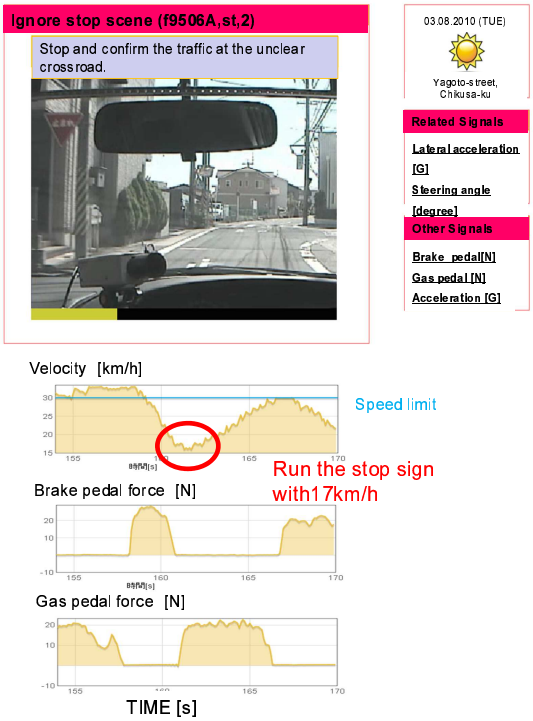

Fig. 17. Example of interface diagnosing hazardous situation at an intersection

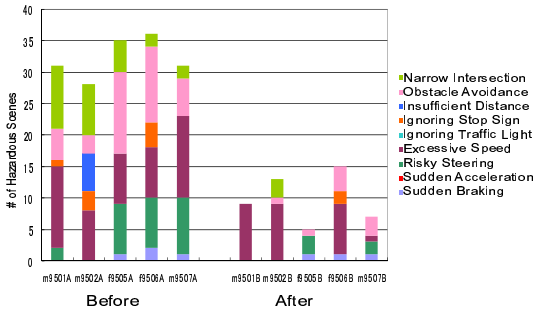

Fig. 18. Comparison of number of hazardous situations before and after using the system.

to allow the subjects to get familiar with the vehicle during the first session. After the second session, the subjects used the driving diagnosis browser and received feedback before taking part in the third session. We compared the number of hazardous situations detected during the second and third sessions. Fig. 18 compares the number of detected hazardous situations for five drivers before and after using the system. The number of hazardous scenes detected for the non-expert drivers decreased by 50% after using the system.

VII. Summary and Future Work

We have presented the behavior signal processing and modeling approaches with a focus on the interaction between driver, vehicle, and environment. The experimental evaluations using realistic driving behavior have shown promising outcomes with a wide range of vehicle applications such as recognizing driver identity, predicting driver maneuver, detecting driver state, and assessing driving behavior. Our future work will focus on identify variations within and between drivers (e.g., between different countries).

Acknowledgment

This work was supported by the Strategic Information and Communication R&D Promotion Programme (SCOPE) of Ministry of Internal Affairs and Communications of Japan, and by the Core Research for Evaluational Science and Technology (CREST) of the Japan Science and Technology Agency. We are also grateful to the members and collaborators of these projects for their valuable contribution and comments.

References

[1] P. Angkititrakul, C. Miyajima, K. Takeda, "Modeling and Adaptation of Stochastic Driver-Behavior Model with Application to Car-Following," in *IEEE IV*, pp.814-819, Baden-Baden, Germany, June 2011.

[2] M. Brackstone and M. McDonald, "Car-following: A historical review," *Transportation Res.*, pt. F, vol. 2, no. 4, pp. 181-196, Dec. 1999.

[3] W. Chee and M. Tomizuka, "Vehicle lane change maneuver in automated highway systems," *PATH Project Report: UCB-ITS-PRR,94-22*, UC Berkeley, 1994.

[4] H.C. Gabor, J.A. Hinch, J. Steiner, *Event Data Recorder: A Decade of Innovation*, SAE International, Warrendale, PA, 2008.

[5] Y. Kuroyanagi, C. Miyajima, N. Kitaoka, and K. Takeda, "Analysis and detection of potentially hazardous situations in real-world driving," in *ICICIC*, vol. 2, no. 3, pp. 621-626, June, 2011.

[6] L. Malta, P. Angkititrakul, C. Miyajima, and K. Takeda, "Multi-modal real-world driving data collection, transcription, and integration using Bayesian Network," in *IEEE-IV*, pp. 150–155, 2008.

[7] L. Malta, C. Miyajima, N. Kitaoka, and K. Takeda, "Analysis of Real-World Driver's Frustration," *IEEE ITS*, vol. 12(1), pp. 109-118, Mar. 2011.

[8] J.C. McCall and M.M. Trivedi, "Driver Behavior and Situation Aware Brake Assistance for Intelligent Vehicles," *IEEE ITS*, vol. 95(2), pp. 374-387, 2007.

[9] C. Miyajima, Y. Nishiwaki, K. Ozawa, T. Wakita, K. Itou, K. Takeda, and F. Itakura, "Driver Modeling Based on Driving Behavior in Driver Identification," *Proceedings of the IEEE*, vol. 95, no. 2, Feb. 2007.

[10] C. Miyajima, T. Kusakawa, N. Kitaoka, K. Itou, and K. Takeda, "On-going Data Collection for Driver Behavior Signal," in *DSPINCARS*, 2007.

[11] K.P. Murphy, "Dynamic Bayesian Networks: Representation, Inference, and Learning," *PhD Dissertation*, Univ. of California at Berkeley, 2002.

[12] Y. Nishiwaki, C. Miyajima, H. Kitaoka, K. Itou, K. Takeda, "Generation of Pedal Operation Patterns of Individual Drivers in Car-Following for Personalized Cruise Control," in *IEEE-IV*, pp. 823–827, Taiwan, 2007.

[13] Y. Nishiwaki, C. Miyajima, H. Kitaoka, K. Takeda, "Stochastic modeling of vehicle trajectory during lane-changing," in *IEEE-ICASSP*, pp. 1377-1380, Taiwan, 2009.

[14] N. Oliver and N.P. Pentland, "Driver behavior recognition and prediction in a SmartCar," in *Proc. SPIE Aerosense, Enhanced and Synthetic Vision*, vol. 4023, pp. 2280–290, Apr. 2000.

[15] A. Pentland and A. Liu, "Modeling and prediction of human behavior," *Neural Comput.*, vol. 11, pp. 229-242, 1999.

[16] L. Rabiner and B. Juang, *Fundamentals of Speech Recognition*, Englewood Cliffs, NJ: Prentice-Hall, Apr. 1993.

[17] D.A. Reynolds, T.F. Quatieri, R.B. Dunn, "Speaker Verification using Adapted Gaussian Mixture Models," *Digital Signal Processing*, vol. 10(1), pp. 19–41, 2000.

[18] D.D. Salvucci, E.P. Boer, and A. Liu, "Toward an Integrated Model of Driver Behavior in a Cognitive Architecture," *Transportation Research Record*, 2001.

[19] K. Takeda, C. Miyajima, T. Suzuki, K. Kurumida, Y. Kuroyanagi, H.Ishikawa, P. Angkititrakul, R. Terashima, M. Oikawa, and Y. Komada, "Improving Driving Behavior by Allowing Drivers to Browse Their Own Recorded Driving Data," in *IEEE-ITSC*, Washington, Oct. 2011.

[20] K. Tokuda, T. Yoshimura, T. Masuko, T. Kobayashi, and T. Kitamura, "Speech parameter generation algorithms for HMM-based speech synthesis," in *ICASSP*, pp. 1315-1318, June 2000.

[21] P.I.J. Wouters and J.M.J. Bos, "Traffic accident reduction by monitoring driver behaviour with in-car data recorders," *Accident Anal. Prev.*, vol.32(5), pp. 643-650, 2000.